



HAL
open science

Teacher recruitment and management : current practices and future challenges

Mélina Hillion

► **To cite this version:**

Mélina Hillion. Teacher recruitment and management : current practices and future challenges. Economics and Finance. Université Paris sciences et lettres, 2018. English. NNT : 2018PSLEH118 . tel-03168270

HAL Id: tel-03168270

<https://theses.hal.science/tel-03168270>

Submitted on 12 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Ecole des hautes études en sciences sociales

Teacher Recruitment And Management:
Current Practices And Future Challenges

Ecole doctorale n°465

ECOLE DOCTORALE ECONOMIE PANTHEON SORBONNE

Spécialité ANALYSE ET POLITIQUE ECONOMIQUES

Soutenue par Méлина HILLION
Le 15 novembre 2018

Dirigée par **Philippe ASKENAZY**

COMPOSITION DU JURY :

Mme. MEURS Dominique
Université Paris Ouest – Nanterre La
Défense,
Rapporteur

Mme. BACACHE Maya
Télécom Paris Tech,
Rapporteur

M. GURGAND Marc
Ecole d'Economie de Paris, CNRS,
Membre du jury

Mme. BEJEAN Sophie
Université de Bourgogne,
Membre du jury

M. BRYSON Alex
University College London,
Membre du jury

Remerciements

Tout d'abord, je tiens à remercier mon directeur de thèse, Philippe Askenazy, de m'avoir fait confiance dès le début en acceptant de diriger mes recherches, de m'avoir fait bénéficier de ses conseils et de ses brillantes idées, et de m'avoir encouragée et soutenue tout au long de cette aventure. Je le remercie également pour sa disponibilité et pour avoir rendu possible mon séjour à l'Université de Berkeley.

Je remercie Dominique Meurs et Maya Bacache pour leurs nombreux commentaires très constructifs, qui m'ont permis d'améliorer considérablement mon travail. Je leur suis également reconnaissante d'avoir accepté d'être rapporteurs de cette thèse. J'ai également fortement bénéficié des conseils de Marc Gurgand que je remercie d'avoir accepté d'être membre du jury.

I would like to thank Alex Bryson for welcoming me so warmly to the Quantitative Social Science seminar at UCL and for agreeing to be a member of the jury. Je suis très reconnaissante à Sophie Béjean d'avoir également accepté d'être membre du jury.

Je souhaite remercier les personnes qui m'ont accueillie à la DEPP et sans le concours desquelles cette thèse n'aurait pas eu sa forme actuelle. Je suis particulièrement reconnaissante à Fabienne Rosenwald, Xavier Sorbes, Cédric Afsa, Pierette Briant et Julie Solard qui m'ont ouvert les portes de la DEPP et m'ont permis de travailler sur les données du ministère de l'éducation. Je les remercie pour leur confiance et leur bienveillance à chaque étape de mes travaux. Je remercie également Caroline Simonis-Sueur pour son aide précieuse et Olivier Monso pour sa très grande disponibilité. Je dois également beaucoup à Catherine Valette, Edouard Maugendre, Marion Defresne et Yves Dubois qui ont eu la patience de répondre à mes nombreuses sollicitations. Je remercie Daniel Auverlot et Philippe Claus d'avoir pris le temps de me recevoir et de répondre à toutes mes questions.

J'ai eu la chance de passer deux années exceptionnelles au CREST. Les nombreux échanges avec les doctorants et les chercheurs ont été une source constante d'inspiration, de réflexion et d'amélioration. Je tiens en particulier à remercier Francis Kramarz et Xavier d'Haultefoelle qui m'ont accueillie au laboratoire de micro-économétrie du CREST et m'ont fait bénéficier de leurs conseils précieux. Je remercie également l'ensemble des chercheurs du CREST pour leur disponibilité, et tout particulièrement Thomas le Barbanchon, Christian Belzil, Pierre Cahuc, Bruno Crépon, Laurent Davezies, Bertrand Garbinti, Robert Gary-Bobo, Alessandro Iaria, Anett John, Vincent Pons, Anna Simoni et Arne Uhlenborff pour m'avoir apporté leur aide à différents moments de ma thèse. J'ai eu la joie de partager le bureau d'Alicia Marguerie, Arnaud Philippe et Daphné Skandalis qui m'ont énormément apporté et avec qui j'ai passé des moments inoubliables. Je remercie également Simon Georges-Kot, Ivan Ouss et Pierre Pora pour leur soutien et leur amitié. Je suis également particulièrement reconnaissante à Aicha Ben Dhia, Antoine Bertheau, Marianne Blehaut, Edouard Chrestien, Jeanne Commault, Emma Duchini, Sandro Favre, Manon Garrouste, Mathilde Godard, Malka Guillot, Yannick Guyonvarch, Clémence Lenoir, Jeremy L'Hour, Victor Lyonnet, Esther Mbih, Julie Pernaudet, Audrey Rain, Pauline Rossi, Clémence Tricaud, Jérôme Trinh, Benjamin Walter et Meryam Zaiem. Je remercie le personnel administratif du CREST, et tout particulièrement Martine Germond, Arnaud Richet et Murielle Jules pour leur aide précieuse.

Plusieurs personnes rencontrées à PSE ont eu un impact déterminant sur cette thèse. J'ai d'abord eu la chance de partager le bureau de Camille Terrier qui m'a fait découvrir les données du ministère de l'éducation. Son organisation et sa rigueur ont été un exemple pour moi. Ensuite, j'ai eu la chance de travailler avec Thomas Breda. J'ai considérablement appris au cours de notre collaboration et grâce à ses nombreux conseils lors de mon comité de thèse. Je le remercie pour son intuition et son audace qui nous ont permis de publier le premier chapitre de cette thèse.

I would like to thank Stephen Ceci for his incredible support from writing to publishing this article. His kindness, comments, availability and responsiveness were extraordinary and played a key role.

Je suis également particulièrement reconnaissante envers Thomas Piketty et Claudia Senik pour leurs commentaires précieux lors de mes présentations en séminaire. Je

souhaite remercier les participants de l'Applied Economics seminar et du Labour and Public Economics seminar de PSE, et tout particulièrement Asma Benhenda, Clémentine Van Effenterre et Fanny Landaud.

I had the opportunity to visit the IRLE and the University of Berkeley. I am extremely grateful to Jesse Rothstein for his hospitality, availability and for the time he devoted to my research projects. I would also like to thank Patrick Kline for his advice, which influenced my thesis beyond my expectations. C'est au contact de plusieurs chercheurs de l'Université de Berkeley que j'ai entamé une seconde phase de maturité de mon projet de recherche. Je remercie Emmanuel Saez et Gabriel Zucman qui m'ont aidée à prendre du recul et à me poser les bonnes questions. I also had the opportunity to meet and share excellent moments with Stephen Aebischer, Yukiko Asai, Cédric Gorinas, Maxim Massenkof, Claire Montialoux, Tobias Renkin and Michael Siegenthaler. Mon regard sur la recherche en économie n'a plus été le même après ce séjour de quelques mois en Californie.

Je souhaite également remercier Arnaud Maurel qui m'a initiée à la recherche, et Denis Fougère qui m'a encouragée et donné la possibilité d'échanger avec les participants du LIEPP de Sciences Po.

J'ai eu la chance de retrouver à la DARES un environnement de travail stimulant et bienveillant. Je remercie mes collègues pour leur soutien et leur patience, et tout particulièrement Selma Amira, Marilyne Beque, Thomas Coutrot, Cécile Higounenc, Ceren Inan, Amélie Mauroux, Sarah Memmi, Corinne Mette, Maria-Teresa Pignoni, Elodie Rosankis, Aguibou Tall et Nicole Waldren.

J'aimerais remercier mon frère et mes parents, Ronan, Odile et André, qui m'ont toujours soutenue, inspirée et incitée à garder un esprit ouvert et critique.

Je remercie Anis pour sa patience, sa gentillesse, son énergie, son optimisme et sa bonne humeur à toute épreuve. Sa contribution est inestimable.

Merci à tous mes amis pour leurs encouragements et leur compréhension, en particulier lorsque ma disponibilité faisait défaut.

Résumé et mots clés

Résumé

L'organisation de la fonction publique française connaît de profondes mutations depuis le début des années 2000. Cette thèse examine la capacité du système de recrutement et de gestion des enseignants à répondre aux enjeux d'attractivité, d'efficacité et de diversité au cœur des projets actuels de modernisation du système éducatif. Le premier chapitre examine la neutralité du processus de recrutement des enseignants du secondaire vis-à-vis du genre. Il révèle que des biais d'évaluation existent et tendent à favoriser le genre minoritaire, contribuant ainsi à renforcer la mixité au sein des disciplines universitaires. Les deuxième et troisième chapitres examinent la capacité des incitations monétaires, de la demande d'enseignants et des exigences en matière de diplômes à attirer davantage et potentiellement de meilleurs candidats aux postes d'enseignant. Une hausse du niveau de diplôme requis pour enseigner (réforme de la "masterisation") ne semble pas améliorer l'efficacité du recrutement en termes d'attractivité, de profils de compétences et de diversité. Le quatrième chapitre examine la relation entre absentéisme, mobilité et conditions de travail des enseignants. Il montre que les écoles et les directeurs d'école influencent significativement les absences, les départs et le bien-être psychologique des enseignants. Le manque de soutien hiérarchique, les comportements hostiles et l'intensité du travail semblent jouer un rôle particulièrement important.

Mots clés

Recrutement, Biais de genre, Offre d'enseignants, Exigence de diplôme, Absentéisme, Roulement du personnel, Conditions de travail

Abstract and keywords

Abstract

The organization of the French civil service has undergone profound changes since the early 2000s. This thesis examines the capacity of the teacher recruitment and management system to meet the challenges of attractiveness, efficiency and diversity at the heart of current projects to modernize the education system. The first chapter examines the gender neutrality of the recruitment process for secondary school teachers. It reveals that evaluation biases tend to favor the minority gender and contribute to strengthening the gender diversity within university disciplines. The second and third chapters examine the ability of monetary incentives, teacher demand and degree requirements to attract more and potentially better candidates for teaching positions. An increase in the level of qualification required to teach (from bachelor's to master's level) does not seem to improve the effectiveness of recruitment in terms of attractiveness, skills profiles and diversity. The fourth chapter examines the relationship between absenteeism, mobility and working conditions of teachers. It shows that schools and school principals significantly influence teacher absences, turnover and psychological well-being. Lack of hierarchical support, hostile behaviors and work intensity seem to play a critical role.

Keywords

Recruitment, Gender bias, Teacher supply, Degree requirement, Absenteeism, Staff turnover, Working conditions

Contents

Remerciements	i
Resumé et mots clés	iv
Abstract and keywords	v
Introduction	1
1 Teaching accreditation exams in France favor women in male-dominated disciplines and men in female-dominated fields	18
1.1 Introduction	18
1.2 Institutional background	21
1.2.1 Competitive exams to recruit teachers in France	21
1.2.2 Systematic non-anonymous oral and anonymous written tests	21
1.2.3 Exams at three different levels	22
1.2.4 Two to three examiners at each test	24
1.3 Data	25
1.4 Method	27
1.4.1 Percentile ranks	27
1.4.2 Variations in percentile ranks between oral and written tests (DD)	28
1.4.3 Odds ratios and relative risks	29
1.4.4 Using total scores on written and oral tests or keeping only one written and one oral test	29
1.4.5 A simple linear model to derive econometric specifications	29
1.4.6 Statistical models used to assess the gender bias on oral tests in each field and at each level	30

1.4.7	Using initial scores instead of percentile ranks	32
1.4.8	Statistical model to assess how the gender bias on oral test varies from a subject to another one	33
1.4.9	Statistical model to assess how the relationship between subjects' extent of male-domination and gender bias on oral test varies between the medium- and the high-level exams	33
1.4.10	Clustering standard errors	33
1.5	Results	34
1.5.1	Gender differences between oral and written test scores at exams to recruit secondary school and postsecondary professorial teachers	34
1.5.2	A clear pattern of rebalancing gender asymmetries in academic fields, strongest at the highest-level exam, and invisible at the lower-level exam	36
1.5.3	Implications for the gender composition of recruited teachers and professors in different fields	37
1.5.4	Gender of evaluators	37
1.5.5	Comparison of an oral test that is common across all exams	38
1.6	Discussion	39
1.6.1	Handwriting detection	39
1.6.2	Gender differences in the types of abilities that are required on oral and written tests	40
1.6.3	Results from statistical models DD, S, and S+IV at the medium- and higher-level exams	42
1.6.4	Results from statistical models DD, S, and S+IV at the lower-level exams	43
1.6.5	Analysis of the effect of the gender composition of the examiner panels	44
1.7	Conclusion	45
	Tables	49
	Figures	63
2	Understanding the effect of salary, degree requirements and demand on teacher supply and quality: a theoretical approach	70

2.1	Introduction	70
2.2	Related literature	72
2.3	The effect of wage	76
2.3.1	A simple model	76
2.3.2	Implication for teacher supply	78
2.3.3	Implication for teacher skills	78
2.4	The effect of labor demand	80
2.4.1	Model	80
2.4.2	Implication for teacher supply	84
2.4.3	Implication for teacher <i>skills</i>	86
2.5	The effect of degree requirement	87
2.5.1	Educational and professional choices when teachers are required to have a bachelor's degree	87
2.5.2	Educational and professional choices when teachers are required to have a masters' degree	93
2.6	Conclusion	97
	Appendix	102
3	How does the increase in teachers' qualification levels affect their supply and characteristics?	106
3.1	Introduction	106
3.2	How to measure teacher quality?	111
3.3	Institutional background	113
3.3.1	The recruitment of public primary school teachers	113
3.3.2	Reform of teachers' diploma level	114
3.4	Data and descriptive statistics	116
3.4.1	Teacher qualifications	117
3.4.2	Teacher supply and teacher demand	120
3.4.3	Teacher <i>skills</i>	124
3.4.4	Teacher diversity	126
3.4.5	Teacher salaries	127
3.5	Exploratory analysis when teachers are recruited at the bachelor level (be- fore 2011)	129

3.5.1	Relationship between teacher supply, teacher demand and salary .	129
3.5.2	Relationship between teacher characteristics, teacher demand and salary	131
3.5.3	Relationship between diploma level and teacher characteristics . .	134
3.6	Model	135
3.7	Results	137
3.7.1	Effect of a higher degree requirement on teacher supply	137
3.7.2	Effect of a higher diploma requirement on teacher characteristics .	140
3.8	A placebo test	146
3.9	Discussion	148
3.9.1	Is there a link between the qualification reform and teacher shortages?	148
3.9.2	Short-term or long-term effects?	152
3.10	Conclusion	152
	Tables	157
	Figures	174
	Appendix	188
A	Primary school teacher recruitment examinations are not standardized in France	188
B	Scope of the analysis	190
C	Descriptive statistics	190
D	A first method for estimating scoring biases between regions	192
E	A second method for estimating scoring biases between regions and years	195
E.1	Descriptive evidence and preliminary analysis	196
E.2	Model	201
E.3	Results	209
F	Robustness checks	214
	Tables	216
	Figures	227
4	How school context and management influence sick leave and teacher departure	233
4.1	Introduction	233
4.2	Literature review	238

4.3	Context	241
4.3.1	Absences in the French education system	241
4.3.2	Recruitment and wage in (secondary) education	241
4.3.3	Mobility	243
4.3.4	Teachers in deprived areas	245
4.3.5	School principals and vice-principals	245
4.4	Data and descriptive statistics	246
4.4.1	Sources	246
4.4.2	Descriptive statistics	247
4.5	Method	248
4.5.1	Specification	248
4.5.2	Estimation	250
4.6	Results	253
4.6.1	Contribution of teachers, schools and principals to the annual duration of teacher absences	255
4.6.2	What is the overall impact of schools and principals on teachers absences? A counterfactual analysis	260
4.6.3	Relationship between the contribution of principals and schools to teacher absences and turnover	263
4.6.4	Who leaves schools that contribute most to absences?	265
4.6.5	Relationship between school effects and school characteristics	266
4.6.6	What are the working conditions and psychosocial risks factors associated with schools' and principals' effects on teacher absences?	268
4.6.7	What is the relationship between teacher health and school and principal effects on absenteeism?	272
4.6.8	Has the increase in prevention from 2014 been more intense in schools (and among school principals) that increase teacher absences?	275
4.6.9	Are newly recruited teachers assigned to more favourable work environments?	278
4.7	Robustness checks	281
4.7.1	Exogeneous mobility assumption	281
4.7.2	Is the most appropriate model additive or multiplicative?	289

4.7.3	What is the impact of log-linearisation on estimates?	292
4.7.4	Are absences partly explained by peer effects?	293
4.8	Conclusion	298
	Tables	306
	Figures	322
	Conclusion: Policy implications, pending issues and future researches	331

List of Figures

1	Administrative regions and academic regions (<i>académies</i>) in France	3
2	Number of primary and secondary teaching posts over the period 1987-2018	5
3	Average net salary (in euros per month) of university graduates according to the proportion of women among university graduates in the field. . . .	8
1.1	Average gender gap between oral and written examinations, by subject, in the recruitment of teachers at the high- and medium-level.	63
1.2	Average rank difference between women and men on oral and written tests in each subject-specific exam at the high- and medium-level.	64
1.3	Difference between women and men average rank on oral test in the subject "Behave as an ethical and responsible civil servant" in different fields of specialization.	65
1.4	Average rank difference between women and men on the oral test "Behave as an Ethical and Responsible Civil Servant" according to subject feminization.	66
1.5	Odds ratios of admission for women candidates at non-anonymous oral tests versus anonymous written tests.	67
1.6	Average gender gap between oral and written examinations, by subject, for candidates taking both medium- and higher-level exams the same year.	68
1.7	Average gender gap in the recruitment of teachers at the high- and medium-level, by subject, considering only one oral and one written test.	69
2.1	Professional choices for individuals eligible to teaching positions	77
2.2	Professional choices of bachelor's graduates when a bachelor's degree is required to teach.	91

2.3	Professional choices for bachelor's graduates when a master's degree is required to teach.	95
3.1	Recruitment and tenure of public primary school teachers for the period 1989-2017	116
3.2	Number of candidates according to the number of teaching posts during the periods 1996-2009 and 2011-2015	139
3.3	Annual logarithmic variation in the number of candidates compared to 1996	140
3.4	Average score of teachers (corrected for standardization errors) on the recruitment examination during the periods 1996-2009 and 2011-2013	143
3.5	Annual variation in the average examination score (corrected for assessment bias, see Appendix) of teachers compared to 2003	144
3.6	Annual variation in the average classroom observation score (2-3 years after tenure) of teachers recruited over the period 2006-2013 (compared to 2006)	145
3.7	Teacher shortage and market tightness over the period 1996-2017	151
3.8	Number of candidates according to the number of teaching posts over the period 1996-2009	174
3.9	Teacher market tightness according to the number of teaching posts over the periods 1996-2009 and 2011-2015	175
3.10	Average score of teachers (raw data) on the recruitment examination during the periods 1996-2009 and 2011-2013	176
3.11	Annual variation in the average exam score (raw data) of teachers compared to 2003	177
3.12	Annual variation in the average examination score (corrected for assessment bias, see Appendix) of candidates compared to 2003	178
3.13	Annual variation in the average examination score (raw data) of candidates compared to 2003	179
3.14	Annual variation in the proportion of men among candidates compared to 2003	180
3.15	Annual variation in the proportion of men among teachers compared to 2003	181
3.16	Annual variation in the median income in the city of residence of candidates compared to 2003	182

3.17 Annual variation in the median income in the city of residence of teachers compared to 2003	183
3.18 Annual logarithmic variation in the number of candidates for the Regional Institutes of Administration (IRA) compared to 2008	184
3.19 Annual variation in the share of men among candidates for the Regional Institutes of Administration (IRA) compared to 2008	185
3.20 Annual variation in the share of men among recruits for the Regional Institutes of Administration (IRA) compared to 2008	186
3.21 Evolution of the number of candidates by age group over the period 2003-2015	187
3.22 Raw and corrected exam score of teachers by academic region in 2003-2004	195
3.23 Variation in candidates' exam score between the first and second attempts according to the quartile of the region of departure and arrival	198
3.24 Evolution of the average score of candidates between the first and second participation in the competition during the period 2005-2014	201
3.25 Average score gain between the first two attempts for candidates who participate in the competition twice in the same region quartile	206
3.26 Raw and corrected exam score of teachers by academic region over the period 2003-2013	213
3.27 Evolution of teachers' raw and corrected scores on recruitment examinations at national level over the period 2003-2013	214
3.28 Variation in candidates' exam score between the first and second attempts according to the quartile of the region of departure and arrival	227
3.29 Relationship between the average level of knowledge of candidates and the generosity of the assessment in the region	228
3.30 Relationship between the average level of knowledge of candidates and the success rate of the competition in the region	229
3.31 Relationship between the success rate of the competition and the generosity of the assessment in the region	230
3.32 Evolution of candidates' raw and corrected scores on recruitment examinations at national level over the period 2003-2013	231

3.33	Evolution of teachers' raw and corrected exam scores at national level during the period 2003-2013 for different values of the correlation between the level of knowledge and the learning gain between attempts (model 3)	232
4.1	Distribution of the Priority Education Network according to school fixed effects on teachers' absences (in percentiles)	267
4.2	Distribution of civil servant teachers according to experience and school fixed effects on teachers' absences (in percentiles)	280
4.3	Absence duration (in logarithm) for teachers transiting between terciles of school fixed effects in year $t=0$	285
4.4	Absence duration (in logarithm) for schools transiting between terciles of school principal fixed effects in year $t=0$	288
4.5	Teacher turnover in schools transiting between terciles of school principal fixed effects in year $t=0$	289
4.6	Correlation between teacher fixed effects and school fixed effects according to the number of "movers" (teachers) per school	322
4.7	Correlation between teacher fixed effects and school principal fixed effects according to the number of "movers" (teachers) per school principal	323
4.8	Correlation between school fixed effects and school principal fixed effects according to the number of "movers" (principals) per school	324
4.9	Distribution of the Priority Education Network according to school fixed effects on teacher turnover (in percentiles)	325
4.10	Distribution of civil servant teachers according to experience and school fixed effects on teacher turnover (in percentiles)	326
4.11	Distribution of contract teachers according to experience and school fixed effects on teachers' absences (in percentiles)	327
4.12	Distribution of contract teachers according to experience and school fixed effects on teacher turnover (in percentiles)	328
4.13	Absence duration (in days) for teachers transiting between terciles of school fixed effects in year $t=0$ (model 4)	329

List of Tables

1	Share of women by field among university professors and assistant professors, secondary school teachers and university graduates	7
2	Measures of teacher shortages in primary and secondary education in 2016	9
1.1	Description of teachers' recruiting exams	49
1.2	General sample statistics for teaching exams 2006-2013	49
1.3	Sample statistics for the high-level exam (Agrégation) and medium-level exam (CAPES) 2006-2013	50
1.4	Estimates of the bonus for women on oral tests at the higher-level exam in each field. Linear regression models DD, S, and S+IV. 2006-2013 . . .	51
1.5	Estimates of the bonus for women on oral tests at the medium-level exam in each field. Linear regression models DD, S, and S+IV. 2006-2013 . . .	52
1.6	Values taken by Indexes of Feminization	53
1.7	Estimates of the linear relationship $b = \beta + \gamma s$ between the bias towards females on oral tests b and 3 indexes of fields' extent of feminization (s). 2006-2013.	53
1.8	Probability of success by gender, assuming success is either based only on written tests or only on oral tests.	54
1.9	Heterogeneity of the bonus for female candidates at the high-level exams oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2006-2013	55
1.10	Heterogeneity of the bonus for female candidates at the medium-level exams oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2006-2013	55

1.11	Estimates of the bonus for women on oral tests for women at the math and literature tests in the lower-level exam. Linear regression models DD, S, and S+IV. 2006-2013	56
1.12	Heterogeneity of the bonus for female candidates at the lower-level exams math and literature oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2011-2013	56
1.13	Effect of the gender composition of the examiners' panels on oral test scores at the math medium-level exam	57
1.14	Composition of the jury at the Maths medium-level exam in 2014	57
1.15	Bonus for women at one oral versus one written test in each field. Linear regression models DD. 2006-2013.	58
1.16	Description of all tests at the medium-level examination	59
1.17	Description of all tests at the high-level examination	60
1.18	Female mean rank at all tests at the medium-level examination	61
1.19	Female mean rank at all tests at the high-level examination	62
3.1	Highest level of diploma and principal occupation of candidates for primary school teaching posts over the period 2003-2015	119
3.2	Orientation of bachelor's graduates during the periods 2005-2009 and 2012-2017	120
3.3	Number of teaching posts in primary schools, number of applicants and number of higher education graduates in 2003-2017	124
3.4	Gender and social composition of candidates and primary school teachers over the 2003-2015 period	127
3.5	Average salaries of bachelor's graduates, master's graduates and primary school teachers in the 25 academic regions of metropolitan France over the period 2003-2017	129
3.6	Relationship between teacher supply, teacher demand and competing employment opportunities over the period 2003-2009	131
3.7	Relationship between average salaries of bachelor's and master's graduates and teachers' <i>competencies</i> in the period 2003-2009	133
3.8	Relationship between the average salaries of bachelor's and master's graduates and the composition of the teaching force during the period 2003-2009	134

3.9	Share of local recruitment by academic region	157
3.10	Number of general baccalaureate graduates, candidates and teaching positions by academic region and time interval	158
3.11	Teacher market tightness, teacher demand and teacher supply by academic region and time interval	159
3.12	Raw and corrected exam scores of candidates by academic region and time interval	160
3.13	Raw and corrected exam scores of teachers by academic region and time interval	161
3.14	Classroom observation score of teachers 2-3 after tenure by academic region and time interval	162
3.15	Correlation between diploma level and gender of candidates and teachers over the period 2003-2009	163
3.16	Correlation between diploma level and median income in the city of residence of candidates and teachers over the period 2003-2009	163
3.17	Correlation between diploma level and corrected examination score of candidates and teachers over the period 2003-2009	164
3.18	Correlation between diploma level and raw examination score of candidates and teachers over the period 2003-2009	164
3.19	Correlation between diploma level and classroom observation score of teachers over the period 2006-2009	165
3.20	Effect of increasing the degree level required to teach on the number of candidates	166
3.21	Effect of the interaction between the 2011 reform and the demand for teachers on the number of candidates	167
3.22	Effect of increasing the degree level required to teach on the exam score of candidates and teachers	168
3.23	Effect of the interaction between the 2011 reform and the demand for teachers on candidates' and teachers' exam score	169
3.24	Effect of increasing the degree level required to teach on the classroom observation score of teachers	170

3.25	Effect of increasing the diploma level required to teach on the number of candidates	171
3.26	Evolution of the number of candidates and the proportion of men among candidates and recruits of the Regional Institutes of Administration (IRA) over the period 2008-2017	172
3.27	Number of candidates by year and age group	173
3.28	Average score (standardized) and average rank (percentile) of candidates according to the number of times they participated in the competition during the period 2005-2013	192
3.29	Variance decomposition of the exam score over the period 2003-2013	210
3.30	Correlation between individual fixed effects, region fixed effects and year fixed effects estimated by models (2) en (3)	212
3.31	Number of candidates, written score and success rate in the primary school teacher recruitment examination over the period 2005-2013, at the national level and by academic region	216
3.32	Average exam score the first and second time candidates participate in the recruitment competition during the period 2005-2014	217
3.33	Academic regions divided into quartiles according to the generosity of the evaluation standard	217
3.34	Average score of candidates the first and second time they participate in the competition according to the quartile of the region of departure and arrival	218
3.35	Correlation between candidates' gender, age and exam score on the first attempt over the period 2005-2009	219
3.36	Correlation between candidates' gender, age, level of diploma and specialization of diploma over the period 2005-2009	220
3.37	Learning gain between the first two attempts depending on the candidate's level of knowledge	221
3.38	Learning gain between the second and third attempts according to the candidate's level of knowledge	222
3.39	Region fixed effects estimated using models (2) and (3)	223
3.40	Year fixed effects estimated using models (2) and (3)	224

3.41	Learning gain between attempts estimated using models (2) and (3)	225
3.42	Comparison of the region fixed effects obtained using different samples and estimation strategies	226
4.1	Socio-demographic characteristics, professional situation and absenteeism of staff in public educational institutions in 2010	247
4.2	Characteristics of public educational institutions in 2010	248
4.3	Comparison of the characteristics of the two main study samples over the period 2007-2015	255
4.4	Contribution of time-varying covariates and fixed effects to the duration of absence of secondary school teachers (Model A)	259
4.5	Variance decomposition of the absence duration of secondary school teachers (Model A)	260
4.6	Impact of schools and school principals on teachers' absence duration according to their quartile of fixed effects	262
4.7	Average absence duration of secondary school teachers according to different scenarios	263
4.8	Correlations between the effect of schools (resp. school principals) on absences and the effect of schools (resp. school principals) on teacher turnover	264
4.9	Interaction between school (resp. school principal) fixed effects and absence difference between individuals and coworkers	266
4.10	Relationship between working conditions and the effects of schools and principals on absences	272
4.11	Relationship between the effects of schools and principals on absences and teachers' health status	274
4.12	Relationship between the level of prevention in schools and the effects of schools and principals on absences	278
4.13	Contribution of time-varying covariates, fixed effects and peer effects to the duration of absence of secondary school teachers (Arcidiacono et al. 2012) (Model D)	297
4.14	Variance decomposition of the absence duration of secondary school teachers taking into account peers effects (Arcidiacono et al. 2012) (Model D)	298

4.15	Correlation between the fixed effects estimated by the main specifications (models 3 and 4) and the specification that takes into account peer effects (model 5)	298
4.16	Contribution of time-varying covariates and fixed effects to the duration of absence of secondary school teachers (Model B)	306
4.17	Variance decomposition of the absence duration of secondary school teachers (Model B)	307
4.18	Variance decomposition of absence duration for female teachers (Model A)	307
4.19	Variance decomposition of absence duration for male teachers (Model A)	308
4.20	Contribution of time-varying covariates and fixed effects to teacher turnover (Model C)	309
4.21	Relationship between school fixed effects and school characteristics	310
4.22	Relationship between school fixed effects and school characteristics, taking into account department fixed effects	311
4.23	Share of teachers exposed to major psychosocial risk factors for different teacher samples	312
4.24	Exposure score to major psychosocial risk factors for different teacher samples	313
4.25	Percentage of teachers in poor health and percentage of teachers reporting preventive actions in their schools	313
4.26	Absence duration (in logarithm) for individuals transiting between terciles of school fixed effects in year $t=0$ (Model 3)	314
4.27	Absence duration (in days) for teachers transiting between terciles of establishment fixed effects in year $t=0$ (Model 4)	314
4.28	Absence duration (in logarithm) in schools transiting between terciles of school principal fixed effects in year $t=0$ (Model 3)	315
4.29	Absence duration (in days) in schools transiting between terciles of school principal fixed effects in year $t=0$ (Model 4)	315
4.30	Turnover in schools transiting between terciles of school principal fixed effects in year $t=0$ (model i)	316
4.31	Relationship between the conditional variance of absences and the conditional expectation of absences	316
4.32	Comparison of estimates obtained by PPML and OLS after log-linearization	317

4.34	Variance decomposition of the absence duration of secondary school teachers taking into account peers' fixed effect and peer's covariates effects (Arcidiacono et al. 2012) (Model E)	319
4.35	Contribution of time-varying covariates, fixed effects and peer effects to the duration of absence of secondary school teachers (Ichino & Maggi 2000) (Model F)	320
4.36	Variance decomposition of the absence duration of secondary school teachers taking into account peers effects (Ichino & Maggi 2000) (Model F)	321

Introduction

The organization of the French civil service has undergone profound changes since the early 2000s. The Organic Law on Finance Laws (LOLF, 2001), implemented from 2006, and the General Review of Public Policies (RGPP, 2007) intend to modernize government action by moving away from the traditional conception of State organization. These reforms are inspired by the principles of the New Public Management ([Hood, 1991](#)), which has been implemented in most OECD countries since the 1990s. The New Public Management criticizes the functioning of the traditional bureaucratic organization (centralized, hierarchical, impersonal, procedural) that leads to a sub-optimal use of state resources. Based on economic and managerial theories¹, proponents of this concept advocate applying private sector management practices to public administrations. The objective is twofold: to improve the quality of public services and to reduce public spendings. This new paradigm requires that the organization of the State be more oriented towards transparency and performance than towards the resources mobilized, which requires more monitoring and evaluation of results. The New Public Management also advocates the decentralization of state decisions, by giving greater autonomy and responsibilities to managers ([Merrien, 1999](#); [Bezes et al., 2011](#); [Kalimullah et al., 2012](#)).

In France, civil service reform has led to the implementation of performance indicators at different levels of execution, the individualization of the evaluation of public employees, and the implementation of a remuneration system partly based on performance. It has also resulted in an increased use of contracts (employees on fixed-term or permanent contracts, public-private partnerships, delegation of services to private companies) and in the reduction of the number of public employees over the period 2007-2012 ([Chevallier, 2010](#)). In the interests of performance, neutrality and equity, recent legislation also en-

¹Public Choice Theory, Transaction Cost Theory and Principal-Agent Theory initiated by [Black \(1958\)](#) and [Arrow \(1963\)](#), and the Scientific Management Movement ([Merkle, 1980](#); [Hume, 1981](#); [Pollitt, 1990](#))

courages the public service to reflect the diversity of citizens, in particular by diversifying recruitment channels.

This thesis focuses on the recruitment and management of teachers who represent 13% of the civil service employees in France². In 2016, primary and secondary school teachers accounted for 1.3% and 1.7% of the working population³. The majority of them (89% of primary and 80% of secondary school teachers) worked in public schools. The organization of the education system in France is centralized and placed under the supervision of the Ministries of Education, Higher Education and Research. The State is responsible for educational policy (curricula, diplomas, budgets, evaluations) and for the recruitment, remuneration and career management of all teaching staff. Several government functions are delegated to regional representatives. Since the 19th century, the country has been geographically divided into *academies* ("academic regions") to which local management of the education system is entrusted. The administrative division corresponding to the *academies* became closer to that of the administrative regions during the 20th century but they never coincided perfectly. Until 2015⁴, there were 27 administrative regions and 30 academic regions (see Figure 1). Each academic region is headed by a *recteur* appointed by the President of the Republic. The *recteurs* ensure the application of national education policy and regulations. They are responsible for the budget and for the management of staff and schools from primary education to university in the academic regions. School principals have administrative (budget proposal and implementation), pedagogical (team coordination and leadership), supervisory (curriculum, school rules, safety), representational (with the municipality and parents) and mediation (problem solving, communication between teachers and parents) responsibilities. They are also responsible for the health and safety of all education employees under their authority. Education inspectors control and evaluate primary and secondary school teachers. They are assigned to an area that includes several schools within the academic region and they organize in-service teacher training.

²There are 5.7 million civil servants and contract workers in the three components - State, territory and hospital - of the French civil service in 2016. Source: L'emploi dans la fonction publique en 2016, Insee Première N°1691, Mars 2018. Bilan social du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche - 2015-2016. Authors' calculations.

³Source: Bilan social du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche - 2015-2016. Tableaux de l'économie française, Insee Référence, Edition 2018. Authors' calculations.

⁴Several administrative regions merged in 2016 but with no consequences for academic regions. In 2018, there are 18 administrative regions.

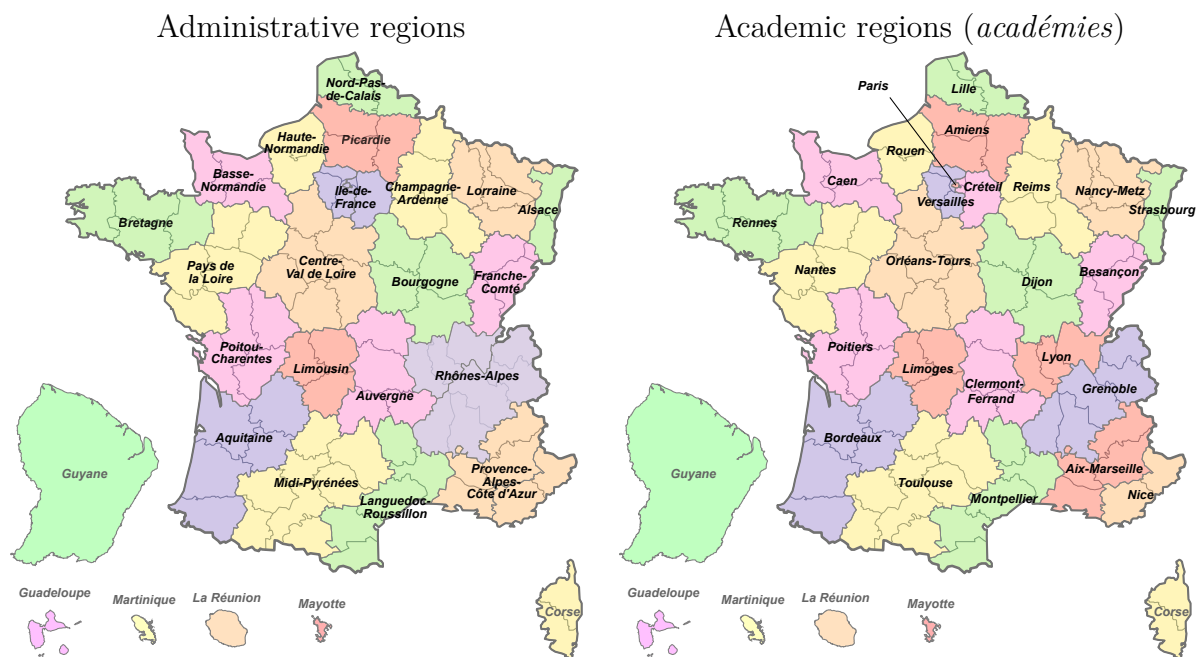


Figure 1 – Administrative regions and academic regions (*académies*) in France

Note: The map on the left shows the administrative regions in France from 1970 to 2015. The map on the right shows the academic regions (*Académies*) since 1996.

Most teachers are recruited through competitive examinations (99% in primary education and 92% in secondary education). The *recteurs* organize teacher competitions in the regions, select the examinations and appoint the members of the juries. Examiners are generally chosen from among teachers and education inspectors for the recruitment of primary school teachers and among teachers and university professors for the recruitment of secondary school teachers. Teachers who pass the competition usually become civil servants after a one year probationary period. Other teachers are hired on the basis of renewable short-term contracts (more rarely on the basis of open-ended contracts) of a maximum duration of one year. Head teachers have little influence on the recruitment and mobility of permanent teachers. However, they are in charge of the recruitment and renewal of non-permanent teachers' contracts. The remuneration of teachers depends mainly on their status (civil servant or under contract) and seniority. Overtime and bonuses represent 7% of primary school teachers' salaries and 16% of secondary school teachers' salaries on average. Evaluations carried out by education inspectors, as well as by school principals in secondary education, have a relatively moderate influence on teachers' career development and salary prospects.

In recent years, the education system has been at the heart of ongoing changes in the

public service. Successive reforms have led to a new organization of school time, a renewal of school curricula, the implementation of pupil assessments in primary education and the increased autonomy of secondary schools in terms of subject provision. Human resources management has also undergone significant changes. The number of tenured teacher positions has decreased considerably over the period 2007-2011 (see Figure 2). Since 2011, primary and secondary school teachers must hold a master's degree (a bachelor's degree was required since 1989). In 2014, the competition tests were modified to give a more important role to the assessment of professional skills (Desbiolles, 2017). An oral interview to assess candidates' motivation, behavioural qualities and adaptability to the teaching profession was added in 2014. In 2017, the professional evaluation of teachers was modified to introduce more transparency and consistency at the national level. Between 2012 and 2014, and from 2018 onwards, the first day of sick leave is no longer paid in the public service. This measure aims to limit excessive absence behaviour and reduce the costs associated with compensation and replacement of absent employees. Recently, discussions have intensified on the autonomy granted to school principals, particularly in terms of staff recruitment and school budget management. Some actors in the education system are questioning the ability of competitive recruitment to meet the specific needs of students. Others argue that the civil servant status limit the possibility of setting up an individualized management system (Boissinot, 2017).

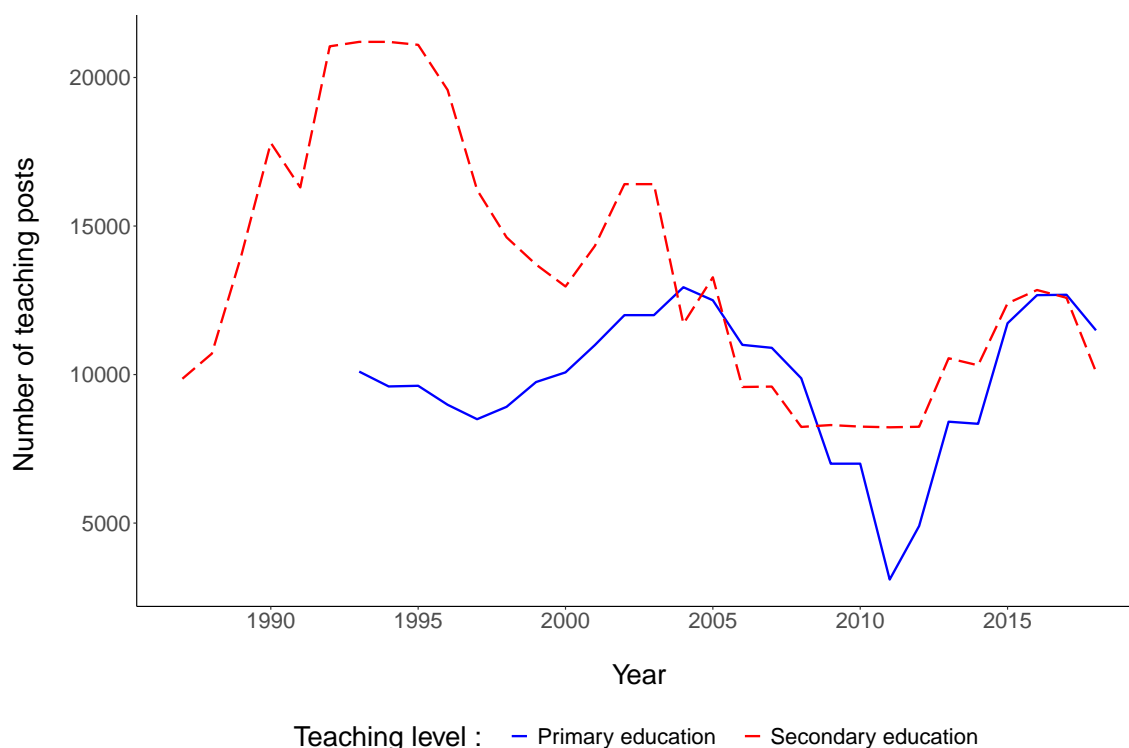


Figure 2 – Number of primary and secondary teaching posts over the period 1987-2018

Note: Evolution of the number of teaching positions in the external recruitment competitions for public primary and secondary education over the period 1987-2018. About 85% of tenured teachers are recruited through external competitions. Source: Ministry of Education (DEPP).

The 2010-2016 period is characterized by a stagnation in public sector wages (*gel du point d'indice*). The increase in teacher qualification levels from 2011 onwards has not led to an increase in teacher salaries⁵, which remain below the OECD average, particularly in primary education (OECD, 2014). In this context of wage moderation and organizational transformations, this thesis examines the capacity of the teacher recruitment and management system to meet the challenges of attractiveness, efficiency and diversity at the heart of ongoing modernization projects. Particular attention is paid to how recent reforms have contributed to the equal treatment of pupils.

The four chapters that compose this thesis exploit the richness and exhaustiveness of the administrative data of the Ministry of Education. The recruitment and personnel management databases make it possible to monitor the careers of hundreds of thou-

⁵An increase in teacher salaries has began for July 2016. It is reflected in an increase in the fixed salary (end of the *gel du point d'indice*), an increase in bonuses (subject to the exercise of specific missions in secondary education) and an increase in the salary at the beginning and at the end of teacher career by 2020.

sands of primary and secondary school teachers, from recruitment, through successive assignments, to departure from the education system. They indicate how individuals were recruited (type of competition, results of recruitment tests), provide detailed socio-demographic information (age, sex, level of diploma, specialization of diploma, place of residence, family and professional situation), characteristics of posts occupied (subject taught, number of hours worked, type of establishment, address), career development (status, seniority, evaluation by education inspectors) and provide information on absences (reason, dates, duration). The studies presented in this thesis largely exploit the longitudinal nature of these data. The rules for managing teachers in France (recruitment, assignment) also make it possible to implement original identification strategies to address international issues: the lack of gender diversity between professions, the discrimination in hiring, the effect of qualification requirements on teacher recruitment, the effects of managerial practices and working conditions on staff health and motivation. This thesis hopes to contribute to the understanding of the challenges of the current education system and shed new light on past and future reforms.

The **first chapter** focuses on the issue of the gender neutrality of the secondary school teacher recruitment examination. The recruitment of civil servants by competitive examination has gradually become widespread in France since the 19th century. The 1983 law defining the "rights and obligations of civil servants" specifies that the competition must guarantee the neutrality of recruitment (principle of equal treatment of candidates), the efficiency of the civil service (by selecting the most competent candidates) and the independence of civil servants from the political authorities (recruitment is based solely on the "merit" of the candidates). These principles also define the rules governing the remuneration and career of civil servants, according to their status and seniority. However, the current recruitment of civil servants is far from being meritocratic. Social science research has shown that competitions largely reproduce the gender, ethno-racial, geographical, economic and social inequalities that accumulate throughout the candidates' academic (and sometimes professional) careers (Duru-Bellat, 2003; Versini, 2004; Pouget, 2005; DiPrete and Eirich, 2006; Eymeri-Douzans, 2012). This first chapter examines whether competitive recruitment, a standardized process that promises candidates to be assessed under the same conditions and according to the same criteria, is neutral, as required by law, or whether it contributes to increasing (or reducing) inequality among

applicants.

The lack of gender diversity in most academic disciplines is a well-known phenomenon that concerns higher education graduates, secondary school teachers and university professors (see Table 1). This gender disparity is coupled with a wage disparity since the most feminized specialties generally offer the lowest income prospects (see Figure 3). Several studies also show that a significant share (between 8 and 20%) of the gender pay gap comes from differences in university specialization (Machin and Puhani, 2003). In addition to the issue of equal access to public employment, gender segregation raises the issue of equal treatment of students. Studies in educational sciences have indeed shown that the gender of the teacher can influence the results, orientation and specialization choices of students. Cases of discrimination have been identified in some studies (Dee, 2005), but recent research has mainly highlighted the role model of teachers (Bettinger et al., 2005; Marx et al., 2013; Paredes, 2014; Porter et al., 2017; Breda et al., 2018). These results, consistent with the theory of representative bureaucracy, suggest ensuring diversity of teaching staff in all disciplines.

Table 1 – Share of women by field among university professors and assistant professors, secondary school teachers and university graduates

Discipline	Proportion of women among professors and assistant professors (%)	Proportion of women among secondary school teachers (%)	Proportion of women among university graduates (%)
Mathematics	20.9	38.4	39.7
Philosophy	27.1	44.8	39.7
Physics	16.8	49.5	24.8
Chemistry	37.4		44.5
Geography	36.5	47.8	43.4
History	41.9		52.3
Economics and social sciences	39.6	65.1	53.3
Biology and geology	45.9	82.5	53.1
Modern and classical literature	55.7	82.5	80.1
Languages	61.9	83.9	78.2

Notes: 20.9% of mathematics professors and assistant professors (respectively 38.4% of mathematics secondary school teachers) are women in 2013. 39.7% of university graduates in mathematics are women in 2005-2013.

Source: Ministry of Higher Education and Research (SIES), Ministry of Education (DEPP), Labor Force Survey (INSEE). Author's calculations.

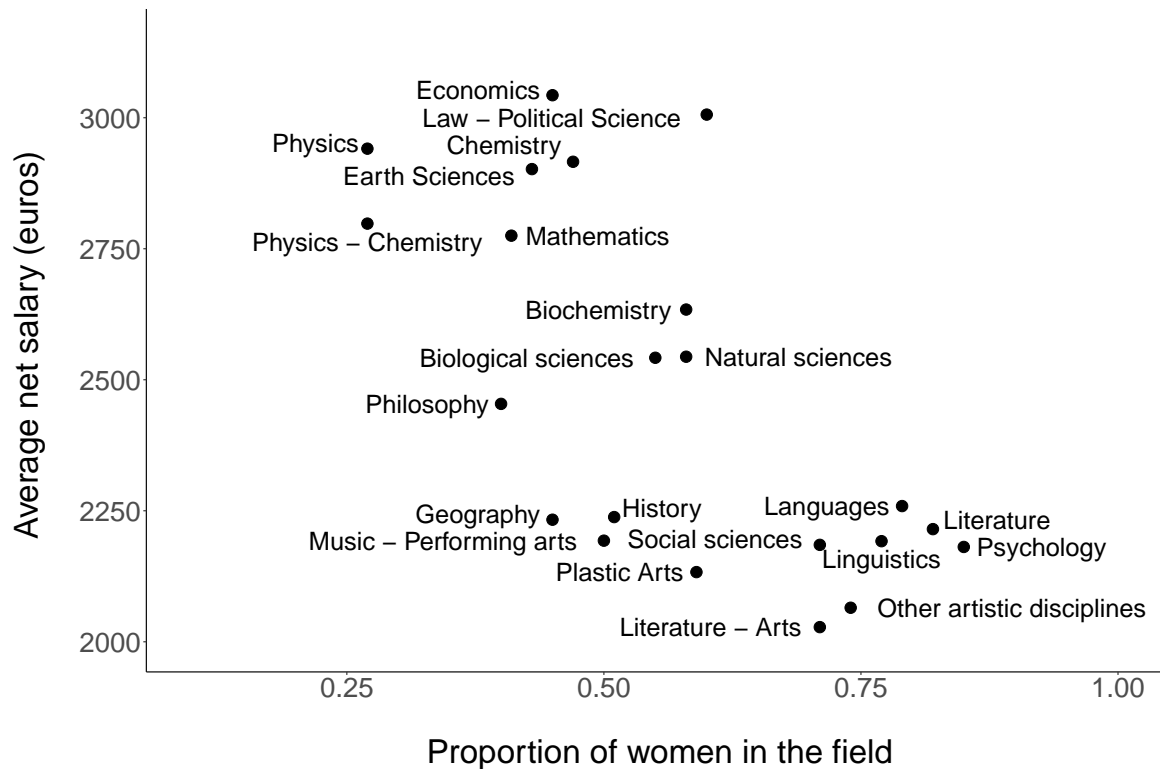


Figure 3 – Average net salary (in euros per month) of university graduates according to the proportion of women among university graduates in the field.

Note: Average net salary (in euros per month) of higher education graduates by academic discipline over the period 2002-2011. Proportion of women among higher education graduates by academic discipline over the period 2002-2011. Source : Labor Force Survey (INSEE). Author's calculations.

The first chapter reveals that the recruitment of secondary school teachers through competitive examinations is not neutral. Evaluation biases tend to favour the minority gender, and thus contribute to strengthening the diversity of teaching staff in all disciplines. Although in contradiction with the principle of neutrality established by law, these recruitment biases are likely to reduce differences in students' specialization in higher education, and thus potentially the gender pay gap. The advantage granted to the minority sex when recruiting teachers partly compensates for school and family inequalities, which is reflected in the lack of gender diversity among candidates. The (uncoordinated) attitude of recruiters towards greater gender diversity in secondary education could reflect their mobilization on diversity issues.

The **second and third chapters** examine the capacity of the current management system (recruitment, remuneration, assignment) to attract and motivate the next generation of teachers. For several years, an increasing number of countries have been facing

teacher shortages to varying degrees (European Commission 2014). This phenomenon is intensifying on an international scale. It concerns France again after having disappeared for several years in secondary education and several decades in primary education. Teacher shortages result in a small number of candidates, difficulties in recruiting appropriate profiles, positions that remain unoccupied for a longer or shorter period, and difficulties in replacing absent teachers. Although difficult to measure, the shortage of teachers in France can be approximated by the number of candidates per post for recruitment competitions, the rate of unfilled posts at the end of the usual recruitment procedure and the proportion of teachers on short-term contracts⁶ In secondary education, teacher shortages are concentrated in a few subjects: mathematics, literature and foreign languages in particular. In primary education, where recruitment is decentralized, shortages are concentrated mainly in two regions: Créteil and Versailles (see Table 2).

Table 2 – Measures of teacher shortages in primary and secondary education in 2016

	Number of candidates per teaching post (%)	Proportion of unfilled teaching posts in the main competitive examinations (%)	Proportion of teachers on short contracts (%)
Primary education			
<i>All academic regions</i>	2.5	4.7	0.6
Bordeaux	3.9	-4.1	0
Lyon	2.8	0	0.7
Aix-Marseille	2.6	0	0.2
Lille	2.6	0	0.1
Grenoble	2.5	0	0
Versailles	1.2	12.9	0.4
Créteil	1.2	25.9	2.5
Secondary education			
<i>All fields</i>	3.2	14.3	6.2
History/Geography	4.9	0	3.9
Physics/Chemistry	4.4	0	3
Biology/Geology	4.3	0	5.6
Economics/Management/Social Sciences	5.6	8.2	9.4
Languages	2.9	15.9	7.6
Literature	1.7	23.3	6.9
Mathematics	2.2	24.6	7.2

Notes: Data are presented for academic regions that have recruited more than 14,000 primary school teachers in 2016 and for academic fields that have recruited more than 11,000 secondary school teachers in 2016. In the Créteil region, there are 1.2 candidates for 1 primary school post in 2016. In 2016, 25.9% of primary school positions in the external competition were not filled and 2.5% of primary teachers were on short-term contracts in the Créteil region.

Source: Ministry of Education (DEPP). Author's calculations.

⁶Teacher shortages could also be approximated by the number of teacher-free days per student during the school year. Teachers on short-term contracts respond to specific needs (replacing an absent teacher) or fill positions that remain vacant at the end of the tenured teacher assignment process.

The **second chapter** draws on the theoretical and empirical literature in economics to propose a simple model that provides a better understanding of the effect of salary, the number of positions and the level of qualification required on the supply and skills of teachers. Based on reasonable assumptions, the model demonstrates an increase in the number of candidates as teacher salary increases, a positive (but decreasing) elasticity in the supply of teachers with the number of positions, and a decrease in the number of candidates as the required qualification level increases. This theoretical framework also helps to understand the mechanisms by which diversity in recruitment is likely to be affected. However, more restrictive assumptions would be needed to remove ambiguity about the effect of salary, demand and diploma level on teachers' competencies. The unexpected increase in teachers' salaries from 2017 onwards (especially for those working in priority areas of education) will make it possible to estimate the elasticity of teacher supply and characteristics with salary.

The **third chapter** examines whether the so-called "masterisation" reform, which raised the level of qualification required to teach from a bachelor's to a master's degree in 2011 in France, has affected the attractiveness of the profession and the characteristics of the teachers recruited. This reform is quite clearly based on the principles of the new public management. First, it aims to improve the quality of education by extending the duration of teachers' studies. Second, it aims to reduce public expenditure by no longer remunerating initial training, which is now integrated into a master's program that must be obtained before recruitment⁷. Although a dozen OECD countries are now recruiting teachers at the master level (European Commission 2014), this type of policy has never been evaluated to my knowledge. This chapter shows that the increase in the level of qualification required has led to a sharp drop in the number of candidates for primary teacher posts. The reform also reduced the level of teachers' knowledge measured by the competitive recruitment examination, but did not affect the pedagogical skills assessed in class by education inspectors. The gender mix of recruitment also decreased after the reform, but the context of stagnating wages may have contributed to the decline in the share of men among candidates. Following the 2011 reform, recruitment inequalities have

⁷The initial training of civil servant teachers lasts two years. The first year focuses on the acquisition of disciplinary and pedagogical knowledge, while the second year focuses on practical training. Before 2011 and from 2014, the first year of training is optional and precedes recruitment. The second year is open only to successful candidates and includes a significant number of hours of classroom instruction. Trainee teachers are paid full-time during this second year of training.

tended to widen between regions. Teacher shortages have emerged in regions where the number of candidates has declined the most and additional competitive examinations were created from 2015 onwards to fill posts that remained vacant following the usual recruitment procedure. The share of contract teachers and the difficulties in replacing absent teachers remain higher in these regions than in the rest of the country.

The **fourth and last chapter** of this thesis examines the link between working conditions and managerial practices on the one hand, and teacher absences and turnover on the other. In most countries, the implementation of the New Public Management has led to increased school autonomy. This autonomy has taken various forms and intensities depending on the country (Scheerens and Maslowski, 2008) but it generally leads to greater margin of decision for the school principal in terms of pedagogy (choice of teaching projects and methods, courses and programmes, timetables), personnel management (recruitment, dismissal, remuneration) and budget management (including the search for private funding sources). Supporters of school autonomy argue that giving schools more responsibility and flexibility to achieve the objectives set by the state will improve the performance of the education system as a whole and reduce educational inequalities. By giving more space to individual initiatives, school autonomy fosters innovation and makes the teaching profession more attractive. By giving school principals the freedom to form the most appropriate pedagogical team, school autonomy makes it possible to meet the specific needs of pupils.

Recent literature in economics of education shows that headteachers have a significant influence on the academic performance of students (Coelli and Green, 2012; Branch et al., 2012). However, the managerial practices, pedagogical choices and conditions associated with student success are still largely unknown (Clark et al., 2009). One reason for this is that the effects of the reforms that have led to greater school autonomy around the world have received very little evaluation. Recent empirical studies show that these initiatives have mixed, and overall rather modest, effects on student performance (Clark et al., 2009; Allen, 2010; Epple et al., 2017). Studies based on international comparisons are more optimistic and generally conclude that school autonomy has a strong positive impact on student achievement. However, the decentralization of responsibilities and decisions is generally accompanied by a strengthening of school monitoring, which can explain a significant part of the estimated gains in school success.

Several actors in the French education system advocate greater autonomy for schools, particularly in the recruitment and evaluation of teachers (Chaix, 2015; Lefèvre, 2015). Many of them agree that the success of such a transformation depends on the managerial skills of school principals and, in particular, on the attention they pay to working conditions. Chapter 4 shows that, despite relatively limited decision latitude⁸, the school principal has a real influence on sick leaves and voluntary departures of teachers. In addition, teachers faced with increased absenteeism and departures from their schools following the arrival of a new headteacher more often report a lack of hierarchical support. This first result raises the question of the recruitment and training of headteachers, particularly in the area of psychosocial risk prevention. In the United States, where schools enjoy a high degree of autonomy, a recent study shows that intensive training of principals in management practices has a positive impact on student success (Fryer et al., 2017). The effect of training headteachers in prevention could be the subject of future research. This fourth chapter also shows that schools influence teacher absences and turnover. In schools that increase absenteeism and departures, teachers are more often confronted with hostile behaviors and high work intensity. The analysis reveals that teachers working in these schools are more likely to develop depressive symptoms and that young and inexperienced teachers are more likely to be assigned to these schools. These results call for an improvement in working conditions in schools. They also question the effectiveness of seniority-based mobility rules that result in inexperienced teachers being assigned to schools with the least favorable working and learning conditions. This chapter suggests that teacher absenteeism can be significantly reduced by addressing the issue of working conditions and managerial practices. The method proposed in this chapter can contribute to identifying the schools and school heads for whom prevention actions should be carried out as a priority.

In the following paragraphs, I present in more detail the context, the methods and the results obtained in the four chapters that compose this doctoral thesis.

Chapter 1, *Teaching accreditation exams in France favor women in male-dominated*

⁸In France, with rare exceptions, school principals are not consulted when recruiting and assigning tenured teachers. In secondary education, principals have authority over teachers and non-teaching staff but are not empowered to take disciplinary action. Moreover, the evaluation of teachers by school principals has practically no impact on their remuneration.

disciplines and men in female-dominated fields, is a joint work with Thomas Breda and investigates whether discrimination against women is a possible cause behind their underrepresentation in certain STEM (Science, Technology, Engineering, and Mathematics) subjects, as suggested by some authors. We show that this is not the case at the competitive exams used to recruit almost all French secondary and postsecondary teachers. Our method is based on the comparisons of oral non gender-blind tests with written gender-blind tests for about 100,000 individuals observed in 11 different fields over the period 2006-2013. We find that the bias in favor of women (respectively men) is strongly increasing with the extent of a field's male (respectively female)-domination, as measured by the share of women (respectively men) among the professors and assistant professors in the field. This bias turns from 3 to 5 percentile ranks for men in literature and foreign languages to about 10 percentile ranks for women in math, physics or philosophy. One implication of this study is that active policies aimed at counteracting stereotypes and discrimination should focus more on early ages, before educational choices are made. In the absence of discrimination in hiring, future research should also focus on understanding why women enroll less often than men in science.

Chapter 2, *Understanding the effect of salary, degree requirements and demand on teacher supply and quality: a theoretical approach*, examines under which conditions degree requirements and financial incentives can help attract more and potentially better candidates for teaching positions. For several years, most developed countries have faced growing problems of teacher shortages in primary and secondary education, and it seems essential to better understand the levers available to stem this phenomenon. After a review of the theoretical and empirical literature, this chapter proposes a simple framework based on classical economic theory. The model shows, with reasonable assumptions, that increasing teacher salaries and demand have positive effects on teacher supply, which is consistent with the results of the empirical literature. On the other hand, increasing the level of qualification required to teach is expected to reduce the number of candidates for teaching posts. The effect of salary, demand and qualifications on teacher quality is much more difficult to predict. It depends heavily on the nature of the correlation between teacher on-the-job performance and the competencies measured by the degrees and selection criteria used in recruitment. The effect also depends on the ability of financial and non-financial incentives to attract the best candidates. Empirical evaluations should

make it possible to test the model's assumptions and predictions.

Chapter 3, *How does the increase in teachers' qualification levels affect their supply and characteristics?*, assesses the reform that requires teachers to be recruited at master level rather than at bachelor level in France since 2011. In recent years, many European countries have increased the qualification level of teachers. However, these reforms have not been evaluated and we do not know whether they have improved recruitment efficiency. Using a first difference approach, I find that the increase in teacher qualifications has resulted in a 50% decrease in the average number of candidates. The accreditation score (subject content knowledge) of teachers fell by 0.55SD and the share of men (minority gender) fell by about 15% following the introduction of the reform. In contrast, the classroom observation score (teaching attitudes and practices) remained relatively unchanged. The recruitment inequalities between regions (level of recruitment and attractiveness) also tended to increase after the reform. These results invite to recruit teachers again at the bachelor level. Another possibility would be to substantially increase teachers' salaries to compensate for the wage gap with the private sector.

Chapter 4, *How school context and management influence sick leave and teacher departure*, examines the influence of secondary schools and school principals on sick leave and teacher turnover in France. Much research in economics suggests that managerial and workplace practices contribute to corporate and state performance but we know little about their effects on workers' behaviour and health. I decompose teacher absences due to illness (respectively teacher turnover) into individual, school and principal contributions using a method similar to the value-added models in economics of education or to the AKM methods in labour economics (high dimensional decomposition methods). The main sample is composed of about 300,000 teachers that I observe on average for 6 consecutive years. I find that schools and school principals substantially contribute to teacher absences: sick leave duration increases by 250-300% (\approx 16-18 days per year) on average for teachers moving from the first to the fourth quartile of school fixed effects, while sick leave duration increases by 170-220% (\approx 12-14 days per year) on average in schools transiting from the first to the fourth quartile of school principal fixed effects. Evidence suggests that school and principal fixed effects are more likely to reflect adverse working conditions than absence norms. I show that school (respectively school principal) effects increase with teacher turnover, and that teachers who are more absent than their

colleagues are also more likely to leave these schools (respectively school principal). I find evidence that school effects are correlated with work intensity and hostile behaviors, while principal effects are correlated with lack of hierarchical support. School fixed effects are also negatively correlated with teachers' psychological well-being as measured by the World Health Organization index (WHO-5). These results suggest that better consideration of psychosocial risk factors in the education sector could reduce teacher absences and associated costs, but also improve teacher health.

Bibliography

- Allen, R. (2010). Does school autonomy improve educational outcomes ? Judging the performance of foundation secondary schools in England. *DoQSS Working Paper*, (10-02).
- Arrow, K. J. (1963). *Social Choice and Individual Values*. New Haven and London, Yale University Press.
- Bettinger, E. P., Long, B. T., Ehrenberg, R., Jacob, B., and Murnane, R. (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, 95(2):152–157.
- Bezes, P., Demazière, D., Le Bianic, T., Paradeise, C., Normand, R., Benamouzig, D., Pierru, F., and Evetts, J. (2011). New Public Management et professions dans l’Etat: au-delà des oppositions, quelles recompositions? *Sociologie du travail*, 53(3):293–348.
- Black, D. (1958). *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.
- Boissinot, A. (2017). Refonder la formation des enseignants : pour quoi faire ? Cinq questions sur la formation des maîtres. *Adminsitration & Education*, 2(154):13–18.
- Branch, G. F., Hanushek, E. a., and Rivkin, S. G. (2012). Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals. *CALDER working paper*, (17803).
- Breda, T., Grenet, J., Monnet, M., and Effenterre, C. V. (2018). Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools. *PSE Working Papers*, 2018 - 06.
- Chaix, G. (2015). Autonomies : pourquoi ? pour qui ? comment ? *Adminsitration & Education*, 147(3):13–32.
- Chevallier, J. (2010). Révision générale des politiques publiques et gestion des ressources humaines. *Revue française d’administration publique*, 136(4):907–918.
- Clark, D., Martorell, P., and Rockoff, J. (2009). School Principals and School Performance. *Calder Working Paper No 38*.
- Coelli, M. and Green, D. A. (2012). Leadership effects: school principals and student outcomes. *Economics of Education Review*, 31(1):92–109.
- Dee, T. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Association*, 95(2):158–165.
- Desbiolles, P. (2017). Quelle évolution des concours de recrutement des enseignants ? *Adminsitration & Education*, 2(154):37–42.
- DiPrete, T. A. and Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. *Annual Review of Sociology*, 32(1):271–297.
- Duru-Bellat, M. (2003). Actualité et nouveaux développements de la question de la reproduction des inégalités sociales par l’école. *L’Orientation scolaire et professionnelle*, 32(4):571–594.
- Epple, D., Romano, R. E., and Urquiola, M. (2017). School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature*, 55(2):441–492.
- Eymeri-Douzans, J.-M. (2012). Les concours à l’épreuve. *Revue française d’administration publique*, 142(2):307–325.
- Fryer, R. G., Angrist, J., Card, D., Charles, K., Dobbie, W., Gibbons, R., Greenstone, M., Katz, L., Hendren, N., Shapiro, J., Reenen, J. V., Noveck, M. H., Gussin, S., Clark, D., Ruebeck, H., Sharma, D., and Williams, D. (2017). Management and Student

- Achievement: Evidence from a Randomized Field Experiment. *NBER Working Paper series*.
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69:3–20.
- Hume, L. J. (1981). *Bentham and bureaucracy*. Cambridge University Press.
- Kalimullah, N. A., Alam, K., and Nour, M. (2012). New Public Management: Emergence and Principles. *BUP Journal*, 1(1):1–22.
- Lefèvre, H. (2015). L'autonomie en entreprise : et si on l'appliquait dans les établissements ? *Adminsitration & Education*, 147(3):151–158.
- Machin, S. and Puhani, P. A. (2003). Subject of degree and the gender wage differential: evidence from the UK and Germany. *Economics Letters*, 79(3):393–400.
- Marx, D. M., Monroe, A. H., Cole, C. E., and Gilbert, P. N. (2013). No Doubt About It : When Doubtful Role Models Undermine Men ' s and Women ' s Math Performance Under Threat. *Basic and Applied Social Psychology*, 38(5):258–268.
- Merkle, J. A. (1980). *Management and Ideology: the legacy of the international scientific management movement*. Berkeley: California University Press.
- Merrien, F.-X. (1999). La Nouvelle Gestion publique : un concept mythique. *Lien social et Politiques*, (41):95–103.
- OECD (2014). Education at a Glance 2014: Highlights. *OECD Publishing*.
- Paredes, V. (2014). A teacher like me or a student like me? Role model versus teacher bias effect. *Economics of Education Review*, 39:38–49.
- Pollitt, C. (1990). *Managerialism and the Public Service: The Anglo American Experience*. Oxford: Blackwell.
- Porter, C., Serra, D., Croson, R., Desmet, K., Lindo, J., Meer, J., Millimet, D., Ozerturk, S., Kuka, E., Salmon, T., and Schechter, L. (2017). Gender Differences in the Choice of Major: The Importance of Female Role Models. *Working Paper*.
- Pouget, J. (2005). La Fonction publique : vers plus de diversité ? *Dossiers de la fonction publique*, pages 143–162.
- Scheerens, J. and Maslowski, R. (2008). Autonomie des établissements scolaires : des moyens à la recherche d'un objectif ? *Revue française de pédagogie*, 164:27–36.
- Versini, D. (2004). Rapport sur la diversité dans la fonction publique.

Chapter 1

Teaching accreditation exams in France favor women in male-dominated disciplines and men in female-dominated fields

JOINT WITH THOMAS BREDA

1.1 Introduction

Why are women underrepresented in most areas of science, technology, engineering, and mathematics (STEM)? One of the most common explanations is that a hiring bias against women exists in those fields ([West and Curtis, 2006](#); [Shalala et al., 2006](#); [Hill et al., 2010](#); [Sheltzer and Smith, 2014](#)). This explanation is supported by a few older experiments ([Swim et al., 1989](#); [Foschi et al., 1994](#); [Steinpreis et al., 1999](#)), a recent testing with fictitious resumes ([Moss-Racusin et al., 2012](#)), and a recent lab experiment ([Reuben et al., 2014](#)), suggesting that the phenomenon still prevails.

However some scholars have challenged this view ([Ceci and Williams, 2011](#); [Ceci et al., 2014](#)) and another recent testing with fictitious resumes finds opposite results, namely a bias in favor of women in academic recruitment ([Williams and Ceci, 2015](#)). Studies based on actual hiring also find that when women apply to tenure-track STEM positions, they are more likely to be hired ([National Research Council, 2010](#); [Wolfinger et al., 2008](#); [Glass and Minnotte, 2010](#); [Irvine, 1996](#)). However, those studies do not control for applicants' quality and a frequent claim is that their results simply reflect the fact that only the best female PhDs apply to these positions while a larger fraction of males do so ([Ceci et al.,](#)

2014; National Research Council, 2010). One study did control for applicants' quality and reported a bias in favor of women in male-dominated fields (Breda and Ly, 2015), but it has limited external validity due to its very specific context. The present analysis is based on a natural experiment over 200,000 individuals participating in competitive exams for primary, secondary and college/university teaching positions in France over the period 2006-2013, and it has two distinct advantages over all previous studies. First, it provides the first large-scale real-world evidence on gender biases in both evaluation and hiring, and how those biases vary across fields and contexts. Second, it offers possible explanations for the discrepancies between existing studies. Those discrepancies may be explained by various factors, ranging from experimental conditions, contexts, type of evaluations made (e.g. grading or hiring), and the math-content of the exams. We hypothesize that two moderators are important to understand what shapes evaluation biases against or in favor of women: the actual degree of female under-representation in the field in which the evaluation takes place and the level at which candidates are evaluated, from lower-level (primary and secondary teaching) to college/university hiring.

Carefully taking into account the extent of under-representation of women in 11 academic fields allows us to extend the analysis beyond the STEM distinction. As pointed out recently (Ceci et al., 2014; Williams and Ceci, 2015; Breda and Ly, 2015; Leslie et al., 2015), the focus on STEM versus non STEM fields can be misleading to understand female underrepresentation in academia as some STEM fields are not dominated by men (e.g. 54% of U.S. Ph.Ds. in molecular biology are women) while some non-STEM fields, including humanities, are male-dominated (e.g. only 31% of U.S. Ph.Ds. in philosophy are women¹). The underrepresentation of women in academia is thus a problem that is not limited to STEM fields. A better predictor of this underrepresentation, some have argued, is the belief that innate raw talent is the main requirement to succeed in the field (Leslie et al., 2015).

The level at which the evaluation takes place matters because stereotypes (or political views) can influence behavior differently if evaluators face already highly skilled applicants (as in Williams and Ceci 2015; Breda and Ly 2015) or moderately skilled ones (as in Ceci and Williams 2011; Ceci et al. 2014). By their mere presence among the pool of applicants for a high-level position, candidates signal their motivation and potential talent, whereas

¹National Science Foundation, Survey of Earned Doctorates (2011); www.nsf.gov/statistics/srvydoctorates/.

this is less true at a lower level, such as primary school teaching. Females who have mastered the curriculum, and who apply to high-skill jobs in male-dominated fields signal that they do not elicit the general stereotypes associating quantitative ability with men. This may induce a rational belief reversal regarding the motivation or ability of those female applicants (Fryer and Levitt, 2010), or a so-called “boomerang effect” (Heilman et al., 1988) that modifies the attitudes towards them. Experimental evidence provides support for this theory by showing that gender biases are lower or even inverted when information clearly indicates high competence of those being evaluated (Heilman et al., 1988; Koch et al., 2015).

To study how both female underrepresentation and candidates’ expected aptitudes can shape skills assessment, we exploit the two-stage design (written then oral tests) of the three national exams used in France to recruit virtually all primary-school teachers (CRPE), middle- and high-school teachers (CAPES and Agrégation), as well as a large share of graduate school and university teachers (Agrégation). A college degree is necessary to take part in those competitive exams. Except for the lower level (CRPE), each exam is subject-specific and typically includes 2 to 3 written and oral tests taken roughly at the same time. Importantly, oral tests are not general recruiting interviews: depending on the subject, they include exercises, questions or text discussions designed to assess candidates’ fundamental skills, exactly as are written tests. All tests are graded by teachers or professors specialized in the subject, except at the lower-level where a non-specialist sometimes serves on a 2-to-3 examiner panel along with specialists. 80% of evaluators at the highest-level exam (Agrégation) are either full-time researchers or university professors in French academia. The corresponding statistics is 30% at the medium level exam (CAPES).

Our strategy exploits the fact that the written tests are “blinded” (candidates’ name and gender are not known by the professor who grades these tests) while the oral tests are obviously not. Providing that female handwriting cannot be easily detected - which we discuss later -, written tests provide a counterfactual measure of students’ cognitive ability in each subject.

The French evaluation data offers unique advantages over previously-published experiments; they provide real-world test scores for a large group of individuals, thus they avoid the usual problem of experiments’ limited external validity. At the same time, these data

present a compelling “experiment of nature” in which naturally-occurring variations can be leveraged to provide controls. A final advantage is to draw on very rich administrative data that allow numerous statistical controls to be applied.

1.2 Institutional background

1.2.1 Competitive exams to recruit teachers in France

Teachers in France are recruited through competitive exams, either internally from already hired civil servants (usually already holding a teaching accreditation) or externally from a pool of applicants who are not yet civil servants. Candidates to private and public schools are recruited through the same competitive exams but they have to specify their choice at the time of the registration. The final rankings are distinct. We have data and therefore focus on the three competitive exams used to recruit teachers externally for positions in public schools or public higher education institutions (such as prep schools and colleges/universities, see below). More than 80% of all new teaching positions in France are filled with candidates that have passed one of these three exams.

1.2.2 Systematic non-anonymous oral and anonymous written tests

The competitive exams for teaching positions first comprise an “eligibility” stage in the form of written tests taken in April. All candidates are then ranked according to a weighted average of all written test scores; the highest-ranked students are declared eligible for the second stage (the eligibility threshold is exam-specific). This second “admission” stage takes place in June and consists of oral tests on the same subjects (see Table 1.1). Importantly, oral test examiners may be different from the written test examiners and they do not know what grades students have obtained on the written tests. Students are only informed about their eligibility for oral tests two weeks before taking them and are also unaware of their scores on the written tests. After the oral tests, a final score is computed as a weighted average of all written and oral test scores (with usually a much higher weight placed on the oral tests). This score is used to create the final ranking of the eligible candidates in order to admit the best ones. The number of admitted candi-

dates is usually equal to the number of positions to be filled by the recruiting body and is known by all in advance.

Competitive exams based on written and oral tests are very common in France: they are typically used to recruit future civil servants, as well as students in France's most prestigious higher education institutions (see details in [Breda and Ly 2015](#)). Each year, hundreds of thousands of French citizens take such exams. Historically, most of these exams only included oral tests or oral interviews, but the growing number of candidates over time led the exams' organizers to add a first stage selection of candidates that is based on written tests, which are less costly to evaluate than the second stage oral exams. These exams are thus widespread in French society, and something most candidates are familiar with.

1.2.3 Exams at three different levels

We exploit data on three broad types of exams: the Agrégation, the CAPES (Certificat d'Aptitude au professorat de l'enseignement du second degré) and the CRPE (Concours de Recrutement des Professeurs des Ecoles). As explained below, the Agrégation exam is partly geared toward evaluating potential candidates for professorial hiring.

Higher level exam: Agrégation

The most prestigious and difficult of those exams is the Agrégation. It has strong historical roots. For example, it dates back to 1679 in Law, 1764 in Arts, and started to spread to other fields in 1808. It is a field-specific exam, meaning that candidates take it in a given subject in order to get the accreditation to teach that subject only. Although there are roughly forty fields of specialization, a dozen of them comprise 80% of both positions and candidates. We focus exclusively on these dozen fields for the present study. Once candidates have chosen a particular subject, they are tested only in that subject, with the exception of a short interview aimed to detect their ability to "behave as an ethical and responsible civil servant" (see below).

Agrégation is highly selective and only well-prepared candidates with a strong background in their field of study have a chance to pass it. Even among those well-prepared candidates, admission rates are around 12.8% (See [Table 1.1](#)). Since the reform of 2011, candidates at Agrégation must hold at least a masters' degree (before that, the Maîtrise

diploma, which is obtained after four completed years of college, was sufficient).

Passing the Agrégation exam is necessary to teach in higher education institutions such as the selective preparatory school that prepare during two years the best high-school graduates for the competitive entrance exams to the French Grandes Ecoles (such as Ecole Polytechnique, Ecole Normale Supérieure, Ecole Centrale, HEC, etc.). They also give access to university full-teaching positions (PRAG). These positions are for example taken by PhDs who did not manage (yet) to get an assistant professor position. In total, about a fourth of the individuals who have passed Agrégation teach in postsecondary education.

Agrégation and CAPES holders both teach in middle and high-school. However, Agrégation holders are rarely appointed to middle schools and have on average much higher wages, fewer teaching hours, and steeper career paths in secondary education.

Although there is no official link between the Agrégation exam and academia, it is well-known that the two are related in practice. First, a large majority of examiners at Agrégation are full-time researchers or professors at university (see statistics in section 1.2.4). Then, on the candidates' side, holding the Agrégation can help for an academic career in some fields and a significant fraction of researchers actually hold this diploma. Conversely, according to the French association of Agrégation holders, about 15% of Agrégation holders who teach in high-school have a PhD. Some of the most prestigious higher education institutions, the Ecoles Normale Supérieure, select the best undergraduate students and prepare them for both a teaching and an academic career. Two of those three institutions command to all their students to take the Agrégation exam, even if they are only interested in an academic careers. The historical role played by the Agrégation and its rankings among the French intellectual elite might be best summarized by an anecdote. In 1929, Jean-Paul Sartre and Simone De Beauvoir both took and passed the philosophy Agrégation exam. Jean Paul Sartre was ranked first while Simone De Beauvoir was ranked second. Both became very famous philosophers and life partners. However many specialists considered that Simone De Beauvoir was scholarly better, and should have been ranked first instead of Jean-Paul Sartre. As a matter of fact, Sartre had already taken and failed this exam in 1928, while De Beauvoir got it at her first try. This illustrates the toughness of this exam, its informal links with academia (it is taken and graded by many (future) academics), and the fact that the patterns observed nowadays

in our data may have not always prevailed.

Medium level exam: CAPES

CAPES is very similar to Agrégation but the success rate is higher (23% against 12.8% for Agrégation, see Table 1.2) due to lower knowledge requirements. CAPES and Agrégation are not exclusive: each year, about 600 individuals take both exams. Only 4.4% of them pass Agrégation, whereas they are a much larger share (18.19%) to pass CAPES (see Table 1.2). Candidates at CAPES also need to hold a Master's degree or a Maîtrise. CAPES holders cannot have access to most positions in higher education and they teach exclusively in secondary education. Finally, and not surprisingly, CAPES is seen as less prestigious than Agrégation.

Lower level exam: CRPE

CRPE is exclusively aimed at recruiting non-specialized primary-school teachers. It is a non-specialized exam with a series of relatively low-level tests in a wide range of fields (maths, french, history, geography, sciences, technologies, art, literature, music and sport). In that sense it is very different from CAPES and Agrégation.

1.2.4 Two to three examiners at each test

All three exams include a series of written and oral tests. By law, each individual test needs to be graded by at least two evaluators. Written tests are usually graded twice, while the examination panel for each oral test typically includes three members, usually not of the same gender (even if it is sometimes hard to respect this rule for practical reasons). At the higher-level (Agrégation) and medium-level (CAPES) exams, examiners are always specialists in the exam field and they usually had passed the exam in the past (at least 50% of them). We collected data on the composition of the examiner panels for every field and exam level over the period 2006-2013. We found that evaluators are typically teachers in secondary or post-secondary schools (15% at the higher-level and 54% at the medium-level exam), full-time researchers, professors or assistant professors at the university (76% at the higher-level and 30% at the medium-level exam) or teaching inspectors (9% at the higher-level and 16% at the medium-level exam). They

know perfectly the program on which candidates are tested, and they grade the tests accordingly.

The lower-level exam is not field-specific but it includes both a written and an oral test in math and in literature since 2011. Each two-to-three examiners panel includes non-specialists and generally at least one specialist in the subject matter.

1.3 Data

The data used in these analyses belong to the French Ministry of Education and is made available on contractual agreement (which defines the conditions of access and use, and ensures confidentiality). The data provide information on every candidate taking the CRPE, CAPES and Agrégation exams over the period 2006-2013. For each and every exam, the data provides the aggregated scores of the candidates on the written and oral examinations. These scores are weighted averages of the scores obtained on all written and all oral tests (the weights are predefined and known by all examiners and candidates in advance). The aggregated score on written tests establishes a first-stage ranking of the candidates that is used to decide who is eligible to take the oral tests. After the oral tests, a final score is computed for eligible candidates as the sum of the oral and written tests aggregated scores. This final score is used to establish a second-stage ranking and decide which candidates are admitted. The data also include information on the socio demographic characteristics of the candidates, including sex, age, nationality, highest diploma, family and occupational status.

The detailed scores for the first six tests in each competitive examination (except for the period 2007-2010 for the CRPE, for which no detailed information is available) are also collected. The reason why only a subset of six test scores is available in addition to the total scores on the oral and written tests is that the Ministry of Education has arbitrarily formatted the data collected each year at each exam in a way that prevents storing more information. This arbitrary truncation implies that we miss some detailed scores in the exams that include more than six tests in total. In practice, between one (e.g. Mathematics) and five (e.g. Modern Literature) oral tests scores are missing for the high-level examination (see Table 1.17).

The data is exhaustive. In particular, it contains about thirty CAPES and Agrégation

exams in small subfields that we have not analyzed, either because the sample sizes are too small (e.g. 10 observations per year at the grammar Agrégation) or because they appear too atypical as compared to traditional academic fields (e.g. jewelry, banking, audiovisual). Out of the 20 different foreign or regional language CAPES and Agrégation exams, we have kept only the four main ones for which we have significant sample sizes (English, Spanish, German and Italian) and grouped them into one single field labeled "Foreign languages". Finally, in each field that we consider, we have retained in the analyses only candidates eligible for the oral tests who indeed took all written and oral tests². However, even after this data cleaning, the sample sizes are still very large: about 18,000 candidates at the Agrégation, 70,000 at the Capes and more than 100,000 at the CRPE. Descriptive statistics are provided in Tables 1.1 to 1.3. Most major academic fields are represented in our final sample (see Table 1.3).

For each competitive examination, candidates take between two and six written tests and between two and five oral tests, depending on the field. Even when they differ across fields, the way those tests are framed share similarities. In Mathematics, Physics and Chemistry, the written tests consist of problems, supplemented by a few questions, to assess the scientific knowledge of the candidate. In Philosophy, History, Geography, Biology, Literature and Foreign languages, the written tests systematically include an "essay". This exercise is very widespread in secondary education and in the recruitment of French civil servants. It consists in a coherent and structured writing test in which the candidates develop an argument based on their knowledge, sometimes using several documents. It is typically based on a general question or citation (Literature and Philosophy), a concept (History and Geography), a phenomenon (Economics and Social sciences), or a statement (Biology and Geology) that needs to be discussed.

Oral tests always include a "lesson". This is the case for all exams and in all fields. The "lesson" is a structured teaching sequence on a given subject. The presentation ends up with an interview in which the examiners challenge the candidate's knowledge and, to some extent, her pedagogical skills. The "lessons" in mathematics and literature were only added to the CRPE after the 2011 reform.

Finally, a test entitled "Behave as an Ethical and Responsible Civil Servant" (BERCS) was introduced in 2011 for all three levels of recruitment (CRPE, Capes, Agrégation). It

²A small fraction of the candidates eligible for the oral test do not take them because of illness, or because they already accepted another position and are no longer interested.

consists of a short oral interview. In the medium- and high-level exams, this interview is a subpart of an oral test that otherwise attempts to evaluate competence in the core subject. It is consequently graded by teachers or professors specialized in the core subject. In the lower-level exam, it is graded as a subpart of the literature test. We only have data on detailed scores on the BERCS test at the lower- and medium-level exams. A description of all tests, all exams and all fields is provided in Tables 1.16 and 1.17.

1.4 Method

1.4.1 Percentile ranks

Oral and written tests are usually scored between 0 and 20. We use the empirical cumulative distribution of the scores for each test, meaning that we transform them into percentile ranks. The percentile rank corresponding to the worst score is 0, while that of the best score is 1. The percentiles are computed by including only candidates eligible for the oral test who indeed took all written and oral tests.

We conduct this transformation for two reasons. First, we focus on a competitive exam for which candidates are not expected to achieve a specific score, but only to be ranked for the predefined number of available places. As only ranks matter in this hiring exams, interpreting our results in terms of gains or losses in rankings makes sense. Second, the initial test score distributions for the written and oral tests are very different. This is because our sample contains only the best candidates upon completion of the series of written tests, all of whom tend to get good grades on these written tests. However, examiners expect a higher average level from these candidates on oral tests, and try to use the full spread of available grades in their marking, such that the distribution of scores in the oral tests has a lower mean and is more spread out between 0 and 20. Transforming scores in percentile ranks is the most natural way of keeping only the ordinal information in an outcome variable and to avoid meaningless quantitative (or cardinal) differences between the units of interest, hence avoiding the possibility that comparisons could reflect the magnitude of these meaningless quantitative differences.

1.4.2 Variations in percentile ranks between oral and written tests (DD)

The main statistics of interest is the difference between women's average percentile ranks on oral and written tests, minus the same difference for men's. This statistics DD can take all values between -1 and 1, no matter the actual share of women among candidates. It is thus comparable across fields with varying shares of female candidates. To see this, note that the average ranking r_F^W and r_F^O that women can get on written or oral tests depends on their p_F is their proportion p_F among the pool of candidates in a given subject. Looking at the 2 extreme cases where females are all ranked above or below the males on written or oral tests, we get:

$$\left\{ \begin{array}{l} \frac{p_F}{2} \leq r_F^O \leq 1 - \frac{p_F}{2} \\ \frac{p_F}{2} \leq r_F^W \leq 1 - \frac{p_F}{2} \end{array} \right. \quad \text{which implies that } -(1 - p_F) \leq r_F^O - r_F^W \leq 1 - p_F$$

Similarly the difference $r_M^O - r_M^W$ between men's average percentile ranks on oral and written tests is also bounded between $-p_F$ and p_F . Combining the bounds for females and males average ranks, we directly get $-1 \leq DD = (r_F^O - r_F^W) - (r_M^O - r_M^W) \leq 1$

Furthermore, it is straightforward to check that the bounds -1 and 1 are indeed attained in the extreme cases where females are all ranked above or below the males. Note that a "simple" difference between women's average percentile ranks on oral and written tests would have bounds that vary according to p_F . For example, if there were (almost) only women, such a difference would be 0, it would vary between -0.5 and 0.5 if there were 50% women, and between -1 and 1 if there were (almost) only men. Our choice to normalize by the rank difference for men is therefore designed to avoid the magnitude of the estimated effects to vary across contexts. To check that it is indeed the case, we have ran simulations in which evaluation biases of the same magnitude occur on oral tests in samples with various shares of women and men. These simulations confirm that DD converges to the same value, regardless of the proportion of women among the candidates.

In terms of interpretation, a variation of 0.1 of DD is compatible, for example, with the following scenarios: 1) all the women overtake 10% of the men between the oral and the written tests and 2) 10% of the women overtake all the men between the oral and the written tests.

1.4.3 Odds ratios and relative risks

To assess to what extent oral tests improve or decrease women's chances of passing the exam, we compare what would have been their admission rates if admission had been based on written tests only, or if it had been based on oral tests only. Odds ratios and relative risk measures are computed to compare the two cases.

1.4.4 Using total scores on written and oral tests or keeping only one written and one oral test

At the medium- and high-level exams in a given field (e.g. math, philosophy), candidates take more than one written test and more than one oral test in the subject corresponding to the exam field. To avoid arbitrary selection of some tests over other ones, the main analysis is based on comparisons of the candidates' aggregated scores on oral tests and on written tests. These scores are weighted averages based on all tests. However, we also reproduce the main results keeping only one written test and one oral test for each medium- and high-level field-specific exam. We have tried to keep the pairs of tests that match most closely in terms of the underlying subtopic or test program on which they were based (see Figure 1.7). We implement this alternative approach to make sure the baseline results are not driven by oral or written tests that are too different to be really comparable (such as the oral test "behave as an ethical and responsible civil servant" introduced in 2011, that has no written test counterpart - but a very small weight in the oral tests aggregated score).

1.4.5 A simple linear model to derive econometric specifications

Suppose that the written tests measure the ability θ_{1i} of individual i with error ϵ_{iw} and that oral tests measure the ability θ_{2i} with error ϵ_{io} . Suppose also that examiners have a gender bias β in favor of women.

Then the scores $Score_i^{Written}$ and $Score_i^{Oral}$ obtained by individual i at the written and oral tests are given by:

$$\begin{cases} Score_i^{Written} = \theta_{1i} + \epsilon_{iw} \\ Score_i^{Oral} = \theta_{2i} + \beta F_i + \epsilon_{io} \end{cases}$$

with F_i a dummy equals to 1 if individual i is a woman, $E[\theta_{1i} \epsilon_{iw}] = 0$, $E[\theta_{2i} \epsilon_{io}] = 0$ and $E[F_i \epsilon_{io}] = 0$

Suppose additionally that abilities θ_{i1} and θ_{i2} are linearly related in the following way: $\theta_{i2} = \rho\theta_{i1} + \nu_i$ where ν_i is an ability component that is exclusively measured on the oral tests and that is independent of θ_{i1} . Then, we derive the relation between the oral and written scores: $Score_i^{Oral} = \rho Score_i^{Written} + \beta F_i + (\epsilon_{iw} + \nu_i - \rho\epsilon_{io})$ (1)

1.4.6 Statistical models used to assess the gender bias on oral tests in each field and at each level

We now lay down the statistical models used to estimate evaluation biases at each exam. Technical discussions are presented here, while the estimation results are left for the next section.

Model DD

Linear regression models are used to check the robustness of the DD statistics (see Tables 1.18 and 1.19) to the inclusion of control variables and to alternative specifications. Such models are also used to statistically assess if the positive relationship between subjects' extent of male domination and female bonuses on oral tests is larger at the higher level (Agrégation) than at the medium level (CAPES).

For each subject and for each exam, a difference-in-difference estimator of the gender bias β can be computed from a DD model of the form: $\Delta Rank_i = \alpha + \beta F_i + \epsilon_i$ where $\Delta Rank_i = Rank_i^{Oral} - Rank_i^{Written}$ is the variation in rank between oral and written tests of candidate i , F_i an indicator variable equal to 1 for female candidates and 0 for males, and ϵ_i an error term.

Coefficients β estimated from those models in each subject-specific medium- and high-level exam are reported in column DD1 in Tables 1.4 and 1.5. Coefficients β estimated in math and literature at the lower-level general exam are reported in column DD1 in Table 1.11.

We then check that results are robust to the inclusion of control variables for candidates' characteristics (age, month of birth, education, department of residence, and nationality) and examinations' characteristics (year and region for the lower-level exam

implemented at a regional and decentralized level) by estimating the following model: $\Delta Rank_i = \alpha + \beta F_i + \gamma X_i + \epsilon_i$ See column DD2 in Tables 1.4, 1.5 and 1.11.

Note that the difference-in-difference (DD) model is widely used to study discrimination. It is the empirical counterpart of model (1) when test scores have been transformed into percentile ranks and when ρ is assumed to be equal to 1.

Model S

Estimates of the coefficient of interest β obtained from the DD model can be biased if both $\rho \neq 1$ and $E[F_i Score_i^{Written}] \neq 0$. To see this, we use (1) to re-write the DD model: $\Delta Rank_i = \alpha + \beta F_i + \tau_i$, with $\tau_i = \epsilon_{iw} + \nu_i - \rho\epsilon_i + (\rho - 1)Rank_i^{Written}$

To address this possible issue, we move to an alternative specification (S) where gender differences on oral tests are estimated conditional on the rank on written tests.

$$Rank_i^{Oral} = \alpha + \beta F_i + \gamma Rank_i^{Written} + \delta X_i + \epsilon_i \quad (2)$$

This model (S), estimates consistently the coefficient β without any assumption on ρ . Results are presented in column S in Tables 1.4, 1.5 and 1.11. Estimates without control variables are not presented but are very similar. To control more flexibly the relationship between written and oral test ranks, we replaced the linear control by a third order polynomial in the written test ranks, or even a set of dummies for different possible written test ranks. Results in that case can be understood as the bonuses obtained on oral tests by women among candidates who got almost exactly the same written test score. Those results are not presented but are very similar to those obtained in column S.

Model S+IV

Model S is more general than model DD as it allows the weight of the candidates' unobserved abilities to be different on oral and written tests. However it has a well-known caveat (see Wooldridge 2002, section 4.4): if the written test score is a noisy measure of candidates' unobserved ability (i.e. $\epsilon_{iw} \neq 0$), then the estimates of the bonus on oral tests for women are likely to be biased. Intuitively, this is because the candidates' differences in ability that are not captured by the noisy written test score can in that case also be captured by gender. To put it differently, gender can play the role of a second imperfect measure of ability that will complement the noisy written test score. This will happen if

abilities are not identically distributed across gender. Formally, if the error term ϵ_{iw} is different from 0, it is mechanically correlated with $Rank_i^{Written}$ in (1), implying that β cannot be consistently estimated with equation (2) when both genders do not have the same abilities in average. For this reason, and because test scores are usually assumed to be noisy measures of ability, applied econometricians tend to favor model (DD) over model (S). There is no practical way, however, to decide which of the two issues accruing with the empirical equations DD and S is empirically the most problematic.

A way to avoid both issues is to instrument the written test rank by an alternative measure of candidates' ability (see [Wooldridge 2002](#), section 5.3.2) when estimating equation (1). Results are presented in column S+IV in Tables 1.4, 1.5 and 1.11. Those results and our choice of instruments are discussed in detail in section 1.6.3.

Note that to consistently estimate the gender bias β on oral tests with model (S+IV), we still need to assume that the oral-specific ability component ν_i is not correlated with gender: $Cov[\nu_i, F_i] = 0$. This is the key assumption behind our strategy: all skills that are specific to oral tests and cannot be captured with written tests should not vary systematically with gender. Otherwise, the gender bias on oral test could simply reflect those differences. We discuss this further in section 1.6.2.

1.4.7 Using initial scores instead of percentile ranks

A drawback with the use of percentile ranks is that it imposes some algebraic constraints. For example, the weighted average of women's and men's percentile ranks has to be equal to 0.5. This can lead to an under-estimation of standard errors when they are based on all candidates, as observations are redundant (the variation in ranks for men can be entirely deduced from the variation in ranks for women). To check that this issue does not alter the significance of the results, we re-estimate all models using the initial candidates' total scores on oral and written tests. The magnitude of the coefficients is then harder to interpret, but their significance remains unchanged.

1.4.8 Statistical model to assess how the gender bias on oral test varies from a subject to another one

We estimate the relationship between subjects' extent of male-domination and female bonuses on oral test directly from regression models of the type:

$$\Delta Rank_{ij} = \alpha_j + \beta F_i + \gamma(S_j.F_i) + \epsilon_{ij} \quad (3)$$

where j is a subscript for subjects and S_j the share of women in academia in subject j . The intercept β and the slope γ are the coefficients of interest that are estimated both at the medium and high-level exams. Estimates obtained using the 3 different measures of subjects' feminization described in Table 1.6 are summarized in Table 1.7.

1.4.9 Statistical model to assess how the relationship between subjects' extent of male-domination and gender bias on oral test varies between the medium- and the high-level exams

In order to get a valid statistical comparison of the medium- and high-level exams, we nest them in a single regression model and estimate:

$$\Delta Rank_{ijl} = \alpha_{jl} + \beta_m(F_i * M_l) + \gamma_m(S_j.F_i * M_l) + \beta_h(F_i * H_l) + \gamma_h(S_j.F_i * H_l) + \epsilon_{ijl}$$

where l is a subscript for the exam level (high or medium) and M_l (resp H_l) is an indicator variable equal to 1 if candidate i is observed at the medium-level (resp high-level) exam.

The estimates obtained for the intercept β and the slope γ at the medium- and high-level obtained with this specification are by definition equal to those obtained with equation (3). For the 3 different measures of subjects' feminization described in Table 1.6, we perform a Chow test of equality between, on the one hand β_m and β_h , and on the other hand γ_m and γ_h . Results of those tests are summarized in Table 1.7.

1.4.10 Clustering standard errors

Standard errors can be correlated for two reasons: 1) Candidate-specific unobserved characteristics can correlated error terms across candidates' test scores and 2) Systematic grading behaviors from examiners and the specific content of each test can correlate error terms within tests.

The first point is to a large extent dealt with by using ranks based on total scores. This implies that we keep only one observation per candidate in the main analysis. This aggregation of the scores leads to a loss of statistical power. However, it avoids any serial correlation in the error terms coming from the use of several oral or written tests for a given candidate ³.

To deal with the second point and compute correct standard errors for β and γ , it is necessary to allow the error terms ϵ_i to be correlated within each cell defined by a type of subject and a given year. We thus cluster standard errors at the year x subject level.

This level of clustering is conservative regarding error correlations that are due to the similar evaluation biases within examiner panels. Indeed each cluster includes many examiner panels. For example, our sub-analysis of the math medium-level exam (for which we have more detailed data) reveals that 48 examiner panels evaluated the oral tests at that exam in 2013. However, errors can also be correlated because of the specific content of a written test for example, which is common across all the examiners panels that are grading the test. Finally, a significant fraction of candidates take both the oral and written tests of CAPES and Agrégation in a given subject, leading to possible error terms correlations across examination levels. To deal with this (which relates to the first point above), we systematically include CAPES and Agrégation in the same cluster for a given subject and year. At the end, we build quite large clusters, but the number of subjects (9) and years (8) is also large enough to have 72 distinct clusters and still get significant results while clustering at a broad level.

1.5 Results

1.5.1 Gender differences between oral and written test scores at exams to recruit secondary school and postsecondary professorial teachers

To assess gender bias in evaluation, we focus on candidates who took all oral and written tests, and rank them according to their total score. We then subtract the difference in

³The only remaining source of error correlation due to the candidates comes from the retakers that are observed two consecutive years. Those can easily be dealt with by simply removing the retakers, which does not affect much the results.

male candidates' average percentile ranks between written and oral tests to the same difference for female candidates. This standardized measure is bounded between -1 and 1, and it is independent of the share of females among the total pool of applicants. It is equal to 1 if all women are below the men on written tests and above them on oral tests. For each subject-specific exam, we computed this measure and its statistical significance using a linear regression model of the type $\Delta Rank_i = a + bF_i + \epsilon_i$. $\Delta Rank_i$ is the variation in rank between oral and written tests of candidate i , F_i is an indicator variable equals to 1 for female candidates and 0 for males, ϵ_i is an error term and b is the measure of interest.

In fields in which women are underrepresented (mathematics, physics, chemistry and philosophy), oral tests favor women over men both on the higher-level (professorial and high-school teaching) and medium-level (secondary school teaching only) exams (Figure 1.1, $P_s < 0.001$ in all cases, see sample sizes in Table 1.3 and detailed results in Tables 1.4 and 1.5). In contrast, but to a lesser extent, oral tests in fields in which women are well-represented (literature and foreign languages) favor men over women (Figure 1.1, $P_s < 0.001$ in both cases, see Tables 1.4 and 1.5). In history, geography, economics (which also includes some other social sciences tests) there are only small gender differences between oral and written tests. Those differences are not significantly different from 0 at the 5% statistical level. In biology, a bias against women is found, but on the high-level exam only. All results are robust to the inclusion of control variables and to the use of two alternative statistical models S and S+IV.

A simple explanation for these results would be that examiners on oral tests try to lower gender differences in ability observed on written tests. Figure 1.2 shows that this is not always the case. Bonuses on oral tests are observed in fields where both genders had similar rankings on the written tests (philosophy, chemistry at the highest-level, classical literature at the medium-level). More strikingly, in cases where there is a significant ranking gap between women and men on written tests, the oral test may even fully invert this gap (physics at the highest-level, math at the medium-level).

1.5.2 A clear pattern of rebalancing gender asymmetries in academic fields, strongest at the highest-level exam, and invisible at the lower-level exam

A clear pattern emerges from Figure 1.1: the more male-dominated a field is, the higher the bonus for women on the non-blind oral tests. To formally capture this pattern, we study how the bonus b on oral tests varies with the share of women s among assistant professors and senior professors in the French academy (we also consider other measures of fields' feminization, see Table 1.6). We find a significant negative relationship at both the higher- and the medium-level exams (see Figure 1.1: $b = 0.19 - 0.42s$ at the high-level exam ; $b = 0.16 - 0.30s$ at the medium-level exam, with $Ps < 0.0001$ for both slopes and intercepts of the fitted lines).

The relationship between the extent of a field's male-dominance and female bonuses on oral tests is about 50% larger at the highest level exams (for high-school teachers and professorial). At that level, switching from a subject as feminine as foreign languages ($s = 0.62$) to a subject as masculine as math ($s = 0.17$) leads female candidates to gain on average 22 percentile ranks on oral tests with respects to written tests. To avoid sample selection bias, this comparison between the medium- and the high-level exam is made on a subsample of about 5,000 individuals that have taken both exams in the same subject the same year (see Table 1.7).

The statistical analysis also reveals an absence of large significant gender biases on oral tests at the lower-level teaching exam. Importantly, this exam is not subject-specific. However, since 2011, all applicants have been required to take an oral and a written test both in math and literature, which make it possible to study the bonus on oral tests for women in those two subjects. We only find a small premium of around 2 percentile rank for women on oral tests, both in math and literature, with no clear difference between those two subjects (see Table 1.11). This finding underscores the importance of distinguishing between selection processes for primary school teachers vs. secondary school teachers and college/university professors.

1.5.3 Implications for the gender composition of recruited teachers and professors in different fields

Given that at each level and in each subject there is a predetermined number of possible hires, the differences in rankings between oral and written tests are likely to influence admission and hiring rates. We compared the likelihood of being hired for women if admissions were based either only on rankings on oral tests, which are non-blind, or only on rankings on written tests, which are gender-blinded. We computed the corresponding relative risk and odds ratio of admissions (Figure 1.5 and Table 1.8). A similar pattern is observed: the probability of admission of women increases by up to 10ppt in the least feminized fields - math, physics and philosophy - on oral tests compared to written tests. This increase is systematically larger at the highest-level exam used to recruit professors and highly qualified teachers. In contrast, women have a significantly lower probability of admission on oral tests compared to written tests in feminized fields, like literature or foreign languages, mostly at the highest-level exam.

Those patterns were found to be remarkably stable across the written ranks' distribution (Tables 1.9 and 1.10), indicating that they concern all candidates, and not only those ranked around the hiring threshold. They are also visible at the most prestigious top ranks, with twice more (resp. 30% less) women ranked first on the oral tests than on the written ones in mathematics, physics, chemistry or philosophy (resp. in literature and foreign languages) at the highest-level exams.

1.5.4 Gender of evaluators

Evaluation biases could reflect an opposite-sex preference by which male evaluators who are more numerous in male-dominated subjects favor female candidates and vice versa. Data on the gender composition of each specific examiner panel is available for the math medium-level exam in 2013. Analysis on this subsample reveals that the gender gap between oral and written test scores is not impacted at all by examiners' gender (Table 1.13). This is in line with previous research (Williams and Ceci, 2015; Breda and Ly, 2015; Bagues et al., 2017) that also reported that the pro-female bonus in academic hiring does not depend on the raters' gender. This suggests that context effects (surrounding gender stereotypes) are much more important than examiners' gender in explaining gender biases

in evaluation.

1.5.5 Comparison of an oral test that is common across all exams

To better understand the origin of the gender biases on oral tests, we exploit a remarkable feature of the teaching exams: since 2011, all of them have included an oral test entitled "Behave as an Ethical and Responsible Civil Servant" (BERCS). BERCS is the only test that is not subject-specific⁴ but is still evaluated by specialists.

Comparisons of gender differences in performance on this oral test across subjects at the medium-level exam reveals that women systematically rank better, and that this bonus b' decreases with the share of women s in the overall field (Figure 1.3, $b' = 0.14 - 0.29s$, with $Ps < 0.0001$ for both the slope and the intercept). This pattern is similar to what is observed in Figure 1.1 and 1.2 using tests based on subject-specific skills, and suggests that examiners favor women who chose to specialize in male-dominated subjects no matter what they are tested on.

However, as we do not have a blind counterfactual measure of ability for the BERCS test, the pattern in Figure 1.3 could also reflect that women who enrolled in the more male-dominated fields have better aptitude for that particular test than women who enrolled in other fields. To refute this interpretation, we used the grade obtained on this oral test at the lower-level exam as a neutral measure of ability for the few candidates who took the same year both the lower-level exam and one of the 9 subject-specific medium-level exams (118 candidates). As the lower-level exam is not subject-specific, it offers a counterfactual measure in a gender-neutral context. Among the group of candidates who took the medium-level exam in a less male-dominated subject (economics, history, geography, biology, literature, foreign languages), men get a significant (at the 10% level) advantage over women on the oral test BERCS at the medium-level compared to what they get at the lower-level exam (see Figure 1.4, $P = 0.091$). The reverse is true (however not statistically significant) among the group that took the medium-level exam in a male-

⁴We check that candidates' score at the test "behave as an ethical and responsible civil servant" for the computation of candidates' rank on oral tests do not impact the main results by restricting the analysis to the period before it was implemented in 2011. We also replicated the analysis keeping only one oral and one written test in each of the middle- and high level exams. We kept the pairs of tests that match the most closely in terms of the subtopic or test program on which they were based. Results are virtually unchanged (see Figure 1.7 and Table 1.15).

dominated subject (math, physics or philosophy).

1.6 Discussion

In natural experiments, the researcher does not have full control on the research design, thus the results usually need to be interpreted with caution. The setting we exploit has two main potential issues: gender may be inferred on written tests from handwriting, and there might be gender differences in the types of abilities that are required on oral and written tests, even on a similar topic based on the same program. We discuss those issues now, before presenting in detail the results of the statistical analysis.

1.6.1 Handwriting detection

Former tests that we conducted have shown that the rate of success in guessing gender from hand-written anonymous exam sheets is on average 68.6% (Breda and Ly, 2015). This percentage is significantly higher than the 50% average that would be obtained from random guess. It is nevertheless closer from random guess than from perfect detection (100%).

To examine to what extent some handwriting could be unambiguously detected, we asked five different assessors to guess the gender of each exam sheet. A joint analysis of their answers reveals that for about a quarter of the exam sheets (26%), the gender of their writer is incorrectly guessed by a majority of assessors (at least 3 out of 5), suggesting that examiners are often uncertain about the candidates' gender on written tests. However, the joint analysis also reveals that in 39% of cases, all five evaluators make correct guesses.

The ability of examiners to detect the gender of some candidates at the written tests with a relatively high degree of confidence could be problematic regarding the interpretation of the paper's results if and only if those examiners are biased in opposite directions on the written and oral tests. In contrast, if evaluators are biased the same way on oral and written tests, the comparison of the two should not lead to large systematic observable differences.

We may also argue that being ambiguously exposed to a presumably female or male handwriting is a much weaker treatment than being exposed to a female or male candidate in the flesh that occurs during an oral test. Hence, partial gender detection on written

tests is likely to attenuate the magnitude of the estimated biases, while still identifying them, unless evaluators have opposite gender biases at oral and written tests. This later hypothesis cannot be rejected empirically but seems unlikely because the same examiners evaluate both the written and oral tests it is hard to think they change their attitude between the written and oral parts.

A last point is that the analysis of the BERCS test described in the next section only relies on comparisons of one oral test across exams' subject and levels. The sensitivity analysis done with the BERCS test is therefore not subject at all to handwriting detection problems and offers an alternative confirmation that our baseline results are not reflecting gender-driven grading behaviors going in opposite direction at oral and written tests.

1.6.2 Gender differences in the types of abilities that are required on oral and written tests

A more fundamental issue is that the gap between a candidate's oral and written test scores in a given subject can capture the effect of gender-related attributes that are visible only on oral or written tests, such as the quality of handwriting or elocution (see 5-8 for surveys on possible sex differences in cognitive abilities, including verbal fluency). If such attributes directly impact test performance, they can undermine the results. In the framework of the formal model in section 1.4.5, those attributes are captured in the term ν_i . If ν_i varies systematically with gender, the gender bias on oral tests cannot be identified, and our results could simply reflect gender differences in the skills that are specific only to the oral or written tests.

The first defense against those alternative interpretations is that our key result is not the absolute gender gap in the oral versus written test score in a given subject, but the variation - and even reversal - of this gap across subjects revealing a systematic pattern. If there are gender-specific differences in abilities between oral and written tests, these differences would need to vary across subjects to explain our results. We now discuss and reject this idea.

A first reason why the present results could reflect skill differences is that the populations tested in the different subjects are not the same, but selected themselves. That is, the women who decided to study math and take the math exams might be especially self-confident in math and perform better on oral tests in math for this reason, whereas

the same self-selection happens for men in literature. There is evidence refuting this argument that sample selection drives the results: on the high-level exam in Physics-Chemistry, the same candidates have to take oral and written tests both in Physics and Chemistry. Among these candidates, the bonus for women on oral tests is 9 percentile points larger in physics than in chemistry, a subject that is less male-dominated according to all indicators (see Table 1.6). The idea that sample selection does not drive the general pattern in Figure 1.1 is also confirmed by a former analysis that is entirely based on identical samples of candidates being tested on different subjects (Breda and Ly, 2015).

Formally, doing the analysis on a single sample of candidates implies that we can allow for each applicant to possess different abilities on oral and written tests. To interpret the variation across subjects of the female bonus on oral tests as an evaluation bias, it is only necessary to assume that the differences between oral and written abilities do not vary systematically across subjects by gender. One could argue that this assumption was violated in some cases: handwriting quality or elocution might be more important in some subjects than others, or perhaps the oral tests in the most male-dominated subjects are framed in a way that makes more visible the qualities that are more prevalent among women. Results obtained on the BERCS test (Figures 1.3 and 1.4) fully refute those possibilities. They indeed reveal that the gender bias according to the gender incongruity of the evaluation context persists on this BERCS test that is common across all contexts. Those results cannot be attributed to differences in the skills required for the test (Figures 1.3 and 1.4), neither to the selection of candidates across contexts (Figure 1.4). As both the test subject and the sample of candidates are held constant in the experiment presented in Figure 1.4, observed differences almost surely reflect examiners' bias according to the extent of male-domination in the candidates' field of specialization. The only alternative hypothesis would be that the candidates pay different efforts when evaluated at the low-level and medium-level exams, and that these differences vary according to gender and the field of specialization chosen at the medium-level exam. As the tests are relatively short but usually require a long preparation, it seems unlikely that the candidates who have already trained for the tests do not pay maximal effort during it.

1.6.3 Results from statistical models DD, S, and S+IV at the medium- and higher-level exams

Tables 1.4 and 1.5 present the results obtained from three different statistical models. Model DD is estimated without any control variable (DD1) or with control variables (DD2). Comparisons of columns DD1 and DD2 shows that the inclusion of control variables for candidates' age, month of birth, nationality, county of residence, and education has only a small effect on the subject-specific gender biases. This is consistent with the idea that systematic (gender) differences between oral and written test scores capture evaluation biases due to gender rather than other types of biases (due to the other control variables), or variations in candidates' ability between oral and written tests. Indeed, if candidates' ability varies between oral and written tests, one might think that the inclusion of controls would capture part of this variation, which would not be captured anymore by the gender indicator.

Estimates obtained from model S are sometimes quite different from those obtained from model DD. However, the general pattern of higher bonus on oral tests for females in more male-dominated subjects can still be observed with model S at both the high- and medium-level exams.

In all cases, model S is subject to measurement error bias (see methods). It is thus probably better to focus on the model S + IV as long as this model passes the usual tests for the validity of the instruments. This is considered to be the case when the F statistics of the test of weak instruments is above 15, and the p-value of the Sargan ⁵ test is above .05. When we use as instruments the candidates' year and month of birth, those conditions are satisfied in all subjects but economics at the medium-level exam, and in foreign languages, biology, physics and math at the highest-level exams. Reassuringly, the estimates obtained in those subjects where the instruments are statistically valid also exhibit the central pattern of a larger bonus on oral tests for females in more male-dominated subjects.

Note that the month of birth is a standard instrument in the economics of education literature (Angrist and Krueger, 1991; Terrier, 2016). However, the statistical analysis

⁵The Sargan test is used for testing the exogeneity of all the instruments when at least two are available to the econometrician, and one is assumed to be exogenous. Under the null assumption, instruments are exogenous.

revealed that it is necessary to use a second instrument to increase the strength of the instruments and pass the Fisher test. Our choice of using age as a second instrument comes from the fact it is a good proxy for that experience, which itself impacts competence positively. A concern, however, is that age might be visible and lead to evaluation bias during oral tests. This would violate the exclusion restriction. The fact that Sargan tests do not reject the exogeneity of the instruments in most cases is reassuring in that respect: assuming, as it is usually the case, that month of birth is a valid instrument (which is a standard assumption), we cannot reject that age is also valid.

A careful examination of the estimates reveals that those obtained using the S + IV model are often very close (and never statistically different) from those obtained with the DD models (DD1 or DD2). This suggests that the additional restriction imposed in the DD model (that a latent ability parameter impacts ranks at the oral and written tests to the same extent) is valid. We investigate this more formally by testing in the S + IV model if we can reject that the correlation between the written and the oral rank is equal to 1. In most subjects, we cannot reject this assumption, which is exactly the one that is made in the DD model (which is formally equivalent to an S model where the effect of the written test score on the oral test score is restricted to be equal to 1). It can also be observed that the correlation between the written and the oral rank jumps up between the S model and the S + IV model. This is consistent with the idea that measurement error is a quantitatively important issue in the S model.

To conclude: all models support the pattern of a higher bonus for females on oral tests in more male-dominated subjects; the S + IV model suggests that the DD model should be preferred over the S model.

1.6.4 Results from statistical models DD, S, and S+IV at the lower-level exams

Results at the lower-level exam are presented in Table 1.11. Both in math and literature, the instruments used at the medium and high-level exams (age and month of birth) do not pass the Sargan test of overidentification, leading us to discard them. Instead, we take advantage of the large sample size at the lower-level exam and restrain the analysis to individuals who took the exam two consecutive years (and have therefore failed during the first year). For those candidates, the written test score obtained the second year is

instrumented by the written test score obtained the first year. This instrument is certainly a more direct and better alternative measure of ability. It also has the advantage to be unobserved on oral test a given year (contrary to candidates' age that is partly visible), so that it cannot have any direct effect on the oral test score (a necessary assumption for the theoretical validity of an instrument) .

Results using the previous year written test score as instrument are presented in column S + IV2 in Table 1.11. The Fisher test of weak identification confirms that this instrument is very strong at the lower-level exam. We see that the hypothesis that the correlation between the written and the oral rank is equal to 1 is strongly rejected, both in math and in literature. The direct implication of this is that the DD model is no longer valid at the lower-level exam. This is also visible in Table 1.12 that re-estimate the DD model after splitting the sample in five quintiles: estimates obtained there are always smaller in math and larger in literature than those obtained on the full sample, which should not happen if the DD model where valid. Focusing instead on the S model, or better, on the S + IV model, we see that women obtain small bonuses on oral tests of about 2%, both in math and literature. The weak correlation between the written and the oral rank - almost null in math and around 0.26 in literature - suggests however that the abilities measured by written and oral tests differ substantially and that the estimated coefficients β should be considered carefully.

1.6.5 Analysis of the effect of the gender composition of the examiner panels

Table 1.13 presents estimates from the following model:

$$\Delta Rank_{ipj}^{Oral} = \alpha_i + \mu_j + \beta N_{pj} + \gamma(F_i * N_{pj}) + \delta X_p + \epsilon_{ipj}$$

where N_{pj} is the number of women in the three-people examiner panel p that evaluated candidate i on oral test j .

The analysis is only run at the math medium-level exam in 2013, the only one for which we have detailed information on the actual interviewers of each single candidate. As candidates take two oral tests, we can include in the model individual α_i and oral tests fixed effects μ_j (model 1 in Table 1.13). The model is thus identified within candidates, i.e. from variations in a candidate's ranking between two oral tests according to the number of women in the examiners' panel at each of the tests. We can also control for

the average observable characteristics X_p of the members of a given examiner panel (main employment position and county of residence). This is done in model 2.

Those controls for panels' characteristics can also be replaced with fixed effects for examiners' panels as in the following equation: $\Delta Rank_{ipj}^{Oral} = \alpha_i + \mu_j + \delta_p + \gamma(F_i * N_{pj}) + \epsilon_{ipj}$

This specification captures unobserved heterogeneity in grading behavior across panels. It is estimated in model 3. The estimated effect of the number of women in the examiners panels on the female candidates test scores are very similar across models and never significantly different from 0 from a statistical point of view.

1.7 Conclusion

In natural experiments, the researcher does not have full control on the research design, thus the results usually need to be interpreted with caution. The data we exploited in these analyses have two potential caveats: gender may be inferred on written tests from handwriting, and there might be gender differences in the types of abilities that are required on oral and written tests, even on a similar topic based on the same program. Based on the paper's evidence - in particular the results at the BERCS test that is common across exams -, neither of these alternative hypotheses is likely to explain the results. Instead, a gender incongruity effect appears to rebalance gender asymmetries in academic fields by favoring the minority gender. For women, this runs counter to the claim of discrimination in recruitment of professors into math-based fields. If anything, women appear to be advantaged on tests by both male and female evaluators. In contrast, men appear to be advantaged in recruitment into the most feminized fields. Those behaviors are not driven by a policy of affirmative action, totally forbidden in scoring these exams. They are also strongest on the highest-level exam, where candidates are more skilled, and where initial gender imbalances between the different fields are largest.

Even if they may not generalize to all recruiting contexts, the present results shed light on the possible causes behind the underrepresentation of women in many academic fields. They confirm evidence from a recent testing ([Williams and Ceci, 2015](#)) that women can be favored in male-dominated fields at high recruiting levels (from secondary school teaching to professorial hiring), once they have already specialized and heavily invested in those fields (candidates on teaching exams hold at least a college or a masters degree)

⁶. In contrast, the study of the recruiting process for primary school teachers shows that pro-women biases in male-dominated fields can disappear in less prestigious and less selective hiring exams, where candidates are not necessarily specialized. Perhaps the bias in favor of women in male-dominated fields would even reverse at lower recruiting levels, as in experiments done with medium-skilled applicants (Moss-Racusin et al., 2012; Reuben et al., 2014). Discrimination may then still impair women's chances to pursue a career in quantitative science (or philosophy), but only at early stages of the curriculum, before or just when they enter the pipeline that leads to a PhD or a professorial position.

However, there is no compelling evidence of hiring discrimination against individuals who already decided against social norms to pursue an academic or a teaching career in a field where their own gender is in the minority. Perhaps the knowledge that they have at least as good an opportunity as their male counterparts at the levels of secondary school teaching and professorial recruiting would encourage talented young women to study in male-dominated fields. Active policies aimed at counteracting stereotypes and discrimination should focus more on early ages, before educational choices are made.

⁶The higher-level teaching exam is held by a significant fraction of researchers and may in some cases accelerate a career in French academia. In that sense, results obtained on this exam can be seen as more closely related to the specific debate on the underrepresentation of women scientists in academia.

Bibliography

- Angrist, J. and Krueger, A. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, 106(4).
- Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review*, 107(4):1207–38.
- Breda, T. and Ly, S. T. (2015). Professors in core science fields are not always biased against women: Evidence from France. *American Economic Journal: Applied Economics*, 7(4):53–75.
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest*, 15(3):75–141.
- Ceci, S. J. and Williams, W. M. (2011). Understanding current causes of women’s underrepresentation in science. *Proceedings of the National Academy of Sciences of the United States of America*, 108(8):3157–62.
- Foschi, M., Lai, L., and Sigerson, K. (1994). Gender and Double Standards in the Assessment of Job Applicants. *Social Psychology Quarterly*, 57(4):326–339.
- Fryer, R. G. J. and Levitt, S. D. (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, 2(2):210–240.
- Glass, C. and Minnotte, K. L. (2010). Recruiting and hiring women in STEM fields. *Journal of Diversity in Higher Education*, 3(4):218–229.
- Heilman, M. E., Martell, R. F., and Simon, M. C. (1988). The vagaries of sex bias: Conditions regulating the undervaluation, evaluation, and overvaluation of female job applicants. *Organizational Behavior and Human Decision Processes*, 41(1):98–110.
- Hill, C., Corbett, C., and St. Rose, A. (2010). Why so Few? Women in Science, Technology, Engineering and Mathematics. (*American Association of University Women, Washington, DC*).
- Irvine, A. D. (1996). Jack and Jill and Employment Equity. *Dialogue*, 35(02):255.
- Koch, A. J., D’Mello, S. D., and Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, 100(1):128–161.
- Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Academic disciplines. *Science*, 347(6219):23–34.
- Moss-Racusin, C. a., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty’s subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41):16474–16479.
- National Research Council (2010). Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty. (*National Academies Press, Washington, DC*).
- Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women’s careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.
- Shalala, D. E., AGonino, A. M., Bailyn, L., Birgeneau, R. J., and Cauce, A. M. (2006). Beyond bias and barriers: fulfilling the potential of women in academic science and engineering. (Summary). *National Academies Press*, pages 1–12.
- Sheltzer, J. M. and Smith, J. C. (2014). Elite male faculty in the life sciences employ fewer women. *Proceedings of the National Academy of Sciences*, 111(28):10107–10112.
- Steinpreis, R. E., Anders, K. A., and Ritzke, D. (1999). The impact of gender on the re-

- view of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles*, 41(7-8):509–528.
- Swim, J., Borgida, E., Maruyama, G., and Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin*, 105(3):409–429.
- Terrier, C. (2016). Boys Lag Behind: How Teachers’ Gender Biases Affect Student Achievement.
- West, M. S. and Curtis, J. W. (2006). AAUP Faculty Gender Equity Indicators 2006. (*American Association of University Professors, Washington, DC*), page 86.
- Williams, W. and Ceci, S. (2015). Williams and Ceci 2015 Supplemental Information. *PNAS online*, pages 1–30.
- Wolfinger, N. H., Mason, M. A., and Goulden, M. (2008). Problems in the Pipeline: Gender, Marriage, and Fertility in the Ivory Tower. *The Journal of Higher Education*, 79(4):388–405.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Number 2. The MIT Press.

Tables

Table 1.1 – Description of teachers' recruiting exams

	Different exams in different subjects?	Teaching level	Admission rate 2006-2013	Date written tests	Date oral tests	Required diploma to apply	
						Period 2006-2010	Period 2011-2013
Higher-level: Agrégation	Yes	Mostly high-school and higher education	12,78%	April	June	College degree (4 years at university)	Master (5 years at university)
Medium-level: CAPES	Yes	Middle school and high-school	23,03%	April	June	College degree (3 years at university)	Master (5 years at university)
Lower-level: CRPE	No, but math and French oral and written tests for all candidates after 2011	Primary school	21,52%	April (September after 2011)	June	College degree (3 years at university)	Master (5 years at university)

Table 1.2 – General sample statistics for teaching exams 2006-2013

	Whole sample	Higher level: Agrégation (all fields*)	Medium level: Capes (all fields*)	Lower level: CRPE
Number of candidates	501,196	67,501	160,575	273,12
Number of candidates eligible for the oral tests	214,78	18,887	77,316	118,577
Number of admitted	104,365	8,629	36,974	58,762
Admission rate	20,82%	12,78%	23,03%	21,52%
Admission rate among those who take both the medium- and high-level exams the same year	-	4,40%	18,19%	-
Share of candidates who take the CAPES and the Agrégation exam the same year	-	66,57%	30,60%	-
Admission rate among candidates eligible for the oral tests	48,59%	45,69%	47,82%	49,56%
Mean age of candidates	27,57	28,57	27,43	27,4
Share of French citizens among all candidates	98,38%	95,24%	97,45%	99,70%
Share of retakers** among all candidates	24,75%	23,17%	25,24%	24,86%
Share of retakers** among candidates eligible for the oral tests	18,75%	17,29%	19,87%	18,26%
Share of women among all candidates	73,38%	56,08%	63,85%	83,26%
Share of women among eligible candidates	74,50%	54,12%	65,97%	83,31%
Share of women among admitted candidates	75,92%	53,26%	67,52%	84,54%

* The 11 fields (over 40 existing fields) considered in this research. ** Retakers are candidates who took but did not pass the exam the previous year.

Table 1.3 – Sample statistics for the high-level exam (Agrégation) and medium-level exam (CAPES) 2006-2013**A. High-level exam (Agrégation)**

	Mathematics	Physics	Philosophy	Chemistry	Economics	Geography	History	Biology	Classical Literature	Modern Literature	Languages	All
Number of candidates	12,634	5,573	4,862	2,302	1,33	1,413	9,326	8,863	1,843	6,218	13,137	67,501
Number of candidates eligible for the oral tests	4,782	1,821	843	679	417	428	1,424	1,589	852	1,812	4,24	18,887
Number of admitted	2,266	821	365	328	213	210	675	679	391	784	1,897	8,629
Admission rate	17,94%	14,73%	7,51%	14,25%	16,02%	14,86%	7,24%	7,66%	21,22%	12,61%	14,44%	12,78%
Share of admitted among eligible	47,39%	45,09%	43,30%	48,31%	51,08%	49,07%	47,40%	42,73%	45,89%	43,27%	44,74%	45,69%
Share of women among all candidates	33,42%	30,81%	40,23%	52,82%	48,50%	49,40%	48,93%	66,51%	75,53%	79,50%	80,73%	56,08%
Share of women among eligible candidates	27,14%	30,48%	32,50%	55,82%	57,79%	51,87%	43,68%	68,66%	74,06%	80,85%	81,23%	54,12%
Share of women among admitted candidates	27,89%	33,86%	35,62%	58,23%	57,75%	58,57%	42,37%	66,42%	69,31%	78,44%	78,86%	53,26%

B. Medium-level exam (CAPES)

	Mathematics	Physics-Chemistry	Philosophy	Economics	History - Geography	Biology	Classical Literature	Modern Literature	Languages	All
Number of candidates	22,031	14,401	5,932	4,921	28,823	16,233	2,423	20,111	45,7	160,575
Number of candidates eligible for the oral tests	13,226	7,547	684	1,206	11,039	5,671	1,92	12,313	23,71	77,316
Number of admitted	6,403	3,402	274	650	5,073	2,475	1,018	6,394	11,285	36,974
Admission rate	29,06%	23,62%	4,62%	13,21%	17,60%	15,25%	42,01%	31,79%	24,69%	23,03%
Admission rate among eligible candidates	48,41%	45,08%	40,06%	53,90%	45,96%	43,64%	53,02%	51,93%	47,60%	47,82%
Share of women among all candidates	45,71%	42,86%	42,30%	47,04%	50,09%	64,63%	81,30%	82,41%	83,13%	63,85%
Share of women among eligible candidates	43,91%	44,27%	32,89%	48,67%	52,02%	65,60%	81,09%	83,26%	83,42%	65,97%
Share of women among admitted candidates	49,45%	48,24%	33,21%	53,08%	51,59%	65,62%	80,75%	82,51%	83,14%	67,52%

* Retakers are candidates who took but did not pass the exam the previous year.

Table 1.4 – Estimates of the bonus for women on oral tests at the higher-level exam in each field. Linear regression models DD, S, and S+IV. 2006-2013

	Bonus for Women				Effect of Written rank		Observations	Weak identification F stat	Sargan Chi2 p-value	Student p-value : written rank = 1
	DD1	DD2	S	S + IV	S	S + IV				
Maths	0.115*** (0.00795)	0.0969*** (0.00817)	0.0377*** (0.00733)	0.136*** (0.0124)	0.541*** (0.0128)	1.298*** (0.0713)	4111	109.584	0.972	0.000
Physics	0.113*** (0.0138)	0.112*** (0.0149)	0.0565*** (0.0133)	0.116*** (0.0162)	0.481*** (0.0234)	1.041*** (0.107)	1708	45.914	0.386	0.702
Philosophy	0.0939*** (0.0246)	0.104*** (0.0269)	0.0646*** (0.0220)	0.160*** (0.0499)	0.256*** (0.0357)	1.973*** (0.580)	829	5.019	0.571	0.094
Chemistry	0.0581*** (0.0203)	0.0366 (0.0235)	0.00303 (0.0211)	0.0475* (0.0252)	0.537*** (0.0377)	1.060*** (0.241)	651	7.790	0.845	0.802
Economics	-0.00661 (0.0319)	0.0155 (0.0398)	0.000762 (0.0326)	0.00300 (0.0349)	0.334*** (0.0540)	1.053*** (0.348)	403	4.548	0.988	0.878
Geography	0.0314 (0.0289)	0.00706 (0.0340)	0.0445 (0.0293)	0.00968 (0.0388)	0.434*** (0.0506)	0.987** (0.437)	424	2.495	0.463	0.976
History	-0.000247 (0.0181)	-0.00717 (0.0190)	-0.00766 (0.0153)	-0.000217 (0.0319)	0.280*** (0.0264)	2.114** (1.074)	1410	1.713	0.831	0.299
Biology	-0.0350** (0.0170)	-0.0461** (0.0181)	-0.0584*** (0.0146)	-0.0451** (0.0196)	0.342*** (0.0237)	1.255*** (0.175)	1571	24.347	0.676	0.146
Classical literature	0.00311 (0.0209)	-0.0135 (0.0239)	-0.0406** (0.0206)	-0.00115 (0.0303)	0.475*** (0.0316)	1.267*** (0.297)	909	7.052	0.346	0.369
Modern literature	-0.0189 (0.0191)	-0.0195 (0.0205)	-0.0411** (0.0168)	-0.00749 (0.0246)	0.354*** (0.0227)	1.338*** (0.390)	1812	5.618	0.648	0.387
Languages	-0.0585*** (0.0120)	-0.0586*** (0.0124)	-0.0707*** (0.0103)	-0.0527*** (0.0133)	0.387*** (0.0144)	1.165*** (0.125)	4114	43.279	0.188	0.187
Controls:										
County	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Nationality	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Age	No	Yes	Yes	No	Yes	No		No	No	No
Month of birth	No	Yes	Yes	No	Yes	No		No	No	No
Diploma	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Year	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes

Note: * p<0.1, ** p<0.05, *** p< 0.01. Standard errors in parenthesis. The number of observations corresponds to the case without control variables. It only decreases marginally after adding controls. It also slightly differs from the number of candidates eligible to oral examination given in table S3b due to a sample restriction to candidates taking both the written and the oral tests. Instrumental variables (IV): Age and Month of birth. The Sargan statistic tests for the exclusion restriction condition. When the p-value of the Sargan test is above 0.05 the exogeneity of the instruments cannot be rejected with a 5% type 1 error. Fisher statistic tests the weakness of instruments. Instruments are typically weak when the Fisher statistic is below 15.

Table 1.5 – Estimates of the bonus for women on oral tests at the medium-level exam in each field. Linear regression models DD, S, and S+IV. 2006-2013

	Bonus for Women				Effect of Written rank		Observations	Weak identification F stat	Sargan Chi2 p-value	Student p-value : written rank = 1
	DD1	DD2	S	S + IV	S	S + IV				
Maths	0.130*** (0.00612)	0.127*** (0.00625)	0.0790*** (0.00518)	0.116*** (0.00640)	0.314*** (0.00928)	0.896*** (0.0603)	11462	172.240	0.474	0.085
Physics-Chemistry	0.0639*** (0.00760)	0.0641*** (0.00784)	0.0444*** (0.00664)	0.0601*** (0.00748)	0.383*** (0.0116)	0.941*** (0.0635)	6683	143.609	0.440	0.354
Philosophy	0.0901*** (0.0321)	0.0857** (0.0368)	0.0555* (0.0287)	0.124* (0.0701)	0.0980** (0.0467)	2.103* (1.255)	577	1.323	0.573	0.379
Economics	0.0631*** (0.0195)	0.0189 (0.0218)	0.0161 (0.0180)	0.0207 (0.0506)	0.320*** (0.0301)	2.955*** (0.883)	1072	4.420	0.840	0.027
History-Geography	0.00539 (0.00617)	0.00230 (0.00631)	-0.00982* (0.00525)	0.00983 (0.00746)	0.345*** (0.00909)	1.287*** (0.122)	10548	57.254	0.372	0.019
Biology	0.0146 (0.00960)	0.00323 (0.00991)	-0.0109 (0.00796)	0.0140 (0.0124)	0.309*** (0.0130)	1.475*** (0.167)	5263	38.607	0.729	0.004
Classical literature	-0.0245 (0.0174)	-0.0250 (0.0189)	-0.0455*** (0.0164)	-0.0221 (0.0185)	0.459*** (0.0227)	1.083*** (0.107)	1792	47.319	0.132	0.439
Modern literature	-0.0390*** (0.00710)	-0.0425*** (0.00726)	-0.0442*** (0.00625)	-0.0443*** (0.00766)	0.453*** (0.00835)	1.101*** (0.0510)	11679	226.815	0.575	0.048
Languages	-0.0145** (0.00566)	-0.0120** (0.00576)	-0.0167*** (0.00479)	-0.0130* (0.00779)	0.374*** (0.00620)	1.569*** (0.0911)	22385	134.891	0.474	0.000
Controls:										
County	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Nationality	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Age	No	Yes	Yes	No	Yes	No		No	No	No
Month of birth	No	Yes	Yes	No	Yes	No		No	No	No
Diploma	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Year	No	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes

Note: * p<0.1, ** p<0.05, *** p< 0.01. Standard errors in parenthesis. The number of observations corresponds to the case without control variables. It only decreases marginally after adding controls. It also slightly differs from the number of candidates eligible to oral examination given in table S3b due to a sample restriction to candidates taking both the written and the oral tests. Instrumental variables (IV): Age and Month of birth. The Sargan statistic tests for the exclusion restriction condition. When the p-value of the Sargan test is above 0.05 the exogeneity of both the instruments cannot be rejected with a 5% type 1 error. Fisher statistic tests the weakness of instruments. Instruments are typically weak when the Fisher statistic is below 15.

Table 1.6 – Values taken by Indexes of Feminization

	Index of Feminization	Alternative measure 1	Alternative measure 2	Alternative measure 1b	Alternative measure 2b
	Proportion of women among professors and assistant professors in the field	Proportion of women among Agrégation holders in the field	Proportion of women among Agrégation candidates in the field over the period 2006-2013	Proportion of women among CAPES holders in the field	Proportion of women among CAPES candidates in the field over the period 2006-2013
Mathematics	20,88%	36,83%	28,53%	51,56%	46,05%
Physics	16,78%	40,71%	31,73%	46,21%	45,25%
Chemistry	37,40%		57,30%		
Philosophy	27,14%	36,20%	32,69%	40,33%	31,89%
Economics	39,64%	45,13%	57,07%	50,98%	49,16%
Geography	36,52%	43,37%	43,83%	52,89%	52,18%
History	41,90%		52,12%		
Biology	45,94%	65,09%	68,75%	65,32%	65,84%
Classical Literature	55,75%	76,36%	74,70%	83,51%	82,59%
Modern Literature	55,50%	77,03%	80,85%	85,55%	83,55%
Languages	61,89%	78,90%	81,40%	84,67%	83,85%

Source: Statistics from the Ministry of higher education and research.

Table 1.7 – Estimates of the linear relationship $b = \beta + \gamma s$ between the bias towards females on oral tests b and 3 indexes of fields' extent of feminization (s). 2006-2013.

	Candidates taking both Capes and Agrégation			All candidates		
	Medium level (N=3488)	High level (N=3488)	Difference	Medium level (N=71460)	High level (N=17766)	Difference
First index of feminization: Proportion of female among assistant professors and professors in each field						
Slope (γ)	-0,28 (0.02)	-0,53 (0.00)	-0,23 (p=.08)	-0,33 (0.00)	-0,41 (0.00)	-0,08 (p=.11)
Intercept (β)	0,13 (0.01)	0,25 (0.00)	0,11 (p=.04)	0,17 (0.00)	0,19 (0.00)	0,02 (p=.28)
Second index of feminization: Proportion of female among the medium-level exam holders in each field						
Slope (γ)	-0,27 (0.03)	-0,42 (0.00)	-0,12 (p=0.37)	-0,27 (0.00)	-0,37 (0.00)	-0,09 (p=0.08)
Intercept (β)	0,2 (0.01)	0,3 (0.00)	0,1 (p=0.22)	0,2 (0.00)	0,25 (0.00)	0,05 (p=0.11)
Third index of feminization: Proportion of female among the high-level exam holders in each field						
Slope (γ)	-0,25 (0.02)	-0,4 (0.00)	-0,12 (p=0.30)	-0,26 (0.00)	-0,35 (0.00)	-0,09 (p=0.05)
Intercept (β)	0,16 (0.01)	0,26 (0.00)	0,09 (p=0.16)	0,17 (0.00)	0,21 (0.00)	0,04 (p=0.09)

Note: All estimated intercepts and slopes are significant at the 5% level. Standard errors clustered at the (subject*year) level are reported in parenthesis (except for the difference between the slopes or intercepts where the p-value of the test of the null hypothesis is reported). Each model includes controls for candidates' characteristics (age, month of birth, nationality, county of residence and education) as well as time and field fixed effects.

Table 1.8 – Probability of success by gender, assuming success is either based only on written tests or only on oral tests.

A. High-level exam (Agréation). 2006-2013.												
	Mathematics	Physics	Philosophy	Chemistry	Economics	Geography	History	Biology	Classical literature	Modern literature	Languages	All
Fictive success rate for women after written tests (a)	44,80%	44,60%	42,40%	49,60%	53,90%	55,70%	47,70%	43,30%	44,90%	42,90%	46,00%	45,50%
Fictive success rate for women after oral tests (b)	57,50%	54,10%	51,70%	52,00%	54,30%	53,40%	46,80%	41,30%	44,60%	42,00%	44,30%	46,80%
Relative risk for women (=b/a)	1,28	1,21	1,22	1,05	1,01	0,96	0,98	0,95	0,99	0,98	0,96	1,03
Odds ratio for women (b/(1-b))/(a/(1-a))	1,67	1,46	1,45	1,1	1,02	0,91	0,96	0,92	0,99	0,97	0,93	1,05
Share of women among hired	27,90%	33,90%	35,60%	58,20%	57,70%	58,60%	42,40%	66,40%	69,30%	78,40%	78,90%	53,30%
Share of women among fictively hired after written test	23,10%	29,40%	31,20%	56,70%	58,20%	58,60%	43,80%	68,90%	70,80%	80,50%	81,50%	52,50%
Share of women among fictively hired after oral test	29,80%	35,60%	38,10%	59,90%	58,70%	57,60%	42,90%	65,70%	70,50%	78,60%	78,50%	54,00%
B. Medium-level exam (CAPES). 2006-2013.												
	Mathematics	Physics-Chemistry	Philosophy	Economics	History-Geography	Biology	Classical literature	Modern literature	Languages	All		
Fictive success rate for women after written tests (a)	52,20%	48,10%	43,30%	61,80%	47,00%	47,00%	55,80%	53,10%	49,80%	50,40%		
Fictive success rate for women after oral tests (b)	62,40%	53,10%	51,30%	65,60%	46,90%	47,70%	55,90%	52,70%	49,40%	51,70%		
Relative risk for women (=b/a)	1,2	1,1	1,19	1,06	1	1,02	1	0,99	0,99	1,03		
Odds ratio for women (b/(1-b))/(a/(1-a))	1,52	1,22	1,38	1,18	1	1,03	1	0,98	0,98	1,05		
Share of women among hired	49,40%	48,20%	33,20%	53,10%	51,60%	65,60%	80,70%	82,50%	83,10%	67,50%		
Share of women among fictively hired after written test	43,20%	44,60%	30,10%	49,10%	51,30%	66,00%	81,20%	84,00%	83,70%	66,60%		
Share of women among fictively hired after oral test	51,60%	48,40%	35,30%	53,20%	51,30%	67,00%	81,10%	82,10%	83,00%	68,00%		
C. Lower-level exam (CRPE). 2011-2013.												
	Mathematics	Literature	All									
Fictive success rate for women after written tests (a)	60,60%	65,20%	62,90%									
Fictive success rate for women after oral tests (b)	64,10%	63,70%	63,90%									
Relative risk for women (=b/a)	1,06	0,98	1,02									
Odds ratio for women (b/(1-b))/(a/(1-a))	1,16	0,94	1,04									
Share of women among fictively hired after written test	81,80%	87,50%	84,60%									
Share of women among fictively hired after oral test	86,50%	87,00%	86,80%									

Notes: Each year in each exam, there is a predefined number of hires h . The success rate of a given population is the share of individuals in that population who rank within the h first ranks.

Table 1.9 – Heterogeneity of the bonus for female candidates at the high-level exams oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2006-2013

	Q1	Q2	Q3	Q4	Q5
Maths	0.0929*** (0.0153)	0.0918*** (0.0154)	0.0882*** (0.0157)	0.0988*** (0.0171)	0.0562*** (0.0210)
Physics	0.121*** (0.0269)	0.116*** (0.0269)	0.134*** (0.0270)	0.0803*** (0.0273)	0.0305 (0.0296)
Philosophy	0.0146 (0.0461)	0.0880* (0.0466)	0.166*** (0.0482)	0.0975** (0.0445)	0.0727 (0.0475)
Chemistry	0.0940** (0.0413)	0.0275 (0.0426)	0.0264 (0.0410)	0.0850* (0.0449)	0.0303 (0.0405)
Economics	0.0308 (0.0601)	0.00857 (0.0623)	0.117* (0.0660)	-0.0233 (0.0617)	-0.0856 (0.0614)
History	0.0876 (0.0554)	0.127** (0.0602)	0.0715 (0.0558)	0.0409 (0.0589)	0.0389 (0.0574)
Geography	-0.0724** (0.0329)	0.0519 (0.0344)	-0.0607* (0.0324)	-0.0216 (0.0337)	0.111*** (0.0338)
Biology	0.0308 (0.0601)	0.00857 (0.0623)	0.117* (0.0660)	-0.0233 (0.0617)	-0.0856 (0.0614)
Classical literature	0.0307 (0.0455)	-0.0809 (0.0496)	-0.114** (0.0457)	0.00493 (0.0417)	-0.0443 (0.0402)
Modern literature	0.0597 (0.0373)	-0.0250 (0.0350)	-0.0789** (0.0358)	-0.0307 (0.0361)	-0.0496 (0.0349)
Languages	-0.0764*** (0.0230)	-0.0385* (0.0231)	-0.0679*** (0.0230)	-0.0456* (0.0240)	-0.0663*** (0.0225)

Notes: Q1 to Q5 indicate subsamples of candidates based on their level on written tests (five quantiles with a fifth of the number of observations given in Table S4a). Standard errors in parenthesis. * p<0.1, ** p<0.05, *** p< 0.01.

Table 1.10 – Heterogeneity of the bonus for female candidates at the medium-level exams oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2006-2013

	Q1	Q2	Q3	Q4	Q5
Maths	0.0809*** (0.0111)	0.0915*** (0.0114)	0.105*** (0.0111)	0.119*** (0.0114)	0.107*** (0.0116)
Physics-Chemistry	0.0701*** (0.0154)	0.0654*** (0.0147)	0.0518*** (0.0151)	0.0571*** (0.0150)	0.0476*** (0.0151)
Philosophy	-0.0230 (0.0553)	0.0468 (0.0602)	0.0341 (0.0606)	0.134** (0.0562)	0.110* (0.0607)
Economics	0.0412 (0.0388)	0.0715* (0.0373)	0.0967*** (0.0366)	0.0861** (0.0374)	0.0599 (0.0381)
History-Geography	0.00471 (0.0118)	-0.00112 (0.0117)	0.000338 (0.0116)	0.00199 (0.0117)	-0.00454 (0.0116)
Biology	0.0412 (0.0388)	0.0715* (0.0373)	0.0967*** (0.0366)	0.0861** (0.0374)	0.0599 (0.0381)
Classical literature	-0.0206 (0.0346)	-0.0333 (0.0351)	-0.0390 (0.0346)	-0.0287 (0.0336)	-0.0271 (0.0337)
Modern literature	-0.0331** (0.0134)	-0.0282** (0.0133)	-0.0397*** (0.0140)	-0.0321** (0.0140)	-0.0248* (0.0140)
Languages	-0.00759 (0.0105)	-0.0146 (0.0108)	-0.0216** (0.0106)	-0.00586 (0.0109)	-0.00121 (0.0105)

Notes: Q1 to Q5 indicate subsamples of candidates based on their level on written tests (five quantiles with a fifth of the number of observations given in Table S4b). Standard errors in parenthesis. * p<0.1, ** p<0.05, *** p< 0.01.

Table 1.11 – Estimates of the bonus for women on oral tests for women at the math and literature tests in the lower-level exam. Linear regression models DD, S, and S+IV. 2006-2013

	DD1	DD2	S	S + IV
Maths				
Bonus for Women	0.185*** (0.00703)	0.169*** (0.00704)	0.0384*** (0.00526)	0.0191 (0.0143)
Rank at written test			0.0663*** (0.00654)	-0.0308 (0.0284)
Observations	24306	24254	24254	2861
Student p-value : written rank = 1			0.000	0.000
Weak identification F stat				2225.331
Literature				
Bonus for Women	-0.0180*** (0.00677)	-0.0359*** (0.00683)	0.0393*** (0.00515)	0.0322** (0.0144)
Rank at written test			0.138*** (0.00630)	0.258*** (0.0503)
Observations	24306	24254	24254	2861
Student p-value : written rank = 1			0.000	0.000
Weak identification F stat				399.954
Controls:				
County	No	Yes	Yes	Yes
Nationality	No	Yes	Yes	Yes
Age	No	Yes	Yes	Yes
Month of birth	No	Yes	Yes	Yes
Diploma	No	Yes	Yes	Yes
Year	No	Yes	Yes	Yes
Region	No	Yes	Yes	Yes
Region X Year	No	Yes	Yes	Yes

Notes: Instrument in model S + IV is the candidates' rank at the test the previous year. Standard errors in parenthesis. * p<0.1, ** p<0.05, *** p< 0.01.

Table 1.12 – Heterogeneity of the bonus for female candidates at the lower-level exams math and literature oral tests. Estimates of the DD model on 5 subsamples based on quantiles of the written test scores. 2011-2013

	Q1	Q2	Q3	Q4	Q5
Maths	0.0357** (0.0155)	0.0568*** (0.0135)	0.0868*** (0.0122)	0.0715*** (0.0111)	0.0628*** (0.00998)
Literature	0.0463*** (0.0101)	0.0697*** (0.0117)	0.0395*** (0.0121)	0.0435*** (0.0126)	0.0599*** (0.0129)

Notes: Q1 to Q5 indicate subsamples of candidates based on their level on written tests (five quantiles). Standard errors in parenthesis. * p<0.1, ** p<0.05, *** p< 0.01.

Table 1.13 – Effect of the gender composition of the examiners' panels on oral test scores at the math medium-level exam

	(1)	(2)	(3)
Number of women among examiners			
0	ref	ref	-
1	-0.0281 (0.0495)	-0.0144 (0.0573)	-
2	-0.112** (0.0564)	-0.101 (0.0639)	-
Number of women among examiners X female candidate			
0	ref	ref	ref
1	0.0766 (0.0853)	0.0803 (0.0858)	0,07917 (.05869)
2	0.0918 (0.0969)	0.0957 (0.0973)	0,0999 (.06723)
Controls:			
Oral test	Yes	Yes	Yes
Examiners	No	Yes	-
Candidates fixed effects	Yes	Yes	Yes
Groups of examiners fixed effects	No	No	Yes
Observations	2276	2276	2276

Note : Oral test scores only. * p<0.1, ** p<0.05, *** p<0.01. Examiner panels controls are county of residence and main employment status.

Table 1.14 – Composition of the jury at the Maths medium-level exam in 2014

Agregated jury	
Number of examiners*	72
Share of women among examiners	41,67%
Number of groups of examiners	48
Number of examiners per group	3
Groups of examiners	
Groups with no woman	2
Groups with one woman	32
Groups with two women	14
Groups with three women	0
Number of candidates evaluated by a group with no woman**	105
Number of candidates evaluated by a group with one woman	1516
Number of candidates evaluated by a group with two women	689
Number of candidates evaluated by a group with three women	0

* Each examiner is member of two examination panels. ** Each candidate is evaluated twice, by two different examination panels.

Table 1.15 – Bonus for women at one oral versus one written test in each field. Linear regression models DD. 2006-2013.**A. Higher-level exam**

	Maths	Physics	Philosophy	Chemistry	Economics	Geography	History	Biology	Classical literature	Languages
Bonus for women	0.119*** (0.00986)	0.128*** (0.0155)	0.0853* (0.0447)	0.0588** (0.0252)	0.0302 (0.0360)	-0.00536 (0.0325)	-0.00183 (0.0202)	-0.0245 (0.0191)	-0.0287 (0.0374)	0.0601** (0.0241)
Observations	4110	1708	320	651	403	424	1410	1571	490	1836

B. Medium-level exam

	Maths	Physics- Chemistry	Philosophy	Economics	History- Geography	Biology	Classical literature	Modern literature	Languages
Bonus for women	0.132*** (0.00648)	0.0586*** (0.00827)	0.101*** (0.0346)	0.0353 (0.0220)	0.0213*** (0.00658)	0.00269 (0.00970)	-0.0663*** (0.0207)	0.00878 (0.00838)	-0.0365*** (0.00612)
Observations	11462	6683	577	1072	10548	5263	1792	11679	22385

Note: Results based on candidates' rank difference between one oral test and one written test in each exam. The selected oral and written tests have been chosen to match as closely as possible in terms of their framing and the subtopic they cover (see Tables S14a and S14b). Standard errors in parenthesis. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$. The number of observation slightly differ from the number of candidates eligible to oral examination given in table S3b due to a sample restriction to candidates taking the oral tests. The number of observations in Table S13a also slightly differs from Table S4a due to missing detailed tests for some years in a few disciplines (see explanations in section 2).

Table 1.16 – Description of all tests at the medium-level examination

	Mathematics	Physics- Chemistry	Philosophy	Economic and social sciences	History- Geography	Biology	Classical Literature	Modern Literature	Languages
	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013
Written tests									
Test 1	One or several problems	Physics : one or several problems + questions and exercices	Essay	Economics : one essay + one question on history or epistemology	History : essay	Essay	Essay in French : literature and art culture	Essay in French : literature and art culture	Text commentary in foreign language
Test 2	One or several problems	Chemistry : one or several problem + questions and exercices	Study of a philosophical text	Sociology : one essay + one question on history or epistemology	Geography : essay	Essay	Translation in ancient language	Grammatical study of texts in French	Translation of a text in foreign language
Oral tests									
Test 1	Lecture : detailed outline on a random subject + questions	Lecture in physics or chemistry: presentation of one or several experiments + questions	Lecture : teaching sequence on a random subject + questions	Lecture : presentation on a random subject + questions	Lecture in history or geography : exposition on a random subject + questions	Lecture : exposition on a random subject + questions	Lecture : analysis of a random text in French or ancient language + questions	Lecture : analysis of a random text in French + questions	Lecture : presentation of documents + questions in foreign language
Test 2a	Questions with documents : knowledge of discipline, teaching programs and pedagogical reflexion	Questions with documents : knowledge of discipline, teaching programs and pedagogical reflexion	Text analysis : knowledge of discipline, teaching programs and pedagogical reflexion	Analysis of documents, questions and exercices : knowledge of discipline, teaching programs and pedagogical reflexion	Analysis of documents : knowledge of discipline, teaching programs and pedagogical reflexion	Analysis of documents : knowledge of discipline, teaching programs and pedagogical reflexion	Analysis of documents : knowledge of discipline, teaching programs and pedagogical reflexion	Analysis of documents : knowledge of discipline, teaching programs and pedagogical reflexion	Presentation of documents in foreign languages and questions
Test 2b	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document	BERCS : question with a document

Note: Official Journal of the Ministry of Education. Tests in red are used for the robustness check provided in Table 13a. A few tests changed slightly over the period 2006-2013. * The discipline (physics or chemistry) is randomly assigned to the candidate. ** In each field, this test aims at evaluating the candidate's knowledge of the discipline, of the teaching programs and her pedagogical skills.

Table 1.17 – Description of all tests at the high-level examination

	Mathematics	Physics	Chemistry	Philosophy	Economic and social sciences	Geography	History	Biology	Classical Literature	Modern Literature	Languages
	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013	2011-2013
Written exams											
Test 1	General math : one or several problems	Physics : one or several problems	Chemistry : one or several problems	Essay in philosophy without program	Essay in economics	Topic in geography	Essay in history	Essay on topic A***	Translation from latin	Essay in french	Essay in foreign language
Test 2	Mathematical analysis and probability : one or several problems	Chemistry : one or several problems	Physics : one or several problems	Essay in philosophy with program	Essay in sociology	Geography of territories	Essay in history	Essay on topic B***	Translation from ancient greek	Grammatical study of a french text before 1500	Translation
Test 3	-	Problème de physique	Problème de chimie	Epreuve d'histoire de la philosophie : commentaire de texte	Composition au choix sur l'histoire et la géographie ou le droit public et la science politique	Epreuve d'option** sur dossier : exercices, analyse de documents, synthèse, carte	Explication de texte en histoire	Essay on topic C***	Translation to latin	Grammatical study of a french text after 1500	Essay in French on foreign literature or civilisation
Test 4	-	-	-	-	-	Dissertation en histoire	Dissertation en géographie	-	Translation to ancient greek	Essay in french	-
Test 5	-	-	-	-	-	-	-	-	Dissertation française	Translation to latin	-
Test 6	-	-	-	-	-	-	-	-	-	Translation to foreign language	-
Oral exams											
Test 1	1) Lecture in algebra and geometry + questions 2) BERCS	1) Lecture in physics + questions 2) BERCS	Lecture in chemistry + questions	Lecture in philosophy	Lecture in economics and social sciences + questions	Analysis of documents + questions 2) BERCS	Lecture in history + questions	Experiment	Lecture + questions	Lecture + questions	Analysis of a text in foreign language + question in foreign language
Test 2	Lecture in mathematical analysis and probability + questions	1) Lecture in chemistry + questions 2) BERCS	1) Lecture in physics + questions 2) BERCS	1) Lecture + questions 2) BERCS	1) Analysis of documents + questions 2) BERCS	Lecture in geography + questions	1) Analysis of documents + questions 2) BERCS	Experiment	1) Analysis of a text in french + questions 2) BERCS	Analysis of a text in french	Translation and grammatical analysis + questions
Test 3	Modeling : presentation with documents	Experiment in physics + questions	Experiment in chemistry + questions	Analysis of a text in french	Math and statistics	History : analysis of documents + questions	Geography : analysis of documents + questions	Presentation in a chosen topic	Analysis of an ancient text + questions	1) Analysis of a text in french + questions 2) BERCS	Exposé en français sur un sujet de littérature étrangère puis entretien
Test 4	-	-	-	Translation and analysis of a text in foreign language	-	-	-	1) Presentation and experiment 2) BERCS	Analysis of a latin text + questions	Commentaire d'un texte de littérature ancienne ou moderne. Entretien sur le contenu présenté.	1) Translation + questions 2) BERCS
Test 5	-	-	-	-	-	-	-	-	Analysis of a greek text + questions	-	-

Note: Official Journal of the Ministry of Education. Tests in red are used for the robustness check provided in Table 13b. Tests in grey are missing data. A few tests changed slightly over the period 2006-2013.
* Candidates choose one between the two possible subjects. Topic A : biology et cell physiology, molecular biology ; Topic B : biology et physiology of organisms et biology of populations ; Topic C : Earth sciences, universe sciences and Earth's biosphere ** Those tests contain two subparts noted 1) and 2) and evaluated by the same group of examiners.

Table 1.18 – Female mean rank at all tests at the medium-level examination

	Mathematics		Physics-Chemistry		Philosophy		Social sciences		History-Geography	
	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013
Written exams										
Test 1	0.479	0.457	0.451	0.431	0.507	0.487	0.512	0.49	0.489	0.496
Test 2	0.479	0.478	0.547	0.541	0.488	0.498	0.5	0.522	0.502	0.5
Oral exams										
Test 1	0.542	0.547	0.532	0.522	0.488	0.528	0.524	0.519	0.506	0.502
Test 2	0.52	0.547	0.522	0.535	0.581	0.51	0.532	0.54	0.494	0.495
Test 3	-	-	-	-	0.566	-	0.546	-	0.496	-
	Biology		Classical Literature		Modern Literature		Languages			
	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013		
Written exams										
Test 1	0.509	0.504	0.504	0.503	0.491	0.495	0.504	0.495		
Test 2	0.493	0.49	0.493	0.491	0.51	0.503	0.5	0.503		
Oral exams										
Test 1	0.5	0.494	0.489	0.51	0.493	0.5	0.498	0.498		
Test 2	0.522	0.488	0.494	0.499	0.494	0.504	0.51	0.495		
Test 3	-	-	0.498	-	-	-	-	-		

Note: Test ranks are standardized between 0 and 1, with mean 0.5. A female mean rank < 0.5 (resp. > 0.5) means that female do worse (resp. better) than male in average.

Table 1.19 – Female mean rank at all tests at the high-level examination

	Mathematics		Physics		Chemistry		Philosophy		Social sciences		Geography	
	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013
Written tests												
Test 1	0.43	0.436	0.457	0.437	0.489	0.5	0.504	0.482	0.523	0.497	0.557	0.538
Test 2	0.405	0.44	0.551	0.527	0.467	0.508	0.478	0.461	0.52	0.498	0.532	0.513
Test 3	-	-	0.431	0.472	0.497	0.514	0.514	0.518	0.522	0.533	0.502	0.53
Test 4	-	-	-	-	-	-	-	-	-	-	0.528	0.553
Test 5	-	-	-	-	-	-	-	-	-	-	-	-
Test 6	-	-	-	-	-	-	-	-	-	-	-	-
Oral tests												
Test 1	0.505	0.512	0.543	0.527	0.495	0.59	0.563	0.519	0.516	0.524	0.53	0.579
Test 2	0.504	0.504	0.553	0.564	0.51	0.576	0.514	0.53	0.492	0.482	0.536	0.583
Test 3	0.496	-	-	-	-	-	-	-	0.498	0.475	-	-
Test 4	-	-	-	-	-	-	-	-	-	-	-	-
Test 5	-	-	-	-	-	-	-	-	-	-	-	-
	History		Biology		Classical Literature		Modern Literature		Languages			
	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013	2006-2010	2011-2013
Written tests												
Test 1	0.518	0.473	0.514	0.514	0.508	0.49	0.489	0.503	0.505	0.505	-	-
Test 2	0.492	0.511	0.508	0.485	0.51	0.489	0.523	0.524	0.504	0.496	-	-
Test 3	0.495	0.491	0.496	0.5	0.503	0.492	0.509	0.519	0.499	0.499	-	-
Test 4	0.502	0.507	-	-	0.476	0.471	0.494	0.49	-	0.494	-	-
Test 5	-	-	-	-	0.49	0.5	0.504	0.489	-	0.504	-	-
Test 6	-	-	-	-	-	-	0.504	0.493	-	-	-	-
Oral tests												
Test 1	0.52	0.409	0.495	0.481	0.482	0.5	-	-	0.498	-	-	-
Test 2	0.516	-	0.5	0.506	-	-	-	-	0.493	-	-	-
Test 3	-	-	0.502	0.482	-	-	-	-	-	-	-	-
Test 4	-	-	-	-	-	-	-	-	-	-	-	-
Test 5	-	-	-	-	-	-	-	-	-	-	-	-

Note: Test ranks are standardized between 0 and 1, with mean 0.5. A female mean rank < 0.5 (resp. > 0.5) means that female do worse (resp. better) than male in average.

Figures

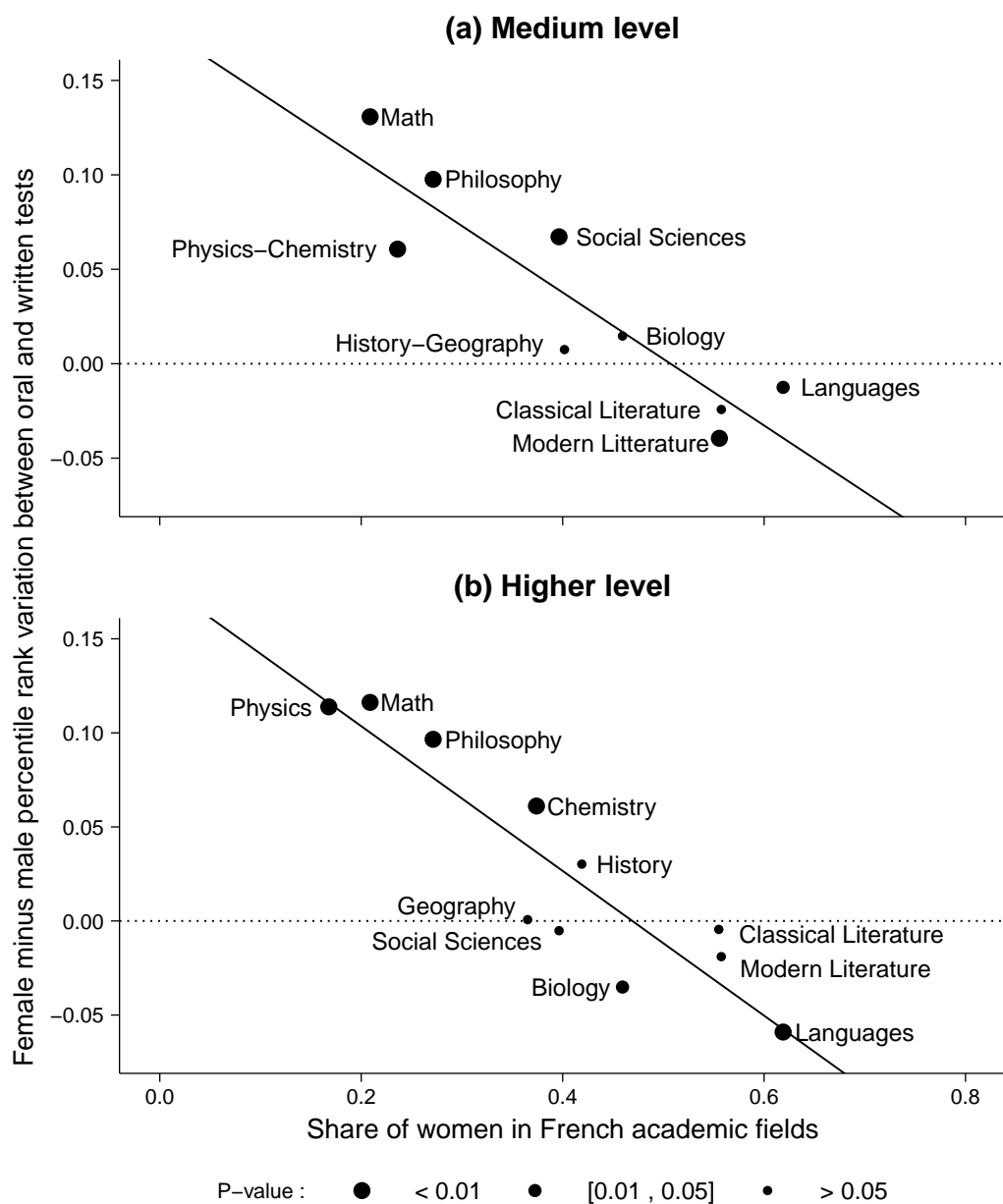


Figure 1.1 – Average gender gap between oral and written examinations, by subject, in the recruitment of teachers at the high- and medium-level.

Note: Differential variation in average percentile ranks of female and male candidates between anonymous written and non-anonymous oral tests. Computed for each subject-specific exam at the high- and medium-level as the gap between females' average percentile rank on oral and written tests, minus the same gap for men. Feminization index is the share of females among professors and assistant professors in each field (see SM for alternative measures).

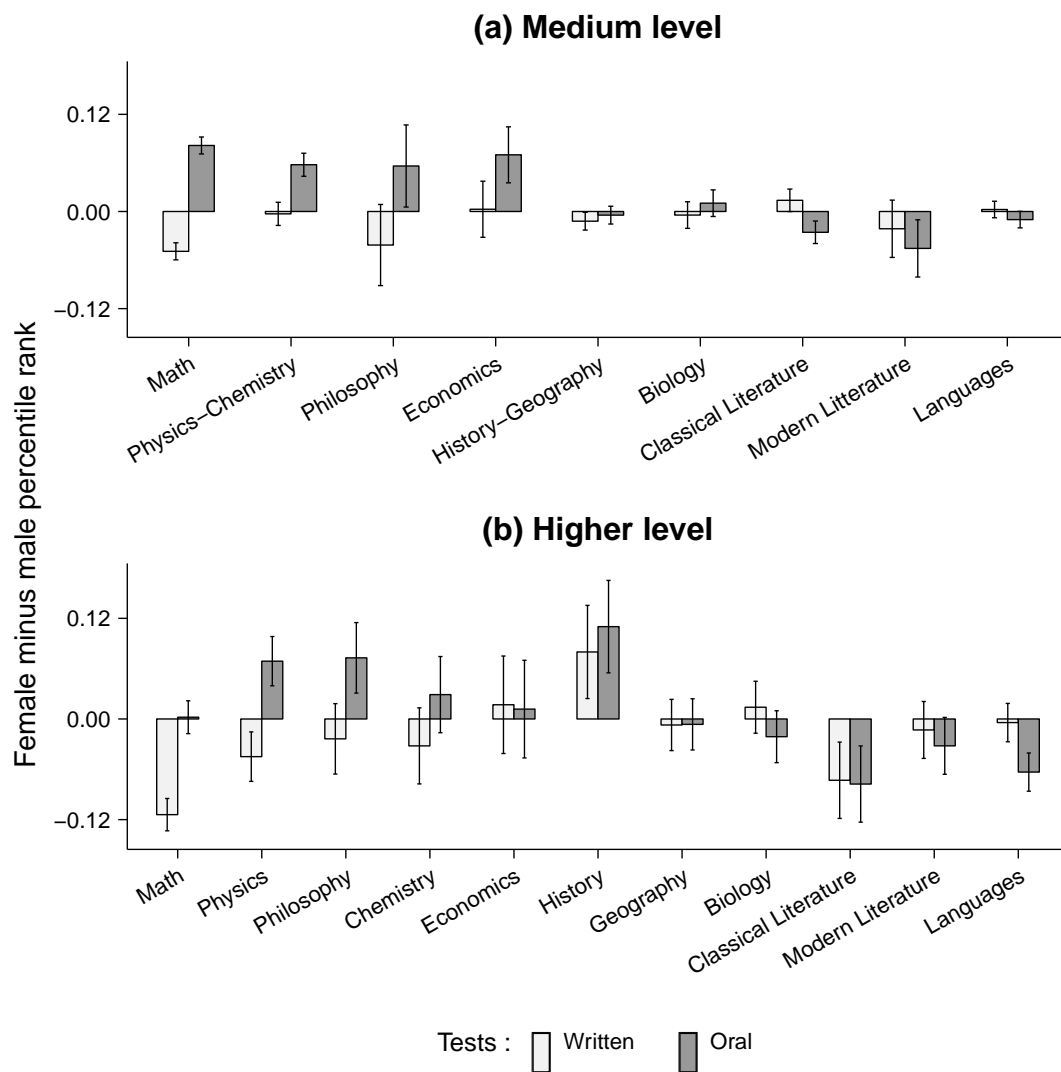


Figure 1.2 – Average rank difference between women and men on oral and written tests in each subject-specific exam at the high- and medium-level.

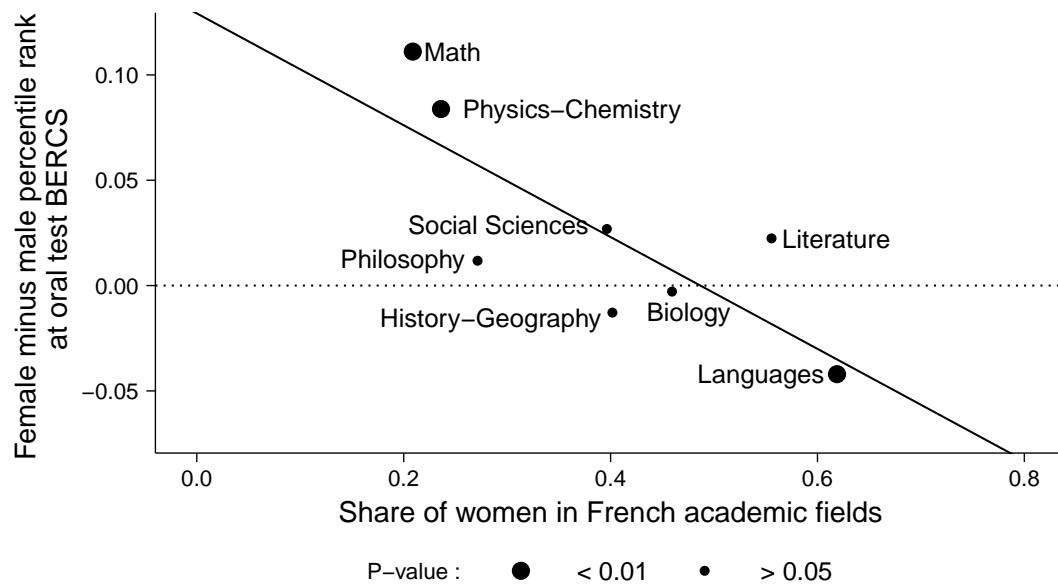


Figure 1.3 – Difference between women and men average rank on oral test in the subject "Behave as an ethical and responsible civil servant" in different fields of specialization.

Note: Computed for each subject-specific exam at the medium level. Feminization index is the share of females among professors and assistant professors in each field (see SM for alternative measures).

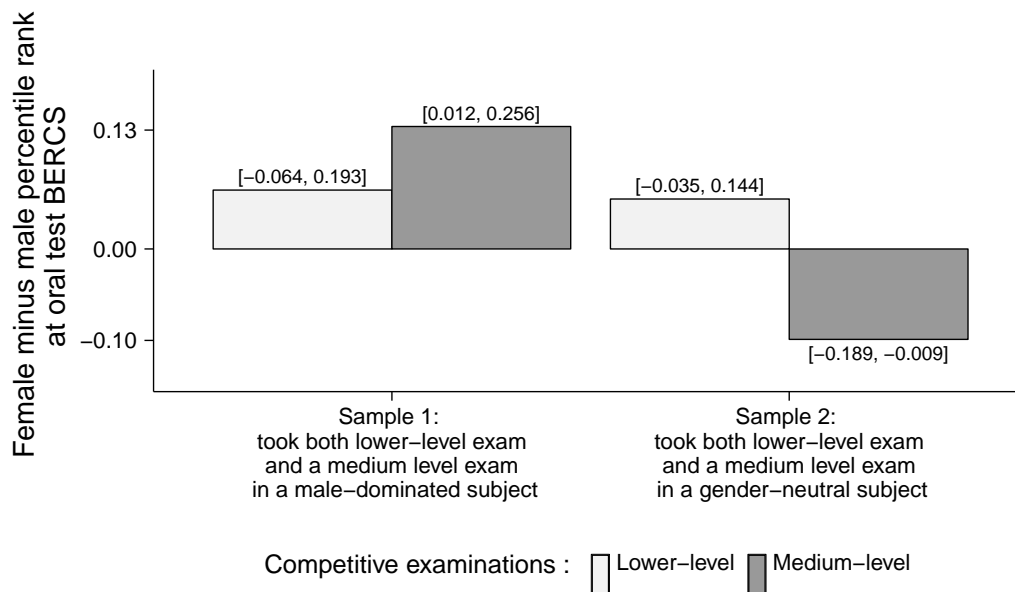


Figure 1.4 – Average rank difference between women and men on the oral test "Behave as an Ethical and Responsible Civil Servant" according to subject feminization.

Note: Rank difference between women and men at the oral test "Behave as an Ethical and Responsible Civil Servant" at the lower-level exam and at the medium-level exam among two samples of candidates: those who took both the lower-level and a medium-level exams in a strongly male-dominated subject (left side, $N=45$), and those who took both the lower-level and a medium-level exams in a more gender neutral subject (right side, $N=73$). Ranks at the tests have been computed within each sample, ignoring other candidates that are not in the sample. Confidence intervals at the 90% level are given in square brackets.

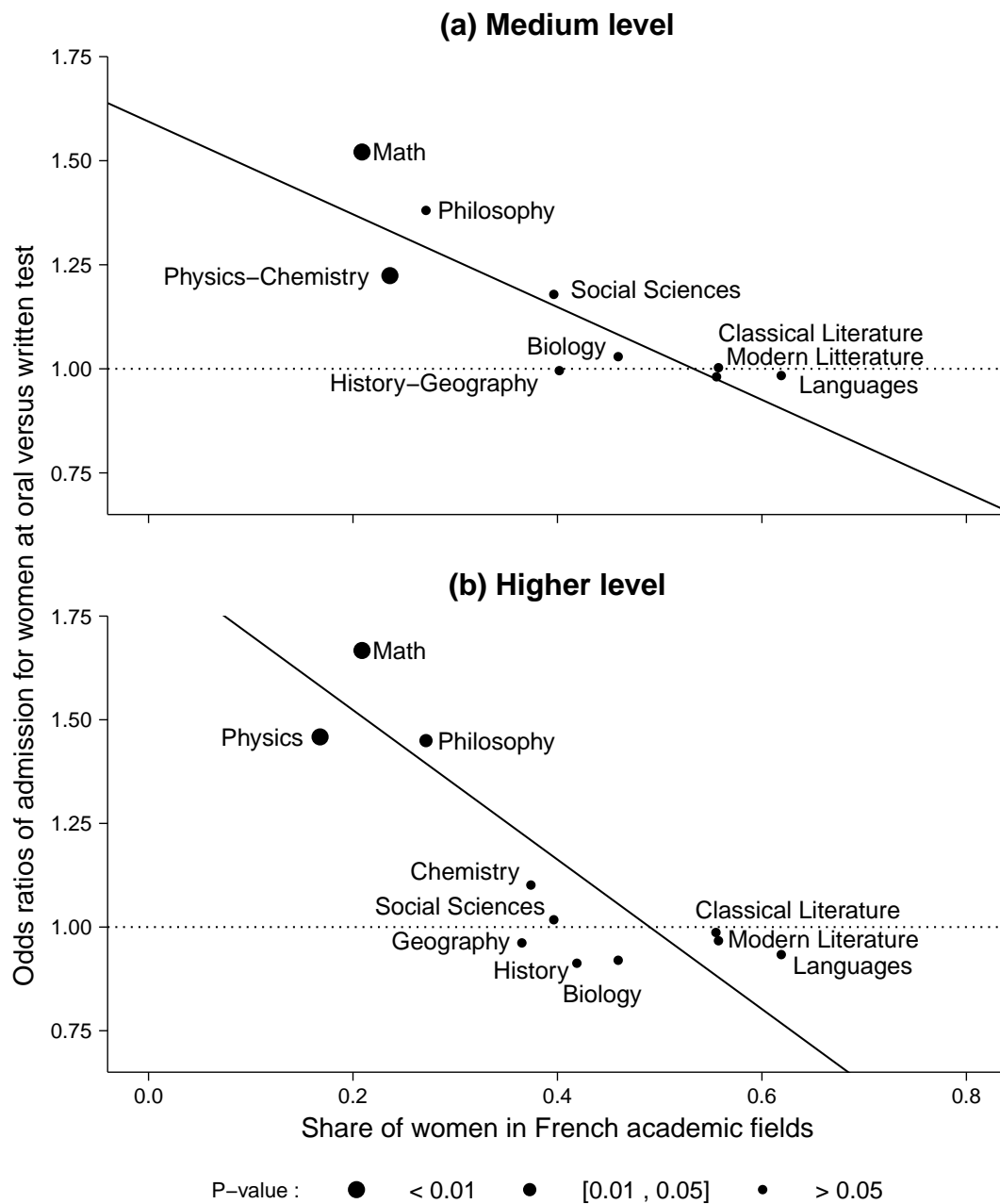


Figure 1.5 – Odds ratios of admission for women candidates at non-anonymous oral tests versus anonymous written tests.

Note: Computed for each subject-specific exam at the high- and medium-level. Feminization index is the share of females among professors and assistant professors in each field (see SM for alternative measures).

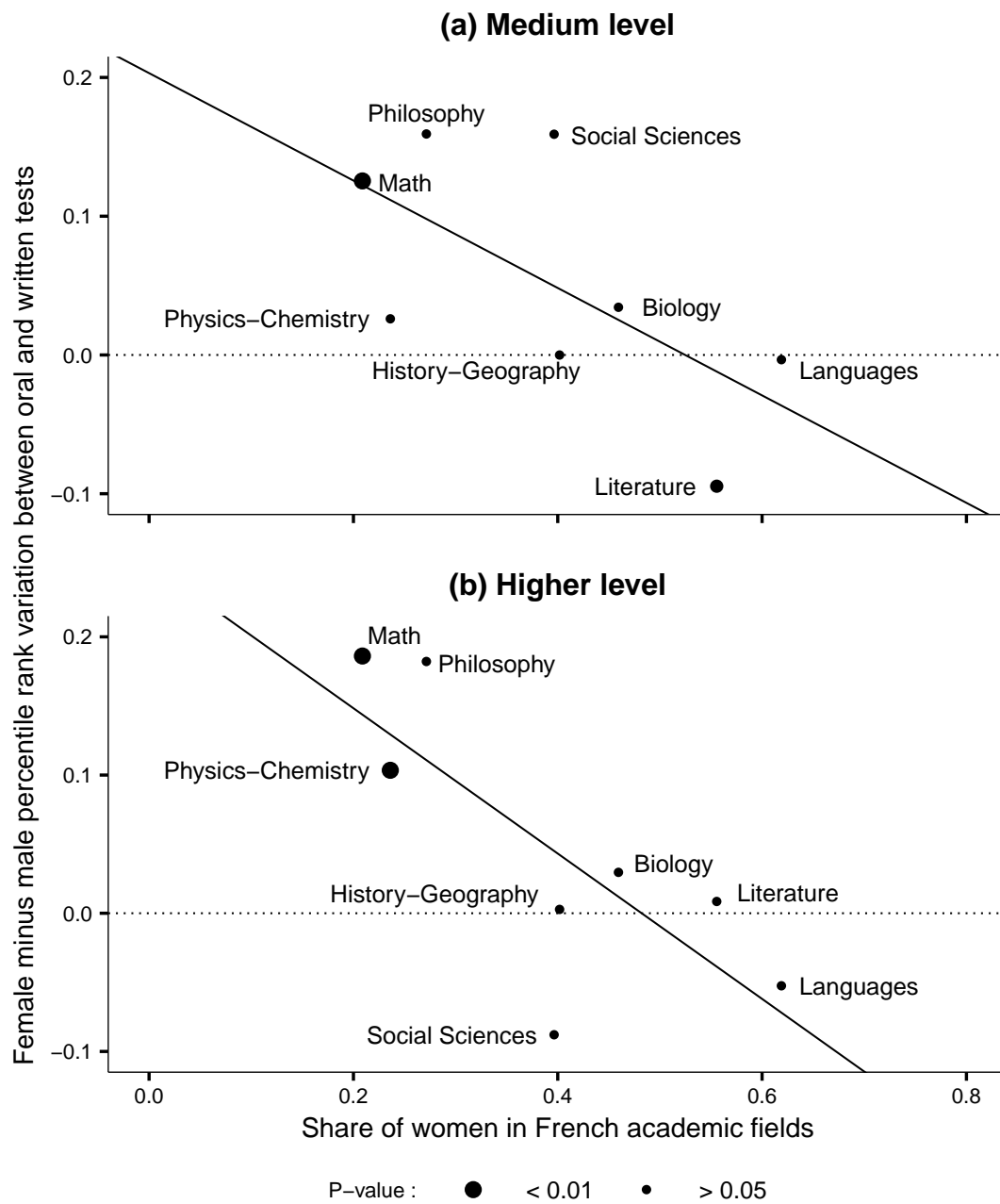


Figure 1.6 – Average gender gap between oral and written examinations, by subject, for candidates taking both medium- and higher-level exams the same year.

Note: Based only on candidates taking both the medium- and higher-level exams the same year. The figure gives the differential variation in average percentile ranks of female and male candidates between anonymous written and non-anonymous oral tests. Computed for each subject-specific exam at the high- and medium-level. Feminization index is the share of females among professors and assistant professors in each field.

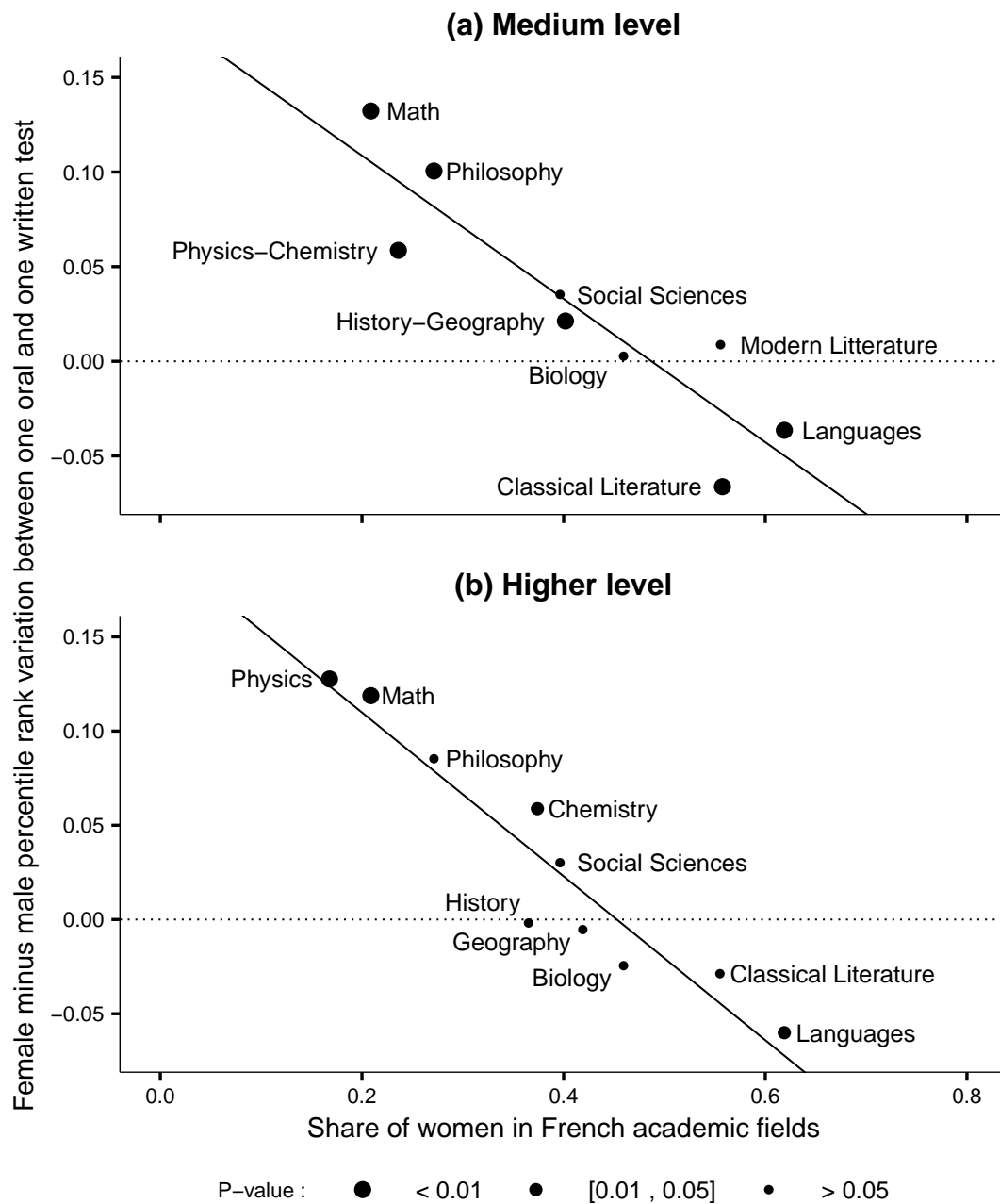


Figure 1.7 – Average gender gap in the recruitment of teachers at the high- and medium-level, by subject, considering only one oral and one written test.

Note: Based only on one written test and one oral test in each subject (instead of total scores as in Figure 1). The figure gives the differential variation in average percentile ranks of female and male candidates between anonymous written and non-anonymous oral tests. Computed for each subject-specific exam at the high- and medium-level. Feminization index is the share of females among professors and assistant professors in each field.

Chapter 2

Understanding the effect of salary, degree requirements and demand on teacher supply and quality: a theoretical approach

2.1 Introduction

Improving teacher quality is a major international concern, which explains why teacher selection methods have long been at the heart of educational research. In recent years, the lack of attractiveness of the teaching profession has also raised many questions. In France, as in most developed countries, teacher shortages are intensifying in primary and secondary education. Since 2013, the two largest French academic regions (Créteil and Versailles) are no longer able to fill all primary school teacher positions through regular recruitment processes. This is also the case for several subjects (notably mathematics, literature and foreign languages) in secondary education.

The objective of this chapter is to provide a theoretical framework that clarifies the conditions for attracting more and potentially better candidates to teaching positions. In economics, considerable attention has been paid to financial and non-financial incentives, sometimes to the detriment of other potential levers. The model proposed in this chapter examines the effect of compensation but also the effect of public management tools that are less often studied, such as qualification requirements and labor demand.

The theoretical and empirical works in economics have shown that monetary incentives are generally able to attract more candidates (Falch, 2010, 2011; Marinescu and Wolthoff, 2012; Dal Bó et al., 2013). In particular, Rumberger (1987) shows that com-

compensating for the wage gap between engineers and science teachers can eliminate the shortage of teachers in these disciplines. Falch et al. (2009) find a relationship between higher earning opportunities in alternative occupations and teacher shortages. However, empirical studies are less conclusive regarding the effect of financial incentives on teacher quality (Hanushek et al., 1999; Greaves and Sibieta, 2013; Rothstein, 2015). Most studies suggest that strong financial incentives are needed to observe a positive effect on student achievement (Prost, 2013; Greaves and Sibieta, 2013; Rothstein, 2015). I propose a simple model that highlights the mechanisms through which financial incentives operate. The model predicts a positive elasticity of teacher supply with wage under realistic assumptions but the effect on teachers' skills is much more ambiguous and will depend on the context.

The effect of demand on labour supply has received little attention in the economic literature. In France, primary and secondary school teachers account for 1.5% and 1.9% of the working population respectively. Approximately 27,000 tenured primary and secondary public school teachers are recruited each year¹, representing approximately 9% of the general baccalaureate graduates, 16% of the general bachelor's graduates and 22.4% of the general master's graduates (excluding engineers and business school graduates) of a generation². The recruitment of teachers therefore represents a massive selection process. Daussin-Bénichou et al. (2015) show that, on average, each teaching position attracts an additional 1.5 candidates. However, recruitment difficulties in some regions suggest a much lower elasticity when the demand for teachers increases. The model that I propose predicts a positive elasticity of the number of candidates with the number of posts. However, the elasticity decreases with teacher demand, which may contribute to explain why some academic regions are facing teacher shortage in France. The model also shows that teacher skills, measured during the recruitment process, tend to decrease with teacher demand.

A significant proportion of recruiters pre-select candidates based on qualifications. As skills are not fully revealed in job interviews, qualification is an attractive criteria for reducing uncertainty about the competencies of candidates. However, increasing the

¹Less than 2% of tenured teachers resign or are dismissed during their career.

²Source: SIES, DEPP, INSEE. 16% of primary and secondary school teachers work in private education. Public and private primary and secondary school teachers recruited each year represent about 10.7% of the general baccalaureate graduates, 19% of the bachelor's graduates and 26.7% of the master's graduates (excluding engineering and business school graduates) of a generation.

degree level of teachers can be counter-productive. In particular, the most competent individuals may renounce to apply as higher qualifications are associated with better outside opportunities (Berger and Toma, 1994; Angrist and Guryan, 2008).

Although it is difficult to predict the quality of a future teacher on the basis of his or her observable characteristics, many countries have recently decided to increase the minimum qualification required to become a teacher. These initiatives are supported by a recent report commissioned by the European Commission, which calls for a higher level of recruitment of teachers in order to ensure the "attractiveness and excellence" of the profession (European Commission, 2014). These provisions have been extended to a large number of European countries over the last decade, but their effects on teacher supply and quality have never been evaluated. In 2013, sixteen European countries require a master's degree to teach in secondary education and eleven countries require a similar diploma level to teach in primary education³. The model presented in this chapter predicts that an increase in the level of qualification required will tend to decrease the number of candidates to teaching positions. However, the effect on teacher quality is ambiguous and will namely depends on the correlation between academic skills and on-the-job teacher performance.

The rest of this chapter proceeds as follows. Section 4.2 review the literature and Section 3.4.5 introduces a simple model to examine the effect of wage on teacher supply and skills. Section 2.4 examines the effect of teacher demand and Section 2.5 studies the effect of diploma requirement on teacher supply and skills. Section 4.8 concludes.

2.2 Related literature

Several economic studies have examined the possibility of increasing the number and quality of applicants through financial incentives to compensate for the monetary and non-monetary benefits of outside opportunities. The empirical literature generally confirms that higher salaries help attract more candidates (Falch, 2010, 2011; Marinescu and Wolthoff, 2012; Dal Bó et al., 2013). Based on a randomized experiment in Mexico⁴, Dal Bó et al. (2013) estimate that the elasticity of labour supply with wage is around

³France, Portugal, Slovenia, Slovakia, Finland, Island, Croatia, Czech Republic, Germany, Estonia

⁴The salary offered to the treatment group was 33% higher than the salary offered to the control group.

2.15 in the public sector, which means that a 10% wage increase leads to an increase in the number of candidates of about 22%. Using two different methods, [Falch \(2010\)](#) and [Falch \(2011\)](#) estimate that the elasticity of teacher supply with salary is about 1.25-1.4 in Norway. Several articles also suggest that a salary differential between teachers and other professions ("alternative earning opportunities") reduces labor supply and contributes to teacher shortages. In the US, [Rumberger \(1987\)](#) shows that compensating for the wage gap between engineers and science teachers would help eliminate the shortage of science teachers. [Murnane and Olsen \(1989\)](#) shows that secondary school teachers are more likely to drop out of teaching when salaries associated with their subject specialty increase in the private sector (business and industry). Similarly, [Dolton and van der Klaauw \(1999\)](#) find that higher opportunity wages and lower teacher salaries contribute to teacher turnover in the UK. In Norway, [Falch et al. \(2009\)](#) finds that teacher shortages increase sharply as the regional unemployment rate decreases. These results suggest that teacher supply is affected by the relative pay of teachers compared to other occupations.

The economic literature is much less conclusive regarding the effect of financial incentives on the quality of recruitment ([Hanushek, 1997](#); [Hanushek et al., 1999](#); [Bénabou et al., 2009](#); [Marinescu and Wolthoff, 2012](#); [Dal Bó et al., 2013](#); [Greaves and Sibieta, 2013](#); [Rothstein, 2015](#)). [Marinescu and Wolthoff \(2012\)](#) show that higher wages allow to attract more educated and experienced applicants in the private sector, while [Dal Bó et al. \(2013\)](#) find positive effect of higher wages on personality traits (motivation, integrity, prosocial inclinations) and skills (IQ, past earning and past occupation) in the public sector. The effectiveness of financial incentives in education depends on the correlation between the determinants of teacher quality and the skills valued in non-teaching labour markets ([Hanushek and Rivkin, 2004](#)). [Card and Krueger \(1992\)](#) use the variation in teachers' salaries across states in the US and find that a 10% increase in teachers' salaries is associated with a 0.1 percentage point increase in student achievement. Similarly, [Loeb and Page \(2000\)](#) use the variation between states in teachers' salaries relative to alternative occupational opportunities and observe that a 10% increase in teachers' salaries reduces high school dropout rate from 3% to 4% ten years later. However, studies based on individual longitudinal data usually find that teachers' salaries have little or negligible impact on students' academic performance and future outcomes. Using the National Longitudinal Survey of Youth, [Betts \(1995\)](#) shows that teachers' salaries are

not correlated with students' future earnings. [Grogger \(1996\)](#) uses the High School and Beyond survey and finds that school expenses (including teacher salaries) have a positive but limited effect on students' after-school incomes⁵. The authors argue that studies using inter-state variation suffer from measurement errors and omitted variables, which distorts the estimation of the impact of school spending on student achievement. In a meta-analysis of 119 studies, [Hanushek \(1997\)](#) identifies 24 studies finding a significant positive relationship between teacher salaries and student achievement, 8 studies finding a significant negative relationship and 87 studies finding a non-significant positive or negative relationship. [Hanushek et al. \(1999\)](#) also finds that the salary offered has little effect on teacher certification test scores (evaluation of pedagogical knowledge). On the other hand, salaries appear to have a positive effect on student achievement in mathematics and reading. A 10 percentage point increase in the starting salary of teachers would increase mathematics achievement by 0.12 percentage points and reading achievement by 0.08 percentage points. However, the author notes that causal interpretation of these results is not guaranteed. Using sharp geographical discontinuities in teacher salaries, [Greaves and Sibieta \(2013\)](#) find little evidence supporting that an increase in teacher salaries improves pupils' performance in national assessments at age 11, although the authors rely on accurate estimates. The authors conclude that salary changes of the observed magnitude (5%) are unlikely to attract and retain good teachers. Overall, research in the field of education tends to show that strong wage increases are needed to significantly improve student achievement, particularly in socio-economically disadvantaged areas ([Prost, 2013](#); [Greaves and Sibieta, 2013](#)).

In a different approach, performance pay has aroused increasing interest among economists, employers and policy-makers in recent decades. Again, the effectiveness of this incentive is mixed and strongly depends on the performance indicators selected ([Bacache, 2009](#); [Bryson et al., 2017](#)). According to the theory of agency ([Ross, 1973](#); [Jensen and Meckling, 1976](#)), indexing remuneration to individual performance increases the level of effort by reducing moral hazard ; it also helps to attract (increase signalling) and retain (decrease adverse selection) the most competent people by promising a fairer remuneration of productivity. A first well known problem is that individual performance is rarely fully observed, which can lead individuals to neglect little or poorly observed tasks in favour

⁵A 10% increase in school spending would only increase the average salary of adult students by 0.68%.

of easily measurable tasks (e. g. quantity versus quality of production) (Holmstrom and Milgrom, 1991; Baker, 1992). Another risk is to discourage individuals by substituting a monetary value for the symbolic or intrinsic value of work (Frey S. and Jegen, 2000; Osterloh and Frey, 2000). Incentive pay is also likely to increase unequal treatment of users when the service is co-produced (for example, if teachers focus their efforts on the students who learn most quickly) (Firestone and Pennell, 1993; Bacache-Beauvallet, 2006; Jones, 2013). Specifically, in the field of education, pay for performance typically involves making teachers' salaries dependent on student achievement on standardized tests. Over the past decade, many states and school districts have implemented this type of compensation programs in the United States. Some authors have reported a decrease in collaboration between teachers (Jones, 2013) and a decrease in intrinsic motivation for teaching (Belfield and Heywood, 2008) following competition. Other authors point to the existence of cheating behaviour (Jacob and Levitt, 2003), teaching to the test, and show that teachers tend to focus their efforts on tested topics at the expense of other forms of knowledge and skills (Jacob, 2005). Overall, the results of studies that assessed the impact of pay-for-performance on student achievement are mixed (Jacob, 2005; Atkinson et al., 2009; Glewwe et al., 2010; Fryer, 2013; Jones, 2013; Balch and Springer, 2015). According to Rothstein (2015), the costs of individual financial incentives even outweighs their benefits in terms of student achievement and future incomes.

The effect of demand on labour supply has received much less attention in the economic literature. (Hanushek et al., 1999) and (Leigh, 2012) point out that failing to control for teachers demand in schools (often unobserved) makes it difficult to interpret causally the effect of salaries on the supply and quality of teachers obtained in most studies. In France, the demand for civil servants (85% of teachers) is determined by the number of places in each recruitment examination, while salaries are determined by a salary scale that varies mainly according to profession and seniority. (Daussin-Bénichou et al., 2015) show that the number of candidates for public service competitions generally increases with the number of positions offered. The authors estimate that an additional (primary or secondary) teaching position is associated with an additional 1.5 candidates over the 1980-2012 period. The authors also observe that the number of candidates for teaching posts is strongly and positively correlated with the unemployment rate during this period.

The empirical literature in economics of education is skeptical about the effectiveness of raising the degree requirement of teachers. The effect on recruitment depends on (1) the correlation between the level of education and the quality of a teacher, (2) the number of candidates discouraged by increased opportunity costs related to the level of education, and (3) the quality of these discouraged candidates. Most empirical studies report a positive but weak correlation between the quality of teachers and their degree level (Hanushek and Rivkin 2004, although: Jacob et al. 2016). Berger and Toma (1994) find a negative correlation between the degree level required to teach and students' SAT scores in the US. A limitation of these studies is that they rely on cross-sectional analysis and do not exploit exogenous change in the degree required to teach.

2.3 The effect of wage

2.3.1 A simple model

Let's start with a very simple model to examine how changes in teachers' and competing occupations wages theoretically alter teacher supply and characteristics.

Consider all individuals who meet the eligibility criteria for teaching positions (for example, the level of diploma required). Let's introduce u_i the utility of the teaching profession for the individual i , $c_i > 0$ the training cost to become a teacher (including preparation for the recruitment process), p_i the probability of being recruited as a teacher⁶ and v_i the expected utility of alternative professions. To simplify, assume that $c_i = \text{constant} = c$.

Suppose the expected utility of participating in the teacher recruitment process can be written $E[U^{trp}] = u_i p_i + v_i (1 - p_i) - c$. In other words, individuals get the utility $u_i - c$ if they succeed in becoming teachers and $v_i - c$ if they fail. If individuals are risk neutral, they participate in the recruitment process when $E[U^{trp}] > v_i$ i.e. $p_i (u_i - v_i) > c$. Because $p_i \in [0, 1]$ and $c > 0$, a necessary condition is that $u_i > v_i$. This condition simply requires that candidates for teaching posts have a greater preference for teaching than for other professions.

Let us call "potential candidates" individuals who satisfy a more restrictive condition: $u_i - c > v_i$. Potential candidates refer to individuals who participate in the teacher

⁶In the French context, it corresponds to the probability of success in the competitive examination

recruitment process when $p_i = 1$, i.e. when recruitment is certain. The number of potential candidates indicates the maximum number of candidates that recruiters could attract to teaching positions in a given context. Since in practice $p_i < 1$, "candidates" for teaching positions are actually potential candidates who satisfy $p_i > \frac{c}{u_i - v_i}$. On the contrary, all individuals eligible for teaching positions that meet $u_i - c < v_i$ will never consider becoming a teacher (in a given context), even if it is certain they will be recruited.

Figure 2.1 shows how eligible individuals are divided into "never candidates", "potential candidates" and "candidates" according to the value of the parameters p_i , $u_i - v_i$ and c .

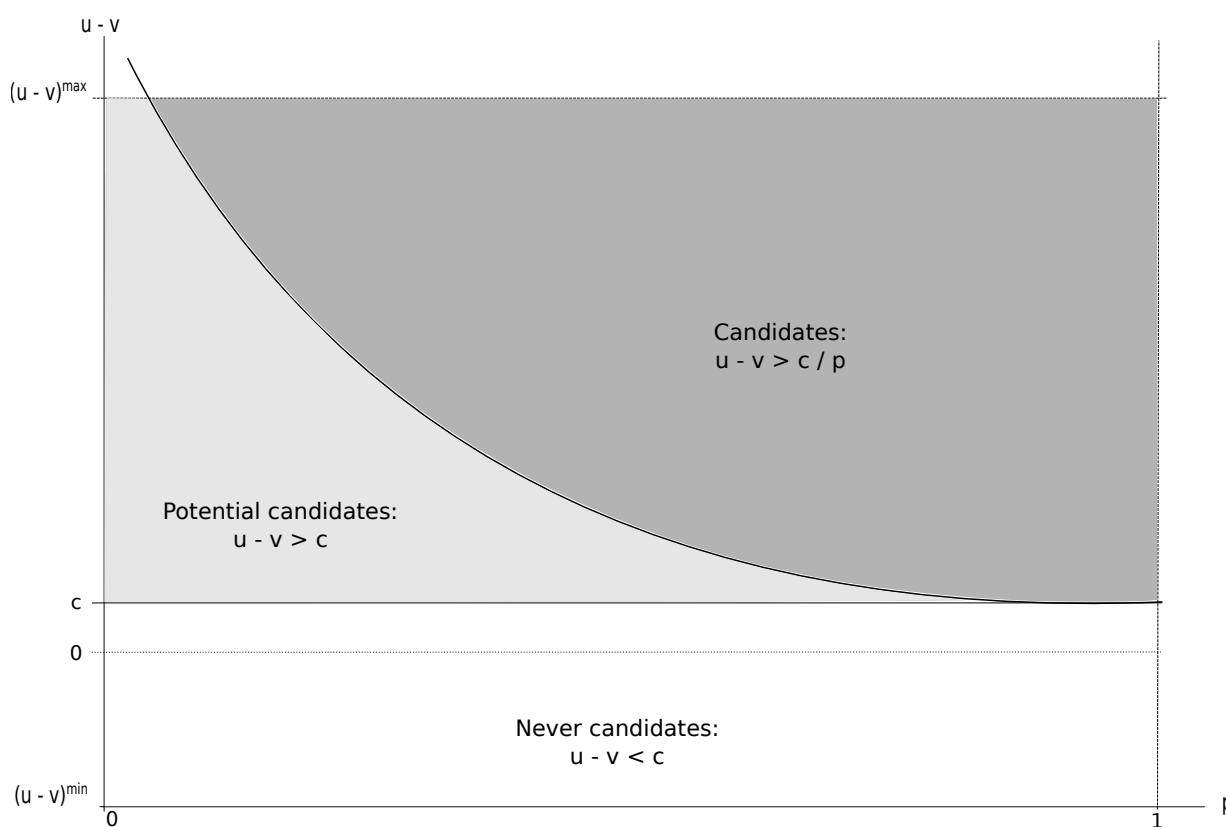


Figure 2.1 – Professional choices for individuals eligible to teaching positions

Note: Among persons eligible for teaching posts, this figure shows 1) individuals who are candidates (dark grey part - "candidates") because their relative utility for teaching is positive and their probability of recruitment is high, 2) individuals who are not candidates (white part - "never candidates") because their relative utility for teaching is negative and 3) individuals who are not candidates (light grey part) because their probability of recruitment is low although their relative utility for teaching is positive. With the notation of the model, u is the utility of the teaching profession, c is the opportunity cost, v is the expected utility of alternative professions and p is the probability of being recruited as a teacher.

2.3.2 Implication for teacher supply

To see more clearly how teacher wage affect teacher supply, assume that the utility of the teaching profession can be written $u_i = w^t + \nu_i$, and that the expected utility of alternative occupations can be written $v_i = w + \mu_i$, with w^t the average wage of teacher over the professional life⁷, ν_i a non-monetary preference for teaching, w the average wage (over the professional life) of alternative occupations and μ_i an expected (monetary and non-monetary) individual specific preference for alternative professions.

Let's note h the joint distribution function of random variables $(\nu - \mu, p)$ and $H_{\nu-\mu}$ the marginal cumulative distribution function of $\nu - \mu$. The share of potential candidates among eligible individuals is given by $P(u_i - v_i > c) = P(\nu_i - \mu_i > c - w^t + w) = 1 - H_{\nu-\mu}(c - w^t + w)$. Because $H_{\nu-\mu}$ is by definition an increasing function, the share of potential candidates among eligible individuals increases with the difference $w^t - w$ between the average salary of teachers and that of other professions.

Similarly, the share of candidates among eligible individuals writes $P(u_i - v_i > \frac{c}{p_i}) = P(\nu_i - \mu_i > \frac{c}{p_i} - w^t + w) = \int_{p=0}^1 \int_{\mu-\nu=\frac{c}{p_i}-w^t+w}^{+\infty} h(p, \nu - \mu) dp d(\nu - \mu) = G_c(w^t - w)$. Because h is a distribution function, G is an increasing function of $w^t - w$ which implies that the share of candidates among eligible individuals is an increasing function of the average wage of teachers and a decreasing function of the average wage of alternative occupations. We can also show that the share of candidates among potential candidates is an increasing function of $w^t - w$.

Therefore, under the assumptions of the model, an increase in teacher salaries, all other things being equal, will unambiguously increase the supply of teachers: eligible individuals will be more likely to be potential candidates and potential candidates will be more likely to be candidates.

2.3.3 Implication for teacher skills

Numerous studies show that the quality of teachers is one of the factors that most strongly influence the academic success, but also the behaviour and professional careers of students (Hanushek, 1992; Rockoff, 2004; Rivkin et al., 2005; Chetty et al., 2014). The determinants of teacher quality are multidimensional and are the subject of active research in the educational sciences. Our theoretical framework introduces two potential

⁷In France, teachers' wage follows a salary scale than essentially depends on seniority

determinants of teaching quality: the non-monetary preference for teaching ν , which may reflect the individual's intrinsic motivation for the profession, and the probability of being recruited p , which refers to the abilities a measured during the recruitment process and likely to correspond to the knowledge needed to teach. Assuming that motivation and knowledge contribute to teacher quality, noted Q , we can write $Q = f(\nu, p, \theta)$, with θ a set of unobserved determinants of teacher quality.

Unlike the case of teacher supply, the effect of salary on the teacher quality is not a simple problem. With the notations and assumptions of the model, the wage effect depends on the relationship between the relative utility of teaching $u_i - v_i$ and the probability of success p_i . But even more, the effect of wage depends on the nature (intensity and functional form) of the relationship between p_i , ν_i and the quality Q_i of teacher i .

In our setting, the candidates' average knowledge is related to the expression:

$E[p_i | candidate] = \int_{p=0}^1 \int_{\mu-\nu=\frac{c}{p_i}-w^t+w}^{+\infty} p h(p, \nu - \mu) dp d(\nu - \mu)$. Assuming that $\nu - \mu \perp p$, $p \sim \mathcal{U}[0, 1]$ and $\nu - \mu \sim \mathcal{U}[(\nu - \mu)^{min}, (\nu - \mu)^{max}]$, we can show that $E[p_i | u_i - v_i > \frac{c}{p_i}]$ is an increasing function of $w^t - w$ for $w^t - w < c - (\nu - \mu)^{min}$ and a decreasing function of $w^t - w$ for $w^t - w > c - (\nu - \mu)^{min}$ (all other things being equal). However, the result may be completely different if we change one or more of these simplifying assumptions.

The effect of wage on the quality of teachers recruited is even more complex to predict. With the simplifying assumptions above, we can show that teachers' average knowledge increases (up to a ceiling) with teachers' salary. This result, which seems to contradict the decrease in the average knowledge of candidates with salary, comes from the fact that a higher wage attracts very competent candidates. Thus, while candidates' average knowledge decreases with salary due to the arrival of large numbers of less qualified candidates, recruiters are more likely to fill teaching positions with highly qualified individuals (the most important aspect for a recruiter is the number of candidates at the top of the skill distribution).

This very simple model already gives an overview of the theoretical effect of an increase in teachers' salaries. Although the effect on the number of candidates is clearly positive, the effect on the quality of teachers recruited depends on the nature of the relationship between the aptitude criteria measured at recruitment (a_i) and the actual quality of teachers. The wage effect will also depend on the relationship between the relative utility

of the individual for teaching $u_i - v_i$ and the probability of recruitment p_i , which depends in part on how the labor market values the skills a_i measured during the recruitment process. Nevertheless, a change in hypothesis can lead to very different results, which underlines the importance of empirical evaluation before drawing a conclusion. The following section presents the results obtained when the determinants of the recruitment probability p_i are made more explicit.

2.4 The effect of labor demand

Let us consider the case where teachers are recruited through competitive examinations once a year. This section examines the effect of teacher demand on teacher supply and *skills*.

2.4.1 Model

Notations

With the notations of the previous section, the expected utility of participating in the teacher recruitment process is $U^{trp} = u_i p_i + v_i (1 - p_i) - c$, while the alternative expected utility is v_i . Remember that p_i is the probability of success in the teacher recruitment exam, u_i is the utility for teaching, v_i is the expected utility for alternative professions and c is the cost of preparation to become a teacher.

Consider N_E the random variable that indicates the number of individuals eligible for teaching positions (for example, all bachelors' graduates in a given year), N_{PC} the number of potential candidates, N_C the number of candidates and N_P the number of teaching positions in a given year. Assume also that $N_{PC} \rightarrow \infty$ when $N_E \rightarrow \infty$.

I note a_i the skills measured at recruitment and F_a^∞ the corresponding asymptotic cumulative distribution function for potential candidates⁸. I note F_a the empirical cumulative distribution function of skills for potential candidates in a given year.

Probability of success p_i

Suppose the recruitment exam assesses the potential candidates' skill a_i with a zero mean and variance σ_η measurement error η_i . I assume that each individual knows his or her skill

⁸In this section, all distributions with the symbol ∞ are assumed to be continuous.

level (a_i) as well as the variance of the measurement error (σ_η), but ignores the specific shock η_i he or she will receive upon recruitment. I note x_i the score the individual i gets when participating in the exam: $x_i = a_i + \eta_i$. I also note F_η^∞ the asymptotic cumulative distribution of η_i and F_x^∞ the asymptotic cumulative distribution of the exam score x_i (hypothetically) obtained by potential candidates⁹. Teacher positions (N_P in total for a given year) are allocated to the highest ranked candidates on the recruitment test. This means that all candidates who score higher than the last recruited candidate are necessarily recruited.

To clarify the situation, consider x_1, \dots, x_{N_C} the scores obtained by N_C candidates who participate in the teacher recruitment process in a given year. $\forall(i, j)$ such that $x_i \neq x_j$, there is (almost surely) a unique permutation σ^{-1} of $[1, N_C]$ that orders the x_i such that:

$$\underbrace{x_{\sigma^{-1}(1)} > x_{\sigma^{-1}(2)} > \dots > x_{\sigma^{-1}(N_P)}}_{\text{recruited : } x_i > s} > \dots > x_{\sigma^{-1}(N_C)}.$$

Transformation $\sigma(i)$ provides the rank of individual i and $\sigma^{-1}(k)$ provides the identity of the k^{th} individual ranked by decreasing values of x . Therefore, the recruitment threshold s for a given year is defined as the score of the last candidate recruited: $s = x_{\sigma^{-1}(N_P)}$ ¹⁰. The probability that individual i passes the exam conditional on s and ability a_i is: $p_{i,s} = P(x_{\sigma^{-1}(\sigma(i))} \geq x_{\sigma^{-1}(N_P)} | x_{\sigma^{-1}(N_P)}, a_i) = P(x_i > s | s, a_i) = P(a_i + \eta_i > s | s) = 1 - F_\eta^\infty(s - a_i)$ and the expected probability that individual i passes the exam conditional on a_i writes: $p_i = E[p_{i,s} | a_i] = E[1 - F_\eta^\infty(s - a_i) | a_i]$.

Let's note $F_x^{c \infty}$ the asymptotic cumulative distribution of the exam score of candidates and note x_P^∞ the unique solution of $F_x^{c \infty}(x_P^\infty) = E[\frac{N_P}{N_C}]$, with $E[\frac{N_P}{N_C}]$ the expected share of positions per candidate for a given year. The asymptotic convergence of an empirical quantile ensures that $s \xrightarrow[N_E \rightarrow \infty]{a.s.} x_P^\infty = F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}])$.

Therefore, the (expected) probability that the individual i will pass the examination (conditional on his or her skill a_i and conditional on the empirical distribution of candidates' skills in a given year) is: $p_i = E[p_{i,s} | a_i] = E[1 - F_\eta^\infty(s - a_i) | a_i] = E[1 - F_\eta[F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) - a_i] | a_i] = 1 - F_\eta[F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) - a_i]$.

$$p_i = 1 - F_\eta^\infty[F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) - a_i] \quad (1)$$

⁹This is the score that potential candidates would get if they participated in the recruitment exam. In practice, we saw in the previous section that only a fraction of potential candidates actually participate in the recruitment process (these are the candidates).

¹⁰When there are k distinct individuals such that $x = s$, replace N_P by $N_P + k$

For a given year, the probability of passing the exam increases with the individual's skill a_i , increases with the number of teaching positions N_P and decreases with the number of candidates N_C . The probability of being recruited also depends on the asymptotic cumulative distribution function of the examination score for candidates $F_x^{c\infty}$ (entirely determined by the asymptotic cumulative distribution functions of the examination score for potential candidates F_x^∞ and by the expected number of positions per potential candidate $E[\frac{N_P}{N_{PC}}]$) and the error term F_η^∞ . Note that expression (1) is similar to the expression obtained by Dal Bó et al. (2013).

Strategic behavior

Suppose individuals anticipate that the probability of passing the exam depends on the decision of other potential candidates. Intuitively, if the number of (good) candidates N_C increases for a given number of teaching positions N_P , competition will intensify and the least competent individuals will be discouraged from participating. As a result, the number of candidates will decrease and the average skill level of candidates will tend to increase. If individuals act strategically, the decision to participate in the recruitment process will depend on the distribution $F_x^{c\infty}$ of the scores of candidates when the game is reproduced a large number of times.

The decision to apply to teacher recruitment process

Suppose individuals are strategic and risk neutral. The potential candidate i participates in the teacher recruitment process if $u_i p_i + w_i (1 - p_i) - c > v_i$ i.e. $p_i > \frac{c}{u_i - w_i}$.

As in the previous section, I assume that the utility of teaching and the expected utility of alternative professions can be written $u_i = w^t + \nu_i$ and $v_i = w + \mu_i$, with w^t the average wage of teacher over the professional life, ν_i a non-monetary preference for teaching, w the average wage (over the professional life) of alternative occupations and μ_i an expected (monetary and non-monetary) individual specific preference for alternative professions.

To simplify, I suppose that all individuals face the same relative utility for teaching: $u_i - v_i = w^t + \nu - w - \mu = u - v$ ¹¹. This additional assumption combined with expression

¹¹The results in the following paragraphs can also be interpreted conditionally to the value of $\nu_i - \mu_i$, which allows to remain in a more general framework

(1) allows us to write that the potential candidate i participates in the recruitment process if $a_i > F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) - F_\eta^{-1 \infty}(1 - \frac{c}{u-v}) = x_P^\infty - F_\eta^{-1}(1 - \frac{c}{u-v}) = B^\infty$. All applicants in a given year meet the condition $a_i > B^\infty$.

From the statistician's perspective, the share of potential candidates involved in the teacher recruitment process satisfies: $P(a_i > B^\infty | B^\infty) = 1 - F_a^\infty(B^\infty) = E[\frac{N_C}{N_{PC}}]$, where F_a^∞ is the asymptotic cumulative distribution of potential candidates' competencies. This relationship allows to derive the expression:

$$F_a^{-1 \infty}(1 - E[\frac{N_C}{N_{PC}}]) = F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) - F_\eta^{-1 \infty}(1 - \frac{c}{u-v}) \quad (2)$$

Expression (2) relates the expected number of candidates per potential candidate to the expected number of positions per candidate. This expression also involves the inverse cumulative distribution functions $F_a^{-1 \infty}$, $F_x^{c-1 \infty}$, F_η^{-1} , the cost of preparing individuals for the teacher recruitment process c and the relative utility for teaching $u - v$.

Additional assumptions for distribution functions

I assume that F_a^∞ and F_η^∞ are normal cumulative distributions: $N(m_a, \sigma_a^2)$ and $N(0, \sigma_\eta^2)$. This condition implies that the examination score of potential candidates follows a normal distribution F_x^∞ , which can be noted $N(m_x, \sigma_x^2)$, with $m_x = m_a$ and $\sigma_x^2 = \sigma_a^2 + \sigma_\eta^2$.

With these additional assumptions, the asymptotic cumulative distribution of candidates' examination scores is written:

$$\begin{aligned} F_x^{c \infty}(z) &= P(x_i < z | a_i > B^\infty, B^\infty) = P(a_i + \eta_i < z | a_i > B^\infty, B^\infty) \\ &= \frac{1}{P(a_i > B^\infty | B^\infty)} \int_{B^\infty}^{+\infty} F_\eta^\infty(z - a) f_a(a) da \end{aligned}$$

The cumulative distribution of candidates' competencies becomes simply:

$$F_a^{c \infty}(z) = P(a_i < z | a_i > B^\infty, B^\infty) = \frac{F_a^\infty(z) - F_a^\infty(B^\infty)}{1 - F_a^\infty(B^\infty)} = \frac{F_a^\infty(z) - 1}{E[\frac{N_C}{N_{PC}}]} + 1$$

Approximation

In order to obtain simple analytical expressions for the expected number of candidates, the expected examination score of candidates and the expected examination score of recruited teachers, I approximate the term $x_P^\infty = F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}])$ that appears in relations (1) and (2). I show in the Appendix that for $\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$, $x_P^\infty = F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) \approx F_x^{-1 \infty}(1 - E[\frac{N_P}{N_{PC}}])$.

A first implication of the condition $\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$ is that potential candidates participate in the recruitment process even when the difference between their skill level a_i and the expected recruitment threshold x_P^∞ far exceeds the standard deviation of the measurement error: $a_i - x_P^\infty \gg \sigma_\eta$.

A second implication is that $P(a_i < B^\infty | x_i \geq x_P^\infty) \ll 1$, which means that only potential candidates with a very low probability of success give up participating in the recruitment competition (which corresponds to the contraposition of the previous implication).

A third implication is that $F^{-1}(1 - \frac{c}{u-v}) \gg 1$ (with F the normal cumulative distribution), which is equivalent to $\frac{u-v}{c} \gg 1$. This means that the difference between the utility of the teaching profession and the expected utility of alternative professions far exceeds the cost of the competition.

It follows from the above considerations that the condition $\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$ is satisfied when the teaching profession is "sufficiently" attractive compared to competing professions. In practice, this means that there are significantly more candidates than teaching posts, which can be tested empirically.

2.4.2 Implication for teacher supply

Expected number of candidates per eligible individual

Using the expression (2) and the approximation in Section 2.4.1, the expected number of candidates per potential candidate is related to the expected number of positions per potential candidate according to:

$$F_a^{-1} \infty (1 - E[\frac{N_C}{N_{PC}}]) \approx F_x^{-1} \infty (1 - E[\frac{N_P}{N_{PC}}]) - F_\eta^{-1} (1 - \frac{c}{u-v})$$

Using the assumptions about the forms of distributions formulated in the Section 2.4.1, we obtain:

$$E[\frac{N_C}{N_{PC}}] \approx F[1 - \frac{\sigma_x}{\sigma_a} F^{-1}(1 - E[\frac{N_P}{N_{PC}}]) - \frac{\sigma_\eta}{\sigma_a} F^{-1}(1 - \frac{c}{w^t - w + \nu - \mu})] \quad (3)$$

with F^{-1} the inverse of the cumulative distribution function of the normal law $N(0, 1)$.

Conditionally to the value of $\nu_i - \mu_i$, the expected number of candidates per potential candidate $\frac{N_P}{N_{PC}}$ increases with the expected number of teaching positions per potential

candidate $\frac{N_P}{N_{PC}}$. Consequently, the elasticity of the expected number of candidates with the number of positions is always positive. However, it decreases with the number of teaching positions, meaning that the ability to attract more candidates (per potential candidate) by increasing the demand for teachers decreases with the number of teaching positions (per potential candidate).

The model also illustrates the effect of teacher wage on teacher supply based on the assumptions in this section. First, remember that potential candidates satisfy $u_i - c < v_i$ and that $H_{\nu-\mu}$ is the asymptotic cumulative distribution function of $\nu_i - \mu_i$. Therefore, we saw in Section 3.4.5 that the expected share of potential candidates in the eligible population is given by $E[\frac{N_{PC}}{N_E}] = 1 - H_{\nu-\mu}(c - w^t + w)$, which implies that an increase in teachers' salary w^t , all things being equal, increases the share of potential candidates among eligible persons. Second, conditional on the value of $\nu_i - \mu_i$, expression (3) explicitly shows that an increase in teachers' wage w^t will increase the expected number of candidates per potential candidate $E[\frac{N_C}{N_{PC}}]$ (see Section 2.3.2 for a generalization of this result). Therefore, an increase in teachers' wage w^t (for a fixed number of teaching positions N_P and conditional on preferences (ν_i, μ_i) and average salary w in alternative occupations) will increase the share of candidates among eligible individuals $E[\frac{N_C}{N_E}]$:

$$E[\frac{N_C}{N_E}] \approx K(w^t - w - c) F\left[1 - \frac{\sigma_x}{\sigma_a} F^{-1}\left(1 - E\left[\frac{N_P}{N_{PC}}\right]\right) - \frac{\sigma_\eta}{\sigma_a} F^{-1}\left(1 - \frac{c}{w^t - w + \nu - \mu}\right)\right] \quad (3')$$

with K an increasing function of $w^t - w - c$ ¹². Note that expressions (3) and (3') are valid as long as $\forall N, N_P \leq N_C \leq N_{PC}$.

Teacher shortage

According to expression (3), when the number of positions N_P increases for a given (expected) number of potential candidates $E[N_{PC}]$, the (expected) number of candidates increases until $E[N_C] = E[N_{PC}]$. If $N_P > N_{PC}$ then $N_P > N_C$ which means there is a shortage of teachers. The (expected) number of unfilled positions is therefore $N_P - E[N_C]$.

In other words, the model predicts that $E[\frac{N_C}{N_E}] \rightarrow K(w^t - w - c) = \alpha$, where α represents the share of potential candidates in the eligible population, when the number of teaching positions increases.

¹²Expression (3') applies conditionally to the value of $\nu_i - \mu_i$ or assuming that $\nu_i - \mu_i = \text{constant} = \nu - \mu$ for potential candidates

2.4.3 Implication for teacher *skills*

Expected exam score of candidates

The expected exam score of candidates is equal to the expected skills of candidates. With the previous assumptions we obtain (see proof in the Appendix):

$$E[x_i|a_i > B^\infty] = E[a_i|a_i > B^\infty] = m_a + \frac{\sigma_a}{E[\frac{N_C}{N_{PC}}]} f(F^{-1}(1 - E[\frac{N_C}{N_{PC}}])) \quad (4)$$

First, we note that the expected examination score of candidates increases with the average skills of potential candidates m_a .

Second, after replacing $E[\frac{N_C}{N_{PC}}]$ by expression (3'), we show that the expected examination score of candidates decreases, all things being equal, with the expected number of positions per eligible person $E[\frac{N_P}{N_E}]$. The intuition behind this result is that an increase in the number of positions (for a fixed number of eligible individuals) increases the probability p_i of being recruited for all individuals. As a result, a higher demand for teachers attracts less qualified candidates who previously declined to participate in the competition.

Third, the expected examination score of candidates decreases with the wage of teachers conditional on the value of $\nu_i - \mu_i$ ¹³. This result comes from the fact that a higher salary will attract a higher fraction of low-skilled candidates to teaching positions.

Expected exam score of teachers

Using the approximation in section 2.4.1, the expected exam score of recruited teachers is (see proof in the Appendix):

$$E[x_i|a_i \geq B^\infty, x_i \geq x_P^\infty] \approx m_a + \frac{\sqrt{\sigma_a^2 + \sigma_\eta^2}}{E[\frac{N_P}{N_{PC}}]} f(F^{-1}(1 - E[\frac{N_P}{N_{PC}}])) \quad (5)$$

It is not possible to obtain a simple expression giving the expected level of competence of the teachers recruited.

Expression (5) first reveals that the expected exam score for recruited teachers in-

¹³The expected examination score of candidates decreases unconditionally with the salary of teachers when $\nu_i - \mu_i \perp a_i$. However, if $\nu_i - \mu_i \not\perp a_i$ an increase in teachers' salary is likely to change the distribution of potential candidates' competencies, which in turn may change the conclusion regarding candidates' expected score

creases with the average skill level of potential candidates m_a . Second, it reveals that the expected examination score for recruited teachers decreases, all other things being equal, with the expected number of positions per eligible person $E[\frac{N_E}{N_E}]$. The intuition is that the first positions are allocated to the most competent candidates; then recruiters select less competent candidates until all teaching positions are filled. As the number of highly competent candidates is limited, the average skill level of teachers decreases with the number of posts to be filled.

Expression (5) also indicates that the expected examination score for teachers increases with salary conditional on the value of $\nu_i - \mu_i$. This result remains unconditionally valid as long as $\nu_i - \mu_i \perp a_i$ ¹⁴. This result is explained by the fact that higher salaries attract more qualified candidates who are then recruited to become teachers¹⁵.

2.5 The effect of degree requirement

The models presented in Sections 3.4.5 and 2.4 provide indications of how a change in the diploma required to teach can affect the supply and characteristics of teachers. First, an increase in the level of qualification is likely to change the expected number of eligible individuals and the distribution of their preferences and skills. These changes may then affect the expected number of potential candidates as well as the distribution of their preferences and skills. Finally, the supply and characteristics of teachers are likely to change according to expressions (3) to (5).

The next section will formalize these mechanisms. In the baseline scenario, a bachelor's degree is required to teach, while in the scenario of interest, teachers must have a master's degree. The comparison of the results obtained in these two scenarios will make it possible to conclude on the theoretical impact of the reform of teacher qualification implemented in France in 2011.

2.5.1 Educational and professional choices when teachers are required to have a bachelor's degree

In what follows, candidates to teaching positions must have a bachelors' degree at least.

¹⁴On the other hand, the effect of salary on teachers' average score could be different if $\nu_i - \mu_i \not\perp a_i$.

¹⁵Remember that only the number of candidates at the top of the skill distribution affects the average skill level of teachers.

Notations

Consider all the students who recently obtained a bachelor's degree. Note v_i^b the expected utility of the individual i for bachelor-level occupations (other than teaching), v_i^m the expected utility of occupations that require a master's degree, c_i^m the (opportunity) cost of preparing for a master's degree and $q_i \in]0, 1[$ the probability of obtaining a master's degree. The q_i probability measures the individual's ability to obtain a master's degree, which depends on his skills, effort and risk of dropping out of school.

Using the same notations as in the previous sections, the individual i is now characterized by three random variables (u_i, p_i, q_i) , where u_i represents the utility for teaching, $p_i \in]0, 1[$ is the probability of being recruited as a teacher and $q_i \in]0, 1[$ is the probability of obtaining a master's degree.

To further simplify the problem, I examine the effect of the degree required to teach on the distribution of potential candidates, which is equivalent to considering the simple case where the probability of being recruited is 1 (see Section 2.3.1). In a second step, we can use expressions (3) to (5) to infer the effect on teachers' supply and skills.

Assumptions

To simplify, I assume that individuals do not change professions during their lifetime and do not return to university once they have made a career choice. In France, this assumption is reasonable given that less than 2% of teachers leave their profession and less than 2% of bachelor graduates return to school after employment at this level of education. I also assume that people who graduate with a master's degree do not intend to become teachers. This result is consistent with the data as only 6% of teacher candidates have a master's degree or higher¹⁶. I assume that those who choose to prepare for the teacher recruitment process do not enroll in a master's program in the meantime (although 20% of candidates do so in practice). Consequently, those who fail to be recruited as teachers must accept employment at the bachelor's degree level according to the model.

As in the previous sections, I suppose that the utility of the teaching profession can be written $u_i = w^t + \nu_i$, where w^t is the average teacher's salary (over a lifetime) and ν_i is a non-monetary preference for teaching. Similarly, I suppose that the expected utility for alternative professions can be written $v_i^b = w^b + \mu_i$ at the bachelors' level and $v_i^m = w^m + \mu_i$

¹⁶During the 2003-2009 period when a bachelor's degree is required to teach

at the master's level, where w^b and w^m are the average salaries of bachelor's and master's graduates (over a lifetime), while μ_i is the individual-specific expected utility (monetary and non-monetary) for alternative occupations. I also consider that all individuals face the same opportunity cost when they enroll in a master's program $c_i^m = c^m$.

To simplify the notations, I note $u'_i = w^t + \nu_i - \mu_i - c^t = w^t + \nu'_i$ the relative utility of teaching. Variables u'_i and q_i are assumed to be uniformly distributed, with $q_i \in [0, 1]$ and $\nu'_i \in [\nu'_{min}, \nu'_{max}]$. I also assume that the probability of obtaining a master's degree does not depend on the relative utility of teaching. ($u'_i \perp q_i$) and that $u'_{min} < w^b < w^m < u'_{max}$, with $u'_{min} = w^t + \nu'_{min}$ and $u'_{max} = w^t + \nu'_{max}$.

The decision to apply to teacher recruitment process

Under the previous assumptions, new bachelor's graduates have two educational options:

- 1) Stop their studies and get the expected utility: $\max(v_i^b, u_i - c)$
- 2) Enroll in a masters' program and get the expected utility:

$$q_i v_i^m + (1 - q_i) \max(v_i^b, u_i - c) - c^m$$

More specifically, individuals who complete a masters' degree (with probability q_i) obtain the utility $v_i^m - c^m$ while those who fail (with probability $1 - q_i$) get the utility $\max(v_i^b, u_i) - c^m$.

Assuming that bachelor's graduates maximize their expected utility and are risk neutral, we identify two types of people who intend to become teachers:

- 1) Individuals who leave school and intend to enter the teaching profession directly (they satisfy $u_i - c > \max(v_i^b, v_i^m - \frac{c^m}{q_i})$ i.e. $u'_i > \max(w^b, w^m - \frac{c^m}{q_i})$)
- 2) Individuals who are continuing their education but are not obtaining a master's degree and who intend to enter the teaching profession later (they satisfy $w^b < u'_i < w^m - \frac{c^m}{q_i}$ and $q_i > \frac{c^m}{w^m - w^b}$)

Figure 2.2 illustrates the theoretical educational and occupational choices of a cohort of bachelor's graduates based on the values of (u'_i, q_i) . The dark grey portion represents individuals who may be involved in the teacher recruitment process directly after graduation (i.e. potential candidates). The white part such as $q_i < \frac{c^m}{w^m - w^b}$ corresponds to individuals who take a job at the bachelor level and will never be teachers ($u_i - c < v_i^b$ i.e. $u'_i < w^b$). The white and light grey portions such as $q_i > \frac{c^m}{w^m - w^b}$ represents bachelor's graduates enrolling in a master's program. The white part corresponds to individuals

who will never become teachers: they take a master's level job if they succeed (probability q_i) and a bachelor's level job if they fail (probability $1 - q_i$). The light grey portion corresponds to those who take a job at the master's level if they succeed (probability q_i) and intend to become teachers if they fail (probability $1 - q_i$).

Note that without the opportunity to become a teacher, people who satisfy $q_i > \frac{c^m}{w^m - w^b}$ would systematically enroll in a master's program. Consequently, some individuals (those who belong to the dark grey portion and satisfy $q_i > \frac{c^m}{w^m - w^b}$) give up enrolling in master's programs because they have a high relative utility for teaching and a relatively low probability of obtaining a master's degree. In this context, the opportunity cost c^m associated with preparing for a master's degree and the uncertainty of obtaining a degree favor applications for teaching positions.

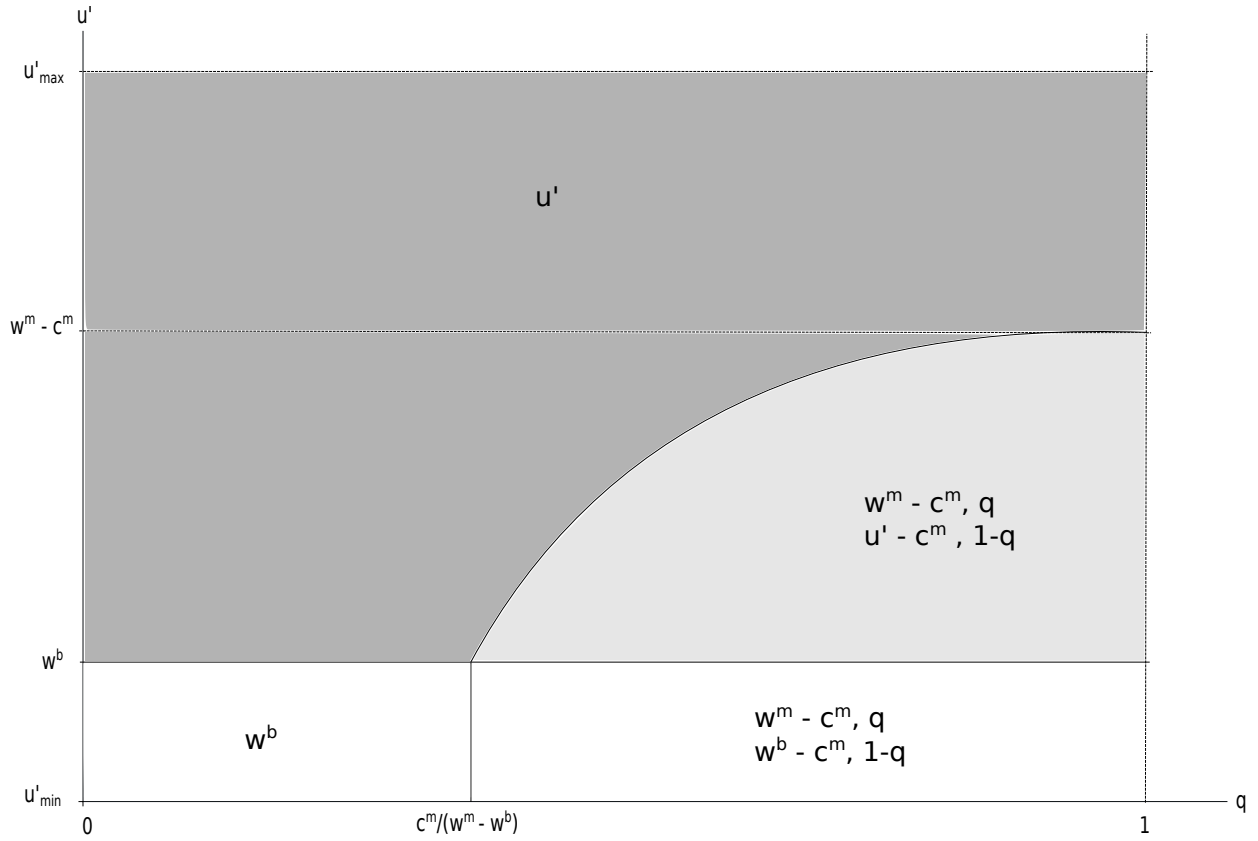


Figure 2.2 – Professional choices of bachelor's graduates when a bachelor's degree is required to teach.

Note: Among bachelors' graduates, this figure shows 1) individuals who are potential candidates (dark grey part) because their utility for teaching is greater than their utility for professions at the master's level, or because their probability of obtaining a masters' degree is low, 2) individuals who are not candidates (white part) because their utility for teaching is less than their utility for professions at the bachelor's level and 3) individuals who are potential candidates (light grey part) if they fail to obtain a masters' degree because their utility for teaching is less than their utility for professions at the master's level. With the notation of the model, u' is the utility of the teaching profession (including training cost), w^m is the expected wage of professions at the masters' level, c^m is the opportunity cost of preparing a masters' degree, w^b is the expected wage of professions at the bachelors' level and q is the probability of obtaining a masters' degree.

Implications for teacher supply

With the assumptions $u' \perp q$, $u' \sim \mathcal{U}[u'_{min}, u'_{max}]$ and $q \sim \mathcal{U}[0, 1[$, we get the share of potential candidates ($p_i = 1$) among bachelor's graduates P_t^b :

$$P_t^b = \frac{\nu'_{max} + c^m}{\nu'_{max} - \nu'_{min}} + \frac{2w^t - w^b - w^m}{2(\nu'_{max} - \nu'_{min})} - \frac{(c^m)^2}{2(\nu'_{max} - \nu'_{min})} \frac{1}{w^m - w^b} \quad (6)$$

It is clear from this model that the number of potential candidates for teaching positions increases with teachers' wage and decreases with the average salaries of bachelor's and master's graduates. When the salaries of bachelor's and master's graduates increase by one unit, teachers' salaries must also increase by one unit in order to maintain the same number of potential candidates for teaching positions.

Relationship (6) makes it possible to deduce the elasticity of teacher supply with wage:

$$\epsilon_t^b = \frac{w^t}{P_t^b} \frac{dP_t^b}{dw^t} = \frac{1}{\nu'_{max} - \nu'_{min}} \frac{w^t}{P_t^b}$$

Implications for teacher *skills*

This theoretical framework introduces two potential components of teaching quality: the non-monetary preference for teaching ν and the probability of obtaining a master's degree q , which measures the academic skills of individuals. Assuming that motivation and academic skills contribute to teacher quality Q , we can write $Q = f(\nu, q, \theta)$, with θ a set of unobserved determinants of teacher quality.

With the assumptions of the model, we can show that the average academic knowledge of potential candidates $E(q|potential\ candidate)$ increases with teachers' salary when a bachelor's degree is required. Assuming that $\nu_i \perp q_i$ and $\mu_i = constant$, we also show that the average motivation (non-monetary preference) of potential candidates decreases with teachers' wage. However, when these assumptions are relaxed, the effect of a salary increase on the motivation and academic skills of potential candidates becomes ambiguous (see for example the theoretical discussion of (Dal Bo et al., 2013)¹⁷).

We can also show that for a fixed number of positions (and assuming that $\nu_i \perp q_i$), teachers' academic knowledge increases with teachers' salary. As we already mentioned in Section 2.3.3, the real effect of the salary depends on the relationship between ν , q and the actual determinants of teacher quality.

¹⁷Empirically, the authors find a positive relationship between the salary and motivation of candidates for positions in the Mexican public sector. This result suggests that skills and motivation are actually positively correlated, which would tend to invalidate the $\nu_i \perp q_i$ hypothesis.

2.5.2 Educational and professional choices when teachers are required to have a masters' degree

In this subsection, teachers are required to have a master's degree. The reform of teacher qualification levels introduced in 2011 in France allows candidates to participate in the recruitment process as soon as they are enrolled in a master's program (before they graduate). Applicants must have completed a master's degree before being officially recruited.

The decision to apply to teacher recruitment process

When a master's degree is required to teach, recent bachelor's graduates have two options:

- 1) Stop their studies at the bachelor level and accept a job with the expected utility v_i^b
- 2) Enroll in a masters' program and obtain the expected utility:

$$q_i \max(v_i^m, u_i - c) + (1 - q_i)v_i^b - c^m$$

Among the individuals who enroll in a master's program, those who complete a master's degree (with probability q_i) intend to become teachers (if $u_i - c > v_i^m$) or choose another job of expected utility v_i^m (if $u_i - c < v_i^m$). Those who fail to obtain a master's degree (with probability $1 - q_i$) accept a job at the bachelor's level with the expected utility v_i^b .

Assuming that bachelor's graduates maximize their expected utility and are risk neutral, those who enroll in a master's program satisfy: $q_i > \min(\frac{c^m}{v_i^m - v_i^b}, \frac{c^m}{u_i - v_i^b})$ i.e. $q_i > \min(\frac{c^m}{w^m - w^b}, \frac{c^m}{u_i' - w^b})$. In particular, individuals who want to become teachers satisfy $q_i > \frac{c^m}{u_i' - w^b}$ and $u_i' > w^m$. However, only master's graduates (at the end of the school year) will eventually be recruited as teachers.

Figure 2.3 summarizes the decisions of bachelor's graduates based on the values of (u_i', q_i) . The dark grey portion represents individuals who are enrolled in master's programs and intend to become teachers upon graduation. The light grey portion corresponds to individuals who are enrolling in a master's program but do not intend to become teachers. The white part represents people who stop their studies after obtaining a bachelor's degree and take a job at this level of qualification.

Compared to the situation where a bachelor's degree is required to become a teacher, all individuals who meet $q_i > \frac{c^m}{w^m - w^b}$ are now enrolled in a master's program. Therefore, the model predicts that increasing the level of qualification required to teach will lead to

an increase in the number of master's students.

On the other hand, the share of bachelor's graduates who intend to become teachers decreases when a master's degree is required. First, all individuals who satisfy $q_i > \frac{c^m}{w^m - w^b}$ and $u' < w^m$ are no longer candidates for teaching positions as they prefer alternative professions at the master's level. Second, bachelor's graduates who satisfy $q_i < \frac{c^m}{w^m - w^b}$, that is, who have a low probability of obtaining a master's degree, now interrupt their studies after graduation and no longer intend to become teachers. Third, individuals who satisfy $u' > w^m$ now face the risk of failing a master's degree and being unable to become teachers.

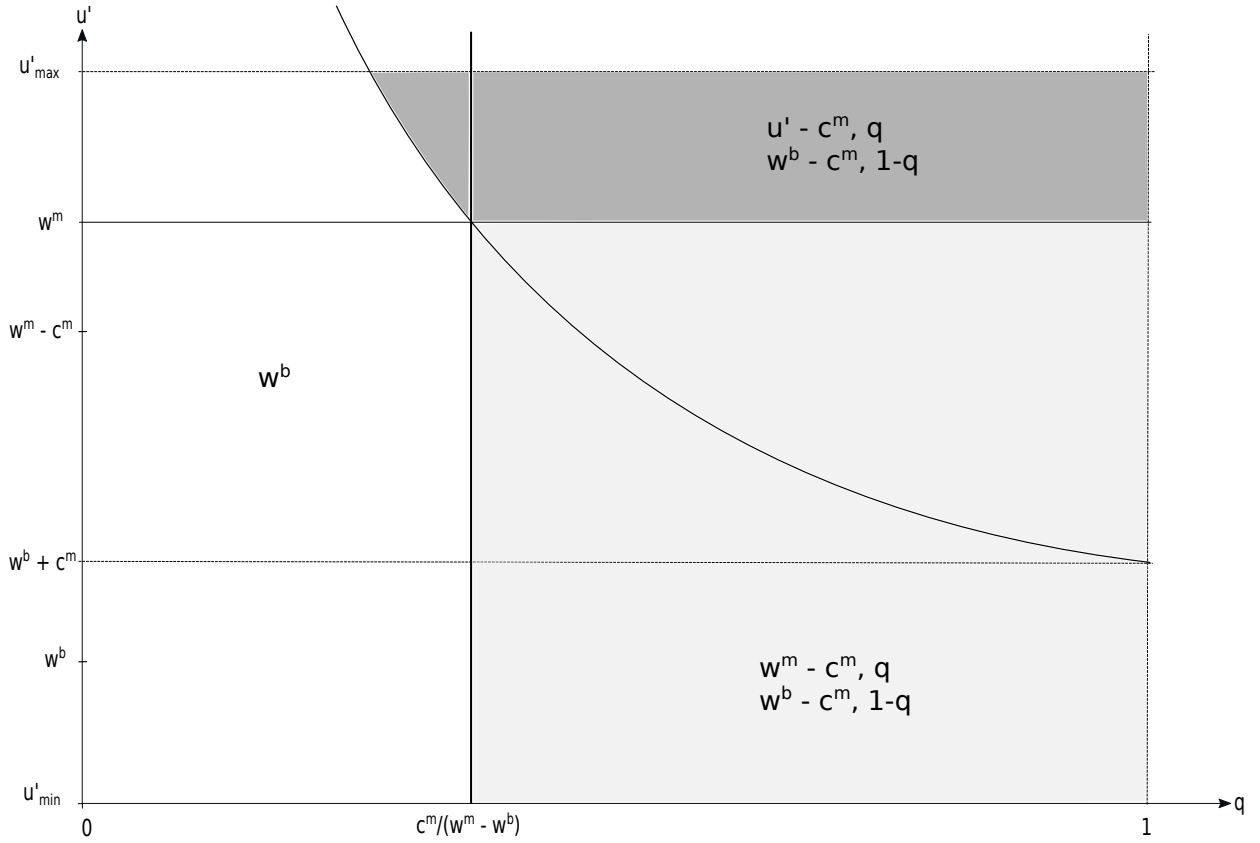


Figure 2.3 – Professional choices for bachelor's graduates when a master's degree is required to teach.

Note: Among bachelors' graduates, this figure shows 1) individuals who are potential candidates (dark grey part) because their utility for teaching is greater than their utility for professions at the master's level, and because their probability of obtaining a masters' degree is high, 2) individuals who are not candidates (white part and light grey part) because their utility for teaching is less than their utility for professions at the master's level, or because their probability of obtaining a masters' degree is low. With the notation of the model, u' is the utility of the teaching profession (including training cost), w^m is the expected wage in professions at the masters' level, c^m is the opportunity cost of preparing a masters' degree, w^b is the expected wage in professions at the bachelors' level and q is the probability of obtaining a masters' degree.

Implications for teacher supply

According to the model, the maximum proportion of bachelor's graduates applying for teaching positions (dark grey area in figure 2.3) is given by:

$$P_t^{m1} = \frac{\nu'_{max} + w^t - w^m}{\nu'_{max} - \nu'_{min}} - \frac{c^m}{\nu'_{max} - \nu'_{min}} \ln\left(\frac{\nu'_{max} + w^t - w^b}{w^m - w^b}\right) \quad (7)$$

However, only master's graduates (with probability q_i) are actually eligible for teach-

ing positions:

$$P_t^{m1'} = \frac{\nu'_{max} + w^t - w^m}{2(\nu'_{max} - \nu'_{min})} - \frac{(c^m)^2}{2(\nu'_{max} - \nu'_{min})} \frac{1}{w^m - w^b} + \frac{(c^m)^2}{2(\nu'_{max} - \nu'_{min})} \frac{1}{\nu'_{max} + w^t - w^b} \quad (7')$$

The elasticity¹⁸ of the supply of teachers with salary is therefore:

$$\epsilon_t^{m1} = \frac{1}{\nu'_{max} - \nu'_{min}} \left[1 - \frac{c^m}{\nu'_{max} + w^t - w^b} \right] \frac{w^t}{P_t^{m1}}$$

First, expressions (7) and (7') again confirm that the supply of teachers increases with teachers' salary when a master's degree is required to teach, all other things being equal, and decreases with the average salaries of bachelor's and master's graduates. Second, the comparison of expressions (6) and (7) (or (7')) shows that an increase in the qualification level of teachers leads to a decrease in the number of potential candidates. This result appears clearly when we compare the dark grey areas of Figures 2.2 and 2.3. Third, according to expression (3) in section 2.4, a decrease in the share of potential candidates results in a decrease in the share of candidates (teacher supply).

Implications for teacher skills

Assuming that q_i measures individuals' academic knowledge, the model implies that the average knowledge of potential candidates increases when a master's degree is required to teach: $E(q|potential\ candidate)|_m > E(q|potential\ candidate)|_b$.

Assuming that $\nu_i \perp q_i$ and $\mu_i = constant$, the average preference (or motivation) of potential candidates for teaching also increases with the level of qualification required: $E(\nu|potential\ candidate)|_m > E(\nu|potential\ candidate)|_b$.

¹⁸To obtain a simplified expression of the elasticity of teacher supply with salary when a master's degree is required, consider all students enrolled in a master's program. Master's students applying for teaching positions satisfy: $w^t + \nu^t - c^m > w^m - c^m$. Therefore, the proportion of master's students who wish to become teachers is given by:

$$P_t^{m2}(t) = \frac{\nu'_{max} + w^t - w^m}{\nu'_{max} - \nu'_{min}} \quad (8)$$

The result is a much simpler expression of the elasticity of teacher supply with salary:

$$\epsilon_t^{m2} = \frac{w^t}{P_t^{m2}} \frac{dP_t^{m2}}{dw^t} = \frac{1}{\nu'_{max} - \nu'_{min}} \frac{w^t}{P_t^{m2}}$$

Note that P_t^{m1} represents the share of potential candidates for teaching positions among bachelor's graduates while P_t^{m2} represents the share of potential candidates for teaching positions among students enrolled in master's programs.

Therefore, assuming that q_i and p_i are positively correlated, expression (3) of Section 2.4 suggests that an increase in degree requirements has an ambiguous effect on candidates' competencies. On the one hand, a better selection of potential candidates (through an increase of m_a and a decrease of σ_a) should improve the average skill level of candidates. On the other hand, the increase in the number of candidates per potential candidate ($E[\frac{N_C}{N_{PC}}]$) tends to decrease the average skill level of candidates. The effect of a higher degree requirement on candidates' knowledge will be positive or negative depending on whether the first or second mechanism dominates. It should also be noted that these mechanisms depend on strong assumptions that may not be valid in practice.

Similarly, expression (4) suggests that raising the qualification level of teachers has an ambiguous effect on the competence of recruited teachers. On the one hand, the increase in the number of positions per potential candidate ($E[\frac{N_P}{N_{PC}}]$) tends to decrease teachers' average knowledge. On the other hand, increasing the average skill level (m_a) of potential candidates tends to increase the average skill level of teachers. The effect of a higher degree requirement on teachers' knowledge can therefore be positive or negative depending on the dominant effect. The effect on teacher quality will eventually depend on the relationship between the probability of obtaining a master's degree, the probability of being recruited as a teacher, and the actual on-the-job performance of teachers.

2.6 Conclusion

The model presented in this chapter has shown, under reasonable assumptions, a positive effect of wage and teacher demand, and a negative effect of degree requirements, on the supply of teachers, which is consistent with the results of the empirical literature. However, the effect on teacher quality is much more difficult to predict. The nature of the correlation between teachers' performance and the competencies measured by degrees and recruitment processes plays a crucial role. The quality of teachers also depends on how degree requirements and salary affect candidates at the top of the skill distribution. Empirical evaluations are essential to conclude on the effect of these potential levers on teacher quality.

The recruitment of teachers in France has undergone several reforms in recent years, which could make it possible to test certain predictions of the model. First, the number

of teaching posts has changed considerably over the period 2007-2016. The sharp drop in the number of posts in 2007-2011 and the equally significant increase in 2012-2016, following the change of government in 2012, make it possible to examine the correlation between teacher supply and demand. Second, the increase in the degree level required (from bachelor to master level) to teach in primary and secondary schools from 2011 onwards can be used to examine the effect of a higher degree requirement on teacher supply and characteristics. This evaluation will be the subject of the next chapter. Third, it might be interesting to examine whether the 11% salary increase from 2018 for teachers working in the most disadvantaged schools (priority education areas), concentrated in a few academic regions, attracted more teachers in a context of shortage. The introduction of standardized tests at the entry and exit of pupils from primary education from 2017 onwards should also make it possible to establish a link between future reforms and pupils' academic progress.

Bibliography

- Angrist, J. D. and Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5):483–503.
- Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., and Wilson, D. (2009). Evaluating the impact of performance-related pay for teachers in England. *Labour Economics*, 16(3):251–261.
- Bacache, M. (2009). *Les Stratégies absurdes, comment faire pire en croyant faire mieux*. Paris, Seuil.
- Bacache-Beauvallet, M. (2006). How incentives increase inequality. *Labour*, 20(2):383–391.
- Baker, G. (1992). Incentive Contracts and Performance Measurement. *The Journal of Political Economy*, 100(3):598–614.
- Balch, R. and Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review*, 44:114–125.
- Belfield, C. R. and Heywood, J. S. (2008). Performance pay for teachers: Determinants and consequences. *Economics of Education Review*, 27(3):243–252.
- Bénabou, R., Kramarz, F., and Prost, C. (2009). The French zones d’éducation prioritaire: Much ado about nothing? *Economics of Education Review*, 28(3):345–356.
- Berger, M. C. and Toma, E. F. (1994). Variation in State Education Policies and Effects on Student Performance. *Journal of Policy Analysis and Management*, 13(3):477.
- Betts, J. R. (1995). Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth. *The Review of Economics and Statistics*, 77(2):231.
- Bryson, A., Forth, J., and Stokes, L. (2017). How much performance pay is there in the public sector and what are its effects? *Human Resource Management Journal*, 27(4):581–597.
- Card, D. and Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, 100(1):1.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679.
- Dal Bo, E., Finan, F., and Rossi, M. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *Quarterly Journal of Economics*, pages 1169–1218.
- Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service (online appendix). *Quarterly Journal of Economics*, 128(3):1169–1218.
- Daussin-Bénichou, J.-m., Idmachiche, S., Leduc, A., and Pouliquen, E. (2015). Les déterminants de l’attractivité de la fonction publique de l’État. *Document de travail de l’INSEE*.
- Dolton, P. and van der Klaauw, W. (1999). The Turnover of Teachers: A Competing Risks Explanation. *Review of Economics and Statistics*, 81(3):543–550.
- Falch, T. (2010). The Elasticity of Labor Supply at the Establishment Level. *Journal of Labor Economics*, 28(2):237–266.
- Falch, T. (2011). Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment. *American Economic Review*, 101(3):460–465.
- Falch, T., Johansen, K., and Strøm, B. (2009). Teacher shortages and the business cycle.

- Labour Economics*, 16(6):648–658.
- Firestone, W. A. and Pennell, J. R. (1993). Teacher Commitment, Working Conditions, and Differential Incentive Policies. *Review of Educational Research*, 63(4):489–525.
- Frey S., B. and Jegen, R. (2000). Motivation Crowding Theory: a Survey of Empirical Evidence. *Working Paper Series*, (26).
- Fryer, R. G. (2013). Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. *Journal of Labor Economics*, 31(2):373–407.
- Glewwe, P., Ilias, N., Kremer, M., American, S., Journal, E., Economics, A., and July, N. (2010). Teacher Incentives. *American Economic Journal: Applied Economics*, 2(3):205–227.
- Greaves, E. and Sibieta, L. (2013). The effect of teacher pay on pupil attainment: a regression discontinuity design. *Working Paper*, (October).
- Grogger, J. (1996). School expenditures and post-schooling earnings: Evidence from high school and beyond. *Review of Economics & Statistics*, 78(4):628.
- Hanushek, E., Kain, J., and Rivkin, S. (1999). Do higher salaries buy better teachers? *Annual Meeting of the American Economic Association*.
- Hanushek, E. A. (1992). The trade-off between Child Quantity and Quality. *Journal of Political Economy*, 100(1):84—117.
- Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19(2):141.
- Hanushek, E. A. and Rivkin, S. G. S. G. (2004). How to Improve the Supply of High-Quality Teachers. *Brookings Papers on Education Policy*, 2004(1):7–25.
- Holmstrom, B. and Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization*, 7(special):24–52.
- Jacob, B., Rockoff, J., Taylor, E., Lindy, B., and Rosen, R. (2016). Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools. *NBER Working Paper series*.
- Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5-6):761–796.
- Jacob, B. A. and Levitt, S. D. (2003). Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *The Quarterly Journal of Economics*, 118(3):843–877.
- Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.
- Jones, M. D. (2013). Teacher behavior under performance pay incentives. *Economics of Education Review*, 37:148–164.
- Leigh, A. (2012). Teacher pay and teacher aptitude. *Economics of Education Review*, 31(3):41–53.
- Loeb, S. and Page, M. E. (2000). Examining the Link between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation. *Review of Economics and Statistics*, 82(3):393–408.
- Marinescu, I. and Wolthoff, R. (2012). Wages, Applications, and Skills. *Working Paper*.
- Murnane, R. J. and Olsen, R. J. (1989). The Effect of Salaries and Opportunity Costs on Duration in Teaching : Evidence from Michigan. *The Review of Economics and Statistics*, 71(2):347–352.
- Osterloh, M. and Frey, B. S. (2000). Motivation, Knowledge Transfer, and Organizational Forms. *Organization Science*, 11(5):538–550.

- Prost, C. (2013). Teacher Mobility: Can Financial Incentives Help Disadvantaged Schools to Retain Their Teachers? *Annals of Economics and Statistics*, (111/112):171–191.
- Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.
- Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review*, 94(2):247–252.
- Ross, S. A. (1973). The Economic Theory of Agency : The Principal ' s Problem. *American Economic Review*, (63):134– 139.
- Rothstein, J. (2015). Teacher Quality Policy When Supply Matters. *American Economic Review*, 105(1):100–130.
- Rumberger, R. W. (1987). The impact of salary differentials on teacher shortages and turnover: The case of mathematics and science teachers. *Economics of Education Review*, 6(4):389–399.

Appendix

Proof of approximation 1

$$F_x^{c-1 \infty}(1 - E[\frac{N_P}{N_C}]) \approx F_x^{-1 \infty}(1 - E[\frac{N_P}{N_{PC}}]) \quad \text{for } \frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$$

Consider the following decomposition:

$$\underbrace{P(x_i \geq x_P^\infty)}_{(a)} = \underbrace{P(x_i \geq x_P^\infty, a_i \geq B^\infty)}_{(b)} + \underbrace{P(x_i \geq x_P^\infty, a_i < B^\infty)}_{(c)}$$

Let us develop a simple approximation for the terms (a), (b) and (c).

- $P(x_i \geq x_P^\infty, a_i \geq B^\infty)$ (b)

By definition of the quantile x_P^∞ we have: $P(x_i \geq x_P^\infty, a_i \geq B^\infty) = 1 - F_x^{c-1 \infty} = \frac{N_P^\infty}{N_{PC}^\infty}$

- $P(x_i \geq x_P^\infty, a_i < B^\infty)$ (c)

$$\begin{aligned} P(x_i \geq x_P^\infty, a_i < B^\infty) &= P(x_i \geq x_P^\infty) P(a_i < B^\infty | x_i \geq x_P^\infty) \\ &= P(x_i \geq x_P^\infty) \int_{-\infty}^{B^\infty} (1 - F_\eta^\infty(x_P^\infty - a)) f_a(a) da \\ &= P(x_i \geq x_P^\infty) \int_{-\infty}^{B^\infty} (1 - F(\frac{x_P^\infty - a}{x_P^\infty \sigma_\eta})) f(\frac{a - m_a}{x_P^\infty \sigma_a}) \frac{da}{x_P^\infty \sigma_a} \\ &= P(x_i \geq x_P^\infty) \int_{\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta}^{+\infty}} (1 - F(x)) f(\frac{x_P^\infty - m_a - x_P^\infty \sigma_\eta x}{x_P^\infty \sigma_a}) \frac{x_P^\infty \sigma_\eta}{x_P^\infty \sigma_a} dx \end{aligned}$$

For $x \gg 1$, $F(x) = 1 - \frac{f(x)}{x}(1 + o(\frac{1}{x}))$, therefore:

$$\begin{aligned} \text{for } \frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1, \quad P(a_i < B^\infty | x_i \geq x_P^\infty, x_P^\infty, B^\infty) &= \int_{\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta}^{+\infty}} (1 - F(x)) f(\frac{x_P^\infty - m_a - x_P^\infty \sigma_\eta x}{x_P^\infty \sigma_a}) \frac{x_P^\infty \sigma_\eta}{x_P^\infty \sigma_a} dx \\ &\leq \frac{f(\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta})}{\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta}} \int_{-\infty}^{+\infty} f(\frac{x_P^\infty - m_a - x_P^\infty \sigma_\eta x}{x_P^\infty \sigma_a}) \frac{x_P^\infty \sigma_\eta}{x_P^\infty \sigma_a} dx \\ &\leq \frac{f(\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta})}{\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta}} = o(1) \end{aligned}$$

Finally, $P(x_i \geq x_P^\infty, a_i < B^\infty) = P(x_i \geq x_P^\infty) o(1)$

- $P(x_i \geq x_P^\infty)$ (a) :

Therefore, the expression becomes:

$$\begin{aligned} P(x_i \geq x_P^\infty) &= P(x_i \geq x_P^\infty, a_i \geq B^\infty) + P(x_i \geq x_P^\infty, a_i < B^\infty) \\ &= \frac{N_P^\infty}{N_{PC}^\infty} + P(x_i \geq x_P^\infty) o(1) \quad \text{for } \frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1 \end{aligned}$$

i.e. $P(x_i \geq x_P^\infty)(1 + o(1)) = \frac{N_P^\infty}{N_{PC}^\infty}$ for $\frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$

And then:

$$P(x_i \geq x_P^\infty) = \frac{1}{1+o(1)} \frac{N_P^\infty}{N_{PC}^\infty} \quad \text{for } \frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1$$

Using the fact that $P(x_i \geq x_P^\infty) = 1 - F_x(x_P^\infty)$, we get:

$$x_P^\infty = F_x^{-1} \left(1 - \frac{1}{1+o(1)} \frac{N_P^\infty}{N_{PC}^\infty} \right) \approx F_x^{-1} \left(1 - \frac{N_P^\infty}{N_{PC}^\infty} \right) = F_x^{-1} \left(1 - E\left[\frac{N_P}{N_{PC}}\right] \right)$$

Finally, with the definition $x_P^\infty = F_x^{c-1} \left(1 - E\left[\frac{N_P}{N_C}\right] \right)$, we obtain that:

$$x_P^\infty = F_x^{c-1} \left(1 - E\left[\frac{N_P}{N_C}\right] \right) \approx F_x^{-1} \left(1 - E\left[\frac{N_P}{N_{PC}}\right] \right) \quad \text{for } \frac{x_P^\infty - B^\infty}{x_P^\infty \sigma_\eta} \gg 1.$$

Proof of approximation 2

$$E[x_i | a_i \geq B^\infty, x_i \geq x_P^\infty] \approx m_a + \frac{\sqrt{\sigma_a^2 + \sigma_\eta^2}}{E\left[\frac{N_P}{N_{PC}}\right]} f\left(F^{-1}\left(1 - E\left[\frac{N_P}{N_{PC}}\right]\right)\right)$$

Consider the following decomposition:

$$E[x_i \mathbb{1}_{\{x_i \geq s\}} | s] = E[x_i \mathbb{1}_{\{x_i \geq s, a_i < B^\infty\}} | s] + E[x_i \mathbb{1}_{\{x_i \geq s, a_i \geq B^\infty\}} | s]$$

- $E[x_i \mathbb{1}_{\{x_i \geq s, a_i < B^\infty\}} | s]$:

$$\begin{aligned} E[x_i \mathbb{1}_{\{x_i \geq s, a_i < B^\infty\}} | s] &= P(x_i \geq s, a_i < B^\infty | s) E[x_i | x_i \geq s, a_i < B^\infty, s] \\ &= P(x_i \geq s | s) o(1) E[x_i | x_i \geq s, a_i < B^\infty, s] \quad \text{for } \frac{s-B}{\sigma_\eta} \gg 1 \end{aligned}$$

- $E[x_i \mathbb{1}_{\{x_i \geq s, a_i \geq B^\infty\}} | s]$:

$$\begin{aligned} E[x_i \mathbb{1}_{\{x_i \geq s, a_i \geq B^\infty\}} | s] &= P(x_i \geq s, a_i \geq B^\infty | s) E[x_i | x_i \geq s, a_i \geq B^\infty, s] \\ &= \frac{N_P}{N_{PC}} E[x_i | x_i \geq s, a_i \geq B^\infty, s] \end{aligned}$$

- $E[x_i \mathbb{1}_{\{x_i \geq s\}} | s]$:

$$E[x_i \mathbb{1}_{\{x_i \geq s\}} | s] = P(x_i \geq s | s) E[x_i | x_i \geq s, s]$$

Therefore,

$$\begin{aligned} P(x_i \geq s | s) E[x_i | x_i \geq s, s] &= P(x_i \geq s | s) o(1) E[x_i | x_i \geq s, a_i < B^\infty, s] \\ &+ \frac{N_P}{N_{PC}} E[x_i | x_i \geq s, a_i \geq B^\infty, s] \quad \text{for } \frac{s-B}{\sigma_\eta} \gg 1 \end{aligned}$$

Then,

$$\begin{aligned} &P(x_i \geq s | s) \{E[x_i | x_i \geq s, s] - o(1) E[x_i | x_i \geq s, a_i < B^\infty, s]\} \\ &= \frac{N_P}{N_{PC}} E[x_i | x_i \geq s, a_i \geq B^\infty, s] \quad \text{for } \frac{s-B^\infty}{\sigma_\eta} \gg 1 \end{aligned}$$

We have shown previously that: $P(x_i \geq s|s) = \frac{1}{1-o(1)} \frac{N_P}{N_{PC}}$ for $\frac{s-B^\infty}{\sigma_\eta} \gg 1$

So we derive:

$$\{E[x_i|x_i \geq s, s] - o(1) E[x_i|x_i \geq s, a_i < B^\infty, s]\} \frac{1}{1-o(1)} = E[x_i|x_i \geq s, a_i \geq B^\infty, s] \quad \text{for } \frac{s-B^\infty}{\sigma_\eta} \gg 1$$

which can be written: $E[x_i|x_i \geq s, s] \approx E[x_i|x_i \geq s, a_i \geq B^\infty, s]$ for $\frac{s-B^\infty}{\sigma_\eta} \gg 1$

$s \rightarrow x_P^\infty$ allows to conclude that:

$$E[x_i|x_i \geq x_P^\infty] \approx E[x_i|x_i \geq x_P^\infty, a_i \geq B^\infty] \quad \text{for } \frac{x_P^\infty - B^\infty}{\sigma_\eta} \gg 1$$

Expected exam score and expected ability of candidates

When a is normally distributed, the expected exam score of candidates is:

$$E[a_i|a_i > B^\infty] = m_a + \sigma_a \frac{f(\frac{B^\infty - m_a}{\sigma_a})}{1 - F(\frac{B^\infty - m_a}{\sigma_a})}, \quad \text{with } f \text{ and } F \text{ respectively the density and the cumulative function of the normal distribution } N(0, 1),$$

with $B^\infty = x_P^\infty - F_\eta^{-1}(1 - \frac{c}{u-v}) \approx F_x^{-1}(1 - E[\frac{N_P}{N_{PC}}]) - F_\eta^{-1}(1 - \frac{c}{u-v})$ for $\frac{x_P^\infty - B^\infty}{\sigma_\eta} \gg 1$.

The expected score of candidates is $E[x_i|a_i > B^\infty] = E[a_i|a_i > B^\infty]$.

Using the relation $B^\infty = F_a^{-1}(1 - E[\frac{N_C}{N_{PC}}]) = m_a + \sigma_a F^{-1}(1 - E[\frac{N_C}{N_{PC}}])$, I derive a simple expression for $E[a_i|a_i > B^\infty]$:

$$\begin{aligned} E[a_i|a_i > B^\infty] &= m_a + \sigma_a \frac{f(\frac{B^\infty - m_a}{\sigma_a})}{1 - F(\frac{B^\infty - m_a}{\sigma_a})} \\ &= m_a + \sigma_a \frac{f(F^{-1}(1 - E[\frac{N_C}{N_{PC}}]))}{E[\frac{N_C}{N_{PC}}]} \\ E[a_i|a_i > B^\infty] &= m_a + \frac{\sigma_a}{E[\frac{N_C}{N_{PC}}]} f(F^{-1}(1 - E[\frac{N_C}{N_{PC}}])) \end{aligned} \quad (5)$$

Expected recruitment threshold

Under the previous assumptions, the expected recruitment threshold satisfies: $x_P^\infty = E[s] =$

$$F_x^{c-1} \infty(1 - E[\frac{N_P}{N_C}]) \approx F_x^{-1} \infty(1 - E[\frac{N_P}{N_{PC}}]) = m_a + \sqrt{\sigma_a^2 + \sigma_\eta^2} F^{-1}(1 - E[\frac{N_P}{N_{PC}}])$$

This provides the expression: $E[s] \approx m_a + \sqrt{\sigma_a^2 + \sigma_\eta^2} F^{-1}(1 - E[\frac{N_P}{N_{PC}}])$ (6)

Expected exam score of teachers

The expected score of recruited teachers is:

$$E[E[x_i|a_i \geq B^\infty, x_i \geq s, s]] = E[x_i|a_i \geq B^\infty, x_i \geq x_P^\infty] = \int_{x_P^\infty}^{+\infty} x f_x^{c-1}(x) dx$$

The expected exam score of recruited teachers is:

$$E[a_i|x_i \geq x_P^\infty] = E[x_i|x_i \geq x_P^\infty] - E[\eta_i|x_i \geq x_P^\infty]$$

Under the assumption $\frac{x_P^\infty - B^\infty}{\sigma_\eta} \gg 1$, we show (see proof below) that:

$$E[x_i|a_i \geq B^\infty, x_i \geq x_P^\infty] \approx E[x_i|x_i \geq x_P^\infty]$$

Using the relation $x_P^\infty \approx m_x + \sigma_x F^{-1}(1 - E[\frac{N_P}{N_{PC}}])$, we have:

$$\begin{aligned} E[x_i|x_i \geq x_P^\infty] &= m_x + \sigma_x \frac{f(\frac{x_P^\infty - m_x}{\sigma_x})}{1 - F(\frac{x_P^\infty - m_x}{\sigma_x})} \\ &\approx m_x + \sigma_x \frac{f(F^{-1}(1 - E[\frac{N_P}{N_{PC}}]))}{E[\frac{N_P}{N_{PC}}]} \\ &\approx m_a + \frac{\sqrt{\sigma_a^2 + \sigma_\eta^2}}{E[\frac{N_P}{N_{PC}}]} f(F^{-1}(1 - E[\frac{N_P}{N_{PC}}])) \quad \text{for } \frac{x_P^\infty - B^\infty}{\sigma_\eta} \gg 1 \end{aligned}$$

Therefore, the expected score of recruited teachers for $\frac{x_P^\infty - B^\infty}{\sigma_\eta} \gg 1$ is approximately:

$$E[x_i|a_i \geq B^\infty, x_i \geq x_P^\infty] \approx m_a + \frac{\sqrt{\sigma_a^2 + \sigma_\eta^2}}{E[\frac{N_P}{N_{PC}}]} f(F^{-1}(1 - E[\frac{N_P}{N_{PC}}])) \quad (7)$$

Chapter 3

How does the increase in teachers' qualification levels affect their supply and characteristics?

3.1 Introduction

Over the past decade, many countries have raised the level of teacher qualifications. In 2013, sixteen European countries require a master's degree to teach in secondary education, while eleven countries require a similar level in primary education. These reforms are part of an international movement to "professionalize" teachers ([Hildebrandt and Eom, 2011](#); [Lapostolle, 2013](#)) in order to improve the performance of the education system. Although, a recent report advocates improving teachers' qualifications to ensure the "attractiveness and excellence" of the profession ([European Commission, 2014](#)), these reforms have not been evaluated and their effect on recruitment efficiency has not yet been demonstrated.

In the public like in the private sector, many recruiters screen candidates based on their degree level. As skills are not fully revealed during job interviews, qualification is an attractive criterion for reducing uncertainty about candidates' competencies. Nevertheless, the literature in education is rather skeptical about the effectiveness of raising the degree level of teachers. First, most empirical studies report a positive but weak correlation between teacher quality and degree level¹ ([Hanushek and Rivkin 2004](#), al-

¹Chapter 2 also discusses the fact that the effect of the diploma depends on the correlation between

though: [Jacob et al. 2016](#)). The evidence remains weak, however, as these studies are primarily based on cross-sectional analyses and may be subject to interpretation bias. Results are often obtained by comparing the performance of undergraduate and graduate teachers, without taking into account unobserved characteristics (e.g., motivation) between these two populations. Second, the number of candidates for teaching positions is expected to decrease as the required level of qualification increases (see Chapter 2 of this thesis for a discussion of the mechanisms involved). On the one hand, individuals with a low probability of obtaining a master's degree will no longer apply to teaching posts. On the other hand, the most competent applicants may also give up teaching because higher qualifications are associated with better alternative job opportunities ([Berger and Toma, 1994](#); [Angrist and Guryan, 2008](#)). In addition, evidence from the occupational licensing literature suggests that the difficulty of recruiting qualified teachers can be particularly acute in socio-economically disadvantaged areas when constraints on qualifications increase ([Angrist and Guryan, 2008](#)). Third, raising the degree level of teachers may have a negative impact on the diversity of the profession. In particular, the increase in opportunity costs associated with higher requirements can affect the professional choices of students from the most modest socio-economic backgrounds. Men are also less likely to apply for teaching positions if their career prospects with a higher degree level are better than those of women, as recent data suggest². The empirical literature has shown that diversity of educational personnel is important to ensure equal treatment of students. Several studies have highlighted the existence of interactions between the characteristics of students and those of teachers (effects of role models, stereotype threat, discrimination or favouritism) that are likely to contribute to differences in academic performance, orientation and the pursuit of higher education according to gender, social or cultural origin ([Meier, 1984, 1993](#); [Dee, 2004, 2005](#); [Carrington et al., 2008](#); [Grissom et al., 2009](#); [Feld et al., 2016](#); [Paredes, 2014](#); [Egalite et al., 2015](#); [Porter et al., 2017](#)). In France, the diversity of teachers also raises the question of access to public employment, which is still difficult for certain categories of the population ([Versini, 2004](#); [Fougère and Pouget, 2004](#); [Calvès, 2005](#); [Meurs et al., 2006](#); [L'Horty, 2016](#)).

To examine whether a higher level of qualification can contribute to improving the academic skills and professional competences of teachers.

²National partnership for women & families, « America's women and the wage gap », fact sheet 2018. See also evidence for France in Section 3.4.4 of this study.

iciency of teacher recruitment, I consider the reform that modified the diploma required to teach in primary and secondary education in France in 2011. Like most countries, France had recruited primary school teachers and most secondary school teachers (in the public and private sectors) at the bachelor level since 1989. Public school teachers were recruited through competitive examinations, and then assigned to a training institute called the *Institut Universitaire de Formation des Maitres (IUFM)* for a period of one year. During that year, future teachers received both theoretical and practical training. They taught about one-third of the year and were paid full-time. Starting in 2011, all teachers must have a master's degree (or equivalent) in any specialty to be eligible for recruitment. In addition, initial teacher training is integrated into a master's degree in education called *Master métiers de l'enseignement, de l'éducation et de la formation (MEEF)* that replaces the training institutes IUFM. The objective of the reform is twofold. First, it aims to improve the quality of recruitment by extending the duration of teachers' studies. Second, it aims to reduce public expenditure by no longer funding initial teacher training³. In addition, this reform did not lead to a significant increase in teachers' salaries⁴.

The objective of this study is to estimate the effect of the 2011 qualification reform on teacher supply, skills, and gender and social diversity. In France, the majority of public sector teachers are recruited through competitive examinations (99% in primary education and 92% in secondary education). I focus on primary school teachers because their recruitment is decentralized (30 regions set the number of posts offered each year and organize recruitment through competitive examinations) and very local (more than 80% of candidates apply in the region where they live), which offers the possibility of exploiting regional variations to precisely control the effect of unemployment rates, wages and teacher demand during the period surrounding the reform (see chapter 2 for a discussion of these mechanisms). This aspect is important insofar as the 2011 reform is part of a context of falling public spending which is reflected in particular by the stagnation of teachers' salaries over the 2010-2016 period (*gel du point d'indice*) and by the sharp

³As a result, teachers who have obtained a master's degree in a specialty other than education (MEEF master's degree) will no longer receive initial training from 2011. The reform will be made more flexible from 2014 onwards (see section 3.2) and initial training will again be paid full-time and offered to all teachers recruited.

⁴The reform was followed by a slight salary increase at the beginning of the career: 5% increase in the first year, then 3% over the next two years, 1% after 4 and 5 years of seniority, then no salary increase beyond 5 years of seniority.

decline in the number of teaching posts over the 2005-2011 period (see Figure 1 in the introduction to this thesis).

The data used in this study come mainly from the Ministry of Education's administrative files. Recruitment databases contain individual information on all candidates who have participated in the recruitment competition since 2003. For each candidate, they provide sociodemographic information (sex, age, nationality, level of diploma, discipline, family situation, place of residence), the results obtained in the various tests constituting the recruitment competition and the recruitment decision. However, the administrative data do not indicate the social origin of the applicants. I use the median income in the candidates' city of residence to get a rough measure of their socio-economic background. Personnel management databases provide information on teachers throughout their careers since 2005. Data include teacher assignments in schools, number of hours worked, location of schools and results of teacher evaluations by education inspectors. I have detailed information on over 500,000 applicants and over 130,000 primary school teachers for the period 2003-2017. I also have the number of positions (teacher demand) and candidates (teacher supply) by region since 1996.

In France, the quality of primary school teachers cannot be estimated using value-added models (Rivkin et al., 2005; Rothstein, 2010; Chetty et al., 2014) because there are no standardized tests for primary school pupils. Instead, I consider two alternative measures of teaching quality: the accreditation score obtained in the recruitment examination and the classroom observation score assigned by education inspectors on average 2-3 years after tenure. The accreditation score measures the candidates' disciplinary knowledge and pedagogical skills in six subjects taught in primary school: mathematics, French, history, geography, science and technology. The classroom observation score measures teachers' pedagogical skills and attitudes, but it also takes seniority into account because it is used by the administration for teacher promotion. Most educational research finds a positive relationship between classroom observation score, accreditation score and student learning.

The results show that the increase in the level of qualifications required to teach has led to a sharp drop in the number of candidates for teaching posts. In public primary education, the supply of teachers decreased by 50% on average, but the decline was much more pronounced in regions with the highest demand for teachers. Similarly, the teacher

accreditation score decreased by an average of 0.55SD point after the introduction of the reform, and this decline was more pronounced in regions with high teacher demand. The share of men (minority gender) among candidates and teachers has also tended to decrease from 2011 onwards. However, the classroom observation score (two to three years after tenure) remained relatively stable over the study period. Similarly, the reform does not appear to have affected the social composition of teachers (measured by median income in the city of residence). The effect of the reform of teachers' qualifications (decline in the number of candidates without improvement in their academic knowledge) combined with the increase in teacher posts from 2012 may have contributed to shortages in some regions. Recruiters, who consider that the level of candidates is too low, prefer to leave a large number of vacant posts from 2013 in the two largest French academic regions. This situation has generated difficulties in replacing absent teachers and has led to the introduction of additional competitions in these regions from 2015.

The validity of these first-difference estimates is based on the assumption that teacher qualification reform is exogenous. However, this reform takes place in a context of low attractiveness of the civil service, in particular because of the stagnation of salaries between 2010 and 2016. These factors may have had a deterrent effect on some applicants, thus changing the number and composition of teachers during the period studied. The sharp drop in the number of candidates and recruitment scores in 2011, after several years of relative stability, allows us to be relatively confident about our identification strategy. However, the decline in the share of men among candidates is more in line with a downward trend that could result from the external context. In order to verify this hypothesis, we examine the evolution of the number of candidates and the share of men among the applicants for the recruitment competition of the Regional Institutes of Administration (*Instituts Régionaux d'Administration, IRA*) during the period 2008-2017. This recruitment competition is decentralized in 5 IRAs (located in Bastia, Lille, Lyon, Metz and Nantes) and shares many points in common with that of teachers. IRAs' recruits have responsibilities in French administrations and earn salaries close to those of teachers. The IRA competition is open to undergraduate students throughout the period before and after teacher qualification reform, making it an interesting counterfactual. The results show a downward trend in the number of candidates for IRAs posts, of the order of 5% per year over the period 2009-2015. However, this decrease remains modest compared

to the variations observed for primary school teachers. In particular, we do not see any significant changes in 2011. The results also show a downward trend in the share of men among IRAs applicants over the period 2009-2013. The order of magnitude is close to that observed for teachers, which invites us to moderate the effect of teacher qualification reform on the diversity of this profession.

This Chapter is organized as follows. Section II briefly reviews the different ways of measuring teacher quality. Section III presents the context of this study and section IV presents the data used. Section V tests some of the predictions in Chapter 2 and provides an initial exploratory analysis. Section VI presents the model and Section VII describes the results. Section VIII tests the robustness of the estimates and Section IX examines the link between qualification reform and teacher shortages in some regions. Section X concludes.

3.2 How to measure teacher quality?

Most research finds a positive relationship between teachers' pedagogical skills, knowledge of subject matter and students' performance on standardized tests. The literature generally distinguishes between two types of knowledge: content knowledge (CK) and pedagogical content knowledge (PCK) (Ball et al., 2005; Hill et al., 2005, 2004; Baumert et al., 2010). Content knowledge (i. e., subject matter) is by far the most studied and refers to the level of proficiency required to teach a subject. A substantial number of studies use GPA, SAT, classroom or certification test scores to measure candidates' knowledge of a discipline (Hill et al., 2005; Goldhaber and Anthony, 2007; Baumert et al., 2010; Rockoff et al., 2011; Campbell et al., 2014). The authors generally find a moderate but positive association between teacher content knowledge and student learning, in particular when they are measured by certification scores (Ferguson, 1991; Ehrenberg and Brewer, 1994). The results are also robust in the use of tests specifically designed to assess teachers' basic or advanced knowledge of the discipline (Mullens et al., 1996; Rowan et al., 1997; Hill et al., 2004). Pedagogical content knowledge has recently received increasing attention. The concept was developed after quantitative and qualitative research established that content knowledge alone is not sufficient to guarantee student learning. The PCK measure refers to disciplinary knowledge adapted to the context of teaching. It in-

cludes the teacher's ability to understand the student learning process, identify students' difficulties and propose appropriate teaching methods. The PCK measure also takes into account knowledge of the subject matter, in line with the idea that a teacher who is more competent in the subject being taught is also better able to develop pedagogical strategies, representations and explanations that facilitate student learning. Previous research generally find that pedagogical content knowledge predicts teacher quality (ability to advance student achievement on standardized tests) better than content knowledge (Hill et al., 2005; Jacob et al., 2016). The written tests in the French competition comprise both the CK (first part of the tests) and the PCK (second part) measures.

Increasingly, researchers consider that assessments based on classroom observations capture dimensions of teacher quality that are not always reflected in other performance measures (Gallagher, 2004; Kimball et al., 2004; Milanowski, 2004; Kersting et al., 2013; Cantrell and Kane, 2013; Jacob et al., 2016). Teacher assessment scores from classroom observations and student learning gains (on standardized tests) are positively correlated around 0.2-0.4, which is well above the correlations obtained using CK, PCK and most other quality indicators available. In addition, Rockoff and Speroni (2011) show that the classroom observation score of a new teacher predicts future student learning gains as accurately as a value-added model that estimates teacher quality based on a single year of observation (progression of students over the teacher's first year of experience). The authors conclude that these measures are complementary rather than substitutive. Araujo et al. (2016) recently confirmed that an increase of one standard deviation of the teacher classroom assessment score, as measured by the Classroom Assessment Scoring System, increases students' scores in mathematics and language in kindergartens by 0.07 to 0.11 standard deviation points. Finally, Benhenda (2017) shows that classroom observation scores obtained by secondary school teachers in France are (slightly) positively correlated with students' academic achievement (a one standard deviation increase in the classroom assessment score is associated with a 0.02 standard deviation increase in students' academic performance).

3.3 Institutional background

3.3.1 The recruitment of public primary school teachers

In France, 89% of primary school teachers work in the public sector. Most of them are recruited through competitive examinations (98%), which gives them the status of officials after a one-year trial period. Primary school teachers who are not civil servants (2% of the workforce) work on renewable one-year contracts⁵ and are not tenured. They are usually recruited directly by the school principal after having been pre-selected by the regional education authority (*rectorat*).

There are three main competitions, called *CRPE*⁶, to become full teachers in the public sector. The *external* competition is open to candidates who are not yet civil servants. It corresponds to approximately 89% of the teaching posts and 87% of the candidates. The *internal* competition is open to candidates who have taught for at least five years on fixed-term contracts and who are not yet civil servants (approximately 7% of teaching posts and 5% of candidates). A third competition is open exclusively to civil servants who are not yet teachers (approximately 4% of teaching posts and 8% of candidates). In this study, I am interested in teachers recruited by *external* competition because it is the main access route to teaching in France.

The CRPE is a two-stage competition that takes place once a year. The examination is the same for all candidates, but recruitment is decentralized in thirty academic regions (called *academies*), which means that the number of teaching positions is specific to each region. The number of positions is publicly advertised during the registration period, approximately six months before the recruitment process begins. Candidates can only take the exam in one region.

The competition assesses the candidates' knowledge and pedagogical skills in the main subjects taught in primary education. All candidates take written tests in mathematics, French, history, geography, science and technology⁷. The best applicants are eligible for the second phase of the competition, which consists of three oral examinations and a professional interview. The candidates with the highest marks in the written and oral

⁵The contracts can be renewed an unlimited number of times

⁶Concours de Recrutement des Professeurs des Ecoles

⁷Since 2014, candidates take only two written tests, one in mathematics and one in French.

tests are assigned to a one-year trial period⁸.

During the probationary period, trainee teachers are supported by a more experienced teacher. At the end of the school year, a Commission composed of three to six teaching professionals (mainly pedagogical inspectors) decides to grant tenure, extend the probationary period or dismiss trainees in case of serious misconduct. The Commission's decision is based on the classroom observations provided by education inspectors. Most teacher trainees are tenured after one year (99%).

After the trial period, the tenured teacher may request a transfer to another region. However, priority in mobility is given to experienced teachers. As a result, young recruits usually work for at least five years in the same region before moving to another region (see Chapter 4 for details).

3.3.2 Reform of teachers' diploma level

Until 2009, all bachelor graduates (three years of university holders) can participate in the primary school teacher recruitment competition. Registration for the competition takes place in September-October and the tests run from April to June. From October to April, all candidates can take part in a preparation program at a specialized institute called *IUFM*⁹. Teachers recruited in July are assigned to schools in September and are paid full time from that date. During the probationary year, the working time of the trainee teachers is divided between teaching (33% of the time) and theoretical training in pedagogy and psychology at an IUFM institute (67% of the time).

In July 2009, the government increased by decree the level of diploma required to become a teacher (from year 2010 onwards) and modified the initial training schedule (from 2011). The reform has two phases. In 2010, first-time entrants to the competition must be enrolled in a fourth year of university (first year of a master's program or equivalent) or in the preparation program offered by IUFM institutes. Candidates who take the exam for the second time or more are not affected by the first stage of the reform and can still write the exam with a bachelor's degree. Teachers recruited in 2010 are paid full time during the probationary year and continue to spend one third of their time teaching and the rest of their time training at an IUFM institute.

⁸The number of teachers recruited is limited by the number of posts available in each region.

⁹Institut Universitaire de Formation des Maitres

From 2011 onwards, all applicants must have completed four years of university and be enrolled in the second year of a master's program or equivalent to be eligible for the teacher recruitment competition. Candidates may enroll in any master's program, including the Master in Education offered by ESPE schools¹⁰ that replace the IUFM institutes from 2012 onwards. The first year of this master's program prepares students for the recruitment examination (disciplinary and didactic knowledge) while the second year includes part-time courses and part-time teaching in a school (practical training). The ESPE master's program is very close to the initial training provided by the IUFM before 2012. However, second year master's students are no longer paid and must obtain a degree to be eligible for the competition. In addition, students enrolled in a master's program in a discipline other than education do not receive specific preparation for teaching. After passing the competitive recruitment examination, teachers are assigned full-time to a school for a probationary year.

In 2014, the new government maintains teacher education at the master's level (five years of university) but candidates have the possibility to take the competition during the first year of a master's program (fourth year of university) instead of the second year. Teachers with four years of university are paid full-time while they complete a master's degree (half of the time) and teach in a school (half of the time) during the probationary year. This situation is close to the pre-reform period, except that trainee teachers must now validate a master's degree. New teachers who already have a master's degree can also teach part-time in a school and take education training the rest of the time. For all teachers, the tenure is subject to (1) the validation of a master's degree or equivalent (theoretical part) and (2) obtaining a favourable opinion from the Pedagogical Commission (practical part). If one of these two conditions is not met, the trainee teachers may be granted an additional year to satisfy both conditions (diploma and tenure), after which they lose the benefit of the competition.

Figure 3.1 summarizes the main stages of the recruitment process over the 2000-2017 period. Primary school teachers must have a bachelor's degree until 2009 and a master's degree from 2011. Candidates may take part in the recruitment competition in the year following the completion of a bachelor's degree until 2009, during their final year of a master's program between 2011 and 2013 and during their first year of a master's

¹⁰Ecoles Supérieures du Professorat et de l'Éducation

program from 2014 onwards. During the period 2011-2013, initial teacher training takes place before recruitment and is unpaid.

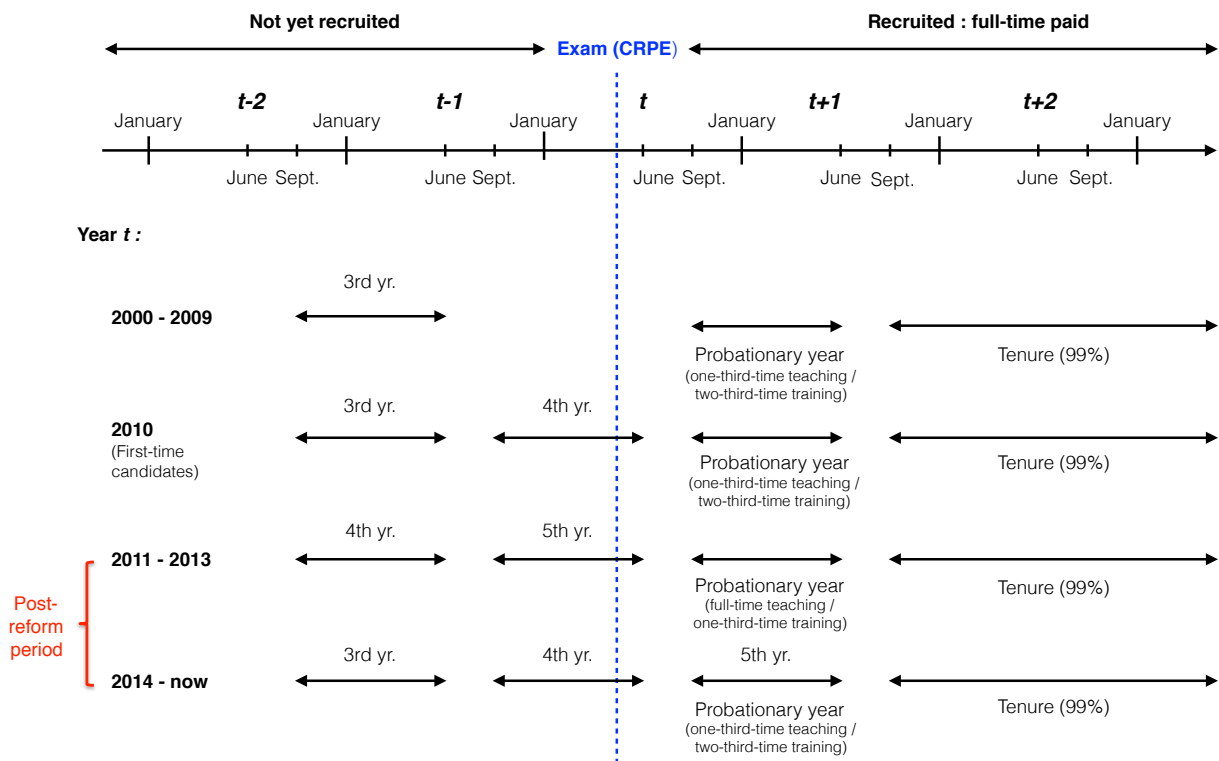


Figure 3.1 – Recruitment and tenure of public primary school teachers for the period 1989-2017

Note: Until 2009, candidates must hold a bachelor's degree (three year of university studies) to be eligible to write the primary school teacher recruitment exam. In 2010, candidates participating in the competition for the first time must be enrolled in the first year of a master's degree or be enrolled in a training institute (IUFM). Between 2011 and 2013, all applicants must be enrolled in the fifth year of university studies to be eligible for the competition. They must obtain a master's degree at the end of the school year to be recruited. From 2014, candidates must be enrolled in the fourth year of university studies. They must obtain a master's degree the following year to be recruited. After a one-year trial period (which is divided into training and teaching in varying proportions depending on the year) the vast majority of teachers are tenured 99%.

3.4 Data and descriptive statistics

The data used in this study come from different sources. Since 1996, the Ministry of Education has published annually the number of teaching posts, the number of candidates

for the written examinations, the number of candidates eligible for the oral examinations and the number of teachers recruited by each academic region for the CRPE competitions. The Ministry also provides databases containing detailed information on all candidates registered for recruitment competitions since 2003. These administrative files provide socio-demographic information, in particular the sex, age, level of education, specialty of the diploma, place of residence, family and professional situation of more than 200,000 candidates to the *external* CRPE competition. They also provide information on all marks obtained by candidates in the written and oral tests of the competition.

The Ministry of Education also provides access to personnel management databases that provide information on teachers throughout their careers. Since 2005, the data specify the school in which the teacher works, the postal code and status of the school (primary school, nursery school, special school), the teacher's contract (civil servant, fixed-term contract), the number of hours worked and the teacher's place of residence. The management records also provide the classroom observation score that corresponds to teachers' evaluation by the pedagogical inspectors every five years on average.

The Ministry of Higher Education provides aggregate data on the number of students and graduates of higher education at national level since 1985 (annually since 2005). The Labor Force Survey allows to estimate annually the average wages of teachers and post-secondary graduates since 2003. The National Institute of Statistics and Economic Studies (INSEE) has also provided the unemployment rate at the national level since 1985. These variables will be used to control for changes in teacher supply and characteristics that are not directly related to the increase in the level of qualification required to teach from 2011 onwards. The following sections present the main variables that are considered in this study and provide descriptive statistics.

3.4.1 Teacher qualifications

Table 3.1 indicates the degree level and principal occupation of candidates at the time of enrollment in the primary teacher recruitment competition (approximately 6 months before the start of the examination) for the periods 2003-2009 (column 1), 2011-2013 (column 2) and 2014-2015 (column 3). Between 2003 and 2009, 72% of applicants hold a bachelor's degree or equivalent, 17% completed four years of university studies and 6% have five or more years of university studies. Table 3.1 also reveals that preparation for

the teacher recruitment examination is not always compatible with university attendance. Until 2009, only 21% of candidates are enrolled in university (*students*) during the school year they take the exam. A significant proportion of candidates (32%) are enrolled in a specialized institute called the *Institut Universitaire de Formation des Maîtres* (IUFM) which offers a one-year preparation for the competition but does not deliver a university degree. In other words, the training year spent in an IUFM institute is not recognized and does not give rise to equivalence within the university system. In addition, 34% of applicants are employed at the time of registration, of which 0.3% have a one-year (renewable) contract as primary school teachers, and 13% are unemployed. Because almost one out of every two candidates takes part in the competition at least twice, 47% of applicants prepared the competition for at least one year in an IUFM training institute.

From 2011, a master's degree or equivalent becomes compulsory for primary and secondary education. Between 2011 and 2013, all applicants must have completed at least 4 years of university studies and be enrolled in the 5th academic year to participate in the teacher recruitment competition. In addition, the IUFM training institutes are replaced in 2012 by schools called the *Ecoles Supérieures du Professorat et de l'Éducation* (ESPE) which are empowered to deliver a master's degree in education. Candidates for the teacher recruitment competition are free to take a master's degree in education at an ESPE school or in another discipline at the university. Candidates who have passed the competition but have not validated their second year of master's degree will retain the benefit of the competition for one year. They must obtain a master's degree the following year to be recruited. Column 2 of Table 3.1 indicates that 42% of applicants completed four years of university studies and 50% of applicants obtained a master's degree or more between 2011 and 2013. In addition, the proportion of students among applicants has increased significantly: almost 59% in 2011-2013 compared to 21% in 2003-2009. In contrast, the share of employed persons decreased (27% in 2011-2013 compared to 34% in 2003-2009), while the share of unemployed persons remained stable (14%).

From 2014 onwards, master's degrees are still compulsory for teaching, but candidates are now allowed to take the recruitment exam in the fourth academic year. At the end of the competition, candidates must have completed 4 years of university studies and enroll in the 5th year of university. The following year, they will have to obtain a master's degree (in an ESPE school or university) alternating with part-time teaching. Tenure

at the end of this first year of teaching is conditional on obtaining a master's degree. Column 3 of Table 3.1 indicates that 54% of applicants have completed four years of university and 38% of applicants have a master's degree or more. In addition, 53% of candidates are students in the year they write the exam, of whom 26% are enrolled in a Master in Education program at an ESPE school and 27% are enrolled or have completed a Master's degree in another discipline ¹¹. 34% of applicants are employed, of whom 1.2% hold a temporary teaching position in primary school, and 14% are unemployed.

Table 3.1 – Highest level of diploma and principal occupation of candidates for primary school teaching posts over the period 2003-2015

	2003-2009	2011-2013	2014-2015
Highest degree level (%)			
Unknown	3.5	7.2	7.8
3 years of university	71.9	0.2	0.9
4 years of university	17.2	42	53.7
5 years of university or more	6.2	50.3	37.5
Main occupation (%)			
First year of IUFM	32.4	15.9	0
Student, of which :	20.7	42.7	52.9
Student ESPE	-	-	26.1
Other student	20.7	-	26.7
Teacher on fixed term contract	0.3	0.3	1.2
Other employed	33.7	27.1	33
Unemployed	12.9	13.9	13.7

Notes: Highest level of diploma of the candidates at the time of registration for the *external* primary school teacher recruitment examination. Diploma requirement does not apply for high performance athletes and parents of three or more children. Candidates who meet these criteria are not required to indicate their degree level ("unknown" in the Ministry of Education's management databases). High-performance athletes and parents of three or more children are more likely to register in this category when a master's degree becomes mandatory in 2011. Candidates also specify their main occupation at the time of registration for the competition. 52.9% of candidates are students over the period 2014-2015, against 20.7% over the period 2003-2009.

Source: Ministry of Education (DEPP). *Scope:* Metropolitan France. Authors' calculations.

The 2011 reform, which consisted in upgrading the degree level of teachers from undergraduate to master's level, concerns all teachers in primary and secondary education ¹². Since the number of candidates for the primary and secondary school competitions is relatively high (see Section 3.4.2), the reform may be responsible for a significant increase in the rate of continuation of studies after graduation from a bachelor's degree and in

¹¹The nomenclature of the Ministry of Education's competition database changes from 2014 onwards, making it possible to distinguish students in ESPE schools from other students. This was not the case for the years 2012 and 2013, although this training program already existed.

¹²Prior to 2010 in secondary education, candidates for the "Aggregation" exam already had to have completed 4 years of university, whereas candidates for the "CAPES" exam had to have completed 3 years of university. See Chapter 1 of this thesis for more information on the different statutes and recruitment competitions for secondary school teachers.

the number of master's degree holders. Table 3.2 shows the orientation choices of new bachelors' graduates for the 2005-2009 (column 1) and 2012-2016 (column 2) periods ¹³. A comparison of columns 1 and 2 shows that the continuation rate after a bachelors' degree increased from 69% before 2010 to 79% after the 2011 reform. Moreover, from 2011 onwards, 76% of bachelor's degree holders enroll in the first year of a master's program and 49% obtain a master's degree two or three years later, compared to 66% and 37% respectively before the reform. Section 3.4.2 will provide additional statistical evidence regarding these aspects.

Table 3.2 – Orientation of bachelor's graduates during the periods 2005-2009 and 2012-2017

	2005-2009	2012-2017
Percentage of bachelor's graduates who :		
Permanently or temporarily stop their studies (leading to a diploma) (%)	31	21
Enroll in another bachelor's degree program (%)	3	3
Enroll in the first year of a master's program* (%)	66	76
Complete a master's degree in two or three years (%)	37	49
Apply for public positions in primary and secondary schools** (%)	33	20

* Including management and engineering schools

** Authors' computations using the Ministry of Education competition data.

Notes: This table shows the educational and professional orientation of bachelor's graduates during the 2005-2009 and 2012-2017 periods. Data published by the Ministry of Higher Education slightly underestimate the proportion of bachelor's graduates enrolled in master's programs because the data cover 90% of post-secondary schools. In addition, the data do not take into account the fact that some students temporarily interrupt their studies and return to university later. 20% of bachelor's graduates applied for public primary or secondary teaching positions in 2012-2017, against 33% over the period 2005-2009.

Source: : Ministry of Education (DEPP) - Ministry of Higher Education (SIES)

3.4.2 Teacher supply and teacher demand

In France in 2016, primary and secondary school teachers account for 1.3% and 1.7% of the working population respectively¹⁴. Approximately 27,000 tenured primary and secondary public school teachers are recruited each year ¹⁵, representing approximately 9% of the general baccalaureate graduates, 16% of the general bachelor's graduates and 22.4% of the general master's graduates (excluding engineers and business school graduates) of a

¹³"La poursuite d'études des diplômés de licence en première année de master", Note d'information, SIES, Juin 2017, N°7. "Parcours et réussite aux diplômes universitaires: les indicateurs de la session 2014", Note Flash, SIES, Février 2016, N°1. "Parcours et réussite en licence et en master à l'université", Note d'information, SIES, Avril 2013, N°2. "Parcours et réussite en licence des inscrits en L1 en 2004", Note d'information, SIES, Novembre 2009, N°23

¹⁴Source: Bilan social du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche - 2015-2016. Tableaux de l'économie française, Insee Référence, Edition 2018.

¹⁵Less than 2% of tenured teachers resign or are dismissed during their career.

generation¹⁶. The number of candidates in a generation is about twice as many as the number of positions.

Table 3.3 shows, on average by year, the number of positions in primary schools, the number of applicants and the number of higher education graduates at national level and by region (mean value, standard deviation, minimum and maximum value in the twenty-five metropolitan regions when data are available). A distinction is made between the period 2003-2009, which precedes the increase in the level of qualification required to teach, the period 2011-2013, during which candidates must be enrolled in the second year of a master's degree, and the period 2014-2017, during which candidates must be enrolled in the first year of a master's degree.

Table 3.3 shows that the number of positions in primary schools decreases by 50% over the period 2011-2013 compared to the periods 2003-2009 and 2014-2017. The number of candidates¹⁷ follows a similar trend but is almost twice as low in 2014-2017 than in 2003-2009. The table also shows a strong heterogeneity between regions: the smallest region has an average of 100 teaching posts per year, while the largest region has an average of 1350 teaching posts during the period studied.

Table 3.3 also shows the number of general baccalaureate graduates three years before the teaching competition, the number of bachelor's graduates and the number of master's graduates during the three periods considered. The general baccalaureate (*baccalauréat général*) is a secondary school leaving examination which gives access to the university and preparatory classes for the *Grandes Ecoles* in France¹⁸. The vast majority of teachers obtained a general baccalaureate 3-4 years before taking the teacher recruitment exam. The number of general baccalaureate holders is relatively stable over the 2003-2017 period

¹⁶Source: SIES, DEPP, INSEE. 16% of primary and secondary school teachers work in private education. Public and private primary and secondary school teachers recruited each year represent about 10.7% of the general baccalaureate graduates, 19% of the bachelor's graduates and 26.7% of the master's graduates (excluding engineering and business school graduates) of a generation.

¹⁷Most of the candidates take the competition at least twice. As a result, the number of unique candidates is about half the number shown in the table 3.3

¹⁸There are three types of secondary school leaving examinations that give access to higher education. The general baccalaureate gives access to long university programmes (bachelor's and master's degrees for example) and to preparatory classes for the *Grandes Ecoles*. The technology baccalaureate mainly provides access to shorter specialized programs, although the best students may go on to university. The professional baccalaureate gives access to the shortest specialized programs, which generally last two years. The success rate for the baccalaureate varies by region, year and type of baccalaureate. Overall, the success rate is close to 90% (Le Laidier & Thomas 2013). In 2016, approximately 40% of high school students obtained a general baccalaureate, 16% obtained a technology baccalaureate, 22% obtained a professional baccalaureate, and 22% stopped their studies at a level below the baccalaureate (DEPP 2016, Repère et références statistiques).

(270,000 people per year on average). However, it varies greatly from one region to another; the smallest region has an average of 2,800 general baccalaureate graduates and the largest region has an average of 27,000 general baccalaureate graduates each year. The number of bachelor's and master's graduates is only available at national level. The number of graduates tends to increase over time, but the number of master's graduates increases more rapidly from 2011 onwards¹⁹. On average, there are 5.9 master's graduates for every 10 bachelor's graduates in 2005-2009 and 7.1 master's graduates for every 10 bachelor's graduates in 2011-2017. In other words, the number of master's graduates increased by 12 percentage points more than the number of bachelor's graduates over the same period. This substantial increase (which corresponds to approximately 23,000 additional master's graduates each year after adjusting for time trends) may be related to the increase in the level of qualification required to teach in primary and secondary education from 2011 onwards²⁰.

The last part of Table 3.3 shows the number of positions in primary schools and the number of applicants divided by the number of general baccalaureate holders at the national level and by region over the period 2003-2017. This standardization of labour demand (number of teaching positions) and labour supply (number of candidates) neutralizes the size effects and makes the regions more comparable. Table 3.9 shows that primary school teacher recruitment is very local: 82% of candidates take the exam in the region where they live (although they are not required to).

Table 3.3 shows that demand for teachers is halved in the 2011-2013 period compared to previous and subsequent periods. Although the number of positions per general baccalaureate graduate is rigorously the same at the national level in 2003-2009 and 2014-2017, the number of candidates per general baccalaureate holders is twice as low in 2014-2017 as in 2003-2009. As with the sudden increase in the number of master's graduates over the 2011-2017 period, we suspect that the increase in the level of qual-

¹⁹The original data are annual since 2005 and show that the relative increase in the number of master's graduates is taking place in 2011. Source: Bilan Formation-Emploi. *L'école et ses sortants - Diplômes de l'enseignement supérieur*. CEREQ-DEPP-SIES. <https://www.insee.fr/fr/statistiques/2522723?sommaire=2526273>

²⁰Note that the number of bachelor's and master's graduates in Table 3.3 does not refer to the population eligible for teaching positions. First, about one in ten undergraduate students and two in ten master's students are foreign nationals. Outside the European Union, foreign nationals are not eligible for teaching posts in France. Second, the numbers in Table 3.3 include both graduates of generalist programs that lead to teaching and graduates of professional programs far removed from teaching. Graduates of specialized schools such as engineering schools, business schools, journalism schools, veterinary schools are excluded from Table 3.3.

ifications required to teach from 2011 onwards is responsible for this sharp decline in the supply of teachers. The demand for teachers also differs greatly from one region to another. Over the period 2014-2017, demand varies between 1.5 primary education posts per 100 general baccalaureate graduates and 10 primary education posts per 100 general baccalaureate graduates. These differences between regions are likely to reflect differences in demographic trends (number of school-age children) as well as differences in attractiveness (teacher assignments and mobility) between regions (Hilary & Louvet 2013). Tables 3.10 and 3.11 provides detailed figures for the twenty-five metropolitan regions.

Table 3.3 – Number of teaching posts in primary schools, number of applicants and number of higher education graduates in 2003-2017

	2003-2009	2011-2013	2014-2017
Primary school positions			
All metropolitan regions*	9450	4925	10025
Mean	378	197	401
SD	269	202	330
Min	89	34	120
Max	1244	865	1456
Candidates to primary school positions			
All metropolitan regions*	47825	16575	25450
Mean	1913	663	1018
SD	912	361	437
Min	496	163	444
Max	4401	1635	1990
General baccalaureate graduates year t-3			
All metropolitan regions*	262850	272700	277200
Mean	10514	10908	11088
SD	4782	5376	5628
Min	2894	2708	2658
Max	23837	27561	29058
Bachelor's graduates			
All regions	161000	168000	175000
Mean	6440	6720	7000
Masters' graduates			
All regions	93500	123000	124000
Mean	3740	4920	4960
Primary school positions / General baccalaureate graduates year t-3			
All metropolitan regions*	0.036	0.018	0.036
Mean	0.035	0.016	0.037
SD	0.011	0.008	0.018
Min	0.017	0.008	0.015
Max	0.076	0.047	0.099
Candidates / General baccalaureate graduates year t-3			
All metropolitan regions*	0.182	0.061	0.092
Mean	0.186	0.061	0.095
SD	0.036	0.012	0.017
Min	0.083	0.026	0.052
Max	0.245	0.084	0.128

* Excluding Corsica

Notes: Number of positions and number of candidates for the external primary school teacher recruitment competition. Mean, standard deviation, minimum and maximum values for metropolitan academic regions, excluding Corsica, over the periods 2003-2009, 2011-2013 and 2014-2017. The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. Number of bachelor's and master's degrees is only available at the national level and rounded to the nearest thousand in the publications of the Ministry of Higher Education. Master's and Bachelor's degrees were created in 2005 in France.

Source: Ministry of Education (DEPP) and Ministry of Higher Education (SIES). Author's calculations.

3.4.3 Teacher *skills*

This study relies on two measures of teacher competence. The first measure corresponds to the average score obtained in the written tests of the teacher recruitment competition. The written examination consists of four tests in six subjects taught at elementary school: mathematics, French, history, geography, science and technology. The first part of the

tests assesses candidates' knowledge and understanding of the school curriculum. The second part assesses teaching skills (for example, candidates should identify students' mistakes and propose appropriate teaching strategies).

I consider the scores obtained in the 2003-2013 competitive recruitment examination because the written tests changed in 2014. I normalize the exam scores (mean zero and variance one) and I propose a corrected score that takes into account the imperfect harmonization of rating practices across regions and over time. I explain in the Appendix how I correct the score, which allows to estimate an upper bound of the effect of the teacher degree level reform. Tables 3.12 and 3.13 present the average score obtained by candidates and teachers over three periods: 2003-2006, 2007-2010 and 2011-2013. The first three columns correspond to the exam score observed in the data, while the last three columns correspond to the score corrected by the method proposed in the Appendix. Table 3.12 shows that the average score of candidates increased by a maximum of 0.2 SD points over the period 2011-2013 compared with previous periods. Conversely, Table 3.13 reveals that the average score of teachers decreased by at least 0.15 SD points over the period 2011-2013 compared to previous periods.

The second skill measure used in this study is the classroom observation score, or pedagogical score, assigned by education professionals (*inspecteurs pédagogiques*) who are mandated by the local education authorities of the academic regions (*rectorats*). On average, this score is revised every four years and is used by the administration to determine teacher promotion. Until 2017, the classroom observation score was not based on a harmonized evaluation grid and depended heavily on seniority. The evaluation criteria were left entirely to the inspectors' discretion²¹. Consequently, the comparison of teachers' pedagogical scores only makes sense for those working in the same inspection area (each pedagogical inspector is assigned to a geographical area comprising ten to fifteen schools) and who have the same level of experience.

I consider the classroom observation score obtained two-three years after tenure, which maximizes the number of observations at a given seniority level. I normalize these scores (mean zero and variance one) over the period 2004-2015. Table 3.14 shows the average pedagogical score obtained by the teachers recruited during the periods 2006-2009 (column 1) and 2010-2013 (column 2). The table shows that the classroom observation score

²¹"L'évaluation des enseignants", Inspection générale de l'éducation nationale. Rapport 2013

increases slightly (0.02 SD points) on average over the 2010-2013 period.

Currently, France does not have standardized tests in primary schools. It is therefore not possible to measure the ability of teachers to advance students on standardized tests (value-added models).

3.4.4 Teacher diversity

In 2008-2009, children of blue-collar workers and employees accounted for 27.1% of undergraduate students and 17.9% of master's students²². The increase in opportunity costs associated with raising the level of qualifications required to teach can affect the professional choices of students from the most modest backgrounds. Inequalities in access to educational posts according to social origin are therefore likely to increase after the reform in 2011.

In the same academic year, men accounted for 42.6% of bachelors' students and 42.3% of master's students. This small difference in the continuation rate is not expected to significantly change the gender composition of teachers after the 2011 reform. However, the Labor Force Surveys show that men in full-time employment earn on average 12.7% more than women in full-time employment at the bachelor's level, while men earn on average 15.7% more than women at the master's level²³. Given that non-teaching salaries are higher for men than for women at the master's level rather than at the bachelor's level, the increase in the degree level required to teach in 2011 should reduce the share of men among primary school teachers.

The Ministry of Education's competition database provides information on gender, but does not contain information on the social background of applicants. Therefore, I consider the median income in the city of residence²⁴ to approximate the social origin of applicants. Table 3.4 shows that men represented 18.3% of candidates and 16.2% of teachers during the period 2003-2009. In contrast, they represent only 14.6% of applicants and 13.6% of teachers in the period 2011-2015, a decrease of about 18%. Table 3.4 shows that the median net income in the teachers' city of residence is slightly higher than in

²²Source: "Les étudiants", *Repères et références statistiques*, édition 2009. DEPP-SIES.

²³The gender pay gap is estimated for non-teaching staff working full-time over the period 2010-2016. The estimates are obtained by neutralizing the effects of age and weighting the observations so as to reproduce the distribution of primary school teachers' diploma specializations.

²⁴This is the median income in euros per consumption unit (i.e. divided by the number of inhabitants in the household). To avoid problems related to income growth over time, this is the median net income in cities in 2012. Source: INSEE.

the applicants' city of residence. In contrast, median income in the city of origin remains stable over the 2003-2015 period for both applicants and teachers.

Table 3.4 – Gender and social composition of candidates and primary school teachers over the 2003-2015 period

	2003-2009	2011-2015
Share of men (%)		
Candidates	18.3	14.7
Recruited teachers	16.2	13.7
Median net income in city of residence (in euros in 2012)		
Candidates	20,527	20,560
Recruited teachers	20,921	20,981

Note: Proportion of men among candidates and teachers recruited for the primary school examination during the 2003-2009 and 2011-2015 periods. For each candidate and each teacher recruited during the 2003-2009 and 2011-2015 periods, I calculate the median income after taxes (in euros per consumption unit) in his/her city of residence. I consider the median income in 2012 in order to neutralize the effect of income variations over time. Median income in the city of residence is considered to be an approximate measure of individuals' social origin.

Source: Ministry of Education (DEPP) and Institute of Official Statistics (INSEE). Author's calculations.

3.4.5 Teacher salaries

The theoretical arguments and empirical evidence presented in Chapter 2 suggest that salaries of teachers and competing professions can influence teacher supply and characteristics. Failure to take into account the effect of wages may bias the assessment of the diploma level reform in 2011. First, this reform was implemented in a context of wage stagnation for all civil servants between 2010 and 2016 (*gel du point d'indice des fonctionnaires*). The inflation reaches 6.5% while average earnings for undergraduate and graduate occupations may have continued to increase over this period. Second, the reform was followed by a slight increase in teachers' salaries at the beginning of their careers: an increase of 5% in the first year, 3% in the second and third years, 1% in the fourth and fifth years and no increase beyond five years of service. Since September 2016, primary school teachers have also received a premium of 89 euros net per month, which corresponds to an average wage increase of 4.5%. More recently, a gradual wage increase has been introduced from 2017 onwards and is expected to continue until 2020.

I use the Labour Force Survey to estimate the salaries of primary school teachers and those in other bachelor's and master's level occupations over the 2003-2016 period. The number of bachelor's and master's graduates is sufficient to estimate salaries by region, but teachers' wages can only be estimated at the national level. This is not a major issue

as teachers' salaries are entirely governed by a national salary scale²⁵. The teacher salary scale is publicly available and the Ministry of Education publishes teacher salaries by status and age group once a year.

Table 3.5 presents the average salaries of full-time primary school teachers at the national level and the average salaries of full-time bachelor's and master's graduates (excluding teachers) at the national level and by region during the periods 2003-2009, 2011-2013 and 2014-2016. Salaries are estimated at age 41 (average age of the active population with a bachelor's or master's degree according to the Labor Force Survey), which neutralizes age differences between regions, diploma levels and occupations. In addition, the salaries of bachelor's and master's graduates are weighted so that the degree specialties correspond (in distribution) to those of primary school teachers over the 2003-2016 period. In other words, Table 3.5 presents an estimate of the salary teachers could expect in another profession at the bachelor's or master's level at age 41.

Table 3.5 shows that elementary school teachers receive on average a lower salary than those with a bachelor's and master's degree in the same specialty at the same age, regardless of the period considered²⁶. The table also shows wage disparities between regions by degree level. In 2014-2016, the average salary of a bachelor's graduate (at age 41) is 2,312 euros net per month and varies between 2,110 euros and 2,513 euros depending on the region. Over the same period, the average salary of a master's graduate is 2,844 euros net per month and varies between 2,314 euros net per month and 3,262 euros net per month depending on the region. In comparison, the average salary of a full-time primary school teacher is 2,129 euros net per month over the period 2014-2016 (calculated at national level at age 41). Naturally, these estimates are subject to measurement (relatively small sample sizes) and reporting errors. They must therefore

²⁵On average, the salary of primary school teachers is 95% determined by status (type of contract) and experience. 2.3% of the salary corresponds to overtime and approximately 3% corresponds to bonuses related to residence, specific functions (pedagogical advice or supervision of trainee teachers, for example) or assignment to a Priority Education Network school (*Réseau d'Éducation Prioritaire*). Pupils in the Priority Education Network schools are mainly from disadvantaged socio-economic backgrounds. These schools receive additional financial and human resources to reduce social inequalities in educational attainment. Primary school teachers assigned to Priority Education Network schools receive a salary premium of 5% of the average salary. About 13% of primary school teachers are assigned to a school in the priority education network. Source: Bilan social du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche - 2015-2016.

²⁶The average salary reported by teachers in the Labour Force Survey is about 7% lower than that calculated from Ministry of Education administrative data. Since the estimated salaries of bachelor's and master's graduates are subject to the same biases, I keep the salaries of teachers calculated from the Labour Force Survey in the rest of the study

be considered with caution.

Table 3.5 – Average salaries of bachelor’s graduates, master’s graduates and primary school teachers in the 25 academic regions of metropolitan France over the period 2003-2017

	2003-2009	2011-2013	2014-2016
Primary school teachers (euros)			
At the metropolitan level	1931	2074	2129
Bachelors’ graduates (euros)			
At the metropolitan level	2166	2307	2312
Mean	2096	2266	2283
SD	164	192	128
Min	1855	1955	2110
Max	2407	2686	2513
Masters’ graduates (euros)			
At the metropolitan level	2762	3010	2844
Mean	2654	2871	2707
SD	204	290	273
Min	2324	2331	2314
Max	3066	3384	3262

Note: Average (self-reported) net monthly salary of primary school teachers at the metropolitan level for the periods 2003-2009, 2011-2013 and 2014-2016. Average (self-reported) net monthly salary of bachelor’s and master’s degree holders at the metropolitan level. Average, standard deviation, minimum and maximum values for the 25 academic regions of metropolitan France.

Source: Labor Force Survey. Author’s calculations. *Scope:* Metropolitan France excluding Corsica

3.5 Exploratory analysis when teachers are recruited at the bachelor level (before 2011)

3.5.1 Relationship between teacher supply, teacher demand and salary

Figure 3.8 shows how teacher supply varies with teacher demand over the period 1996-2009. The Y axis indicates the number of candidates per *baccalaureate* graduates three years prior to the competition and the X axis corresponds to the number of positions per *baccalaureate* graduates three years before the competition. The number of *baccalaureate* graduates is used to standardize teacher supply and demand, which allows to compare regions with very different demographics. Each cross represents a region of a given year and the solid line corresponds to a polynomial adjustment with a 5% confidence interval. The light blue crosses correspond to the years 2003-2004 when candidates were exceptionally allowed to take the examination in two different regions, which artificially increased the number of candidates. The figure shows that the supply of teachers increases with the demand for teachers. However, the elasticity of the number of candidates with

the number of posts decreases with the demand for teachers, which means that each new additional teaching position attracts slightly fewer candidates.

Table 3.6 shows how the number of applicants varies in the twenty-five metropolitan areas with the number of teaching positions, the unemployment rate and the average wage in the concurrent occupations (other than teaching) at bachelor and master level²⁷. The first column does not include control variables, column 2 includes year fixed effects and column 3 includes year and region fixed effects. The variable explained in column 4 is the logarithm of the number of candidates. To standardize the regions, I divide the number of positions and candidates by the number of baccalaureate graduates three years before the teacher competition, which is proportional to the size of the eligible population to teaching positions. Table 3.6 confirms that the elasticity of the number of candidates with the number of positions is positive, but decreases with labor demand. It also shows that the number of candidates decreases significantly with the average salary of master's graduates, but not with the average salary of bachelor's graduates when control variables are added. Moreover, teacher supply is positively associated with the unemployment rate in the region. These results are consistent with the empirical evidence and theoretical predictions presented in Chapter 2. The increase in the unemployment rate and the decrease in average (non-teaching) wages contribute to the decline in the attractiveness of alternative occupations²⁸, which has a positive effect on teacher supply.

²⁷Data are aggregated at the regional level. The salary estimates take into account the effect of age and ensure that the diploma specialization are representative of those of primary school teachers. Source: French Labor Force Survey.

²⁸In the theoretical model developed in Chapter 2, we can replace the expected wage in alternative occupations w by the expected work income to take into account the risk of unemployment: $w * q + w_0 * (1 - q) = (w - w_0) * q + w_0$, where q is the unemployment rate and w_0 is the unemployment benefit associated with wage w . If the replacement salary is proportional to the wage, $w_0 = \alpha w$, then we get the expected income: $(1 - \alpha) * q * w + \alpha * w$.

3.5. Exploratory analysis when teachers are recruited at the bachelor level (before 2013)

Table 3.6 – Relationship between teacher supply, teacher demand and competing employment opportunities over the period 2003-2009

	(1)	(2)	(3)	(4)
				Logarithm of number of candidates per baccalaureate holder t-3
			Number of candidates per baccalaureate holder t-3	
Teacher demand				
Number of positions per baccalaureate holder t-3	6.731*** (0.704)	9.049*** (0.814)	7.421*** (1.204)	
(Number of positions per baccalaureate holder t-3) ²	-46.30*** (7.851)	-62.26*** (9.442)	-49.16*** (10.03)	
Logarithm of number of positions per baccalaureate holder t-3				0.560*** (0.0913)
Competing job opportunities				
Unemployment rate (%)	0.00717*** (0.00205)	0.0139*** (0.00384)	0.0129 (0.00797)	0.0585* (0.0317)
Net salary of master's graduates / 100	0.00349*** (0.000838)	-0.000737 (0.000801)	-0.00126* (0.000756)	-0.00550* (0.00313)
Net salary of bachelor's graduates / 100	0.00378*** (0.00127)	-0.00149 (0.00137)	-0.0000850 (0.00127)	0.00238 (0.00525)
Control variables				
Year FE	No	No	Yes	Yes
Region FE	No	Yes	Yes	Yes
Observations	200	200	200	200
Adjusted R-squared	0.58	0.73	0.78	0.88

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parenthesis.

Notes: Data are aggregated at the regional level. Each column explains the supply of teachers by teacher demand, the unemployment rate and the average salary at the bachelor's and master's level in the region over the period 2003-2009. Column (4) is a log-log specification of teacher supply and demand. The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. I divide the number of candidates and the number of positions by the number of baccalaureate graduates 3 years before the competition date in order to make the regions comparable. Estimates of salaries at the bachelor's and master's level take into account the effect of age and reflect diploma specializations of the primary school teachers. Wages are divided by 100 to make the coefficients more readable.

Source: French Labor Force Survey. Ministry of Education (DEPP). National Institute of Statistics (INSEE).

3.5.2 Relationship between teacher characteristics, teacher demand and salary

Table 3.7 shows how the examination score and classroom observation score vary with teacher demand, the unemployment rate in the region, and the salaries of bachelor's and master's graduates over the 2003-2009 period. The models presented in Table 3.7 include year and region fixed effects. The results show that teacher demand is negatively correlated with teacher recruitment score but is not significantly associated with candidate recruitment score or classroom observation scores. This finding suggests that increasing the number of teaching positions is not attracting enough candidates to the top of the knowledge distribution, which leads examiners to be somewhat less selective in filling

teaching positions. On the other hand, the demand for teachers does not seem to affect the teaching skills of recruits.

The table also shows that the unemployment rate in the region is positively associated with candidate and teacher examination scores, but negatively associated with classroom observation scores. This finding suggests that the decline in the attractiveness of competing professions tends to increase the average knowledge of applicants and teachers; however, this leads to the recruitment of less pedagogically qualified teachers. The effect of the salaries of competing occupations at the master's and bachelor's levels is generally insignificant. However, we note that the salary of bachelor's graduates is significantly and positively associated with the classroom observation score, which is consistent with the negative effect of the unemployment rate on the classroom observation score.

Table 3.8 shows how the share of men and median income in the city of residence²⁹ vary according to teacher demand, the unemployment rate and the wages of bachelor's and master's graduates. Columns (1) and (2) show that the proportion of men among candidates and among teachers is not associated with the demand for teachers, nor with the unemployment rate. However, the share of men among teachers decreases considerably with the average salary of master's graduates. This finding suggests that the share of men decreases at the top of the knowledge distribution as master's level occupations become more attractive, which is consistent with the fact that the wage gap between men and women increases with the level of educational attainment. Columns (3) and (4) show that median income in the city of residence of applicants and teachers increases with teacher demand, but is not significantly associated with unemployment rates or wages in competing occupations.

²⁹This is the median income in euros per consumption unit (i.e. divided by the number of inhabitants in the household). In order to neutralize the effects of the increase in median income over time, we consider the median income that was calculated for the city in 2012.

3.5. Exploratory analysis when teachers are recruited at the bachelor level (before 2013)

Table 3.7 – Relationship between average salaries of bachelor’s and master’s graduates and teachers’ *competencies* in the period 2003-2009

	(1)	(2)	(3)
	Candidates	Teachers	
	Exam score	Exam score	Classroom observation score
Teacher demand			
Logarithm of number of positions per baccalaureate holder t-3	-0.0151 (0.0542)	-0.480*** (0.123)	0.407*** (0.149)
Alternative job opportunities			
Unemployment rate (%)	0.0404** (0.0188)	0.104** (0.0426)	-0.174*** (0.0604)
Net salary of master’s graduates / 100	0.000565 (0.00186)	-0.000855 (0.00420)	0.00951** (0.00430)
Net salary of bachelor’s graduates / 100	0.00336 (0.00312)	0.0101 (0.00706)	0.0183** (0.00769)
Control variables			
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	200	200	125
Adjusted R-squared	0.87	0.90	0.91

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis.

Notes: Data are aggregated at the regional level. Column (1) explains the exam score of candidates by teacher demand, the unemployment rate and the average salary at the bachelor’s and master’s level in the region over the period 2003-2009. Column (2) explain the exam score of teachers and column (3) explain the classroom observation score of teachers (2-3 years after tenure) by teacher demand, the unemployment rate and the average salary at the bachelor’s and master’s level in the region over the period 2003-2009. The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. I divide the number of candidates and the number of positions by the number of baccalaureate graduates 3 years before the competition date in order to make the regions comparable. Estimates of salaries at the bachelor’s and master’s level take into account the effect of age and reflect diploma specializations of the primary school teachers. Wages are divided by 100 to make the coefficients more readable.

Source: French Labor Force Survey. Ministry of Education (DEPP). National Institute of Statistics (INSEE).

Table 3.8 – Relationship between the average salaries of bachelor's and master's graduates and the composition of the teaching force during the period 2003-2009

	(1)	(2)	(3)	(4)
	Share of men		Median net income in city of residence	
	Candidates	Teachers	Candidates	Teachers
Teacher demand				
Number of positions per baccalaureate holder t-3	0.281 (0.229)	-0.480 (0.438)	6392.1** (2846.5)	13589.4*** (3651.4)
Alternative job opportunities				
Unemployment rate (%)	-0.00366 (0.00340)	-0.00420 (0.00652)	-19.50 (42.36)	-39.77 (54.34)
Net salary of master's graduates / 100	-0.000267 (0.000336)	-0.00147** (0.000644)	-1.306 (4.182)	2.494 (5.365)
Net salary of bachelor's graduates / 100	-0.000248 (0.000564)	-0.000620 (0.00108)	6.447 (7.022)	10.22 (9.007)
Control variables				
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	200	200	200	200
Adjusted R-squared	0.73	0.33	0.99	0.98

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parenthesis.

Notes: Data are aggregated at the regional level. Columns (1) and (2) explain the share of men among candidates and teachers by teacher demand, the unemployment rate and the average salary at the bachelor's and master's level in the region over the period 2003-2009. Columns (3) and (4) explain the median income (in euros per year) in the city of residence of candidates and teachers by teacher demand, the unemployment rate and the average salary at the bachelor's and master's level in the region over the period 2003-2009. To neutralize the effect of income growth over time, I consider the median income in euros per consumption unit (i.e. divided by the number of inhabitants in the household) in cities in 2012. The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. I divide the number of candidates and the number of positions by the number of baccalaureate graduates 3 years before the competition date in order to make the regions comparable. Estimates of salaries at the bachelor's and master's level take into account the effect of age and reflect diploma specializations of the primary school teachers. Wages are divided by 100 to make the coefficients more readable.

Source: French Labor Force Survey. Ministry of Education (DEPP). National Institute of Statistics (INSEE).

3.5.3 Relationship between diploma level and teacher characteristics

Tables 3.15 and 3.16 show how gender and median income in the city of residence correlate with the degree level of applicants over the period 2003-2009. Table 3.15 shows that the proportion of men among candidates (column 1-3) and among teachers (column 4-6) increases with the level of the degree: they represent 18% of candidates with a bachelor's degree, 20% of candidates with a four-year university degree and 22% of candidates with a master's degree or higher. Although on average they have a higher level of education, men score lower than women on written tests (-0.17SD point), suggesting that the selection processes that lead men and women to the teaching profession are different. Table 3.16 shows that the median income in the candidates' (respectively teachers') city of residence

increases with their level of education. For example, the median income in the cities of candidates (teachers) with a master's degree is 650 euros (790 euros) higher than the median income in the cities of candidates (teachers) with a bachelor's degree. Candidates who are not subject to a degree requirement (high performance athletes and parents of three or more children) are also more likely to live in cities with a high median income (+930 euros compared to candidates with a bachelor's degree).

Table 3.17 reveals that the score on the written tests tends to increase with the degree level of candidates (columns 1 to 3) and teachers (columns 4 to 6)³⁰. These results are consistent with the empirical literature (Hanushek and Rivkin, 2004; Jacob et al., 2016) and suggest that an increase in the level of qualification required to teach can have a positive impact on teachers' knowledge. Table 3.19 shows that the classroom observation score does not systematically vary with the qualification level of teachers. The model in column 4 includes inspection area fixed effects and the model in column 5 includes school fixed effects. This result suggests that teaching skills are not directly related to the qualification level of teachers once they have successfully completed the recruitment process and validated the trial period.

This preliminary analysis suggests that raising the qualification level of teachers could contribute to improving their knowledge without compromising their pedagogical skills. Teacher gender diversity could also improve but at the expense of social diversity. However, these results are likely to be endogenous because teachers have made the choice of whether or not to continue their studies after a bachelor's degree. Section 4.6 will examine how these relationships evolve when a master's degree becomes compulsory for teaching in primary schools.

3.6 Model

To estimate the impact of a higher degree requirement on teacher supply and characteristics, I consider the following specification:

$$Y_{jt} = \sum R_j 1_j + \sum \delta_t 1_t + \alpha N_{P_{jt}} + \gamma X_{jt} + \beta Z_{i(j)t} + \epsilon_{jt} \quad (1)$$

³⁰The results are very similar whether we consider the raw score or the corrected examination score, see Table 3.18

where Y_{jt} is alternatively the number of candidates, the exam score of candidates (respectively teachers), the share of men among candidates (teachers) or the median income in the city of residence of candidates (teachers) for region j and year t . R_j is a region fixed effect that captures the time-invariant specificities of region j , δ_t is a year fixed effect, $N_{P_{jt}}$ is the number of positions in region j and year t , X_{jt} is a vector of time-varying characteristics in region j (unemployment rate, average salaries of bachelors' and masters' graduates), $Z_{i(j)t}$ is a vector of individual characteristics in region j (number of times they participate in the competition) and ϵ_{jt} is a mean zero error term. The coefficients of interest are the year fixed effects δ_t . The identification of the effect of the degree level required to teach is based on the assumption that the 2011 reform agenda is exogenous conditional on control variables.

To estimate the impact of the reform on the classroom observation score,³¹ I prefer the following specification:

$$Y_{kt} = \sum I_k 1_k + \sum \delta_t 1_t + \alpha N_{P_{j(k)t}} + \gamma X_{kt} + \beta Z_{i(k)t} + \epsilon_{kt} \quad (2)$$

where Y_{kt} is the classroom observation score of teachers working in inspection area k (or school k depending on the specification) at year t and I_k an inspection area (or school) fixed effects. The coefficients of interest are the year fixed effects δ_t .

Several arguments justify using the expression (2) rather than (1) when the classroom observation score is the variable explained. First, education inspectors are assigned to inspection areas that cover several schools in the academic region, but there are no standard criteria (until 2017) for evaluating teachers. Therefore, fixed effects of inspection areas will capture variations in assessment practices between education inspectors. Second, the classroom observation score is used by the administration for teacher promotion. As a result, teachers who work in difficult and unattractive schools are likely to receive higher grades that encourage them to stay a little longer. School fixed effects will neutralize these effects.

³¹I consider the classroom observation score two years after tenure. At this time, 99.6% of teachers still work in the academic region where they were recruited

3.7 Results

3.7.1 Effect of a higher degree requirement on teacher supply

Figure 3.2 shows how the number of candidates for the competition varies according to the number of teaching positions in primary schools. Each point represents a region x year, the blue crosses correspond to the period 1996-2009 and the red triangles to the period 2011-2018. I divide the number of candidates (Y-axis) and the number of positions (X-axis) by the number of general baccalaureate graduates in the region three years before the competition date. This standardization makes the twenty-five metropolitan regions more comparable.

Figure 3.2 shows that the number of candidates increases with the number of positions before (1996-2009) and after (2011-2018) teacher qualification reform. However, for a given number of posts, the number of candidates is much lower after the reform. For example, regions with 4 teaching posts (for around 100 baccalaureate holders) had on average 20 candidates (for around 100 baccalaureate holders) before the reform but fewer than 10 candidates (for around 100 baccalaureate holders) after 2011. Moreover, the elasticity of teacher supply with teacher demand is lower in the period 2011-2018 than in the period 1996-2009. The theoretical model in Chapter 2 predicts that the number of candidates converges towards the number of potential candidates (i.e. the number of people whose utility for teaching is greater than the utility provided by a competing profession) when the number of positions increases. Before the reform, the number of candidates converges towards 25 per 100 general baccalaureate graduates when the demand for teachers exceeds 6 posts per 100 general baccalaureate graduates. After the reform, the number of candidates converges to 10 per 100 general baccalaureate graduates when the demand for teachers exceeds 3 posts per 100 general baccalaureate graduates. These results suggest a much more rapid saturation of the pool of potential candidates (and a much smaller pool of potential candidates) as a result of the increase in the level of diploma required to teach starting in 2011³². However, the results suggested by Figure 3.2 can be subject

³²The number of candidates artificially increases in 2003 and 2004 due to the exceptional authorization to participate in the competition in two regions during the same year. This exception translates into a number of "outliers" on Figure 3.2 (in the area around 3 to 5 positions per 100 general baccalaureate graduates). More generally, a lot of candidates take the primary teacher recruitment examination twice or more. The number of *unique* candidates is obtained by approximately halving the official number of candidates.

to different forms of bias related to regional heterogeneity and temporal shocks.

Table 3.20 shows the results of the model (1) estimation. Column (1) controls the number of teaching positions in the region and reproduces the results of Figure 3.2. Column (2) includes region fixed effects and column (3) adds the number of positions squared and the unemployment rate as additional control variables. Column (4) shows the results when I consider a logarithmic version of the model (1):

$$\log(Y_{jt}) = \sum R_j 1_j + \sum \delta_t 1_t + \alpha \log(N_{P_{jt}}) + \gamma X_{jt} + \beta Z_{i(j)t} + \epsilon_{jt} \quad (1')$$

Model (1') makes it possible to study multiplicative effects rather than additive effects on teacher supply.

The first three specifications lead to similar estimates: on average, the regions lose 670 candidates after the reform (for about 1900 candidates over the period 1996-2009 and 1600 candidates over the recent period 2007-2009). The fact that the estimates are robust to the inclusion of region fixed effects and control variables suggests that the interpretation of Figure 3.2 is not seriously biased by heterogeneity across regions. We also note that the number of candidates increased considerably during the 2003-2004 period due to the possibility of taking the examination in two different regions (+ 590 candidates on average). The effect seems to have spread to the years 2005-2007 even though candidates are no longer allowed to take the exam in two regions (+ 290 candidates on average).

The log specification is used to examine the relative change in the number of candidates over time for a given number of positions. Column (4) confirms that the number of candidates (for a given number of teaching posts) decreases on average by 50% over the period 2011-2013 and by 40% over the period 2014-2018 compared to 1996. To clarify the reading of the results, I report on Figure 3.3 the coefficients and confidence intervals at the 5% level of column (4). The figure shows that it is reasonable to exclude the assumption of a downward trend in the number of candidates over time. It also confirms that the most significant change occurs at the time of the 2011 reform³³. To examine if the effect of the 2011 reform increases with teacher demand as Figure 3.2 suggests, we interact the number of teaching posts and the post-reform period. The estimates are reported in Table 3.21 and confirm that the decline in the number of candidates is

³³The effect of the exceptional authorization to pass the competition in two regions in 2003 and 2004 is visible on the Figure 3.3, as well as the contamination effect for the years 2005-2007.

significantly more pronounced in regions where the demand for teachers is higher.

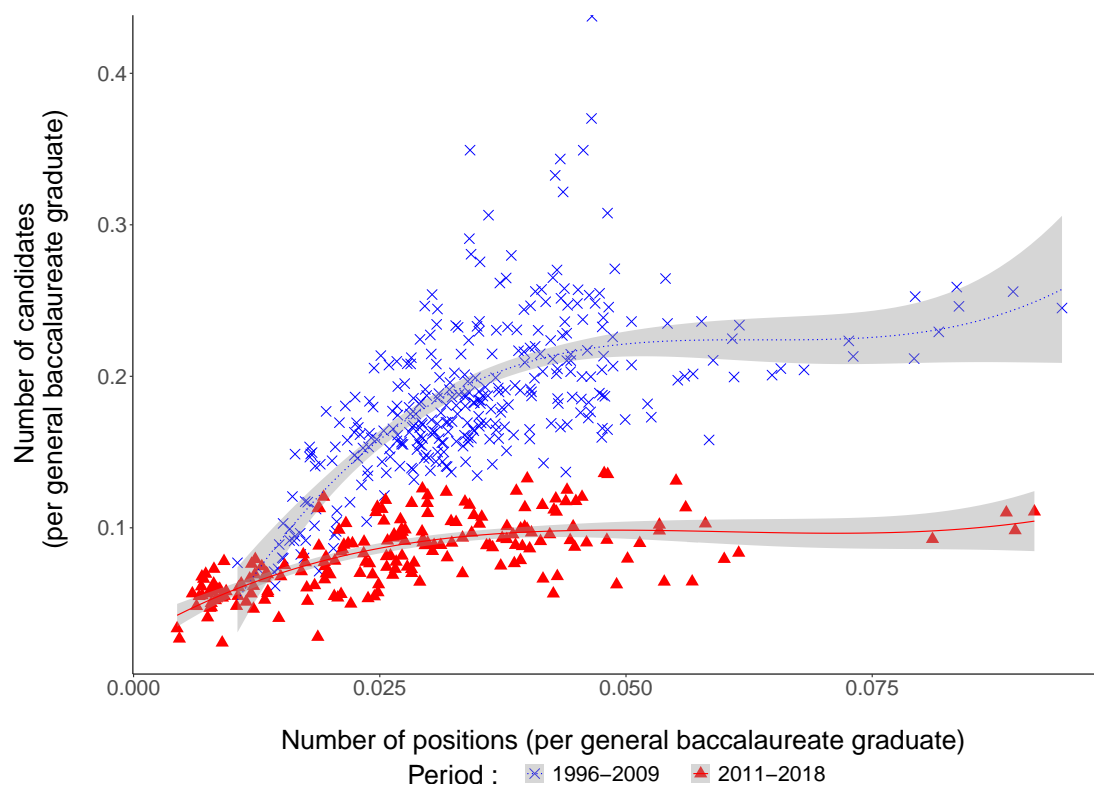


Figure 3.2 – Number of candidates according to the number of teaching posts during the periods 1996-2009 and 2011-2015

Note: Each symbol represents a region X year. I divide the number of candidates (Y -axis) and the number of positions (X -axis) by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable.

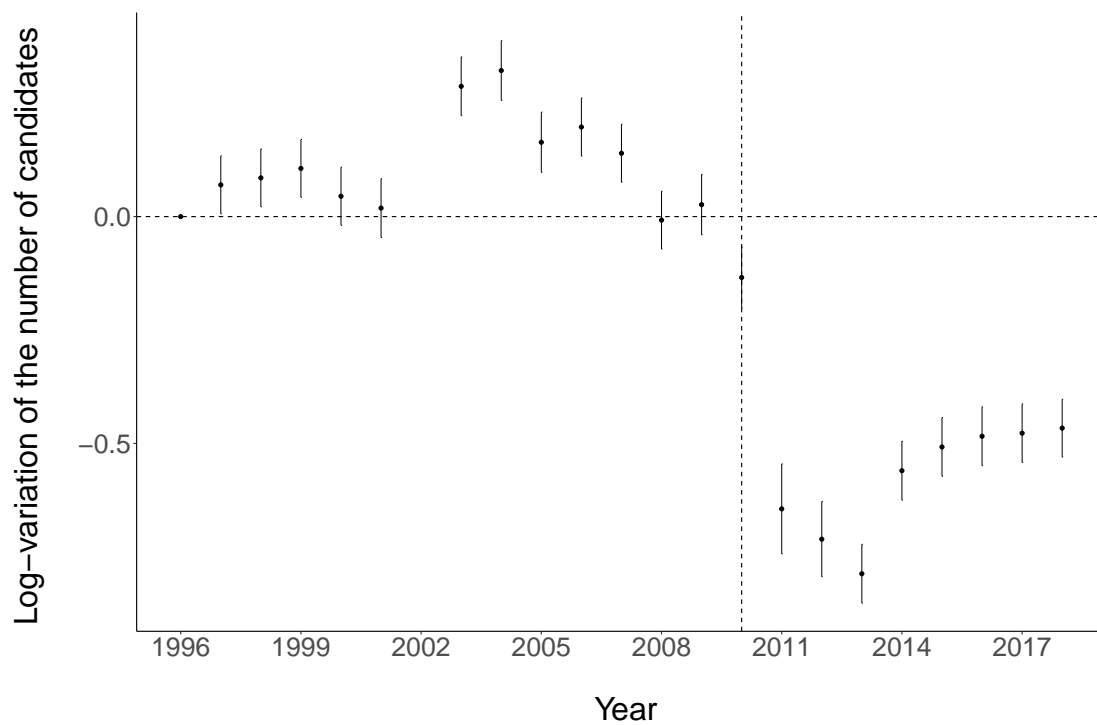


Figure 3.3 – Annual logarithmic variation in the number of candidates compared to 1996

Note: Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1') where the number of candidates and the number of positions are specified in logarithm. A decrease in the logarithm of the number of candidates of 0.7 (respectively 0.5) corresponds to a decrease in the number of candidates of 50% (respectively 40%). Column (4) of Table 3.20 reports the coefficients and standard errors.

3.7.2 Effect of a higher diploma requirement on teacher characteristics

Examination score

Figure 3.4 shows on the Y-axis the average corrected exam score of teachers³⁴ by region x year for the periods 2003-2009 (blue crosses) and 2011-2013 (red triangles). To make the regions more comparable, the X-axis is the number of positions divided by the number of general baccalaureate holders three years before the competition date. Figure 3.10 shows the result when we use raw data instead of the corrected score. Figures 3.4 and 3.10 illustrate that teachers' exam score decreases with the number of teaching positions in the region (see Section 3.5.2). The figures also suggest that, for a given number of

³⁴The corrected exam score is normalized mean zero and variance one for teachers over 2003-2009.

teaching posts, the examination score of teachers decreases significantly after the reform. Moreover, the gap between regions increases between the pre-reform and post-reform periods with the demand for teachers. These results suggest that the increase in the level of qualifications required to teach has mainly affected regions where the demand for teachers is greatest.

Table 3.22 shows the results estimated from the model (1). The first two columns correspond to the candidates' examination scores (raw and corrected) and the last two columns correspond to the teachers' examination scores (raw and corrected). The examination scores are standardized separately for candidates and teachers over the 2003-2009 period, so that the coefficients are interpreted as standard deviation points for the population considered in each column. The results of Table 3.22 are obtained by including region fixed effects and controlling the number of positions (in logarithm) and the unemployment rate in the region.

Table 3.22 shows that the corrected score (columns 2 and 4) leads to more optimistic conclusions than the raw data (columns 1 and 3). The results of columns (2) and (4) thus provide an upper limit of the impact of the 2011 reform on the knowledge of candidates and teachers.

The estimates of column (1) are presented in Figure 3.13 and show that the average score of candidates (raw data) tends to decrease over the period 2011-2013. However, this decrease can hardly be attributed to the reform as it corresponds to a downward trend since 2008. The results in column (2) are presented in Figure 3.12 and show an increase in the candidates' corrected examination score after the reform. However, the coefficients after the reform are not statistically different from those for the period 2003-2009. Columns (1) and (2) show that the 2011 reform does not appear to have significantly affected the candidates' level of knowledge.

The estimates in column (3) are presented in Figure 3.11 and show that the teachers' examination score (raw data) decreases significantly by 1.4 SD points in 2011-2013. The estimates obtained by considering the corrected score are presented in column (4) and illustrated in Figure 3.6. They confirm a sharp drop in the teacher examination score (0.55SD) after the reform. In addition, this effect is concentrated over the period 2011-2013, which makes it possible to reject the hypothesis of a downward temporal trend. A plausible interpretation of these results is that the decline in the number of candidates,

with the same average level of knowledge, forced recruiters to be less selective in filling the same number of teaching posts. As a result, the average teacher examination score decreased for a given number of positions.

To test if the teacher examination score decreased further after the reform in regions with high teacher demand ("slope difference" in Figure 3.4), I consider the following specification:

$$Y_{jt} = R_j + \alpha N_{P_{jt}} + \delta 1_{t>2010} + \mu 1_{t>2010} * N_{P_{jt}} + \gamma X_{jt} + \beta Z_{i(j)t} + \epsilon_{jt} \quad (3)$$

where $1_{t>2010}$ is a dummy variable that is 1 if $t > 2010$ and 0 otherwise. I present the estimates of the coefficients α , δ and μ in Table 3.23 for candidates (column 1 and 2) and teachers (column 3 and 4). The table reveals that the difference in slope is significant for teachers, suggesting that the higher the demand for teachers in the region, the lower the average examination score after the reform. This result suggests that increasing the level of teacher qualification contributes to increasing recruitment disparities between regions.

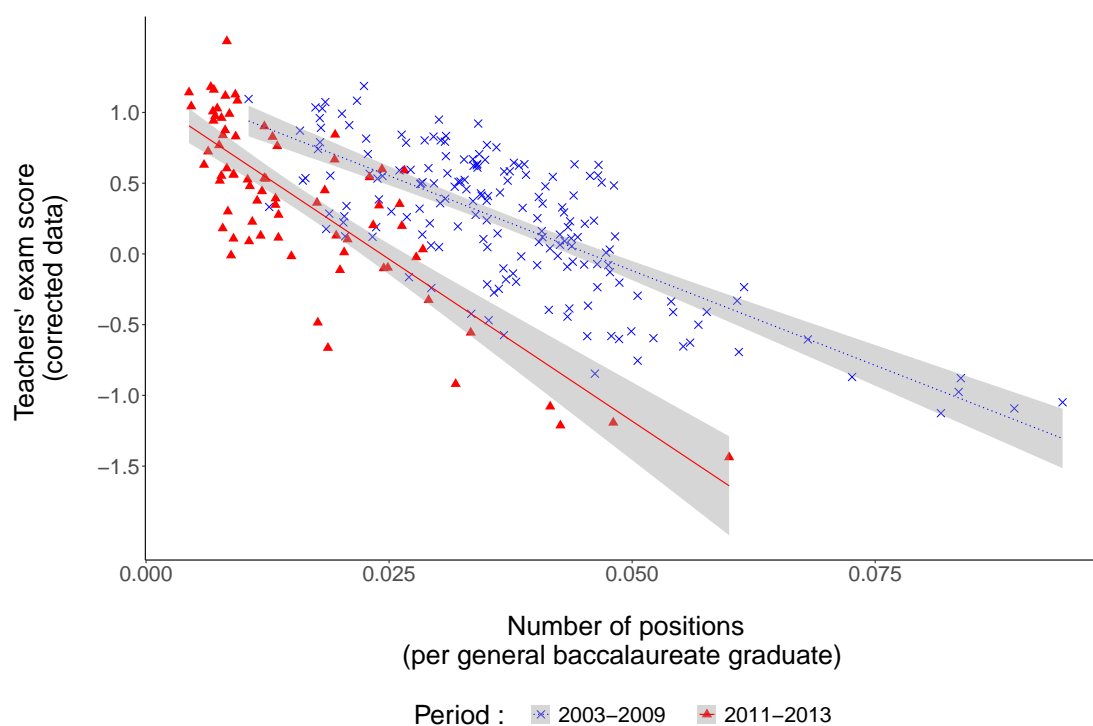


Figure 3.4 – Average score of teachers (corrected for standardization errors) on the recruitment examination during the periods 1996-2009 and 2011-2013

Note: Each symbol represents a region X year. The X -axis presents the number of positions divided by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable. The Y -axis shows the average recruitment score (corrected for assessment bias, see Appendix) of teachers for each region and year. The corrected teacher score is normalized (mean 0, variance 1) over the period 2003-2013.

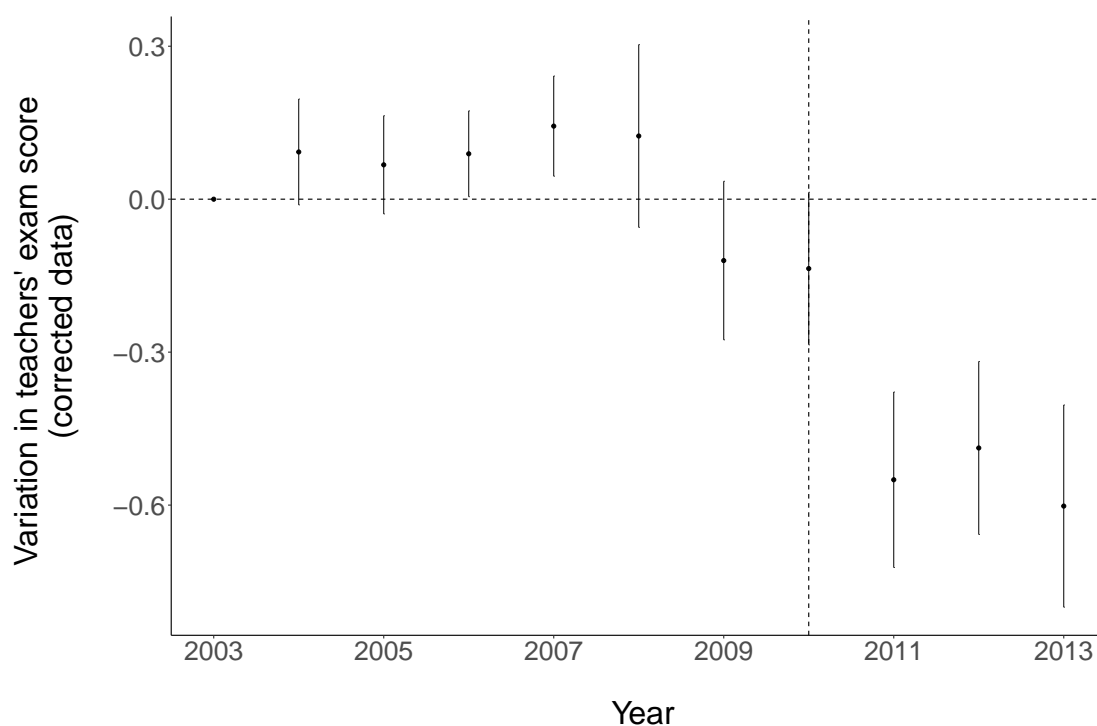


Figure 3.5 – Annual variation in the average examination score (corrected for assessment bias, see Appendix) of teachers compared to 2003

Note: The Figure shows that the examination score (corrected for assessment bias, see Appendix) of teachers decreases on average by 0.5 SD point from 2011. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of positions (in logarithm) and the unemployment rate in the region as control variables. Column (4) of Table 3.22 reports the coefficients and standard errors.

Classroom observation score

I estimate the effect of the reform on the classroom observation score after two or three years of tenure using model (2) and I present the results in Table 3.24 and Figure 3.6. Column (1) controls the number of teaching positions (in logarithm), the unemployment rate and the average salaries of bachelors' and masters' graduates in the region in the year of teacher recruitment. Column (2) includes regions fixed effects and column (3) includes school fixed effects.

The three specifications lead to different results, which underlines the importance of controlling the school's effects on assessment practices. Column (3) shows that the teachers' classroom observation score increases by 0.15 SD points after the reform. However, this increase is not statistically significant, which may be due to a lack of statistical

power.

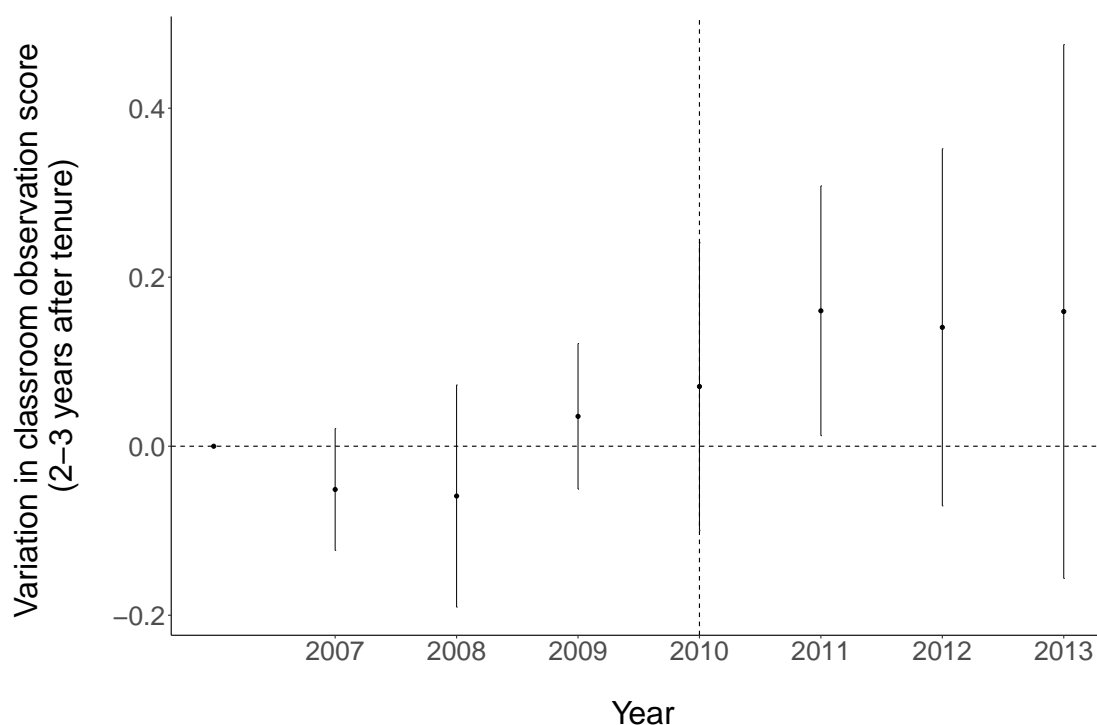


Figure 3.6 – Annual variation in the average classroom observation score (2-3 years after tenure) of teachers recruited over the period 2006-2013 (compared to 2006)

Note: Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (2) including school fixed effect, the number of positions (in logarithm), the unemployment rate and the average salaries of bachelors' and masters' graduates in the region (the year of recruitment) as control variables. Column (3) of Table 3.24 reports the coefficients and standard errors.

Teacher composition

Table 3.25 shows the impact of the reform on the gender diversity of candidates (column 1) and teachers (column 2), as well as on the median income in the candidates' (column 3) and teachers' (column 4) cities of residence. The coefficients presented in Table 3.25 are estimated by model (1), controlling for region fixed effects, the number of teaching positions, the unemployment rate, and the average wages of bachelor's and master's graduates in the region. The coefficients and 5% confidence intervals are reported on Figures 3.14 to 3.17.

Columns (1) and (2) of Table 3.25 show that the share of men decreases on average by 2.8 percentage points for applicants and 1.7 percentage points for teachers. Since men represented 13.7% of teachers before the reform, this corresponds to a relative decline of

12.4% on average. The fact that the reform contributes to reducing the proportion of men is consistent with the evidence that men with masters' degrees have better career opportunities (other than teaching) than women (see discussion in section 3.5.2). However, Figures 3.14 and 3.15 show that the share of males among applicants and teachers has a slight downward trend over time, which implies that the results should be viewed with caution.

Columns (3) and (4) indicate a decrease in the median income in the city of residence of candidates and teachers after the reform. However, like for the share of men, Figures 3.16 and 3.17 show a general downward trend and no marked break in 2011. Therefore, the effect of the reform on the social origin (assuming that our proxy measure is correct) of candidates and teachers is not clearly established.

3.8 A placebo test

The reform of the level of qualification required to teach is part of a context of falling public expenditure, which is reflected in particular by a fall in the recruitment of civil servants over the period 2008-2011 (no replacement of one civil servant out of two retired) and by a very limited increase in salaries over the period 2010-2016 (absence of any increase in the fixed part of salary - *gel du point d'indice* - as well as indexed bonuses on the fixed part of salary). This context may have discouraged some candidates for the public service and thus changed the number and composition of candidates for this type of position starting in 2008.

In order to verify whether this particular context has favored the sharp drop in the number of candidates (and the gender mix) for primary school teacher posts from 2011, we examine the evolution of the number of candidates and the gender diversity of the recruitment examination for the Regional Institutes of Administration (*Instituts Régionaux d'Administration*, IRAs). This competition leads to administrative functions (management of human, financial and material resources, management of teams and projects) in central (ministries) and regional administrations (prefectures, educational establishments, etc.) and has several points in common with the primary school teacher recruitment competition.

First, there are three entrance examinations in the IRAs: the external competition,

open to non-civil servant personnel, the internal competition, open to candidates already in post in the administration (civil servants or not) for at least 4 years, and a third competition reserved for candidates with at least 5 years of experience in the private sector as local elected officials or as heads of associations. The competition is decentralized in five IRAs located in Bastia, Lille, Lyon, Metz and Nantes. Candidates may take part in the competition only in one IRA where they will be assigned during initial training.

The IRA competition is a two-stage competition (registration in October, written tests in April and oral tests in June) and offers approximately 600 positions per year. Each year, there are approximately 6,000 candidates, 46% of whom are men. The salaries of civil servants recruited in the IRA competition are similar to those of primary school teachers. In 2018, the remuneration (excluding bonuses) is 1760 euros gross monthly in the first year for IRA officials (2067 euros gross monthly for primary school teachers) and around 3660 euros gross monthly at the end of their career (3777 euros gross monthly for primary school teachers). However, the bonuses represent about 30% of IRA civil servants' salaries and only 5% of teachers' salaries³⁵.

The IRA competition recruits at the bachelor level over the period 1996-2018. Internal competitions (40% of posts) and external competitions (55% of posts) are the main channels for access to IRAs. We have data on the five IRA recruitment competitions (internal and external) over the 2008-2017 period. We know the number of positions, the number of candidates and the number of people recruited over the 2008-2017 period. The proportion of men among candidates and recruits is also publicly available for the period 2008-2015. It is thus possible to reproduce the results of the sections 3.7.1 and 3.7.2 for the IRA competition.

Table 3.26 shows how the logarithm of the number of candidates (column 1), the share of men among candidates (column 2) and the share of men among recruits (column 3) varies each year compared to the 2008 reference year. Estimates are obtained by controlling the logarithm of the number of posts and including competition (internal, external) x IRA (Bastia, Lille, Lyon, Metz and Nantes) fixed effects. The figures 3.18 to 3.20 reproduce the coefficients and confidence intervals of interest of columns (1) to (3) of the Table 3.26.

Column (1) of Table 3.26 and Figure 3.18 show that the number of candidates in

³⁵Source: Rapport annuel sur l'état de la fonction publique, *Les rémunérations dans la fonction publique*, 2012.

IRA competitions decreases by about 5 percentage points each year over the 2009-2015 period and then increases from 2016 onwards. However, we do not see a sharp drop in the number of candidates in 2011 as is the case for primary school teachers³⁶ (see Table 3.20 and Figure 3.3). In addition, the gradual decline in the number of candidates for the IRA competition (for a given number of posts) between 2008 and 2017 is around 20%, compared to 50% over the same period for the primary school teacher recruitment examination. Column (2) and Figure 3.19 show that the proportion of men among the IRA candidates tends to decrease during the period 2009-2013, then to increase from 2014. On the other hand, column (3) and Figure 3.20 show that the proportion of men among IRA recruits does not vary significantly over the 2008-2015 period. These results invite us to moderate the role of the 2011 reform on the drop in the share of men among candidates in the teacher recruitment competition (see Table 3.22 and Figure 3.14).

A comparison of the results of the teacher recruitment examination and the IRA examination suggests that the sharp decline in the number of candidates for teaching posts from 2011 can reasonably be attributed to the reform of the level of qualifications required to teach. On the other hand, the decline in gender diversity among teachers appears to be part of a more general context of declining attractiveness of the civil service to men. The increase in the level of qualification required may have intensified the phenomenon, but the available data do not provide clear empirical evidence.

3.9 Discussion

3.9.1 Is there a link between the qualification reform and teacher shortages?

For the first time in 2013³⁷, several positions in the external primary teacher competition were not filled. Over the 2013-2017 period, the average shortage rate is 3%, but two regions account for the majority of vacant posts: Créteil and Versailles. These are the two largest regions in terms of the number of teaching posts, those with the highest

³⁶Unlike the case of teachers, the data show that there was no significant decrease in the number of posts in internal and external IRA competitions over the 2008-2017 period.

³⁷Data are not available before 1996

population growth and the lowest attractiveness³⁸. These two regions represent 27% of teaching posts but only 16% of candidates and 15% of general baccalaureate graduates (three years before the competition). In 2013-2017, 20% of the teaching posts in the Créteil region were not filled in the external recruitment competition³⁹. The Versailles region experienced a first shortage in 2013, then in 2016 (13% of posts remained vacant) and 2017 (5% of posts remained vacant).

The shortage of teachers in these regions has made it difficult to replace absent teachers. Some classes did not have a teacher at the beginning of the school year, and others had a succession of substitute teachers during the school year. This situation has led regions to recruit more short-term (non-tenure) teachers, who are most often assigned full-time to a classroom without initial training. In a report published in 2014, the rights defender described the situation experienced by some students in the Créteil region as "a violation of the constitutional principle of equality of users before the public service"⁴⁰.

To overcome this problem, the Ministry of Education set up a special competition in 2015 to fill 500 additional posts in the Créteil region. This special competition is held shortly after the external examination to allow candidates to participate in both competition. In 2016 and 2017, 500 positions were again offered in the Créteil special competition, in addition to the external competition. In 2018, the special competition was extended to the Versailles region (250 posts) and maintained in Créteil (400 posts).

To what extent is this situation related to the increase in the level of qualification required to teach in 2011? Figure 3.7.A shows how the number of positions per candidate (Y-axis) varies with the number of positions divided by the number of general baccalaureate graduates in the region (X-axis). The crosses correspond to the period 1996-2009 while the triangles correspond to the period 2011-2017. The years and academic regions where not all teaching positions were filled during the external recruitment process are indicated in red. Since almost half the candidates take the external competition twice or more times, I indicate on the Y-axis of Figure 3.7.B the (approximate) number of new

³⁸Attractiveness is measured by the ratio between the number of applications for leaving and the number of applications for entry into the region made by tenured teachers. Source: http://cache.media.education.gouv.fr/file/2013/82/4/DEPP_NI_2013_21_attractivite_academies_enseignants_second_degre_public_2012_272824.pdf

³⁹321 posts out of 1540 remained vacant in 2015, 424 posts out of 1635 in 2016 and 489 posts out of 1600 in 2017

⁴⁰Source: <https://www.defenseurdesdroits.fr/sites/default/files/atoms/files/2015-262.pdf>.

candidates each year⁴¹.

First, we note that in regions where there is a shortage of teachers, the number of candidates is nevertheless higher than the number of teaching positions (Figure 3.7.A). This is due to the fact that examiners prefer to leave vacant posts rather than recruit candidates whose level of knowledge is not considered sufficient. The emergence of teacher shortages from 2013 onwards therefore tends to support the sharp decline in the level of recruitment observed after the 2011 reform (see section 3.7.2).

Second, Figure 3.7 suggests that teacher shortages occur when there are fewer than 10 candidates for 7 positions (Figure A) or when there are more positions than new candidates (Figure B). This situation becomes much more frequent from 2011, as the number of posts per candidate increases considerably in most regions (triangles).

This study provides a plausible explanation for the emergence of shortages in primary education in some regions in France. First, the number of candidates falls sharply, but the level of knowledge does not increase significantly after the reform. Second, the number of posts increases from 2012 onwards and examiners must recruit more teachers. Teacher shortages occur when recruiters give up filling all positions to ensure a minimum level of academic knowledge among recruits. Interestingly, Figure 3.7 and Section 3.7.1 suggest that all the posts in the external competition of the Créteil and Versailles regions would have been filled if the required degree had remained at the bachelor's level (crosses).

⁴¹The applicants taking part in the competition for the first time represent approximately 48% of the candidates over the period 2005-2009 and approximately 58% of the candidates over the period 2011-2017. The data do not allow me to identify new candidates during the period 1996-2004. Therefore, I extrapolate the rate of 48% calculated over the period 2005-2009 to estimate the number of candidates who participated in the competition for the first time in 1996-2004.

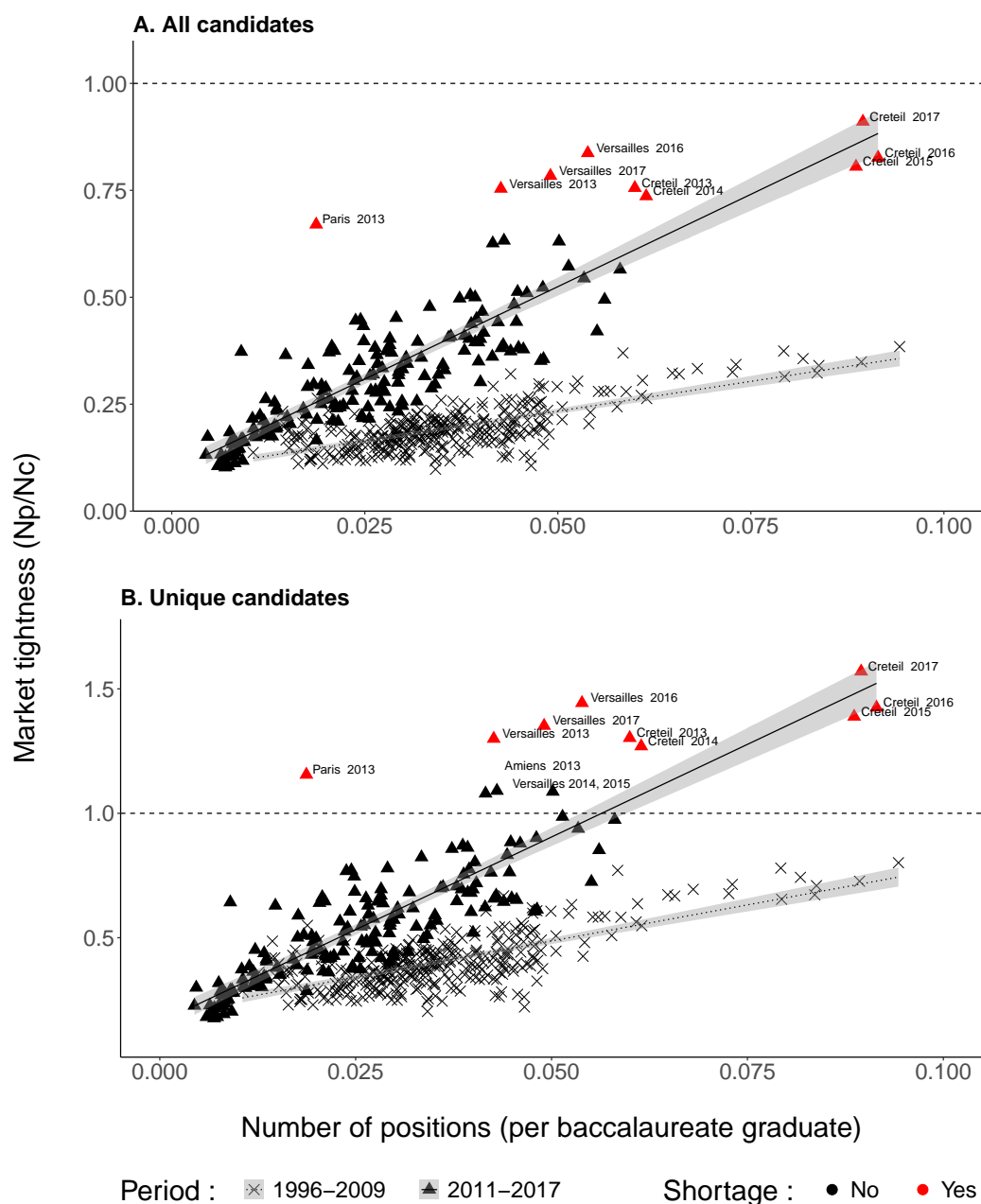


Figure 3.7 – Teacher shortage and market tightness over the period 1996-2017

Note: Each symbol represents a region X year. The X-axis presents the number of positions divided by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable. The Y-axis shows the number of teaching positions per candidate (market tightness) for each region and year over the period 1996-2017. In panel A, the market tightness is calculated on the basis of the number of candidates for external competitions. In panel B, the market tightness is computed on the basis of the number of new candidates only. The markers in red indicate the regions and years in which not all primary teacher positions were filled during external competitive recruitment (teacher shortage).

3.9.2 Short-term or long-term effects?

The reform was implemented gradually (between 2010 and 2011), providing candidates with the time needed to obtain the required degree. However, the reform may have strongly discouraged employed individuals who must return to university (and eventually give up a salary) to obtain a master's degree. The sharp fall in the number of applicants from 2011 could therefore be partly explained by the fact that professional reorientation has become more difficult in the short term.

To examine this question, Figure 3.21 and Table 3.27 show how the number of applicants varies by age group over the period 2003-2015. Individuals under 29, 30 and 34 year olds represent respectively 70%, 85% and 95% of the candidates. The evolution of the number of candidates over time follows a similar trajectory regardless of age group. Only candidates over the age of 44 seem relatively spared by the 2011 reform (see Table 3.27). The slight increase in the number of candidates aged between 18 and 23 in 2014 is due to the possibility of taking the competition during the first year of a master's degree.

Thus, the increase in the level of qualification required to become a teacher does not seem to have significantly slowed down professional retraining. This result suggests that the estimates presented in this study reflect the medium- to long-term impact of the 2011 reform.

3.10 Conclusion

This study reveals that the increase in the level of qualification required to teach has had a rather counterproductive effect in France with regard to the effectiveness of recruitment. The reform has led to a sharp decline in the number of candidates, making it more difficult to fill teaching posts, especially in regions where the demand for teachers is greatest. The characteristics of teachers have also changed following the implementation of the reform. The accreditation score decreased by 0.5 SD points on average, especially in regions with the highest number of teaching posts, and the share of men tended to decline after the reform. However, the classroom observation score, which measures teachers' skills and attitudes, has not changed significantly.

Nevertheless, the context in which the reform took place in France may have played an important role. First, French teachers' salaries are lower than the OECD average,

particularly in primary education (OECD 2014). Increasing teachers' qualifications without increasing their salaries may have exacerbated the negative effect of the reform on the number of candidates. According to our estimates, the average salary of primary school teachers is about 25% lower than that of master's graduates with identical specialties, which makes this profession particularly unattractive from a financial point of view. Second, Chapter 2 has shown the importance of considering potential candidate pools to anticipate the effect of such reform. The regions where the number of candidates has declined the most are those where the ratio between the number of positions and the number of general baccalaureate graduates is among the highest. In 2016, I estimated that there is about 1 primary school teacher post for every 3 master graduates in the Créteil region and 1 primary school teacher post for every 5 master graduates in the Versailles region. By contrast, there is about 1 primary school teacher for every 5 general baccalaureate graduates in the Créteil region and 1 primary school teacher for every 10 general baccalaureate graduates in the Versailles region⁴². As the needs for secondary school teachers are almost identical to those for primary education, there is saturation of eligible populations in these two regions, which may explain the significant drop in the level of recruitment and the emergence of teacher shortages. The additional competitions set up from 2015 to recruit teachers from other regions seem to partly solve this shortage problem. Nevertheless, it is questionable whether these teachers recruited outside their home regions will remain in the Versailles and Créteil regions for a long time, as these regions are not their first wishes for assignment.

⁴²Recent statistics from the Ministry of Higher Education reveal that on average 40% of general baccalaureate graduates obtain a bachelor's degree 3 or 4 years later and that about 50% of bachelor's graduates complete a master's degree 2 or 3 years later.

Bibliography

- Angrist, J. D. and Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5):483–503.
- Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., and Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *IDB Working Paper Series*, (665).
- Ball, D. L., Hill, H. C., and Bass, H. (2005). Knowing mathematics for teaching. *American Educator*, (Fall):14–22.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, a., Klusmann, U., Krauss, S., Neubrand, M., and Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1):133–180.
- Benhenda, A. (2017). How to Identify Good Teachers ? Teacher Evaluations and Student Achievement. *Working Paper*.
- Berger, M. C. and Toma, E. F. (1994). Variation in State Education Policies and Effects on Student Performance. *Journal of Policy Analysis and Management*, 13(3):477.
- Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2013). Measuring Test Measurement Error: A General Approach. *Journal of Educational and Behavioral Statistics*, 38(6):629–663.
- Broccolichi, S., Ben-ayed, C., Mathey-pierre, C., and Trancart, D. (2007). Fragmentations territoriales et inégalités scolaires : des relations complexes entre la distribution spatiale, les conditions de scolarisation et la réussite des élèves. *Education et Formations*, (74):31–48.
- Calvès, G. (2005). Renouveau démographique de la fonction publique de l'Etat, vers une intégration prioritaire des Français issus de l'immigration. *Rapport*.
- Campbell, P. F., Nishio, M., Smith, T. M., Clark, L. M., Conant, D. L., Rust, A. H., Neumayer DePiper, J., Jones Frank, T., Griffin, M. J., and Choi, Y. (2014). The Relationship Between Teachers' Mathematical Content and Pedagogical Knowledge, Teachers' Perceptions, and Student Achievement. *Journal for Research in Mathematics Education*, 45(4):419–459.
- Cantrell, S. and Kane, T. J. (2013). Ensuring Fair and Reliable Measures of Effective Teaching. *The Education Digest*, pages 59–63.
- Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.
- Carrington, B., Tymms, P., and Merrell, C. (2008). Role models, school improvement and the 'gender gap' - Do men bring out the best in boys and women the best in girls? *British Educational Research Journal*, 34(3):315–327.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679.
- Dee, T. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Association*, 95(2):158–165.
- Dee, T. S. (2004). Teachers, Race, and Student Achievement in a Randomized Experiment. *Review of Economics and Statistics*, 86(1):195–210.
- DEPP (2014). Géographie de l'école. *DEPP-DGESIP-DGRI, Rectorat de Bordeaux*.
- Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45:44–52.

- Ehrenberg, R. G. and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review*, 13(1):1–17.
- Feld, J., Salamanca, N., and Hamermesh, D. S. (2016). Endophilia or Exophobia: Beyond Discrimination. *Economic Journal*, 126(594):1503–1527.
- Ferguson, R. F. (1991). Paying for public education. *Harvard Journal on Legislation*, 28(2):465–498.
- Fougère, D. and Pouget, J. (2004). L’emploi public s’est-il diversifié? Sexe, niveau d’étude, origine sociale et origine nationale des salariés de la fonction publique et des collectivités territoriales. *Working Paper*.
- Gallagher, H. A. (2004). Vaughn Elementary ’s Innovative Teacher Evaluation System: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4):79–107.
- Goldhaber, D. and Anthony, E. (2007). Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching. *Review of Economics and Statistics*, 89(1):134–150.
- Grissom, J. A., Nicholson-Crotty, J., and Nicholson-Crotty, S. (2009). Race, region, and representative bureaucracy. *Public Administration Review*, 69(5):911–919.
- Hanushek, E. A. and Rivkin, S. G. S. G. (2004). How to Improve the Supply of High-Quality Teachers. *Brookings Papers on Education Policy*, 2004(1):7–25.
- Hildebrandt, S. A. and Eom, M. (2011). Teacher professionalization: Motivational factors and the influence of age. *Teaching and Teacher Education*, 27(2):416–423.
- Hill, H., Rowan, B., and Ball, D. (2005). Effects of teachers’ mathematical knowledge for teaching on student achievement. *American Education Research Journal*, 42(2):371–406.
- Hill, H. C., Schilling, S. G., and Ball, D. L. (2004). Developing Measures of Teachers’ Mathematics Knowledge for Teaching. *The Elementary School Journal*, 105(1):11–30.
- Jacob, B., Rockoff, J., Taylor, E., Lindy, B., and Rosen, R. (2016). Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools. *NBER Working Paper series*.
- Kersting, N. B., Chen, M.-k., and Stigler, J. W. (2013). Value-added Teacher Estimates as Part of Teacher Evaluations: Exploring the Effects of Data and Model Specifications on the Stability of Teacher Value-added Scores. *Education Policy Analysis Archives*, pages 1–39.
- Kimball, S. M., White, B., Milanowski, A. T., and Borman, G. (2004). Examining the Relationship Between Teacher Evaluation and Student Assessment Results in Washoe County. *Peabody Journal of Education*, 79(4):54–78.
- Lapostolle, G. (2013). La création des Écoles supérieures du professorat et de l’éducation s’inscrit-elle dans un modèle supranational de professionnalisation des enseignants ? *Tréma*, (40):60–75.
- L’Horty, Y. (2016). Les discriminations dans l’accès à l’emploi public: Rapport au Premier Ministre. *La documentation française*.
- Meier, K. J. (1984). Teachers, Students, and Discrimination: The Policy Impact of Black Representation. *The Journal of Politics*, 46(1):252–263.
- Meier, K. J. (1993). Latinos and Representative Bureaucracy Testing the Thompson and Henderson Hypotheses. *Journal of Public Administration Research and Theory*, 3(4):393–414.
- Meurs, D., Pailhé, A., and Simon, P. (2006). Persistance des inégalités entre générations liées à l’immigration : l’accès à l’emploi des immigrés et de leurs descendants en France.

- Population*, 61(5):763.
- Milanowski, A. (2004). The Relationship Between Teacher Performance Evaluation Scores and Student Achievement: Evidence From Cincinnati. *Peabody Journal of Education*, 79(4):4–32.
- Mullens, J. E., Murnane, R. J., and Willet, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 40(2):139–157.
- Paredes, V. (2014). A teacher like me or a student like me? Role model versus teacher bias effect. *Economics of Education Review*, 39:38–49.
- Porter, C., Serra, D., Croson, R., Desmet, K., Lindo, J., Meer, J., Millimet, D., Ozerturk, S., Kuka, E., Salmon, T., and Schechter, L. (2017). Gender Differences in the Choice of Major: The Importance of Female Role Models. *Working Paper*.
- Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., and Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1):43–74.
- Rockoff, J. E. and Speroni, C. (2011). Subjective and objective evaluations of teacher effectiveness: Evidence from New York City. *Labour Economics*, 18(5):687–696.
- Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *The Quarterly Journal of Economics*, 125(1):175–214.
- Rothstein, J. (2015). Teacher Quality Policy When Supply Matters. *American Economic Review*, 105(1):100–130.
- Rowan, B., Chiang, F.-S., and Miller, R. J. (1997). Using Research on Employees' Performance to Study the Effects of Teachers on Students' Achievement. *Sociology of Education*, 70(4):256.
- Versini, D. (2004). Rapport sur la diversité dans la fonction publique. *Rapport*.

Tables

Table 3.9 – Share of local recruitment by academic region

Region of application	Share of candidates living in the region where they apply (by region of application)	Region of residence	Share of candidates applying in the region where they live (by region of residence)
All	84.7 %	All	82.1 %
Paris	56.5%	Paris	50.3%
Creteil	58.9%	Clermont Ferrand	73.3%
Versailles	67.9%	Limoges	74.4%
Guyane	70.8%	Rennes	76.6%
Lyon	77.1%	Montpellier	77.0%
Clermont Ferrand	77.4%	Versailles	79.9%
Nice	78.7%	Caen	80.2%
Poitiers	81.2%	Toulouse	80.4%
Orleans Tours	81.3%	Dijon	80.7%
Limoges	83.0%	Creteil	81.8%
Grenoble	83.1%	Nice	83.0%
Amiens	83.1%	Grenoble	83.0%
Dijon	83.6%	Orleans Tours	83.8%
Caen	84.5%	Reims	84.0%
Aix Marseille	84.7%	Amiens	84.2%
Nantes	86.2%	Lyon	84.8%
Reims	88.0%	Poitiers	85.2%
Montpellier	90.4%	Bordeaux	85.2%
Toulouse	91.1%	Nantes	85.7%
Bordeaux	91.3%	Aix Marseille	86.6%
Rennes	92.7%	Besancon	86.9%
Rouen	93.3%	Corse	87.0%
Besancon	93.6%	Rouen	87.8%
Strasbourg	93.9%	Nancy Metz	90.5%
Guadeloupe	96.2%	Strasbourg	91.6%
La Reunion	96.9%	Martinique	91.8%
Lille	97.2%	Lille	93.9%
Martinique	97.2%	Guadeloupe	96.5%
Nancy Metz	97.2%	La Reunion	99.0%
Corse	97.4%	Guyane	99.2%

Notes: The "share of candidates living in the region where they apply" is obtained for each academic region by dividing the number of candidates who live in the region by the total number of candidates who apply in that region. The "share of candidates applying in the region where they live" is obtained for each academic region by dividing the number of candidates who live and apply in the region by the number of candidates living in the region (but may apply to another region). For example, 56.5% of candidates applying for the Paris region competition live in the Paris region, while the remaining 45.5% live in another region. 50.3% of candidates residing in the Paris region apply in this region, and the remaining 49.7% apply in another region. These two measures indicate the extent to which primary teacher recruitment is local.

Source: Ministry of Education (DEPP).

Table 3.10 – Number of general baccalaureate graduates, candidates and teaching positions by academic region and time interval

	Number of positions				Number of candidates				Number of general baccalaureate graduate t-3			
	1996-2000	2001-2005	2006-2010	2011-2015	1996-2000	2001-2005	2006-2010	2011-2015	1996-2000	2001-2005	2006-2010	2011-2015
All	8,732	10,840	9,367	6,547	43,907	55,036	47,322	21,988	269,511	256,861	259,590	274,018
Limoges	85	102	89	60	460	594	506	227	3067	2826	2633	2685
Besancon	149	224	161	120	910	1158	998	483	5287	4922	4800	4795
Caen	174	250	177	114	1069	1553	1090	500	6415	6066	5967	6170
Clermont-Ferrand	177	184	158	101	1116	1451	929	430	6178	5612	5344	5143
Dijon	191	257	250	144	1151	1192	1093	479	7212	6771	6581	6584
Reims	202	239	210	135	866	932	975	458	6242	5623	5508	5450
Paris	203	236	232	187	988	1237	1041	517	13217	12259	12366	13541
Poitiers	241	313	234	147	1148	1921	1342	608	7410	6621	6474	6626
Rennes	252	273	228	148	1506	1971	1548	779	15616	14001	13892	14475
Strasbourg	253	324	284	153	1324	1410	1352	626	6874	6977	7250	8062
Rouen	264	276	286	186	1296	1340	1301	616	7570	7524	7520	7982
Montpellier	294	446	295	208	2021	3131	2149	1010	9528	9341	9731	10423
Nice	295	325	255	151	1713	1885	1518	639	7210	7635	8299	9054
Nantes	298	431	404	268	1606	2670	2271	1124	17038	15453	15111	15698
Toulouse	315	388	305	206	2398	2982	2280	1014	11803	10838	10881	11871
Amiens	344	375	367	213	1447	1410	1492	625	7930	7622	7274	7370
Bordeaux	346	471	398	247	2230	2853	2600	1226	12885	12071	12204	13256
Orleans-Tours	355	424	362	240	1700	2389	1758	793	11069	10091	10104	10563
Grenoble	364	432	357	350	2070	3015	2115	1147	14054	13797	14159	15047
Aix-Marseille	388	479	336	272	2336	3026	2388	1061	11697	11236	11936	12534
Nancy-Metz	406	471	350	161	2047	2005	1930	835	10644	9877	9822	9850
Lyon	425	497	430	359	2095	3164	2351	1390	13653	13471	13382	14152
Lille	596	840	766	403	2900	3891	4165	1606	18485	18250	17849	17238
Creteil	1000	1187	1161	978	3163	3336	3636	1733	14640	14596	15571	17296
Versailles	1113	1397	1271	998	4344	4520	4493	2064	23786	23379	24931	28157

Notes: Number of general baccalaureate graduates three years before the exam date, number of positions and number of candidates for the external primary school teacher recruitment competition over the periods 1996-2000 2001-2005, 2006-2010 and 2011-2015. The data are provided at national level and by academic region of metropolitan France (excluding Corsica). The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. The number of candidates increased significantly (by 25%) over the period 2001-2005 because applicants were exceptionally allowed to write the exam in more than one region in the same year in 2003 and 2004 (see Appendix).

Source: Ministry of Education (DEPP). *Scope:* Metropolitan France.

Table 3.11 – Teacher market tightness, teacher demand and teacher supply by academic region and time interval

	100 * Positions / Candidates				100 * Positions / General baccalaureate graduates t-3				100 * Candidates / General baccalaureate graduates t-3			
	1996-2000	2001-2005	2006-2010	2011-2015	1996-2000	2001-2005	2006-2010	2011-2015	1996-2000	2001-2005	2006-2010	2011-2015
Total	19.9	20	19.7	30	3.3	4.2	3.6	2.4	16.6	21	18.3	8
Toulouse	13.1	13.1	13.4	20.3	2.7	3.6	2.8	1.7	20.6	27.5	20.9	8.4
Montpellier	14.4	14.6	13.6	20.4	3.1	4.8	3.1	2	21.5	32.9	22.8	9.8
Bordeaux	15.7	16.8	15.2	19.7	2.7	3.9	3.3	1.9	17.2	23.2	21.7	9.6
Clermont-Ferrand	16	13.6	16.9	22.9	2.9	3.3	3	2	18.1	24.3	17.8	8.7
Caen	16.4	17.3	16.1	22	2.7	4.1	3	1.8	16.5	23.7	18.6	8.2
Dijon	16.6	21.6	23.3	30.9	2.7	3.8	3.8	2.2	16.3	17.6	16.3	7.1
Aix-Marseille	16.6	16.2	13.8	26	3.3	4.3	2.8	2.2	19.9	26.5	20.3	8.5
Besancon	16.7	19.5	16.1	24.7	2.8	4.6	3.4	2.5	16.8	23.6	21.1	10.1
Rennes	16.9	14.3	14.8	18.8	1.6	2	1.7	1	9.5	14	11.5	5.3
Nice	17.3	17.4	16.6	23.8	4.1	4.3	3.1	1.7	23.7	24.7	18.7	7.1
Grenoble	17.6	14.8	16.8	29.7	2.6	3.1	2.6	2.3	14.8	20.9	15.5	7.7
Nantes	18.6	16.5	17.7	23.3	1.8	2.8	2.7	1.7	9.7	17	15.3	7.3
Limoges	18.7	17.7	17.4	25.1	2.8	3.6	3.4	2.2	15	20.3	19.5	8.8
Strasbourg	19.1	23	20.8	23.5	3.7	4.6	4	1.9	19.4	20	19.2	8.1
Nancy-Metz	19.8	23.6	18.1	19.2	3.8	4.8	3.6	1.6	19.2	20.3	19.9	8.3
Lyon	20.3	16.7	18.2	25.5	3.1	3.7	3.2	2.5	15.3	22.2	17.6	9.8
Rouen	20.5	20.8	22	29.9	3.5	3.7	3.8	2.3	17.1	17.8	17.3	7.7
Lille	20.6	21.9	18.2	25.2	3.2	4.6	4.3	2.3	15.5	21	23.6	9.1
Paris	20.6	19.3	22.3	37.2	1.5	1.9	1.9	1.4	7.3	9.8	8.5	3.8
Orleans-Tours	20.9	19.4	20.5	31.3	3.2	4.2	3.6	2.3	15.3	21.6	17.6	7.3
Poitiers	21	18.8	17.3	24.4	3.3	4.7	3.6	2.2	15.7	25	20.8	9
Reims	23.4	25.8	21.6	29.6	3.3	4.3	3.8	2.5	14.1	16.7	17.6	8.4
Amiens	24	26.8	24.7	35.7	4.4	4.9	5.1	2.9	18.3	18.3	20.6	8.1
Versailles	25.9	31.5	28.1	51.1	4.7	6	5.2	3.5	18.1	19	18.5	6.8
Creteil	31.6	35.7	31.9	58.9	6.9	8.1	7.6	5.6	21.8	22.7	23.8	9.5

Notes: Number of general baccalaureate graduates three years before the exam date, number of positions and number of candidates for the external primary school teacher recruitment competition over the periods 1996-2000 2001-2005, 2006-2010 and 2011-2015. The number of baccalaureate holders 3 years before the recruitment examination is proportional to the size of the population eligible for teaching positions. Teacher market tightness is defined as the number of teaching positions per candidates, teacher demand is proportional to the number of positions per general baccalaureate graduates three years before the exam date and teacher supply is proportional to the number of candidates per general baccalaureate graduates three years before the exam date. The data are provided at national level and by academic region of metropolitan France (excluding Corsica). The number of candidates increased significantly (by 25%) over the period 2001-2005 because applicants were exceptionally allowed to write the exam in more than one region in the same year in 2003 and 2004 (see Appendix).

Source: Ministry of Education (DEPP). *Scope:* Metropolitan France.

Table 3.12 – Raw and corrected exam scores of candidates by academic region and time interval

	(1)	(2)	(3)	(4)	(5)	(6)
	Raw score			Corrected score		
	2003-2006	2007-2010	2011-2013	2003-2006	2007-2010	2010-2013
All	-0.025	0.035	-0.21	-0.007	0.028	0.203
Paris	-0.092	0.045	-0.126	-0.104	0.03	0.268
Aix-Marseille	0.017	-0.083	-0.258	0.041	-0.066	0.169
Besançon	0.008	-0.014	-0.398	0.165	0.141	0.166
Bordeaux	-0.059	0.115	-0.213	0.109	0.273	0.366
Caen	0.055	0.232	-0.007	-0.071	0.101	0.278
Clermont-Ferrand	-0.181	-0.035	-0.347	0.167	0.31	0.414
Dijon	0.032	0.088	-0.016	0.02	0.072	0.373
Grenoble	-0.018	0.006	-0.003	0.18	0.203	0.609
Lille	-0.286	-0.136	-0.369	-0.211	-0.071	0.112
Lyon	-0.354	-0.044	-0.254	0.175	0.485	0.683
Montpellier	-0.002	0.076	-0.282	0.12	0.194	0.249
Nancy-Metz	-0.02	-0.088	-0.242	-0.048	-0.121	0.13
Poitiers	0.213	0.241	-0.318	0.193	0.227	0.083
Rennes	0.072	0.293	-0.024	0.09	0.305	0.405
Strasbourg	0.048	0.204	-0.001	-0.064	0.083	0.292
Toulouse	-0.236	-0.01	-0.139	0.025	0.244	0.53
Nantes	0.175	0.288	-0.124	0.08	0.185	0.193
Orleans-Tours	0.324	0.18	-0.182	0.093	-0.054	-0.006
Reims	0.035	0.035	-0.244	-0.1	-0.107	0.026
Amiens	0.026	0.124	-0.208	-0.359	-0.269	-0.191
Rouen	0.012	0.079	-0.101	-0.159	-0.101	0.133
Limoges	-0.327	-0.11	-0.439	0.05	0.261	0.346
Nice	-0.113	0.095	-0.105	0.036	0.237	0.452
Créteil	0	-0.066	-0.401	-0.331	-0.405	-0.33
Versailles	0.114	-0.042	-0.199	-0.09	-0.253	0.001

Notes: Columns (1) to (3) indicate the raw score of candidates for the 2003-2013 recruitment examinations. Columns (4) to (6) indicate the examination score of candidates corrected for assessment bias using the method presented in the Appendix. The data are provided at national level and by academic region of metropolitan France (excluding Corsica). The candidates' raw and corrected scores are normalized (mean 0 and variance 1) over the period 2003-2013.

Source: Ministry of Education (DEPP). Author's calculations.

Table 3.13 – Raw and corrected exam scores of teachers by academic region and time interval

	(1)	(2)	(3)	(4)	(5)	(6)
	Raw score			Corrected score		
	2003-2006	2007-2010	2011-2013	2003-2006	2007-2010	2010-2013
All	-0.041	0.05	-0.942	-0.013	0.042	-0.141
Paris	0.144	0.291	-0.842	0.192	0.309	0.023
Aix-Marseille	0.08	0.154	-0.932	0.187	0.241	-0.003
Besancon	0.146	-0.017	-0.822	0.471	0.321	0.319
Bordeaux	0.175	0.389	-0.428	0.518	0.697	0.685
Caen	0.01	0.454	-0.367	-0.127	0.268	0.251
Clermont-Ferrand	0.032	0.226	-0.667	0.695	0.857	0.798
Dijon	0.073	0.001	-0.285	0.121	0.037	0.508
Grenoble	0.281	0.11	-0.376	0.671	0.51	0.793
Lille	-0.493	-0.194	-1.158	-0.242	0.011	-0.122
Lyon	-0.459	-0.033	-0.298	0.564	0.93	1.407
Montpellier	0.277	0.551	-0.361	0.532	0.771	0.678
Nancy-Metz	0.133	-0.056	-0.252	0.142	-0.037	0.5
Poitiers	0.484	0.654	-0.926	0.497	0.649	-0.054
Rennes	0.957	1.133	0.122	0.971	1.118	0.928
Strasbourg	-0.112	0.199	-0.144	-0.215	0.05	0.482
Toulouse	0.026	0.439	-0.096	0.539	0.901	1.149
Nantes	0.386	0.682	-0.478	0.273	0.525	0.199
Orleans-Tours	0.251	0.225	-0.62	-0.082	-0.115	-0.15
Reims	-0.328	-0.295	-0.678	-0.448	-0.43	-0.07
Amiens	-0.084	0.033	-1.284	-0.649	-0.564	-0.993
Rouen	0.005	0.174	-0.526	-0.206	-0.077	0.028
Limoges	-0.449	0.052	-0.689	0.312	0.75	0.803
Nice	-0.076	0.34	-0.428	0.262	0.625	0.664
Creteil	-0.44	-0.416	-1.858	-0.875	-0.866	-1.463
Versailles	-0.139	-0.376	-1.604	-0.391	-0.621	-1.006

Notes: Columns (1) to (3) indicate the raw score of teachers for the 2003-2013 recruitment examinations. Columns (4) to (6) indicate the examination score of teachers corrected for assessment bias using the method presented in the Appendix. The data are provided at national level and by academic region of metropolitan France (excluding Corsica). The teachers' raw and corrected scores are normalized (mean 0 and variance 1) over the period 2003-2013.

Source: Ministry of Education (DEPP). Author's calculations.

Table 3.14 – Classroom observation score of teachers 2-3 after tenure by academic region and time interval

	(1)	(2)
	Classroom observation score after 2-3 years of experience	
	2006-2009	2010-2013
All	-0.009	0.023
Paris	0.333	0.622
Aix-Marseille	-0.144	-0.128
Besançon	0.553	0.47
Bordeaux	-0.016	0.307
Caen	-0.177	-0.112
Clermont-Ferrand	-0.117	-0.309
Dijon	-0.253	-0.177
Grenoble	-0.419	-0.361
Lille	0.087	-0.128
Lyon	-0.024	0.011
Montpellier	-0.246	-0.14
Nancy-Metz	0.916	0.82
Poitiers	-0.209	-0.124
Rennes	0.127	-0.001
Strasbourg	-0.523	-0.648
Toulouse	-0.026	0.126
Nantes	-0.203	-0.192
Orleans-Tours	-0.467	-0.48
Reims	-0.25	-0.15
Amiens	-0.686	-0.742
Rouen	-0.467	-0.41
Limoges	0.104	-0.304
Nice	0.191	0.073
Créteil	-0.085	0.139
Versailles	0.34	0.424

Notes: This Table indicates the average classroom observation score after 2-3 years of experience for teachers recruited over the periods 2006-2009 (column 1) and 2010-2013 (column 2). The data are provided at national level and by academic region of metropolitan France (excluding Corsica). The classroom observation scores 2-3 years after tenure are normalized (mean 0 and variance 1) over the period 2006-2013.

Source: Ministry of Education (DEPP). Author's calculations.

Table 3.15 – Correlation between diploma level and gender of candidates and teachers over the period 2003-2009

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability that the candidate is male			Probability that the teacher is male		
No diploma condition	-0.140*** (0.00337)	-0.140*** (0.00337)	-0.140*** (0.00336)	-0.0979*** (0.0104)	-0.100*** (0.0104)	-0.101*** (0.0104)
3yrs of college	-	-	-	-	-	-
4yrs of college	0.0212*** (0.00176)	0.0214*** (0.00176)	0.0233*** (0.00176)	0.0328*** (0.00380)	0.0306*** (0.00380)	0.0326*** (0.00382)
5yrs of college and more	0.0400*** (0.00272)	0.0394*** (0.00272)	0.0436*** (0.00273)	0.0314*** (0.00503)	0.0333*** (0.00503)	0.0389*** (0.00506)
Control variables:						
Unemployment rate	Yes	Yes	Yes	Yes	Yes	Yes
Number of attempts	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes
Nb of positions	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	345957	345957	345957	68476	68476	68476

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are clustered at the region level.

Notes: This Table explains the gender of candidates (columns 1 to 3) and teachers (columns 4 to 6) by diploma level over the period 2003-2009. Column (3) and (6) include region fixed effects, the number of times candidates take the exam, the unemployment rate and the number of teaching posts in the region X year as control variables.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.16 – Correlation between diploma level and median income in the city of residence of candidates and teachers over the period 2003-2009

	(1)	(2)	(3)	(4)	(5)	(6)
	Median income in the city of candidates			Median income in the city of teachers		
No diploma condition	637.6*** (32.09)	632.0*** (32.07)	561.7*** (29.55)	1033.9*** (111.4)	1070.4*** (111.3)	926.4*** (101.0)
3yrs of college (reference)	-	-	-	-	-	-
4yrs of college	206.6*** (16.72)	215.9*** (16.71)	119.7*** (15.47)	335.1*** (40.68)	368.4*** (40.68)	134.4*** (37.12)
5yrs of college and more	891.6*** (25.84)	859.9*** (25.86)	652.7*** (23.92)	1111.6*** (53.83)	1083.3*** (53.79)	789.3*** (49.16)
Control variables:						
Unemployment rate	Yes	Yes	Yes	Yes	Yes	Yes
Number of attempts	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes
Nb of positions	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	339789	339789	339789	67380	67380	67380

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are clustered at the region level.

Notes: This Table explains the median income in the city of residence of candidates (columns 1 to 3) and teachers (columns 4 to 6) by diploma level over the period 2003-2009. To neutralize the effect of income growth over time, I consider the median income in euros per consumption unit (i.e. divided by the number of inhabitants in the household) in cities in 2012. Column (3) and (6) include region fixed effects, the number of times candidates take the exam, the unemployment rate and the number of teaching posts in the region X year as control variables.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.17 – Correlation between diploma level and corrected examination score of candidates and teachers over the period 2003-2009

	(1)	(2)	(3)	(4)	(5)	(6)
	Candidates' exam score			Teachers' exam score		
No diploma condition	-0.537*** (0.00871)	-0.533*** (0.00868)	-0.533*** (0.00854)	-0.0786*** (0.0252)	-0.0695*** (0.0252)	-0.0762*** (0.0216)
3yrs of college	-	-	-	-	-	-
4yrs of college	-0.0271*** (0.00454)	-0.0308*** (0.00453)	-0.0538*** (0.00448)	-0.00508 (0.0103)	0.000810 (0.0103)	-0.0288*** (0.00882)
5yrs of college and more	0.195*** (0.00703)	0.209*** (0.00702)	0.177*** (0.00693)	0.310*** (0.0137)	0.305*** (0.0137)	0.189*** (0.0117)
Control variables:						
Unemployment rate	Yes	Yes	Yes	Yes	Yes	Yes
Number of attempts	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes
Nb of positions	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	345957	345957	345957	68476	68476	68476

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Notes: This Table explains the exam score (corrected for assessment bias, see appendix) of candidates (columns 1 to 3) and teachers (columns 4 to 6) by diploma level over the period 2003-2009. Column (3) and (6) include region fixed effects, the number of times candidates take the exam, the unemployment rate and the number of teaching posts in the region X year as control variables.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.18 – Correlation between diploma level and raw examination score of candidates and teachers over the period 2003-2009

	(1)	(2)	(3)	(4)	(5)	(6)
	Candidates' exam score			Teachers' exam score		
No diploma condition	-0.538*** (0.00842)	-0.534*** (0.00840)	-0.536*** (0.00857)	-0.104*** (0.0246)	-0.0917*** (0.0246)	-0.0855*** (0.0239)
3yrs of college (reference)	-	-	-	-	-	-
4yrs of college	-0.0454*** (0.00440)	-0.0488*** (0.00439)	-0.0544*** (0.00449)	-0.0428*** (0.0100)	-0.0342*** (0.0100)	-0.0331*** (0.00975)
5yrs of college and more	0.173*** (0.00688)	0.186*** (0.00688)	0.176*** (0.00695)	0.256*** (0.0134)	0.248*** (0.0134)	0.204*** (0.0129)
Control variables:						
Unemployment rate	Yes	Yes	Yes	Yes	Yes	Yes
Number of attempts	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes
Nb of positions	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	365976	365976	345957	71662	71662	68476

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Notes: This Table explains the raw exam score of candidates (columns 1 to 3) and teachers (columns 4 to 6) by diploma level over the period 2003-2009. Column (3) and (6) include region fixed effects, the number of times candidates take the exam, the unemployment rate and the number of teaching posts in the region X year as control variables.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.19 – Correlation between diploma level and classroom observation score of teachers over the period 2006-2009

	(1)	(2)	(3)	(4)	(5)
	Classroom observation score				
No diploma condition	-0.0199 (0.0551)	-0.0289 (0.0554)	-0.0198 (0.0520)	0.00614 (0.0495)	-0.0601 (0.0736)
3yrs of college (reference)	-	-	-	-	-
4yrs of college	0.0527*** (0.0200)	0.0421** (0.0200)	0.0540*** (0.0188)	0.0250 (0.0178)	0.0276 (0.0246)
5yrs of college and more	-0.233 (0.183)	-0.222 (0.182)	-0.134 (0.171)	0.0243 (0.169)	-0.0731 (0.213)
Control variables:					
Unemployment rate	Yes	Yes	Yes	Yes	Yes
Nb of positions	No	Yes	Yes	Yes	Yes
Region FE	No	No	Yes		
Teaching inspector FE	No	No		Yes	
School FE	No	No			Yes
Constant	Yes	Yes	Yes	Yes	Yes
Observations	19052	18927	18927	15558	18753

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Notes: This Table explains the classroom observation score of teachers 2-3 years after tenure (normalized mean 0, variance 1) by diploma level over the period 2006-2009. Column (2) includes the unemployment rate and the number of teaching posts in the region X year of recruitment as control variables. Column (3) includes region fixed effects, column (4) includes teaching inspector area fixed effects and column (5) includes school FE.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.20 – Effect of increasing the degree level required to teach on the number of candidates

	(1)	(2)	(3)	(4)
Year effects (ref:1996)	Number of candidates			Logarithm of number of candidates
1997	149.7 (138.9)	124.2 (94.06)	138.7 (91.35)	0.0710** (0.0322)
1998	133.8 (138.9)	124.2 (94.00)	129.8 (91.26)	0.0858*** (0.0321)
1999	139.6 (138.9)	173.3* (94.11)	153.7* (91.43)	0.105*** (0.0322)
2000	22.20 (138.9)	78.77 (94.33)	40.65 (91.83)	0.0427 (0.0323)
2001	-56.82 (139.0)	47.55 (95.13)	-20.85 (93.18)	0.0148 (0.0327)
2003	510.4*** (139.0)	620.1*** (95.24)	553.9*** (93.24)	0.283*** (0.0327)
2004	552.5*** (139.2)	710.4*** (96.57)	618.1*** (95.24)	0.316*** (0.0334)
2005	172.0 (139.3)	353.2*** (97.37)	261.6*** (95.98)	0.158*** (0.0337)
2006	308.0** (138.9)	380.1*** (94.53)	354.9*** (91.89)	0.195*** (0.0323)
2007	219.6 (138.9)	289.4*** (94.50)	267.4*** (91.83)	0.138*** (0.0323)
2008	-5.668 (138.9)	16.14 (94.04)	12.44 (91.29)	-0.00819 (0.0321)
2009	135.2 (139.1)	15.55 (95.48)	118.3 (94.55)	0.0328 (0.0337)
2010	-108.8 (139.1)	-232.5** (95.58)	-124.7 (94.83)	-0.128*** (0.0338)
2011	-490.7*** (140.1)	-800.4*** (103.6)	-490.5*** (115.2)	-0.619*** (0.0517)
2012	-600.4*** (139.5)	-821.2*** (98.97)	-592.6*** (104.7)	-0.694*** (0.0427)
2013	-843.7*** (138.9)	-893.1*** (94.25)	-827.0*** (92.28)	-0.783*** (0.0328)
2014	-650.2*** (138.9)	-702.6*** (94.28)	-634.3*** (92.36)	-0.556*** (0.0328)
2015	-795.3*** (139.1)	-675.7*** (95.48)	-716.0*** (92.98)	-0.511*** (0.0326)
2016	-836.5*** (139.2)	-680.4*** (96.51)	-721.4*** (93.99)	-0.488*** (0.0327)
2017	-801.5*** (139.1)	-668.8*** (95.82)	-704.8*** (93.25)	-0.481*** (0.0326)
Control variables				
Number of positions	Yes	Yes	Yes	No
(Number of positions) ²	No	No	Yes	No
Logarithm of number of positions	No	No	No	Yes
Region FE	No	Yes	Yes	Yes
Unemployment rate	No	No	Yes	Yes
Observations	525	525	525	525
Adjusted R-squared	0.74	0.88	0.89	0.97

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parenthesis.

Notes: The variable explained is the number of candidates for the external primary teacher recruitment competition by region and year. Year fixed effects are estimated using model (1) in columns (1) to (3) and model (1') in column (4). The number of candidates increases significantly in 2003 and 2004 because applicants were exceptionally allowed to write the exam in two regions the same year. In 2010, new candidates must be enrolled in the fourth year of university. In 2011-2013, candidates must be enrolled in the fifth year of university. In 2014-2017, candidates must be enrolled in their fourth year of university and must obtain a master's degree to be tenured. Column (3) and (4) include region fixed effects, the number of teaching positions and the unemployment rate in the region X year as control variables. Column (4) is a log-log specification for the number of candidates and positions. The coefficients and confidence intervals of column (4) are reported in Figure 3.3.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.21 – Effect of the interaction between the 2011 reform and the demand for teachers on the number of candidates

	(1)	(2)	(3)	(4)
	Number of candidates			Logarithm of number of candidates
Period 2011-2017	-237.2*** (60.46)	-281.0*** (41.28)	44.76 (69.62)	-0.199* (0.120)
Number of positions	3.080*** (0.0887)	2.922*** (0.125)	5.066*** (0.284)	
Number of positions x Period 2011-2017	-1.894*** (0.130)	-1.785*** (0.0875)	-3.109*** (0.286)	
Number of positions squared			-0.00156*** (0.000187)	
Number of positions squared x Period 2011-2017			0.000981*** (0.000195)	
Logarithm of number of positions				0.490*** (0.0256)
Logarithm of number of positions x Period 2011-2017				-0.0781*** (0.0209)
Region FE	No	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes	Yes
Observations	525	525	525	525

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parenthesis.

Notes: The variable explained is the number of candidates for the external primary teacher recruitment competition by region and year. The model considered is : $Y_{jt} = \sum R_j 1_j + \alpha f(N_{Pjt}) + \delta 1_{t>2010} + \beta f(N_{Pjt}) * 1_{t>2010} + \gamma X_{jt} + \epsilon_{jt}$, where $f(N_{Pjt})$ is a quadratic or a logarithmic function of the number of teaching positions in region j and year t , and $1_{t>2010}$ is a dummy variable equal to 1 if $t > 2010$ (period 2011-2017) and 0 otherwise. The coefficients of interest β , which reflect the effect of the interaction between the number of positions and the post-reform period, are indicated in bold in the Table. Column (2) to (4) include region fixed effects, the number of position and the unemployment rate in the region X year as control variables. Column (4) is a log-log specification for the number of candidates and positions.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.22 – Effect of increasing the degree level required to teach on the exam score of candidates and teachers

	(1)	(2)	(3)	(4)
	Candidates' exam score		Teachers' exam score	
	Raw data	Corrected data	Raw data	Corrected data
Year effects (ref: 2003)				
2004	0.0368 (0.0739)	0.118*** (0.0242)	-0.0948 (0.117)	0.0926* (0.0530)
2005	-0.0649 (0.0494)	0.0260 (0.0240)	-0.0866 (0.0991)	0.0674 (0.0491)
2006	-0.0509 (0.0686)	0.0751*** (0.0235)	-0.188 (0.119)	0.0890** (0.0430)
2007	0.100** (0.0423)	0.149*** (0.0375)	0.0623 (0.0698)	0.143*** (0.0501)
2008	0.162** (0.0617)	0.161*** (0.0566)	0.135 (0.0976)	0.124 (0.0914)
2009	0.0199 (0.0958)	0.0676** (0.0320)	-0.306** (0.126)	-0.120 (0.0793)
2010	-0.0188 (0.0864)	0.151*** (0.0351)	-0.504*** (0.115)	-0.136* (0.0763)
2011	-0.189 (0.118)	0.195*** (0.0623)	-1.295*** (0.196)	-0.550*** (0.0878)
2012	-0.239** (0.106)	0.223*** (0.0578)	-1.332*** (0.150)	-0.488*** (0.0865)
2013	-0.405** (0.151)	0.194*** (0.0474)	-1.706*** (0.259)	-0.602*** (0.101)
Control variables				
Region FE	Yes	Yes	Yes	Yes
Logarithm of number of positions	Yes	Yes	Yes	Yes
Number of attempts	Yes	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes	Yes
Observations	427570	427570	89174	89174

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis are clustered at the region level.

Notes: The variable explained is the exam score of candidates (columns 1 and 2) and teachers (columns 3 and 4) for the external primary teacher recruitment competition. Year fixed effects are estimated using model (1). The results obtained by considering the raw examination score are presented in columns (1) and (3) while the results obtained by considering the examination score corrected for evaluation bias (see Appendix) are presented in columns (2) and (4). Each column includes region fixed effects, the number of times candidates take the exam, the number of teaching positions and the unemployment rate in the region X year as control variables. The coefficients and confidence intervals of column (3) and (4) are reported in Figures 3.11 and 3.6 respectively.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.23 – Effect of the interaction between the 2011 reform and the demand for teachers on candidates' and teachers' exam score

	(1)	(2)	(3)	(4)
	Candidates' exam score		Teachers' exam score	
	Raw data	Corrected data	Raw data	Corrected data
Period 2011-2013	-0.112** (0.0518)	0.148*** (0.0209)	-0.396*** (0.112)	-0.0335 (0.0712)
Number of positions	0.0000603 (0.000174)	-0.000167*** (0.0000475)	-0.000320 (0.000258)	-0.000829*** (0.000147)
Number of positions x period 2011-2013	-0.000186*** (0.0000561)	0.0000530* (0.0000318)	-0.00110*** (0.000123)	-0.000681*** (0.0000772)
Control variables				
Region FE	Yes	Yes	Yes	Yes
Number of attempts	Yes	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes	Yes
Observations	394544	394544	82934	82934

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parenthesis are clustered at the region level.

Notes: The variable explained is the exam score of candidates (columns 1 and 2) and teachers (columns 3 and 4) for the external primary teacher recruitment competition. The model considered is : $Y_{jt} = \sum R_j 1_j + \alpha N_{Pjt} + \delta 1_{t>2010} + \beta N_{Pjt} * 1_{t>2010} + \gamma X_{jt} + \epsilon_{jt}$, where N_{Pjt} is the number of teaching positions in region j and year t , and $1_{t>2010}$ is a dummy variable equal to 1 if $t > 2010$ (period 2011-2013) and 0 otherwise. The coefficients of interest β , which reflect the effect of the interaction between the number of positions and the post-reform period, are indicated in bold in the Table. Each column includes region fixed effects, the number of times candidates take the exam and the unemployment rate in the region X year as control variables. The results obtained by considering the raw examination score are presented in columns (1) and (3) while the results obtained by considering the examination score corrected for evaluation bias (see Appendix) are presented in columns (2) and (4).

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.24 – Effect of increasing the degree level required to teach on the classroom observation score of teachers

	(1)	(2)	(3)
	Classroom observation score		
Year effects (ref: 2006)			
2007	-0.0604 (0.0404)	-0.188*** (0.0456)	-0.0512 (0.0368)
2008	-0.0474 (0.0714)	-0.254*** (0.0759)	-0.0589 (0.0671)
2009	0.0511 (0.0657)	0.107 (0.0723)	0.0354 (0.0439)
2010	0.0519 (0.0811)	0.130* (0.0738)	0.0707 (0.0868)
2011	0.167 (0.185)	0.310* (0.155)	0.160** (0.0754)
2012	0.105 (0.178)	0.281** (0.127)	0.141 (0.108)
2013	-0.0526 (0.161)	0.268** (0.125)	0.159 (0.161)
Control variables			
Logarithm of number of positions	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes
Net salary of masters' graduates	Yes	Yes	Yes
Net salary of bachelor's graduates	Yes	Yes	Yes
Region FE	No	Yes	-
School FE	No	No	Yes
Observations	39686	39686	39686

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis are clustered at the region level.

Notes: The variable explained is the classroom observation score (normalized mean 0, variance 1) 2-3 years after the tenure of teachers recruited over the period 2006-2013. Year fixed effects are estimated using model (2). Each column includes the number of teaching positions (in logarithm), the unemployment rate and the average net salary at the bachelor's and master's level in the region X year as control variables. Column (2) includes region fixed effects and column (3) includes school fixed effects. The coefficients and confidence intervals of column (3) are reported in Figure 3.6.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.25 – Effect of increasing the diploma level required to teach on the number of candidates

	(1)	(2)	(3)	(4)
	Share of men		Median income in city of residence	
	Candidates	Teachers	Candidates	Teachers
Year effects (ref: 2003)				
2004	-0.00124 (0.00299)	-0.00398 (0.00687)	-38.78 (48.12)	-73.41 (64.24)
2005	0.00566 (0.00403)	0.00532 (0.00697)	12.47 (75.01)	101.4 (95.12)
2006	0.00321 (0.00355)	0.00347 (0.00686)	-20.19 (65.53)	93.63 (109.0)
2007	0.00167 (0.00468)	0.00794 (0.00498)	-46.33 (42.38)	56.01 (45.42)
2008	-0.00149 (0.00510)	0.0107 (0.00863)	-32.85 (63.91)	182.6** (84.22)
2009	-0.00451 (0.00524)	0.0172* (0.00954)	-117.0 (74.81)	-189.6** (73.62)
2010	-0.0215*** (0.00535)	-0.00677 (0.00767)	-90.54 (81.13)	-226.6** (101.5)
2011	-0.0198** (0.00816)	-0.00000325 (0.0147)	15.17 (98.88)	-221.0** (90.85)
2012	-0.0315*** (0.00722)	-0.0215** (0.0104)	-70.70 (129.8)	-307.1*** (105.9)
2013	-0.0411*** (0.00580)	-0.0250*** (0.00752)	-69.11 (140.3)	-431.5*** (81.29)
2014	-0.0315*** (0.00512)	-0.0210*** (0.00633)	-124.8 (134.5)	-487.3*** (103.4)
2015	-0.0248*** (0.00442)	-0.0177** (0.00648)	-159.2 (125.7)	-514.1*** (121.7)
Control variables				
Region FE	Yes	Yes	Yes	Yes
Number of positions	Yes	Yes	Yes	Yes
Number of positions squared	Yes	Yes	Yes	Yes
Number of attempts	Yes	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes	Yes
Net salary of bachelors' graduates	Yes	Yes	Yes	Yes
Net salary of masters' graduates	Yes	Yes	Yes	Yes
Observations	476462	107357	474172	106904
Adjusted R-squared	0.00	0.00	0.17	0.19

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis are clustered at the region level.

Notes: The variables explained is the gender of candidates (column 1) and teachers (column 2), and the median income (in euros per year) in the city of residence of candidates (column 3) and teachers (column 4) for the external primary teacher recruitment competition over the period 2003-2015. To neutralize the effect of income growth over time, I consider the median income in euros per consumption unit (i.e. divided by the number of inhabitants in the household) in cities in 2012. Year fixed effects are estimated using model (1). Each column includes region fixed effects, the number of times candidates take the exam, the number of teaching positions, the unemployment rate and the average salary at the bachelors' and masters' level in the region X year as control variables. The coefficients and confidence intervals of column (1) to (4) are reported in Figures 3.14, 3.15, 3.16 and 3.17 respectively.

Source: Ministry of Education (DEPP) and National Institute of Statistics (INSEE).

Table 3.26 – Evolution of the number of candidates and the proportion of men among candidates and recruits of the Regional Institutes of Administration (IRA) over the period 2008-2017

	(1)	(2)	(3)
	Logarithm of number of candidates	Share of men among candidates	Share of men among recruited persons
Year effects (ref: 2008)			
2009	0.182*** (0.0472)	0.0571*** (0.0159)	0.0367 (0.0553)
2010	0.151*** (0.0368)	0.0362*** (0.0119)	-0.00625 (0.0414)
2011	0.115*** (0.0332)	0.0154 (0.0105)	-0.0394 (0.0364)
2012	0.0797** (0.0332)	0.0146 (0.0105)	-0.0438 (0.0364)
2013	0.0372 (0.0353)	-0.0279** (0.0113)	0.00455 (0.0393)
2014	-0.0442 (0.0316)	-0.000289 (0.00979)	-0.0347 (0.0341)
2015	-0.131*** (0.0316)	0.0182* (0.00979)	-0.0276 (0.0341)
2016	-0.0853 (0.0617)	-	-
2017	-0.0622 (0.0617)	-	-
Control variables			
IRA x competition FE	Yes	Yes	Yes
Logarithm of number of positions	Yes	Yes	Yes
Unemployment rate	Yes	Yes	Yes
Observations	100	80	80
Adjusted R-squared	0.74	0.83	0.39

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis.

Notes: The variables explained are the number of candidates (column 1) and the share of men among candidates (column 2) and recruits (column 3) in the external and internal Regional Institutes of Administration (IRA) competitive examinations by region and year. Public data do not provide the proportion of men among candidates and recruits in 2016 and 2017. Year fixed effects are estimated using model (1') in column (1), and model (1) in column (2) and (3). Five IRAs located in Bastia, Lille, Lyon, Metz and Nantes are recruiting through competitive examinations. The data are aggregated by type of competition (internal or external) at the IRA level (2 X 5 observations each year). Each column includes IRA X type of competition fixed effects, the number of positions (in logarithm) and the unemployment rate in the IRA X year as control variables. The coefficients and confidence intervals of column (1) to (3) are reported in Figures 3.18, 3.19 and 3.20 respectively.

Source: Ministry of Public Service (DGAFP) and National Institute of Statistics (INSEE).

Table 3.27 – Number of candidates by year and age group

Age group:	18-23	24-28	29-33	34-38	39-43	44-48	49-53	54-58
2003	13644	29421	11237	4649	1949	493	103	18
2004	13868	29802	10302	4531	1980	515	93	17
2005	11849	25713	9778	5001	2227	605	128	12
2006	12089	24549	8587	4419	1868	556	94	16
2007	12610	22757	7106	3739	1702	473	101	16
2008	11564	20977	6243	3372	1505	500	92	20
2009	10736	19115	6135	3452	1684	531	81	23
2010	9199	15982	4854	2786	1457	483	89	13
2011	2348	8482	2381	1573	952	377	86	17
2012	2891	9167	2156	1275	869	354	80	15
2013	2415	10058	2632	1628	1150	453	105	24
2014	5966	9934	3402	2352	1676	687	202	46
2015	6201	10957	4390	3108	2146	956	308	76

Notes: Number of candidates for the external primary school teacher recruitment competition by age group over the period 2003-2015.

Source: Ministry of Education (DEPP). Authors' calculations.

Figures

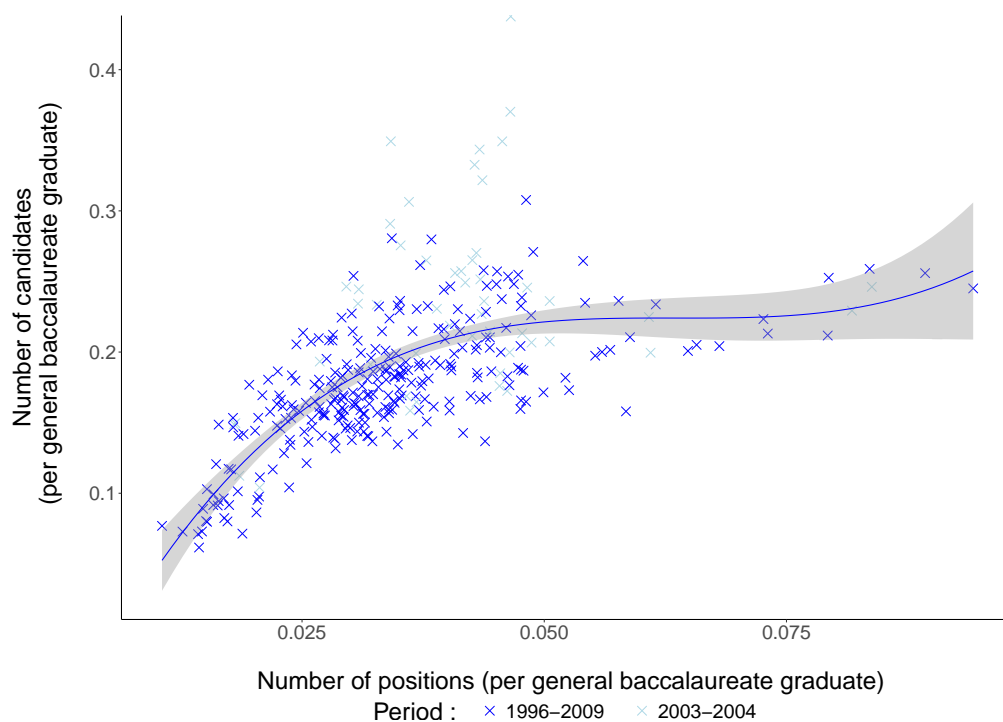


Figure 3.8 – Number of candidates according to the number of teaching posts over the period 1996-2009

Note: Each symbol represents a region X year. I divide the number of candidates (Y-axis) and the number of positions (X-axis) by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable. The number of candidates increases significantly in 2003 and 2004 (light blue markers) because candidates were exceptionally allowed to write the exam in two regions in the same year.

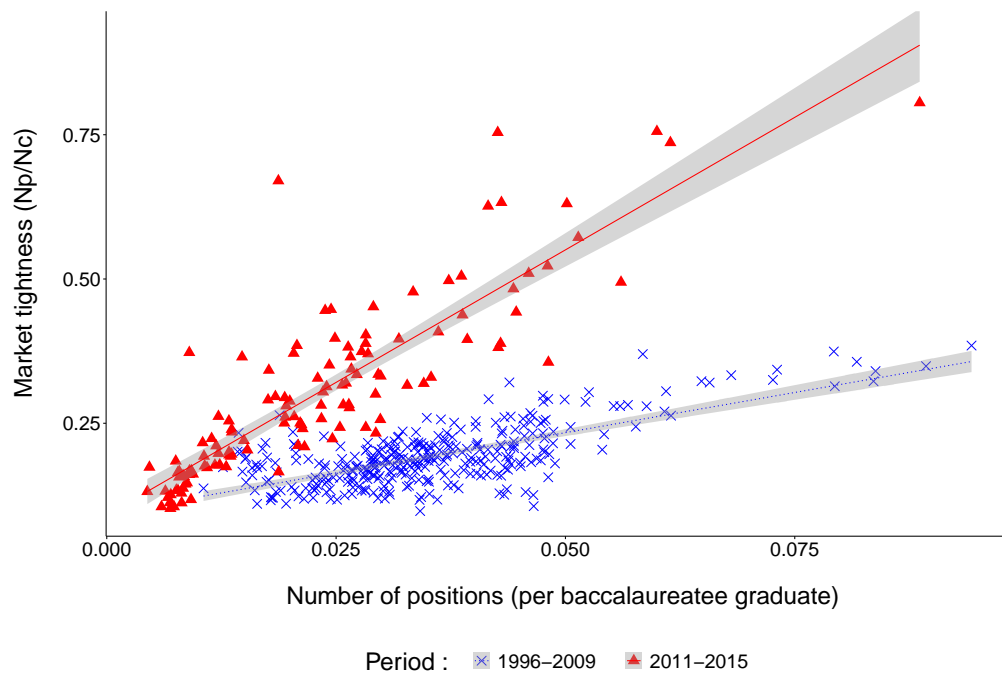


Figure 3.9 – Teacher market tightness according to the number of teaching posts over the periods 1996-2009 and 2011-2015

Note: Each symbol represents a region X year. Teacher market tightness on the Y-axis corresponds to the number of teaching positions per candidate in the region. A higher value indicates a greater difficulty in recruiting teachers. I divide the number of positions (X-axis) by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable.

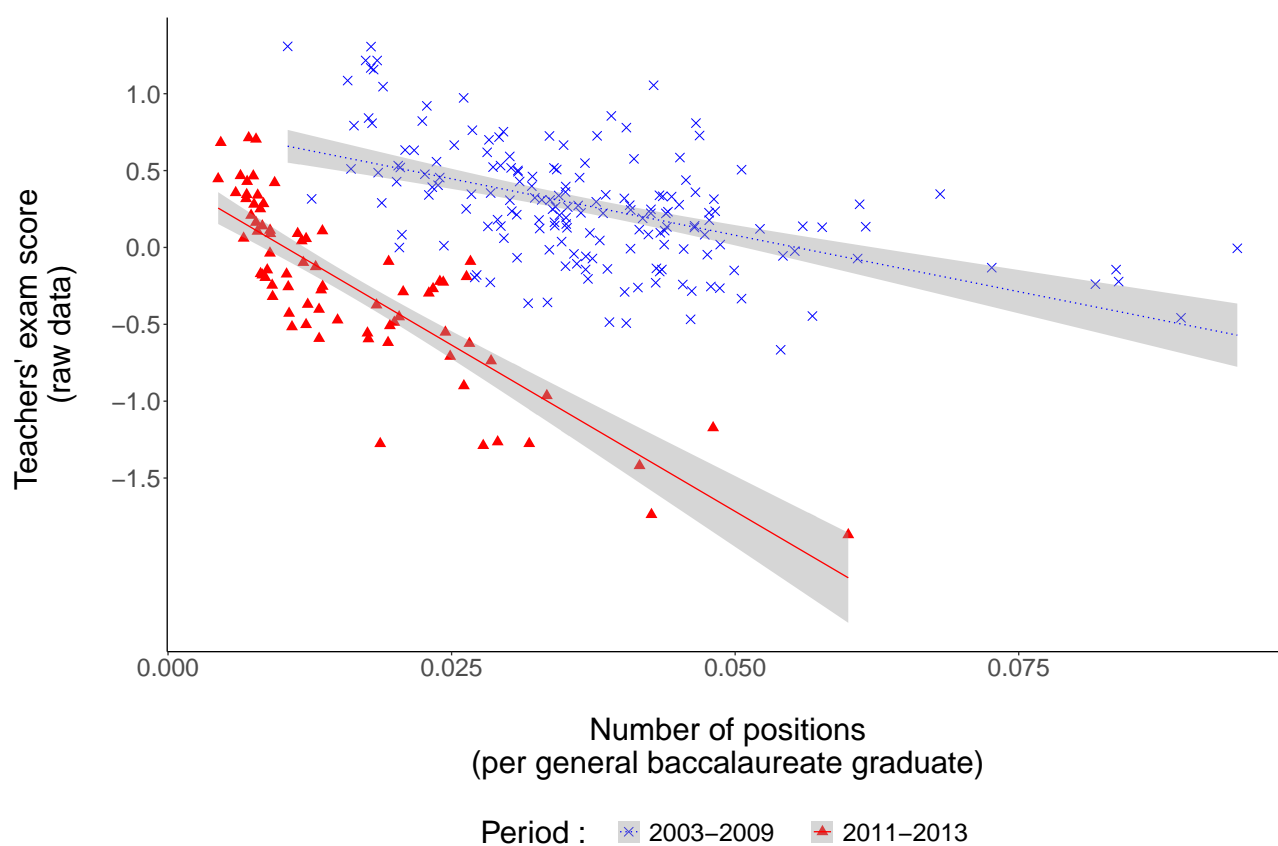


Figure 3.10 – Average score of teachers (raw data) on the recruitment examination during the periods 1996-2009 and 2011-2013

Note: Each symbol represents a region X year. The X-axis presents the number of positions divided by the number of general baccalaureate graduates 3 years before the competition date in order to make the regions more comparable. The Y-axis shows the average recruitment score (raw data) of teachers for each region and year. The raw teacher score (i.e. the score before correction by the method proposed in the Appendix) is normalized (mean 0, variance 1) over the period 2003-2013.

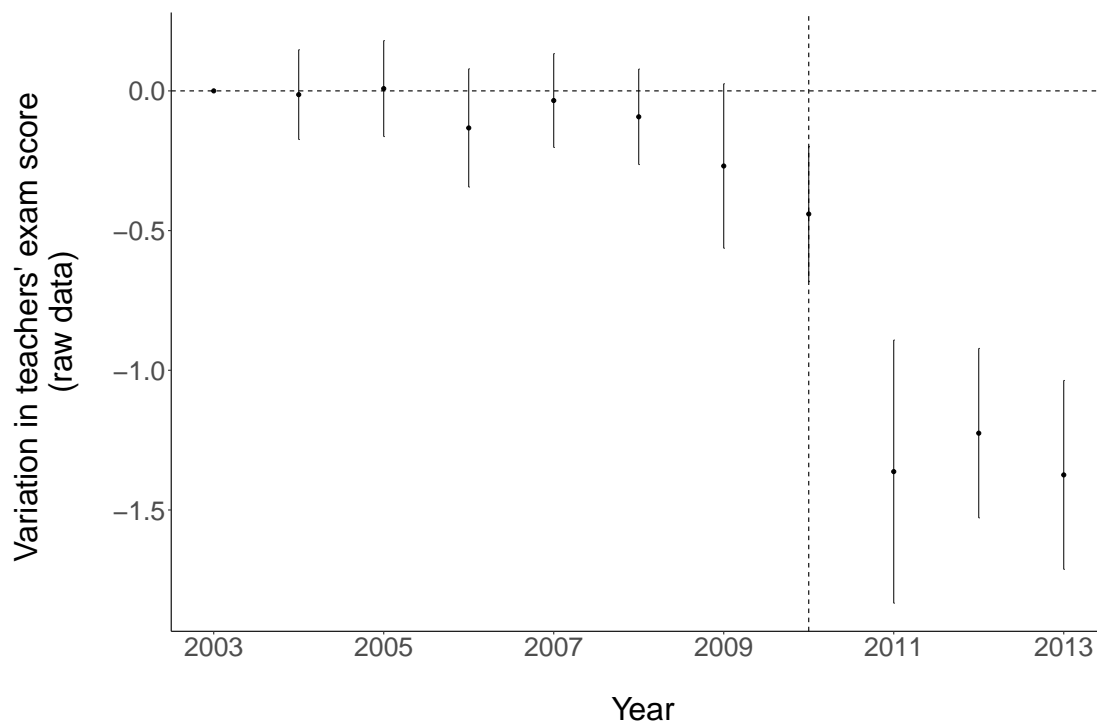


Figure 3.11 – Annual variation in the average exam score (raw data) of teachers compared to 2003

Note: The Figure shows that the raw examination score of teachers decreases on average by 1 SD point from 2011. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of positions (in logarithm) and the unemployment rate in the region as control variables. Column (3) of Table 3.22 reports the coefficients and standard errors.

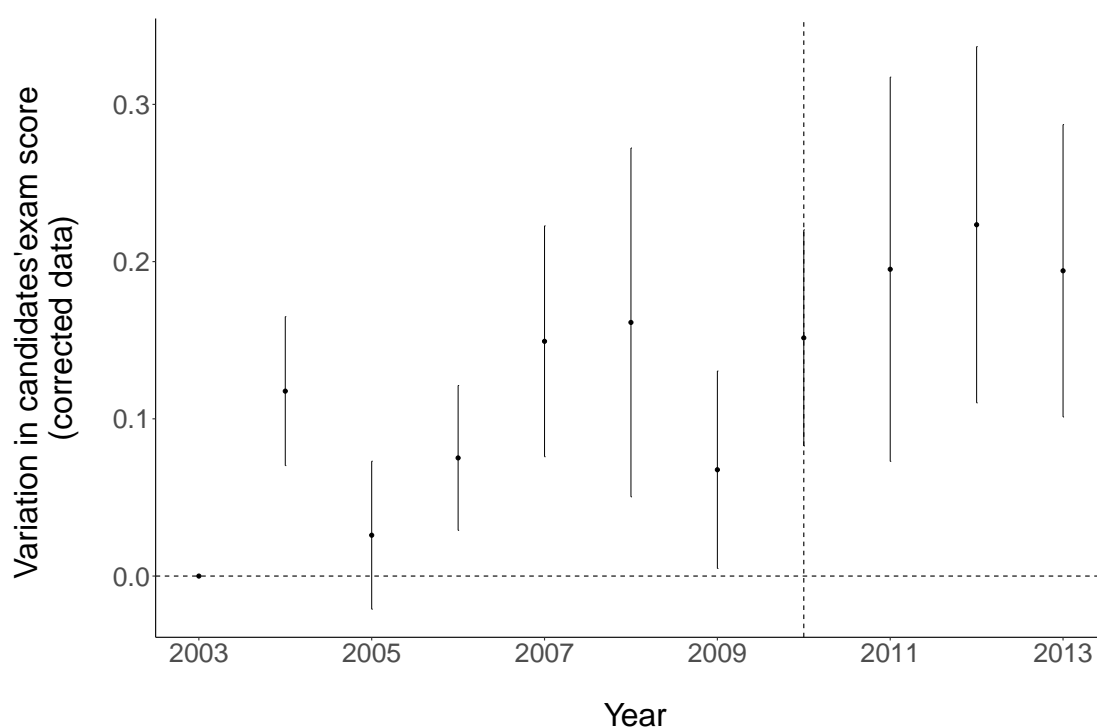


Figure 3.12 – Annual variation in the average examination score (corrected for assessment bias, see Appendix) of candidates compared to 2003

Note: The Figure shows how the examination score (corrected for assessment bias, see Appendix) of candidates varies over time compared to 2003. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of positions (in logarithm) and the unemployment rate in the region as control variables. Column (2) of Table 3.22 reports the coefficients and standard errors.

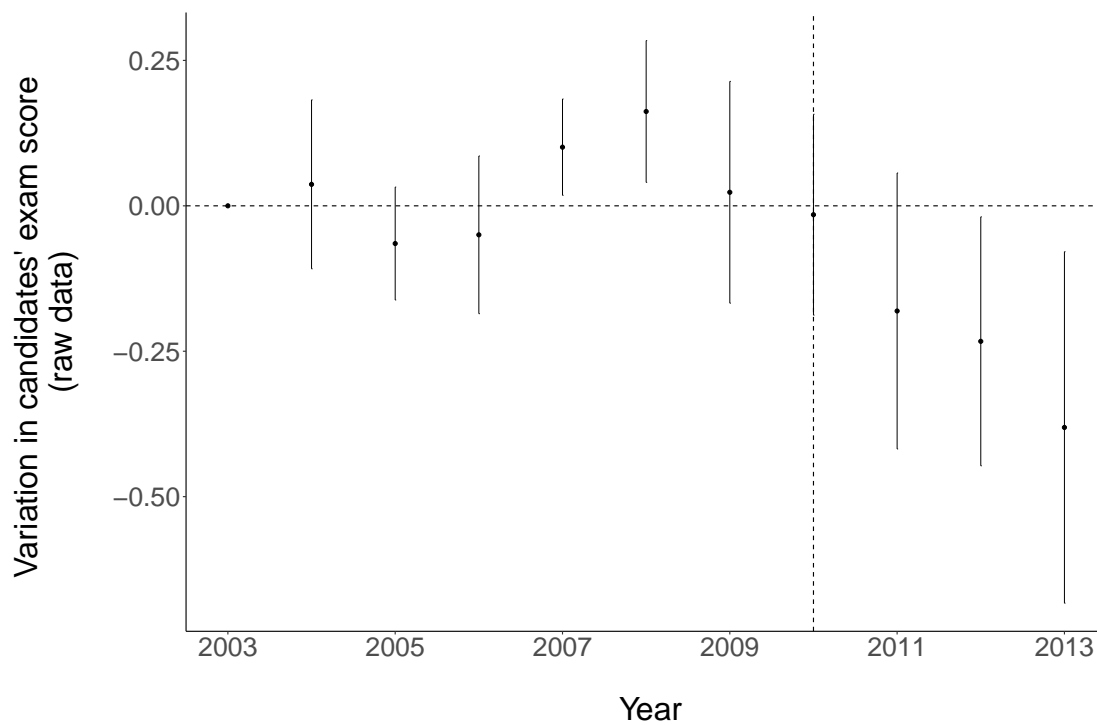


Figure 3.13 – Annual variation in the average examination score (raw data) of candidates compared to 2003

Note: The Figure shows how the raw examination score of candidates varies over time compared to 2003. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of positions (in logarithm) and the unemployment rate in the region as control variables. Column (1) of Table 3.22 reports the coefficients and standard errors.

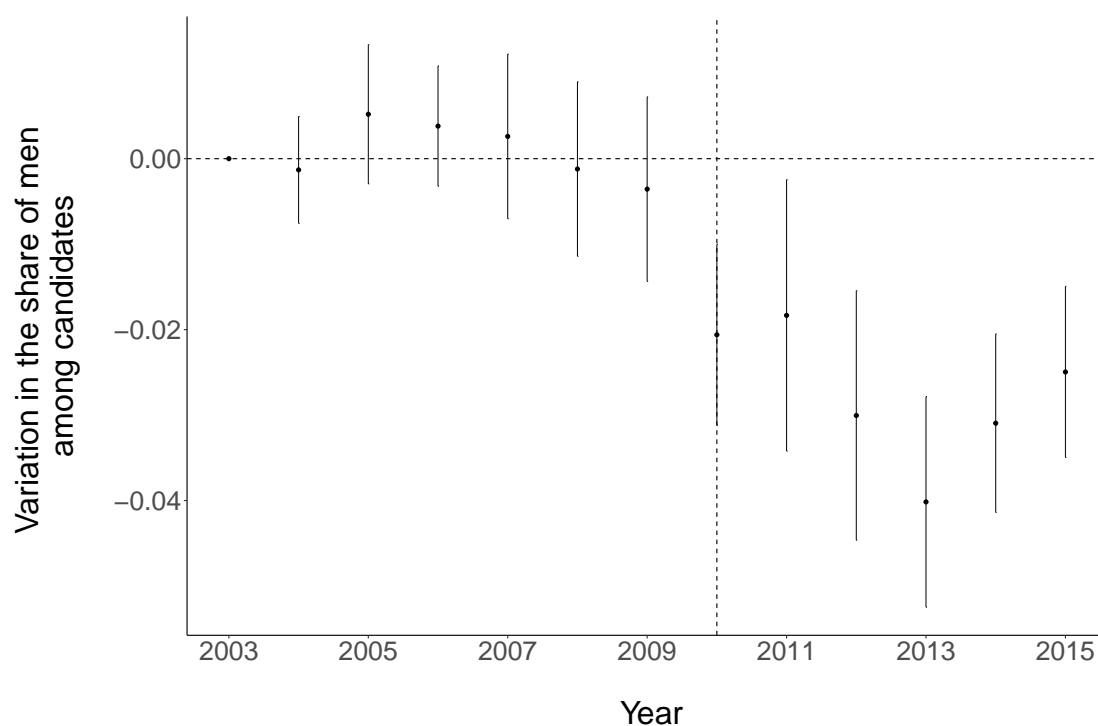


Figure 3.14 – Annual variation in the proportion of men among candidates compared to 2003

Note: The Figure shows how the share of men among candidates varies over time compared to 2003. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of teaching positions, the number of times candidates write the exam, the unemployment rate and the average salary at the bachelors' and masters' level in the region as control variables. Column (1) of Table 3.25 reports the coefficients and standard errors.

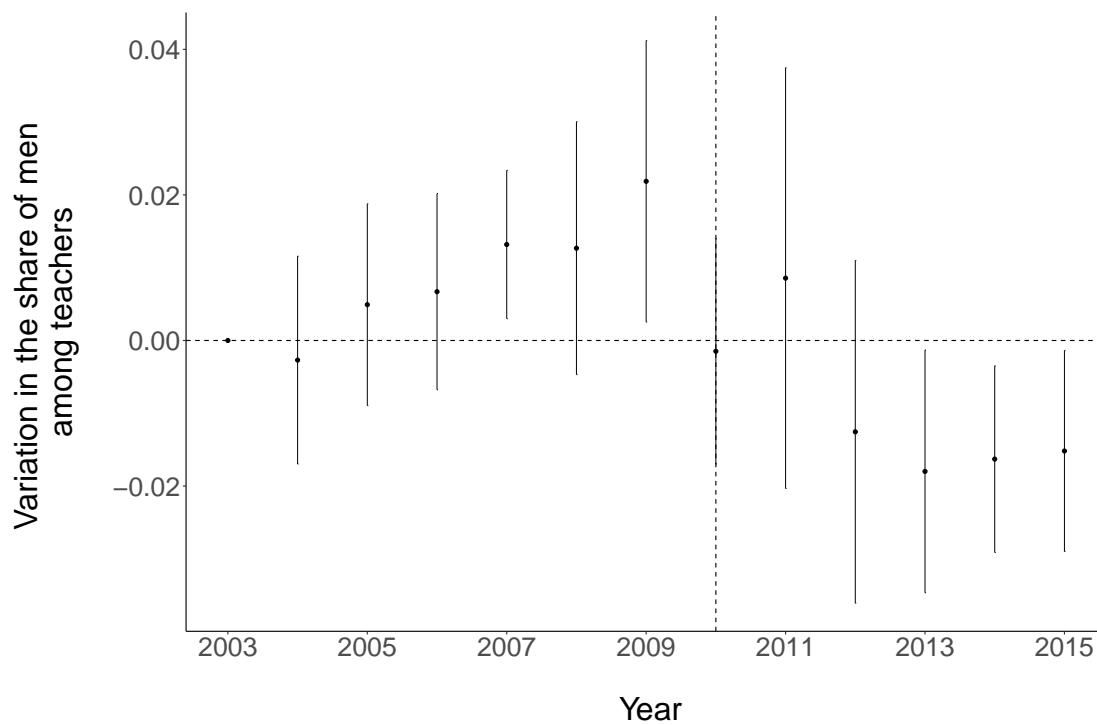


Figure 3.15 – Annual variation in the proportion of men among teachers compared to 2003

Note: The Figure shows how the share of men among teachers varies over time compared to 2003. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of teaching positions, the number of times candidates write the exam, the unemployment rate and the average salary at the bachelors' and masters' level in the region as control variables. Column (2) of Table 3.25 reports the coefficients and standard errors.

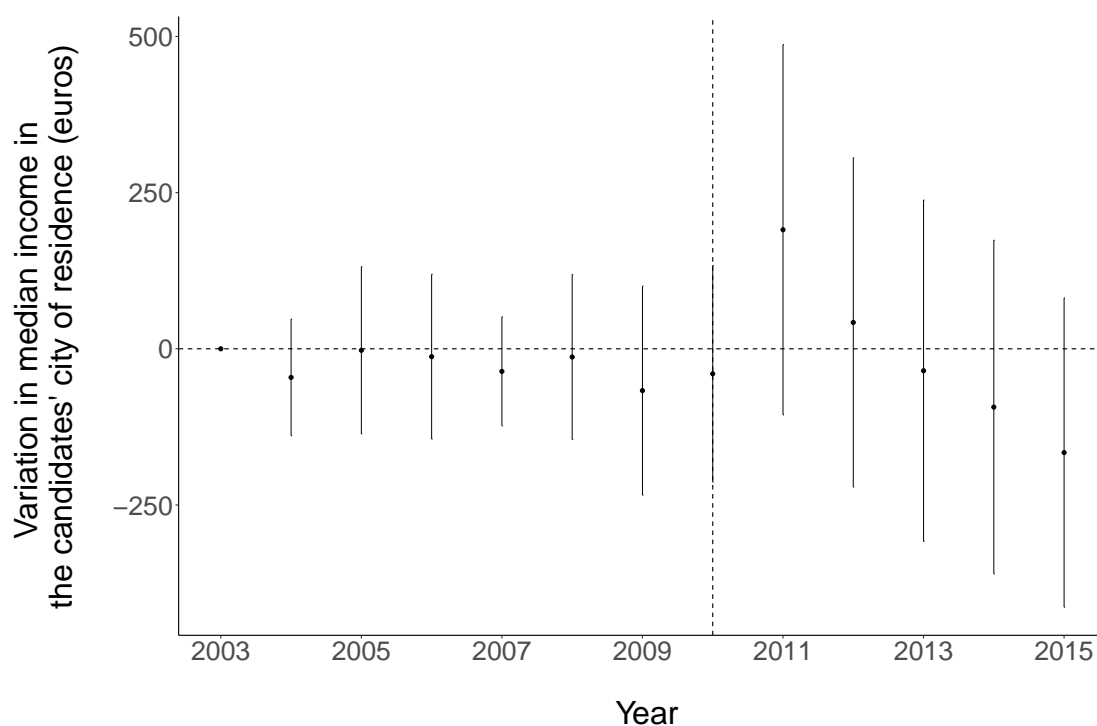


Figure 3.16 – Annual variation in the median income in the city of residence of candidates compared to 2003

Note: The Figure shows how the median income (in euros per year and per consumption unit) in the city of residence of candidates varies over time compared to 2003. To neutralize the effect of income growth over time, I consider the median income in cities in 2012. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of teaching positions, the number of times candidates write the exam, the unemployment rate and the average salary at the bachelors' and masters' level in the region as control variables. Column (3) of Table 3.25 reports the coefficients and standard errors.

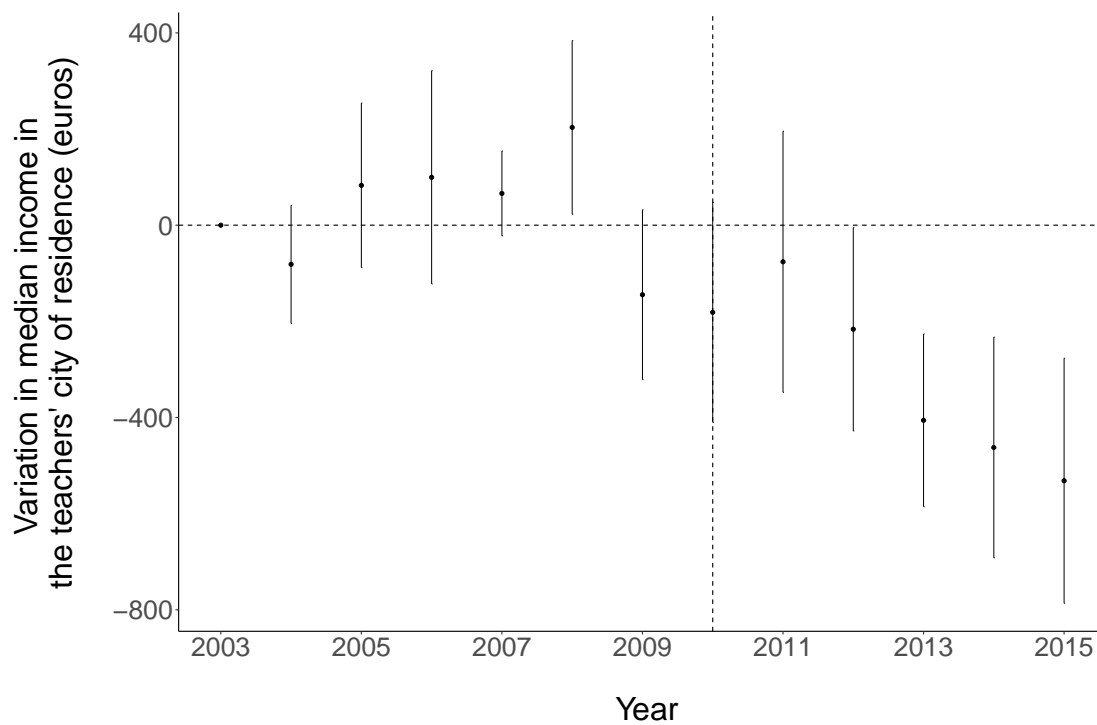


Figure 3.17 – Annual variation in the median income in the city of residence of teachers compared to 2003

Note: The Figure shows how the median income (in euros per year and per consumption unit) in the city of residence of teachers varies over time compared to 2003. To neutralize the effect of income growth over time, I consider the median income in cities in 2012. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including region fixed effect, the number of teaching positions, the number of times candidates write the exam, the unemployment rate and the average salary at the bachelors' and masters' level in the region as control variables. Column (4) of Table 3.25 reports the coefficients and standard errors.

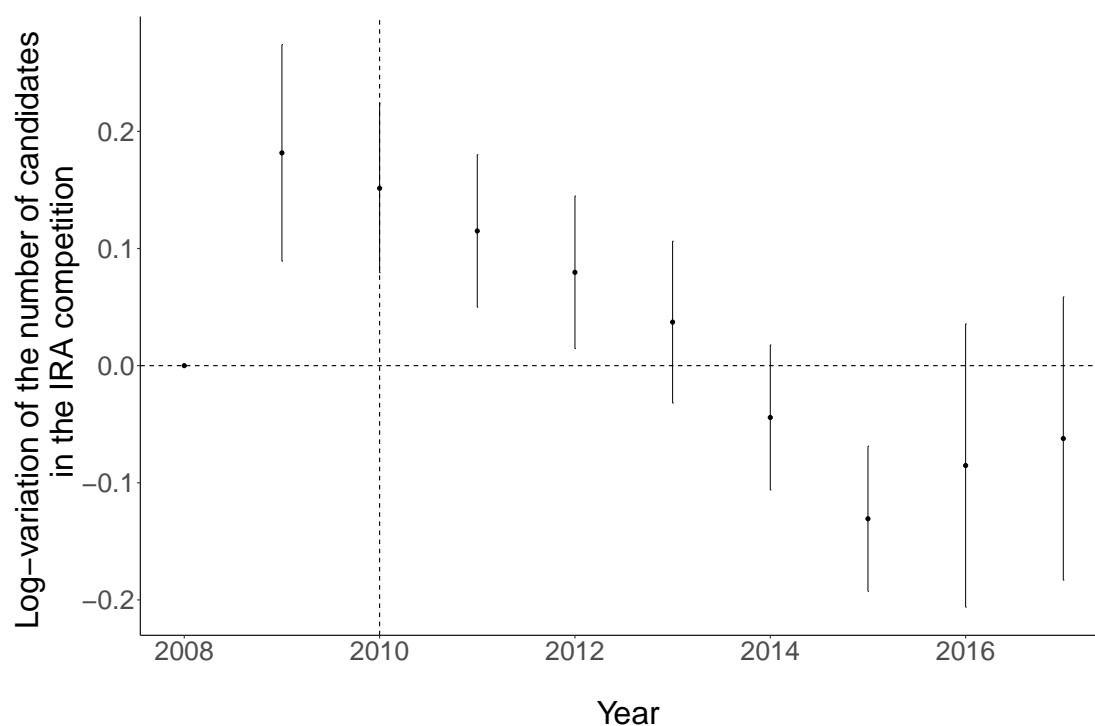


Figure 3.18 – Annual logarithmic variation in the number of candidates for the Regional Institutes of Administration (IRA) compared to 2008

Note: Five IRAs located in Bastia, Lille, Lyon, Metz and Nantes are recruiting through competitive examinations. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1') where the number of candidates and the number of positions for the external and internal Regional Institutes of Administration (IRA) competitive examinations are specified in logarithm. The model also includes IRA X type of competition (internal or external) fixed effects, and the unemployment rate in the IRA X year as control variables. Column (1) of Table 3.26 reports the coefficients and standard errors.

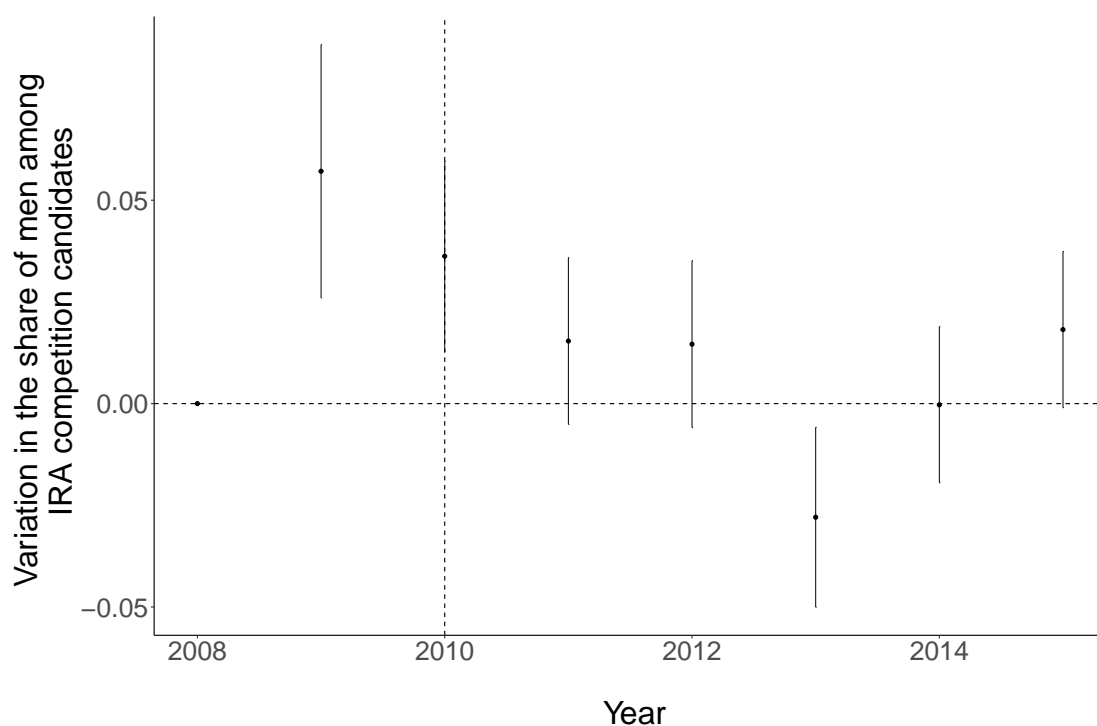


Figure 3.19 – Annual variation in the share of men among candidates for the Regional Institutes of Administration (IRA) compared to 2008

Note: Five IRAs located in Bastia, Lille, Lyon, Metz and Nantes are recruiting through competitive examinations. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including IRA X type of competition (internal or external) fixed effects, the number of IRA positions (in logarithm) and the unemployment rate in the IRA X year as control variables. Column (2) of Table 3.26 reports the coefficients and standard errors.

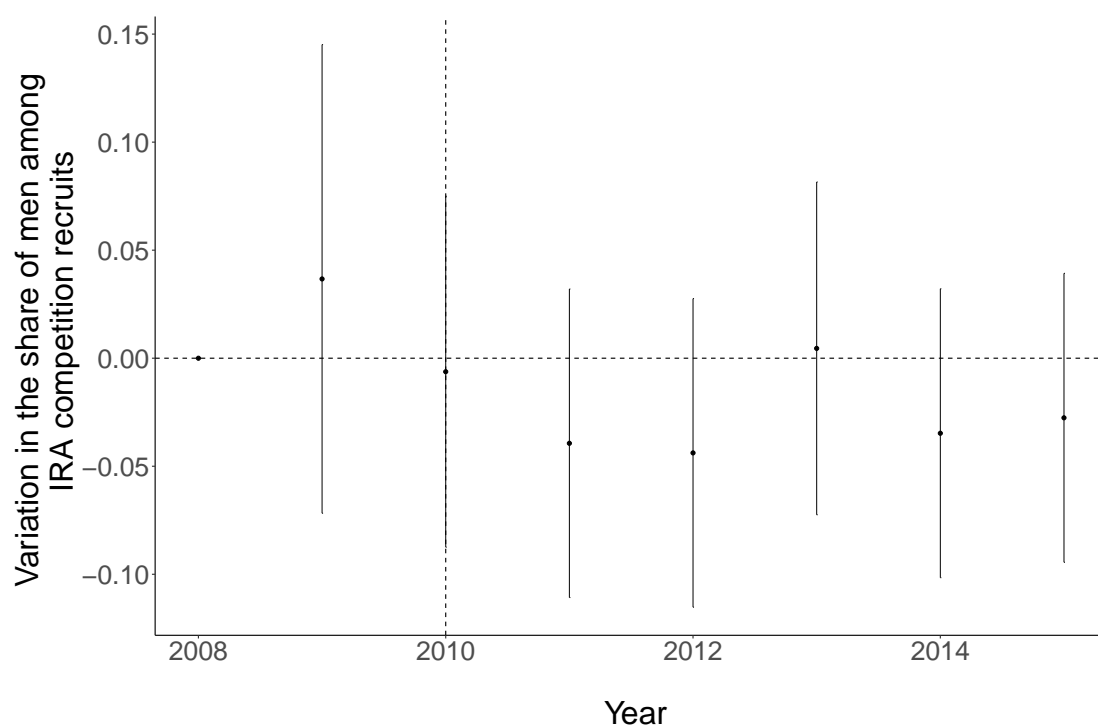


Figure 3.20 – Annual variation in the share of men among recruits for the Regional Institutes of Administration (IRA) compared to 2008

Note: Five IRAs located in Bastia, Lille, Lyon, Metz and Nantes are recruiting through competitive examinations. Coefficients (points) and confidence intervals at the 5% level (segments) are estimated from model (1) including IRA X type of competition (internal or external) fixed effects, the number of IRA positions (in logarithm) and the unemployment rate in the IRA X year as control variables. Column (3) of Table 3.26 reports the coefficients and standard errors.

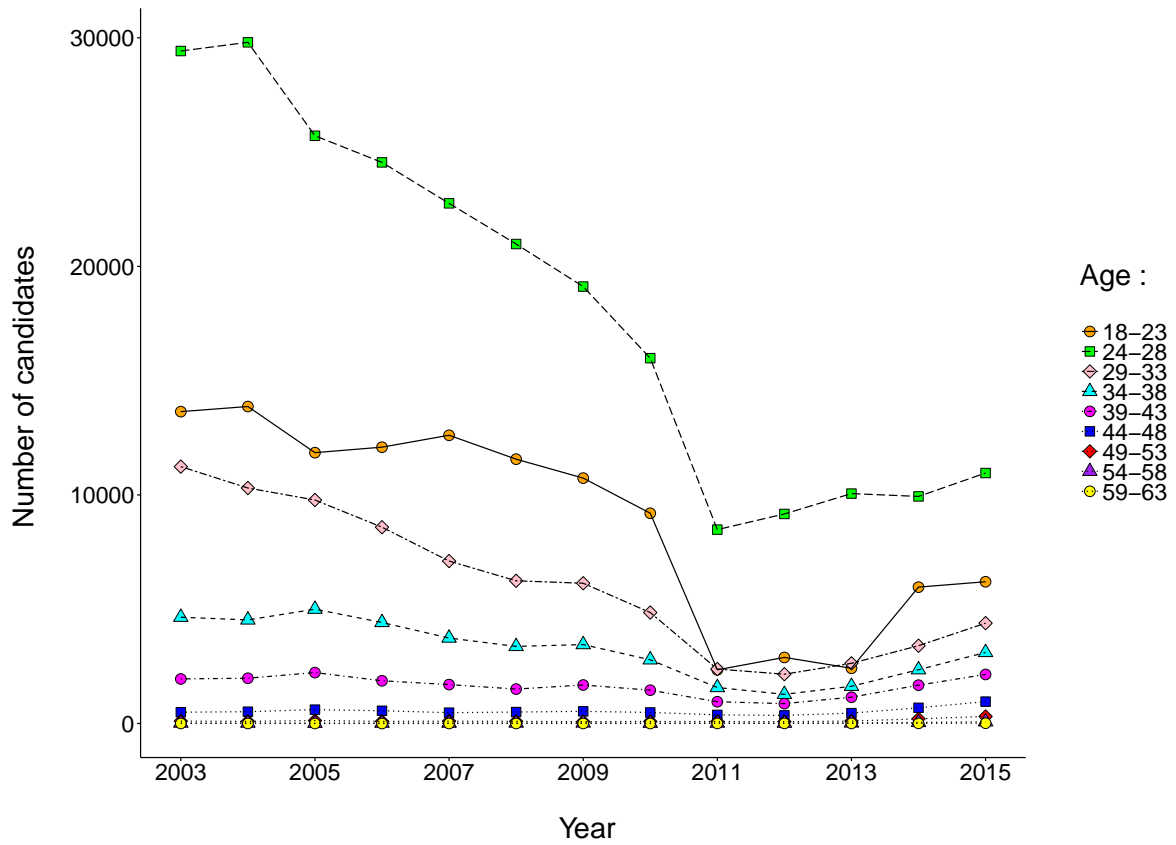


Figure 3.21 – Evolution of the number of candidates by age group over the period 2003-2015

Note: Trend in the number of candidates for the external primary school teacher recruitment competition by age group over the period 2003-2015.

Appendix

This Appendix examines the extent to which teachers' recruitment test scores can be compared over time and across regions.

A Primary school teacher recruitment examinations are not standardized in France

Primary school teachers are recruited through competitive examinations in 30 academic regions. Recruitment tests take the form of open-ended questions (e.g. essays, problems) and are designed to assess the same competencies in all regions. However, examination topics vary from year to year and often from region to region. Reviewers are provided with detailed guidelines for assessing candidates and harmonize their rating practices within regions. For example, they jointly evaluate a selection of test sheets, which allows to calibrate the assessment of other candidates. Each written test is evaluated by two different examiners. This harmonization process is intended to ensure equity among candidates for the same teaching positions in the regions.

But, is it possible to compare the scores obtained by candidates in different regions and different years? Several studies show that the French regions are characterized by significant socio-economic disparities (parents' level of education and social background, unemployment rate, family situation, household income) and significant inequalities in educational attainment (high school graduation rate, educational delay, share of non-graduates, higher education continuation rate)⁴³⁴⁴. These differences in socio-economic background and educational achievement are likely to generate differences in candidates' characteristics between regions (the vast majority of candidates participate in the competition in the region where they live). As a result, the tests used to calibrate the assessment are likely to vary from one region to another. In addition, boards of examiners are composed of teachers and professors who work in the region where the recruitment takes place. They may share assessment practices related to teaching experience that differ between regions. Therefore, rating practices are likely to be heterogeneous across the

⁴³DEPP, *Géographie de l'école 2014*, n° 11.

⁴⁴Broccolichi, S, Ben-Ayed, C, Mathey-Pierre, C, Trancart, D (2007). Fragmentations territoriales et inégalités scolaires: des relations complexes entre la distribution spatiale, les conditions de scolarisation et la réussite des élèves, *Education et Formation*, n° 74.

country.

Similarly, recruitment scores may not be comparable from year to year. In particular, the 2011 reform, which raised the level of qualification required to become a teacher, may have changed the profile of candidates and the expectations of assessors, who are potentially more demanding with master's degree holders than with bachelor's degree holders⁴⁵.

To verify the existence of bias due to the imperfect standardization of primary teacher recruitment examinations, I propose two methods. The first strategy exploits the fact that candidates were exceptionally allowed to take the competitive exams in the same year in two different regions in 2003 and 2004⁴⁶. The rating biases, assimilated to region fixed effects in 2003 and 2004, can be estimated by difference, which neutralize the effect of candidates' skills on exam score. This method does not allow to estimate evaluation biases between years. The second approach exploits the fact that most candidates participate in the competition two consecutive years and that a significant proportion of them change region on the second attempt. Under certain assumptions, the rating biases, assimilated to region x year effects, can be estimated by double difference, which makes it possible to neutralize the effect of candidates' skills and the effect of learning between attempts (typically because of the similarity of the tests).

The results obtained using both approaches suggest that assessment biases exist between regions and tend to underestimate the extent of recruitment disparities. Between the most and least selective region (metropolitan France), the average accreditation score increases from 1.5 SD point before correction to 1.8 SD point after correction of evaluation bias. The second method also suggests a significant change in scoring practices in 2011, which tends to overestimate the negative impact of the reform. I propose a corrected score that allows to estimate a lower bound (in absolute terms) of the effect of the 2011 reform.

⁴⁵It is important to note that no formal recommendations have been made to encourage reviewers to modify the assessment of candidates. The composition of the examination committees (mainly primary school teachers, secondary school teachers and education inspectors) was also not significantly changed after the reform. In 2011, many presidents and jury members have already been involved in the recruitment of primary school teachers in previous years.

⁴⁶With the exception of 2003 and 2004, candidates apply to only one region each year. The written tests are held on the same day in all regions. In 2003 and 2004, the academic regions were divided into two groups and two test dates were organized, one week apart

B Scope of the analysis

As in Chapter 3, I consider the scores obtained by candidates in the written tests because the oral tests were modified several times during the period under review (in 2006 and 2011).

The methods presented in this annex exploit the fact that a significant number of candidates apply for the competition several times in different regions (the same year or two different years), which generates many "connections" between regions. An exploratory analysis reveals that the overseas departments (Martinique, Guadeloupe, Reunion Island and French Guiana) and Corsica are much less "connected" to each other and to the metropolitan regions, which can be explained in two ways. First, the number of positions and candidates is much lower (with the exception of Reunion Island) in these regions than in the academic regions of metropolitan France. Second, the geographic mobility of candidates is lower (97% of candidates live in the regions where they apply) in these academic regions than in metropolitan areas (probably due to geographic distance). As the number of connections between the metropolitan regions and the French overseas departments and Corsica is low, I exclude the latter from the analysis.

C Descriptive statistics

I standardize the average (weighted) scores obtained in the written examinations (mean zero and variance one) over the period 2005-2013. Table 3.31 indicates at the national level and for the 25 academic regions of metropolitan France the number of candidates for the period 2005-2013⁴⁷ (column 1), the average score of candidates (column 2), the standard deviation of the score of candidates (column 3), the (average) score of the last candidate eligible for oral examinations (column 4), the proportion of candidates eligible for oral examinations (column 5) and the proportion of candidates recruited (column 6). Table 3.31 suggests the existence of significant disparities between regions. First, the average success rate in the competition (number of teachers recruited divided by the number of candidates) over the period 2005-2013 varies from 17% in the Bordeaux region to 36% in the Créteil region (column 6). Similarly, the share of candidates eligible for oral

⁴⁷The fact that candidates were able to take the competition in two regions during the 2003 and 2004 school years artificially inflated the number of candidates, so I do not present the results for these years in Table 3.31

examinations (second stage of the competition) is generally twice as high as the success rate in the region⁴⁸ (column 5). The average written examination score of candidates (column 2), as well as the score of the last candidate eligible for the oral examinations (column 4), also vary greatly from one region to another. One of the objectives of this study is to determine whether these disparities in exam scores reflect actual differences in teachers' knowledge or whether they result from differences in grading practices (or in complexity of recruitment tests) between regions.

On average, candidates take the competition 1.8 times during the period 2005-2013 (185,592 unique individuals for 329,304 candidates). Specifically, 89,361 individuals take the competition once and 96,232 individuals take the competition twice or more. Table 3.28 shows how the average score (column 2) varies depending on the number of times candidates participate in the competition. Candidates who participate in the competition only once have a score significantly higher than candidates who participate several times in the competition, at least on their first attempt. Table 3.28 also presents the average percentile rank of candidates (column 4), normalized by region x year. On average, candidates move up 0.08 percentile rank on their second participation in the competition. The ranking of candidates improves with each new attempt, but with decreasing performance.

⁴⁸Examiners set the eligibility threshold for oral examinations so that there are approximately two candidates eligible for a vacant teaching position

Table 3.28 – Average score (standardized) and average rank (percentile) of candidates according to the number of times they participated in the competition during the period 2005-2013

	(1)	(2)	(3)	(4)
	Observations	Average written score	Standard-deviation	Average written rank
All candidates	329304	0	1	0.5
Candidates who take the exam once	89361	0.051	1.22	0.528
Candidates who take the exam twice				
First attempt	39146	-0.148	0.857	0.444
Second attempt	39146	0.208	1.049	0.528
Candidates who take the exam three times				
First attempt	18111	-0.227	0.828	0.41
Second attempt	18111	0.022	0.786	0.492
Third attempt	18111	0.156	0.973	0.554
Candidates who take the exam four times				
First attempt	8220	-0.32	0.796	0.38
Second attempt	8220	-0.058	0.749	0.463
Third attempt	8220	-0.012	0.764	0.486
Fourth attempt	8220	0.015	0.911	0.524

Note: This table presents the score and ranking of candidates according to the number of times they participated in the teacher recruitment competition during the period 2005-2013. The interval between attempts is at least one year and results are shown for the first 4 attempts only. The years 2003 and 2004 are excluded because candidates were exceptionally allowed to take the exam twice in the same year (in two different regions). The period after 2013 is also excluded because the written tests were modified in 2014 and no longer measure the same knowledge as for the period 2003-2013. The examination score of candidates in column (2) is normalized (mean 0 and variance 1) over the period 2005-2013. Column (3) shows the standard deviation of candidates' exam score on each attempt. Column (4) indicates the percentile ranking of candidates in each region X year. It is calculated by dividing the rank of each candidate in a given region and year by the number of candidates in that region and year. For each region and year, the top ranked candidate has a percentile rank of 1 while the last ranked candidate has a percentile rank of 0.

Source: Ministry of Education (DEPP).

D A first method for estimating scoring biases between regions

Exceptionally, in 2003 and 2004, candidates were allowed to take the examination in two different academic regions in the same year. Regions were divided in two groups according to geographical criteria (western-southern academic regions and eastern-northern academic regions of mainland France) and two examination dates were organized one week apart⁴⁹. The written tests were identical within the same group of regions x year. Approximately 25% of applicants participated in the competition in two different regions in 2003 or 2004 (about 5,500 candidates each year).

⁴⁹In 2003, the academic regions, which take the name of their largest city, were divided between the two groups as follows. Group 1 (first exam date): Paris, Créteil, Versailles, Lille, Amiens, Rouen, Strasbourg, Reims, Nancy-Metz, Besançon, Orléans-Tours, Dijon, Lyon. Group 2 (second exam date): Grenoble, Aix-Marseille, Bordeaux, Caen, Clermont-Ferrand, Montpellier, Poitiers, Rennes, Toulouse, Nantes, Limoges, Nice. In 2004, the Grenoble academic region was moved to Groupe 1.

The objective is to estimate the rating bias specific to each region over the 2003-2004 period. The fact that some candidates take the exam in two different regions in the same year allows to neutralize the effect of the individual's knowledge on the exam score. To simplify the problem, I pool the observations in 2003 and 2004⁵⁰ and I assume that region-specific scoring biases are constant over the 2003-2004 period. Therefore, the estimation method is based on the assumption that any systematic score gain or loss for the same individuals in one region relative to another (chosen as a reference) reflects a variation in rating practices. I consider the following additive model:

$$S_{ijdt} = a_i + r_j + \gamma_{dt} + \epsilon_{ijdt} \quad (1)$$

where S_{ijdt} is the (written) exam score of candidate i in year t in region j that belongs to examination group d , a_i the level of knowledge of candidate i , r_j a region fixed effect⁵¹ that captures the time-invariant specificity of region j , γ_{dt} a group x year fixed effect that captures group specific shock on test score in year t (e.g. difficulty of tests in a specific examination group, learning between exam dates), and ϵ_{ijdt} an error term that captures unobserved transitory shocks on exam score (e.g. health status of a candidate) and that satisfies $E[\epsilon_{ijdt} | a_i, r_j, \gamma_d, \delta_t] = 0$.

Necessary conditions for region fixed effects r_j to capture differences in scoring practices across regions are 1) tests measure the same competencies, 2) mobility of candidates between regions is exogenous and 3) unobserved shocks on test scores ϵ_{ijdt} are independent of year, region and individual fixed effects. The first conditions is ensured by the fact that i) tests are identical within each group of regions for a given year and ii) γ_{dt} captures variation in difficulty of the tests (and learning between exam dates) over time and between groups of regions. The exogenous mobility condition is also reasonable given that candidates register in regions several months before the exam dates, which ensures that the decision to take the competition a second time (in another region) does not depend on the shocks experienced in the first competition. In addition, candidates do not choose the order in which they take the examinations, as the dates of the tests are set in advance by the administration. This condition ensures that candidates do not systematically take the exam in the first place in the region where they prefer to be recruited and where they

⁵⁰At the moment, I do not exploit the fact that some candidates are taking the competition both in 2003 and 2004.

⁵¹Because the academic region of Grenoble has changed groups between 2003 and 2004, we are able to estimate region fixed effects relative to a single reference region (Paris academic region).

put more effort.

Column (1) of Table 3.42 presents the scoring biases (in standard deviation points) of each academic region with reference to the Paris region. The rating bias between the academic regions of Créteil (most generous rating) and Lyon (least generous rating) is about 1 standard deviation point. This result suggests that the scores obtained in the primary teacher recruitment competitions are not directly comparable from one region to another.

Figure 3.22 shows the relationship between the corrected and raw exam score of teachers by region over the period 2003-2004. The two exam scores are standardized (mean 0 and variance 1) for the population of teachers recruited in 2003-2004. Figure 3.22 reveals that recruitment disparities (in terms of teacher exam score) between academic regions increase from 1.47 SD points before correction to 1.84 SD points after correction by model (1). The correlation between the raw score and the adjusted score of teachers is $\rho = 0.91$.

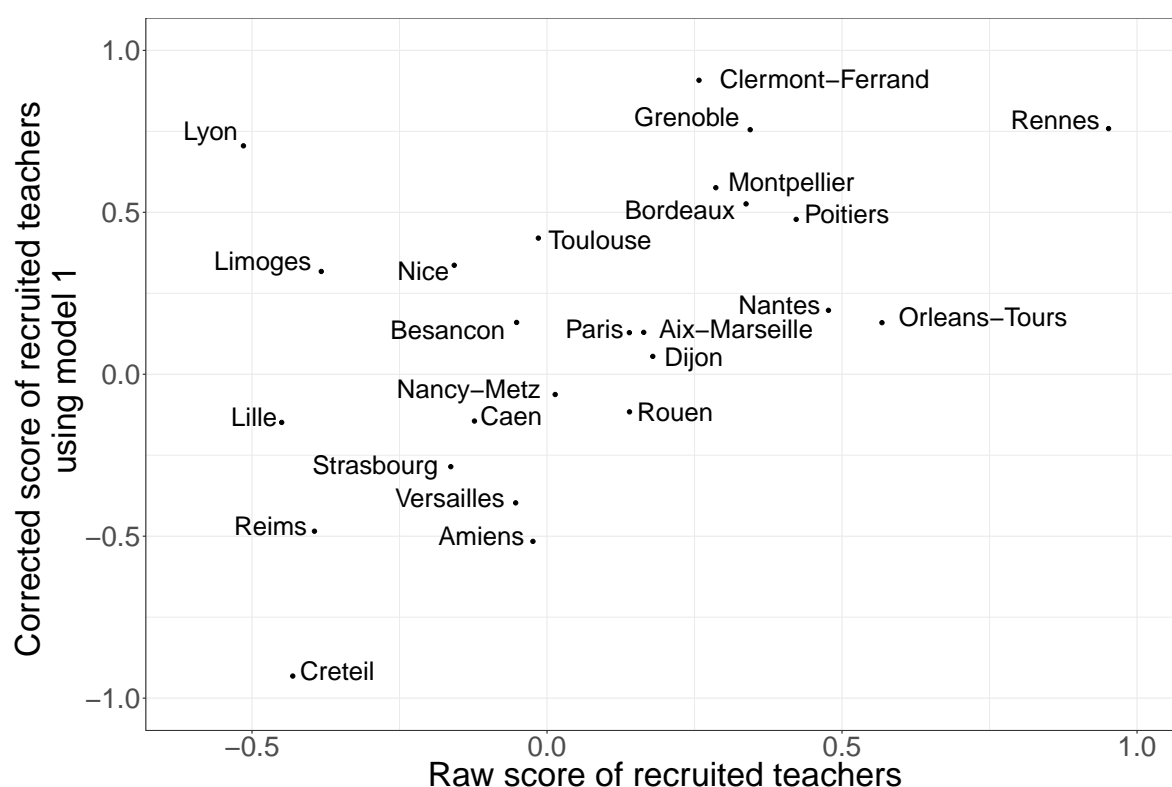


Figure 3.22 – Raw and corrected exam score of teachers by academic region in 2003-2004

Note: Each point represents an academic region in metropolitan France (excluding Corsica). The raw examination score corresponds to the examination score observed in the data. The corrected examination score is obtained by neutralizing the evaluation biases (region fixed effects) estimated from model (1). The raw and corrected scores are normalized (mean 0 and variance 1) for teachers recruited over the 2003-2013 period. This Figure shows that the raw score and the corrected score of teachers are positively correlated.

E A second method for estimating scoring biases between regions and years

The objective is to estimate how evaluation biases vary over time and across regions for the primary school teacher competition. I take advantage of the fact that 52% of new applicants fail on their first attempt and take the competition a second time, generally a year later. Most of them stay in the same academic region, but a significant fraction (10%) take the competition in two different regions. Examination topics are not necessarily identical between regions over the same year.

E.1 Descriptive evidence and preliminary analysis

Comparison of test scores between regions

In order to examine whether there are systematic differences in assessment between regions⁵², I focus on candidates who take the competition at least twice during the period 2005-2013. I consider an event study, as proposed by [Card et al. \(2013\)](#), that compares the scores of candidates who write the exam twice in the same region with those of candidates who take the exam twice in two different regions.

I start with dividing the 25 metropolitan regions into four groups that define four levels of assessment "generosity" based on the estimates presented in section [E.3](#). More precisely, candidates are expected to be assessed more generously in regions that belong to higher quartiles. [Table 3.33](#) shows the distribution of academic regions between assessment quartiles. Then, I divide the candidates into 16 groups according to the quartiles of the regions where they participate in the competition for the first and second time.

Columns (1) and (2) of [Table 3.34](#) confirm that the vast majority of candidates take the exam twice in the same quartile (and generally in the same region). Of the 10% of candidates who change region between two attempts, it is interesting to note that the majority leave a lower quartile region (potentially more severe assessment) and move to a higher quartile region (potentially more generous assessment). For example, 1.2% of candidates who take the competition for the first time in a fourth quartile region take the competition for the second time in a first quartile region. Conversely, 6.1% of applicants who take the exam for the first time in a region that belongs to quartile 1 take the competition for the second time in a region that belongs to quartile 4. The rate of inter-quartile mobility is also higher among candidates who enter the competition in the first quartile (12% change quartile the following year) than among candidates who enter the competition in the fourth quartile (3.1% change quartile the following year).

[Table 3.34](#) also shows the average scores obtained by candidates on the first attempt (column 3) and second attempt (column 4) for each of the 16 possible transitions between quartiles. To illustrate how candidates' scores vary from one region to another, [Figure 3.23.A](#) shows the average scores of candidates who participate in the competition for the first time in a region that belongs to the first (dotted lines) or to the fourth (solid

⁵²Assessment biases refer both to the heterogeneity of examiners' grading practices and to the heterogeneity of the level of complexity of tests across regions.

lines) quartile. For those candidates, Figure 3.23.A presents the average score on the first attempt (left part) and second attempt (right part). The Figure confirms that candidates' score increases on average between attempts. However, it increases much more for candidates who transition from regions of the first quartiles to regions of the last quartiles (dotted lines) than for those who remain in the same quartile (red markers). Conversely, the score of candidates who move from regions of the last quartiles to regions of the first quartiles (solid lines) increases much less between the two attempts. The Figure also shows that the score gain between attempts is similar for candidates who take the competition twice in a region of quartile 1 (0.32 points) or twice in a region of quartile 4 (0.27 points). Table 3.34 confirms a constant score gain for candidates who remain in the same quartile: the progression between attempts is about 0.29SD points on average.

To clarify the effect of transiting between regions that belong to different quartiles, I report in column (5) of Table 3.34 the average score of candidates who take part in the competition for the second time minus the learning gain between attempts (calculated from candidates who remain in the same quartile). In other words, column (5) is obtained by subtracting 0.29SD point from values in column (4). I report on the right part of Figure 3.23.B the adjusted score of column (5). Figure 3.23.B shows strong variations associated with transitions between quartiles of regions. The average score increases by 0.26SD points when candidates move from a quartile 1 region to a quartile 2 region, and increases by more than 0.40SD points when they move from a quartile 1 region to a quartile 3 or quartile 4 region. In contrast, candidates in regions of the fourth quartile lose 0.19SD points when they move to a region of the third quartile, and up to 0.32SD points when they move to a region of the first quartile. These large variations contrast with the relative stability of the scores of candidates who remain in regions that belong to the same quartile (red markers). The relative symmetry of the bonuses and losses associated with transitions appear even more clearly on Figure 3.28 where score differences between groups on the first attempt are neutralized.

The symmetry of the average score variations between quartiles suggest that an additive model that decomposes candidates' scores into individual fixed effects (knowledge), learning between attempts effects and region fixed effects would make it possible to estimate the region-specific evaluation biases.

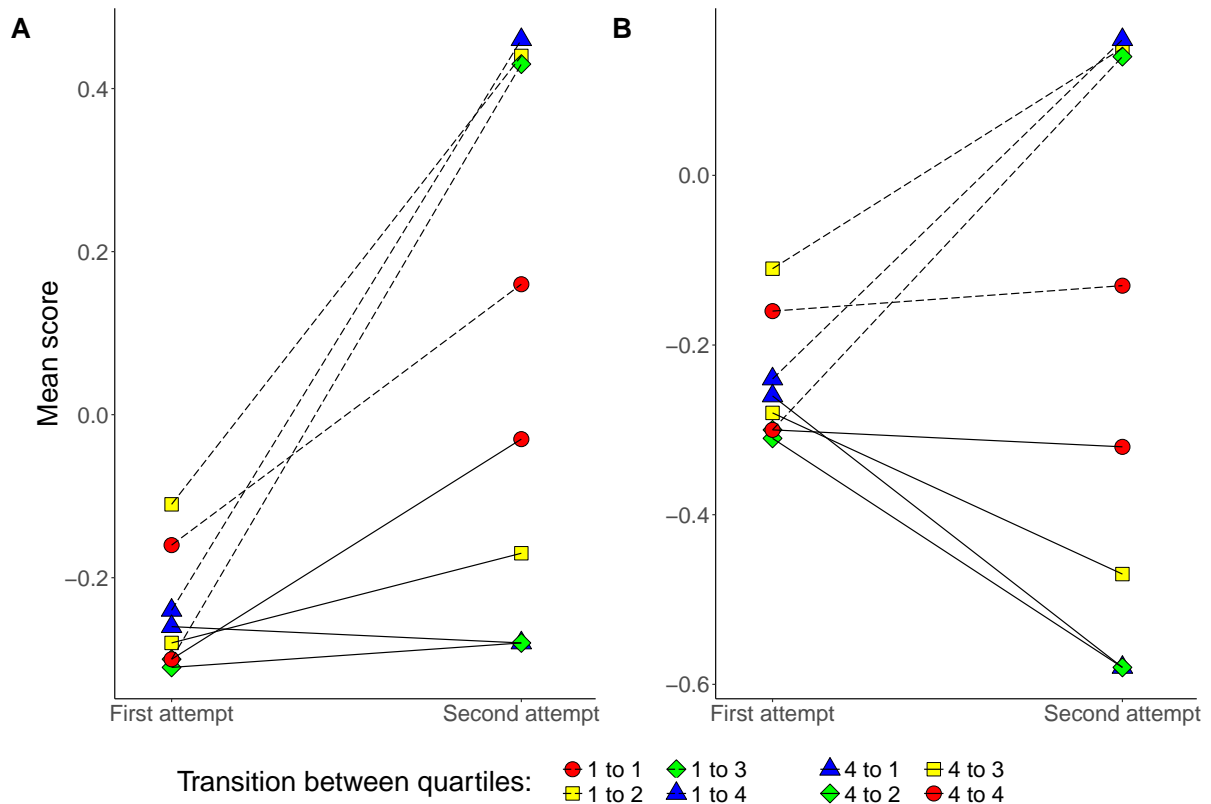


Figure 3.23 – Variation in candidates' exam score between the first and second attempts according to the quartile of the region of departure and arrival

Note: Academic regions are divided into four quartiles according to the extent of assessment bias. Quartile 1 regions are likely to have the least generous evaluation standard. Examination score of candidates are standardized (mean 0 and variance 1) over the period 2003-2013. Panel A shows the average score of candidates who participate in the competition for the first time in a region of quartile 1 or quartile 4 (first attempt) and who participate in the competition for the second time in a region of quartile 1, 2, 3 or 4 (second attempt). In panel B, I (approximately) eliminate the effect of learning by subtracting the average variation in the score of candidates who take the exam twice in the same quartile from the score obtained on the second attempt (for all candidates). Panel B allows to distinguish more clearly the effect of transitions between quartiles on candidates' scores. Average scores on the first attempt are reported in column (3) of Table 3.34. Average scores on the second attempts are reported in column (4) for panel A and in column (5) for panel B.

Comparison of test scores between years

Section C shows that candidates' scores increase with the number of times they write the exam. I study whether this variation reflects candidates' learning between attempts or changes in scoring practices over time. I focus on candidates who participate in the competition more than once and limit the analysis to the first two consecutive attempts (one year apart), which avoids attrition problems. This sub-sample consists in 70,024

candidates for whom two scores are observed over the period 2005-2013. I divide these candidates into nine cohorts based on the year they write the exam for the first time. For example, the "2005-2006" cohort includes 12,979 candidates who write the exam for the first time in 2005 and for the second time in 2006.

Table 3.32 and Figure 3.24 show the average score of applicants the first and second time they participate in the competition for each of the nine cohorts. They reveal that the exam score of applicants from cohorts 2005-06 to 2008-09 increases by 0.4SD point on average between the first and second attempts. This score gain is relatively stable for these four cohorts, as is the average score of candidates on the first attempt. However, changes are occurring starting with the 2009-2010 cohort. First, candidates who write the exam for the first time in 2009 obtain a significantly better score on their first attempt than candidates from previous cohorts. We exclude a change in assessment practices (or tests difficulty) in 2009 because the average score on the second attempt of the 2008-2009 cohort is close to the value observed in previous years. A better explanation is that candidates who write the exam for the first time in 2009 have a higher level of knowledge than candidates from previous cohorts (self-selection argument).

Figure 3.24 also shows that the score gain between attempts is much lower for the 2009-2010 cohort than for previous cohorts, and even slightly negative for the 2010-2011 cohort. One possible explanation is that the learning gain between attempts is a decreasing function of knowledge, i.e. a decreasing function of the expected score on the first attempt. However, this hypothesis is unable to explain the decrease of the exam score between the first and second attempts for candidates in the 2010-2011 cohort (compared to candidates in the 2009-2010 cohort in particular). In addition, the score gain between attempts of the 2011-2012 to 2013-2014 cohorts is similar to that of the 2005-2006 to 2008-2009 cohorts, while the first attempt score is much lower for cohorts from 2011-2012 to 2013-2014 than for cohorts from 2005-2006 to 2008-2009.

A more plausible interpretation is that the assessment is more demanding from 2010-2011, which coincides with the first (2010) and second (2011) stages of the reform that raised the level of qualification required to become a teacher in France. First, the increase in the educational level of candidates registered for the first time in 2010 may have discouraged the least qualified candidates who did not pass the competition in 2009 from re-registering in 2010. This hypothesis is compatible with the increase in the first attempt

score for the 2009-2010 cohort of candidates. Second, a more demanding assessment from 2010 onwards is compatible with the (temporary) decrease in the score gain between two attempts for the 2009-2010 and 2010-2011 cohorts.

Several mechanisms are likely to explain the change in assessment standards starting in 2010-2011. First, examiners' expectations and requirements regarding candidates' knowledge may have increased with the required level of diploma. Second, the level of reference tests used to calibrate the assessment in each region may have increased with the degree level of candidates, resulting in an overall increase in the level of assessment requirements. Third, the difficulty of tests may have increased in all regions from 2010-2011. The latter hypothesis is not confirmed either by the reviewers or by the nature of the subjects.

According to our results, the raw recruitment scores presented in Table 3.32, as well as in Figure 3.24, tend to underestimate the level of candidates' knowledge from 2010 onwards. The decrease in the average scores of candidates in the 2011-2012 to 2013-2014 cohorts, compared to those in the 2005-2006 to 2008-2009 cohorts, is thus likely to reflect partly a change in the assessment standard.

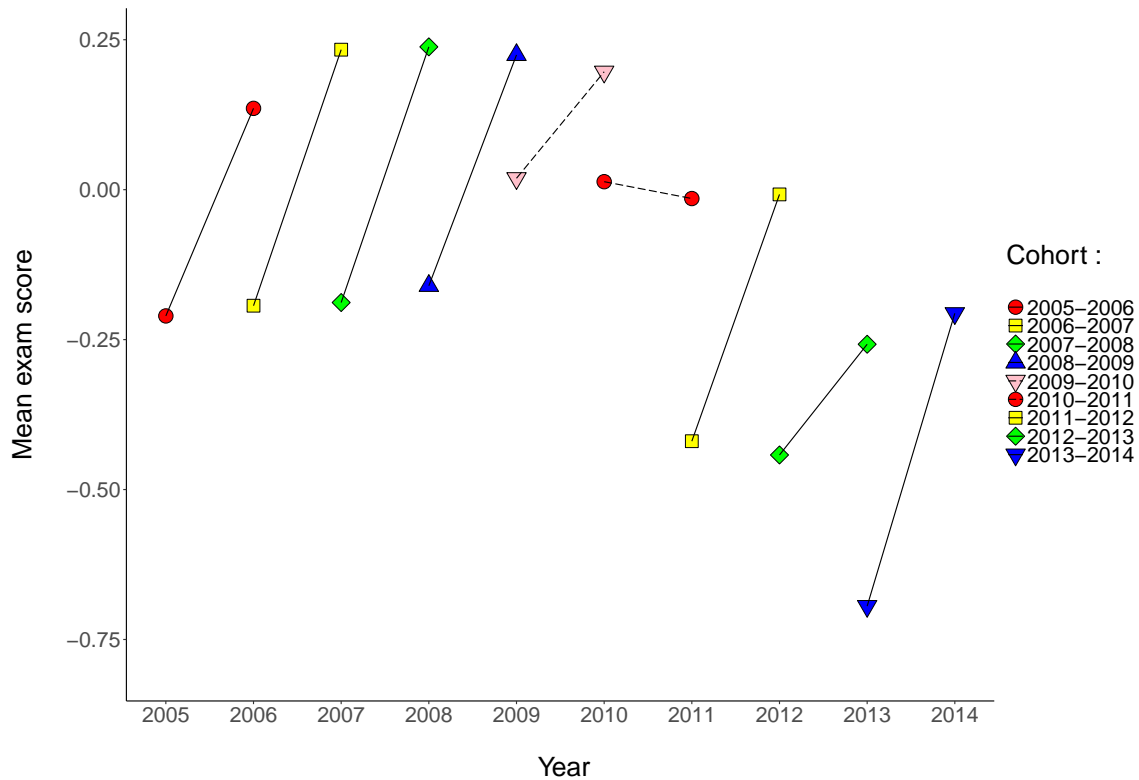


Figure 3.24 – Evolution of the average score of candidates between the first and second participation in the competition during the period 2005-2014

Note: This Figure shows how the score gain between attempts varies over the period 2005-2014. A $t/t+1$ cohort consists of candidates who write the examination for the first time in year t and for the second time in year $t+1$. The candidates' examination score is standardized (mean 0 and variance 1) over the period 2003-2014. The average scores for the first and second attempts are shown in columns (2) and (3) of Table 3.32. The number of observations per cohort is indicated in column (1).

E.2 Model

During the period 2003-2013, the subjects of the tests vary from year to year and from region to region. They may be systematically easier in some regions or years. Therefore, the objective of the method is to estimate the assessment biases for each region and each year, i.e. both the effect of the examiners' scoring practices and the difficulty of the tests specific to each region and year.

Based on the descriptive evidence in Section E.1, the model, in its simplest specification, assumes 1) that there are region-specific and time-constant evaluation biases and 2) that there are evaluation biases that vary over time and are common to all regions. This simple model also assumes that the learning gain between attempts does not depend on the regions, years and characteristics of the candidates.

According to the model's assumptions, a significant deviation from the expected learning gain between two attempts (in all regions) indicates a change in the standard of assessment between the two years. For example, if the scores of candidates who write the exam twice in the same region increase significantly between t and $t + 1$ compared to the average learning gain of previous and subsequent periods, the assessment is probably more generous from the year $t + 1$. The difference in scores between the first and second attempts for candidates who write the exam twice in the same region allows to estimate the learning gain between attempts. The, the score variation for candidates who write the examination twice in two different regions allows to estimate the difference in assessment practices between regions, after eliminating the learning gain and year effect by double difference. In other words, if candidates systematically improve their scores more when they move to a given region, the method concludes that the assessment is probably more generous in that region.

For the method to be valid, tests must measure the same competencies, mobility of candidates between years and regions must be exogenous, and unobserved shocks on test scores must be independent of year, region and individual effects. I detail the assumptions and the method used in the following sections.

Main specification

Let's note S_{ijtk} the exam score of individual i on his/her k^{th} attempt in region j and year t . I consider the additive model (2) $S_{ijtk} = a_i + b_k + r_j + \delta_t + \epsilon_{ijtk}$, where a_i represents the level of knowledge of individual i , b_k the score gain associated with the k^{th} attempt ("learning gain"), r_j a region fixed effect that captures the time-invariant evaluation bias in region j , δ_t a year fixed effect that captures the evaluation on year t and ϵ_{ijtk} an error term that captures the transitory shocks on exam score (difficulty of a test subject, health shock for candidate i) and that satisfies $E[\epsilon_{ijtk}|a_i, b_k, r_j, \delta_t] = 0$. For the model to be identified, I impose $b_1 = 0$ (the first attempt is taken as a reference), $\delta_{2003} = 0$ (2003 is the reference year) and $r_{Paris} = 0$ (Paris is the reference academic region).

I note N^* the number of scores (total number of observations), N the number of unique candidates, K the maximum number of attempts over 2003-2013, J the number of regions and T the number of years. In matrix notation, model (2) can be written as follows: $S = Ia + Lb + Mr + P\delta + \epsilon$, where S is a $N^* \times 1$ vector of scores, I a $N^* \times N$

matrix of candidate indicators, L a $N^* \times K$ matrix of attempt indicators, M a $N^* \times J$ matrix of region indicators, P a $N^* \times T$ matrix of year indicators and ϵ a $N^* \times 1$ vector of unobserved shocks.

Validity conditions

The parameters of model (2) are correctly estimated by ordinary least squares (OLS) under validity conditions that I present in the following paragraphs.

A connected set of regions and years

The individual fixed effect a_i (level of knowledge) is not directly observable in the data but can be removed by difference. To do this, I restrict the estimation to the candidates who write the exam at least twice during the 2003-2013 period. The mobility of applicants between regions and years allows to estimate the coefficients b_k , r_d and δ_t of model (2).

To ensure that the effects of regions (respectively years) are estimated relative to the same reference, a necessary condition is that the regions and years are linked together, directly or indirectly, by the mobility of at least one candidate. This condition is met in this study, but the intensity of connections between regions is very heterogeneous. As already mentioned in section B, regions outside metropolitan France are weakly linked to each other and to other regions (recruitment is very local and the number of candidates is relatively low in these regions). Therefore, they are excluded from the analysis⁵³. The main sample includes 110,682 unique candidates, i.e. about 95% of the candidates who wrote the primary school teacher recruitment examination at least twice during the period 2003-2013.

Exogeneity assumption: unobserved shocks and mobility decision

The orthogonality between the error term and the variables of model (2) is a necessary condition for OLS consistency, which means in matrix formulation that $E[I'\epsilon] = 0_{N^* \times 1}$, $E[L'\epsilon] = 0_{K \times 1}$, $E[M'\epsilon] = 0_{J \times 1}$, $E[P'\epsilon] = 0_{T \times 1}$. In other words, the unobserved shocks ϵ on exam scores must be independent of the individual knowledge a_i , the number of attempts k , the year t and the region j where the individual i takes the exam. Estimates are unbiased if the positive and negative shocks received each year, in each region and by each candidate, cancel each other out on average over the period considered.

⁵³The effects of these regions would otherwise be estimated from a small number of observations

A credible source of bias stems from the correlation between the error term and the dummy variables that indicate the number of times candidates write the exam. More specifically, the hypothesis of exogeneity is violated if the mobility of candidates (i. e. the decision to take the examination several times and the choice of regions) depends on the shocks received in the past. A plausible hypothesis is that candidates who receive a negative shock in year t (e. g. sick candidate) are more likely to write the exam in year $t+1$. As a result, the model tends to overestimate the learning effect between attempts⁵⁴. However, if the bias in estimating learning gains is constant over time, the estimation of year fixed effects (with reference to 2003) is unbiased.

Another threat to identification appears if candidates who receive a negative shock in year t are more likely to take the competition the following year in regions where assessment practices are more generous (e. g., regions with higher average success rates). In this case, the model would tend to overestimate the fixed effect of these regions, i. e. the generosity of assessment practices in these regions. It is partly possible to test this second hypothesis by examining Figure 3.23.B. If candidates who experience negative shocks on their first attempt are more likely to move to a higher quartile region, we should note that the score gain increases more for candidates who move from a lower quartile to a higher quartile than for candidates who move in the opposite direction. In other words, an asymmetry of mobility decisions should result in an asymmetry on Figure 3.23.B. However, Figure 3.23.B shows that the variation in score associated with inter-quartile mobility is fairly symmetrical (which is confirmed by Table 3.34). This result suggests that the potentially endogenous mobility of candidates is not a major source of bias in estimating region fixed effects.

Misspecifications

A first source of misspecification in model (2) comes from the assumption that year and region effects are additive and separable. Instead, we could consider model (2bis) $S_{ijtk} = a_i + b_k + r_{jt} + \epsilon_{ijtk}$, where the effects of regions and years interact with each other

⁵⁴A Tobit model would not solve this problem of endogeneity, because the scores of candidates who take the competition several times are neither right- nor left-censored. Indeed, it is the final score (weighted sum of the written and oral test scores) that determines recruitment, not the score obtained in the written tests. If S_{ijt1}^w is the written score on the first attempt, S_{ijt1}^o the oral score on the first attempt and δ_{tj} the recruitment threshold in year t in region j , candidates who participate in the competition for the second time necessarily satisfy $a S_{ijt1}^w + b S_{ijt1}^o < \delta_{tj}$ i.e. $a \epsilon_{ijt1}^w + b \epsilon_{ijt1}^o < f(\delta_{tj})$, but not necessarily $\epsilon_{ijt1}^w < g(\delta_{tj})$.

(r_{jt}). In practice, the estimation of models (2) and (2bis) leads to very similar results (the correlation of average region effects is $\rho = 0.985$ and the correlation of average year effects is $\rho = 0.999$). The scores of the candidates corrected by the models (2) and (2bis) are also highly correlated ($\rho = 0.99$). Therefore, in the following sections, we assume that regions and years are separable as in model (2).

A second source of misspecification comes from the assumption that the learning gain b_k between attempts is constant over time in all regions and for all applicants. However, if the learning gain varies with the initial knowledge a_i of candidates, estimates are likely to be biased. An alternative specification could be model (3) $S_{ijtk} = a_i + b_k(ai) + r_j + \delta_t + \epsilon_{ijtk}$, with $b_k(ai)$ a function of the level of competence a_i .

Table 3.34 and Figure 3.23 suggest that the best candidates (high values of a_i) tend to self-select in the regions of the first quartiles. In particular, Figure 3.23 shows that candidates who write the exam twice in a quartile 1 region have on average a higher score than those who write the exam twice in a quartile 4 region. If the learning gain between attempts varies with a_i , we should observe different score gains for candidates who take the exam twice in the same region depending on their quartile.

Figure 3.25 shows that the score gain between attempts decreases with the region's quartile, suggesting that the learning gain is an increasing function of the initial knowledge a_i . Variations in learning gain between quartiles are rather moderate (+0.32 SD point for the first quartile, +0.35 SD point for the second quartile, +0.28 SD point for the third quartile and +0.27 SD point for the fourth quartile).

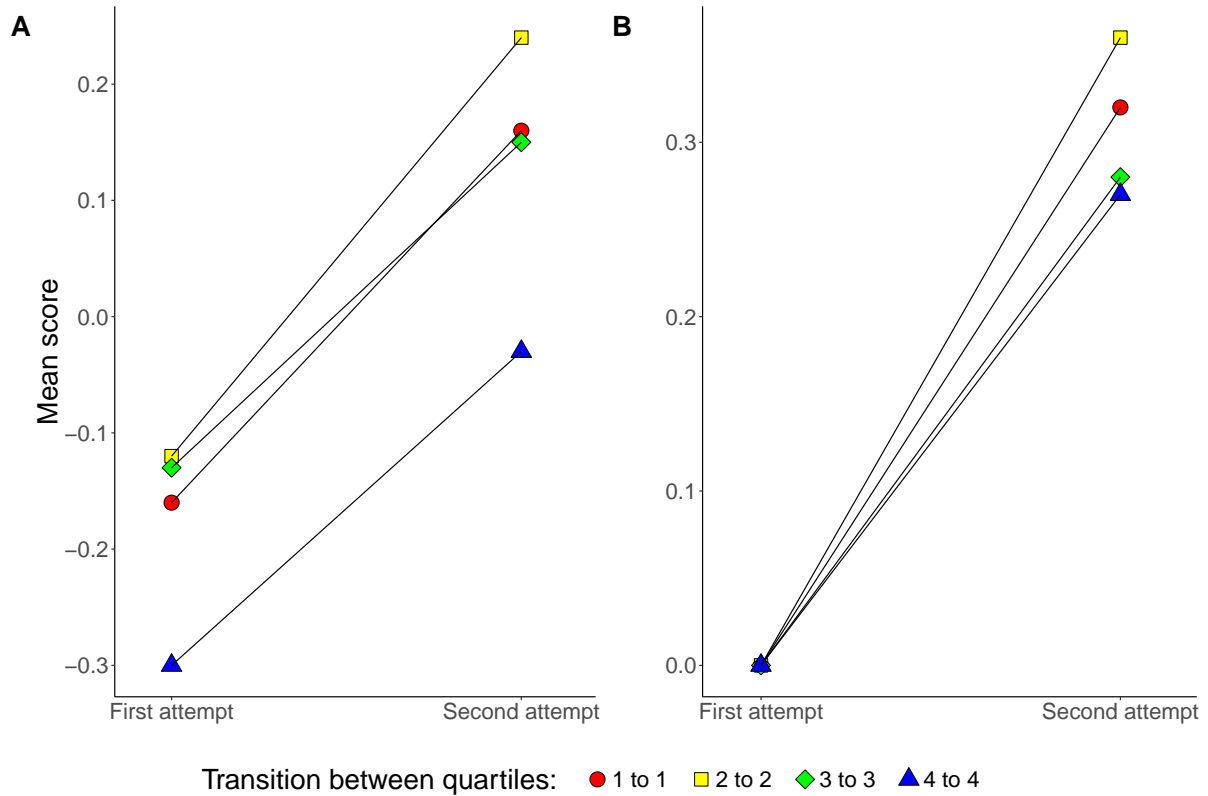


Figure 3.25 – Average score gain between the first two attempts for candidates who participate in the competition twice in the same region quartile

Note: This Figure shows how the score gain between the first two attempts for candidates who take the exam twice in the same region quartile varies according to that quartile. The candidates' examination score is standardized (mean 0 and variance 1) over the period 2003-2013. In Panel A, the average score on the first and second attempts is based on raw data. In panel B, I subtract the score of the first attempt from the score of the first and second attempts so that only the score gain between attempts is visible.

It is important to carefully examine the relationship between learning and candidates' level of knowledge, as the 2010-2011 reform may have changed candidates' average competencies. If model (2) is incorrectly specified, the estimates of the year fixed effects will be biased. To simplify, I suppose that the relationship between learning gain between attempt and knowledge can be written as $b_k(ai) = b_k^0 + b_k^1 a_i$, with $b_1^0 = b_1^1 = 0$. Model (2) is a special case of model (3) where $\forall k, b_k^1 = 0$. To identify the specification that best matches the data, I test the null hypothesis $H0 : b_k^1 = 0$.

According to model (3), the score of candidate i in region j on his/her first attempt writes $S_{i,j,t,k=1} = a_i + r_j + \delta_t + \epsilon_{i,j,t,k=1}$ (a), while the score on his/her second attempt a year later in region j' writes $S_{i,j',t+1,k=2} = a_i + (b_2^0 + b_2^1 a_i) + r_{j'} + \delta_{t+1} + \epsilon_{i,j',t+1,k=2}$ (b). I combine (a) and (b) to eliminate the individual competence a_i and I get a testable

relationship: $S_{i,j',t+1,k=2} - S_{i,j,t,k=1} = b_2^0 + b_2^1 S_{i,j,t,k=1} + (r'_j - r_j - b_2^1 r_j) + (\delta_{t+1} - \delta_t - b_2^1 \delta_t) + \nu_{i,j,j',t,t+1,k=1,k=2}$ (c), with $\nu_{i,j,j',t,t+1,k=1,k=2} = \epsilon_{i,j',t+1,k=2} - \epsilon_{i,j,t,k=1} - b_2^1 \epsilon_{i,j,t,k=1}$

A limitation of the relationship (c), however, is that the score at the first attempt $S_{i,j,t,k=1}$ is measured with error, which generates an estimation bias. By construction, $Corr(S_{i,j,t,k=1}, \epsilon_{i,j,t,k=1}) \neq 0$, so $Corr(S_{i,j,t,k=1}, \nu_{i,j,j',t,t+1,k=1,k=2}) \neq 0$ and \hat{b}_2^1 underestimates the coefficient b_2^1 . A classical method to overcome this type of bias is to use an instrumental variable that is correlated with $S_{i,j,t,k=1}$ but orthogonal to the error term $\nu_{i,j,j',t,t+1,k=1,k=2}$. In other words, a valid instrument must be correlated with the level of knowledge a_i but not with the error term $\epsilon_{i,j,t,k}$.

Table 3.35 shows that the sex and age of candidates are significantly correlated with the scores on the first attempt $S_{i,j,t,k=1}$. The fact that men and women obtain different results in scientific and literary tests is a well-known fact that we have developed in detail in Chapter 1. In Chapter 1, we have also shown that women consistently outperform men on written examinations for primary school teacher recruitment in France, both on scientific and literary tests (see Table 8 in Chapter 1). The fact that age is correlated with written test scores may be due to a decrease in academic knowledge and examination habits over time after graduation. Since the majority of candidates are under 30, it is also possible that age may reflect school delay and reduced mastery of academic knowledge. But to be valid instruments, sex and age must also be uncorrelated with the unobserved shocks $\epsilon_{i,j,t,k}$. This means that on average, candidates (who take the exam several times) receive similar shocks the first time they take the exam, regardless of their gender and age. Since written exams are anonymous, it is unlikely that examiners are biased according to the age or gender of the candidate. However, candidates may be differently sensitive to the difficulty of a test subject depending on their gender or age. Table 3.36 shows that the degree level and specialization field (humanities, sciences, social sciences, sports and others) vary according to the sex and age of the candidates. In particular, men are more likely to have a master's degree in a scientific or sport discipline than women. Older candidates are more likely to have a master's degree in a literary discipline. Since the competition includes several tests in mathematics, science, history and literature, a shock on the difficulty of a test subject in a given discipline may affect candidates differently depending on their specialization and level of qualification, and therefore according to their age and gender. Therefore, I include the level of diploma and field of specialization

as control variables when I instrument $S_{i,j,t,k=1}$ by gender and age of candidates in model (c).

Table 3.37 presents the estimate of the coefficient b_2^1 without instrument (columns (1) to (3)) and using age and sex as instrumental variables (columns (4) to (6)). The instruments are strong (F-statistic of weak identification = 561 \gg 30) and pass the Sargan test of joint exogeneity (p-value = 0.94), which means that it cannot be rejected that they are valid instruments assuming that at least one of them is exogenous.

The last three columns suggest that when the score on the first attempt increases by 1 SD point, the score gain between the first two attempts increases significantly by 0.22 SD point. The sign of the relationship between knowledge and learning between attempts obtained by the instrumental variable method is consistent with what we observed on data aggregated by quartiles (Figure 3.25). This result rejects the null hypothesis that model (2) is correctly specified and rather supports model (3). I repeat the procedure to check if the score gain between the second and third attempt b_3^1 depends on a_i . I present the results in Table 3.38. I find no evidence supporting that the score gain between the second and third attempt varies with the level of competencies: $b_3^1 = 0$. I find a similar result for the score gain between the third and fourth, and between subsequent attempts.

Alternative specification

Evidence presented in the previous section suggest estimating region and year fixed effects using model (3) $S_{ijtk} = a_i + b_k(a_i) + r_j + \delta_t + \epsilon_{ijtk}$ instead of model (2). To simplify, I assume that the relationship between the learning gain between attempts and candidate's knowledge is linear: $b_k(a_i) = b_k^0 + b_k^1 a_i$, with $b_1^0 = b_1^1 = 0$ by convention.

The unobserved individual fixed effect a_i can be eliminated by difference, which leads to an equivalent expression of model (3): $S_{ijtk} - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{S_{ijtk}} = b_k^0 - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{b_k^0} + r_j - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{r_j} + \delta_t - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{\delta_t} + \epsilon_{ijtk}$, with $\overline{S_{ijtk}} = \overline{b_k^0} + a_i (1 + \overline{b_k^1}) + \overline{r_j} + \overline{\delta_t}$ the average score of candidate i and $\overline{x} = \frac{1}{n} \sum_i^n x_i$.

In matrix notation, with the notations of section E.2, it can be written:

$$S - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{S} = [L^0 - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{L^0}] b^0 + [M - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{M}] r + [P - \left(\frac{1+b_k^1}{1+b_k^0}\right) \overline{P}] \delta + \epsilon \quad (3)$$

with L^0 a $N^* \times K$ vector, b^0 a $K \times 1$ vector and $\bar{X} = \begin{bmatrix} \overline{row1(X)} & \dots & \overline{row1(X)} \\ \overline{row2(X)} & \dots & \overline{row2(X)} \\ \dots & \dots & \dots \\ \overline{rowN(X)} & \dots & \overline{rowN(X)} \end{bmatrix}$

I estimate model (3) by replacing the parameters b^1 with their estimated values \hat{b}^1 obtained in Section E.2: $\hat{b}_2^1 = 0.22$ and $\hat{b}_k^1 = 0 \quad \forall k \geq 3$.

In the next section, I compare the estimates of the year and region fixed effects obtained using models (2) and (3). Since the estimates of b_k^1 are based on the assumption that age and sex are valid instrumental variables (exogeneity condition), which can be discussed, I will also present the results obtained for different values of b_k^1 in the last section.

E.3 Results

To estimate model (2) and model (3), I consider the 110,682 individuals (half of the unique candidates) who take the exam at least twice over the period 2003-2013. On average, there are 2.8 observations (written scores) per candidate.

Comparison of models (2) and (3)

Table 3.29 summarizes the estimation results of models (2) and (3). Both models have the same explanatory power (R square = 0.73) and lead to a very similar decomposition of the variance of the exam score. Due to the large number of parameters (110,726) relative to the number of observations (312,154), the variance of the residual underestimates the variance of the error term. After correction, the adjusted R-squared is 0.57 and the variance of the error term represents 35% of the total variance of the candidates' score. This means that there is a relatively large measurement error in assessing the knowledge of candidates for the primary school teacher recruitment competition⁵⁵.

The fixed effects estimated by models (2) and (3) are highly correlated: $\rho = 0.995$ for individual fixed effects (competencies a_i), $\rho = 0.998$ for region fixed effects, $\rho = 0.893$ for year fixed effects and $\rho = 0.838$ for the average effect of learning between attempts

⁵⁵By comparison, Boyd et al. (2013) estimated a measurement error of about 25% for the English Language Arts (ELA) test and 20% for the math test taken by third to eighth grade students in New York State. When teacher quality is assessed on the basis of student achievement in standardized tests (value-added models), the measurement error is much greater, about 60% (Rothstein, 2015).

($\text{corr}(\hat{b}, \hat{b}^0)$). Table 3.29 reveals that most of the explained variance of the written score comes from individual effects. In particular, differences in assessment practices over time and across regions explain a small share of the total variance. Tables 3.39, 3.40 and 3.41 compare the values of the region, year and learning fixed effects estimated by models (2) and (3). Table 3.41 confirms that candidates significantly improve their score between the first and second attempt, and to a lesser extent between the second and third attempt. However, the score gain between subsequent attempts is no longer significant.

Table 3.29 – Variance decomposition of the exam score over the period 2003-2013

	Model 2		Model 3	
Number of parameters				
Candidate fixed effects	110,681		110,681	
Region fixed effects	24		24	
Year fixed effects	10		10	
Learning fixed effects	10		10	
Model fit				
Number of observations	312,154		312,154	
R-squared	0.725		0.724	
Adjusted-R2	0.573		0.573	
RMSE	0.591		0.591	
	Value	Percent of Total	Value	Percent of Total
Variance of score	0.991	100%	0.991	100%
Variance decomposition				
Var(Individual FE + learning FE)	0.787	79%	0.78	79%
Var(Region FE + Year FE)	0.07	7%	0.057	6%
2 Cov(Individual FE + learning FE, Region FE + Year FE)	-0.09	-9%	-0.071	-7%
Var(Residual)	0.225	23%	0.226	23%
Var(Error)	0.349	35%	0.35	35%

Note: The variable explained in model (2) and (3) is the exam score of candidates for the external primary teacher recruitment competition. The exam score is standardized (mean 0 and variance 1) over the period 2003-2013. FE stands for "fixed effects".

Source: Ministry of Education (DEPP).

Sorting of candidates

Table 3.30 shows the correlation between individual fixed effects a_i , region fixed effects (column 1) and year fixed effects (column 2) estimated from models (2) and (3) respectively. It confirms that candidates with lower levels of knowledge (low values of a_i) tend to self-select in regions where the assessment is more generous ($\rho = -0.24$). For each region of metropolitan France, I report on the Y axis of Figure 3.29 the average individual fixed effects over the 2003-2013 period. On the X axis, I indicate the generosity of the evaluation in the region (defined as the score bonus that candidates receive in this region). Since only the ranking is important within each region, candidates have

no reason to self-select based on the generosity of the assessment process in the region. Another explanation is that examiners may tend to adopt a more generous assessment as candidates' skill level decreases.

Several mechanisms are likely to explain that the candidates' level of knowledge differs from one region to another. First, the recruitment of primary school teachers is very local (more than 80% of candidates apply and are recruited in the region where they live). A first possibility is that the level of education in the regions (Broccolichi et al., 2007; DEPP, 2014) influences the candidates' level of knowledge. A second possibility is that individuals take into account the success rate in the region and their own skills when deciding to participate in the recruitment competition (see Chapter 2 of this thesis). Figure 3.30 shows that the more selective a region is (lower number of positions per candidate), the higher the average skills of candidates. The heterogeneity of candidates' level of knowledge between regions may lead to heterogeneity in assessment practices between regions.

Model (2) suggests a slightly positive correlation between individual fixed effects and year effects, but this result is contradicted by model (3). As model (3) is better specified, it is likely that low-skilled candidates are less likely to write the exam in years when the assessment is more demanding. This was also suggested in Figure 3.24 which showed that the low-skilled candidates in 2009-2010 were less likely to take the competition a second time in 2010-2011. The increase (actual or anticipated) in candidates' skill levels as a result of the increase in diploma levels starting in 2010-2011 may have discouraged the low-skilled candidates from participating in the competition a second time (in 2010 and in 2011).

Table 3.30 – Correlation between individual fixed effects, region fixed effects and year fixed effects estimated by models (2) en (3)

	(1)	(2)
	Region FE	Year FE
Model 2		
Individual FE	-0.2489***	0.0044**
Model 3		
Individual FE	-0.2419***	-0.0483***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis are clustered at the region level.

Note: This Table shows the correlation between the individual fixed effects and region fixed effects estimated from models (2) and (3) in column (1), and the correlation between the individual fixed effects and year effects estimated from models (2) and (3) in column (2).

Source: Ministry of Education (DEPP).

Score variation over time and between regions

From models (2) and (3), we can predict the score that candidates would obtain in the absence of heterogeneous rating standards over time and across regions. I calculate two corrected scores from the estimates of models (2) and (3), which I standardize respectively for candidates and teachers over the period 2003-2013.

Figure 3.26 shows, at the regional level, the correlation between the raw score and the corrected score of teachers over the period 2003-2013. The two exam scores are positively correlated ($\rho = 0.92$). However, the corrected score is more dispersed across regions than the raw score. The difference in the average recruitment score between the Créteil region (lowest recruitment score) and the Rennes region (highest recruitment score) rose from 1.41 SD points before correction to 1.85 SD points after correction by model (3). This result suggests that the heterogeneity of grading practices across regions tends to underestimate the regional differences in recruitment measured by the written examination of primary school teachers.

Figures 3.32 and 3.27 show how the raw score (red points) of candidates and teachers changes over time. The Figures also show the evolution of the corrected score by model (2) (blue squares) and by model (3) (green diamonds). The Figures reveal that biases in assessment standards from 2010 tend to overestimate the decline in the accreditation score of candidates and teachers from 2010 onwards. The score corrected by model (2) provides the most optimistic evolution of the average accreditation score. Model (3) leads to an intermediate situation: the corrected score decreases more than with model (2) but

less than the raw score. These results suggest that the corrected score obtained by model (2) can be used to estimate a lower limit of the (negative) effect of the 2010-2011 reform on test scores.

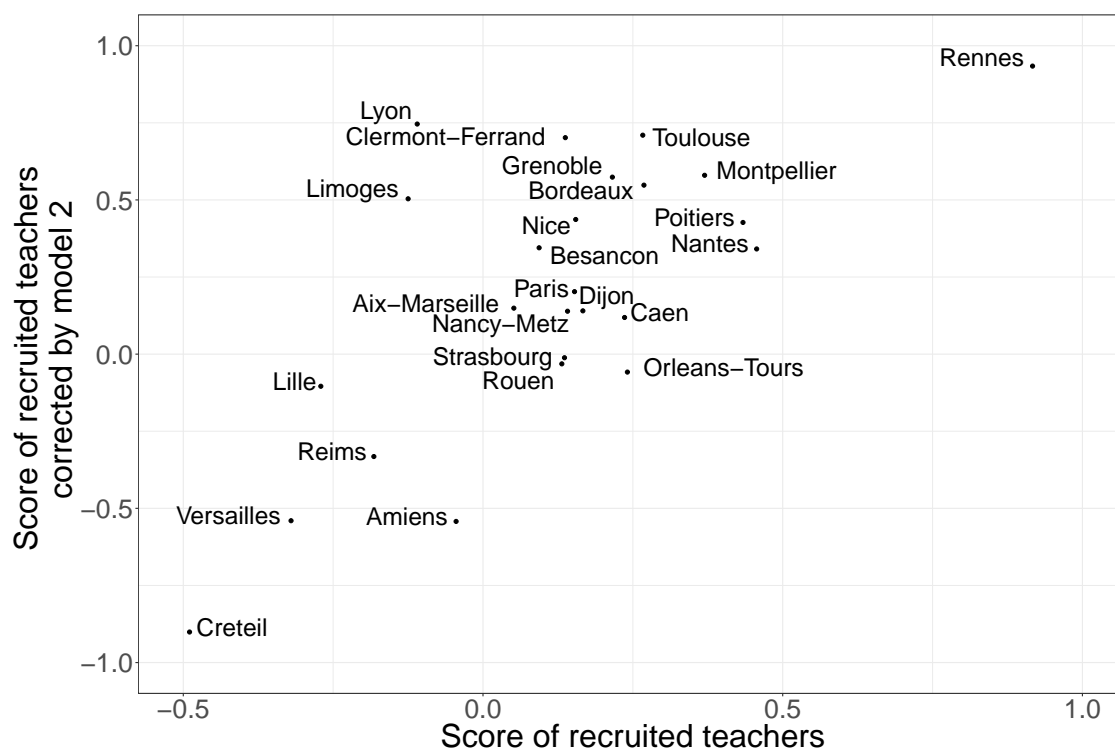


Figure 3.26 – Raw and corrected exam score of teachers by academic region over the period 2003-2013

Note: Each point represents an academic region in metropolitan France (excluding Corsica). The raw examination score corresponds to the average examination score observed in the data for each region over the period 2003-2013. The corrected examination score is obtained by neutralizing the evaluation biases (region fixed effects) estimated from model (2). The raw and corrected scores are normalized (mean 0 and variance 1) for teachers recruited over the 2003-2013 period. This Figure shows that the raw score and the corrected score of teachers are positively correlated.

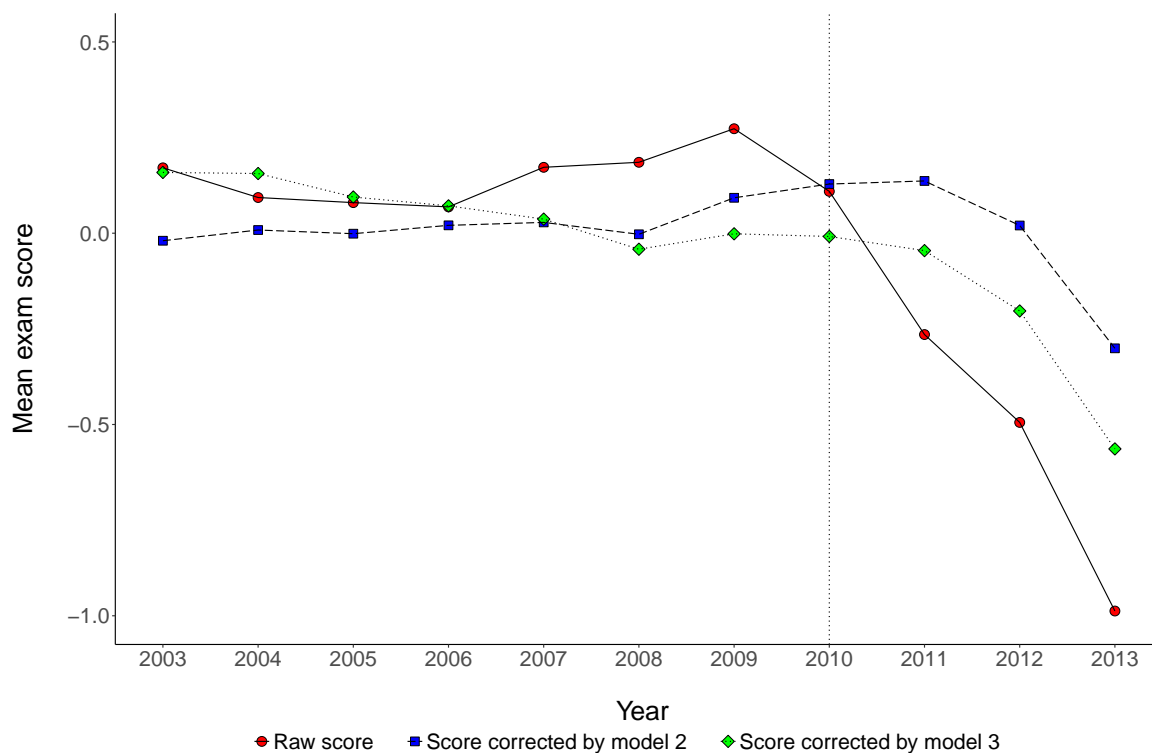


Figure 3.27 – Evolution of teachers’ raw and corrected scores on recruitment examinations at national level over the period 2003-2013

Note: Academic regions in metropolitan France (excluding Corsica). The raw examination score corresponds to the average examination score of teachers observed in the data over the period 2003-2013. The corrected examination scores are obtained by neutralizing the evaluation biases (region fixed effects and year effects) estimated from models (2) and (3). The raw and corrected scores are normalized (mean 0 and variance 1) for teachers recruited over the 2003-2013 period. This Figure shows that the teacher examination scores corrected by models (2) and (3) decrease significantly less after the 2011 reform than the raw exam score.

F Robustness checks

To test the robustness of the results obtained with model (3), Figure 3.33 shows the evolution of the corrected exam score of teachers for different values of β_2^1 (correlation between knowledge and learning between attempts). The Figure shows that the corrected score obtained using model (2) (i.e. assuming that $\beta_2^1 = 0$) provides a lower bound of the (negative) effect of the 2011 reform on teacher recruitment score as long as learning between attempts is an increasing function of the level of knowledge (i.e. as long as $\beta_2^1 > 0$). This assumption seems reasonable given the results presented in section E.2.

Table 3.42 compares the estimates of region fixed effects r_j obtained using model (1)

over the 2003-2004 period (column 1), model (2) over the 2005-2013 period (column 2) and model (2) over the 2003-2013 period (column 3). Column (1) exploits the fact that candidates can take the competition in two different regions in the same year in 2003 and 2004. Column (2) exploits the fact that candidates can take the examination several times in different years. Column (3), our main specification, exploits the fact that candidates can take the exam twice in the same year (in 2003 or 2004) or several times in different years. Table 3.42 shows that estimates of region fixed effects are close despite differences in the samples and estimation methods considered. The correlation between the region fixed effects is $\rho = 0.88$ for columns (1) and (2), $\rho = 0.95$ for columns (2) and (3) and $\rho = 0.96$ for columns (1) and (3).

These results show the robustness of the estimates of region fixed effects over the period 2003-2013 by models (1), (2) and (3). They also support the hypothesis that regional effects (evaluation standards) are roughly constant over time.

Tables

Table 3.31 – Number of candidates, written score and success rate in the primary school teacher recruitment examination over the period 2005-2013, at the national level and by academic region

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of candidates	Average written score	Standard-deviation	Written score of the last candidate eligible to oral examinations	Share of candidates eligible to oral examinations (%)	Share of teachers recruited (%)
Metropolitan France	329304	0	1	0.137	43.4	21.4
Academic regions						
Limoges	3277	-0.21	1.028	0.088	38.2	18.5
Lille	26287	-0.156	0.948	0.083	40.1	19.4
Clermont-Ferrand	6260	-0.095	0.998	0.199	39.6	17.7
Lyon	15163	-0.094	0.991	0.216	41.1	18.2
Creteil	24832	-0.047	0.984	-0.523	62.6	36
Aix-Marseille	15414	-0.043	0.979	0.357	33.6	16.4
Toulouse	14602	-0.043	1.056	0.502	32.5	15.4
Nancy-Metz	12733	-0.025	0.966	0.363	36	17.3
Besançon	6747	0.019	0.95	0.308	35.9	17.1
Versailles	30216	0.02	0.946	-0.277	59	32
Rouen	8633	0.032	1.016	0.209	43.6	22.9
Paris	6934	0.039	1.116	0.001	49.1	24.7
Bordeaux	16856	0.041	0.975	0.499	33.3	17
Montpellier	13597	0.051	1.043	0.066	30.9	15.6
Reims	6585	0.056	0.86	-0.023	55.2	22.2
Nice	9597	0.068	0.975	0.407	35.2	18.4
Amiens	9786	0.068	0.957	0.071	49.4	25.6
Grenoble	14467	0.068	0.985	0.17	46.6	19.6
Dijon	7199	0.095	0.959	0.128	50	23.1
Orleans-Tours	11506	0.148	0.956	0.114	52.6	22.1
Strasbourg	8656	0.171	0.958	0.349	48.4	22
Poitiers	8865	0.175	1.037	0.352	39.8	19.2
Caen	7118	0.201	0.953	0.288	44.4	17.3
Nantes	14975	0.217	1.016	0.415	43.5	19.1
Rennes	10077	0.224	1.109	0.725	32.5	16.2

Note: Years 2003 and 2004 are excluded because candidates were exceptionally allowed to take the exam twice in the same year (in two different regions). The period after 2013 is also excluded because the written tests were modified in 2014 and no longer measure the same knowledge as for the period 2003-2013. The exam score is standardized (mean 0 and variance 1) at metropolitan level (excluding Corsica) over the period 2005-2013. Columns (2) and (3) indicate the mean and standard deviation of the exam score of candidates by academic region. Column (4) reports the average written score of the last candidate recruited over the period 2005-2013. For each academic region, column (5) indicates the share of candidates eligible for oral examinations (second stage of recruitment) and column (6) indicates the share of candidates recruited (success rate).

Source: Ministry of Education (DEPP). *Scope:* Metropolitan France.

Table 3.32 – Average exam score the first and second time candidates participate in the recruitment competition during the period 2005-2014

	(1)	(2)	(3)
Cohort	Observations	First attempt	Second attempt
2005-2006	12975	-0.21	0.136
2006-2007	11736	-0.193	0.233
2007-2008	10484	-0.188	0.238
2008-2009	9499	-0.16	0.225
2009-2010	9620	0.019	0.197
2010-2011	6229	0.013	-0.015
2011-2012	841	-0.419	-0.008
2012-2013	3496	-0.442	-0.258
2013-2014	3199	-0.694	-0.206

Note: A $t/t+1$ cohort consists of candidates who write the examination for the first time in year t and for the second time in year $t+1$. Column (1) indicates the number of candidates in each cohort. For each cohort, columns (2) and (3) indicate the scores the first time and second time candidates write the competition. The exam score is standardized (mean 0 and variance 1) over the period 2005-2014. The values of columns (2) and (3) are shown in Figure 3.24.

Source: Ministry of Education (DEPP).

Table 3.33 – Academic regions divided into quartiles according to the generosity of the evaluation standard

First quartile	Second quartile	Third quartile	Fourth quartile
Lyon	Nice	Poitiers	Strasbourg
Limoges	Besançon	Paris	Reims
Clermont-Ferrand	Montpellier	Dijon	Rouen
Toulouse	Lille	Nancy-Metz	Versailles
Grenoble	Aix-Marseille	Nantes	Orleans-Tours
Bordeaux	Rennes	Caen	Créteil
			Amiens

Note: Academic regions are divided into four quartiles according to the extent of assessment bias. Quartile 1 regions are likely to have the least generous evaluation standard. Estimates of evaluation bias are based on the results presented in section 4.6.

Source: Ministry of Education (DEPP).

Table 3.34 – Average score of candidates the first and second time they participate in the competition according to the quartile of the region of departure and arrival

Transition between quartiles of region	(1)	(2)	(3)	(4)	(5)
	Observations		Mean exam score		
	N	%	First attempt	Second attempt	Second attempt corrected from learning gain
1 to 1	14059	87.2	-0.16	0.16	-0.13
1 to 2	667	4.1	-0.11	0.44	0.15
1 to 3	419	2.6	-0.3	0.43	0.14
1 to 4	983	6.1	-0.24	0.46	0.16
2 to 1	652	4.4	-0.07	0.29	0
2 to 2	12536	84.9	-0.12	0.24	-0.06
2 to 3	441	3	-0.16	0.33	0.04
2 to 4	1143	7.7	-0.28	0.4	0.11
3 to 1	394	2.9	-0.15	0.06	-0.23
3 to 2	319	2.3	-0.21	0.13	-0.16
3 to 3	12282	89.2	-0.13	0.15	-0.14
3 to 4	780	5.7	-0.26	0.2	-0.09
4 to 1	258	1.2	-0.26	-0.28	-0.58
4 to 2	351	1.6	-0.31	-0.28	-0.58
4 to 3	300	1.4	-0.28	-0.17	-0.47
4 to 4	21123	95.9	-0.3	-0.03	-0.32

Note: Academic regions are divided into four quartiles according to the importance of evaluation bias. Quartile 1 regions are likely to have the least generous evaluation standard. Examination score of candidates are standardized (mean 0 and variance 1) over the period 2005-2013. The table only considers candidates who have taken the examination at least twice during the period 2005-2013. Column (1) indicates the number of candidates in each of the 16 possible transitions between the quartiles of regions. For each quartile of the region of departure (first attempt), column (2) shows the distribution of candidates among quartiles at the second attempt. Columns (3) and (4) indicate the average score of the candidates on the first and second attempt. Column (5) shows the average score of candidates on the second attempt (column 4) minus the average gain between the first and second attempts for candidates who write the exam twice in the same quartile of regions (0.29 SD points). The learning gain between attempts is approximately neutralized in column (5). The values of columns (1) to (3) are shown in Figure 3.23.

For example, 14,059 candidates have taken the exam twice in the first quartile regions during the period 2005-2013. They represent 87.2% of the candidates who write the exam for the first time in regions of quartile 1. Their average score is -0.16 points on the first attempt and +0.16 points on the second attempt. Taking into account the learning between attempts, the exam score on the second attempt should be approximately -0.13.

Source: Ministry of Education (DEPP).

Table 3.35 – Correlation between candidates' gender, age and exam score on the first attempt over the period 2005-2009

	(1)	(2)
	Score on first attempt	Score on first attempt
Age	-0.0454*** (0.000564)	-0.0576*** (0.000728)
Women (reference)	-	-
Men	-0.169*** (0.00783)	-0.181*** (0.00837)
Level of diploma		
2 years of university (reference)		-
3 years of university		0.0378 (0.0381)
4 years of university		0.0862** (0.0386)
5 years of university and more		0.626*** (0.0412)
Specialization of diploma		
Humanities (reference)		-
Sciences		0.469*** (0.0111)
Social sciences		-0.0337*** (0.00820)
Sports		-0.0389*** (0.0121)
Other		0.164*** (0.0103)
Constant	Yes	Yes
Observations	117870	104766

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This Table explains the exam score on the first attempt by age and gender of candidates during the period 2005-2009. Column (2) includes the level of the diploma and the field specialization of the diploma as control variables.

Source: Ministry of Education (DEPP).

Table 3.36 – Correlation between candidates' gender, age, level of diploma and specialization of diploma over the period 2005-2009

	(1)	(2)
	Age	Men
Level of diploma		
2 years of university (reference)	-	-
3 years of university	-4.869*** (0.162)	0.0127 (0.0141)
4 years of university	-1.187*** (0.165)	0.0580*** (0.0143)
5 years of university and more	1.846*** (0.176)	0.0408*** (0.0153)
Specialization of diploma		
Humanities (reference)	-	-
Sciences	-0.478*** (0.0470)	0.134*** (0.00409)
Social sciences	-0.157*** (0.0349)	0.0563*** (0.00304)
Sports	-1.038*** (0.0496)	0.397*** (0.00432)
Others	-0.183*** (0.0438)	0.0467*** (0.00381)
Constant	Yes	Yes
Observations	104766	104766

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This Table explains the age (column 1) and the gender (column 2) of candidates by the level of diploma and by the specialization of the diploma.

Source: Ministry of Education (DEPP).

Table 3.37 – Learning gain between the first two attempts depending on the candidate's level of knowledge

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	IV
Learning effect b_2^1	-0.404*** (0.00437)	-0.420*** (0.00437)	-0.443*** (0.00465)	0.203*** (0.0321)	0.191*** (0.0346)	0.221*** (0.0348)
Control variables:						
Region	No	Yes	Yes	No	Yes	Yes
Year	No	Yes	Yes	No	Yes	Yes
Specialization of diploma	No	No	Yes	No	No	Yes
Level of diploma	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45160	45160	40816	45160	45160	40816
Weak identification F-stat				612.7	525.6	561.0
Sargan Chi2 p-value				0.22	0.45	0.94

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table presents the estimation results of model (c) where the explained variable is the score obtained at the second attempt minus the score obtained at the first attempt for candidates who take the exam more than once. The coefficient of interest b_2^1 represents the learning gain between the first two attempts. Columns (1) to (3) provide OLS estimates, and columns (4) to (6) provide IV estimates where the age and sex of candidates are used as instrumental variables. Columns (3) and (6) include region fixed effects, year effects, degree level and degree specialization as control variables. Columns (3) to (6) show that the instruments pass the weak identification test ($F - stat \gg 30$) and the Sargan joint exogeneity test ($p - value \gg 0.05$).

Source: Ministry of Education (DEPP).

Table 3.38 – Learning gain between the second and third attempts according to the candidate’s level of knowledge

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	IV
Learning gain b_3^1	-0.577*** (0.00785)	-0.598*** (0.00772)	-0.620*** (0.00814)	0.0159 (0.0576)	-0.0416 (0.0589)	0.0535 (0.0643)
Control variables:						
Region	No	Yes	Yes	No	Yes	Yes
Year	No	Yes	Yes	No	Yes	Yes
Specialization of diploma	No	No	Yes	No	No	Yes
Level of diploma	No	No	Yes	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15987	15987	14426	15987	15987	14426
Weak identification F-stat				206.7	184.6	173.3
Sargan Chi2 p-value				0.00	0.01	0.14

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table presents the estimation results of model (c) where the explained variable is the score obtained at the third attempt minus the score obtained at the second attempt for candidates who take the exam more than twice. The coefficient of interest b_3^1 represents the learning gain between the second attempt and the third attempt. Columns (1) to (3) provide OLS estimates, and columns (4) to (6) provide IV estimates where the age and sex of candidates are used as instrumental variables. Columns (3) and (6) include region fixed effects, year effects, degree level and degree specialization as control variables. Column (6) show that the instruments pass the weak identification test ($F - stat \gg 30$) and the Sargan joint exogeneity test ($p - value > 0.05$) when degree level and degree specialization are included as control variables.

Source: Ministry of Education (DEPP).

Table 3.39 – Region fixed effects estimated using models (2) and (3)

	(1)	(2)
	Model 2	Model 3
Region FE (ref: Paris region)		
Aix-Marseille	-0.0351** (0.0148)	-0.0151 (0.0143)
Besançon	-0.171*** (0.0216)	-0.136*** (0.0208)
Bordeaux	-0.181*** (0.0147)	-0.157*** (0.0143)
Caen	0.113*** (0.0170)	0.0978*** (0.0166)
Clermont-Ferrand	-0.364*** (0.0159)	-0.343*** (0.0155)
Dijon	0.00329 (0.0179)	0.0158 (0.0175)
Grenoble	-0.215*** (0.0148)	-0.196*** (0.0144)
Lille	-0.0829*** (0.0185)	-0.0702*** (0.0174)
Lyon	-0.551*** (0.0146)	-0.522*** (0.0141)
Montpellier	-0.137*** (0.0145)	-0.114*** (0.0141)
Nancy-Metz	0.0195 (0.0208)	0.0267 (0.0199)
Poitiers	-0.00313 (0.0152)	0.0194 (0.0148)
Rennes	-0.0311* (0.0161)	-0.0176 (0.0157)
Strasbourg	0.104*** (0.0226)	0.118*** (0.0216)
Toulouse	-0.275*** (0.0150)	-0.262*** (0.0145)
Nantes	0.0812*** (0.0145)	0.0880*** (0.0142)
Orleans-Tours	0.219*** (0.0146)	0.217*** (0.0143)
Reims	0.128*** (0.0220)	0.131*** (0.0209)
Amiens	0.381*** (0.0197)	0.356*** (0.0190)
Rouen	0.165*** (0.0195)	0.129*** (0.0189)
Limoges	-0.392*** (0.0197)	-0.379*** (0.0193)
Nice	-0.163*** (0.0179)	-0.159*** (0.0173)
Créteil	0.326*** (0.0125)	0.333*** (0.0122)
Versailles	0.196*** (0.0124)	0.204*** (0.0121)
Constant	Yes	Yes
Observations	310632	310632

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table presents the region fixed effects estimated by models (2) and (3) where the variable explained is the exam score of candidates (standardized mean 0 and variance 1) over the period 2003-2013. The two fixed effects are highly correlated: $\rho = 0.998$.

Source: Ministry of Education (DEPP).

Table 3.40 – Year fixed effects estimated using models (2) and (3)

	(1)	(2)
	Model 2	Model 3
Year FE (ref: 2003)		
2004	-0.0644*** (0.00566)	-0.0443*** (0.00478)
2005	-0.0749*** (0.00814)	-0.0204*** (0.00657)
2006	-0.102*** (0.0106)	-0.0179** (0.00838)
2007	-0.0457*** (0.0132)	0.0663*** (0.0105)
2008	-0.0146 (0.0158)	0.129*** (0.0126)
2009	-0.0234 (0.0184)	0.156*** (0.0149)
2010	-0.146*** (0.0214)	0.0604*** (0.0176)
2011	-0.407*** (0.0246)	-0.171*** (0.0209)
2012	-0.453*** (0.0271)	-0.190*** (0.0221)
2013	-0.535*** (0.0301)	-0.244*** (0.0261)
Constant	Yes	Yes
Observations	310632	310632

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table presents the year effects estimated by models (2) and (3) where the variable explained is the exam score of candidates (standardized mean 0 and variance 1) over the period 2003-2013. The two effects are highly correlated: $\rho = 0.893$.

Source: Ministry of Education (DEPP).

Table 3.41 – Learning gain between attempts estimated using models (2) and (3)

	(1)	(2)
	Model 2	Model 3
	b	b^0
Learning gain (ref: first attempt)		
2nd attempt	0.359*** (0.00400)	0.378*** (0.00350)
3rd attempt	0.463*** (0.00731)	0.404*** (0.00604)
4th attempt	0.505*** (0.0108)	0.410*** (0.00912)
5th attempt	0.533*** (0.0147)	0.404*** (0.0127)
6th attempt	0.553*** (0.0195)	0.390*** (0.0173)
7th attempt	0.556*** (0.0252)	0.361*** (0.0230)
8th attempt	0.581*** (0.0363)	0.357*** (0.0342)
9th attempt	0.741*** (0.0519)	0.486*** (0.0502)
10th attempt	0.692*** (0.0797)	0.408*** (0.0783)
11th attempt	0.570*** (0.168)	0.253 (0.167)
Constant	Yes	Yes
Observations	310632	310632

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table presents the effect of learning between attempts estimated by models (2) and (3) where the variable explained is the exam score of candidates (standardized mean 0 and variance 1) over the period 2003-2013. Column (3) reports the mean value b^0 estimated by model (3). The two effects are highly correlated: $\rho = 0.838$.

Source: Ministry of Education (DEPP).

Table 3.42 – Comparison of the region fixed effects obtained using different samples and estimation strategies

	Region fixed effect		
	Estimate based on candidates who took the exam twice in 2003 or 2004	Estimate based on candidates who took the exam twice in two different years between 2005 and 2013	Estimate based on candidates who took the exam twice between 2003 and 2013
Academic region (ref: Paris)			
Paris	0	0	0
Aix-Marseille	-0.140** (0.0669)	-0.164*** (0.0318)	-0.0495** (0.0223)
Besançon	-0.122* (0.0691)	-0.186*** (0.0445)	-0.169*** (0.0315)
Bordeaux	-0.278*** (0.0649)	-0.240*** (0.0304)	-0.194*** (0.0220)
Caen	-0.134* (0.0687)	0.0968** (0.0407)	0.0967*** (0.0256)
Clermont-Ferrand	-0.544*** (0.0652)	-0.349*** (0.0364)	-0.378*** (0.0238)
Dijon	0.0643 (0.0584)	-0.00676 (0.0377)	0.00635 (0.0261)
Grenoble	-0.326*** (0.0537)	-0.205*** (0.0345)	-0.215*** (0.0221)
Lille	-0.156** (0.0649)	-0.0336 (0.0367)	-0.0821*** (0.0269)
Lyon	-0.690*** (0.0482)	-0.311*** (0.0332)	-0.544*** (0.0213)
Montpellier	-0.335*** (0.0650)	-0.165*** (0.0308)	-0.151*** (0.0218)
Nancy-Metz	0.0436 (0.0732)	-0.00376 (0.0403)	0.0209 (0.0303)
Poitiers	-0.204*** (0.0640)	0.0916** (0.0386)	-0.0176 (0.0229)
Rennes	-0.0827 (0.0660)	-0.108*** (0.0353)	-0.0457* (0.0241)
Strasbourg	0.0785 (0.0841)	0.155*** (0.0448)	0.107*** (0.0329)
Toulouse	-0.406*** (0.0659)	-0.288*** (0.0312)	-0.288*** (0.0225)
Nantes	-0.0108 (0.0637)	-0.0373 (0.0321)	0.0665*** (0.0219)
Orleans-Tours	0.213*** (0.0477)	0.161*** (0.0331)	0.223*** (0.0214)
Reims	0.0706 (0.0832)	0.160*** (0.0426)	0.128*** (0.0321)
Amiens	0.288*** (0.0761)	0.392*** (0.0375)	0.381*** (0.0288)
Rouen	0.143** (0.0603)	0.164*** (0.0443)	0.167*** (0.0284)
Limoges	-0.544*** (0.0720)	-0.443*** (0.0447)	-0.405*** (0.0291)
Nice	-0.434*** (0.0723)	-0.0723* (0.0380)	-0.177*** (0.0267)
Créteil	0.312*** (0.0514)	0.279*** (0.0237)	0.324*** (0.0182)
Versailles	0.203*** (0.0472)	0.137*** (0.0240)	0.195*** (0.0182)
Observations	117186	311418	428604

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level.

Note: This table compares the region fixed effects estimated by model (1) over the period 2003-2004 (column 1), model (2) over the period 2005-2013 (column 2) and model (2) over the period 2003-2013. Column (1) exploits the fact that candidates can take the competition in two different regions in the same year in 2003 and 2004. Column (2) exploits the fact that candidates can take the examination several times in different years. Column (3), our main specification, exploits the fact that candidates can take the exam twice in the same year (in 2003 or 2004) or several times in different years. The correlation between the region fixed effects is $\rho = 0.88$ for columns (1) and (2), $\rho = 0.95$ for columns (2) and (3) and $\rho = 0.96$ for columns (1) and (3). *Source:* Ministry of Education (DEPP).

Figures

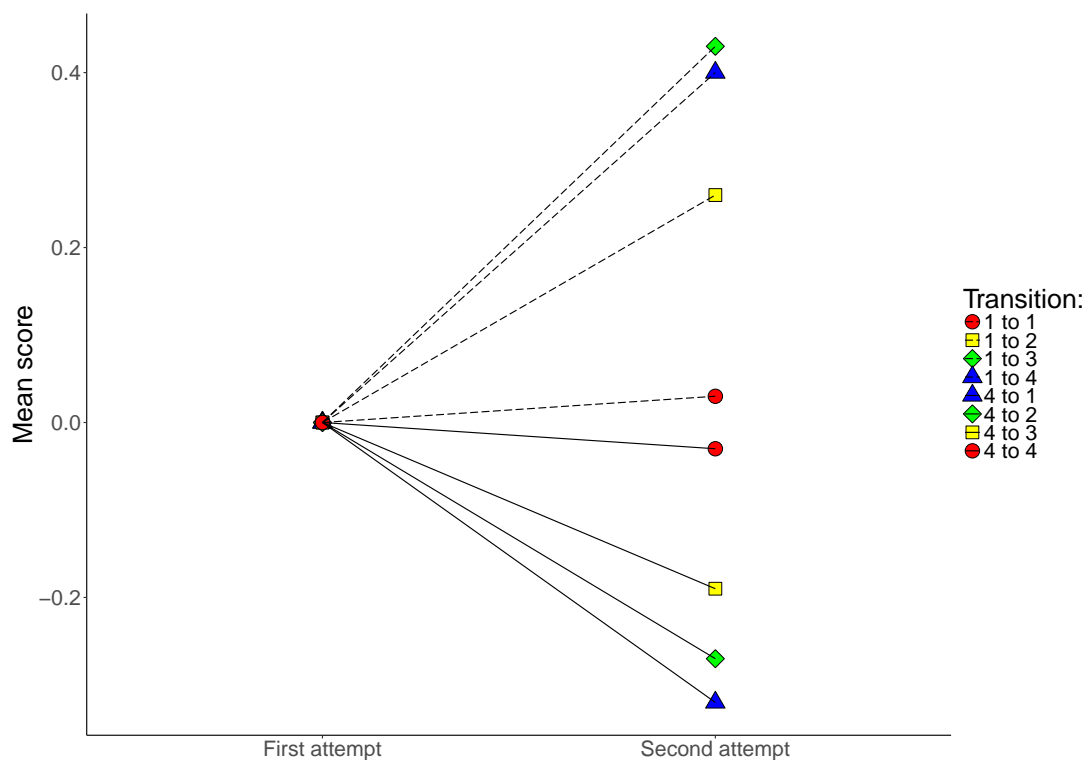


Figure 3.28 – Variation in candidates' exam score between the first and second attempts according to the quartile of the region of departure and arrival

Note: Academic regions are divided into four quartiles according to the extent of assessment bias. Quartile 1 regions are likely to have the least generous evaluation standard. Examination score of candidates are standardized (mean 0 and variance 1) over the period 2003-2013. This Figure shows the average score of candidates who participate in the competition for the first time in a region of quartile 1 or quartile 4 (first attempt) and who participate in the competition for the second time in a region of quartile 1, 2, 3 or 4 (second attempt). To identify more clearly the effect of transitions between quartiles on candidates' scores, I subtract the average score gain for candidates who take the exam twice in the same quartile from the score obtained on the second attempt. In addition, I subtract the average score on the first attempt from the score of the first and second attempts so that only the effect of transitions between quartiles is visible.

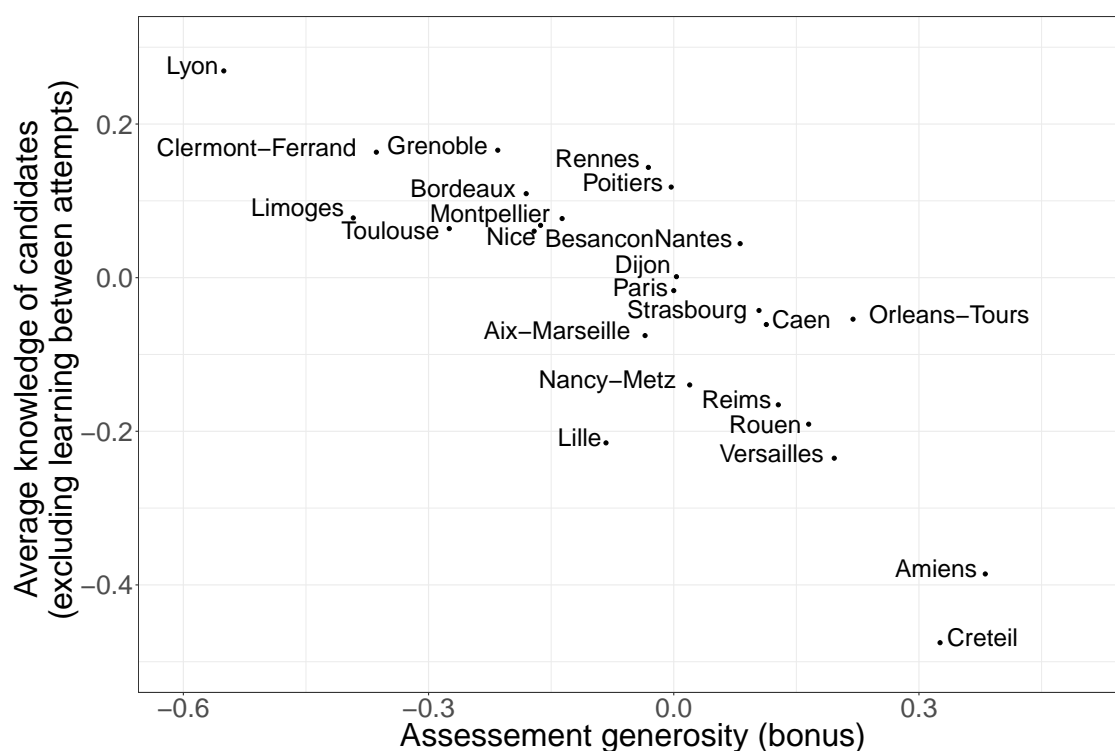


Figure 3.29 – Relationship between the average level of knowledge of candidates and the generosity of the assessment in the region

Note: Each point represents an academic region in metropolitan France (excluding Corsica). The average level of knowledge of candidates on the Y-axis corresponds to the average individual fixed effects estimated by model (2) over the period 2003-2013. The generosity of the evaluation on the X-axis corresponds to the region fixed effects estimated by model (2) over the period 2003-2013. A higher region fixed effect means that candidates obtain on average a score gain when they write the exam in that region. The Figure shows that the candidates' level of knowledge and the generosity of assessment practices in the region are negatively correlated.

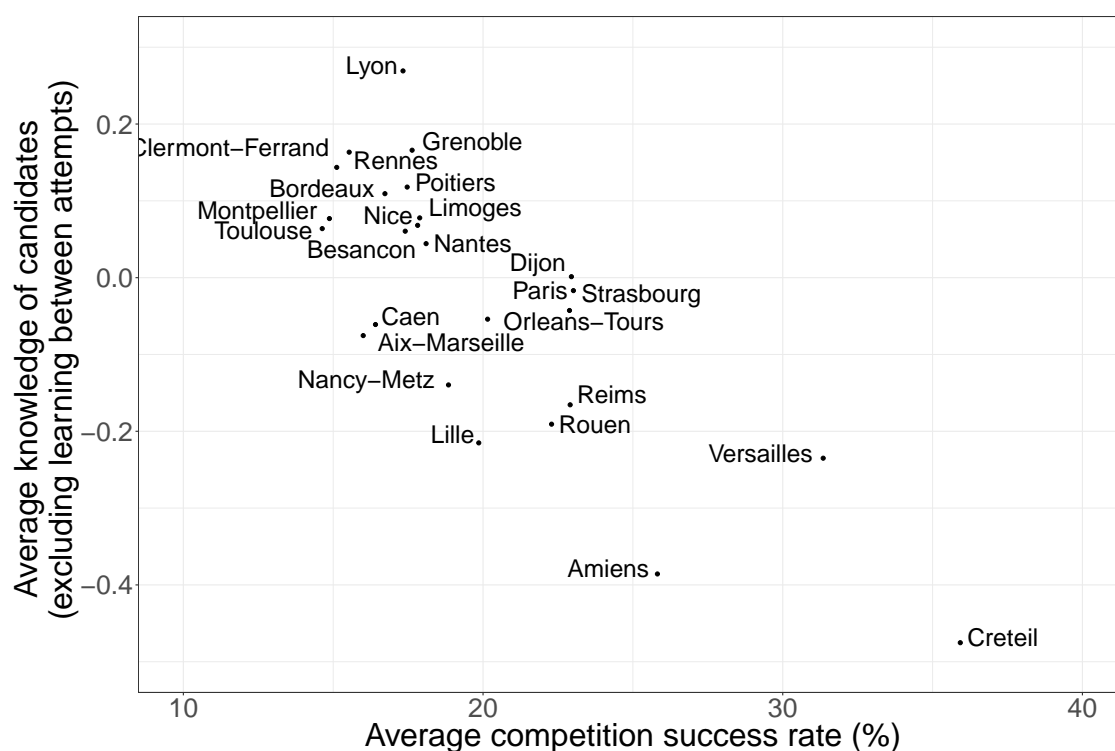


Figure 3.30 – Relationship between the average level of knowledge of candidates and the success rate of the competition in the region

Note: Each point represents an academic region in metropolitan France (excluding Corsica). The average level of knowledge of candidates on the Y-axis corresponds to the average individual fixed effects estimated by model (2) over the period 2003-2013. The average success rate on the X-axis corresponds to the average number of recruited teachers divided by the average number of candidates in the region over the period 2003-2013. The Figure shows that the candidates' level of knowledge and the success rate in the region are negatively correlated.

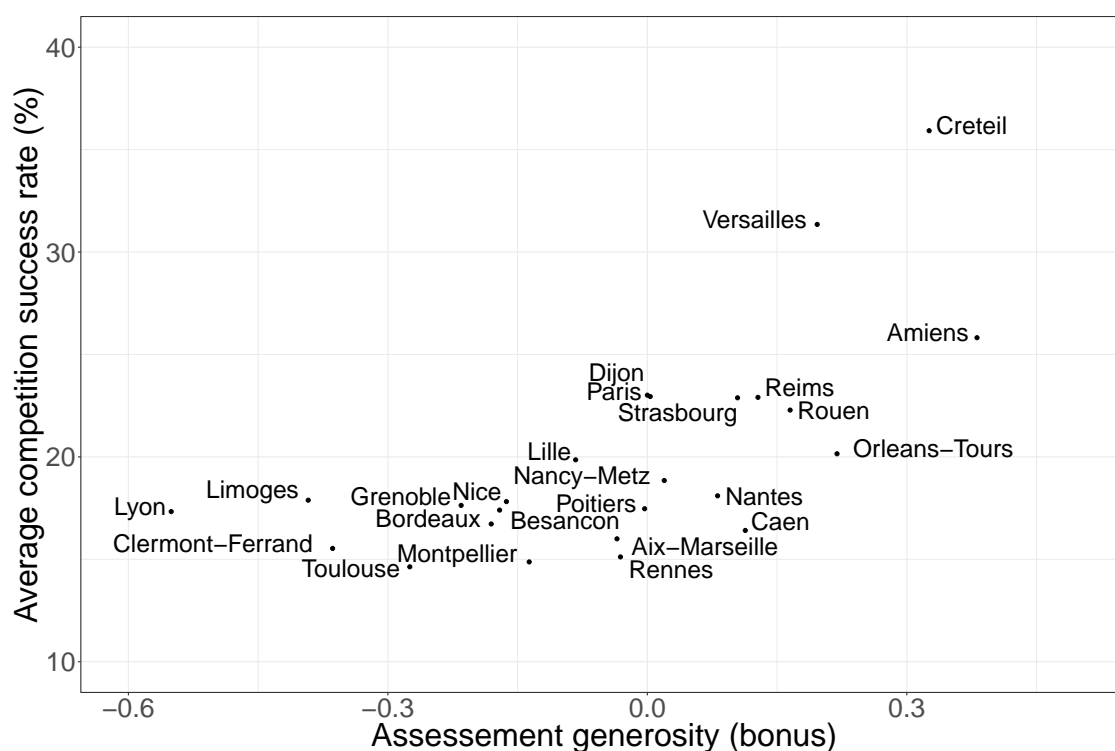


Figure 3.31 – Relationship between the success rate of the competition and the generosity of the assessment in the region

Note: Each point represents an academic region in metropolitan France (excluding Corsica). The average success rate on the Y-axis corresponds to the average number of recruited teachers divided by the average number of candidates in the region over the period 2003-2013. The generosity of the evaluation on the X-axis corresponds to the region fixed effects estimated by model (2) over the period 2003-2013. The Figure shows that the success rate in the region and the generosity of assessment practices in the region are positively correlated.

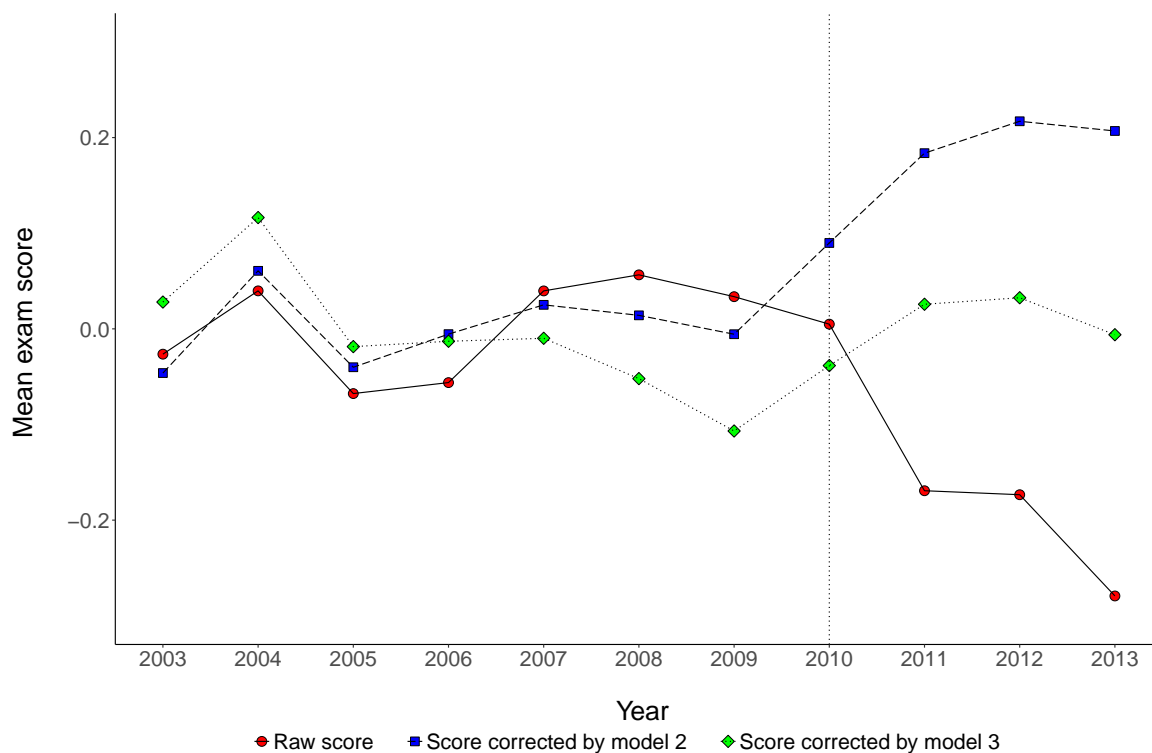


Figure 3.32 – Evolution of candidates’ raw and corrected scores on recruitment examinations at national level over the period 2003-2013

Note: Academic regions in metropolitan France (excluding Corsica). The raw examination score corresponds to the average examination score of candidates observed in the data over the period 2003-2013. The corrected examination scores are obtained by neutralizing the evaluation biases (region fixed effects and year effects) estimated from models (2) and (3). The raw and corrected scores are normalized (mean 0 and variance 1) for candidates over the 2003-2013 period. This Figure shows that candidates’ examination scores corrected by models (2) and (3) increase after the 2011 reform, while the raw score decreases after the 2011 reform.

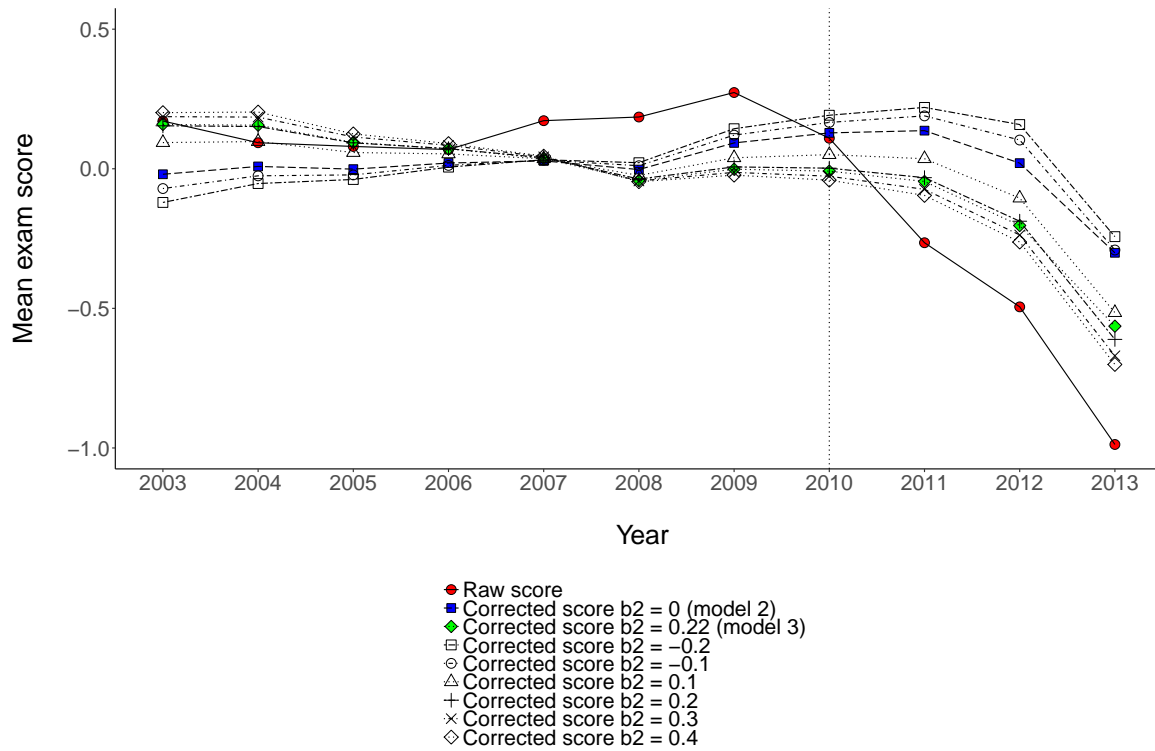


Figure 3.33 – Evolution of teachers' raw and corrected exam scores at national level during the period 2003-2013 for different values of the correlation between the level of knowledge and the learning gain between attempts (model 3)

Note: Academic regions in metropolitan France (excluding Corsica). This Figure tests the robustness of the score corrected by model (3). It presents the evolution of the corrected score of teachers using different values for the coefficient b_2^1 which characterizes the correlation between the candidates' level of knowledge and the learning gain between the first two attempts. The value obtained by estimating model (c) using the sex and age of the candidates as instrumental variables is $b_2^1 = 0.22$. The value tested are $b_2^1 \in \{-0.2, -0.1, 0.1, 0.2, 0.3, 0.4\}$. The score corrected by model (3) seems relatively robust to the choice of the coefficient b_2^1 .

Chapter 4

How school context and management influence sick leave and teacher departure

4.1 Introduction

Research in economics suggests that managerial and workplace practices contribute to corporate and state performance ([Black and Lynch, 2001](#); [Bertrand and Schoar, 2003](#); [Bloom and Van Reenen, 2007](#); [Branch et al., 2012](#); [Bloom et al., 2014](#); [Bryson and Green, 2018](#)) but we know little about their effects on workers' behaviour and health. Absences have received considerable attention in economics because they generate significant costs for employers (work breakdown, loss of working hours, wage compensation, replacement) and welfare states (sickness benefits). They also reveal organizational dysfunctions and provide an opportunity to examine how managers and workplaces influence worker performance ([Coelli and Green, 2012](#); [Branch et al., 2012](#)). A large strand of the literature shows that absences are related to socio-demographic characteristics (namely age, gender, family situation, education, occupation), health and motivation (effort). However, few articles have analyzed the role of workplaces and managerial practices in shaping absenteeism. A better understanding of these effects would help design appropriate measures to reduce absences and resolve the underlying problems they reflect (health, organization or productivity).

I examine this issue by studying the effect of schools (workplaces) and school principals

(managers) on teacher absences. I exploit a unique matched employer-employee dataset that contains all the absence spells of French educational staff over nine years¹. I focus on secondary school teachers who represent approximately 40% of all the educational staff². In France, teachers can be employed on fixed-term or permanent contracts. I exclude part-time teachers (12%) and substitute teachers who work in more than one school during the year (8%). The final dataset includes 8,621 (junior and senior) high schools, 12,548 school principals and 470,443 teachers that I observe on average 5.7 times over the period 2007-2015. The administrative files provide various socio-demographic information (gender, age, family situation, number of children), a description of the professional situation (contract, seniority, subject taught) and some information on schools (number of teachers, location, main occupation of students' parents). I focus on the annual duration of absences for sickness reason³ and I exclude absences due to a work accident, due to a family event (death of a close relative, marriage, sick child) as well as maternity leaves, paternity leaves and parental leaves.

The mobility of teachers and principals between schools allows to decompose teacher absences due to illness into individual, work-related and residual contributions. The individual component includes teacher fixed effects and time-varying teacher characteristics. The work-related contributions include school fixed effects, time-varying school characteristics and principal fixed effects. The results show that schools and principals contribute to about 11% of teacher absenteeism. On average, the absence duration of teachers increases by 250-300% (\approx 16-18 days) when they move from the first to the fourth quartile of school fixed effects, and increases by 170-220% (\approx 12-14 days) when they transition between the first and the fourth quartile of principal fixed effects. Robustness checks show that the endogenous sorting of teachers and principals does not generate a substantial bias in our estimates.

Several mechanisms can explain the contribution of schools and principals to teacher absences. First, school fixed effects may suffer from omitted variables bias (endogeneity problem). Differences between localities in terms of climate, pollution or access to health infrastructures (distance, equipment, number of doctors per inhabitant) for instance, can

¹This dataset belongs to the French Ministry of Education.

²In France, the primary school principal is not the superior of the other teachers in the school. Since we are interested in the effect of managers on teacher absences, we exclude primary school teachers (about 40% of education staff) from the analysis.

³Note that in France, employees must present a medical certificate to be on sick leave.

lead to fallacious correlations between schools and absenteeism. However, the disparity in school effects remains significant after the inclusion of local dummy variables in the model, suggesting that the external environment does not explain most of the school differences in teacher absenteeism. Then, I examine the role of factors related to the work environment. The social science literature identifies two main mechanisms that can explain absenteeism from work: moral hazard, which translates into minimizing effort, and health problems. As in most empirical studies, the available data do not allow to separate absences due to health problems from absences that reflect a lack of effort. However, the data allow to examine whether the effects of schools and school principals on teacher absenteeism are more related to one or the other mechanism.

A first line of economic research suggests that variations in absenteeism between workplaces reflect differences in absence norms (Harvey and Nicholson, 1999; Gazieli, 2004; Bamberger and Biron, 2007; Hausknecht et al., 2008). For example, school effects may result from the fact that teachers in some schools consider absenteeism to be a more acceptable behavior than teachers in other schools. The model used in this study takes into account teachers' fixed effects and variable (observable) characteristics. It is therefore unlikely that the school effects I estimate reflect the sorting of teachers. However, they may reflect peer interactions (social pressure, peer monitoring) that contribute to exacerbating or reducing individuals' absence behavior. Ichino and Maggi (2000) and Bradley et al. (2007) find that peer behavior influences individual absences in the private and public sectors. I test this hypothesis using the recent model proposed by (Arcidiacono et al., 2012) which circumvents a number of issues in peer effects estimation. The results confirm slightly positive peer effects on absences, but these effects contribute very little to the variance of the estimated school (and principal) fixed effects.

A second line of research suggests that schools may differ in terms of working conditions. Several studies show that teachers are highly exposed to professional stress (*job strain*) (Hakanen et al., 2006; Jégo and Guillo, 2016), which have consequences on the health and effort of individuals. Research in epidemiology (Karasek, 1979; Siegrist, 1996), and more recently in economics, show that adverse working conditions can have a detrimental effect on health, motivation and work engagement. Stress at work in particular is associated with mental health disorders (Griffin et al., 2002; Hakanen et al., 2006; Cottini and Lucifora, 2010; Madsen et al., 2017) and cardiovascular diseases (Belkic et al., 2004;

Kuper et al., 2002; Kuper and Marmot, 2003). People who lack resources and support are also more dissatisfied with their job; they are more likely to be absent to escape or protest against their working conditions (Sagie et al., 2002; Schaufeli et al., 2009). In this study, I find evidence that teacher absences are more likely to reflect unfavorable working conditions than permissive absence standards. First, I notice that teacher turnover increases with school fixed effects on absenteeism, and that teachers who leave schools with a high fixed effect are more absent than their colleagues on average. If high fixed effects meant more lax absence standards in schools, the teachers who benefit most from this situation, by being absent more or longer than their colleagues, would also tend to stay longer in these schools. On the contrary, teachers who disagree with these permissive norms, who are less absent than their colleagues, should leave schools with a high fixed effect more quickly. It is precisely the opposite situation that we observe, which suggests that an increase in the fixed effect of the school reflects a situation that is particularly unfavorable for the teachers who are absent the most. Second, I match 707 schools with the 2013 and 2016 working conditions surveys, which allows to study the correlation between school fixed effects and the working environment described by a random sample of 733 teachers. I find that work intensity and hostile behavior are significantly more prevalent in schools with a high fixed effect. This result is consistent with the evidence highlighted in other articles (Skaalvik and Skaalvik, 2009; Pas et al., 2012). I also show that the psychological well-being of teachers, measured by the WHO-5 index in the 2013 and 2016 working conditions surveys, decreases significantly when they move to schools with a higher fixed effect. All these results are robust to the inclusion of teacher fixed effects⁴ and suggest that school effects on absenteeism operate through the deterioration of teachers' working conditions and health.

Similarly, several factors are likely to explain the fixed effects of principals on teacher absenteeism. A first range of explanations refers to the problem of moral hazard. School principals can influence teachers' level of effort by controlling the frequency of absences and exerting moral pressure, for example. However, when the risk of sanction is low, some people reduce their efforts and take more time off work. If this mechanism dominates, principals with a high fixed effect are more likely to manage staff in a lax manner. Another possibility is that the managerial practices of school principals may affect the health and

⁴Since 2013, working conditions surveys have followed a panel of individuals, which makes it possible to introduce teacher fixed effects into the main specifications.

motivation of teachers. Several studies show that lack of hierarchical support, lack of recognition of work, lack of autonomy and conflicting demands lead to job strain and dissatisfaction, which generates disengagement from work and health problems (Farrell and Stamm, 1988; Hakanen et al., 2006; Schaufeli et al., 2009). In this study, I find evidence that principals' effects on absences are more likely to be related to psychosocial risk factors than to laxity or moral hazard. First, I note that principals who tend to increase teacher absences (principals with a high fixed effect) also tend to increase teacher turnover. More precisely, teachers who are absent longer than their colleagues are more likely to apply for a new assignment after the arrival of a principal with a high fixed effect⁵. This result suggests that school principals who tend to increase absenteeism generate an unfavorable environment for the teachers who are absent the most. The analysis of the subsample of schools matched with the 2013 and 2016 working conditions surveys reveals that principals' effects are associated with a lack of hierarchical support, after controlling for school characteristics. My preferred specification also indicates that teachers' psychological well-being (WHO-5 index) decreases with the arrival of a head teacher with a high fixed effect on absenteeism. These results are robust to the inclusion of teacher fixed effects and suggest that principals' effects on absenteeism translate into lower teacher motivation and health.

One of the limitations of this study is that the identification strategy in place does not allow the long-term consequences of repeated exposures to adverse working conditions and management practices to be examined. It is therefore reasonable to assume that the results presented in this study provide a lower bound of the effects of work environment and management practices on teacher absenteeism.

These results suggest that the implementation of awareness-raising activities in schools and better training of head teachers on psychosocial risk prevention could contribute to improving teachers' health and school system performance. The methodology proposed in this study can help target schools and principals who are particularly "at risk". It should be recalled that 8% of education personnel are substitute teachers and that they

⁵In France, the possibility of school principals to encourage and sanction staff is relatively limited. In particular, they do not recruit, dismiss or sanction tenured teachers (who represent 92% of secondary school teachers). In the most serious cases, the principal may establish a disciplinary file that will be reviewed by a board headed by an education inspector. The school principal's assessment also has relatively little influence on teacher remuneration. However, school principals allocate classes, courses and pedagogical projects among teachers. They can reward or sanction staff by deciding whether or not to take into account teachers' requests for organization, for example.

mainly replace colleagues on sick leave. Improving the working environment of teachers would certainly contribute to reduce the number of substitute positions.

The rest of the paper proceeds as follows: Section 2 reviews the social science literature linking absences, work environment and management practices. Section 3 presents the context to this study and section 4 describes the data. Section 5 presents the method and section 6 comments the results. Section 7 tests the key assumptions of the model and Section 8 concludes.

4.2 Literature review

In neoclassical economic approaches, absenteeism is considered as a pure labor supply phenomenon (Brown and Sessions, 1996). In the model developed by Allen (1981), work absences result from a trade-off between labor and leisure. Absences compensate for the difference between the contractual working hours and the working time desired by individuals. In the "shirking models" based on the efficiency wage theory proposed by Shapiro and Stiglitz (1984), work absences result from a trade-off between work utility and effort. The objective of workers is to minimize effort while avoiding associated sanctions that would reduce the utility of work (loss of wages, for example). According to both theories, absences increase with the value of leisure - respectively with the cost of effort (e.g. complex tasks, poor working conditions) - and decrease with the value of work (e.g. wage, job satisfaction). An important implication is that individuals modulate the duration and frequency of absences according to the incentives and sanctions they receive in the workplace. The role of managers is therefore essential to limit absences from work. There is considerable empirical evidence that tend to support this mechanism. First, many authors have shown that absences are more frequent in workplaces where absence standards are more permissive (Harvey and Nicholson, 1999; Gaziel, 2004; Bamberger and Biron, 2007; Hausknecht et al., 2008). They argue that weak sanctions and controls encourage opportunistic behavior by reducing the cost of shirking. Similarly, studies have highlighted the existence of peer effects on the absence behavior of individuals (Ichino and Maggi, 2000; Bradley et al., 2007). The authors' interpretation is that the risk of being caught and punished decreases when colleagues also shirk. Several studies have shown that individual absences increase with job security (e.g., type of contract, job opportu-

nities) (Ose, 2005; Engellandt and Riphahn, 2005; Ichino and Riphahn, 2005; Riphahn and Thalmaier, 1999; Grignon and Renaud, 2007; Hausknecht et al., 2008; Scoppa, 2010). Other authors have examined changes in sick leave compensation systems and conclude that absences also increase with the wage replacement rate and decreases with the intensity of absence monitoring (e. g. require a medical certificate) (Meyer et al., 1995; Bolduc et al., 2002; Johansson and Palme, 2005; Henrekson and Persson, 2004; Frick and Malo, 2008; Ziebarth and Karlsson, 2010; D'Amuri, 2011; Ziebarth, 2013; De Paola et al., 2014).

But absences also reveal the health status of individuals. In most countries, sick leaves - the vast majority of reasons for absence - must be justified by a medical certificate. Absences vary greatly with the prevalence of seasonal illnesses (Marmot et al., 1995; DARES, 2013) and are significantly associated with pathologies, chronic illnesses and self-reported health (Jenkins, 1985; Marmot et al., 1995; Collins et al., 2005). Other studies show that the history of sick leave is a good predictor of future health and mortality risk after controlling for age, occupational status and employment contract (Kivimäki et al., 2003; Vahtera et al., 2004). Recent research reveals that sick people respond to incentives to return to work. In particular, Pichler and Ziebarth (2015) shows that the implementation of sickleave benefits contributes to reducing the spread of communicable diseases in companies and cities by reducing "contagious presenteeism". These results suggest that reactions to sanctions are not always synonymous with opportunistic behavior.

In epidemiology, two theories provide a framework for understanding the influence of workplace and management practices on work stress and health outcomes. In the job demand-resource model (Karasek, 1979), job strain is generated by high psychological demands (time pressure, work intensity, conflicting requests, work interruptions), combined with insufficient resources (low decisional latitude, high constraints, weak autonomy) and low social support from hierarchy and colleagues. In the effort-reward model (Siegrist, 1996), work stress derives from an imbalance between effort (working overtime, personal engagement, responsibility, work intensity) and reward (remuneration, esteem, gratitude, career opportunities). Some authors distinguish between an "energetical process" and a "motivational process" that lead to absenteeism (Steers and Rhodes, 1978; Hakanen et al., 2006). In the "energetical process", unfavorable working conditions affect people's health (exhaustion, long-term illness) and lead to "involuntary" absences. In the "motivational

process", the lack of resources and support contribute to reducing job satisfaction and commitment. Workers are "voluntarily" absent from work to escape (or protest against) adverse working conditions (Farrell and Stamm, 1988; Mathieu and Zajac, 1990; Sagie, 1998; Sagie et al., 2002; Schaufeli et al., 2009). Empirical evidence support that long-time exposure to job strain is associated with mental health disorders (Griffin et al., 2002; Nourry et al., 2014; Madsen et al., 2017), cardiovascular diseases (Belkic et al., 2004; Kuper et al., 2002; Kuper and Marmot, 2003), musculoskeletal problems (Bongers et al., 1993; Davis and Heaney, 2000; Hauke et al., 2011) and (long-term) absences (Niedhammer et al., 1998; Vahtera et al., 2000; Melchior et al., 2003; Schaufeli et al., 2009; Janssens et al., 2014; Wang et al., 2014). The total cost of stress at work is also particularly high for society (Béjean and Sultan-Taïeb, 2005).

Recently, the empirical economic literature has also confirmed that working time and time constraints affect the health of individuals. In particular, long hours of work and involuntary part-time work are associated with a deterioration of physical and mental health (van der Hulst, 2003; Santin et al., 2009; Robone et al., 2011). Fixed-term contract workers are less likely to be absent (Benavides et al., 2000) but they report more health problems and lower psychosocial well-being, especially when they are poorly educated (Robone et al., 2011). Irregular working hours increase sickness absence (Afsa and Givord, 2006) while working in shifts, work intensity, work complexity and lack of autonomy increase mental health problems (Cottini and Lucifora, 2010). Finally, biological, chemical and physical exposures (noise, vibration, postural constraints, outdoor work, night work) are significantly associated with more frequent and longer periods of absence (Melchior et al., 2005; Niedhammer et al., 2017).

Several studies show that teachers are particularly exposed to work stress. They face a high risk of burnout (occupational exhaustion) compared to most professions (Hakanen et al., 2006). Farber (1991) estimates that between 5% and 20% of all teachers in the United States will suffer from burnout at least once during their career. In France, teachers report a lack of recognition of work, a lack of support from management, a lack of relationships with colleagues and a high psychological demand compared to most executives in the private and public sectors (Jégo and Guillo, 2016). Authors have highlighted that time constraints, work intensity, student behavior and lack of support from colleagues and school leaders are significantly associated with teacher turnover, health

problems and absences (Olivier and Venter, 2003; Pomaki and Anagnostopoulou, 2003; Skaalvik and Skaalvik, 2009; Pas et al., 2012). However, the causal interpretation of these studies is often limited and further work is needed to confirm their results.

4.3 Context

4.3.1 Absences in the French education system

In France, any absence for health reasons must be justified by a medical certificate. Employees in the education sector (teachers and non-teachers) who meet this condition receive a full salary for up to three months ; then, they receive half of their salary for up to nine months⁶. The administration may send a doctor at any time to check the health of the absent person. If she refuses the control visit, the administration may interrupt the compensatory wage and take disciplinary action. Similarly, employees lose 1/30 of their wages per day of unjustified absence and are liable to disciplinary sanctions. Usually, absences for health reasons cannot exceed one year. After this period, the person becomes non-active (she loses her position, receives no salary and ceases to accumulate seniority. However, she can return to the profession once she has recovered health), is assigned to a new job or is retired for disability. However, certain illnesses and health conditions ensure full or partial wage compensation up to 5 years after the beginning of the absence.

Absences may also be granted for family events (birth, adoption, death of a close relative, civil union), education, trade union activity, or to care for a sick child. They are generally subject to the approval of the headteacher and are not necessarily remunerated.

4.3.2 Recruitment and wage in (secondary) education

The French education system consists of private schools (about one-third) and public schools (about two-thirds). 78% of pupils in general or vocational secondary education are enrolled in public schools. Public secondary schools employ 81% of the secondary school workforce, which includes teachers (74%), teaching assistants (9%), administrative staff

⁶An amendment to the salary compensation rule for sick leave was introduced between January 2012 and December 2013: during this period, the first day of each new absence is unpaid, meaning that employees lose 1/30 of their salary. The effect of this reform on the frequency and duration of absences is examined in another paper.

(5%), managers (1%), educational advisors (1%), nurses and physicians (1%)⁷. Most public education employees are civil servants: 98% of primary and elementary school teachers, 92% of secondary school teachers and 60% of non-teaching staff (mainly because teaching assistants are never civil servants). Those who are not civil servants are either on fixed-term contracts or on permanent contracts with the Ministry of Education. In what follows, I describe how secondary school teachers are recruited and paid in public schools, but it is generalizable to most public education officials.

Recruitment

Employees in the public education sector must pass a competitive examination and validate a one-year probationary period to become public servants. Access to each profession, and to each status within the profession (called "corps"), involves a specific competition. For example, teachers must pass the "CAPES" or "Agrégation" examination to become civil servants, and these two competitions lead to a slightly different status in the profession, with notably a difference in salary and number of hours worked ("corps des professeurs certifiés" and "corps des professeurs agrégés" respectively)⁸. The "Agrégation" examination is reputed more difficult and selective than the CAPES examination; teachers with Agrégation can teach in secondary or post-secondary education, while teachers with CAPES teach only in high and middle schools. Another difference is that face-to-face (full-time) teaching lasts 15 hours for teachers with Agrégation and 18 hours for teachers with CAPES. Each competition is subject to specific access conditions (level of diploma, professional experience, past civil servant status) and several competitions may lead to the same status within a profession. Civil servant teachers may be assigned to permanent (one year at least) or substitute posts in one or more schools.

Non-civil servant teachers are recruited in two stages. First, the academic region (a division close to the administrative region) centralizes and preselects applications for vacant positions. Candidates are then interviewed by a pedagogical inspector and/or the school principal. Non-civil servant teachers are recruited on short-term contracts for a period of one year, renewable annually. Generally, the same degree level is required to become a civil servant or teacher on fixed-term contracts, except when schools are faced with a shortage of qualified teachers. In recent years, some schools have recruited short-

⁷Repères et références statistiques sur les enseignements, la formation et la recherche, DEPP, 2017

⁸See [Breda and Hillion \(2016\)](#) for a complete description of the CAPES and Agrégation exams

term teachers with three years of university, and in some cases two years of university, while civil servants are required five years of university. Non-civil servant teachers shall fill substitute or permanent posts which remain vacant after the assignment of all civil servant teachers. After six years of uninterrupted service, they must obtain an open-ended contract or stop working in public education. They may also take a competition (a special CAPES or Aggregation competition) to become permanent officials.

Wage

Specific rules govern the level and evolution of salaries of public education employees. The salaries of civil servants depend essentially on their profession, their status (which depends entirely on the competition they have passed) and their seniority according to a detailed salary scale. It consists of a fixed part (about 85% of the salary of secondary teachers), which is the same for everyone, and a bonus part which varies according to the individual situation (school type - middle school, high school, level taught, responsibilities, overtime). On average, teachers with Aggregation earn 30% more than teachers with CAPES⁹.

The salary of an employee on a fixed-term or permanent contract depends on his/her profession and seniority, but also on his/her diploma and professional experience (before entering public education). The salary also depends on the specificities of the teaching position and varies according to the professional evaluation that takes place every three years (taken by the school principal and a pedagogical inspector). The salary of a non-civil servant is always lower (between 15 and 25% depending on his/her degree, experience and seniority) than the salary of a civil servant performing exactly the same work.

4.3.3 Mobility

Teacher mobility between public schools is organised once a year by the Ministry of Education. The aim of the administration is to best meet the mobility aspirations of teachers and fill vacancies throughout the country. This centralised mobility procedure concerns only civil servant teachers (remember that teachers on fixed-term contracts are recruited and managed directly by the schools). All newly recruited teachers (non-permanent) and full time teachers who wish to change schools must participate in the assignment process.

⁹Bilan social du ministère de l'éducation nationale, du ministère de l'enseignement supérieur et de la recherche - Année 2012-2013.

The mobility procedure consists of two stages. The first stage concerns all newly recruited teachers and full teachers who wish to change regions. Applicants for mobility must rank regions according to their preferences. In the second stage, teachers newly assigned to a region and teachers who wish to change schools in their current region rank up to twenty schools in order of preference. For its part, the administration defines priority rules and calculates a mobility score to prioritize requests for teacher assignments. Then, an algorithm is used to match teachers' preferences with the needs of the school according to these priority rules¹⁰. The mobility score takes into account several criteria, including spousal reunification, disability and work in a dangerous or socially and economically disadvantaged area. The score also increases as a function of the length of separation of spouses, time spent in a difficult school, overall seniority in teaching and seniority in the current school. The scale is transparent and the Ministry of Education provides an online questionnaire that allows teachers to calculate their score before applying for mobility. The administration also announces the thresholds that allowed entry into each region, department or school for each discipline in previous years. It is important to note that tenured teachers may withdraw from the assignment process at any time if they are not satisfied with their new assignment. In this case, they stay in their initial school. However, probationary teachers and teachers on temporary positions may not refuse the final assignment decision.

The mobility procedure allows to fill most of the vacant posts, i.e. full-time teaching positions in one school or substitute positions in several schools in the same geographical area.¹¹ However, a minority of positions are created or become vacant after the beginning of the mobility procedure (e. g. due to a teacher's long illness or new needs). A "late mobility procedure" temporarily offers these posts (usually for one year) to teachers who have been assigned a substitute position during the normal mobility procedure. Teachers in temporary positions must participate in the mobility process the following year. Any vacancies remaining after normal and late mobility processes are offered to teachers on fixed-term contracts.

¹⁰For more information on public teacher mobility and assignment procedures, see [Combe et al. \(2016\)](#) and to the Matching in Practice website: <http://www.matching-in-practice.eu/matching-practices-of-teachers-to-schools-france/>

¹¹Substitute teachers (assigned to a geographical area and not to a school) account for about 8% of the teaching force in secondary education. Schools usually use a substitute teacher when a regular teacher is absent for more than two weeks and less than a year

4.3.4 Teachers in deprived areas

Schools in which pupils are mainly from disadvantaged socio-economic backgrounds receive increased financial and human resources. These schools are part of the priority education network whose objective is to reduce social inequalities in educational attainment¹². In 2017, about 20% of junior high school students¹³ are enrolled in a priority education network school. The administration regularly calculates and reviews a social indicator that allows schools to be assigned to this network. Since 2015, the social indicator has taken into account four criteria: the share of pupils who are disadvantaged in socio-economic terms (whose parents are blue-collar workers, retired workers, retired employees, unemployed persons who have never worked, or have no professional activity), the share of scholarship students (depending on their parents' income), the share of pupils living in a disadvantaged and dangerous urban area, the share of pupils who have repeated at primary school.

In practice, there are fewer pupils per teacher (approximately two pupils less for an average class of twenty-five), additional training periods and a salary bonus for teachers (which represents about 6% of the average net salary) working in the priority education network. Teachers are also supported by additional teaching assistants, educators, social and medical staff (Stefanou, 2013). It is also the only case where teachers are compensated for their difficult working conditions. However, schools in the priority education network are rarely the first choice of teachers and the turnover of education staff remains very high (Prost, 2013).

4.3.5 School principals and vice-principals

The principal in secondary school is the immediate supervisor of all education personnel working in the school. He is responsible for the safety of people and property and can make important decisions such as closing the school. Secondary school supervisory staff (principals and vice-principals) are recruited among education and inspection staff by competitive examination or, to a lesser extent, from a list of qualified candidates. To be admitted to the public school administration competition, candidates must have at least

¹²See Garrouste and Prost (2016) for a review of the priority education network and its impact on school achievement.

¹³Since 2015, senior high schools are no longer included in the priority education network.

five years' experience in the field of education; they must have at least ten years' seniority in order to be admitted to a shortlist of candidates.

Requests for mobility of secondary school principals and vice-principals are processed at the national level. Unlike teachers, each application is examined individually to take into account the candidate's profile and the complexity of the position to be filled. Only persons who have been working in the same establishment for more than three years may apply to change schools. Applications for mobility may exceptionally be taken into consideration after two years spent in the same school in the event of disability or spousal reunification. In addition, principals and vice-principals cannot work in the same school for more than six years. Participants in the mobility process cannot refuse their assigned position.

4.4 Data and descriptive statistics

4.4.1 Sources

The data used in this study come mainly from the Personnel Management System of the Ministry of Education. It concerns all educational personnel, i. e. teachers, principals and vice-principals, as well as all public educational institutions. The data contain individual information such as age, sex, marital status, occupation, level of diploma, employment contract, weekly hours, seniority, length of absence and reason for absence (health, adoption, maternity, work accident). It also provides information on schools, such as their type (junior high school, senior high school, vocational college), their address and their inclusion in the priority education network. We also know the proportion of pupils from "disadvantaged" socio-professional backgrounds in each school, the turnover of teaching staff and the unemployment rate in each locality.

The data are longitudinal and cover the period 2007-2015. I focus on secondary school teachers (lower and upper secondary schools) because primary school principals are not the immediate supervisors of primary school teachers. I exclude from the analysis part-time teachers (12%) and substitute teachers who work in several schools during the year (8%). I focus on absences for health reasons accumulated over a school year, which exclude parental leave, absences for family events and absences due to accidents at work. The final data includes 8,625 (junior and senior) high schools, 12,549 school principals

and 481,624 teachers that I observe on average 5.7 times over the period 2007-2015.

The 2013 Working Conditions Survey carried out by the Ministry of Labour describes workers' exposure to physical constraints and psychosocial risks (social relations, job satisfaction and difficulties), as well as work organization (intensity, autonomy) and working time organization (duration, time, hours) in public and private establishments¹⁴. The administrative data from the Ministry of National Education can be linked to the 2013 Working Conditions Survey using the name and address of the employer. Of the 939 secondary school teachers who participated in the 2013 Working Conditions Survey, 828 agreed to provide the name and address of their school. In the end, we managed to (fuzzy) match 632 schools, for which we have both administrative information and working environment information from the 2013 Working Conditions Survey.

4.4.2 Descriptive statistics

Table 4.1 – Socio-demographic characteristics, professional situation and absenteeism of staff in public educational institutions in 2010

	(1)	(2)	(3)	(4)
	All education staff	Primary school teachers	Secondary school teachers	Management staff
Number	933382	349407	375805	14157
Proportion (%)	100	37.4	40.3	1.5
Socio-demographic characteristics				
Age	41.9	40.5	42.8	50.4
Women (%)	69.8	80.4	57.4	45.5
Job characteristics				
Civil servant (%)	87.3	99.4	93.7	99.9
Fixed-term contract (%)	10.9	0.6	5	0.1
Permanent contract (%)	1.8	0	1.3	0
Seniority (Years)	9.5	8.4	13	6.6
Full-time contract (%)	80.7	88.3	87.9	99.9
Substitute member (%)	7.4	10.3	8.5	0.1
Priority Education Network (%)	25.8	27.3	18.8	16.8
Absences during the school year				
At least one sick leave (%)	45.2	52	48.8	13.1
Average duration of sick leave (days)	31.2	28.4	29.1	49.6
Average duration of sick leave < 90 days	11.8	12.1	11.1	17.1
Average duration of sick leave ≥ 90 days	320	304.8	304.4	284.6
At least one maternity leave (%)	4.1	6.4	3.3	0.5
Average duration of maternity leave (days)	110	110.7	109.9	102.3
At least one adoption leave (%)	1.1	1	1.7	0.3
Average duration of adoption leave (days)	15.1	17.6	12.8	12.6
At least one work accident leave (%)	0.9	0.8	1	0.5
Average duration of work accident leave (days)	44.6	38.7	47.5	61.4

Note: Public educational institutions in 2010 includes primary schools, junior high schools, senior high schools, and, to a lesser extent, apprenticeship training centres, distance learning centres, administrative services, teacher training schools. Column (1) describes the characteristics of public education personnel, which includes teachers (78%), school principals (1.5%) but also administrative staff, education and guidance counsellors, library staff, inspection staff and medico-social staff (nurses, psychologists). Columns (2) and (3) describe the characteristics of the primary and secondary school teachers. Column (4) presents the characteristics of management staff which consists mainly in school principals and vice principals. Priority Education Network schools receive increased financial and human resources to compensate for the fact that pupils come from disadvantaged socio-economic backgrounds. Absences are cumulative at the school year level (2010-2011). The average duration per type of absence is calculated for persons absent on at least one day (for this reason) during the school year.

Source: Ministry of Education (DEPP)

¹⁴The 2013 Working Conditions Survey is representative of the employed persons and was conducted between October 2012 and February 2013 in face-to-face interviews with approximately 34,000 people.

Table 4.2 – Characteristics of public educational institutions in 2010

	(1)	(2)	(3)	(4)
	All education establishments	Primary schools	Junior high schools	Senior high schools
Number of establishments	63180	48137	5286	2576
Share of establishments (%)	100	76.2	8.4	4.1
Average number of employees	12.7	5.5	39.9	80.2
Average number of full-time employees	9.8	4.4	29.4	64.2
Schools in the Priority Education Network (%)	13	13.3	20.5	1.6
City location				
Isolated city (%)	1.4	1.3	1.5	1.8
Small-sized urban area (%)	5.1	5.1	4.8	5
Medium-sized urban area (%)	6.7	6.7	6.3	7.3
Large-sized urban area (%)	72	73.3	69.2	63.8

Note: Public educational institutions in 2010 includes primary schools (76%), junior high schools (8.4%), senior high schools (4%), and, to a lesser extent, apprenticeship training centres, distance learning centres, administrative services, teacher training schools. Public institutions employees refer to teachers, school principals, administrative staff, education and guidance counsellors, library staff, inspection staff and medico-social staff (nurses, psychologists). Priority Education Network schools receive increased financial and human resources to compensate for the fact that pupils come from disadvantaged socio-economic backgrounds.

Source: Ministry of Education (DEPP). Institute of Official Statistics (INSEE).

4.5 Method

In this section, we propose a model that aims to disentangle the contributions of the individuals (teachers), workplaces (schools) and managers (school principals) to the duration of absences. Our method is based on the seminal works (Abowd et al., 1999, 2002) that identified the contributions of firms and individuals to the formation of wage inequalities. We also refer to the recent works that have enriched this approach and that have proposed specification tests (Card et al., 2013, 2015).

4.5.1 Specification

A natural way to model a positive variable, like the duration of absence, is to use a model of count data¹⁵ (Cameron and Triverdi, 2005). It assumes that the expected value of the variable of interest Y can be written: $E[Y|X] = \exp(\beta X)$, where X is a set of covariates. In our case, it is particularly interesting to consider this type of multiplicative model because the factors that contribute to absence are likely to interact with each other in a non-linear fashion.

Specifically, we note Y_{ijmt} the annual absence duration of individual i who works at school j , managed by principal m at year t . We write $E[Y_{ijmt}|\alpha_i, \phi_j, \psi_m, X_{it}, \gamma Z_{jt}] = \exp(\alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t)$, where α_i is a time-invariant individual component, ϕ_j a time-invariant school component, ψ_m a time-invariant school principal component,

¹⁵Typically a Poisson regression model or a Negative Binomial regression

X_{it} a vector of time-varying individual characteristics, Z_{jt} a vector of time-varying school characteristics and δ_t a year dummy. To take into account the deviation from the mean, we consider the stochastic version of this model:

$$Y_{ijmt} = \exp(\alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t) \eta_{ijmt} \quad (1),$$

with η_{ijmt} an error term that satisfies $E[\eta_{ijmt} | \alpha_i, \phi_j, \psi_m, X_{it}, Z_{jt}, \delta_t] = 1$.

As mentioned above, sick leave in France must be certified by a doctor. Therefore, the individual fixed effect α_i is supposed to capture primarily the permanent health capital of person i . But this term may also reflect the behavior towards work in general, such as individual i 's effort level or response to moral hazard. The school fixed effect ϕ_j captures any time-invariant component specific to the school j or its immediate environment. These include the level of safety at school, the attitudes of students and parents, the absence norms at school, the climate, the level of pollution or access to health facilities in the city. The school principal fixed effect ψ_m captures the systematic contribution of the school principal m to teacher absences. It reflects the impact of managerial choices, qualities and attitudes. Based on the literature linking absences and psychosocial risks, we expect that principals who provide attention, support and assistance to teachers in difficulty will have a low fixed effect on absences. On the contrary, those who are lax, discouraged or disinvested in their work are more likely to favor the absence of their subordinates, resulting in a higher fixed effect. The time-varying component βX_{it} captures the variation in absences that is related to the change in individual characteristics. It includes the effect of age, seniority, number of children, employment contract (civil servant, open-ended contract, fixed-term contract, probationary period) and a dummy variable equal to one if the person takes an adoption or maternity leave during the year. The vector Z_{jt} contains time-varying characteristics of the establishment j . It includes the size of the school in terms of full-time teachers, the number of managers (principals and vice-principals) per teacher, and the average length of sick leave for school managers. The year fixed effect δ_t captures the annual trend of absences over the time period considered and the error term η_{ijmt} captures any unobserved transitory shock.

In order to compare all the fixed effects of model (1), we must limit the estimation to schools linked by the mobility of teachers and principals. This ensures that fixed effects are estimated based on a single reference group (one teacher, one school and one school

principal) arbitrarily selected from the data. We restrict the estimation to the largest "connected set" of teachers, schools and principals, which represents 59 % of teachers, 61 % of schools and 69 % of school principals¹⁶. In addition, we exclude from the analysis the schools that had only one principal during the reporting period (normally, principals are assigned for a minimum of three years and a maximum of six years) and teachers who are observed only once (for instance teachers recruited in 2015) because their fixed effects cannot be dissociated from temporary shocks η_{ijmt} . At the end, we exploit a sample that consist of 1,721,429 observations, corresponding to 5,245 schools, 8,707 school principals and 283,933 teachers, observed on average 6.1 times over the period 2007-2015¹⁷.

We will discuss the validity of this specification and will provide various robustness checks in section 4.7.

4.5.2 Estimation

We now turn to the estimation strategy of model (1). First of all, we assume that model (1) is correctly specified and that there is no endogeneity problem, which means that $E[\eta_{ijmt} | \alpha_i, \phi_j, \psi_m, X_{it}, Z_{jt}, \delta_t] = 0$. These assumptions will be discussed in more detail in Section 4.7.

The most satisfactory way of estimating model (1) is certainly to use the Pseudo Maximum Likelihood method (PML) proposed by [Gourieroux et al. \(1984\)](#). This method provides robust and unbiased estimates even in the presence of heteroscedasticity, which is not the case with most of the methods used in the literature - for instance, the Non-linear Least Squares estimator (NLS) or the Ordinary Least Squares estimator (OLS) after log-linearization are biased in general ([Silva and Tenreiro, 2006](#); [Santos Silva and Tenreiro, 2011](#)). In this case, the only condition for the Poisson Pseudo Maximum Likelihood (PPML) estimator to be consistent is the correct specification of the conditional mean: $E[Y|X] = \exp(\beta X)$.

The problem is that, to the best of our knowledge, there are no algorithms that propose a feasible and unbiased estimation of three levels of high-dimensional fixed effects (the

¹⁶These relatively low percentages are due to the fact that a number of schools do not have principals in our data. Instead, they usually have assistant directors, but I do not take account of vice-principals in this analysis.

¹⁷I also present result with individual and school fixed effects only. Once I restrict the sample to the largest connected set of individuals and schools, I rely on 2,661,107 observations corresponding to 8,603 schools and 410,553 teachers

vectors α , ϕ and ψ) for non-linear models¹⁸.

The estimation of model (1) after logarithmic linearization seems to be an interesting alternative from this point of view because there are nowadays many algorithms that allow to approach the OLS solution in a feasible and consistent way (usually based on the fixed point method), even when the linear model includes several levels of large scale fixed effects (see for instance the algorithms proposed by [Abowd et al. 1999](#); [Guimarães and Portugal 2010](#); [Arcidiacono et al. 2012](#); [Gaure 2013](#); [Correia 2016](#). Formally, this approach considers the following model:

$$\ln(Y_{ijmt}) = \alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t + \ln(\eta_{ijmt}) \quad (2)$$

However, it should be noted that the models (1) and (2) are not equivalent because of Jensen's inequality. This appears clearly with the expected values of the two previous expressions:

$$\text{Model (1): } E[Y_{ijmt}] = \exp(\alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t)$$

$$\text{Model (2): } E[\ln(Y_{ijmt})] = \alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t$$

while the Jensen inequality imposes that: $\ln(E[Y_{ijmt}]) > E[\ln(Y_{ijmt})]$

Therefore, if model (1) is correctly specified, the coefficients estimated via model (2) (by OLS or using an approximation method) are generally inconsistent. The size of the bias depends on the shape of the heteroscedasticity. Santos [Silva and Tenreiro \(2006\)](#) shows that when model (1) is correctly specified, with $E[\eta_{ijmt}\alpha_i, \phi_j, \psi_m, X_{it}, \gamma Z_{jt}] = 1$, and when the conditional variance of Y_{ijmt} is proportional to the the square of its conditional expectation (i.e. when $\text{Var}(Y_{ijmt}) \propto \exp(\alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t)^2$), then estimating the coefficients of model (1) by PPML or the coefficients of model (2) by OLS leads to the exact same solution. But this is the only case where the two specifications are equivalent. Although it is not entirely possible to test this hypothesis, it is possible to examine the relationship between the variance and the mean of absence duration Y_{ijmt} for well-selected subsets ([Manning and Mullahy, 2001](#)). For example, we can examine the relationship between the average and the variance of absences within schools, which is equivalent to examining the relationship between $\widehat{\text{Var}}[Y_{ijmt}|j]$ and $\widehat{E}[Y_{ijmt}|j]$. Table 4.31 shows the results of the regression of $\ln(\widehat{\text{Var}}[Y_{ijmt}|j])$ on $\ln(\widehat{E}[Y_{ijmt}|j])$ under various conditions. It suggests that the variance of absences is close to being proportional to the

¹⁸Guimaraes (2014) proposes a stata module to estimate a Poisson regression model with up to two high-dimensional fixed effects. We use this module to estimate individual and school fixed effects.

square of the expected value. In this case, the estimation of the model (1) or model (2) with our data should in principle lead to close results.

Another problem is that log linearization is not possible for values of Y_{ijmt} equal to zero. Since every year 52 % of individuals do not take sick leave, there are a majority of nil values in our data. This problem is well known in international trade where the estimation of gravity equations by OLS after log-linearization has long been the dominant approach. However, log-linearization is impossible when there is no exchange between countries. A first solution is to ignore zero values and treat them as missing values. In this case, the coefficients of the linearized model are only estimated for strictly positive Y_{ijmt} values. A second approach is to add a "small constant" to all observations so that the domain of Y becomes strictly positive. Nevertheless, [Silva and Tenreyro \(2006\)](#) and [Santos Silva and Tenreyro \(2011\)](#) have shown that both approaches lead to biased estimates in general, even when the variance of Y_{ijmt} is proportional to the square of its mean. In addition, the authors explain that estimates are not robust to the constant choice and may suffer from significant bias. Intuitively, estimates vary with the value of the "small constant" because of the logarithmic curvature around zero. The estimation of the exponential model by PPML treats small values indiscriminately, while the estimation of the log-linearized model by OLS tends to overweight values close to zero - that is, very negative in logarithm - because the OLS method is very sensitive to extreme values¹⁹. The lower the value of the "small constant", the more negative the logarithm and the greater the distortion on the mean.

In view of the above considerations, the PPML estimator will be preferred wherever possible (i. e. when the model has no more than two high-dimensional fixed effects). In the other cases (when we want to estimate three large fixed effects or peer effects using the algorithm proposed by [Arcidiacono et al. \(2012\)](#) for estimating peer effects for example), we estimate the log-linearized form after adding a constant. To ensure that the choice of the constant does not change our estimates too much, we will present the results obtained by PPML with those obtained by OLS after log-linearization for different values of the constant. We show that a constant between 0.01 and 1 leads for most coefficients to results close to those obtained by PPML; the sign and the order of magnitude are always retained. Therefore, our second preferred method is to estimate the following model by

¹⁹The density diagram of $\ln(Y_{ijmt} + C)$ for different values of the constant C clearly shows that the more you choose a small value for C , the more outliers the zero values are.

OLS:

$$\ln(Y_{ijmt} + C) = \alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t + \ln(\eta_{ijmt}) \quad (3)$$

with $C \in [0.01, 1]$.

In order to make our results comparable with the literature (Ichino and Maggi, 2000; Ose, 2005; Bradley et al., 2007; Puhani and Sonderhof, 2010), I also report in the paper the estimates obtained after log linearization and suppression of zero values (which results in the loss of just over half of the sample) as well as those obtained by estimating an additive model by OLS:

$$Y_{ijmt} = \alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t + \epsilon_{ijmt} \quad (4)$$

with ϵ_{ijmt} an error term that satisfies $E[\epsilon_{ijmt} | \alpha_i, \phi_j, \psi_m, X_{it}, Z_{jt}, \delta_t] = 0$.

Although the estimation of the model (4) allows us to bypass a number of difficulties mentioned in the discussion above, we show later in the paper that our data reject the hypothesis of an additive and separable model. The results are therefore presented for comparison purposes only.

4.6 Results

In this section, I present the results obtained from the estimation methods presented in Section 4.5. The dependent variable is the length of teachers' annual sick leave for model (1), estimated by Poisson Pseudo Maximum Likelihood (PPML), and for model (4), estimated by Ordinary Least Squares (OLS). The explained variable is the logarithm of the length of teachers' annual sick leave for models (2) and (3), estimated by OLS (after eliminating the zero values and adding a constant respectively). I present the estimation results for three high-dimensional fixed effects - individual, principal and school fixed effects - limiting the estimation to models (2) to (4). I also present the results for two high-dimensional fixed effects - individual and school fixed effects - allowing to compare the estimation of models (1) to (4).

I consider several time-varying explanatory variables. I distinguish between four employment statuses: civil servant teachers, civil servant teachers in training (probationary year), teachers on fixed-term contracts (one year renewable every year for six years) and

teachers on open-ended contracts. Civil servants and teachers on open-ended contracts hold tenure but different status (including differences in pay and mobility procedures, see section 4.3 for more details). The time-varying individual characteristics include age, squared age, seniority in current employment status, number of children, a dummy variable indicating whether maternity leave was taken during the year, and a dummy variable indicating whether adoption leave was taken during the year. The time-varying characteristics of schools and principals are the number of national education employees in the school (teachers and non-teachers), the number of principals and vice-principals ("managers") per employee, and the average annual sick leave duration for principals and vice-principals.

Table 4.3 compares for the period 2007-2015 the sample of all full-time teachers assigned to a school during the year (column 1), the sample of all teachers who are observed at least twice and who belong to the largest group of connected teacher-school (column 2), and the sample of all teachers observed at least twice who belong to the largest group of connected teacher-school-principal (column 2). For each of the three samples, Table 4.3 provides descriptive statistics for the dependent and explanatory variables. It reveals that the samples in columns (2) and (3) are relatively representative of full-time teachers assigned to a school (column 1).

Table 4.3 – Comparison of the characteristics of the two main study samples over the period 2007-2015

	(1)		(2)		(3)	
	All full-time teachers assigned to one school during the year		Teachers observed for at least two years who belong to the largest set of connected schools and teachers		Teachers observed for at least two years who belong to the largest set of connected schools, principals and teachers	
Number of observations	2,692,369		2,625,001		1,698,173	
Number of teachers	470,443		403,304		278,781	
Number of schools	8,621		8,609		5,246	
Number of school principals	12,548		12,548		8,707	
Duration of sick leave for teachers (days)						
Mean	10.8		10.6		10.3	
Standard deviation	41.4		40.8		40.1	
Median	0		0		0	
3rd quartile	5		5		5	
9th decile	18		18		17	
99th percentile	282		275		271	
Explanatory variables	Mean	SD	Mean	SD	Mean	SD
Employment contract						
Regular civil servant	0.94	0.24	0.95	0.23	0.95	0.21
Civil servant on probation	0.02	0.12	0.01	0.11	0.01	0.1
Fixed-term contract	0.04	0.2	0.03	0.18	0.03	0.17
Open-ended contract	0.01	0.08	0.01	0.08	0.01	0.08
Characteristics of teachers						
Age (years)	43.6	10.0	43.7	9.9	44.0	9.8
Seniority (years)	13.7	8.5	13.9	8.4	14.1	8.3
Number of children	1.4	1.2	1.4	1.2	1.5	1.2
Maternity leave (%)	3	17	3	17	3	16
Adoption leave (%)	2	13	2	13	2	13
Characteristics of schools and principals						
Number of full-time employees	58.7	38.0	58.8	38.0	63.7	39.2
Number of managers per employee	0.04	0.02	0.04	0.02	0.04	0.02
Duration of sick leave for managers (days)	5.0	24.1	5.0	24.1	5.1	22.8

Note: Column (1) presents the characteristics of secondary school teachers working in public schools over the period 2007-2015, excluding part-time teachers (12%) and substitute teachers who work in more than one school during the year (8%). Column (2) restricts the sample to teachers who are observed at least twice and who belong to the largest connected set of teachers and schools. Column (3) restricts the sample to teachers who are observed at least twice and who belong to the largest connected set of teachers, schools and school principals. Regular civil servants are teachers recruited through competitive examinations who have validated the one-year probationary period. Managers refer to school principals and vice-principals.

Source: Ministry of Education (DEPP).

4.6.1 Contribution of teachers, schools and principals to the annual duration of teacher absences

Tables 4.4 and 4.5 report the estimation results of models (2), (3) and (4) including fixed effects for teachers, schools and school principals. Tables 4.16 and 4.17 show the estimation results of models (1) to (4) including teacher and school fixed effects only.

Table 4.4 (and Table 4.16) presents the estimates of the coefficients for the time-varying variables. All models suggest that individuals are more absent when the security of their employment contract increases. However, individuals with a high fixed effect (a strong propensity to be absent) are more likely to be on open-ended contracts, and even more likely to be on fixed-term contracts²⁰. Overall, civil servant teachers are on average absent two days more than teachers on fixed-term or permanent contracts. As might be

²⁰A result not presented here shows that the fixed component of teachers on fixed-term contracts is one third of a standard deviation higher than the fixed component of civil servants' teachers

expected, the results also show that the duration of teachers' absence increases with age, especially from the age of 50 years onwards (see column 4 in Table 4.16, which corresponds to the estimation of model (1)). The length of absence also increases with seniority. The fact that the length of sick leave increases substantially when a person takes maternity leave during the school year is consistent with Rieck and Telle (2013) that shows that women take substantially more sick leave during pregnancy²¹. This contrasts with the fact that parents who take adoption leave are less likely to take sick leave during this year (partly mechanically since adoption leave reduces the number of calendar days the person may be absent) and with the fact that the length of sick leave decreases with the number of children. The effect of an increase in the number of full-time employees in the school on the duration of absences is small and not robust. The increase in the number of principals per full-time employee in the school appears to be negatively associated with the length of teacher absence, but the result is not always statistically significant. Finally, the increase in the length of sick leave for management staff is not significantly associated with the length of absence of teachers (although there is a slight positive and significant effect for model (1) estimated by PPML. See column 4 of Table 4.16).

Table 4.4 (and Table 4.16) presents a variance decomposition of the length of teacher absences. It shows that despite the large number of fixed effects, models explain at most 38% of the variation in absenteeism duration in models (2) and (3) and 43% in model (1) (Table 4.4 shows the adjusted R-squared statistic because the R-squared statistic is inflated by the large number of parameters considered). This result contrasts sharply with wage decompositions where the explained variance is greater than 90% (e. g. Abowd et al. 1999; Card et al. 2013). It suggests that unobserved time-varying components and random shocks (e. g. health shocks) plays a much more important role in determining the duration of absences than in determining wage²². Tables 4.5 and 4.17 show that the explained share of the variance of absences derives mainly from individual fixed effects. The second largest contribution comes from the school fixed effects, then from

²¹I find that the length of absence does not change significantly in the year preceding or following maternity leave

²²The share of variance explained by the fixed effects of workers, principals and establishments declines mechanically with the length of the period considered and with the number of "movers" between establishments. In this study, teachers are observed on average six times during the 2007-2015 period, which is comparable to (Card et al., 2013) where workers are observed on average five times. In addition, 88% of teachers change schools at least once ("movers") and 70% of principals change schools at least once ("movers") over that period.

the principal fixed effects, and finally from the effect of the observed characteristics that vary over time.

Table 4.4 reports a negative correlation between the fixed effects of schools and principals ($\rho = -0.8$), between the fixed effects of schools and teachers ($\rho = -0.1$) and, to a much lesser extent, between the fixed effects of teachers and principals ($\rho = -0.01$). This result may reflect the sorting of teachers and principals between schools, but it may also be subject to estimation bias. A well-known problem with panel data estimation is that fixed effects estimators become noisier as the number of parameters in the model increases (relative to the number of observations). As a result, the variance of the fixed effects is generally overestimated while the variance of the error term is underestimated (Krueger and Summers, 1988). Moreover, Andrews et al. (2008) demonstrate that the covariance between fixed effects is negatively biased when fixed effects are estimated with error. The authors show that the size of this bias, known as *limited mobility bias*, can be substantial and decreases with the number of individuals who move between establishments (see also Andrews et al. 2012 for an illustration on matched employer-employee data).

To examine the extent of this bias, Figure 4.6 shows the correlation between teacher fixed effects and school fixed effects (Y-axis) as a function of the average number of teachers entering or leaving the school over the 2007-2015 period (X-axis). On the X-axis, schools with the fewest movers over the period considered are gradually eliminated (percentile by percentile by moving to the right along the X-axis) and the number of observations in the sample (school x teacher x year) decreases. The correlation between school fixed effects and teacher fixed effects increases from $\rho = -0.1$ for the entire sample (about 100 "movers" on average per school) to $\rho = -0.04$ when about half of the initial sample is eliminated²³ (schools in this new sample have 150 "movers" on average). We can therefore estimate that the correlation between the effects of schools and teachers is negative and close to $\rho = -0.04$ on average. This result suggests that teachers with the highest propensity to be absent (high fixed effect) are slightly more likely to work in schools that tend to reduce absenteeism (low fixed effect). Figures 4.7 and 4.8 present a similar analysis of the correlations between the fixed effects of teachers and school principals, and between the fixed effects of schools and principals. Figure 4.7 shows that the correlation between the fixed effects of teachers and principals varies very slightly

²³This consists of eliminating 75% of the schools in the sample.

and remains close to $\rho = 0$ when the average number of teachers ("movers") per principal increases over the period 2007-2015. The limited mobility bias is negligible in this case because the average number of movers per school principal is already greater than 150 in the entire sample. Figure 4.8 shows that the correlation between the fixed effects of schools and principals remains negative ($\rho = -0.56$ for schools with 5 principals over 2007-2015) but decreases (in absolute values) as schools run by a small number of principals over the period 2007-2015 are eliminated. This result suggests that the number of principals per school over the period 2007-2015 is too small to know the extent, and potentially the sign, of the correlation between the fixed effects of schools and principals.

Despite these limitations, it is interesting to note that the magnitude of the estimated correlations between teacher, school and principal fixed effects is consistent with the amount of information available during the process of assigning teachers and principals to schools. Teacher assignment is carried out using a central algorithm that takes into account teachers' requests (school ranking) as well as priority criteria such as seniority, family situation and disability status. Teachers can obtain information about schools on the Ministry of Education's website (including the school's ranking in terms of academic performance) to make their choices. However, teachers do not have official information on the head teachers of each school. Therefore, it is expected that teacher fixed effects on absenteeism are more correlated with school fixed effects than with principal fixed effects. The administration studies the mobility requests of school principals individually. Assignment to schools is made on a case-by-case basis, taking into account the profile of school heads and the characteristics of the school. Therefore, it is expected that the fixed effects of schools are more correlated with the fixed effects of principals than with the fixed effects of teachers.

Finally, Tables 4.18 and 4.19 present the decomposition of the variance of the duration of absences for female and male teachers respectively. The results reveal that the contribution of individual, school and principal fixed effects are very similar regardless of the sex of the teacher. The work environment and individual factors therefore seem to play a comparable role in teacher absenteeism, both for men and women. The correlation between the fixed effects of schools (respectively principals) estimated from the male teacher sample and the fixed effects of schools (respectively principals) estimated from the female teacher sample is close to 0.2. This value is probably underestimated due to

estimation errors.

Therefore, in the following study, we will not distinguish the effects of schools and principals on the absences of men and women. Nevertheless, it may be interesting, in future work, to examine whether the occupational risk factors that favour absences are the same for men and women.

Table 4.4 – Contribution of time-varying covariates and fixed effects to the duration of absence of secondary school teachers (Model A)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration $\ln(Y)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model A1	Model A2	Model A3
Employment contract (ref : civil servants)			
Civil servants on probationary year	-3.607*** (0.256)	-0.165*** (0.0184)	-0.310*** (0.0177)
Short-term contracts	-5.077*** (0.358)	-0.287*** (0.0312)	-0.799*** (0.0282)
Open-ended contracts	-4.420*** (0.562)	-0.141*** (0.0444)	-0.519*** (0.0403)
Characteristics of individuals			
Age	-2.393*** (0.0681)	-0.0488*** (0.00408)	-0.0328*** (0.00385)
Age squared	0.0381*** (0.000754)	0.00110*** (0.0000414)	0.000543*** (0.0000395)
Seniority	0.148*** (0.0141)	0.00867*** (0.00134)	0.00675*** (0.000936)
Maternity leave	22.15*** (0.171)	1.428*** (0.00823)	1.930*** (0.0119)
Adoption leave	-1.131*** (0.110)	-0.0600*** (0.0123)	-0.105*** (0.0117)
Number of children	-1.638***	-0.137***	-0.122***
Characteristics of schools and principals			
Number of full-time employees	0.0196*** (0.00714)	-0.000709 (0.000456)	-0.00233*** (0.000409)
Number of managers per employee	-13.17*** (4.887)	-1.049*** (0.278)	-0.211 (0.291)
Average length of sick leave for managers	0.000889 (0.00155)	0.0000652 (0.0000883)	0.0000279 (0.0000896)
Year effects	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes
Principal fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Observations	1697818	719786	1697818
R-squared	0.531	0.548	0.478
Adjusted R-squared	0.433	0.382	0.369

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1), (2) and (3) present the estimation results of the models (4), (2) and (3) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

Table 4.5 – Variance decomposition of the absence duration of secondary school teachers (Model A)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration $\ln(Y)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model A1	Model A2	Model A3
Observations	1697818	719786	1697818
R-squared	0.531	0.548	0.478
Adjusted R-squared	0.433	0.382	0.369
Number of individuals	278755	178688	278755
Number of schools	5243	5236	5243
Number of principals	8702	8690	8702
Var(Y)	1612.29	1.92	4.82
Var(Individual FE)	909.69	1.09	2.18
Var(School FE)	197.83	0.59	0.74
Var(Principal FE)	140.57	0.44	0.58
Var(β x Covariates)	120.92	0.31	0.17
Cov(Individual FE; School FE)	-53.28	-0.14	-0.12
Cov(Individual FE; Principal FE)	-3.08	-0.01	-0.00
Cov(Individual FE; β x Covariates)	-62.83	-0.12	-0.01
Cov(Principal FE; School FE)	-130.49	-0.42	-0.55
Cov(School FE; β x Covariates)	-6.31	-0.01	-0.00
Cov(Principal FE; β x Covariates)	-0.50	0.00	0.00
Var(residual)	756.26	0.87	2.52

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1), (2) and (3) present the estimation results of the models (4), (2) and (3) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

4.6.2 What is the overall impact of schools and principals on teachers absences? A counterfactual analysis

How do schools and principals influence the cumulative number of days of absence in public education? It is difficult to answer this question without a precise estimate of the fixed effects of schools and principals. We know that when the number of fixed effects becomes large, estimates are particularly noisy and dispersion increases artificially. To obtain more accurate estimates, I proceed in two steps. First, I classify schools (respectively principals) into quartiles of fixed effects based on the estimates of model (2) to (4). This first step divides schools (school principals) into four groups that are expected to have a very different impact on teacher absences. Second, I estimate models (2) to (4) by replacing the dummy variables for each school (each principal) with a dummy variable for each quartile of fixed effects. I present the results in the first three columns

of Table 4.6. This procedure guarantees a large number of observations per group in the second stage and leads to more accurate estimates. However, the second stage neglects the heterogeneity of schools and principals within quartiles, which is a potential source of bias. Since quartile segmentation in the first stage is noisy, the estimation of school effects (school principals) is also downward biased in the second stage (attenuation bias). Estimates of school and principal quartile effects for columns (4) and (5) of Table 4.6 are obtained by PPML using model (1). The quartiles in column (4) are based on fixed effects estimated from model (4) and the quartiles in column (5) are based on fixed effects estimated from model (3). Columns (3) to (5) show relatively similar results. All models confirm a substantial increase in the length of absence when teachers move between quartiles of fixed effects. Depending on the model, the length of absence increases by 250-450% when teachers move from the first to the fourth quartile of school fixed effects. Similarly, the duration of absences increases by 170-300% when teachers move from the first to the fourth quartile of principal fixed effects.

I now examine how a change in the effects of schools and principals (e. g. by introducing preventive measures) would affect the annual absence of teachers. For each model in Table 4.6, I simulate the effect of 1) replacing all school effects (respectively principal effects) by their mean effect²⁴, 2) replacing all school effects (principal effects) by the effect of quartiles 1, 2, 3 and 4 respectively, 3) replacing the effect of schools (principals) in quartile 4 by the mean effect, and 4) replacing the effect of schools (principals) in quartile 4 by the effect of schools (principals) in quartile 3. I present the results of this simulation in table 4.7. The first line shows the actual number of days teachers are absent each year: an average of 10.3 days. The additive property of model (4) (result in column 1) implies that the first scenario (replacing all effects of schools and principals by their average effect) leaves the average absence of teachers unchanged. However, model (3) (result in column 2) and model (1) (results in columns 3 and 4) are multiplicative, implying that the school and the principal do not have the same impact for all teachers. In particular, the effects of schools and head teachers increase with the fixed effect of teachers. As a result, replacing all effects of schools and principals with their average effect removes the "protective effect" of teacher sorting, which increases the average absence by 1 to 2 days depending on the model. The latter scenarios (which consist in replacing the

²⁴weighted by the number of full-time teachers in the school

effect of schools (principals) in quartile 4 by their average effect or by the effect of schools (principals) in quartile 3), lead to a decrease in the absence of teachers by an average of one day. This simulation exercise suggests that even a modest intervention to reduce the effects of schools and principals would significantly reduce teacher absences.

Table 4.6 – Impact of schools and school principals on teachers’ absence duration according to their quartile of fixed effects

	(1)	(2)	(3)	(4)	(5)
	Absence duration Y	Log of absence duration ln(Y)	Log of absence duration after transformation ln(Y + 0.1)	Absence duration Poisson(Y)	Absence duration Poisson(Y)
Quartile of establishment FE (ref: first quartile)					
Second quartile	8.691*** (0.248)	0.439*** (0.0124)	0.474*** (0.0128)	0.604*** (0.0241)	0.442*** (0.0278)
Third quartile	15.79*** (0.289)	0.808*** (0.0133)	0.900*** (0.0137)	1.110*** (0.0256)	0.811*** (0.0284)
Fourth quartile	24.99*** (0.366)	1.256*** (0.0153)	1.410*** (0.0153)	1.699*** (0.0287)	1.266*** (0.0320)
Quartile of school principal FE (ref: first quartile)					
Second quartile	6.664*** (0.151)	0.350*** (0.00764)	0.399*** (0.00783)	0.471*** (0.0146)	0.339*** (0.0162)
Third quartile	12.54*** (0.196)	0.664*** (0.00927)	0.746*** (0.00929)	0.891*** (0.0174)	0.638*** (0.0190)
Fourth quartile	19.33*** (0.251)	1.013*** (0.0113)	1.151*** (0.0114)	1.377*** (0.0208)	0.990*** (0.0235)
Employment contract (ref : civil servants)					
Trial year for public servants	-3.903*** (0.230)	-0.196*** (0.0168)	-0.396*** (0.0168)	-0.562*** (0.0364)	-0.537*** (0.0374)
Short-term contracts	-5.665*** (0.329)	-0.300*** (0.0290)	-0.875*** (0.0274)	-0.919*** (0.0598)	-0.916*** (0.0668)
Open-ended contracts	-4.522*** (0.536)	-0.149*** (0.0421)	-0.580*** (0.0391)	-0.552*** (0.0805)	-0.560*** (0.0868)
Individual characteristics					
Age	-2.320*** (0.0627)	-0.0505*** (0.00370)	-0.0864*** (0.00362)	-0.0960*** (0.00796)	-0.103*** (0.00830)
Age ²	0.0369*** (0.000715)	0.00109*** (0.0000386)	0.00132*** (0.0000382)	0.00224*** (0.0000824)	0.00233*** (0.0000859)
Seniority	0.154*** (0.0137)	0.00885*** (0.00129)	0.00913*** (0.000935)	0.00880*** (0.00252)	0.0102*** (0.00263)
Maternity leave	22.14*** (0.169)	1.429*** (0.00802)	1.925*** (0.0118)	1.421*** (0.0132)	1.414*** (0.0132)
Adoption leave	-1.098*** (0.106)	-0.0620*** (0.0120)	-0.113*** (0.0116)	-0.260*** (0.0233)	-0.277*** (0.0237)
Number of children	-1.626*** (0.0914)	-0.137*** (0.00610)	-0.115*** (0.00668)	-0.230*** (0.0122)	-0.227*** (0.0125)
Schools and principals characteristics					
Number of full-time employees	0.0106*** (0.00264)	-0.000902*** (0.000151)	-0.00210*** (0.000146)	0.000352 (0.000304)	-0.000995*** (0.000335)
Number of managers per employee	-5.321 (3.522)	-0.613*** (0.193)	-0.00374 (0.204)	-0.597 (0.368)	-0.736* (0.395)
Average length of management sick leave	0.00156 (0.00119)	0.0000687 (0.0000680)	0.00000309 (0.0000702)	0.000197* (0.000119)	0.000146 (0.000120)
Year effects	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1721429	726105	1721429	1721429	1721429
R-squared	0.523	0.541	0.508		
Adjusted R-squared	0.429	0.388	0.411		
	Quartiles based on FE from model A1	Quartiles based on FE from model A2	Quartiles based on FE from model A3	Quartiles based on FE from model A1	Quartiles based on FE from model A3

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. The variable explained is the total duration (in number of days) of teachers’ sick leave during the school year in column (1), (4) and (5), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3).

To obtain more accurate estimates of school and principal effects on absence duration, I proceed in two steps. First, I classify schools (respectively principals) into quartiles of fixed effects based on the estimates obtained from models (2) to (4) (See columns (1) to (3) of Tables 4.4 and 4.5). This first step allows to divide schools (respectively principals) into four groups that are expected to have a very different impact on teacher absences. Second, I estimate models (2) to (4) by replacing the dummy variables for each school (respectively principal) by a dummy variable for each quartile of fixed effects. I present the results in column (1) to (3) of this Table. Column (4) and (5) report the estimation results by PPML (model 1) when the school and principal fixed effects are replaced by quartile fixed effects estimated from models (4) and (3) respectively. Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in columns (1) to (3) are obtained by OLS, while estimates in column (4) and (5) are obtained by the Poisson Pseudo Maximum Likelihood (PPML) method.

Source: Ministry of Education (DEPP).

Table 4.7 – Average absence duration of secondary school teachers according to different scenarios

	(1)	(2)	(3)	(4)
	Absence duration Y	Log of absence duration after transformation $\ln(Y + 0.1)$	Absence duration Poisson(Y)	Absence duration Poisson(Y)
Annual absence of teachers	10,3	10,3	10,3	10,3
A. School fixed effects (FE)				
Scenarios without teacher sorting				
Replacing school FE by the average FE	10,3	11,5	12,2	11,4
Replacing school FE by the average FE in quartile 1	-1,4	5,8	8	6,2
Replacing school FE by the average FE in quartile 2	7,3	9,3	10,7	9,6
Replacing school FE by the average FE in quartile 3	14,4	14,3	14,1	13,9
Replacing school FE by the average FE in quartile 4	23,6	23,9	19,1	21,9
Scenarios with teacher sorting				
Replacing school FE in quartile 4 by the average FE	7,5	9	9	9,3
Replacing school FE in quartile 4 by the average FE in quartile 3	8,3	9,3	9,2	9,6
B. Principal fixed effects (FE)				
Scenarios without teacher sorting				
Replacing principal FE by the average FE	10,3	11,1	11,6	11,1
Replacing principal FE by the average FE in quartile 1	0,6	6,2	8,3	6,8
Replacing principal FE by the average FE in quartile 2	7,3	9,3	10,5	9,5
Replacing principal FE by the average FE in quartile 3	13,2	13,2	12,6	12,8
Replacing principal FE by the average FE in quartile 4	20	19,8	16,5	18,4
Scenarios with teacher sorting				
Replacing principal FE in quartile 4 by the average FE	8,1	9,2	9,2	9,5
Replacing principal FE in quartile 4 by the average FE in quartile 3	8,7	9,5	9,3	9,7
	Quartiles based on FE of model A1	Quartiles based on FE of model A3	Quartiles based on FE of model A1	Quartiles based on FE of model A3

Note: School and principal quartiles are based on school and principal fixed effects estimated from models (3) and (4) (see columns (3) and (1) of Tables 4.4 and 4.5). The different scenarios consist in replacing the effect of one or more quartiles with that of other quartiles. The scenarios 'without teacher sorting' replace all school and principal effects with their average effects, eliminating the possibility for teachers to self-select into schools based on their own fixed effect.

Source: Ministry of Education (DEPP).

4.6.3 Relationship between the contribution of principals and schools to teacher absences and turnover

We now examine several hypotheses that may explain the impact of schools and principals on teacher absences. We begin by examining the contribution of schools and principals to teacher turnover. We consider the following additive and linear model:

$$T_{ijmt} = \pi_j + \theta_m + \sigma X_{it} + \omega Z_{jt} + \mu_t + \epsilon_{ijmt} \quad (i)$$

where T_{ijmt} represents the ratio of the number of teachers who leave school j at time $t + 1$ divided by the number of teachers in school j at time t , π_j a school fixed effect, θ_m a principal fixed effect, σX_{it} a vector of time-varying individual characteristics, Z_{jt} a time-varying vector of school characteristics, μ_t a year effect and ϵ_{ijmt} an error term that satisfies $E[\epsilon_{ijmt} | \pi_j, \theta_m, \mu_t, X_{it}, Z_{jt}] = 0$.

The results of the estimation are presented in Table 4.20. Model C1 (column 1) includes fixed effects of schools and model C2 (column 2) includes fixed effects of schools and principals. Errors are clustered at the departmental level²⁵. The effects of schools

²⁵There are 100 departments in France

and principals explain about 50% of the variance in teacher turnover. We note that the correlation between the effects of schools and principals on turnover is, as was already the case for absences, strongly negative ($\rho = -0.8$). This correlation is unlikely to be exclusively due to estimation errors since the number of parameters in this model is reasonable relative to the number of observations. Rather, the result suggests that the more the principal encourages teacher turnover, the more likely he or she is to be assigned to a school that tends to reduce teacher turnover. As in the case of absences, this selection is consistent with the assignment of principals on a case-by-case basis. The administration appears to be able to partially balance the effects of schools and principals on teacher (absences and) turnover.

But are the schools (respectively principals) that favor absences the ones that encourage teachers to leave? To answer this question, Table 4.8 presents the correlations between the effects of schools (principals) on absences and the effects of schools (principals) on teacher turnover. The first two lines show that the effects of schools on turnover are positively and significantly associated with the effects of schools on absences, regardless of the estimation method considered (column 1, 2 and 3). The first line corresponds to models with school and principal fixed effects and the second line corresponds to models with school fixed effects only. The third row shows that the effect of school principals on teacher absence is positively and significantly associated with the effect of school principals on teacher turnover. These correlations are relatively low, but noise in fixed effects estimates is likely to generate a significant attenuation bias.

Table 4.8 – Correlations between the effect of schools (resp. school principals) on absences and the effect of schools (resp. school principals) on teacher turnover

	(1)	(2)	(3)
	FE from Model A1/B1	FE from Model A2/B2	FE from Model A3/B3
Corr(School FE on absences ; School FE on turnover) Model A x Model C2	0.0913***	0.1004***	0.0443***
Model B x Model C1	0.1440***	0.1380***	0.1241***
Corr(Principal FE on absences ; Principal FE on turnover) Model A x Model C2	0.0529***	0.0707***	0.0276*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This Table shows the correlation between the effects of schools (respectively principals) on teacher absences and the effects of schools (respectively principals) on teacher turnover for different samples and estimation methods. Models A (for teacher absences) and C2 (for teacher turnover) include school and principal fixed effects, while models B (for teacher absences) and C2 (for teacher turnover) include school fixed effects only. Models A1, A2, A3 (B1, B2, B3) refer to the estimates presented in Table 4.4 (Table 4.16), while models C1 and C2 refer to the estimates presented in Table 4.20.

Source: Ministry of Education (DEPP).

4.6.4 Who leaves schools that contribute most to absences?

What does the correlation between the contribution of schools (respectively principals) to absences and teacher turnover mean? Are teachers who leave schools (head teachers) with a high fixed effect also those who are the most absent? Are they, on the contrary, the least absent? Remind that the rules on teacher mobility do not take into account absences or health problems of teachers, with the exception of a recognized handicap. Assuming that a higher separation rate (for teachers with similar characteristics and occupying the same positions) reflects an unfavorable work environment, the answers to these questions provide us with a first interpretation of the effect of schools and principals on absences. In the first case, schools (school principals) with a high fixed effect "drive away" absent teachers. These schools are not attractive to the teachers most likely to be absent. Therefore, an increase in the fixed effects of schools (principals) is more likely to reflect a deterioration in the health or motivation of these teachers rather than a decrease in their effort. In the second case, high fixed-effect schools cause the least absent teachers to flee. Therefore, an increase in the fixed effects of schools (principals) is likely to reflect absence standards and a permissive work environment. This situation is not suitable for teachers with few absenteeism, who are therefore more likely to leave school than their colleagues who are often absent.

To identify which of these two effects dominates, I consider the following linear probability model:

$$Q_{ijmt} = a + b \widehat{\alpha}_i + c \widehat{\phi}_j + d \widehat{\psi}_m + e (Y_{ijmt} - \overline{Y_{-ijmt}}) + f \widehat{\phi}_j * (Y_{ijmt} - \overline{Y_{-ijmt}}) + g \widehat{\psi}_m * (Y_{ijmt} - \overline{Y_{-ijmt}}) + \sigma X_{it} + \omega Z_{jt} + \mu_t + \epsilon_{ijmt} \quad (iii)$$

where Q_{ijmt} is the probability that teacher i working in school j with principal m on year t moves to another school on year $t + 1$, $\widehat{\alpha}_i$, $\widehat{\phi}_j$ and $\widehat{\psi}_m$ the individual, school and principal fixed effects estimated in a first step, Y_{ijmt} the annual absence duration for individual i , $\overline{Y_{-ijmt}}$ the average absence duration of individual i 's coworkers on year t , X_{it} a time-varying vector of individual characteristics, Z_{jt} a time-varying vector of school characteristics, μ_t the effect of year t and ϵ_{ijmt} an error term that satisfies $E[\epsilon_{ijmt} | \widehat{\alpha}_i, \widehat{\phi}_j, \widehat{\psi}_m, \mu_t, X_{it}, Z_{jt}, Y_{ijmt}] = 0$ ²⁶. Model (iii) allows to examine how the absence differential with colleagues ($Y_{ijmt} - \overline{Y_{-ijmt}}$) interacts with school and principal fixed effects estimated from model (2) to (4).

²⁶Note that $E(Y_{ijmt} - \overline{Y_{-ijmt}}) = (\alpha_i - \overline{\alpha_{-i}}) + (\beta X_{it} - \beta \overline{X_{-it}})$

Table 4.9 shows the estimation result of model (iii). Errors are clustered at department level. It reveals a positive and significant interaction between the difference in absence with colleagues and the fixed effect of the school (respectively the principal). It means that teachers who are absent more than their colleagues are more likely to leave schools (school principals) with high fixed effects. This result supports the hypothesis of detrimental working conditions in schools where the fixed effects of the school or principal are high.

Table 4.9 – Interaction between school (resp. school principal) fixed effects and absence difference between individuals and coworkers

	(1)	(2)	(3)
	FE from Model A1	FE from Model A2	FE from Model A3
Teacher FE on absences (estimated in a 1st step)	0.000420*** (0.0000308)	0.0105*** (0.000893)	0.00717*** (0.000694)
School FE on absences (estimated in a 1st step)	0.000542*** (0.000106)	0.0145*** (0.00124)	0.0132*** (0.00180)
Principal FE on absences (estimated in a 1st step)	0.000556*** (0.000113)	0.0154*** (0.00115)	0.0128*** (0.00146)
Absence difference between individuals and coworkers	0.000711*** (0.0000295)	0.000853*** (0.0000273)	0.000732*** (0.0000258)
School FE on absences (estimated in a 1st step) x Absence difference	0.00000853*** (0.00000168)	0.000214*** (0.0000292)	0.000280*** (0.0000329)
Principal FE on absences (estimated in a 1st step) x Absence difference	0.00000781*** (0.00000205)	0.000188*** (0.0000390)	0.000295*** (0.0000324)
Constant	Yes	Yes	Yes
Year effects	Yes	Yes	Yes
Time varying control variables	Yes	Yes	Yes
Observations	1473925	1018752	1476588

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the department level (standard errors are slightly smaller when estimates are robust to heteroscedasticity only).

Note: This Table shows the estimation results of model (iii). The variable explained is a dummy variable that is equal to 1 for the teachers who change schools the following year, and 0 otherwise. The coefficients of interest (f and g in model (iii)) are in bold in the Table and correspond to the effects of the interaction between the fixed effects of the school (respectively the principal) on teacher absences (estimated from models (2) to (4)) and the absence differential between a teacher and colleagues. To ensure the robustness of the results, the first and last percentiles of the fixed effects of principals and schools are removed from the sample.

Source: Ministry of Education (DEPP).

4.6.5 Relationship between school effects and school characteristics

I now examine whether certain characteristics of schools are significantly associated with their contribution to absences. Table 4.21 shows that the schools that contribute most to absences are also more likely to belong to the Priority Education Network. Figure 4.1 shows the distribution of the Priority Education Network between percentiles of school fixed effect on absences and Figure 4.9 show its distribution among percentiles of school fixed effects on staff turnover. On average, schools that contribute the most to absences (last percentiles) are eight times more likely to be in the Priority Education Network than

schools that contribute the least (first percentiles). Similarly, schools that contribute the most to teacher turnover (last percentiles) are twice as likely to be in the Priority Education Network as schools that contribute the least (first percentiles). This results echoes a recent survey that show that Priority Education students report more serious violent incidents than other high school students and have a somewhat less favourable view of the school climate (DEPP, 2011).

Other school characteristics observed in the administrative data are not systematically associated with school fixed effects (some relationships are significant but not robust between models). However, lower secondary schools and schools in a medium-sized urban area appear to increase teacher absences.

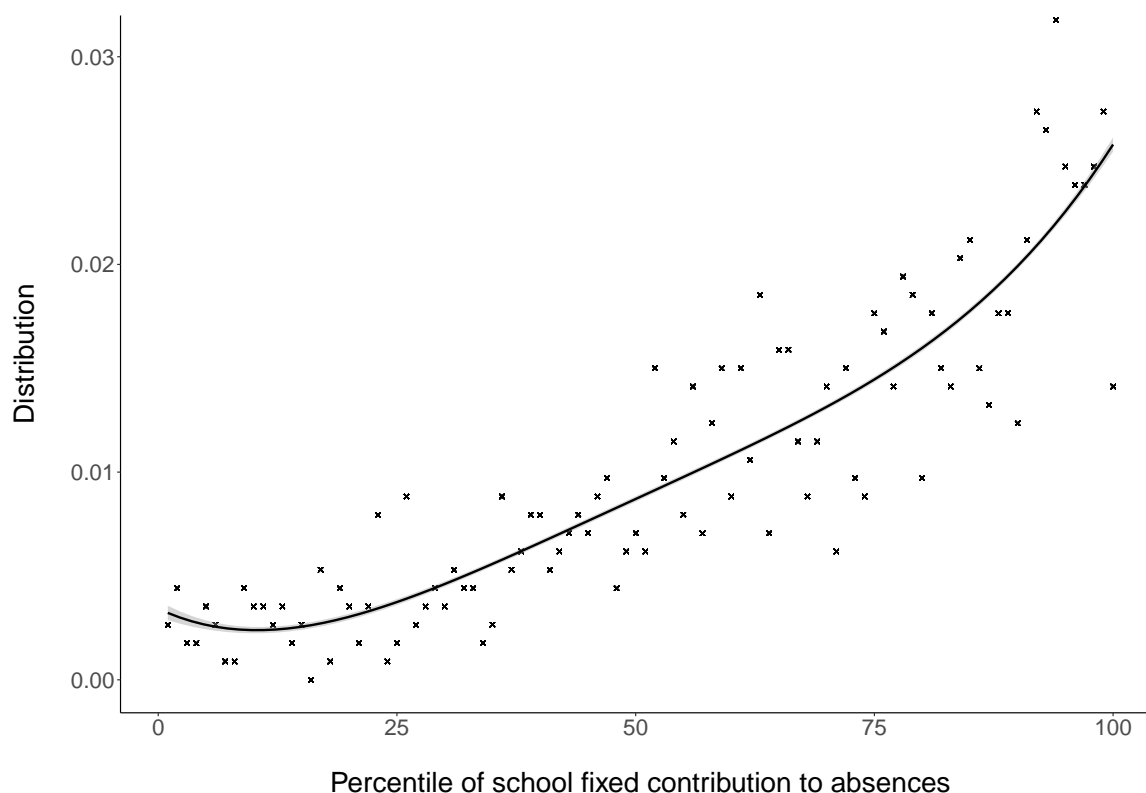


Figure 4.1 – Distribution of the Priority Education Network according to school fixed effects on teachers' absences (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on absences, estimated from model (3). I calculate the number of schools in the Priority Education Network in each school percentile that I divide by the total number of schools in the Priority Education Network. The few schools that joined or left the Priority Education Network during the period 2007-2015 are counted twice. The Y-axis shows the distribution (density) of the Priority Education Network among school percentiles.

4.6.6 What are the working conditions and psychosocial risks factors associated with schools' and principals' effects on teacher absences?

Can schools and principals influence teacher absences through the working conditions they create? Working Conditions surveys provide a very detailed description of the working environment for a sample of employed persons and are particularly well suited to examining this issue. In particular, surveys conducted in 2013 and 2016 in France have three major advantages over surveys of the same type in Europe and the United States. First, they interview a panel of people every 3 years since 2013, which makes longitudinal studies possible (the first wave of interviews took place in 2013 and the second in 2016). Secondly, educational staff (80% of whom are teachers) have been over-represented since 2013 to enable specific studies for this population. Finally, since 2013, the working conditions surveys have interviewed employed persons and, where applicable, their employers. The implementation of this double survey required the collection of information from employees on their employer (name of the establishment, address, activity) so that he could be contacted later. This information can also be used to match the establishments with other data sources.

The working conditions surveys include 1034 public secondary school teachers, of whom 503 were interviewed in 2013 and 2016, and 531 in 2013 or 2016. The attrition rate (non-response) between the two surveys is 40%, which means that 190 more people are in the system in 2016.

To match schools between Ministry of Education administrative records and working conditions surveys, I rely on common information about the employer's establishment. In more than 99% of the cases, teachers interviewed in the working conditions surveys agreed to provide the name of their school. All secondary school teachers interviewed provided the department of their school, 99% provided the postal code and 86% provided the full address of their school (street, number, postal code, city). The surveys also specify the nature of the establishment (primary school, lower secondary school, upper secondary school). This textual information is also present in national education management files. However, some information is missing and reporting errors are possible in both sources.

As a result, I was able to match 707 schools between the two sources ²⁷. These 707 schools correspond to 1113 observations (733 teachers) in the working conditions surveys, which means that the matching rate based on the information provided on the employer is about 72%.

The 2013 and 2016 working conditions surveys include several dozen questions relating to psychosocial risks. To summarize the amount of information available, I begin by grouping the questions around the 6 main risk factors identified in the education literature (Hakanen et al., 2006; Jégo and Guillo, 2016): work intensity, lack of interest (motivation) for work, hostile behaviors, tension with the public (students and their families), lack of support from hierarchy and lack of support from colleagues. Each question is coded by a dummy variable that takes the value 1 if the individual feels exposed or very exposed and 0 if he feels little or not exposed. Table 4.23 presents these variables grouped by psychosocial risk factors. Columns 1 to 4 of Table 4.23 present the share of teachers exposed ²⁸ to each variable for four sub-populations. Column 1 corresponds to all secondary school teachers interviewed by the 2013 and 2016 working conditions surveys. Column 2 corresponds to the teachers surveyed whose school could be matched with the Ministry of Education's administrative records. Column 3 corresponds to the teachers interviewed, whose school could be matched and for whom a school fixed effect could be estimated by model B ²⁹. Lastly, column 4 corresponds to the teachers interviewed, whose school have been matched with administrative records and for whom a school and a principal fixed effect have been estimated by model A ³⁰. The number of observations decreases significantly as the sample restrictions increase (columns 1 to 4). The final sample in column 4 contains 667 observations, 443 teachers and 433 schools, compared to 1537 for the initial sample in column 1. However, Table 4.23 shows that teachers' average exposure to psychosocial risk factors does not change systematically as sample constraints increase and the number of observations decreases.

To summarize teachers' exposure to the 6 major psychosocial risk factors, I consider

²⁷For 86% of the schools, a very large amount of information was consistent between the two sources, which makes us confident about the matching procedure.

²⁸Raw data not adjusted for sampling errors and non-response. To know the share of teachers exposed after weighting refer to Jégo and Guillo (2016)

²⁹These are schools belonging to the largest connected set of schools and teachers during the period studied

³⁰These are schools belonging to the largest connected set of schools, school principals and teachers during the period studied

the unweighted average of the variables that make up each factor. The 6 exposure scores are then normalized (mean zero and unit variance) so that the relative variations within each factor are comparable. The sample resulting from the matching between the working conditions surveys and the administrative sources (column 2 of Table 4.23) is taken as a reference during the standardization procedure. Table 4.24 presents the average teacher score for each of the 6 factors and for the four samples considered (columns 1 to 4). The main lesson is that teachers whose schools have been matched with Ministry of Education management records are more exposed to work intensity, lack of interest in work and tensions with the public (students, parents) than teachers whose schools have not been found in administrative files. However, with the exception of tensions with the public which are significantly higher for the sample in column 3, the differences in exposure are not statistically significant between the samples in columns 2, 3 and 4.

The correlation between teachers' exposure to psychosocial risk factors and the effects of schools and principals on absences (estimated in a first step) is estimated using first difference models. This approach makes it possible to eliminate the time-invariant individual component likely to modify the assessment of working conditions. The subjective component of the description of the work environment, for example because of the individual's optimistic or pessimistic character, can thus be partially neutralized. While this approach has the advantage of limiting selection bias, it requires that the sample be restricted to individuals who were interviewed twice, resulting in an additional loss of observations (this sample includes 414 observations for 2,017 teachers). Moreover, since the effects of schools and principals are fixed over the 2006-2015 period, the only source of identification comes from the fact that teachers change schools (14% of teachers are involved) and/or principals (73% of teachers are involved) between 2013 and 2016.

Specifically, I consider the following models:

$$\widehat{\phi}_j = \alpha_{i(j)} + r_1 W_1 + r_2 W_2 + r_3 W_3 + r_4 W_4 + r_5 W_5 + r_6 W_6 + p \widehat{\psi}_m + \delta_{1t} + \epsilon_{jmi(j)}$$

$$\widehat{\psi}_m = \gamma_{i(j)} + s_1 W_1 + s_2 W_2 + s_3 W_3 + s_4 W_4 + s_5 W_5 + s_6 W_6 + q \widehat{\phi}_j + \delta_{2t} + \nu_{jmi(j)}$$

where $\widehat{\phi}_j$ is the fixed effect of school j , $\widehat{\psi}_m$ is the fixed effect of school principal m working in school j at the time of the working conditions survey, $\alpha_{i(j)}$ and $\gamma_{i(j)}$ are individual fixed effect for teachers interviewed, W is a vector that summarizes exposure to psychosocial risks in the school j at the time of the working conditions survey, δ_{1t} and δ_{2t} are time fixed effects, ϵ_{jmi} and ν_{jmi} are errors terms that I assume independent of the covariates.

Finally, r_1-r_5 and s_1-s_5 are the coefficients of interest.

Table 4.10 presents the coefficients that are statistically significant at the 10% level. The results obtained using the school and principal fixed effects estimated from models (4), (2) and (3) are presented in columns (1), (2) and (3) respectively. Table 4.10 reveals that the school fixed effects tend to increase with the intensity of work as well as with hostile behaviors. On the other hand, they are positively correlated with teachers' motivation for work. When school effects are taken into account, teachers whose principal tends to increase absenteeism (high principal fixed effect) are more likely to report lack of hierarchical support, tensions with the public (students and families), and hostile behaviour. However, they are less likely to feel a lack of interest in work.

These estimates should be viewed with caution due to the absence of a source of causal identification but they tend to support that school and principal effects on absenteeism are more related to difficult working conditions than to teachers' lack of effort or investment in their work.

Table 4.10 – Relationship between working conditions and the effects of schools and principals on absences

	(1)	(2)	(3)
A. Determinant of school FE estimates			
	School FE from model A1	School FE from model A2	School FE from model A3
Work intensity	0.0161 (0.0356)	0.0377 (0.0310)	0.0430* (0.0245)
Lack of interest in work	-0.0492 (0.0375)	-0.0219 (0.0309)	-0.0430* (0.0251)
Hostile behaviors	0.0590 (0.0496)	0.0232 (0.0321)	0.0368 (0.0268)
Control variables			
School principal FE from model A (estimate)	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	472	472	472
B. Determinant of principal FE estimates			
	Principal FE from model A1	Principal FE from model A2	Principal FE from model A3
Lack of support from hierarchy	0.0615* (0.0301)	0.0413 (0.0302)	0.0720*** (0.0250)
Tensions with students and their families	0.0688* (0.0413)	0.0473 (0.0409)	0.0352 (0.0268)
Lack of interest in work	-0.0632 (0.0628)	-0.0331 (0.0493)	-0.0314 (0.0367)
Hostile behaviors	0.0615* (0.0304)	0.0155 (0.0385)	0.0103 (0.0264)
Control variables			
School FE from model A (estimate)	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	472	472	472

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This table explains the fixed effects of schools (respectively principals) on the duration of teacher absence through exposure to psychosocial risk factors. The fixed effects of schools and principals considered in columns (1), (2) and (3) are estimated from models (4), (3) and (2) respectively. The sample is composed of secondary school teachers, randomly selected in 2013, who responded to the Working Conditions Survey in 2013 and 2016, and whose school (in 2013 and 2016) could be matched with the administrative records of the Ministry of Education. The six psychosocial risk factors considered are: work intensity, lack of interest in work, hostile behaviour, tensions with the public, lack of support from the hierarchy and lack of support from colleagues. Each psychosocial risk indicators is standardized (mean 0 and variance 1). The Table shows the variables that are statistically significant at the 10% level. Each column includes teacher fixed effects and year effects as control variables. When the variable explained is the effect of school on the duration of absences (upper part of the Table), I add the fixed effects of principals to the control variables. When the variable explained is the effect of principal on the duration of absences (lower part of the Table), I add the fixed effects of schools to the control variables.

Source: Ministry of Education (DEPP) and Working Conditions Surveys, Ministry of Labor (DARES)

4.6.7 What is the relationship between teacher health and school and principal effects on absenteeism?

The 2013 and 2016 working conditions surveys make it possible to examine how the fixed effects of schools and principals on absenteeism are related to teachers' subjective health and well-being.

To measure general health status, working conditions surveys ask respondents to judge

whether their health status is "very good", "good", "fair", "poor" or "very poor".

In addition, five questions are used to construct the World Health Organization's psychological well-being index (WHO-5): « I have felt cheerful and in good spirits », « I have felt calm and relaxed », « I have felt active and vigorous », « I woke up feeling fresh and rested » and « My daily life has been filled with things that interest me ». Each of the 5 items is scored from 0 (never) to 5 (all the time).

The WHO-5 index measures psychological well-being (sometimes called mental health) during the 14 days preceding the survey on a scale from 0 (minimum well-being) to 25 (maximum well-being). Several studies have shown that WHO-5 is able to detect minor (WHO-5 < 8) and major (WHO-5 < 6) depressive episodes (Henkel et al., 2003; Krieger et al., 2014). A low WHO-5 score also correlates with cardiovascular risk (Birket-smith et al., 2009).

To examine the relationship between school, principal and teacher's subjective health, I construct a binary variable equal to 1 if the respondent teacher judges his or her overall health to be "poor" or "very poor", and 0 if not. To identify the degree of psychological malaise, I construct 3 binary variables that take the value 1 if the respondent has a subjective well-being score of less than 10, 8 and 6 respectively, and 0 otherwise. According to the literature, these three variables can detect increasing levels of risk and severity of depression (Krieger et al., 2014). For each health variable, I consider the following model:

$$Y_{ijmt} = \alpha_i + \gamma X_{it} + \beta_m \widehat{\psi}_m + \beta_j \widehat{\phi}_j + \delta_t + \epsilon_{ijmt}$$

where Y_{ijmt} is a binary variable that represents the general health status or psychological well-being of teacher i working in school j managed by principal m at time t , α_i is a teacher fixed effect, X_{it} is a vector of teacher age and age squared, $\widehat{\psi}_m$ is principal m 's fixed effect on absenteeism (estimated in a first step), $\widehat{\phi}_j$ is school j 's fixed effect on absenteeism (estimated in a first step), δ_t is a year fixed effect and ϵ_{ijmt} is the error term.

Column 1 of Table 4.11 shows that school effects ϕ_j and principal effects ψ_m on absenteeism are not significantly associated with teachers' overall (self-reported) health status, regardless of the model used to estimate them. Similarly, columns 2 to 4 of Table 4.11 reveal that teachers' psychological well-being is generally not significantly associated with principals' effects on absences. However, the correlation is positive when principal effects are estimated by model A3 (preferred specification). Finally, columns 2 to 4 of Table 4.11 show that the effects of schools on absenteeism are strongly and positively associated with

teachers' psychological discomfort for the three WHO-5 thresholds considered, regardless of the model used (A1, A2, A3). However, these effects are generally insignificant at the usual levels when I take into account the fact that the fixed effects of schools and principals are estimated in a first step (that is, when standard deviations are robust to heteroscedasticity).

These results suggest that the effects of schools and, to a lesser extent, principals on absences more likely reflect a deterioration in psychological well-being than a deterioration in teachers' subjective health.

Table 4.11 – Relationship between the effects of schools and principals on absences and teachers' health status

	(1)	(2)	(3)	(4)
	Self-assessed health status: poor or very poor	WHO 5 subjective psychological well-being index ≤ 6	WHO 5 subjective psychological well-being index ≤ 8	WHO 5 subjective psychological well-being index ≤ 10
Model A1				
School FE	-0.0193 (0.0499)	0.0894 (0.0713)	0.0885 (0.0761)	0.104 (0.0797)
School principal FE	-0.00708 (0.0271)	-0.0149 (0.0236)	-0.00146 (0.0341)	0.00636 (0.0359)
Model A2				
School FE	-0.0344 (0.0681)	0.0964 (0.0812)	0.0426 (0.0970)	0.218** (0.0944)
School principal FE	0.0240 (0.0313)	-0.0142 (0.0540)	0.00507 (0.0434)	0.0935 (0.0644)
Model A3				
School FE	-0.0127 (0.0338)	0.165* (0.0810)	0.141 (0.112)	0.186 (0.119)
School principal FE	-0.00630 (0.0324)	0.00728 (0.0617)	0.0491 (0.0505)	0.0951 (0.0744)
Control variables				
Teacher FE	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes
Age squared	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	472	472	472	472

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This table explains self-reported health of secondary school teachers by the fixed effects of schools and principals on absences. The sample is composed of secondary school teachers, randomly selected in 2013, who responded to the Working Conditions Survey in 2013 and 2016, and whose school (in 2013 and 2016) could be matched with the administrative records of the Ministry of Education. Self-reported health indicators are: general health status judged "poor" or "very poor", WHO-5 psychological well-being index less than 6, WHO-5 psychological well-being index less than 8 and WHO-5 psychological well-being index less than 10. Each column includes age, square age, teacher fixed effects and year effects as control variables.

Source: Ministry of Education (DEPP) and Working Conditions Surveys, Ministry of Labor (DARES)

4.6.8 Has the increase in prevention from 2014 been more intense in schools (and among school principals) that increase teacher absences?

On 22 October 2013, representatives of trade unions and public service employers signed an agreement on the prevention of psychosocial risk factors in public institutions. It requires each employer to assess occupational risks and implement a prevention plan between 2014 and 2015. In addition, the government launched in 2014 a national occupational risk awareness campaign. Training for public employers and workers' representatives were strengthened (awareness of psychosocial risks, provision of tools to diagnose and prevent psychosocial risks, counselling) and employers were encouraged to train staff in psychosocial risk prevention.

In public secondary schools, principals are responsible for the safety and protection of the health of public employees³¹. Since the October 2013 agreement, school principals have been responsible for implementing a psychosocial risk prevention plan that involves employee representative bodies (trade unions, health and safety committees, prevention delegates) and prevention professionals. It can therefore be expected that there were a general improvement in knowledge and practices in psychosocial risk prevention between 2013 and 2016.

Since 2013, staff turnover and sick leave have been among the main indicators used by trade unions and public employers to monitor exposure to psychosocial risks. In the previous sections, we observed a link between the effects of schools and principals on teacher absences, psychosocial risk factors and psychological well-being. In this section, I examine whether the prevention measures put in place from 2014 onwards have tended to target schools and headteachers who appear to pose a risk to the health and motivation of teachers. The aim is not to evaluate the effectiveness of the preventive actions implemented because their effect (on reducing absences, staff turnover, exposure to psychosocial risks) is unlikely to appear in the short term, as confirmed by the lack of improvement in teachers' working conditions and health between 2013 and 2016³². The

³¹Since 2001, all public and private companies have had to record occupational exposures in a single document made available to all employees.

³²These results are not presented in Tables 4.10 and 4.11. I note a significant decrease in hostile behaviour between 2013 and 2016 but this concerns all professions (see Ministry of Labour publication: <https://dares.travail-emploi.gouv.fr/IMG/pdf/2017-082v3.pdf>)

objective is to examine whether the increase in psychosocial risk prevention actions from 2014 onwards has been particularly significant in schools and among principals who seem to need them most.

Three questions relating to occupational risk prevention are common to the 2013 and 2016 working conditions surveys: "In the past 12 months, have you had any knowledge of a document written by management describing the risks associated with working in your institution?", "In the past 12 months, have you received information about the risks your work poses to your health or safety?" and "In the past 12 months, have you received security training from your administration?".

Occupational physicians provide medical supervision for educational staff. They are independent and carry out random medical examinations or at the request of the staff. Doctors and prevention assistants play an advisory role to the school management with regard to working conditions and staff protection. The 2013 agreement provided for an increase in the number of doctors and the strengthening of their autonomy. We consider the following question in the working conditions surveys: "When was the last time you consulted an occupational doctor or a preventive doctor?"

We code the four questions of interest into binary variables (equal to 1 if the respondent answers "yes", 0 if not) and we consider the following model:

$$Y_{ijmt} = \alpha_i + \beta_{jt}\widehat{\phi}_j + \beta_{mt}\widehat{\psi}_m + \delta_t + \epsilon_{ijmt}$$

where Y_{ijmt} is equal to 1 if teacher i answers "Yes" and 0 otherwise, α_i is a teacher fixed effect, $\widehat{\phi}_j$ and $\widehat{\psi}_m$ are school and principal fixed effects on absenteeism, δ_t is a year effect and ϵ_{ijmt} an error term. In this specification, the coefficients of interest β_{mt} and β_{jt} vary with time t , meaning that the coefficients associated with school and principal effects are estimated for the years 2013 and 2016 separately.

To avoid endogeneity problems, the effects of schools $\widehat{\phi}_j$ and school principals $\widehat{\psi}_m$ on teacher absences are estimated for the period 2007-2013 (and no longer for the period 2007-2015). Thus, the values of the variables of interest are not affected by prevention actions implemented from 2014.

Table 4.12 presents the estimation results for the 4 prevention variables considered. For each model A1, A2, A3 used to estimate the fixed effects of schools and school principals on teacher absences, I report the value of the coefficient δ_{2016} that captures the change in prevention practices between 2013 and 2016, the coefficient β_{j2013} (respec-

tively β_{m2013}) that captures the correlation in 2013 between prevention and school effects (respectively principal effects) on teacher absenteeism, the coefficient $\beta_{j2016} - \beta_{j2013}$ (respectively $\beta_{m2016} - \beta_{m2013}$) that captures the change in the correlation between prevention and school effects (respectively principal effects) between 2013 and 2016.

Table 4.12 reveals a significant increase between 2013 and 2016 in the proportion of teachers who are aware of a written document listing occupational hazards in their school (+12 percentage points) and who have received information on the impact of professional practice conditions on their health and safety (+7 percentage points). There was also a slight (but statistically insignificant) improvement in the proportion of teachers who received safety training (+4.5 percentage points) and medical examination by an occupational physician in the last 24 months (+1.5 percentage points).

However, prevention practices between 2013 and 2016 do not seem to have increased any further in schools that increase teacher absenteeism during the 2007-2013 period. Similarly, the increase in prevention between 2013 and 2016 is not correlated with the effects of school principals on teacher absenteeism during the 2007-2013 period. Despite the priority given to psychosocial risks in the public service since 2014, these results suggest that current prevention measures do not specifically target schools and principals who contribute most to increasing teacher absenteeism and turnover.

Table 4.12 – Relationship between the level of prevention in schools and the effects of schools and principals on absences

	(1)	(2)	(3)	(4)
	Have had knowledge in the last 12 months of a document written by management describing the risks associated with working in their institution	Have received information in the past 12 months about the health or safety risks related to their work	Have received security training from their administration within the past 12 months	Have undergone a medical examination by an occupational physician within the last 24 months
Model A1				
Year 2016 (ref: Year 2013)	0.123*** (0.0294)	0.0685** (0.0283)	0.0455 (0.0361)	0.0142 (0.0213)
School FE x Year 2013	-0.0283 (0.0813)	0.138* (0.0785)	0.0306 (0.0999)	0.0254 (0.0589)
School principal FE x Year 2013	-0.132* (0.0714)	0.0658 (0.0689)	0.0706 (0.0877)	-0.0153 (0.0517)
School FE x (Year 2016 - Year 2013)	0.0403 (0.0529)	-0.0415 (0.0511)	0.0167 (0.0651)	0.0185 (0.0383)
School principal FE x (Year 2016 - Year 2013)	0.0646 (0.0532)	0.0347 (0.0513)	-0.00159 (0.0653)	0.0292 (0.0385)
Model A2				
Year 2016 (ref: Year 2013)	0.121*** (0.0296)	0.0724** (0.0283)	0.0493 (0.0361)	0.0166 (0.0213)
School FE x Year 2013	-0.160* (0.0874)	0.0492 (0.0837)	-0.0834 (0.107)	0.00961 (0.0630)
School principal FE x Year 2013	-0.102 (0.0683)	0.107 (0.0653)	0.0467 (0.0832)	-0.0138 (0.0492)
School FE x (Year 2016 - Year 2013)	0.0277 (0.0501)	-0.0881* (0.0480)	0.0749 (0.0611)	0.0296 (0.0361)
School principal FE x (Year 2016 - Year 2013)	0.0384 (0.0538)	-0.0133 (0.0515)	0.0528 (0.0655)	0.0549 (0.0387)
Model A3				
Year 2016 (ref: Year 2013)	0.125*** (0.0296)	0.0667** (0.0287)	0.0463 (0.0360)	0.0131 (0.0209)
School FE x Year 2013	0.0201 (0.111)	0.0715 (0.108)	0.173 (0.135)	0.128 (0.0784)
School principal FE x Year 2013	-0.00154 (0.0860)	-0.0340 (0.0833)	0.0854 (0.105)	-0.0113 (0.0607)
School FE x (Year 2016 - Year 2013)	0.000922 (0.0583)	0.101* (0.0565)	0.0428 (0.0711)	-0.00549 (0.0412)
School principal FE x (Year 2016 - Year 2013)	0.0468 (0.0590)	0.107* (0.0571)	0.0402 (0.0719)	0.0287 (0.0417)
Control variables				
Teacher FE	Yes	Yes	Yes	Yes
Observations	472	472	472	472

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: This table examines whether the increase in psychosocial risk prevention between 2013 and 2016 was more pronounced in schools with a high fixed effect on teacher absences, or in schools run by principals with a high fixed effect on teacher absences. The sample is composed of secondary school teachers, randomly selected in 2013, who responded to the Survey of Working Conditions in 2013 and 2016, and whose school (in 2013 and 2016) could be found in the administrative records of the Ministry of Education. The variables explained in columns (1) to (4) are indicators of preventive actions that take the value 1 if the teachers interviewed answered "Yes", and 0 if not. The coefficients of interest are in bold in the Table and reflect the interaction between the effects of schools (respectively principals) on absences and the prevention differential between 2013 and 2016. Each column includes teacher fixed effects.

Source: Ministry of Education (DEPP) and Working Conditions Surveys, Ministry of Labor (DARES)

4.6.9 Are newly recruited teachers assigned to more favourable work environments?

Newly recruited civil servant teachers are first assigned to a school for a period of one year. This is a trial period at the end of which the vast majority (98%) of them obtain

their tenure and must participate in the mobility process described in section 4.3.3. This initial assignment process consists of two steps. First, new teachers rank schools according to their preferences. Second, an algorithm assigns trainee teachers to schools based on priority criteria such as disability status, spousal reunification and competition rankings. In principle, newly recruited teachers are assigned to schools offering favourable conditions for initial training.

Figure 4.2 shows the distribution of teachers by percentiles of the school's effect on absences (X-axis) for different levels of teaching experience. It shows that teachers on probation (solid lines) are more likely to be assigned to low-effect schools. Only 17% of pre-service teachers are assigned to a school in the fourth quartile of fixed effects and less than 6% to a school in the last decile. On the other hand, tenure teachers with five years of experience or less are more likely to be assigned to schools that have a high impact on absences than to schools that have a low impact. The situation is then gradually reversed with experience, as the more experienced teachers are more likely to work in a school with little effect on absences. The distribution of teachers with 15 years' experience is similar to that of pre-service teachers. After 20 years of experience, teachers have a 15% chance of being assigned to a school in the fourth quartile and only 5% chance of being assigned to a school in the last decile. In other words, teachers with 20 years of experience are twice as likely (respectively half as likely) to be assigned to a first quartile school (respectively last quartile school) as teachers with one year of experience.

Figure 4.10 shows a comparable situation when considering the percentiles of school contributions to teacher turnover rather than absences. Pre-service teachers are relatively uniformly assigned in schools, as are teachers with ten years of experience. However, tenured teachers with five years of experience or less are more likely to be assigned to a school with high staff turnover. After 20 years of experience, teachers are 60% more likely to attend a first quartile school than teachers with one year of experience. As a result, the more experienced teachers are, the more likely they are to obtain an assignment in a supportive work environment. This means that teachers have information about the working context in schools and tend to avoid schools that contribute most to absences and turnover. Assignment rules that give priority to experienced teachers favour the concentration of young and less experienced teachers in schools where working conditions are more difficult (although pre-service teachers are better preserved).

Figures 4.11 and 4.12 show how non-civil servant teachers (short contracts or open-ended contracts) are distributed between the percentiles of school fixed effects on absences and staff turnover respectively. Both Figures show that throughout their careers, non-civil servant teachers are evenly distributed in schools, regardless of their contribution to absences and turnover. This result confirms that it is the priority rules of the mobility procedure for civil servant teachers that lead to unequal exposure in terms of working conditions during teachers' careers.

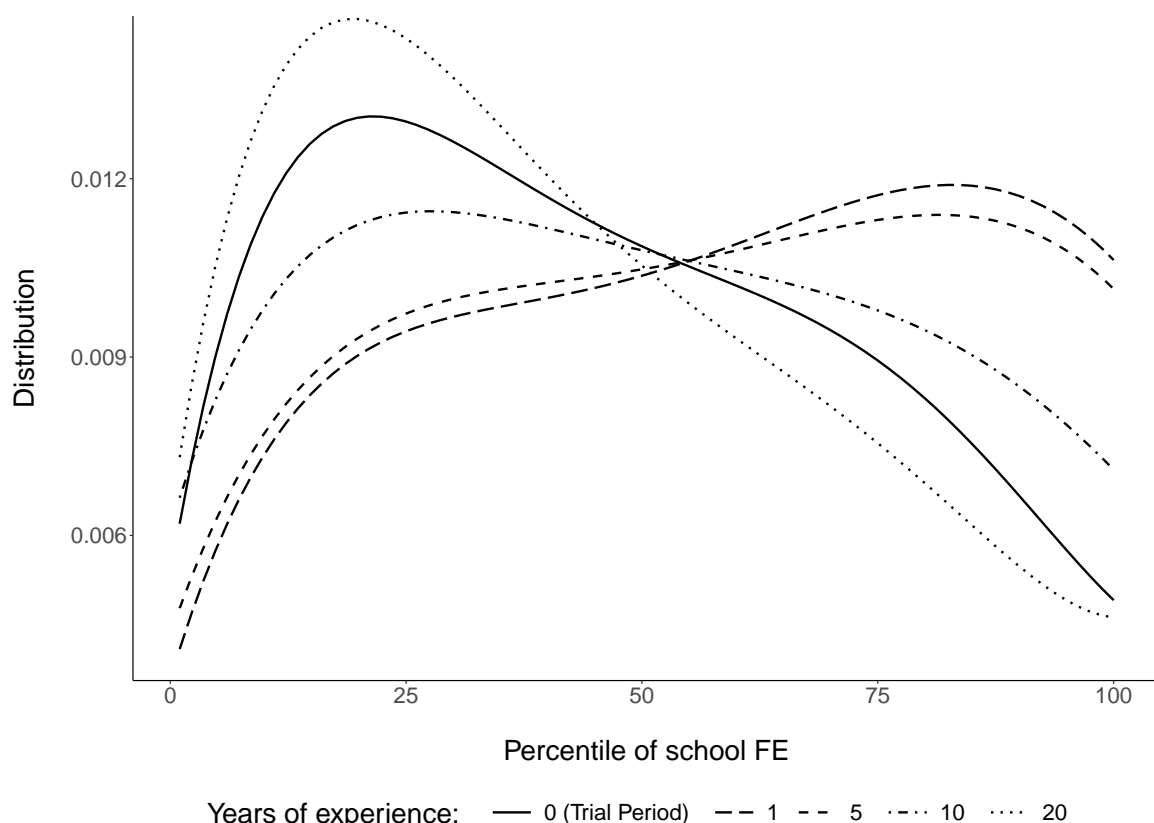


Figure 4.2 – Distribution of civil servant teachers according to experience and school fixed effects on teachers' absences (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on absences, estimated from model (3). I consider 5 levels of professional experience in teaching: probationary period (1 year), between 1 and 5 years of experience, between 6 and 10 years of experience and more than 10 years of experience. For each level of experience, I calculate the number of civil servant teachers in each school percentile that I divide by the total number of teachers with that level of experience. The Y-axis shows the distribution (density) of civil servant teachers among school percentiles according to their level of experience in teaching.

4.7 Robustness checks

The identification of the fixed effects of models (1) to (4) is based on the mobility of teachers and head teachers between schools over the period 2007-2015. Consequently, three hypotheses are necessary for the parameters of models (1) to (4) to be consistently estimated: 1) the model is correctly specified, 2) the mobility of teacher and school principals is exogenous, and 3) the unobserved shocks are conditionally independent of the explanatory variables. The first two assumptions are discussed in this section.

4.7.1 Exogeneous mobility assumption

Teacher mobility and school fixed effects

Recall the three stages of the mobility process for civil servant teachers: 1) Newly recruited teachers (in a trial period) and those in temporary positions are obliged to change schools the following year; the other tenure teachers decide whether or not to apply for a new assignment. 2) Teachers rank schools according to their preferences and an algorithm assigns teachers according to priority criteria (mainly seniority, family status and disability status). 3) Newly recruited teachers and teachers in temporary positions are automatically assigned to their new schools ; the other tenure teachers can refuse their new assignment and remain in their current school. There are therefore several possibilities for teacher self-selection: when they decide to change schools, when they classify schools, and when they decide to accept a new assignment. However, the self-selection of teachers at the beginning of their careers is much more limited: they are more likely to be forced to change schools (because they are more likely to have a temporary position), they are not given priority in the assignment process (they also have little knowledge about the reputation of schools) and they are more likely to have to accept a new assignment.

To examine whether teacher self-selection [sorting] contributes to distorting the estimation of school fixed effects in the multiplicative model, we consider an event study such as (Card et al., 2013). We start by dividing schools into three terciles according to their contribution to absences (based on the fixed effects estimated in a first step). If the model is correctly identified, the schools in the first tercile (respectively in the last tercile) are supposed to reduce (respectively increase) teacher absences compared to the average of the schools (second tercile). Next, we consider all teachers observed for six

consecutive years ($t=-3, t=-2, t=-1, t=0, t=1, t=2$), who change schools in the year $t=0$, but who remain in the same school tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. We assign these teachers to nine transition groups according to the school tercile in $t=-1$ and the school tercile in $t=0$. We then calculate the logarithm of teachers' average absences each year in these nine groups. The results for the nine cohorts are presented in the Table 4.26. Figure 4.3 illustrates the results for teachers who leave the first tercile (low fixed effects on absences: dotted lines) or the third tercile (high fixed effects on absences: solid lines) of school fixed effects. First, Figure 4.3 shows that teachers' absence behaviour is relatively homogeneous within school terciles, regardless of the type of transition occurring in $t=0$. On the other hand, Figure 4.3 shows a strong variation in the absence of teachers when they move between school terciles in $t=0$. The average duration of absences is almost divided by two (-0.6 logarithmic points) for teachers who leave the third school tercile to reach the first school tercile, and it almost doubles (+0.6 logarithmic points) for those who move in the opposite direction (from the first to the third tercile). Similarly, absences increase by 28% (+0.25 logarithmic points) when teachers move from the first to the second tercile and decrease by 26% (-0.3 logarithmic points) when teachers move from the third to the second tercile. These results contrast with the relative stability of absences for teachers who change schools in $t=0$ but remain in the same school tercile³³ (first and third terciles on Figure 4.3). These absence gains and losses are confirmed for all transitions in Table 4.26. The fact that teacher absences vary considerably between terciles of school fixed effects for the same cohort of individuals suggests that our estimation of school fixed effects is not completely biased by teacher selection. We look at this issue in more detail.

Using the expressions of (Card et al., 2013), we consider three main sources of endogenous mobility: a "shock effect", a "drift effect" and a "match effect". We examine their impact on our estimates in light of the results presented in Table 4.26 and Figure 4.3.

A first risk of endogeneity arises if teachers change schools more often in $t=0$ after receiving a positive (respectively negative) shock in $t=-1$, which results in a transitory increase (respectively a decrease) in absences in $t=-1$ ("Shock effect" assumption). In this case, our estimates will tend to underestimate (respectively overestimate) the high-

³³Note that the mean absences shown in Table 4.26 and in Figure 4.3 are not corrected for the effect of temporal trends, age or experience.

fixed effects and overestimate (respectively underestimate) the low-fixed effects. In other words, this type of bias will distort the distribution of establishment fixed effects. Table 4.26 and Figure 4.3 show that there is no systematic increase (or decrease) in absences in $t=-1$ compared to years $t=-2$ and $t=-3$, even for teachers who remain in the same school throughout the period considered. Therefore, the assumption that teacher mobility is caused by systematically positive or negative shocks seems to be rejected. Also, the nature of the shock in $t=-1$ does not predict the type of transition in $t=0$.

A second risk of endogeneity arises if teachers who are increasingly (respectively less and less) absent over time have different mobility paths than other teachers ("drift effect" hypothesis). For example, if teachers who tend to be absent a little longer (respectively a little shorter) each year are more likely to leave schools with a high (low) fixed effect to join schools with a lower (higher) fixed effect, then our estimates will tend to reduce the dispersion of school effects; in other words, we underestimate the contribution of schools to absences. Conversely, if teachers who tend to be absent a little longer (respectively a little shorter) each year are more likely to leave schools with a low (high) fixed effect to join schools with a higher (lower) fixed effect, then our estimates will tend to increase the dispersion of school effects ; that is, we overestimate the contribution of schools to absences. This type of situation could occur, for example, if teachers whose health status deteriorates over time are more likely to be assigned to schools with a low fixed effect on absences. A first counter-argument is that absences and health status are not taken into account in teachers' assignment rules. However, health status may be correlated with some variables considered in the assignment algorithm, such as disability status and seniority (growing with age). It is therefore necessary to examine in more detail whether the past pattern of absences predicts the type of transition between terciles of school fixed effects. Figure 4.3 shows that absences tend to decrease slightly in the three years preceding the date $t=0$, but this trend is shared by all terciles regardless of the type of transition in $t=0$. After adjusting for annual trends (not shown in the paper), we see no trend difference before date $t=0$ for the nine possible transitions between terciles of establishments.

A third risk of endogeneity arises if teachers select themselves in schools that offer them a better "fit", regardless of the schools' tercile ("match effect" hypothesis). This is a credible risk insofar as teachers rank schools according to their preferences and may

or may not accept the school assigned to them. The consequences on our estimates differ according to whether absences mainly reflect a situation that teachers are fleeing (for example because it contributes to deteriorating health status) or are looking for (low hierarchical control). In the (most likely) case where high fixed effects reflect a degraded working environment, teachers may transit more often to schools that further reduce their absences thanks to a better match component. If the estimation of our fixed effects is biased by the existence of this matching effect, we should observe an asymmetry between the absence gains and losses associated with the transitions between terciles (Card et al., 2013). Assuming that better matching results in a systematic decrease in absences ($Match;effect < 0$), the absences of teachers who move from the third tercile to the first tercile of school fixed effects should decrease further in absolute terms ($FE_{1st\ tercile} - FE_{3th\ tercile} + Match\ effect$) than increase the absences of teachers who transit between the first tercile and the third tercile of school fixed effects ($FE_{3th\ tercile} - FE_{1st\ tercile} + Match\ effect$). Table 4.26 and Figure 4.3 contradict this hypothesis. First, the absences of teachers who change schools in $t=0$ but remain in the same tercile do not diminish, contrary to what is predicted by the hypothesis of mobility in favor of the best matches. Second, absences vary symmetrically when teachers move from a lower to a higher tercile of school fixed effect and vice versa. These results suggest that the matching effect, if any, is low enough to ensure that the estimation of the school fixed effects is not unduly affected.

We are therefore relatively confident that the estimates of school fixed effects do not suffer from the endogeneity of the teacher mobility process.

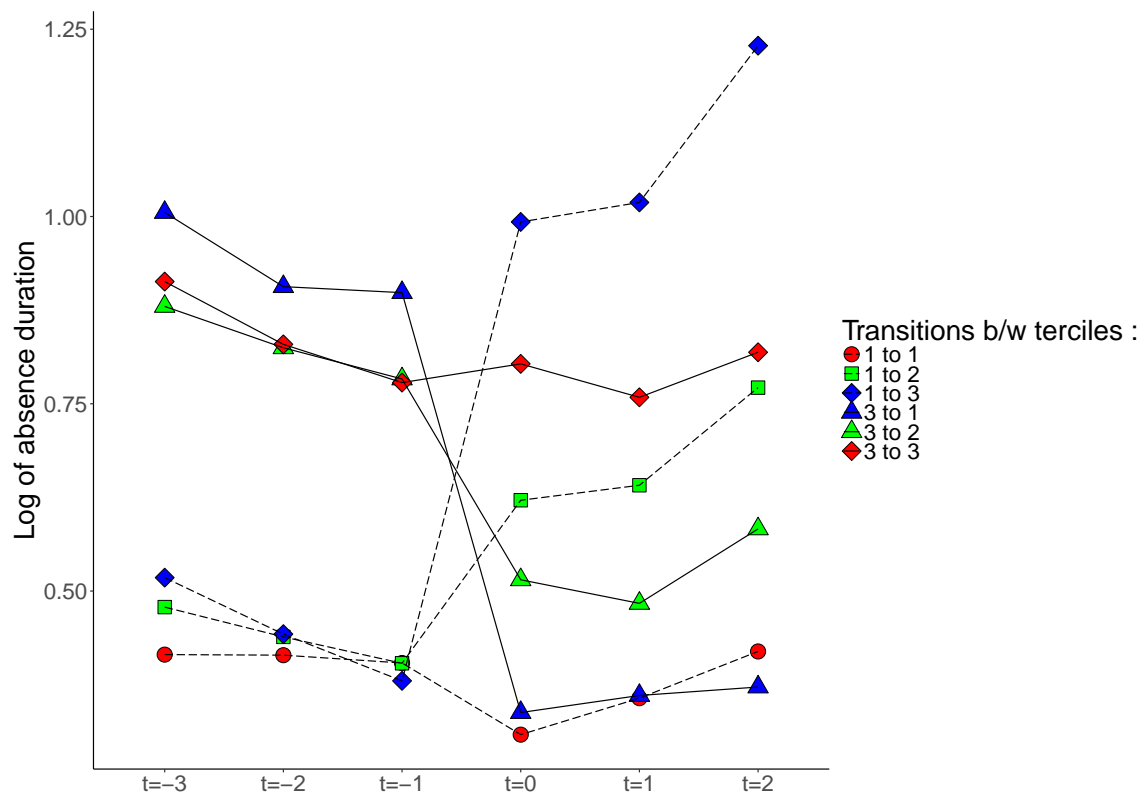


Figure 4.3 – Absence duration (in logarithm) for teachers transiting between terciles of school fixed effects in year $t=0$

Note: Schools are divided into terciles of school fixed effects based on the estimates of model (3). Schools in the third tercile are expected to increase teacher absences the most significantly. I consider all teachers observed for six consecutive years ($t=-3, t=-2, t=-1, t=0, t=1, t=2$), who change schools in year $t=0$, but remain in the same tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. This Figure shows how the logarithm of absence duration varies for teachers leaving in $t=0$ a school from the first tercile (low fixed effects on absences: dotted lines) or from the last tercile (high fixed effects on absences: solid lines). For each of the nine transition groups between terciles of school fixed effects, Table 4.26 shows the number of observations (teachers) and the logarithm of the duration of teachers' absences between $t=-3$ and $t=2$.

head teachers mobility and head teachers fixed effects

It should be recalled that head teachers may apply for a new assignment after having spent three years in the same school. After six years in the same school, mobility is compulsory. Principals rank schools according to their preferences. Then the administration assigns principals on a case-by-case basis, taking into account their wishes, profile and the characteristics of the schools requested. There is therefore both self-selection of head teachers and selection by the administration at the time of assignment.

To examine whether school principals sorting contributes to distorting the estimation

of headteacher fixed effects, we consider an event study similar to that proposed in the previous section (Card et al., 2013). We begin by dividing school principals into three terciles according to their fixed effects estimated in a first step. Principals of the first tercile (respectively of the last tercile) are supposed to reduce (respectively increase) the absences of teachers in the schools they head compared to the average of principals (second tercile). Then, we consider schools that change principal in the year $t=0$, and that keep a principal of the same tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. We divided the schools into nine transition groups according to principal's tercile in $t=-1$ and $t=0$. We compute the logarithm of teachers' absence in the schools of the nine cohorts each year. The results are reported in Table 4.28. Figure 4.4 shows the results for schools "leaving" the first tercile (solid lines) and for schools "joining" the third tercile of headteacher fixed effect in $t=0$ (dashed lines). This time, Figure 4.4 shows that the absence of teachers in schools is heterogeneous within terciles of principals, and this heterogeneity predicts in part the type of change that occurs at date $t=0$. Specifically, the higher the absences in schools before the date $t=0$, the greater the chance that a principal with a low fixed effect will arrive in $t=0$. This means that the administration has information on the managerial skills of school principals and allocates them according to the situation of the schools. Such a level of selection was not observed for the teacher assignment process, which seemed to leave much more room for chance.

Figure 4.4 also shows a strong variation in the average absence of teachers in schools moving from one headteacher' tercile to another in $t=0$. The average duration of absences is divided by 2 (-0.65 logarithmic points) in schools that leave the third tercile to reach the first tercile of principal FE, and increases by 50% (+0.4 logarithmic points) for the schools that "move" in the opposite direction (from the first to the third tercile of principal FE). Similarly, absences increase by 16% (+0.15 logarithmic points) when schools move from the first to the second tercile and decrease by 22% (-0.25 logarithmic points) when schools move from the third to the second tercile of headteacher FE. Average absences remain relatively stable in schools that remain in the same tercile of principal fixed effects over the period considered (first and third terciles on Figure 4.4). The absence gains and losses in schools are confirmed for all transitions in Table 4.28.

Once again, we examine three main sources of endogeneity related to the mobility of head teachers between schools: a "shock effect", a "drift effect" and a "match effect".

A first risk of endogeneity arises if head teachers are more likely to leave a school that receives a positive shock on absences in $t=-1$ ("Shock effect" hypothesis). Indeed, it is possible that when an incident occurs in $t=-1$ and results in an increase in teacher absences, the teaching community, as well as students and their parents, may exert pressure on the administration to have the principal replaced in $t=0$. Table 4.28 and Figure 4.4 show that there is no systematic increase in absences in $t=-1$ compared to years $t=-2$ and $t=-3$ when a new headteacher arrives in $t=0$, which tends to reject the "shock effect" hypothesis.

A second risk of endogeneity arises if trends in schools in terms of absences predict the fixed effect of the next principal ("drift effect" hypothesis). For example, if schools where absences increase (respectively decrease) over time are more likely to have a principal with a low (high) fixed effect, we will tend to underestimate principals' contribution to absences. This could happen if the administration, for example, takes into account previous absences in the school to assign principals. It can also be imagined that it takes time for the teaching community, students and their parents to get the administration to replace a principal who is a problem, and that absences may increase during this period. Figure 4.4 shows no trend in absences before $t=0$, regardless of the transition that occurs in $t=0$, suggesting that the "drift effect" is not a serious threat to our estimates.

A third risk of endogeneity arises if the administration assigns head teachers to schools in order to create a better match ("match effect" hypothesis). As mentioned in the discussion about teacher mobility, the existence of a match component should result in an asymmetry between the absence gains and losses associated with the transition of schools between terciles of principal fixed effects. Precisely, Figure 4.4 shows a slight asymmetry in absolute terms between the schools moving up and down in the distribution of principal fixed effects. This result is consistent with the assumption that a better match between schools and school leaders systematically reduces absences. A second fact supporting the existence of a match effect is that absenteeism decreases in schools that change principals while remaining in the same tercile in $t=0$. Results from Table 4.28 and Figure 4.4 suggest that the matching effect reduces absences by an average of 0.06 logarithmic points. Although this effect is not negligible, it remains small compared to the variations of absences when schools move between terciles of head teachers.

We are relatively confident that our estimates of the fixed effects of principals do not

suffer too much from the endogenous mobility of head teachers between schools.

Finally, we examine whether the mobility of principals tends to distort our estimates of teacher turnover. We replace the logarithm of absences with school turnover and report the results in Table 4.30. The analysis of Table 4.30 and Figure 4.5 shows a large variation in turnover in schools that "transit" between terciles of principals. The results suggest that we can reasonably rule out the hypothesis of a bias associated with headteacher mobility.

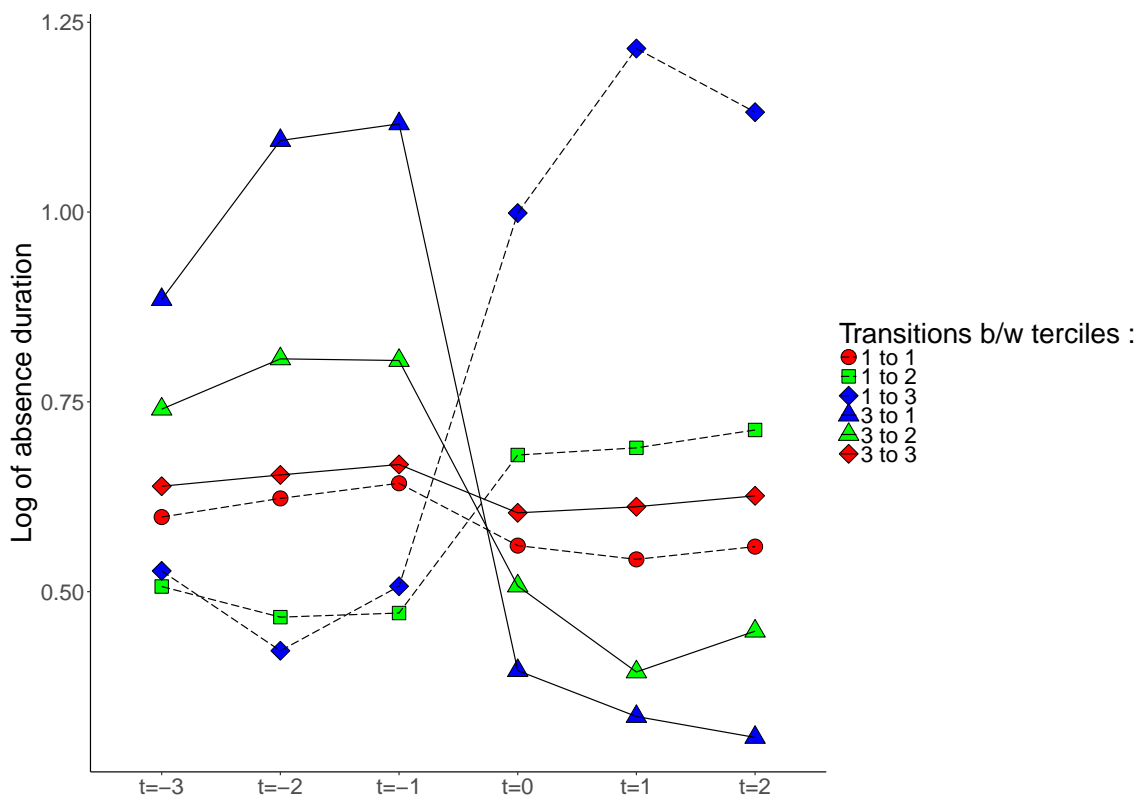


Figure 4.4 – Absence duration (in logarithm) for schools transiting between terciles of school principal fixed effects in year $t=0$

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (3). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years ($t=-3, t=-2, t=-1, t=0, t=1, t=2$), who are run by a new principal in year $t=0$, but remain in the same tercile of school principal between $t=-3$ and $t=-1$, and between $t=0$ and $t=2$. This Figure shows how the logarithm of the mean absence duration varies for schools "leaving" in $t=0$ a principal from the first tercile (low fixed effects on absences: dotted lines) or from the last tercile (high fixed effects on absences: solid lines). For each of the nine transition groups between terciles of principal fixed effects, Table 4.28 shows the number of observations (schools) and the logarithm of the mean duration of absences between $t=-3$ and $t=2$.

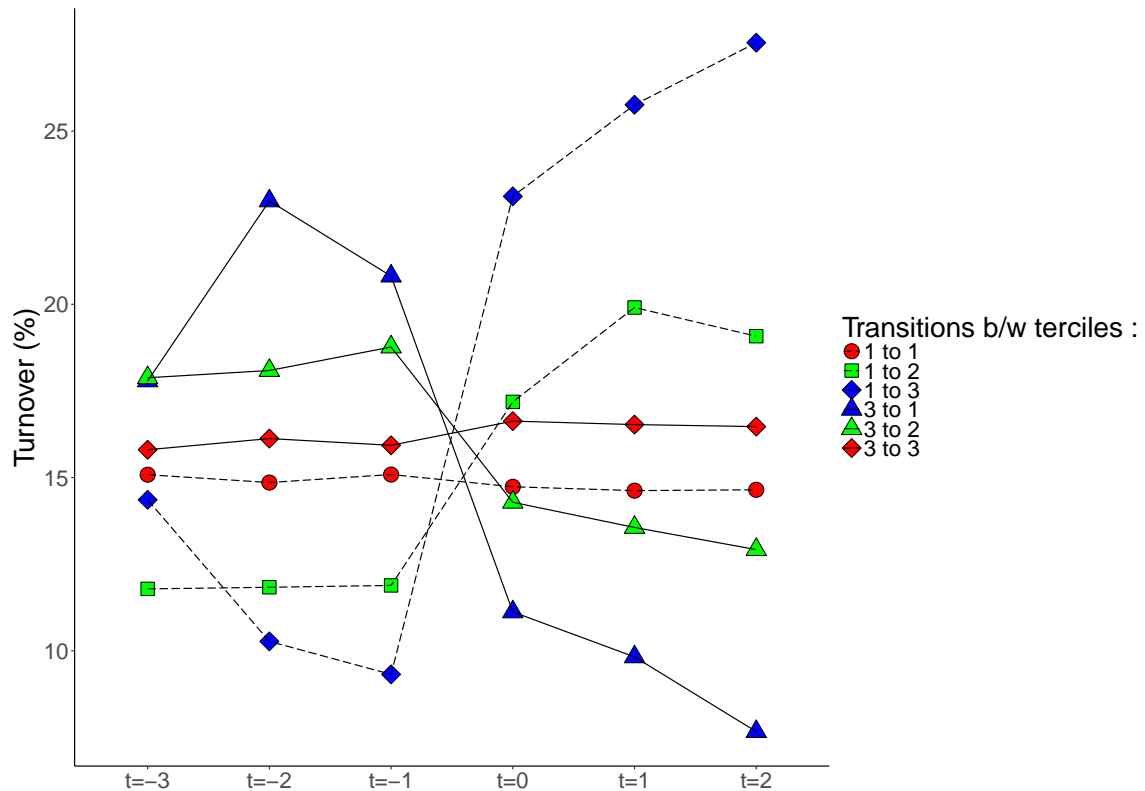


Figure 4.5 – Teacher turnover in schools transiting between terciles of school principal fixed effects in year $t=0$

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (3). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who are run by a new principal in year $t=0$, but remain in the same tercile of school principal between $t=-3$ and $t=-1$, and between $t=0$ and $t=2$. This Figure shows how teacher turnover varies for schools "leaving" in $t=0$ a principal from the first tercile (low fixed effects on absences: dotted lines) or from the last tercile (high fixed effects on absences: solid lines). For each of the nine transition groups between terciles of principal fixed effects, Table 4.30 shows the number of observations (schools) and the logarithm of the mean duration of absences between $t=-3$ and $t=2$.

4.7.2 Is the most appropriate model additive or multiplicative?

In this section, we compare specifications (1) and (4) and examine which one is most appropriate for estimating the effects of the work environment and personnel management on the average duration of absences.

Model (4) is additive and separable, meaning that a change in observed or unobserved individual characteristics, a change in school or direction has the same effect on absences for all individuals regardless of context. In other words, there is no possible interaction

between the explanatory variables in this model. To illustrate this, suppose that School A does not affect the average absence duration of teachers, but that School B increases the average absence duration by a factor $\phi > 0$. Model (4) assumes that absences increase by an average of ϕ days for all people who move from school A to school B, whether they are already very absent (a high value of α_i for example) or not very absent (a low value of α_i for example). Another implication of model (4) is the "reversibility" of effects. This presupposes, for example, that the transition from school A to school B or from school B to school A has the same effect on absences in absolute terms but with the opposite sign. Finally, this model assumes that the explanatory variables have an instant effect and no other long-term effects on absences (e. g., the model does not take into account effects that increase over time or occur several years after exposure to a given environment). This is clearly a limitation of this model, which implies that it probably underestimates the effects of the working context on absences.

Model (1) is multiplicative, which presupposes, contrary to model (4), that the effects of the explanatory variables are not identical regardless of the individual or context, but increase precisely with the level of absence of an individual in a given context. This model therefore introduces the possibility that all explanatory variables interact, which may exacerbate or mitigate their effects. For example, this model assumes that the absence duration of a person in poor health (high value of α_i), or who is exposed to a difficult work environment (high value of ϕ_j), increases much more with the arrival of a poor manager (high value of ψ_m) than the absence duration of a healthy person working in a favorable environment. Model (1) can be linearized by considering the logarithm of the expected absence duration: $\log(E[Y_{ijmt}]) = \alpha_i + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t$. Therefore, the main assumption of this model is that the mean absence duration is an additive and separable function of the observed and unobserved variables after log linearization. In other words, the multiplicative effect of each variable does not depend on the context or characteristics of individuals. Again, let's assume a school A that does not affect the average length of teacher absences, and a school B that increases the average length of absences by a multiplier factor ϕ . The model (1) assumes that the average absence duration of individuals is multiplied by a factor $\exp(\phi)$ when they move from School A to School B, regardless of the characteristics of these individuals. As a result, the transition from School A to School B further widens the gap between those who were

already very absent from School A and those who were not. The model (1) implies that the variables have a reversible effect on the logarithm of absences (transition from school A to school B and then from school B to school A has a neutral effect on absences) as well as on absences in value ($\exp(\phi_A - \phi_B) * \exp(\phi_B - \phi_A) = 1$). Finally, this model, like the model (4), assumes that explanatory variables have an immediate and short-term effect on absences.

To test the hypotheses of these two models, we examine again Figures 4.3 and 4.4 obtained by dividing schools and principals into terciles of fixed effects estimated from the multiplicative model (1). We compare these Figures with those obtained by dividing schools and principals into terciles of fixed effects estimated from the additive model (4): Figures 4.13 and 4.14. Tables 4.27 and 4.29 present the mean absences of teacher and school cohorts for all transitions between terciles of fixed effects estimated from the additive model (4).

As pointed out in the previous sections, a remarkable aspect of Figures 4.3 and 4.4 is the quasi-symmetry of increases and decreases in absences (logarithmic) depending on whether transitions are made from low terciles to high terciles or from high terciles to low terciles. This symmetry is observed for all transitions (see Tables 4.26 and 4.28) and satisfies the hypothesis of "reversible effects" implicit in the model (1). In addition, transitions between different fixed-effect terciles result in an immediate change in absences, which validates the presence of short-term effects. The presence of long-term effects is more difficult to assess here because we do not track individuals over a long period of time. However, when annual trends are taken into account, Tables 4.26 and 4.28 do not reveal any particular changes after three years spent in the same school or headteacher tercile. We note that logarithmic variations in absences between terciles do not depend on absences in the starting tercile. This result suggests that the fixed effects of the individuals, schools and head teachers are likely to be separable (independent of individual characteristics and context). Specifically, Tables 4.26 and 4.28 show that absences (in logarithmic form) increase as much between terciles 1 and 2 and between terciles 2 and 3, although those leaving the second tercile are more absent on average than those leaving the first tercile. Similarly, absences (in logarithmic form) decrease as much between terciles 3 and 2 and terciles 2 and 1, whereas those who leave the second tercile are less absent on average than those who leave the third tercile. Taken together, these evidence

suggest that we can reasonably estimate the effects of schools and principals from the model (1).

Our conclusions on the additive model (4) are much more mixed. First, Figures 4.13 and 4.14 show that absences (values) vary asymmetrically when individuals and schools move from one tercile of fixed effects to another. A more detailed analysis suggests that this asymmetry stems from the heterogeneity of effects according to the initial level of absence, which contradicts the separability assumption of model (4). Indeed, Table 4.27 shows that absences decrease more for individuals who move from tercile 3 to tercile 2, than for those who transition from tercile 2 to tercile 1. Precisely, those who depart from the third tercile are already more absent on average than those who leave the second tercile. It is interesting to note that absences increase as much for people who move from tercile 1 to tercile 2 or from tercile 2 to tercile 3. It can be seen that individuals who leave the first and second terciles are as much absent on average. The analysis of Table 4.29 leads to the same observation: the effect of the arrival of a headteacher belonging to a higher tercile is all the stronger as absences were initially high (absences increase more in schools that "transition" from tercile 2 to 3 than in schools that "transition" from tercile 1 to 2. Precisely absences were higher in schools leaving tercile 2 than in schools leaving tercile 1). Once again, we observe that absences vary in the same way between terciles when schools have initially the same level of absence (absences decrease as much in schools that go from tercile 3 to tercile 2 or from tercile 2 to tercile 1. Precisely, the levels of absences were the same in schools leaving terciles 2 and 3). A quick calculation shows that moving to a higher (or lower) tercile always multiplies absences by a constant coefficient. Ultimately, it appears that the separability assumption is not respected in the model (4). A multiplicative model, such as the model (1), seems to be better suited to studying the effects of the work environment on absences for health reasons.

4.7.3 What is the impact of log-linearisation on estimates?

In this section, we empirically test the impact (on the coefficients of interest) of estimating models (2) and (3) by OLS instead of model (1) by PPML. It should be recalled that the estimation of model (2) is obtained after deletion of the zero values and logarithmic linearisation of the model (1). The estimation of model (3) is obtained by adding a constant to all values and after logarithmic linearization of the model (1).

A condition that would mitigate estimation bias when considering models (2) or (3) instead of model (1), is that the conditional variance of absence duration Y_{ijmt} is proportional to the square of its conditional mean (Silva and Tenreyro, 2006). To examine this relationship in our data, we consider the following regression: $\ln(\text{Var}(Y_{ijmt})) = \alpha + \beta \ln(E(Y_{ijmt})) + \epsilon_{ijmt}$

Table 4.31 presents estimates of the β coefficient when we consider the mean and the variance of Y_{ijmt} conditional on 1) individuals only, 2) schools only and 3) schools, principals and years. In all three cases, the β coefficient is relatively close to 2, suggesting that the conditional variance of the absence duration Y_{ijmt} is approximately proportional to the square of its conditional mean.

Table 4.32 compares the coefficients of interest when they are estimated by PPML (column 1), by OLS after removing zero values and log-linearization (column 2) and by OLS after adding a constant C to all observations and log-linearization (column 3-7).

For most variables, the sign of the coefficients is robust to the model under consideration. However, the magnitude of the estimated effects can vary considerably depending on the model. In particular, as the value of the constant C decreases (column 3 to 7), the estimated coefficients increase in absolute value, and the adjusted R-squared decreases. When the constant is equal to 0.1, the estimates obtained by OLS (column 5) are fairly close to those obtained by PPML (column 1) for most of the variables considered.

4.7.4 Are absences partly explained by peer effects?

Among the workplace factors that can influence individual behaviour (effort, productivity), the effects of peers (colleagues) have received increasing attention in recent years. In this section, we examine whether the school effects that we estimated in this study partly reflect peer effects.

Studying the effects of peers on absenteeism has been one way of demonstrating the existence of employees' shirking behaviour. Ichino and Maggi (2000) finds evidence of positive peer effects on the absenteeism of employees of a large Italian bank, while Bradley et al. (2007) finds the same result for teachers in Australian schools. The authors suggest that the influence of peers on individual absence depends, among other things, on the social pressure exerted by co-workers (colleague supervision, absence norms and acceptable behaviours) and the degree of hierarchical monitoring. In general, the estimation of

peer effects poses many problems when the data are derived from natural experiments (Brodaty and Gurgand, 2016). First, the non-random assignment of individuals to peer groups and work contexts fosters spurious correlation between the absenteeism of peers and individuals. Self-selection of individuals (sorting) and contextual effects, when not properly accounted for (omitted variables), can lead to biased estimates of peer effects. Another important source of bias arises when common shocks, received by both individuals and their peers, are wrongly attributed to peer effects. This is called "correlated effects". A final problem is the simultaneity, or "reflection problem", that Manski (1993) first described for peer effects. Indeed, while the average absence behaviour of peers influences an individual's absence, the absence of the individual also influences the absence of peers, and so on. There is therefore potentially a phenomenon of mutual self-influence that inflates the direct effect of peers.

Ichino and Maggi (2000) and Bradley et al. (2007) do not fully address these potential threats to identification. Regarding contextual effects, Bradley et al. (2007) argue that working conditions are homogeneous across schools because Australian schools are comparable in terms of pay, resources and policies. As a result, they include time varying observable, like the size of the school and the local unemployment rate, education district dummy variables, but they omit school fixed effects in their estimation of peer effects (the real reason is that they have relatively few observations per school because they focus on individuals who change schools - "movers" - between the 2001-2002 school years). Estimation of peer effects is highly dependent on the validity of this assumption, as otherwise unobserved characteristics in schools may explain the relationship between co-worker absences and individual absences. The same is true for Ichino and Maggi (2000) that includes province fixed effects instead of branch fixed effects in their model (they lack observations by branch because they estimate a model on "movers" over the 1993-1995 period). Another risk of endogeneity comes from the "reflection problem" in the model considered by the authors (we provide a simplified version with our notations): $Y_{idt} = \alpha_i + \beta \overline{Y_{-it}} + \gamma X_{it} + \delta_t + \mu_d + \epsilon_{idt}$, with Y_{idt} the absence indicator of individual i at time t in province or education district d , α_i an individual fixed effects, $\overline{Y_{-it}}$ the average absence in the peer group of individual i at time t (omitting individual i), δ_t a time effect, μ_d a province or education district fixed effects and ϵ_{idt} an error term with mean zero. Since Y_{idt} appears both on the left and right hand sides of the equation system (because i

is also one of his/her colleagues' peers), the error terms ϵ_{idt} are correlated, which violates one of the conditions of OLS consistency. The authors propose two ways to address this potential issue: they replace $\overline{Y_{-it}}$ by its lagged value $\overline{Y_{-it-1}}$ and they instrument $\overline{Y_{-it}}$ by $\overline{Y_{-it-1}}$. But this does not completely solve the "reflection problem". A final point is that [Ichino and Maggi \(2000\)](#), like [Bradley et al. \(2007\)](#), use an additive model rather than a multiplicative one, which can introduce an additional specification bias on the coefficients of interest.

To circumvent these estimation problems, we adapt the model proposed by [Arcidiacono et al. \(2012\)](#) and consider the following model:

$$\ln(Y_{ijmt} + C) = \alpha_i + \lambda \overline{\alpha_{-i}} + \phi_j + \psi_m + \beta X_{it} + \gamma Z_{jt} + \delta_t + \nu_{ijmt} \quad (5)$$

with $\overline{\alpha_{-i}}$ the average of the fixed contributions of individual i 's peers, λ the parameter that measures peer effects on absence duration and ν_{ijmt} an error term that is assumed to be mean zero and independent of the explanatory variables.

In this model, the social multiplier λ operates only through the fixed effects of peers, which is much more restrictive than in [Ichino and Maggi \(2000\)](#). In particular, the model assumes that the time-varying characteristics of peers do not affect individual absence³⁴. A clear advantage, however, is that the simultaneity problem is eliminated since Y_{ijmt} appears only to the left of equation (5). [Arcidiacono et al. \(2012\)](#) argue that this model is particularly well-suited when the explained variable is a choice (choice of whether or not to be absent at date t) and when individuals have only an approximate knowledge (expectations) of their peers' choices (individuals do not observe the annual length of absence of their peers at date t , they only expect a duration over the year). In more complex situations, $\overline{\alpha_{-i}}$ can be seen as a proxy for peer behaviour, which helps to get rid of the reflection problem. To limit the effect of sorting, we include time-varying observable characteristics and individual fixed effects in model (5). To reduce spurious correlation due to contextual effects, we add school characteristics that vary over time, school fixed effects and headteacher fixed effects. Our approach is therefore similar to that of [Cornelissen et al. \(2017\)](#) who estimate the effect of peers on wages by considering two high-dimensional fixed effects that neutralize observed and unobserved characteristics of individuals and firms. Like [Cornelissen et al. \(2017\)](#), our estimation strategy exploits

³⁴We can mitigate this assumption by allowing for peer effects via observable time-varying characteristics

the variation in peer characteristics that is due to school change for "movers" and that is due to the arrival and departure of colleagues for "stayers".

I implement the iterative algorithm proposed by [Arcidiacono et al. \(2012\)](#) to estimate model (5) with three high-dimensional fixed effects. The authors show that it converges to the nonlinear least squares solution. We present the result in Table 4.13 and 4.14. The first line of Table 4.13 report the estimated peer effect: on average, a one standard deviation increase in peer's fixed component increase individual absence by 0.016 absence days in the additive model (column 1) and by 10% in multiplicative models (column 2 and 3). The difference between the two types of models is substantial with respect to peer effect estimates. One possible explanation is that peer effects are non-linear and decrease with the fixed effect of the individual (remember that log-linearized models give more weight to 0 days of absence). To test this hypothesis, I estimate model (5) by allowing peer effects to vary with the value of the individual fixed effect. The results seem to confirm this hypothesis: the effect of peers is much more pronounced among teachers with a lower than average individual fixed effect, regardless of the model considered³⁵. However, we note that the estimated coefficients for time-varying characteristics are very close to those obtained with the main specification (see Table 4.4. The decomposition of the variance also leads to similar results (see Table 4.5. similarly, the adjusted R-squared remains virtually unchanged after taking peer effects into account. Table 4.15 presents the correlation between the school (principal) fixed effects estimated from the main specification and from the specification that takes into account peer effects. It shows that individual, school and principal fixed effects are strongly correlated between specifications ($\rho > 0.8$ regardless of the model). Specifically, the individual fixed effects are almost the same ($\rho > 0.99$), but school and principal fixed effects appear to be slightly biased when the effects of peers are not taken into account ($\rho > 0.82$ and $\rho > 0.8$ respectively). This result suggests that our primary specification is not seriously biased by the omission of peer interactions.

Finally, Tables 4.35 and 4.36 show peer effects estimates using the method proposed by [Ichino and Maggi \(2000\)](#) instead of the method proposed by [Arcidiacono et al. \(2012\)](#). Peer effects are twice as strong with the [Ichino and Maggi \(2000\)](#) method, suggesting

³⁵I divide the individual fixed effects into two groups according to whether they are above or below the median. I estimate model (5) with the possibility that peer effects may differ from one group to another. I do not present the results in this document. See [Arcidiacono et al. \(2012\)](#) for more details on this method.

that simultaneity bias ("reflection problem") plays an important role in estimating the effects of peers on absences.

Table 4.13 – Contribution of time-varying covariates, fixed effects and peer effects to the duration of absence of secondary school teachers (Arcidiacono et al. 2012) (Model D)

	(1)	(2)	(3)
	Absence duration Y	duration after transformation $\ln(Y + 1)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model D1	Model D2	Model D3
Peer effect	0.0157*** (0.00227)	0.0826*** (0.00236)	0.0965*** (0.00237)
Employment contract (ref : civil servant status)			
Trial period before civil servant status	-3.653*** (0.221)	-0.198*** (0.00738)	-0.240*** (0.00889)
Short-term contract	-5.442*** (0.126)	-0.451*** (0.00422)	-0.562*** (0.00509)
Permanent contract	-4.579*** (0.273)	-0.305*** (0.00912)	-0.378*** (0.0110)
Age	-2.447*** (0.0206)	-0.0542*** (0.000687)	-0.0644*** (0.000828)
Age squared	0.0377*** (0.000226)	0.000856*** (0.00000755)	0.000994*** (0.00000910)
Seniority	0.162*** (0.00361)	0.00653*** (0.000120)	0.00755*** (0.000145)
Maternity leave	22.13*** (0.131)	1.362*** (0.00437)	1.542*** (0.00527)
Adoption leave	-1.120*** (0.168)	-0.0621*** (0.00559)	-0.0771*** (0.00674)
Number of children	-1.617*** (0.0192)	-0.0916*** (0.000642)	-0.0998*** (0.000774)
Number of full-time employees	-0.000724 (0.000803)	-0.00147*** (0.0000270)	-0.00183*** (0.0000326)
Number of managers per employees	-12.73*** (1.675)	-0.366*** (0.0559)	-0.379*** (0.0673)
Duration of sick leave for managers	0.00126 (0.000922)	0.0000371 (0.0000307)	0.0000340 (0.0000371)
Year effects	Yes	Yes	Yes
Establishment fixed effects	Yes	Yes	Yes
Principal fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Observations	1721420	1721420	1721420
R-squared	0.527	0.525	0.523
Adjusted R-squared	0.428	0.426	0.424

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. Estimates in columns (1) to (3) are obtained using the algorithm proposed by Arcidiacono et al. (2012).

Source: Ministry of Education (DEPP).

Table 4.14 – Variance decomposition of the absence duration of secondary school teachers taking into account peers effects (Arcidiacono et al. 2012) (Model D)

	(1)	(2)	(3)
	Absence duration	Log of absence duration after transformation	Log of absence duration after transformation
	Y	$\ln(Y + 1)$	$\ln(Y + 0.1)$
	Model D1	Model D2	Model D3
Observations	1721420	1721420	1721420
R-squared	0.527	0.525	0.523
Adjusted R-squared	0.428	0.426	0.424
Number of individuals	283933	283933	283933
Number of establishments	4782	4782	4782
Number of principals	8120	8120	8120
Var(Y)	1611.22	1.79	2.58
Var(Individual FE)	899.14	0.91	1.30
Var(Establishment FE)	126.76	0.12	0.18
Var(Principal FE)	66.28	0.06	0.09
Var(β x Covariates)	104.94	0.12	0.15
Cov(Individual FE; Establishment FE)	-59.15	-0.05	-0.07
Cov(Individual FE; Principal FE)	-3.53	-0.00	-0.00
Cov(Individual FE; β x Covariates)	-50.69	-0.02	-0.03
Cov(Principal FE; Establishment FE)	-55.96	-0.05	-0.08
Cov(Establishment FE; β x Covariates)	-4.31	-0.01	-0.01
Cov(Principal FE; β x Covariates)	-0.49	-0.00	-0.00
Var(residual)	762.18	0.85	1.23

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. Estimates in columns (1) to (3) are obtained using the algorithm proposed by Arcidiacono et al. (2012).
Source: Ministry of Education (DEPP).

Table 4.15 – Correlation between the fixed effects estimated by the main specifications (models 3 and 4) and the specification that takes into account peer effects (model 5)

	(1)	(2)
Model used :	Y	$\ln(Y + 0.1)$
Correlation between individual FE	0.9998***	0.9901***
Correlation between school FE	0.9146***	0.8223***
Correlation between principal FE	0.8884***	0.8060***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This table shows the correlation between the fixed effects of teachers, schools and principals estimated from model (5), which takes into account the effects of peers, and estimated from the models (3) and (4). The variable explained is the total duration of teachers' sick leave (in days) in column (1) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (2).
Source: Ministry of Education (DEPP).

4.8 Conclusion

This article examines whether the heterogeneity of employee health-related absenteeism can be explained in part by the heterogeneity of workplaces and managers. To this end, I examine the effect of schools and principals on the annual absences of secondary school

teachers. The case of secondary school teachers is interesting in several ways. First, we know that they are particularly exposed to work stress. Second, a number of recent articles have shown that principals have a significant impact on student achievement without clear reasons (Clark et al., 2009; Branch et al., 2012; Coelli and Green, 2012; Bloom et al., 2014). A plausible hypothesis is that some of the effect is attributable to teacher absences, with several studies showing that teacher absences have a significant negative effect on student achievement (Duflo et al., 2012; Herrmann and Rockoff, 2012).

This research is based on management data from the French Ministry of Education that namely collects the reason and the length of teachers absences during the 2007-2015 period. This data also identifies the principal of the school where teachers are assigned. Our identification strategy is based on the mobility of teachers and principals. The results show that schools and principals contribute significantly to absenteeism. On average, teacher absence duration increases by 250% when teachers move from the first to the fourth quartile of school fixed effects and by 170% when they move from the first to the fourth quartile of principal fixed effects. I perform several robustness tests to rule out the hypothesis that the endogenous mobility of teachers and principals explains these results. Several elements suggest that the effect of schools and principals reflects working conditions and management practices.

These findings implies that measures to reduce psychosocial risks in schools and to provide better training for school leaders would reduce teacher absenteeism and could potentially contribute to student performance. We concur with the findings of Fryer et al. (2017) that show that providing management training to principals significantly improves student achievement. The method proposed in this study could help to target the schools and principals who need it most.

Bibliography

- Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. *Longitudinal Employer-Household Dynamics Technical Papers, Center for Economic Studies, U.S. Census Bureau.*, 2002-06.
- Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333.
- Afsa, C. and Givord, P. (2006). Le rôle des conditions de travail dans les absences pour maladie. *Document de travail de la Direction des Études et Synthèses Économiques, INSEE.*
- Allen, S. G. (1981). An Empirical Model of Work Attendance. *The Review of Economics and Statistics*, 63(1):77–87.
- Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3):673–697.
- Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824–827.
- Arcidiacono, P., Foster, G., Goodpaster, N., and Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics*, 3(3):421–470.
- Bamberger, P. and Biron, M. (2007). Group norms and excessive absenteeism: The role of peer referent others. *Organizational Behavior and Human Decision Processes*, 103(2):179–196.
- Béjean, S. and Sultan-Taïeb, H. (2005). Modeling the economic burden of diseases imputable to stress at work. *European Journal of Health Economics*, 6(1):16–23.
- Belkic, K. L., Landsbergis, P. A., Schnall, P. L., and Baker, D. (2004). Is job strain a major source of cardiovascular disease risk? *Scandinavian journal of work, environment & health*, 30(2):85–128.
- Benavides, F. G., Benach, J., and Roman, C. (2000). How Do Types of Employment Relate To Health Indicators Finding From the Second European Survey on Working Conditions. *Journal of Epidemiology and Community Health*, 54:494–501.
- Bertrand, M. and Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, 118(4):1169–1208.
- Birket-smith, M., Hansen, B. H., Hanash, J. A., Hansen, J. F., and Rasmussen, A. (2009). Mental disorders and general well-being in cardiology outpatients - 6-year survival. *Journal of Psychosomatic Research*, 67(1):5–10.
- Black, S. E. and Lynch, L. M. (2001). How to Compete: The Impact of Workplace Practices and Information Technology on Productivity. *Review of Economics and Statistics*, 83(3):434–445.
- Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2014). Does Management Matter in Schools. *IZA Discussion Paper Series*, (8620).
- Bloom, N. and Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.
- Bolduc, D., Fortin, B., Labrecque, F., and Lanoie, P. (2002). Workers' Compensation, Moral Hazard and the Composition of Workplace Injuries. *The Journal of Human Resources*, 37(3):623.

- Bongers, P. M., De Winter, C. R., Kompier, M. A. J., and Hildebrandt, V. H. (1993). Psychosocial factors at work and musculoskeletal disease. *Scandinavian Journal of Work, Environment and Health*, 19(5):297–312.
- Bradley, S., Green, C., and Leeves, G. (2007). Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics*, 14(3):319–334.
- Branch, G. F., Hanushek, E. a., and Rivkin, S. G. (2012). Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals. *CALDER working paper*, (17803).
- Breda, T. and Hillion, M. (2016). Supplementary Materials for Teaching accreditation exams reveal grading biases favor women. *Science*, 353(6298):1–53.
- Brodaty, T. and Gurgand, M. (2016). Good peers or good teachers? Evidence from a French University. *Economics of Education Review*, 54:62–78.
- Brown, S. and Sessions, J. G. (1996). The Economics of Absence: Theory and Evidence. *Journal of Economic Surveys*, 10(1):23–53.
- Bryson, A. and Green, F. (2018). Do Private Schools Manage Better? *National Institute Economic Review*, (243):R17–R26.
- Cameron, A. C. and Triverdi, P. k. (2005). *Microeconometrics Methods and Applications*. Number 9.
- Card, D., Cardoso, A. R., and Kline, P. (2015). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, 131(2):633–686.
- Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.
- Clark, D., Martorell, P., and Rockoff, J. (2009). School Principals and School Performance. *National Center for Analysis of Longitudinal Data In Education Research*, (38).
- Coelli, M. and Green, D. A. (2012). Leadership effects: school principals and student outcomes. *Economics of Education Review*, 31(1):92–109.
- Collins, J. J., Baase, C. M., Sharda, C. E., Ozminkowski, R. J., Nicholson, S., Billotti, G. M., Turpin, R. S., Olson, M., and Berger, M. L. (2005). The assessment of chronic health conditions on work performance, absence, and total economic impact for employers.
- Combe, J., Tercieux, O., and Terrier, C. (2016). The Design of Teacher Assignment : Theory and Evidence. *Working Paper*.
- Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer Effects in the Workplace. *American Economic Review*, 107(2):425–456.
- Correia, S. (2016). A Feasible Estimator for Linear Models with Multi-Way Fixed Effects. *Working Paper*.
- Cottini, E. and Lucifora, C. (2010). Mental Health and Working Conditions in European Countries. *IZA Discussion Papers*.
- D’Amuri, F. (2011). Monetary Incentives vs. Monitoring in Addressing Absenteeism: Experimental Evidence. *SSRN Electronic Journal*, 35(442):1–28.
- DARES (2013). Les absences au travail des salariés pour raisons de santé : un rôle important des conditions de travail. *Dares Analyses*, (9).
- Davis, K. G. and Heaney, C. a. (2000). The relationship between psychosocial work characteristics and low back pain: underlying methodological issues. *Clinical biomechanics*, 15:389–406.
- De Paola, M., Scoppa, V., and Pupo, V. (2014). Absenteeism in the Italian Public

- Sector: The Effects of Changes in Sick Leave Policy. *Journal of Labor Economics*, 32(2):337–360.
- DEPP (2011). Résultats de la première enquête nationale de victimation au sein des collèges publics au printemps 2011. *Note d'information - MEN-DEPP*.
- Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–1278.
- Engellandt, A. and Riphahn, R. T. (2005). Temporary contracts and employee effort. *Labour Economics*, 12(3):281–299.
- Farber, B. A. (1991). *The Jossey-Bass education series. Crisis in education: Stress and burnout in the American teacher*. San Francisco: Jossey-Bass., san franci edition.
- Farrell, D. and Stamm, C. L. (1988). Meta-Analysis of the Correlates of Employee Absence. *Human Relations*, 41(3):211–227.
- Frick, B. and Malo, M. (2008). Labor market institutions and individual absenteeism in the European Union: The relative importance of sickness benefit systems and employment protection legislation. *Industrial Relations*, 47(4):505–529.
- Fryer, R. G., Angrist, J., Card, D., and Charles, K. (2017). Management and Student Achievement: Evidence from a Randomized Field Experiment. *NBER Working Paper Series*, page 66.
- Garrouste, M. and Prost, C. (2016). Éducation prioritaire. Comment l'école amplifie les inégalités sociales et migratoires ? *Rapport CNESEO*.
- Gaure, S. (2013). OLS with multiple high dimensional category variables. *Computational Statistics and Data Analysis*, 66(0167):8–18.
- Gaziel, H. H. (2004). Predictors of absenteeism among primary school teachers. *Social Psychology of Education*, 7(4):421–434.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica*, 52(3):681.
- Griffin, J. M., Fuhrer, R., Stansfeld, S. A., and Marmot, M. (2002). The importance of low control at work and home on depression and anxiety: Do these effects vary by gender and social class? *Social Science and Medicine*, 54(5):783–798.
- Grignon, M. and Renaud, T. (2007). Sickness and Injury Leave in France : Moral Hazard or Strain ? Sickness and injury leave in France : moral hazard or strain ? *Document de travail de l'IRDES*.
- Guimarães, P. and Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10(4):628–649.
- Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of School Psychology*, 43(6):495–513.
- Harvey, J. and Nicholson, N. (1999). Minor illness as a legitimate reason for absence. *Journal of Organizational Behavior*, 20(6):979–993.
- Hauke, A., Flintrop, J., Brun, E., and Rugulies, R. (2011). The impact of work-related psychosocial stressors on the onset of musculoskeletal disorders in specific body regions: A review and metaanalysis of 54 longitudinal studies. *Work and Stress*, 25(3):243–256.
- Hausknecht, J. P., Hiller, N. J., and Vance, R. J. (2008). Work-Unit Absenteeism: Effects of Satisfaction, Commitment, Labor Market Conditions, and Time. *Academy of Management Journal*, 51(6):1223–1245.
- Henkel, V., Mergl, R., Kohnen, R., Maier, W., Möller, H.-j., Hegerl, U., Henkel, V., Mergl, R., Hegerl, U., Kohnen, R., and Maier, W. (2003). Identifying depression in primary care : a comparison of different methods in a prospective cohort study. *British Medical Journal*, 326.

- Henrekson, M. and Persson, M. (2004). The Effects on Sick Leave of Changes in the Sickness Insurance System. *Journal of Labor Economics*, 22(1):87–113.
- Herrmann, M. A. and Rockoff, J. E. (2012). Worker Absence and Productivity: Evidence from Teaching. *Journal of Labor Economics*, 30(4):749–782.
- Ichino, A. and Maggi, G. (2000). Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm. *Quarterly Journal of Economics*, 115(3):1057–1090.
- Ichino, A. and Riphahn, R. T. (2005). The Effect of Employment Protection on Worker Effort: Absenteeism during and after Probation. *Journal of the European Economic Association*, 3(1):120–143.
- Janssens, H., Clays, E., De Clercq, B., Casini, A., De Bacquer, D., Kittel, F., and Braeckman, L. (2014). The relation between psychosocial risk factors and cause-specific long-term sickness absence. *European Journal of Public Health*, 24(3):428–433.
- Jégo, S. and Guillo, C. (2016). Les enseignants face aux risques psychosociaux. *Éducation et formations*.
- Jenkins, R. (1985). Minor psychiatric morbidity in employed young men and women and its contribution to sickness absence. *British Journal of Industrial Medicine*, 42(1985):147–154.
- Johansson, P. and Palme, M. (2005). Moral hazard and sickness insurance. *Journal of Public Economics*, 89(9-10):1879–1890.
- Karasek, R. A. (1979). Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly*, 24(2):285.
- Kivimäki, M., Head, J., Ferrie, J. E., Shipley, M. J., Vahtera, J., and Marmot, M. G. (2003). Sickness absence as a global measure of health: evidence from mortality in the Whitehall II prospective cohort study. *BMJ*, 327.
- Krieger, T., Zimmermann, J., Huffziger, S., Uhl, B., Diener, C., Kuehner, C., and Grosse, M. (2014). Measuring depression with a well-being index : Further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. *Journal of Affective Disorders*, 156:240–244.
- Krueger, A. B. and Summers, L. H. (1988). Efficiency Wages and the Inter-Industry Wage Structure. *Econometrica*, 56(2):259.
- Kuper, H. and Marmot, M. (2003). Job strain, job demands, decision latitude, and risk of coronary heart disease within the Whitehall II study. *Journal of epidemiology and community health*, 57(2):147–53.
- Kuper, H., Singh-Manoux, A., Siegrist, J., and Marmot, M. (2002). When reciprocity fails: effort-reward imbalance in relation to coronary heart disease and health functioning within the Whitehall II study. *Occupational and environmental medicine*, 59(11):777–84.
- Madsen, I. E. H., Nyberg, S. T., Magnusson Hanson, L. L., Ferrie, J. E., Ahola, K., Alfredsson, L., Batty, G. D., Bjorner, J. B., Borritz, M., Burr, H., Chastang, J.-F., de Graaf, R., Dragano, N., Hamer, M., Jokela, M., Knutsson, A., Koskenvuo, M., Koskinen, A., Leineweber, C., Niedhammer, I., Nielsen, M. L., Nordin, M., Oksanen, T., Pejtersen, J. H., Pentti, J., Plaisier, I., Salo, P., Singh-Manoux, A., Suominen, S., ten Have, M., Theorell, T., Toppinen-Tanner, S., Vahtera, J., Väänänen, A., Westerholm, P. J. M., Westerlund, H., Fransson, E. I., Heikkilä, K., Virtanen, M., Rugulies, R., and Kivimäki, M. (2017). Job strain as a risk factor for clinical depression: systematic review and meta-analysis with additional individual participant data. *Psychological Medicine*, 47(08):1342–1356.

- Manning, W. G. and Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics*, 20(4):461–494.
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531.
- Marmot, M., Feeney, A., Shipley, M., North, F., and Syme, S. L. (1995). Sickness absence as a measure of health status and functioning: from the UK Whitehall II study. *Journal of Epidemiology and Community Health*, 49:124–130.
- Mathieu, J. E. and Zajac, D. M. (1990). A Review and Metaanalysis of the Antecedents, Correlates, and Consequences of Organizational Commitment. *Psychological Bulletin*, 108(2):171–194.
- Melchior, M., Krieger, N., Kawachi, I., Berkman, L. F., Niedhammer, I., and Goldberg, M. (2005). Work factors and occupational class disparities in sickness absence: Findings from the GAZEL cohort study. *American Journal of Public Health*, 95(7):1206–1212.
- Melchior, M., Niedhammer, I., Berkman, L. F., and Goldberg, M. (2003). Do psychosocial work factors and social relations exert independent effects on sickness absence? A six year prospective study of the GAZEL cohort. *J Epidemiol Community Health*, (57):285–293.
- Meyer, B. D., Viscusi, W. K., and Durbin, D. L. (1995). Workers’ compensation and injury duration: evidence from a natural experiment. *The American economic review*, 85(3):322–40.
- Niedhammer, I., Bugel, I., Goldberg, M., Leclerc, a., and Guéguen, a. (1998). Psychosocial factors at work and sickness absence in the Gazel cohort: a prospective study. *Occupational and environmental medicine*, 55(11):735–741.
- Niedhammer, I., Lesuffleur, T., Memmi, S., and Chastang, J.-F. (2017). Working conditions in the explanation of occupational inequalities in sickness absence in the French SUMER study. *European Journal of Public Health*, 27(6):1061–1068.
- Nourry, N., Luc, A., Lefebvre, F., Sultan-Taïeb, H., and Béjean, S. (2014). Psychosocial and organizational work environment of nurse managers and self-reported depressive symptoms: Cross-sectional analysis from a cohort of nurse managers. *International Journal of Occupational Medicine and Environmental Health*, 27(2):252–269.
- Olivier, M. a. J. and Venter, D. J. L. (2003). The extent and causes of stress in teachers in the George region. *South African Journal of Education*, 23(3):186–192.
- Ose, S. O. (2005). Working conditions, compensation and absenteeism. *Journal of Health Economics*, 24(1):161–188.
- Pas, E. T., Bradshaw, C. P., and Hershfeldt, P. A. (2012). Teacher- and school-level predictors of teacher efficacy and burnout: identifying potential areas for support. *Journal of school psychology*, 50(1):129–45.
- Pichler, S. and Ziebarth, N. (2015). The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Shirking Behavior. *Working Paper*.
- Pomaki, G. and Anagnostopoulou, T. (2003). A test and extension of the demand/control/social support model: Prediction of wellness/health outcomes in Greek teachers. *Psychology and Health*, 18(4):537–550.
- Prost, C. (2013). Teacher Mobility: Can Financial Incentives Help Disadvantaged Schools to Retain Their Teachers? *Annals of Economics and Statistics*, (111/112):171.
- Puhani, P. A. and Sonderhof, K. (2010). The effects of a sick pay reform on absence and on health-related outcomes. *Journal of Health Economics*, 29(2):285–302.
- Rieck, K. M. E. and Telle, K. (2013). Sick leave before, during and after pregnancy. *Acta Sociologica*, 56(2):117–137.

-
- Riphahn, R. T. and Thalmaier, A. (1999). Behavioral Effects of Probation Periods: An Analysis of Worker Absenteeism. *IZA Discussion Paper Series*, (67):179–201.
- Robone, S., Jones, A. M., and Rice, N. (2011). Contractual conditions, working conditions and their impact on health and well-being. *European Journal of Health Economics*, 12(5):429–444.
- Sagie, A. (1998). Employee absenteeism, organizational commitment, and job satisfaction: Another look. *Journal of Vocational Behavior*, 52(2):156–171.
- Sagie, A., Birati, A., and Tziner, A. (2002). Assessing the Costs of Behavioral and Psychological Withdrawal: A New Model and an Empirical Illustration. *Applied Psychology*, 51(1):67–89.
- Santin, G., Cohidon, C., Goldberg, M., and Imbernon, E. (2009). Atypical jobs and depressive symptoms in France, based on the 2003 Decennial Health Survey. *American Journal of Industrial Medicine*, 52:57–60.
- Santos Silva, J. M. and Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112(2):220–222.
- Schaufeli, W. B., Bakker, A. B., and van Rhenen, W. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior*, 30(7):893–917.
- Scoppa, V. (2010). Worker absenteeism and incentives: Evidence from Italy. *Managerial and Decision Economics*, 31(8):503–515.
- Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review*, 74(3):45–56.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of occupational health psychology*, 1(1):27–41.
- Silva, J. M. C. S. and Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, 88(4):641–658.
- Skaalvik, E. M. and Skaalvik, S. (2009). Does school context matter? Relations with teacher burnout and job satisfaction. *Teaching and Teacher Education*, 25(3):518–524.
- Steers, R. M. and Rhodes, S. R. (1978). Major influences on employee attendance: A process model. *Journal of Applied Psychology*, 63(4):391–401.
- Stefanou, A. (2013). L'éducation prioritaire - Etat des lieux. *Note d'information - MEN-DEPP*, (13.07).
- Vahtera, J., Kivimäki, M., Pentti, J., and Theorell, T. (2000). Effect of change in the psychosocial work environment on sickness absence: a seven year follow up of initially healthy employees. *Journal of epidemiology and community health*, 54(7):484–493.
- Vahtera, J., Pentti, J., and Kivimäki, M. (2004). Sickness absence as a predictor of mortality among male and female employees. *Journal of Epidemiology and Community Health*, 58(4):321–326.
- van der Hulst, M. (2003). Long workhours and health. *Scandinavian Journal of Work, Environment & Health*, 29(3):171–188.
- Wang, M. J., Mykletun, A., Møyner, E. I., Øverland, S., Henderson, M., Stansfeld, S., Hotopf, M., and Harvey, S. B. (2014). Job strain, health and sickness absence: Results from the Hordaland health study. *PLoS ONE*, 9(4).
- Ziebarth, N. R. (2013). Long-term absenteeism and moral hazard-Evidence from a natural experiment. *Labour Economics*, 24:277–292.
- Ziebarth, N. R. and Karlsson, M. (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics*, 94(11-12):1108–1122.

Tables

Table 4.16 – Contribution of time-varying covariates and fixed effects to the duration of absence of secondary school teachers (Model B)

	(1)	(2)	(3)	(4)
	Absence duration Y	Log of absence duration ln(Y)	Log of absence duration after transformation ln(Y + 0.1)	Absence duration Poisson(Y)
	Model B1	Model B2	Model B3	Model B4
Employment contract (ref : civil servants)				
Civil servants on probationary year	-4.080*** (0.178)	-0.208*** (0.0122)	-0.360*** (0.0123)	-0.624*** (0.0219)
Short-term contracts	-6.277*** (0.260)	-0.295*** (0.0208)	-0.878*** (0.0198)	-0.992*** (0.0360)
Open-ended contracts	-5.490*** (0.412)	-0.172*** (0.0298)	-0.584*** (0.0289)	-0.667*** (0.0506)
Age	-2.345*** (0.0503)	-0.0512*** (0.00291)	-0.0781*** (0.00285)	-0.0921*** (0.00511)
Age squared	0.0375*** (0.000576)	0.00110*** (0.0000303)	0.00126*** (0.0000301)	0.00229*** (0.0000516)
Seniority	0.168*** (0.0116)	0.00890*** (0.00104)	0.00870*** (0.000782)	0.0113*** (0.00186)
Maternity leave	22.52*** (0.138)	1.433*** (0.00639)	1.940*** (0.00936)	1.421*** (0.00875)
Adoption leave	-0.987*** (0.0873)	-0.0557*** (0.00944)	-0.0988*** (0.00932)	-0.248*** (0.0193)
Number of children	-1.531*** (0.0718)	-0.125*** (0.00471)	-0.0966*** (0.00520)	-0.224*** (0.00798)
Number of full-time employees	0.0265*** (0.00452)	0.000199 (0.000272)	-0.000846*** (0.000255)	0.000635 (0.000477)
Number of managers per employee	-6.134** (2.429)	-0.202 (0.135)	0.102 (0.143)	-0.418* (0.227)
Average length of management sick leave	0.00154 (0.000958)	0.0000852 (0.0000528)	0.0000646 (0.0000549)	0.000149* (0.0000828)
Year effects	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
Principal fixed effects	No	No	No	No
Individual fixed effects	Yes	Yes	Yes	Yes
Observations	2661107	1153592	2661107	2661107
R-squared	0.508	0.528	0.498	0.572
Adjusted R-squared	0.416	0.376	0.405	0.492

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers and secondary schools linked together by at least one person over the period 2007-2015. Columns (1), (2), (3) and (4) present the estimation results of the models (4), (2), (3) and (1) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and (4), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in column (4) are obtained by the Poisson Pseudo Maximum Likelihood (PPML) method, while estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

Table 4.17 – Variance decomposition of the absence duration of secondary school teachers (Model B)

	(1)	(2)	(3)	(4)
	Absence duration Y	Log of absence duration ln(Y)	Log of absence duration after transformation ln(Y + 0.1)	Absence duration Poisson(Y)
	Model B1	Model B2	Model B3	Model B4
Observations	2661107	1153592	2661107	2661107
R-squared	0.508	0.528	0.498	0.572
Adjusted R-squared	0.416	0.376	0.405	0.492
Number of individuals	410553	272810	410553	410553
Number of schools	8603	8571	8603	8603
Var(Y)	1667.58	1.92	5.33	-
Var(Individual FE)	885.69	1.00	2.56	7.17
Var(School FE)	55.89	0.12	0.17	0.42
Var(β x Covariates)	124.02	0.29	0.26	1.33
Cov(Individual FE; School FE)	-39.41	-0.09	-0.07	-0.24
Cov(School FE; β x Covariates)	-7.42	-0.01	-0.01	-0.06
Cov(Individual FE; β x Covariates)	-62.57	-0.10	-0.09	-1.06
Var(residual)	820.76	0.91	2.68	-

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers and secondary schools linked together by at least one person over the period 2007-2015. Columns (1), (2), (3) and (4) present the estimation results of the models (4), (2), (3) and (1) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and (4), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in column (4) are obtained by the Poisson Pseudo Maximum Likelihood (PPML) method, while estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

Table 4.18 – Variance decomposition of absence duration for female teachers (Model A)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration ln(Y)	Log of absence duration after transformation ln(Y + 0.1)
	Model A1	Model A2	Model A3
Observations	932233	462356	932233
R-squared	0.535	0.550	0.490
Adjusted R-squared	0.430	0.384	0.375
Number of individuals	157051	111214	157051
Number of schools	5234	5224	5234
Number of principals	8695	8652	8695
Var(Y)	1828.74	1.90	5.13
Var(Individual FE)	1094.10	1.18	2.53
Var(School FE)	428.22	0.85	1.39
Var(Principal FE)	308.53	0.62	1.07
Var(β x Covariates)	150.75	0.34	0.36
Cov(Individual FE; School FE)	-119.22	-0.23	-0.27
Cov(Individual FE; Principal FE)	-3.58	-0.01	-0.00
Cov(Individual FE; β x Covariates)	-81.21	-0.14	-0.12
Cov(Principal FE; School FE)	-289.73	-0.58	-1.03
Cov(School FE; β x Covariates)	-8.00	-0.01	0.01
Cov(Principal FE; β x Covariates)	-0.05	0.00	0.00
Var(residual)	851.59	0.86	2.62

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school **female** teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1), (2) and (3) present the estimation results of the models (4), (2) and (3) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for female teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

Table 4.19 – Variance decomposition of absence duration for male teachers (Model A)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration ln(Y)	Log of absence duration after transformation ln(Y + 0.1)
	Model A1	Model A2	Model A3
Observations	765528	257149	765528
R-squared	0.541	0.575	0.443
Adjusted R-squared	0.442	0.381	0.323
Number of individuals	121696	67411	121696
Number of schools	5236	5205	5236
Number of principals	8685	8468	8685
Var(Y)	1338.87	1.92	4.18
Var(Individual FE)	867.74	1.60	2.07
Var(School FE)	431.83	2.16	1.26
Var(Principal FE)	287.03	1.67	0.94
Var(β x Covariates)	100.28	0.33	0.08
Cov(Individual FE; School FE)	-144.75	-0.54	-0.30
Cov(Individual FE; Principal FE)	-6.71	-0.08	-0.01
Cov(Individual FE; β x Covariates)	-55.21	-0.14	-0.05
Cov(Principal FE; School FE)	-268.91	-1.56	-0.88
Cov(School FE; β x Covariates)	-5.18	-0.02	-0.00
Cov(Principal FE; β x Covariates)	-0.54	0.01	-0.00
Var(residual)	614.58	0.81	2.33

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school **male** teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1), (2) and (3) present the estimation results of the models (4), (2) and (3) respectively. They all include teacher fixed effects, school fixed effects and school principal fixed effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1), the logarithm of the total duration of sick leave in column (2) and the logarithm of the total duration of sick leave plus a constant (0.1 day of absence) in column (3). Since the logarithm is not defined for male teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Estimates in columns (1) to (3) are obtained by OLS using the algorithm proposed by Correia (2016).

Source: Ministry of Education (DEPP).

Table 4.20 – Contribution of time-varying covariates and fixed effects to teacher turnover (Model C)

	(1)	(2)
	Model C1	Model C2
Employment contract (ref : civil servants)		
Trial year for public servants	0.0102*** (0.000622)	0.00643*** (0.000452)
Short-term contracts	0.000915 (0.00274)	0.000984 (0.00123)
Open-ended contracts	0.00155 (0.00277)	0.00161 (0.00128)
Age	-0.000649*** (0.000101)	-0.000252*** (0.0000679)
Age ²	0.00000810*** (0.00000113)	0.00000341*** (0.000000757)
Seniority	0.00000826 (0.00000659)	0.0000145*** (0.00000400)
Maternity leave	0.00401*** (0.000347)	0.00264*** (0.000234)
Adoption leave	-0.000301 (0.000304)	0.000368 (0.000318)
Number of children	0.00000325 (0.0000257)	0.00000341 (0.0000174)
Number of full-time employees	0.00190*** (0.000265)	0.00249*** (0.000350)
Number of managers per employee	-0.540*** (0.0615)	-0.759*** (0.0921)
Average length of management sick leave	0.0000526*** (0.0000135)	0.0000523*** (0.0000155)
Temporary position	0.0172*** (0.00303)	0.0107*** (0.00130)
Year fixed effects	Yes	Yes
Establishment fixed effects	Yes	Yes
School principal fixed effects	No	Yes
Observations	2423253	1600163
R-squared	0.422	0.612
Adjusted R-squared	0.420	0.609

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: This Table presents the estimation results of model (i). The sample considered in column (1) is the largest connected set of secondary school teachers and secondary schools over the period 2007-2015. The sample considered in column (2) is the largest set of secondary school teachers, secondary schools and principals over the period 2007-2015. The variable explained is the teacher turnover rate in the school. For each school and year t, I calculate the number of teachers leaving school in year t+1 divided by the total number of teachers in the school in year t. Column (1) includes school fixed effects while column (2) includes school and principal fixed effects.

Source: Ministry of Education (DEPP).

Table 4.21 – Relationship between school fixed effects and school characteristics

	(1)	(2)	(3)
	FE estimated with absences in level: Y	FE estimated with absences in log: ln(Y)	FE estimated with absences in log after transformation: ln(Y+0.1)
Share of economically disadvantaged students in school	4.860** (1.913)	-0.0362 (0.109)	0.108 (0.0932)
Number of full-time employees	-0.0402*** (0.0118)	-0.00129* (0.000633)	0.000114 (0.000362)
School of the priority education network	2.392*** (0.823)	0.126*** (0.0393)	0.194*** (0.0248)
City location (ref: city in a small-sized urban area)			
Isolated city	-0.336 (1.583)	-0.0620 (0.0735)	0.164* (0.0766)
City in a medium-sized urban area	0.604 (1.375)	0.0401 (0.0442)	0.123*** (0.0439)
City in a large-sized urban area	0.0520 (0.692)	0.00700 (0.0306)	0.0459* (0.0218)
Type of school (ref: general senior high-school)			
Junior High-school	0.301 (0.871)	0.0615 (0.0415)	0.0768* (0.0435)
Professional senior high-school	1.202 (0.979)	-0.0159 (0.0576)	0.0507 (0.0386)
Constant	Yes	Yes	Yes
Department FE	No	No	No
Observations	4428	4424	4428
R-squared	0.013	0.011	0.021
Adjusted R-squared	0.012	0.009	0.019

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This Table explains the fixed effect of schools by school characteristics that are constant over the period 2007-2015. The estimates of school fixed effects in columns (1), (2) and (3) are obtained from models (4), (2) and (3) respectively.

Source: Ministry of Education (DEPP).

Table 4.22 – Relationship between school fixed effects and school characteristics, taking into account department fixed effects

	(1)	(2)	(3)
	FE estimated with absences in level: Y	FE estimated with absences in log: ln(Y)	FE estimated with absences in log after transformation: ln(Y+0.5)
Share of economically disadvantaged students in the school	1.840 (1.635)	0.0114 (0.114)	0.0423 (0.108)
Number of full-time employees	-0.0438*** (0.0110)	-0.00137** (0.000619)	-0.000271 (0.000367)
School of the priority education network	2.243*** (0.769)	0.111*** (0.0375)	0.179*** (0.0283)
City location (ref: city in a small-sized urban area)			
Isolated city	0.285 (1.690)	-0.0463 (0.0752)	0.171** (0.0749)
City in a medium-sized urban area	1.465 (1.186)	0.0669 (0.0454)	0.144*** (0.0444)
City in a large-sized urban area	0.587 (0.597)	0.0236 (0.0303)	0.0525** (0.0220)
Type of school (ref: general senior high-school)			
Junior High-school	0.277 (0.814)	0.0616 (0.0398)	0.0634 (0.0437)
Professional senior high-school	1.298 (0.908)	-0.0261 (0.0589)	0.0506 (0.0377)
Constant	Yes	Yes	Yes
Departement FE	Yes	Yes	Yes
Observations	4428	4424	4428
R-squared	0.055	0.034	0.046
Adjusted R-squared	0.047	0.026	0.038

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: This Table explains the fixed effect of schools by school characteristics that are constant over the period 2007-2015. The estimates of school fixed effects in columns (1), (2) and (3) are obtained from models (4), (2) and (3) respectively. There are a hundred departments in France and each column includes department fixed effects.

Source: Ministry of Education (DEPP).

Table 4.23 – Share of teachers exposed to major psychosocial risk factors for different teacher samples

	Secondary school teachers surveyed in the 2013 and/or 2016 working conditions surveys			
	Secondary school teachers (unweighted)	Secondary school teachers whose schools were found in Ministry of Education management records	Secondary school teachers working in schools for which we estimated a school fixed effect (Model B)	Secondary school teachers working in schools for which we estimated both a school and a principal fixed effect (Model A)
Number of observations	1537	1113	1018	667
Number of secondary school teachers	1034	733	667	443
Number of schools	Unknown	707	664	433
Work intensity (%)				
I am asked to do an excessive amount of work	43.5	44.2	44.7	44.1
I have to perform certain operations quickly when they would require more care	31.1	32.4	33.2	33
I feel exploited	12	11.4	11.6	11.7
I feel overwhelmed by too rapid changes	12.4	12.7	13.1	13
I have to hurry	47.1	48.5	49	50.8
I don't have enough time to do my job properly	28.3	28.6	28.7	28.8
I often have to interrupt one task to perform another that was not scheduled	42.7	42.9	42.7	41.2
Lack of interest in work (%)				
I am not proud to work in this organization	36.7	35.2	34.3	35.1
I never feel the pride of having done my job well	35.1	36.9	37.5	36.9
I do not feel like I am doing anything useful for others	18.5	19.9	20.3	19.5
I get bored at work	5.6	5.8	5.7	5.2
I do not have the opportunity to do things I like	21.7	23.4	23.3	22.5
Hostile Behaviors (%)				
Some people unfairly criticize my work	18.2	18	18.1	18.4
Some people keep me from expressing myself	10.1	9.1	8.7	8.8
Some people ignore me and pretend I am not here	23	23.6	23.8	23.7
Some people give me unnecessary or degrading tasks	4.9	5.4	5.3	5.1
Tensions with students and their families (%)				
I encounter situations of tension with the public	62.2	65.3	66.7	64.3
I'm afraid during my work for my safety or the safety of others	6.6	6.6	6.6	6.1
Lack of support from hierarchy (%)				
My supervisor does not help me with my tasks	26.6	25.8	26	25.6
My superior does not pay attention to what I say	18.7	18.9	19.1	19
People who evaluate my work do not know it well	34.2	35.1	35.8	35.8
I encounter situations of tension with my supervisors	26.3	27.3	27.8	28.2
I do not agree with my superiors on how to do my job well	16.5	17.1	17.2	16.6
Lack of support from colleagues (%)				
The colleagues I work with do not help me with my tasks	14.1	13.9	14.4	14.1
I do not feel like I am part of a team	29.9	30.6	30.6	30.4
The colleagues I work with are unfriendly	6.8	6.6	6.5	6.6
I can not discuss with my colleagues when I disagree with them	4.6	4.7	4.6	5.1
My colleagues do not help me when I have difficulty doing complicated work	15.2	15.3	14.7	15.6
I do not have the opportunity to cooperate to do my job properly	8	8.4	8.3	8.7
I encounter situations of tension with my colleagues	26	25.8	26.5	27.3

Note: The sample in column (1) corresponds to all secondary school teachers interviewed in the 2013 and 2016 working conditions surveys. Column (2) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records. Column (3) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher fixed effects. Column (4) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher and principal fixed effects. I group the questions from the 2013 and 2016 Working Condition Surveys according to six psychosocial risk factors. For each sample, the Table shows the proportion of teachers concerned by each type of exposure.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.24 – Exposure score to major psychosocial risk factors for different teacher samples

Secondary school teachers surveyed in the 2013 and/or 2016 working conditions surveys				
	Secondary school teachers (unweighted)	Secondary school teachers whose school was found in Ministry of Education management records	Secondary school teachers working in schools for which we estimated a school fixed effect (Model B)	Secondary school teachers working in schools for which we estimated both a school and a principal fixed effect (Model A)
Number of observations	1537	1113	1018	667
Number of secondary school teachers	1034	733	667	443
Number of schools	Unknown	707	664	433
Working Conditions Score (normalized)				
Work intensity	-0.02	0	0.012	0.01
Lack of interest in work	-0.031	0	0	-0.017
Hostile Behaviors	0.002	0	-0.002	0
Tensions with students and their families	-0.056	0	0.025	-0.026
Lack of support from hierarchy	-0.013	0	0.012	0.008
Lack of support from colleagues	-0.007	0	0.003	0.019

Note: The sample in column (1) corresponds to all secondary school teachers interviewed in the 2013 and 2016 working conditions surveys. Column (2) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records. Column (3) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher fixed effects. Column (4) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher and principal fixed effects. I group the questions from the 2013 and 2016 Working Condition Surveys according to six psychosocial risk factors. Each question is coded by a dummy variable that takes the value 1 if the individual feels exposed or very exposed and 0 if he feels little or not exposed. I summarize the information by calculating the average of the dummy variables that make up each psychosocial risk factor. These indicators are normalized (mean 0 and variance 1) for the sample of teachers interviewed by the Working Condition Survey (in 2013 or 2016) whose school could be matched with the Ministry of Education's administrative records (sample in column 2 is taking as reference). The Table presents the values obtained by each sample considered.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.25 – Percentage of teachers in poor health and percentage of teachers reporting preventive actions in their schools

Secondary school teachers surveyed in the 2013 and/or 2016 working conditions surveys				
	Secondary school teachers (unweighted)	Secondary school teachers whose schools were found in Ministry of Education management records	Secondary school teachers working in schools for which we estimated a school fixed effect (Model B)	Secondary school teachers working in schools for which we estimated both a school and a principal fixed effect (Model A)
Subjective health (%)				
Bad or very bad health	4.6	4.5	4.1	3.9
Well-being index WHO-5 ≤ 6	6.2	6.4	6.5	6.4
Well-being index WHO-5 ≤ 8	10.5	10.6	10.6	11.1
Well-being index WHO-5 ≤ 10	16.5	17.1	17.4	17.5
Prevention (%)				
Have received information in the past 12 months about the health or safety risks posed by their work	15	14.1	13.8	13.3
Have received security training from their administration within the past 12 months	17	17.3	17.1	17.8
Have had knowledge in the last 12 months of a document written by management describing the risks associated with working in their institution	15.7	15.6	15.2	14.5
Have undergone a medical examination by an occupational physician within the last 24 months	10.9	8.4	7.4	6.9

Note: The sample in column (1) corresponds to all secondary school teachers interviewed in the 2013 and 2016 working conditions surveys. Column (2) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records. Column (3) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher fixed effects. Column (4) restricts the sample to secondary school teachers interviewed whose school could be matched with the Ministry of Education's administrative records and for whom a school fixed effect could be estimated using model (2) to (4) including teacher and principal fixed effects. For each sample, the Table shows the proportion of secondary school teachers with self-reported health problems (general health status rated "poor" or "very poor", WHO-5 psychological well-being index less than 6, WHO-5 psychological well-being index less than 8 and WHO-5 psychological well-being index less than 10) and the proportion of secondary school teachers who report preventive actions in their schools.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.26 – Absence duration (in logarithm) for individuals transiting between terciles of school fixed effects in year $t=0$ (Model 3)

Teacher transitions between terciles of school FE	Observations		Logarithm of duration of teacher absences						
	N	%	$t = -3$	$t = -2$	$t = -1$	$t = 0$	$t = 1$	$t = 2$	
1 to 1	4227	55.3	0.4	0.41	0.42	0.31	0.36	0.42	
1 to 2	2338	30.6	0.4	0.44	0.48	0.62	0.64	0.77	
1 to 3	1080	14.1	0.38	0.44	0.52	0.99	1.02	1.23	
2 to 1	3940	39.8	0.57	0.59	0.67	0.27	0.27	0.31	
2 to 2	3829	38.7	0.58	0.6	0.66	0.56	0.58	0.68	
2 to 3	2125	21.5	0.57	0.6	0.67	0.88	0.87	0.98	
3 to 1	3042	29.6	0.9	0.91	1.01	0.34	0.36	0.37	
3 to 2	3714	36.2	0.78	0.82	0.88	0.52	0.48	0.58	
3 to 3	3515	34.2	0.78	0.83	0.91	0.8	0.76	0.82	

Note: Schools are divided into terciles of school fixed effects based on the estimates of model (3). Schools in the third tercile are expected to increase teacher absences the most significantly. I consider all teachers observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who change schools in year $t=0$, but remain in the same tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. There are nine possible transitions between school terciles in year $t=0$, which defines nine transition groups. For example, the transition group "1 to 1" refers to teachers who change schools in year $t=0$ but remain in a school of tercile 1 in $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$. For each group, this Table shows the number of observations (teachers) and the logarithm of the duration of absence of teachers in years $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.27 – Absence duration (in days) for teachers transiting between terciles of establishment fixed effects in year $t=0$ (Model 4)

Teacher transitions between terciles of school FE	Observations		Duration of teacher absences						
	N	%	$t = -3$	$t = -2$	$t = -1$	$t = 0$	$t = 1$	$t = 2$	
1 to 1	5026	54.7	5.57	5.47	4.88	5.35	5.83	5.89	
1 to 2	2725	29.7	5.92	6.19	6.42	8.69	11.1	12.72	
1 to 3	1433	15.6	6.66	7.4	10.32	22.54	30.14	36.19	
2 to 1	4041	41.3	6.87	6.54	6.81	4.04	3.8	3.49	
2 to 2	3547	36.2	5.45	5.74	5.7	6.05	5.63	5.62	
2 to 3	2208	22.5	6.41	6.55	6.28	8.53	9.65	11.56	
3 to 1	2898	33.1	10.5	10.64	10.92	4.65	3.39	2.79	
3 to 2	2954	33.8	7.71	8.5	8.53	4.98	4.71	3.83	
3 to 3	2898	33.1	7.35	7.16	7.8	6.97	6.91	6.86	

Note: Schools are divided into terciles of school fixed effects based on the estimates of model (4). Schools in the third tercile are expected to increase teacher absences the most significantly. I consider all teachers observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who change schools in year $t=0$, but remain in the same tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. There are nine possible transitions between school terciles in year $t=0$, which defines nine transition groups. For example, the transition group "1 to 1" refers to teachers who change schools in year $t=0$ but remain in a school of tercile 1 in $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$. For each group, this Table shows the number of observations (teachers) and the duration of absence (in days) of teachers in years $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.28 – Absence duration (in logarithm) in schools transiting between terciles of school principal fixed effects in year $t=0$ (Model 3)

School transitions between terciles of school principal FE	Observations		Logarithm of mean absence duration in the school					
	N	%	$t = -3$	$t = -2$	$t = -1$	$t = 0$	$t = 1$	$t = 2$
1 to 1	976	76.1	0.64	0.62	0.6	0.56	0.54	0.56
1 to 2	272	21.2	0.47	0.47	0.51	0.68	0.69	0.71
1 to 3	34	2.7	0.51	0.42	0.53	1	1.22	1.13
2 to 1	250	20.9	0.85	0.8	0.75	0.52	0.49	0.49
2 to 2	669	55.9	0.6	0.59	0.58	0.52	0.53	0.54
2 to 3	278	23.2	0.51	0.53	0.56	0.69	0.76	0.74
3 to 1	24	1.7	1.12	1.09	0.88	0.4	0.34	0.31
3 to 2	228	16.2	0.8	0.81	0.74	0.51	0.39	0.45
3 to 3	1152	82.1	0.67	0.65	0.64	0.6	0.61	0.63

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (3). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who are run by a new school principal in year $t=0$, but remain in the same tercile of school principal between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. There are nine possible transitions between school principal terciles in year $t=0$, which defines nine transition groups. For example, the transition group "1 to 1" refers to school who have a new school principal in year $t=0$ but remain in tercile 1 in $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$. For each group, this Table shows the number of observations (schools) and the logarithm of the mean duration of absences of teachers in the school in years $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.29 – Absence duration (in days) in schools transiting between terciles of school principal fixed effects in year $t=0$ (Model 4)

School transitions between terciles of school principal FE	Observations		Logarithm of mean absence duration in the school					
	N	%	$t = -3$	$t = -2$	$t = -1$	$t = 0$	$t = 1$	$t = 2$
1 to 1	990	77.2	10.77	10.95	10.49	10.36	10.27	10.01
1 to 2	262	20.4	7.42	8.03	9.1	12.93	14.05	14.74
1 to 3	31	2.4	8.6	6.03	10.03	21.77	25.54	25.11
2 to 1	238	19	16.19	15.93	14.56	9.71	8.34	7.73
2 to 2	732	58.5	9.76	9.61	9.9	9.44	9.5	9.6
2 to 3	282	22.5	8.04	8.52	9.86	14.26	16.92	15.77
3 to 1	36	2.6	32.89	26.61	18.36	9.9	10	6.28
3 to 2	248	17.9	17.12	15.8	14.13	9.33	8.06	7.91
3 to 3	1102	79.5	11.13	11.06	11.41	11.34	11.16	11.05

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (4). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who are run by a new school principal in year $t=0$, but remain in the same tercile of school principal between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. There are nine possible transitions between school principal terciles in year $t=0$, which defines nine transition groups. For example, the transition group "1 to 1" refers to school who have a new school principal in year $t=0$ but remain in tercile 1 in $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$. For each group, this Table shows the number of observations (schools) and the mean duration of absences of teachers in the school in years $t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$ and $t=2$.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.30 – Turnover in schools transiting between terciles of school principal fixed effects in year t=0 (model i)

School transitions between terciles of school principal FE	Observations		Teacher turnover in the school					
	N	%	t = -3	t = -2	t = -1	t = 0	t = 1	t = 2
1 to 1	693	85.3	0.15	0.15	0.15	0.15	0.15	0.15
1 to 2	97	11.9	0.12	0.12	0.12	0.17	0.2	0.19
1 to 3	22	2.7	0.09	0.1	0.14	0.23	0.26	0.28
2 to 1	275	30.3	0.16	0.17	0.16	0.12	0.11	0.11
2 to 2	512	56.4	0.15	0.14	0.15	0.15	0.16	0.16
2 to 3	120	13.2	0.14	0.14	0.16	0.21	0.22	0.24
3 to 1	31	3.5	0.21	0.23	0.18	0.11	0.1	0.08
3 to 2	208	23.2	0.19	0.18	0.18	0.14	0.14	0.13
3 to 3	657	73.3	0.16	0.16	0.16	0.17	0.17	0.16

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (3). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years (t=-3, t=-2, t=-1, t=0, t=1, t=2), who are run by a new school principal in year t=0, but remain in the same tercile of school principal between t=-3 and t=-1 and between t=0 and t=2. There are nine possible transitions between school principal terciles in year t=0, which defines nine transition groups. For example, the transition group "1 to 1" refers to school who have a new school principal in year t=0 but remain in tercile 1 in t=-3, t=-2, t=-1, t=0, t=1 and t=2. For each group, this Table shows the number of observations (schools) and the rate of teacher turnover in the school in years t=-3, t=-2, t=-1, t=0, t=1 and t=2.

Source: Ministry of Education (DEPP) and Ministry of Labor (DARES).

Table 4.31 – Relationship between the conditional variance of absences and the conditional expectation of absences

	(1)	(2)	(3)
	Explained variable		
	$\ln(\widehat{Var}[Y_{ijmt} i])$	$\ln(\widehat{Var}[Y_{ijmt} j])$	$\ln(\widehat{Var}[Y_{ijmt} j, m, t])$
Explanatory variable :			
$\ln(E[\widehat{Y}_{ijmt} i])$	1.765*** (0.00133)		
$\ln(E[\widehat{Y}_{ijmt} j])$		1.891*** (0.0101)	
$\ln(E[\widehat{Y}_{ijmt} j, m, t])$			2.273*** (0.00474)
R2	0.90	0.81	0.86
Observations	193602	8377	37583
Sample restriction	Individuals observed 9 times only	Schools with more than 30 observations over 2007-2015	Schools x principals x years with more than 30 observations over 2007-2015

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: This Table shows how the logarithm of the variance of absence duration varies with the average absence duration. In column (1), I calculate the average and variance of absence duration of each teacher observed nine times over the 2007-2015 period. In column (2), I calculate the average and variance of absence duration for each school with more than 30 observations over the period 2007-2015. In column (3), I calculate the average and variance of absence duration for each school x school principal x year cell with more than 30 observations over the period 2007-2015.

Source: Ministry of Education (DEPP).

Table 4.32 – Comparison of estimates obtained by PPML and OLS after log-linearization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Poisson(Y)	ln(Y)	ln(Y + 1)	ln(Y + 0.5)	ln(Y + 0.1)	ln(Y + 0.05)	ln(Y + 0.01)
Employment contract (ref : civil servants)							
Trial year for public servants	-0.624*** (0.0219)	-0.208*** (0.0122)	-0.214*** (0.00689)	-0.256*** (0.00838)	-0.360*** (0.0123)	-0.406*** (0.0141)	-0.512*** (0.0185)
Short-term contracts	-0.992*** (0.0360)	-0.295*** (0.0208)	-0.476*** (0.0111)	-0.591*** (0.0135)	-0.878*** (0.0198)	-1.006*** (0.0228)	-1.305*** (0.0298)
Open-ended contracts	-0.667*** (0.0506)	-0.172*** (0.0298)	-0.319*** (0.0163)	-0.395*** (0.0198)	-0.584*** (0.0289)	-0.667*** (0.0331)	-0.864*** (0.0431)
Age	-0.0921*** (0.00511)	-0.0512*** (0.00291)	-0.0467*** (0.00164)	-0.0555*** (0.00198)	-0.0781*** (0.00285)	-0.0883*** (0.00326)	-0.112*** (0.00424)
Age ²	0.00229*** (0.0000516)	0.00110*** (0.0000303)	0.000822*** (0.0000175)	0.000950*** (0.0000210)	0.00126*** (0.0000301)	0.00140*** (0.0000344)	0.00172*** (0.0000446)
Seniority	0.0113*** (0.00186)	0.00890*** (0.00104)	0.00561*** (0.000440)	0.00655*** (0.000535)	0.00870*** (0.000782)	0.00961*** (0.000895)	0.0117*** (0.00117)
Maternity leave	1.421*** (0.00875)	1.433*** (0.00639)	1.376*** (0.00592)	1.557*** (0.00691)	1.940*** (0.00936)	2.097*** (0.0105)	2.454*** (0.0132)
Adoption leave	-0.248*** (0.0193)	-0.0557*** (0.00944)	-0.0545*** (0.00495)	-0.0671*** (0.00615)	-0.0988*** (0.00932)	-0.113*** (0.0108)	-0.146*** (0.0144)
Number of children	-0.224*** (0.00798)	-0.125*** (0.00471)	-0.0821*** (0.00299)	-0.0882*** (0.00360)	-0.0966*** (0.00520)	-0.0990*** (0.00594)	-0.103*** (0.00772)
Number of full-time employees	0.000635 (0.000477)	0.000199 (0.000272)	-0.000209 (0.000146)	-0.000387** (0.000176)	-0.000846*** (0.000255)	-0.00105*** (0.000291)	-0.00154*** (0.000379)
Number of managers per employee	-0.418* (0.227)	-0.202 (0.135)	-0.0130 (0.0825)	0.0158 (0.0993)	0.102 (0.143)	0.143 (0.164)	0.242 (0.213)
Average length of management sick leave	0.000149* (0.0000828)	0.0000852 (0.0000528)	0.0000503 (0.0000316)	0.0000553 (0.0000381)	0.0000646 (0.0000549)	0.0000681 (0.0000628)	0.0000760 (0.0000817)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Principal fixed effects	No	No	No	No	No	No	No
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2661107	1153592	2661107	2661107	2661107	2661107	2661107
R-squared	0.572	0.528	0.513	0.510	0.498	0.493	0.483
Adjusted R-squared	0.492	0.376	0.422	0.418	0.405	0.398	0.387

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers and secondary schools linked together by at least one person over the period 2007-2015. Column (1) presents the results obtained from model (1) estimated by the PPML method. The variable explained is the total duration of absence during the school year. Column (2) shows the estimates obtained by OLS from model (2) where the explained variable is the logarithm of absence duration. Columns (3) to (7) present the OLS estimates from model (3) where the explained variable is the logarithm of the total duration of sick leave plus a constant that ranges from 1 (column 3) to 0.01 (column 7). Since the logarithm is not defined for teachers who have not been absent for a single day during the school year, the sample size in column (2) is reduced by about half. Each column includes teacher fixed effects and school fixed effects.

Source: Ministry of Education (DEPP).

Table 4.33 – Contribution of time-varying covariates, fixed effects, peers' fixed effects and peers' covariates effects to the duration of absence of secondary school teachers (Arcidiacono et al. 2012) (Model E)

	(1)	(2)	(3)
	Absence duration	duration after transformation	Log of absence duration after transformation
	Y	ln(Y + 1)	ln(Y + 0.1)
	Model E1	Model E2	Model E3
Peer (fixed) effect	0.00118	0.0667***	0.0827***
	(0.00229)	(0.00238)	(0.0023)
Employment contract (ref : civil servant status)			
Share of civil servant teachers	-21.88***	-0.484***	-0.539***
	(1.376)	(0.0459)	(0.0554)
Civil servants on probationary year	-3.600***	-0.197***	-0.238***
	(0.222)	(0.00741)	(0.00893)
Share of teachers on probationary year	-16.15***	-0.270***	-0.284***
	(1.680)	(0.0561)	(0.0676)
Short-term contract	-5.178***	-0.443***	-0.554***
	(0.128)	(0.00426)	(0.00513)
Share of teachers on short-term contract	0.360	-0.115**	-0.152**
	(1.518)	(0.0506)	(0.0610)
Permanent contract	-4.328***	-0.296***	-0.370***
	(0.274)	(0.00914)	(0.0110)
Age	-2.498***	-0.0558***	-0.0669***
	(0.0209)	(0.000697)	(0.000840)
Mean age of co-workers	0.562***	0.0195***	0.0236***
	(0.0830)	(0.00277)	(0.00334)
Age squared	0.0378***	0.000855***	0.000994***
	(0.000229)	(0.00000765)	(0.00000922)
Mean square age of co-workers	-0.00553***	-0.000249***	-0.000306***
	(0.000935)	(0.0000312)	(0.0000376)
Seniority	0.152***	0.00628***	0.00742***
	(0.00364)	(0.000122)	(0.000146)
Mean seniority of co-workers	0.0235	0.00718***	0.00902***
	(0.0166)	(0.000556)	(0.000670)
Maternity leave	22.06***	1.361***	1.541***
	(0.131)	(0.00437)	(0.00527)
Adoption leave	-1.130***	-0.0623***	-0.0773***
	(0.168)	(0.00559)	(0.00674)
Number of children	-1.630***	-0.0921***	-0.100***
	(0.0193)	(0.000644)	(0.000776)
Number of full-time employees	-0.0132***	-0.00155***	-0.00196***
	(0.000838)	(0.0000281)	(0.0000340)
Number of managers per employees	-3.357*	-0.268***	-0.286***
	(1.696)	(0.0566)	(0.0682)
Duration of sick leave for managers	0.00148	0.0000395	0.0000389
	(0.000922)	(0.0000308)	(0.0000371)
Year effects	Yes	Yes	Yes
Establishment fixed effects	Yes	Yes	Yes
Principal fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Observations	1721420	1721420	1721420
R-squared	0.527	0.525	0.523
Adjusted R-squared	0.429	0.426	0.424

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. The Table shows in bold the coefficients that capture the effects of peers (fixed effects of peers and characteristics of peers) in the school. Estimates in columns (1) to (3) are obtained using the algorithm proposed by Arcidiacono et al. (2012).

Source: Ministry of Education (DEPP).

Table 4.34 – Variance decomposition of the absence duration of secondary school teachers taking into account peers' fixed effect and peer's covariates effects (Arcidiacono et al. 2012) (Model E)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration after transformation $\ln(Y + 1)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model E1	Model E2	Model E3
Observations	1721420	1721420	1721420
R-squared	0.527	0.525	0.523
Adjusted R-squared	0.429	0.426	0.424
Number of individuals	283933	283933	283933
Number of establishments	4782	4782	4782
Number of principals	8120	8120	8120
Var(Y)	1611.22	1.79	2.58
Var(Individual FE)	894.56	0.90	1.30
Var(Establishment FE)	132.14	0.13	0.19
Var(Principal FE)	74.03	0.07	0.10
Var(β x Covariates)	97.48	0.11	0.14
Cov(Individual FE; Establishment FE)	-59.49	-0.05	-0.07
Cov(Individual FE; Principal FE)	-3.45	-0.00	-0.00
Cov(Individual FE; β x Covariates)	-44.68	-0.02	-0.02
Cov(Principal FE; Establishment FE)	-63.98	-0.06	-0.08
Cov(Establishment FE; β x Covariates)	-2.36	-0.01	-0.01
Cov(Principal FE; β x Covariates)	-0.54	-0.00	-0.00
Var(residual)	761.89	0.85	1.23

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. Estimates in columns (1) to (3) are obtained using the algorithm proposed by Arcidiacono et al. (2012).

Source: Ministry of Education (DEPP).

Table 4.35 – Contribution of time-varying covariates, fixed effects and peer effects to the duration of absence of secondary school teachers (Ichino & Maggi 2000) (Model F)

	(1)	(2)	(3)
	Absence duration Y	duration after transformation $\ln(Y + 1)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model F1	Model F2	Model F3
Peer effect	0.0257*** (0.00329)	0.189*** (0.00303)	0.204*** (0.00302)
Employment contract (ref : civil servant status)			
Trial period before civil servant status	-5.173*** (0.163)	-0.288*** (0.00656)	-0.345*** (0.00800)
Short-term contract	-6.885*** (0.251)	-0.519*** (0.0109)	-0.643*** (0.0134)
Permanent contract	-5.934*** (0.405)	-0.359*** (0.0162)	-0.443*** (0.0197)
Age	-2.386*** (0.0499)	-0.0492*** (0.00163)	-0.0583*** (0.00196)
Age ²	0.0375*** (0.000572)	0.000823*** (0.0000174)	0.000949*** (0.0000209)
Seniority	0.173*** (0.0115)	0.00579*** (0.000439)	0.00677*** (0.000533)
Maternity leave	22.63*** (0.138)	1.380*** (0.00593)	1.562*** (0.00691)
Adoption leave	-0.967*** (0.0854)	-0.0537*** (0.00494)	-0.0661*** (0.00614)
Number of children	-1.623*** (0.0706)	-0.0880*** (0.00297)	-0.0953*** (0.00358)
Number of full-time employees	-0.0160*** (0.00180)	-0.000863*** (0.0000567)	-0.00101*** (0.0000682)
Number of managers per employees	0.236 (2.075)	0.309*** (0.0687)	0.382*** (0.0826)
Duration of sick leave for managers	0.00189* (0.000942)	0.0000419 (0.0000310)	0.0000437 (0.0000373)
Year effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Department fixed effects	Yes	Yes	Yes
Observations	2657547	2657547	2657547
R-squared	0.504	0.510	0.507
Adjusted R-squared	0.414	0.421	0.417

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. Estimates in columns (1) to (3) are obtained using the method proposed by Ichino & Maggi (2000).

Source: Ministry of Education (DEPP).

Table 4.36 – Variance decomposition of the absence duration of secondary school teachers taking into account peers effects (Ichino & Maggi 2000) (Model F)

	(1)	(2)	(3)
	Absence duration Y	Log of absence duration after transformation $\ln(Y + 1)$	Log of absence duration after transformation $\ln(Y + 0.1)$
	Model F1	Model F2	Model F3
Observations	2657547	2657547	2657547
R-squared	0.504	0.510	0.507
Adjusted R-squared	0.414	0.421	0.417
Number of individuals	409408	409408	409408
Number of departments	100	100	100
Var(Y)	1669.04	1.82	2.62
Var(Individual FE)	872.02	0.88	1.26
Var(Department FE)	2.33	0.00	0.00
Var(β x Covariates)	113.50	0.13	0.17
Cov(Individual FE; Department FE)	-2.12	-0.00	-0.00
Cov(Individual FE; β x Covariates)	-51.69	-0.03	-0.04
Cov(Department FE; β x Covariates)	-0.94	-0.00	-0.00
Var(residual)	860.82	0.90	1.31

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust to heteroscedasticity.

Note: The sample used is the largest set of secondary school teachers, secondary schools and principals linked together by at least one person over the period 2007-2015. Columns (1) to (3) present the estimation results of model (5) taking into account peer effects. The variable explained is the total duration (in number of days) of teachers' sick leave during the school year in column (1) and the logarithm of the total duration of sick leave plus a constant in columns (2) and (3). Each column includes teacher fixed effects, school fixed effects and school principal fixed effects. Estimates in columns (1) to (3) are obtained using the method proposed by Ichino & Maggi (2000).

Source: Ministry of Education (DEPP).

Figures

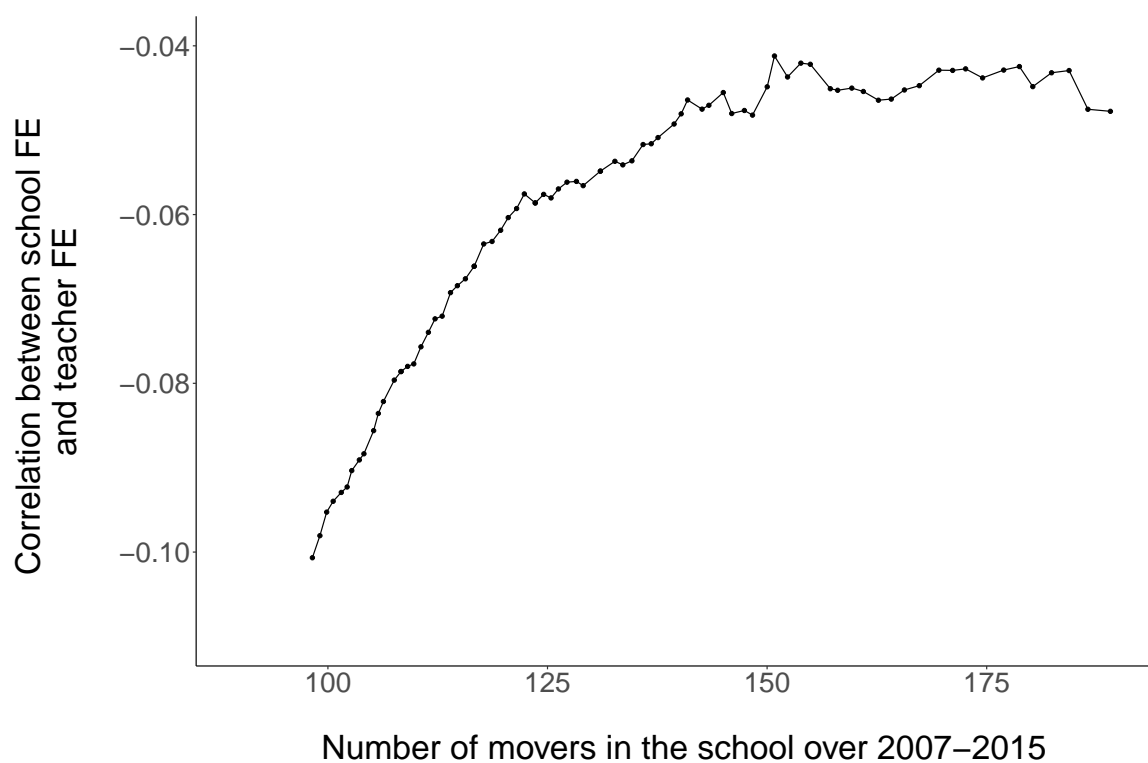


Figure 4.6 – Correlation between teacher fixed effects and school fixed effects according to the number of "movers" (teachers) per school

Note: The fixed effects of schools and teachers are estimated from model (3). Schools are classified according to the number of teachers entering or leaving the school during the period 2007-2015 (movers). Schools with the lowest number of movers are gradually eliminated (percentile by percentile) so that the average number of movers per school increases in the sample (X-axis). The Y-axis shows the correlation between teacher fixed effects and school fixed effects for each sample. The Figure confirms that the limited mobility bias decreases as the number of teachers moving per school increases.

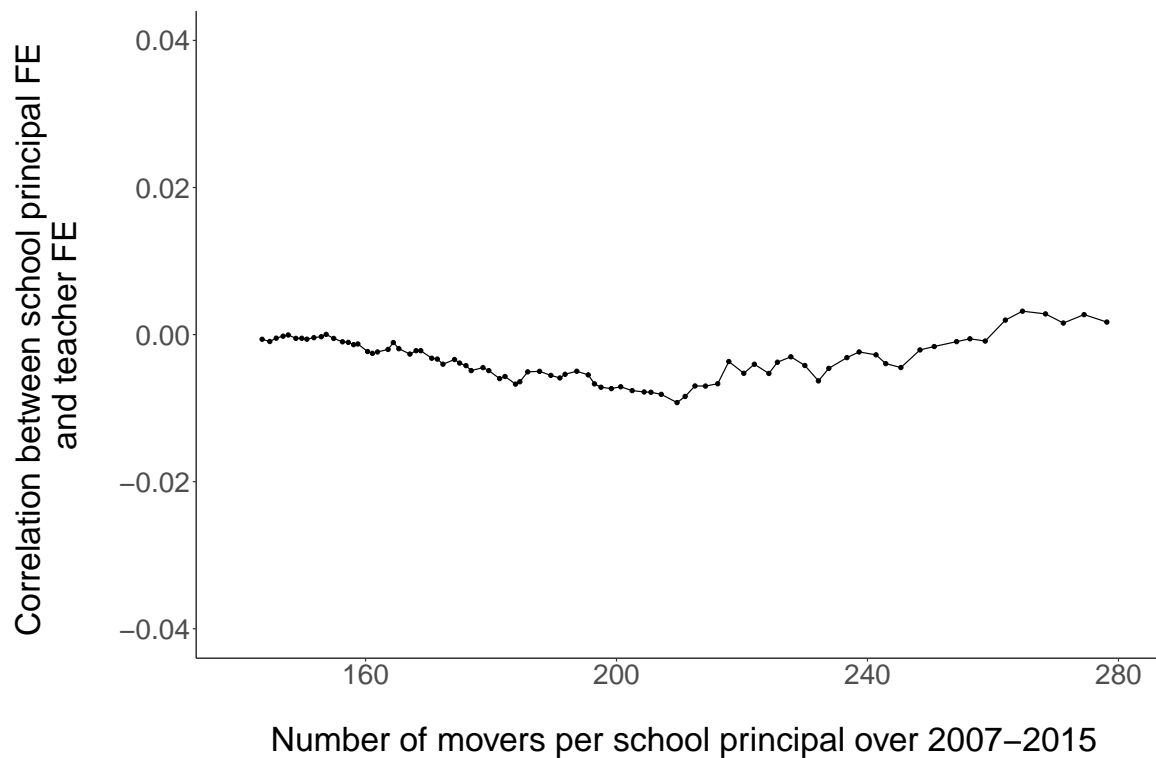


Figure 4.7 – Correlation between teacher fixed effects and school principal fixed effects according to the number of "movers" (teachers) per school principal

Note: The fixed effects of teachers and principals are estimated from model (3). School principals are classified according to the number of teachers entering or leaving their school during the period 2007-2015 (movers). School principals with the lowest number of movers in their schools are gradually eliminated (percentile by percentile) so that the average number of movers per principal increases in the sample (X-axis). The Y-axis shows the correlation between teacher fixed effects and principal fixed effects for each sample. The Figure confirms that the limited mobility bias decreases as the number of teachers moving per school principal increases.

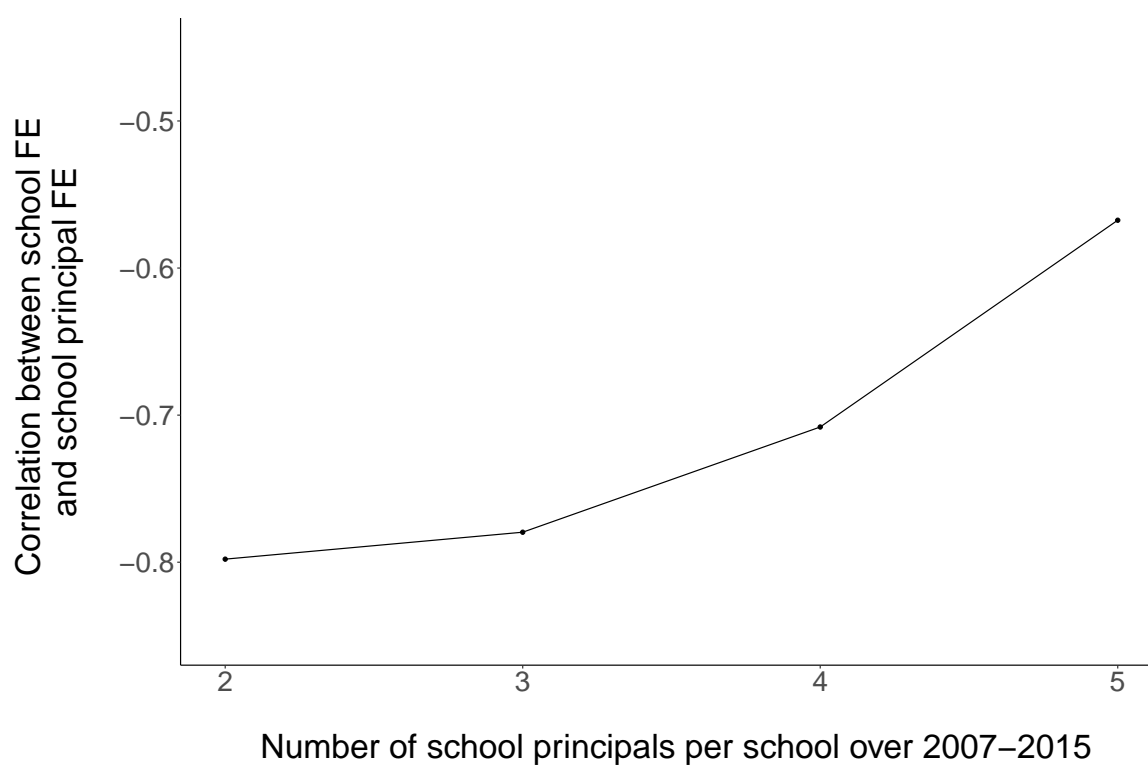


Figure 4.8 – Correlation between school fixed effects and school principal fixed effects according to the number of "movers" (principals) per school

Note: The fixed effects of schools and principals are estimated from model (3). Schools are classified according to the number of school principal entering or leaving the school during the period 2007-2015 (movers). Schools with the lowest number of movers are gradually eliminated (percentile by percentile) so that the average number of movers per school increases in the sample (X-axis). The Y-axis shows the correlation between school principal fixed effects and school fixed effects for each sample. The Figure confirms that the limited mobility bias decreases as the number of principals moving per school increases.

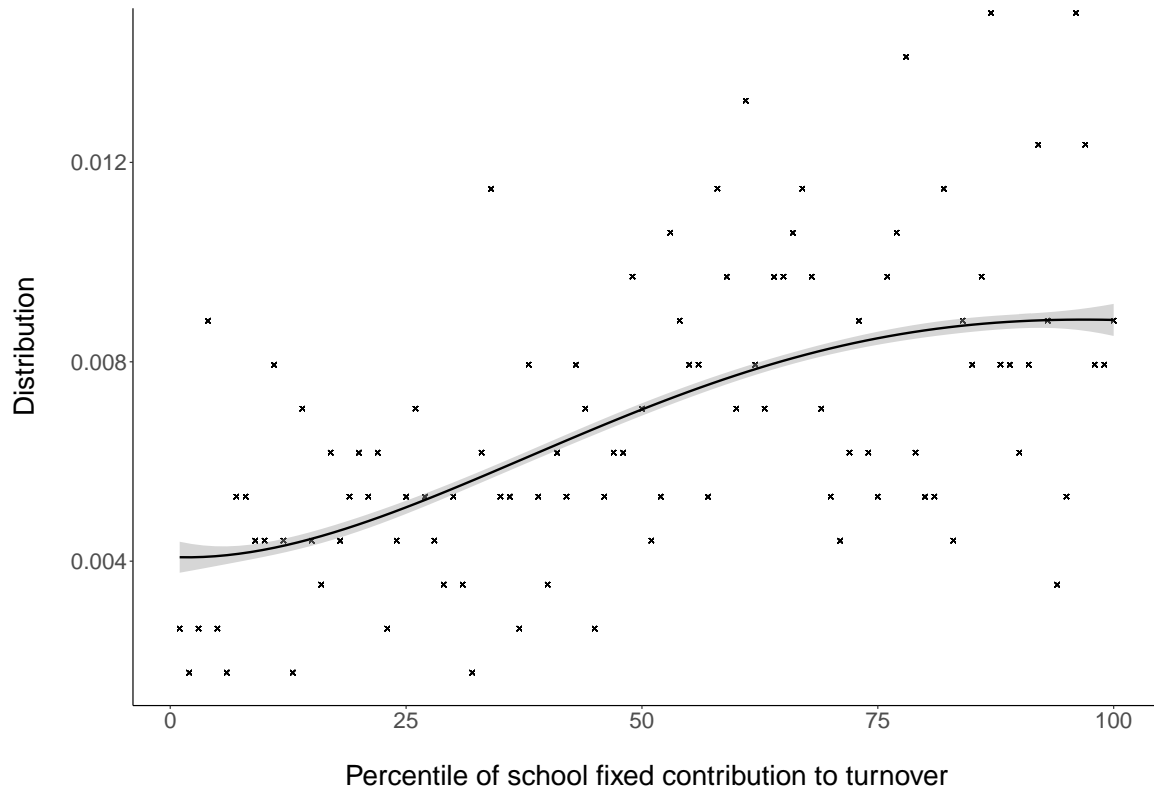


Figure 4.9 – Distribution of the Priority Education Network according to school fixed effects on teacher turnover (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on teacher turnover, estimated from model (i). I calculate the number of schools in the Priority Education Network in each school percentile that I divide by the total number of schools in the Priority Education Network. The few schools that joined or left the Priority Education Network during the period 2007-2015 are counted twice. The Y-axis shows the distribution (density) of the Priority Education Network among school percentiles.

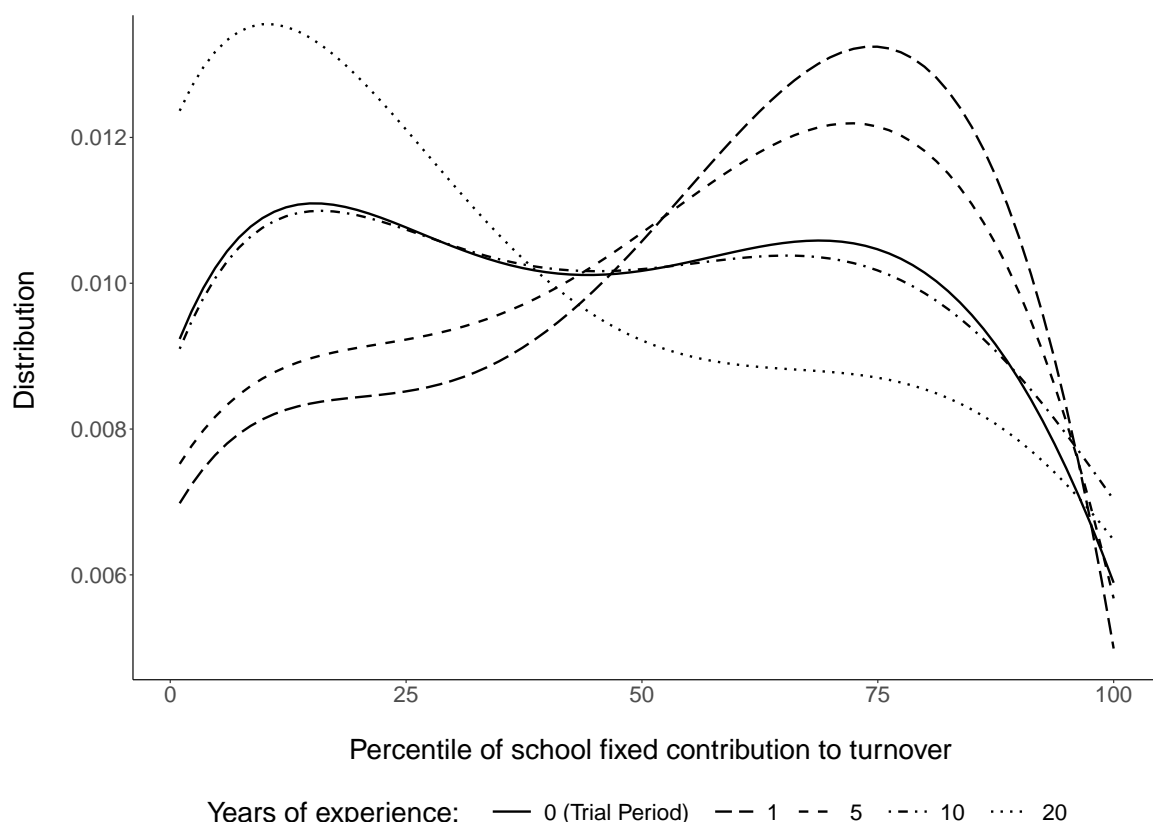


Figure 4.10 – Distribution of civil servant teachers according to experience and school fixed effects on teacher turnover (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on teacher turnover, estimated from model (i). I consider 5 levels of professional experience in teaching: probationary period (1 year), between 1 and 5 years of experience, between 6 and 10 years of experience and more than 10 years of experience. For each level of experience, I calculate the number of civil servant teachers in each school percentile that I divide by the total number of civil servant teachers with that level of experience. The Y-axis shows the distribution (density) of civil servant teachers among school percentiles according to their level of experience in teaching.

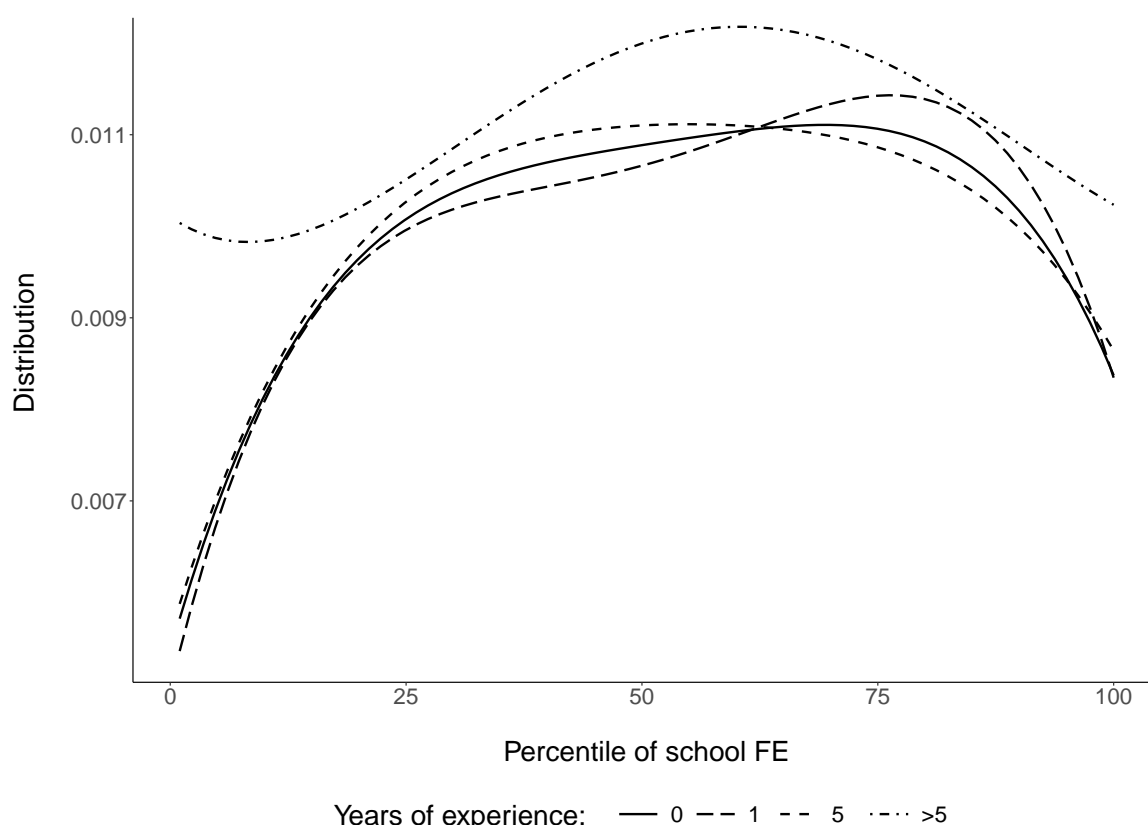


Figure 4.11 – Distribution of contract teachers according to experience and school fixed effects on teachers' absences (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on teachers' absence duration, estimated from model (3). I consider 5 levels of professional experience in teaching: probationary period (1 year), between 1 and 5 years of experience, between 6 and 10 years of experience and more than 10 years of experience. For each level of experience, I calculate the number of contract teachers in each school percentile that I divide by the total number of contract teachers with that level of experience. The Y-axis shows the distribution (density) of contract teachers among school percentiles according to their level of experience in teaching.

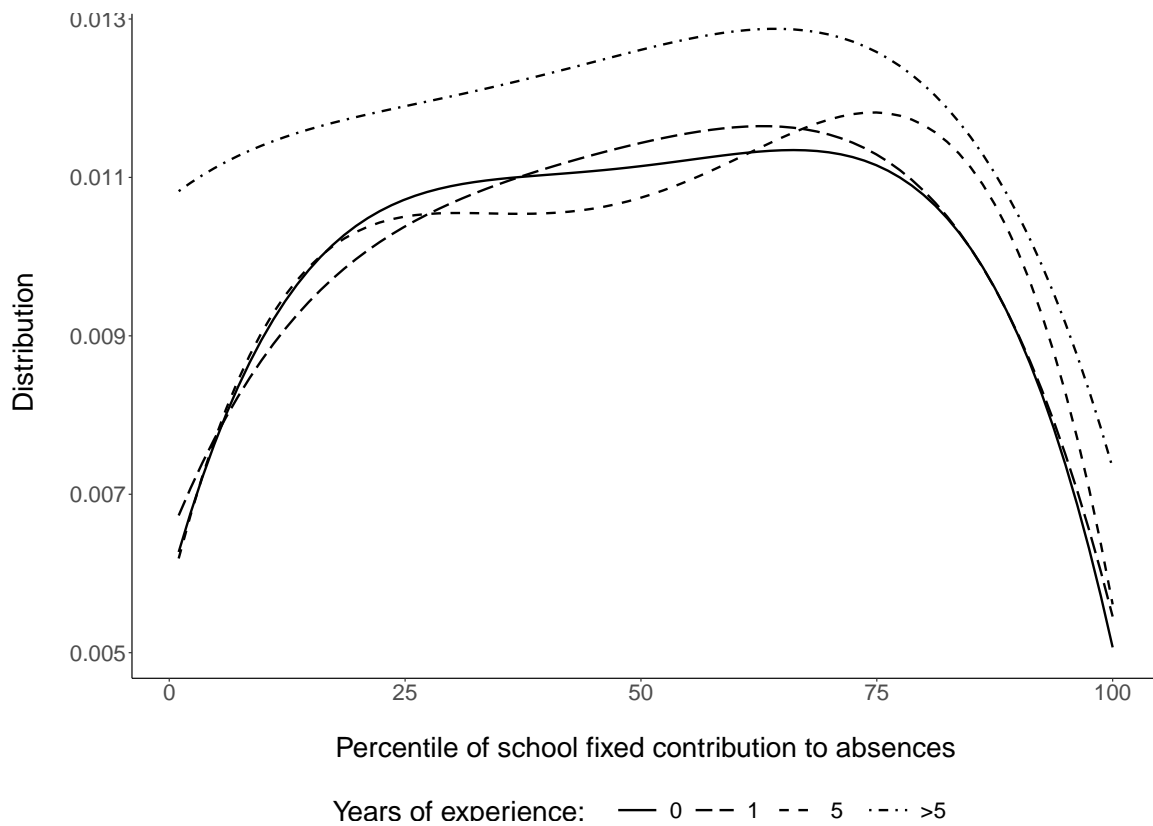


Figure 4.12 – Distribution of contract teachers according to experience and school fixed effects on teacher turnover (in percentiles)

Note: Schools are divided into percentiles according to their fixed effect on teacher turnover, estimated from model (i). I consider 5 levels of professional experience in teaching: probationary period (1 year), between 1 and 5 years of experience, between 6 and 10 years of experience and more than 10 years of experience. For each level of experience, I calculate the number of contract teachers in each school percentile that I divide by the total number of contract teachers with that level of experience. The Y-axis shows the distribution (density) of contract teachers among school percentiles according to their level of experience in teaching..

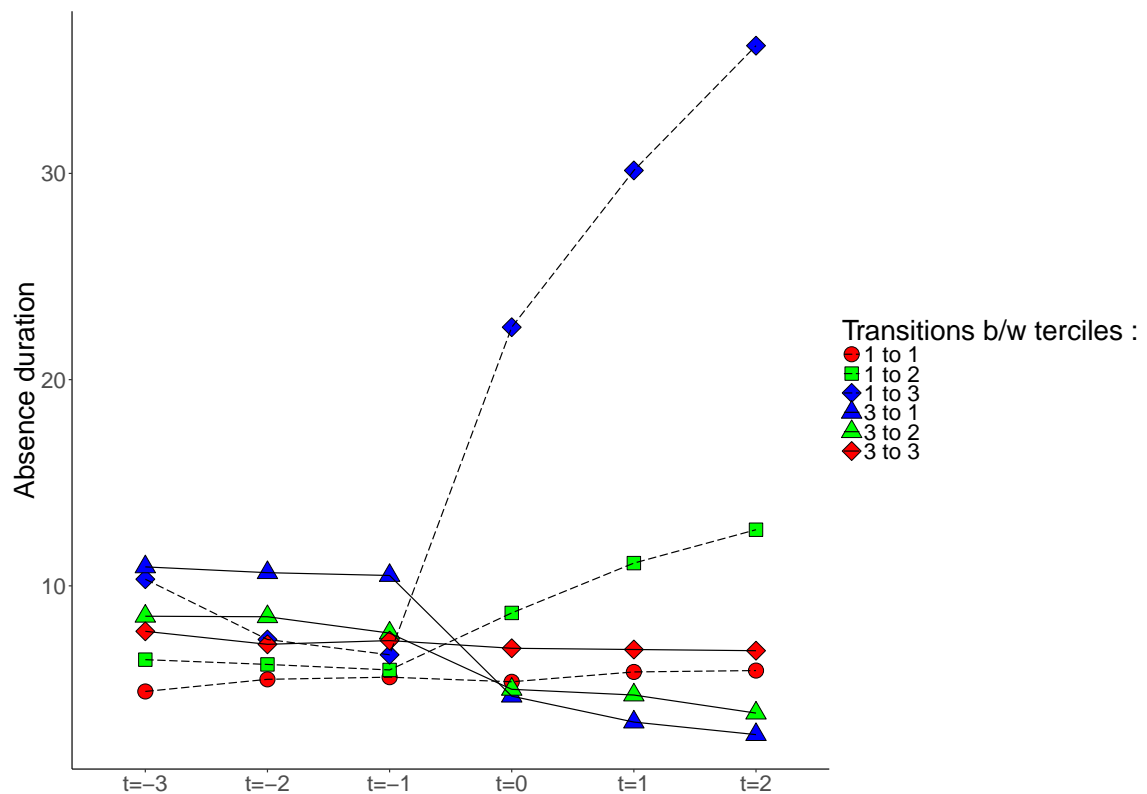


Figure 4.13 – Absence duration (in days) for teachers transiting between terciles of school fixed effects in year $t=0$ (model 4)

Note: Schools are divided into terciles of school fixed effects based on the estimates of model (4). Schools in the third tercile are expected to increase teacher absences the most significantly. I consider all teachers observed for six consecutive years ($t=-3, t=-2, t=-1, t=0, t=1, t=2$), who change schools in year $t=0$, but remain in the same tercile between $t=-3$ and $t=-1$ and between $t=0$ and $t=2$. This Figure shows how the absence duration (in days) varies for teachers leaving in $t=0$ a school from the first tercile (low fixed effects on absences: dotted lines) or from the last tercile (high fixed effects on absences: solid lines). For each of the nine transition groups between terciles of school fixed effects, Table 4.27 shows the number of observations (teachers) and the duration of teachers' absences between $t=-3$ and $t=2$.

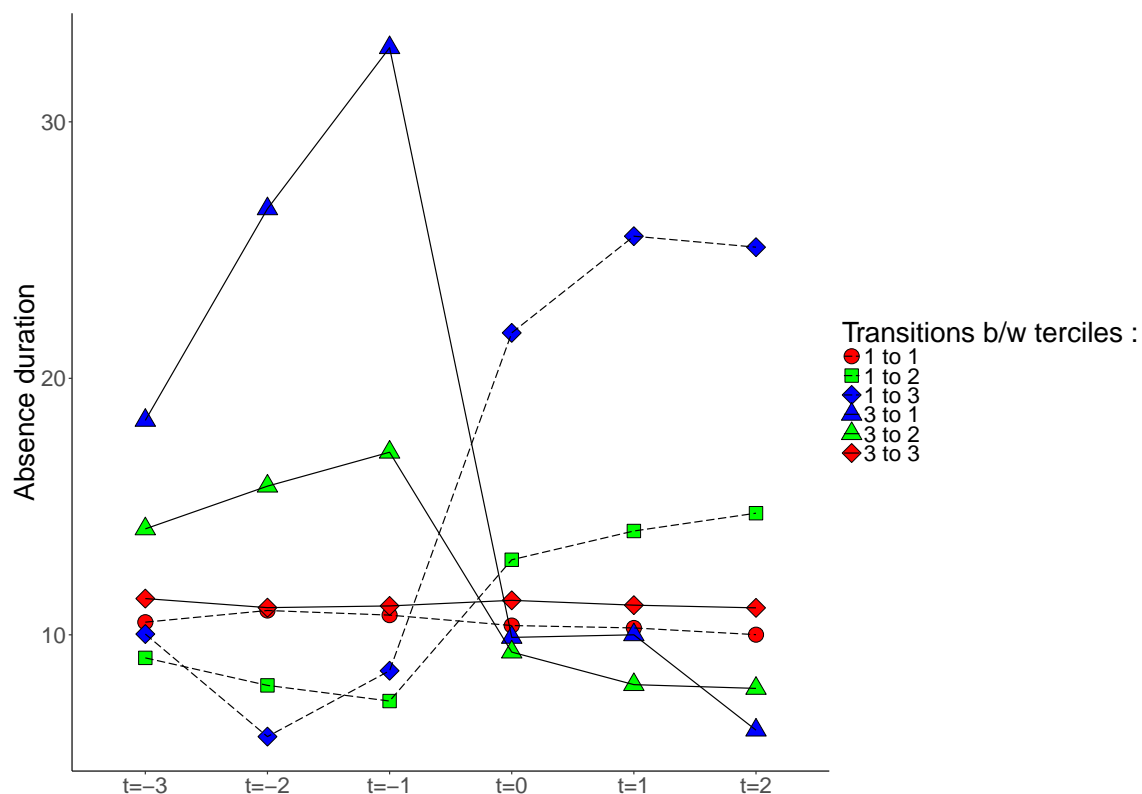


Figure 4.14 – Absence duration (in days) for schools transiting between terciles of school principal fixed effects in year $t=0$ (model 4)

Note: School principals are divided into terciles of principal fixed effects based on the estimates of model (4). School principals in the third tercile are expected to increase teacher absences the most significantly. I consider all schools observed for six consecutive years ($t=-3$, $t=-2$, $t=-1$, $t=0$, $t=1$, $t=2$), who are run by a new principal in year $t=0$, but remain in the same tercile of school principal between $t=-3$ and $t=-1$, and between $t=0$ and $t=2$. This Figure shows how the mean absence duration (in days) varies for schools "leaving" in $t=0$ a principal from the first tercile (low fixed effects on absences: dotted lines) or from the last tercile (high fixed effects on absences: solid lines). For each of the nine transition groups between terciles of principal fixed effects, Table 4.29 shows the number of observations (schools) and the mean duration of absences between $t=-3$ and $t=2$.

Conclusion: Policy implications, pending issues and future researches

At the heart of economic research is the analysis of factors favouring gender disparities between academic disciplines, which foster gender disparities between professions. Many studies have stressed the role of preferences, stereotypes, self-confidence, competition, peer effects and teacher model. Other studies have highlighted the existence of discrimination in hiring against the gender of the minority. The first chapter of this thesis shows that this is not the case in France with regard to the recruitment of secondary school teachers. This result contrasts with those obtained in other countries and calls for efforts to reduce gender disparities between subjects (combating discrimination, stereotypes and lack of self-confidence) much earlier in pupils' school careers.

The absence of discrimination against the minority gender in the recruitment of secondary school teachers is encouraging, but probably not sufficient to ensure the absence of discriminatory behaviors at other stages of the school and professional career, especially when students enter higher education. An ongoing collaboration aims to examine the existence of discrimination on the grounds of gender and social origin for students entering selective higher education institutions (schools and universities) in France.

The third chapter of this thesis shows that increasing the qualification level of teachers in France is not likely to help optimize recruitment. Following the 2011 reform that requires teachers to obtain a master's degree, the number of candidates for primary school teaching posts fell sharply and was not accompanied by an improvement in teachers' content and pedagogical knowledge. The shortage of teachers in some regions has led to the introduction of new competitive examinations, which has generated additional organizational costs.

Given this result, a first solution to attract more candidates would be to recruit teachers at the bachelor level. Chapter 2 suggests two other solutions that are not mutually

exclusive. First, increasing teacher salaries would make it possible to offset (at least partly) the wage gap with the private sector and would contribute to improve the attractiveness of the teaching profession. A salary bonus for teachers working in priority education areas from the 2018 school year onwards³⁶ and the increase in teachers' salaries planned from 2019 onwards, should allow to estimate more precisely the effect of remuneration on the number and quality of candidates. The second proposal is based on the observation that the current boundaries of academic regions are not based on the needs of teachers in these regions. Therefore, merging pools of candidates from regions with high teacher demand with those with lower teacher needs, through joint competition for example, could improve recruitment in less selective regions. This would also avoid the introduction of additional competitions to fill vacant positions. This solution can be seen as an intermediate situation between the regional recruitment of primary school teachers and the national recruitment of secondary school teachers in France.

The introduction of standardized tests at the beginning and end of primary education from the 2018 school year will make it possible to examine the effect of future reforms on pupil achievement. An important limitation of Chapter 3 is the inability to assess the effect of the 2011 teacher qualification reform on student performance. Nevertheless, the educational literature warns against the excessive use of these performance indicators. In particular, numerous experiences in the United States have shown that it is preferable to base the evaluation of teachers, and that of the education system as a whole, on several complementary criteria, including standardized tests but also alternative measures such as classroom observations.

The fourth and final chapter shows that school contexts and school principals play an important role in the absenteeism and transfer demands of secondary school teachers. The positive effect of hierarchical support and the negative influence of work intensity and hostile behavior on teacher absences, leading potentially to a deterioration in teachers' psychological well-being, question principals' training and call for an improvement of the working conditions in some schools. This chapter also confirms that newly recruited teachers are more likely to be assigned to schools where the learning conditions of their profession are less favorable. One possibility would be to give them more time, resources

³⁶From the start of the 2018 school year, teachers working in the reinforced priority education zones (REP+) will receive an additional bonus of 3,000 euros gross per year, which corresponds to a salary increase of 11% on average, and around 14% at the beginning of their careers.

and support when they work in a difficult school environment.

Future work will extend the results of Chapter 4. First, I would like to examine the link between the influence of school principals on teacher absences and the academic performance of secondary school students (at the *Baccalauréat* exam at the end of senior high school and *Brevet* exam at the end of junior high school). Such a study would examine whether principals' management choices are in the best interests of students. Second, I would like to study the heterogeneity of the effects of school context and managerial practices according to the gender of teachers. Consideration could also be given to evaluating some targeted training programs to improve psychosocial risk prevention and management practices of school heads through a randomized experiment. This would make it possible to assess the effect of prevention on teachers' health and absences, as well as on student achievement.

Finally, two ongoing collaborations should complete the fourth chapter. A first study examines the impact of non-compensation of the first day of sick leave on the absenteeism of educational staff³⁷ (teachers and non-teachers). This study assesses the impact of the reform on the frequency and duration of absences and focuses on the reaction of staff exposed to the most difficult working conditions. A second study examines the effect of exposure to psychosocial risk factors and physical constraints on health expenditure based on a representative sample of the French active population.

³⁷The absence of compensation for the first day of absence (called *jour de carence*) concerns all public employees in France. This policy was first introduced in 2012, then abolished in 2014, and then reinstated in 2018.