



**HAL**  
open science

## Overeducation and labour market trajectories

Florian Fouquet

► **To cite this version:**

Florian Fouquet. Overeducation and labour market trajectories. Economics and Finance. Nantes Université, 2023. English. NNT : 2023NANU3013 . tel-04373668

**HAL Id: tel-04373668**

**<https://theses.hal.science/tel-04373668v1>**

Submitted on 5 Jan 2024

**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.

# THESE DE DOCTORAT

NANTES UNIVERSITE

ECOLE DOCTORALE N° 640

*Sciences économiques et sciences de gestion - Pays de Loire*

Spécialité : Sciences Economiques

Par

**Florian FOUQUET**

**Sur-éducation et trajectoires sur le marché du travail**

**Thèse présentée et soutenue à Nantes, le 28 novembre 2023**

**Unité de recherche : Laboratoire d'Economie et de Management de Nantes-Atlantique (LEMNA)**

## **Rapporteurs avant soutenance :**

Julie LE GALLO            Professeur des universités, Institut Agro Dijon  
Ahmed TRITAH            Professeur des universités, Université de Poitiers

## **Composition du Jury :**

Président :            François-Charles WOLFF    Professeur des universités, Nantes Université  
Examineurs :        Arnaud DUPRAY            Ingénieur de recherche, Centre d'études et de recherches sur les qualifications  
                              Christine ERHEL            Professeur du CNAM, Conservatoire National des Arts et Métiers  
                              Camille HEMET            Professeur des universités, Université Paris 1 Panthéon-Sorbonne, PSE  
                              Julie LE GALLO            Professeur des universités, Institut Agro Dijon  
                              Ahmed TRITAH            Professeur des universités, Université de Poitiers  
                              François-Charles WOLFF    Professeur des universités, Nantes Université

Dir. de thèse :        Florent SARI                Professeur des universités, Université Paris-Est Créteil  
Co-dir. de thèse :    Bénédicte ROULAND        Maître de conférences, Nantes Université



---

SUR-ÉDUCATION ET TRAJECTOIRES SUR LE  
MARCHÉ DU TRAVAIL

---

Overeducation and labour market trajectories

---

Florian FOUQUET

Sous la direction de Florent Sari et Bénédicte Rouland

Thèse soutenue le 28 novembre 2023

Nantes Université

# Contents

|   |           |
|---|-----------|
| Remerciements . . . . .   | i         |
| Résumé . . . . .  | iv        |
| Abstract . . . . .  | vi        |
| List of tables . . . . .  | viii      |
| List of figures . . . . .   | ix        |
| <b>General introduction</b>   | <b>1</b>  |
| Educational choices and returns to education . . . . .  | 4         |
| Overeducation, matching frictions and mobility . . . . .  | 9         |
| Overeducation: incidence, consequences, drivers . . . . .   | 13        |
| Road map of the thesis . . . . .  | 16        |
| <b>1 Measuring overeducation using skills requirements</b>  | <b>23</b> |
| 1.1 Introduction . . . . .  | 25        |
| 1.2 Measuring overeducation . . . . .   | 27        |
| 1.3 Building a measure of mismatch based on skills requirements: the<br><i>Skill Analysis</i> measure . . . . . | 30        |
| 1.3.1 The European Qualifications Framework (EQF) . . . . .   | 30        |
| 1.3.2 The O*NET database . . . . .  | 31        |
| 1.3.3 Matching skills, degrees and occupations . . . . .  | 32        |
| 1.3.4 Advantages and drawbacks of the <i>Skill Analysis</i> measure . . . . .                                   | 37        |
| 1.4 Empirical comparison of overeducation measures . . . . .  | 38        |
| 1.4.1 Data and statistical measures . . . . .   | 38        |
| 1.4.2 Overeducation incidence across measures . . . . .   | 39        |
| 1.5 Conclusion . . . . .  | 44        |
| Appendix 1A: Description of EQF levels . . . . .  | 45        |
| Appendix 1B: Description of O*NET skills . . . . .  | 46        |

---

|          |   |           |
|----------|---|-----------|
| <b>2</b> | <b>Local labour markets and spatial determinants of overeducation</b>               | <b>47</b> |
| 2.1      | Introduction . . . . .  | 49        |
| 2.2      | Overeducation, local and spatial drivers . . . . .                                  | 51        |
| 2.2.1    | Literature review . . . . .   | 51        |
| 2.2.2    | Hypotheses . . . . .  | 53        |
| 2.3      | Data and variables . . . . .  | 54        |
| 2.3.1    | Data . . . . .  | 54        |
| 2.3.2    | Variables . . . . .   | 55        |
| 2.3.3    | Descriptive statistics . . . . .  | 58        |
| 2.4      | Empirical strategy . . . . .  | 59        |
| 2.4.1    | Spatial interactions . . . . .  | 59        |
| 2.4.2    | Spatial heterogeneity . . . . .   | 64        |
| 2.4.3    | Endogeneity issue . . . . .   | 65        |
| 2.5      | Results . . . . .   | 65        |
| 2.5.1    | Effects of the local context . . . . .  | 65        |
| 2.5.2    | Heterogeneity between urban and rural areas . . . . .                               | 68        |
| 2.5.3    | Heterogeneity by gender . . . . .   | 68        |
| 2.6      | Conclusion . . . . .  | 72        |
|          | Appendix 2A: Cartography of French employment areas . . . . .                       | 73        |
|          | Appendix 2B: Direct, indirect and total effects in the SAR model . . . . .          | 74        |
|          | Appendix 2C: Estimations results with an SDM model . . . . .                        | 75        |
|          | Appendix 2D: First-stage regression for 2SLS and GS2SLS . . . . .                   | 76        |
|          | Appendix 2E: Spatial regimes SAR model - Estimations results by<br>gender . . . . . | 77        |
| <b>3</b> | <b>Spatial mobility and overeducation of young workers</b>                          | <b>79</b> |
| 3.1      | Introduction . . . . .  | 81        |
| 3.2      | Spatial mobility and overeducation . . . . .  | 84        |
| 3.2.1    | Theoretical links . . . . .   | 84        |
| 3.2.2    | Literature review . . . . .   | 86        |
| 3.3      | Data and variables . . . . .  | 89        |
| 3.3.1    | Data . . . . .  | 89        |
| 3.3.2    | Variables . . . . .   | 91        |
| 3.4      | Descriptive statistics . . . . .  | 94        |
| 3.5      | Estimation strategy . . . . .   | 96        |
| 3.6      | Results . . . . .   | 100       |

|          |   |            |
|----------|---|------------|
| 3.6.1    | Effects of migration on overeducation . . . . .                       | 100        |
| 3.6.2    | Differentiation by migration and educational level . . . . .          | 105        |
| 3.6.3    | Robustness checks . . . . .   | 108        |
| 3.7      | Conclusion . . . . .  | 112        |
|          | Appendix 3A: Statistical overeducation measure . . . . .              | 114        |
|          | Appendix 3B: Geography of French regions . . . . .                    | 115        |
|          | Appendix 3C: Distribution of predicted mismatch within the sample .   | 116        |
|          | Appendix 3D: Selection equation and first stage regression for 2SLS . | 117        |
|          | Appendix 3E: Alternative estimations and robustness checks . . . . .  | 118        |
| <b>4</b> | <b>Educational mismatch and early career trajectories</b>             | <b>123</b> |
| 4.1      | Introduction . . . . .  | 125        |
| 4.2      | Data and variables . . . . .  | 129        |
| 4.2.1    | Sample . . . . .  | 129        |
| 4.2.2    | Mismatches . . . . .  | 130        |
| 4.2.3    | Descriptive statistics . . . . .                                      | 131        |
| 4.3      | Empirical strategy . . . . .  | 133        |
| 4.3.1    | Treatment framework . . . . .   | 133        |
| 4.3.2    | Conditional independence assumption (CIA) . . . . .                   | 135        |
| 4.3.3    | Matching specification and estimation . . . . .                       | 136        |
| 4.4      | Results . . . . .   | 137        |
| 4.4.1    | Effects on employment and job characteristics . . . . .               | 137        |
| 4.4.2    | Effects on earnings . . . . .   | 138        |
| 4.4.3    | Effects on matching status . . . . .                                  | 139        |
| 4.4.4    | Heterogeneity analysis . . . . .                                      | 142        |
| 4.5      | Conclusion . . . . .  | 144        |
|          | Appendix 4A: Statistical measure of overeducation . . . . .           | 147        |
|          | Appendix 4B: Quality of the matching . . . . .                        | 149        |
|          | Appendix 4C: Robustness checks . . . . .                              | 158        |
|          | Appendix 4D: Additional descriptive statistics . . . . .              | 159        |
|          | <b>Conclusion</b>   | <b>161</b> |
|          | <b>Bibliography</b>   | <b>163</b> |





## Remerciements

Avant de vous laisser à la lecture de cette thèse, je me devais d'adresser quelques remerciements. Ils seront probablement beaucoup trop longs et trop sentimentaux, mais il me fallait prendre le temps de remercier chacune des personnes qui auront contribué, de près ou de loin, pour quelques instants ou durant tout ce parcours, par leurs conseils, leurs mots ou leur simple présence, à la réalisation de cette thèse. Par ailleurs, n'y voyez là aucun ordre de préférence, ils n'ont de toute façon pas été écrits dans celui dans lequel ils apparaissent.

Je tiens en premier lieu à remercier Florent Sari et Bénédicte Rouland, qui ont dirigé cette thèse et m'ont encadré et motivé durant ces années de doctorat. Leur soutien et leurs conseils tout au long de ces quatre années auront été précieux pour m'épanouir dans ce travail de thèse. Je suis très chanceux et heureux de vous avoir eu comme directeur et directrice de thèse, ma thèse aurait été très différente si elle ne s'était pas faite avec vous.

J'adresse toute ma reconnaissance à Julie Le Gallo et Ahmed Tritah d'avoir accepté de rapporter cette thèse, ainsi qu'à Arnaud Dupray, Christine Erhel, Camille Hémet et François-Charles Wolff d'avoir accepté d'être membres du jury.

Je suis reconnaissant au LEMNA et à l'IAE de Nantes pour l'environnement de travail dont j'ai pu bénéficier durant la réalisation de cette thèse. J'adresse mes remerciements à toutes les personnes rencontrées dans ce cadre, et en particulier à Alain Martineau, Lionel Richefort et Pierre-Alexandre Mahieu, pour les TD qu'ils m'ont confié et qui ont confirmé mon intérêt déjà bien établi pour l'enseignement; à Camille Aït-Youcef, pour la confiance accordée dans l'organisation du séminaire des doctorants; à Muriel Travers, pour ces quelques mots au début de ma thèse, qui m'auront rassuré sur le fait d'y être à ma place. Merci également aux équipes administratives dont l'efficacité et le soutien m'auront permis de garder l'esprit libéré quand il s'agissait de programmer un déplacement, d'organiser un séminaire, ou simplement de gérer les démarches qui vont avec le statut de doctorant.

I would also like to express my gratitude to Gail and her team for welcoming me to the NZWRI. I feel very lucky to have been given such a nice environment to complete my thesis and discover new areas of research, and to meet all of you. Huge thanks to Lisa, it was a great pleasure working with you, and I hope we will have the opportunity to work together again soon.

Je dois aussi remercier mes camarades de thèse (et quelques personnes dont la thèse est terminée depuis longtemps ou bien commence à peine), doctorantes et doctorants du LEMNA, avec qui j'ai partagé de très bons moments, au bureau comme en dehors.

Pauline, ma colocataire du bureau 215 pendant trois ans, le quotidien de thèse aurait été bien moins agréable sans ta présence. Tes conseils sur la vie en doctorat et la recherche, nos sessions musique, nos débats entre approches théorique et empirique, les soirées jeux de société, les très (parfois trop) longues discussions sur des sujets plus ou moins existentiels et plus ou moins sérieux, et ton soutien sans faille auront été très précieux durant ces quatre années. Je suis heureux d'avoir commencé cette thèse auprès d'une collègue et de la terminer auprès d'une amie.

Clément, ma thèse doit beaucoup à nos pauses café, petits moments de détente au milieu de tout ce sérieux, à nos discussions sur l'économie du travail et l'économétrie, nos études de la *Brick Economy*, ta capacité à débayer mes codes R, et notre petit R5-B12 qui deviendra une relique pour les futures générations de doctorantes et doctorants du LEMNA. Ta bonne humeur communicative a par ailleurs grandement contribué à ma sérénité d'esprit durant cette thèse, et je t'en remercie.

Léo, des préparations de TD de microéconomie aux murs d'escalade, je suis très heureux de t'avoir eu comme partenaire. Pascaline, au plaisir de discuter de nouveau de nos étudiants autour d'un café. Lucie, Raphaël, mes excursions au bureau 218 doivent beaucoup à votre présence.

Hounayda, nos parcours de thèse en parallèle, on aura partagé un certain nombre de joies et de peurs, d'angoisses et de soulagements, notamment à l'approche de la fin. A bientôt pour fêter nos soutenances respectives !

Un mot aussi pour celles et ceux que j'ai certes moins côtoyés, mais qui ont fait partie à un moment ou à un autre de ce parcours : Étienne, Nadim, Jean-Baptiste, Antoine, Victor, Coline, Nada, Xavier-Emmanuel, Guillaume, Emmanuel, Meriem. Pardon à ceux que j'oublie.

Emma, de la tutrice qui n'ose pas me tutoyer à la colocataire d'open-space, on en aura fait du chemin ! Tes questions sur l'économétrie et ta bienveillance auront été plus que bienvenues durant ces derniers mois. C'est désormais à ton tour d'embarquer dans l'aventure de la thèse, tu sais où me trouver si tu en as besoin.

Il me faut aussi adresser toute ma gratitude à un certain nombre de personnes, rencontrées durant mes années de doctorat ou présentes depuis bien plus longtemps, avec lesquelles je suis heureux d'avoir partagé de jolis moments, bien que trop peu

nombreux et trop rares, qui m'auront fait sortir un peu la tête de la thèse, et rappelé qu'autre chose existait en dehors. À bientôt pour les prochains !

Théo, Sophia, Joffrey, nos soirées jeux de société auront été l'occasion de très bons moments et de souvenirs mémorables. Un merci particulier à Bertille, tu es la seule personne à rire à mes blagues et non de mes blagues, et ça c'est beau !

Anne, nos restos et nos chocolats chauds au Bolton sont de ce genre de petits moments parfois trop rares mais si importants, je suis heureux de les avoir partagés avec toi. Louis, nos guitares auront eu du mal à se croiser durant ma thèse, j'espère pouvoir remédier à ça prochainement. Amandine, nos balades ont toujours eu tendance à terminer à la nuit tombée dans les rues du Vieux Mans, merci pour ces moments de calme et d'écoute. Océane, nos fins d'après-midi à l'Atomic's me manquent déjà ! Elora, nos parties de billard auront été des bouffées d'air plus que nécessaires durant ces quatre années.

Marie-Océane, ma binôme durant toute la Licence et le Master, je n'en serai sûrement jamais arrivé jusqu'ici sans ta présence à mes côtés durant toutes ces années. De nos longues séances de révision à la BU (avec sucreries et partage d'écouteurs, *Three Days Grace* et cafés vanille) à nos discussions jusqu'à très tard, de nos prises de tête à nos fous rires, on en aura partagé des choses. Merci pour tout ça, mais aussi et surtout pour tout le reste. Je suis honoré de ton amitié si précieuse depuis plus de dix ans, et je sais qu'elle durera encore longtemps.

J'adresse aussi mes remerciements à ma famille, qui m'a soutenu tout au long de ce parcours, et en particulier à mes soeurs Doryane, Lila et Maëlia.

Enfin, il me faut adresser un immense merci à ma Maman, pour m'avoir transmis son goût pour la culture, le savoir et la connaissance, pour ses relectures de mes papiers (parfois en anglais, parfois remplis d'aspects méthodologiques obscurs, souvent les deux à la fois) et les commentaires, remarques, questions et discussions qui ont pu en découler, et par-dessus tout pour son soutien indéfectible quels que soient mes choix.

## Résumé

Certaines personnes ont un niveau d'éducation supérieur à celui requis pour leur emploi : elles sont sur-éduqués. La sur-éducation touche entre un quart et un tiers des travailleurs dans les pays de l'OCDE, avec des conséquences importantes tant au niveau individuel que collectif. L'objectif de cette thèse est d'explorer les sources et les conséquences de la sur-éducation, notamment en termes de mobilité, qu'elle soit géographique ou professionnelle, afin de tenter de mettre en évidence des leviers pour la réduire ou la limiter. Plus précisément, elle étudie comment la géographie du marché du travail affecte la sur-éducation et le rôle de la mobilité spatiale individuelle, ainsi que l'impact de l'inadéquation emploi-diplôme sur les trajectoires professionnelles futures.

Le premier chapitre propose une mesure nouvelle et actualisée de la sur-éducation, basée sur le contenu des emplois en termes de tâches et les compétences requises pour exercer ces emplois. Elle vise à résoudre les problèmes soulevés par d'autres approches, notamment en ce qui concerne l'évolution du marché du travail et le détail des nomenclatures des professions. En outre, la comparaison empirique de cette nouvelle mesure avec les mesures existantes confirme que les différents types d'approches n'appréhendent pas tout à fait le même phénomène.

Dans le deuxième chapitre, j'utilise des modèles d'économétrie spatiale et de régimes spatiaux pour étudier la dimension géographique de la suréducation et l'influence du contexte du marché du travail local. Je montre que la structure des marchés du travail locaux influe sur la suréducation, mais qu'il existe une forte hétérogénéité de ces déterminants entre zones urbaines et rurales. Je constate également que les femmes sont plus touchées par ces facteurs que les hommes, ce qui tend à confirmer l'idée d'une suréducation différentielle.

En relation avec le chapitre précédent, le troisième chapitre étudie si la mobilité spatiale individuelle (migration résidentielle) permet ou non de réduire le risque de sur-éducation. Les résultats révèlent que la migration interrégionale diminue la probabilité d'être sur-éduqué, avec un effet plus important de la migration vers un grand centre économique (la région parisienne). Une analyse de l'hétérogénéité en termes de niveau d'éducation montre que les travailleurs les plus éduqués bénéficient davantage de la mobilité spatiale.

Le quatrième chapitre s'intéresse à l'inadéquation horizontale, c'est-à-dire des individus travaillant dans un domaine ne correspondant pas à leur diplôme, en plus de la sur-éducation, et cherche à déterminer si le fait d'accepter un emploi inadéquat

en début de carrière permet ou non d'accéder à de meilleures positions sur le marché du travail par la suite. Au lieu d'un effet de tremplin (*stepping stone*), je constate que le *mismatch* dans le premier emploi conduit à rester "piégé" dans des postes similaires au moins en début de carrière, avec des bénéfices limités en termes de protection contre le chômage.

## Abstract

Some workers have an educational level higher than the one required for their job: they are overeducated. Overeducation is a widespread phenomenon, which concerns between a quarter and a third of the workers in OECD countries, leading to detrimental effects both at the individual and collective levels. The purpose of this thesis is to explore the sources and consequences of overeducation, especially in terms of mobility, whether geographical or professional, aiming to find levers to be used to reduce or limit it. More precisely, it investigates how the labour market geography affects overeducation and the role of individual spatial mobility, as well as the impact of educational mismatch on future professional trajectories.

The first chapter proposes a new and up-to-date measure of overeducation, based on the content of jobs in terms of tasks and skills required to perform a job. It aims to address the issues that arise with other approaches, especially related to the evolution of the labour market and the detail of occupational classification. Besides, the empirical comparison of this new measure to existing ones confirms that different types of measures do not exactly capture the same matter.

In the second chapter, I use spatial econometrics and spatial regimes models to study the geographical dimension of overeducation and the influence of the local labour market context. I find that the structure of local labour markets affects overeducation, but with a strong heterogeneity in the drivers of mismatch between urban and rural areas. I also find that women are more affected by those drivers than men, supporting the idea of a differential overeducation.

In relation to the previous one, the third chapter investigates whether individual spatial mobility (residential migration) reduces the risk of being overeducated or not. Results reveal that interregional migration decreases the risk of overeducation, with a stronger effect of migration to a large economic centre (the Paris region). An heterogeneity analysis in terms of educational level show that more educated workers benefit more from spatial mobility.

The fourth chapter extends its focus to horizontal mismatch, *i.e.* the situation of individuals working in a field that does not match their education, in addition to overeducation, and explores whether accepting a mismatched job at the beginning of the career allow to access to better labour market positions thereafter or not. Instead of a “stepping stone” effect, I find that workers mismatched in their first job stay trapped in similar positions at least in their early career, with only limited benefits in terms of protection against unemployment.

# List of Tables

|      |   |     |
|------|---|-----|
| 1.1  | Learning outcomes expressed in skill scores by EQF level . . . . .                              | 34  |
| 1.2  | Overeducation measures . . . . .  | 39  |
| 1.3  | Overeducation rates by measure . . . . .  | 40  |
| 1.4  | Overeducation rates by measure - 15-29 y-o . . . . .  | 41  |
| 1.5  | Correspondence between <i>Skill Analysis</i> and statistical measures - 15-<br>29 y-o . . . . . | 42  |
| 1.6  | Correlation between overeducation measures - 15-29 y-o . . . . .                                | 43  |
| 1.7  | EQF levels with corresponding degrees and learning outcomes . . . . .                           | 45  |
| 1.8  | O*NET skills . . . . .  | 46  |
| 2.1  | Description of the variables . . . . .  | 55  |
| 2.2  | Sample characteristics . . . . .  | 59  |
| 2.3  | Spatial autocorrelation of overeducation rates (Moran's <i>I</i> ) . . . . .                    | 62  |
| 2.4  | Specification tests - LM-tests for inclusion of spatial lags . . . . .                          | 63  |
| 2.5  | Estimation results . . . . .  | 66  |
| 2.6  | Direct, indirect and total effects - SAR model . . . . .  | 67  |
| 2.7  | Estimation results - Spatial regimes models . . . . .   | 69  |
| 2.8  | Direct, indirect and total effects - Spatial regimes SAR model . . . . .                        | 70  |
| 2.9  | Direct, indirect and total effects - Spatial regimes SAR model - by<br>gender . . . . .         | 71  |
| 2.10 | Estimation results - SDM model . . . . .  | 75  |
| 2.11 | First stage regression results . . . . .  | 76  |
| 2.12 | Estimation results by gender - SAR spatial regimes models . . . . .                             | 77  |
| 3.1  | Descriptive statistics . . . . .  | 95  |
| 3.2  | Correspondence between statistical overeducation and subjective mis-<br>match . . . . .         | 96  |
| 3.3  | Linear probability models for statistical overeducation . . . . .                               | 102 |

---

|      |   |     |
|------|---|-----|
| 3.4  | Linear probability models for subjective mismatch . . . . .   | 104 |
| 3.5  | Linear probability models - Effects of migration . . . . .  | 106 |
| 3.6  | Linear probability models - Effects of migration by educational level .                             | 107 |
| 3.7  | Linear probability models - Effects of migration based on (log) distance (measured in km) . . . . . | 109 |
| 3.8  | Correspondence of educational levels between classifications . . . . .                              | 114 |
| 3.9  | Correspondence of occupational levels between classifications . . . . .                             | 114 |
| 3.10 | Occupation-degree correspondence grid - Statistical overeducation . .                               | 114 |
| 3.11 | Employment equation and IV estimation (first stage) . . . . .                                       | 117 |
| 3.12 | Linear probability models for statistical overeducation - Alternative measure . . . . .             | 118 |
| 3.13 | Linear probability models including wage (mismatch equations) . . .                                 | 119 |
| 3.14 | Triprobit model for employment, migration and overeducation . . . .                                 | 120 |
| 3.15 | Triprobit model for employment, migration and subjective mismatch                                   | 121 |
|      |   |     |
| 4.1  | Descriptive statistics - By mismatch status in the first job . . . . .                              | 132 |
| 4.2  | Sample size by treatment status and time of treatment . . . . .                                     | 134 |
| 4.3  | Treatment effects on labour market outcomes . . . . .   | 138 |
| 4.4  | Treatment effects on matching status . . . . .  | 140 |
| 4.5  | Treatment effects on labour market outcomes - Any mismatch . . . .                                  | 142 |
| 4.6  | Treatment effects on matching status - Any mismatch . . . . .                                       | 143 |
| 4.7  | Description of occupational codes - PCS 2020 . . . . .  | 147 |
| 4.8  | Occupation-degree correspondence grid - Statistical overeducation . .                               | 148 |
| 4.9  | Treatment effects on labour market outcomes - Any mismatch . . . .                                  | 158 |
| 4.10 | Treatment effects on matching status - Any mismatch . . . . .                                       | 158 |
| 4.11 | Descriptive statistics - Outcomes . . . . .   | 159 |



# List of Figures

|     |   |     |
|-----|---|-----|
| 2.1 | Overeducation rates by French employment area - 2009 . . . . .                                | 60  |
| 2.2 | Overeducation rates by French employment area - 2013 . . . . .                                | 61  |
| 2.3 | Overeducation rates by French employment area - 2017 . . . . .                                | 61  |
| 2.4 | French employment areas - 2010 zoning . . . . .   | 73  |
| 3.1 | French regions - NUTS2 areas . . . . .  | 115 |
| 3.2 | Predicted probabilities - Statistical overeducation . . . . .                                 | 116 |
| 3.3 | Predicted probabilities - Subjective mismatch . . . . .                                       | 116 |
| 4.1 | Distribution of the propensity score on the matched and unmatched<br>sample - $T_0$ . . . . . | 149 |
| 4.2 | Distribution of the propensity score on the matched and unmatched<br>sample - $T_1$ . . . . . | 150 |
| 4.3 | Distribution of the propensity score on the matched and unmatched<br>sample - $T_2$ . . . . . | 151 |
| 4.4 | Covariate balance on the matched and unmatched sample - $T_0$ . . . .                         | 152 |
| 4.4 | Covariate balance on the matched and unmatched sample - $T_0$ (con-<br>tinued) . . . . .      | 153 |
| 4.5 | Covariate balance on the matched and unmatched sample - $T_1$ . . . .                         | 154 |
| 4.5 | Covariate balance on the matched and unmatched sample - $T_1$ (con-<br>tinued) . . . . .      | 155 |
| 4.6 | Covariate balance on the matched and unmatched sample - $T_2$ . . . .                         | 156 |
| 4.6 | Covariate balance on the matched and unmatched sample - $T_2$ (con-<br>tinued) . . . . .      | 157 |



# General introduction

*« À la différence de la situation qui prévalait il y a peu encore, la première mission des universités et des grandes écoles n'est plus le recrutement des cadres de l'État, qui n'est plus au centre de la vie économique et industrielle, mais de servir les étudiants, de donner à chacun d'eux, quel que soit son milieu d'origine, toutes les chances de trouver son domaine d'excellence, de se préparer aux métiers d'après-demain et de faire progresser le savoir.*

*Tout étudiant devra être assuré de pouvoir quitter l'enseignement supérieur avec un diplôme à valeur professionnelle, s'il est prêt à accomplir les efforts nécessaires pour en obtenir un. »*

Rapport Attali (1998)<sup>1</sup>

*“Qualifications serve a variety of purposes. They signal to employers what their holders in principle know and are able to do (‘learning outcomes’). They may be a prerequisite for accessing certain regulated professions. They help education and training authorities and providers to determine the level and content of learning acquired by an individual. They are also important for an individual as an expression of personal achievement. Therefore qualifications play an important role in raising employability, easing mobility and access to further education.”*

European Commission (2017)<sup>2</sup>

---

<sup>1</sup>« Pour un modèle européen d'enseignement supérieur », Rapport à M. le Ministre de l'Éducation Nationale, de la Recherche et de la Technologie, rapport de la Commission présidée par Jacques Attali, p. 4.

<sup>2</sup>Council Recommendation of 22 May 2017 on the European Qualifications Framework for life-long learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning, Article 1.

Education has always been central to human life. As a “process of imparting or acquiring knowledge, developing physical, intellectual and moral qualities, developing the powers of reasoning and judgment, and generally of preparing oneself or others intellectually for mature life; the knowledge or training acquired by this process; and the resources deployed during this process”<sup>3</sup>, education is indeed at the heart of the children’s transition to adulthood, and more generally of the construction of the human being as an individual. However, in our modern societies, this definition does not cover all aspects of education and needs to be refined. Coombs and Manzoor (1974), echoed by La Belle (1982), suggested a typology of education in three types, namely formal, nonformal and informal education, which is still used nowadays by a number of institutions, such as the Council of Europe, the European Commission or the United Nations Educational, Scientific and Cultural Organization (UNESCO). Nonformal and informal education are defined as the product of any activities providing knowledge or skills experienced outside a specific framework, the main difference being that nonformal education comes from a deliberate will to learn while informal education simply results of everyday life experiences. On the contrary, formal education corresponds to the institutionalised and structured educational system, often provided by states and sanctioned by degrees or diplomas. In this thesis, education will be understood in its formal dimension.

In addition of being a fundamental human right (reaffirmed in Article 26 of the Universal Declaration of Human Rights in 1948), education is often described in our modern countries as a condition for economic prosperity and peace, a central element of the “knowledge economy” and an important contribution to economic growth (Hanushek & Woessmann, 2020), but also as a means of social emancipation and individual fulfillment. The two quotes in the preamble to this introduction illustrate this idea. The first one is drawn from the introduction of the government report which, through an important reform a few years later, led to the higher education system as we know it today in France. The second is Article 1 of the European Commission Recommendation which defines a common framework for educational systems in Europe, as part of the Bologna Process<sup>4</sup>. Two ideas stand out clearly from these quotes. Firstly, education should enable all individuals to reach personal

---

<sup>3</sup>This definition synthesises the definitions of the French dictionaries *Trésor de la Langue Française informatisé* (TLFi) and *Dictionnaire de l’Académie française, 9ème édition* (translated by the author), and the English dictionary *Collins English Dictionary* (HarperCollins Publishers).

<sup>4</sup>Launched in 1999, the Bologna Process is a set of meetings and agreements between European countries, aiming to “bring more coherence to higher education systems across Europe”. See <https://education.ec.europa.eu/> and <https://www.ehea.info/> for more information.

---

achievement and to find their place in the society. Secondly, education should be valued on the labour market, enabling everyone to make the most of their skills and knowledge. Besides, these two objectives are intrinsically related, one contributing to the other.

Following this idea, many advanced countries implemented policies promoting a broader access to education since the 1970s: the *Elementary and Secondary Education Act* in the US as early as 1965, the creation of the professional baccalaureate in France in 1985, the introduction of public allowances to cover tuition fees since the end of the 1980s in the Netherlands or in the UK (among other countries), the implementation of the *Bachelor's-Master's-Doctorate* (BMD) system in many European countries in the 2000s, to cite only a few examples. This resulted in an important educational expansion over the past decades, which continues to this day: in 2021, 80% of the 25-64 year-olds in OECD countries have at least upper secondary education (OECD, 2022). The increase is particularly striking for higher education: on average in OECD countries, 22% of the 25-64 year-olds had tertiary education in 2000, they were 40% in 2021<sup>5</sup>. This share is even higher for the 25-34 year-olds, rising from 27% in 2000 to 48% in 2021 (OECD, 2022). These figures suggest that the goal of personal educational achievement is (at least on the way to being) fulfilled.

The value of education on the labour market, however, is still being challenged. Indeed, the increase in educational attainment coincides in time with the emergence of overeducation (Affichard, 1981; Duncan & Hoffman, 1981; Freeman, 1976; Rumberger, 1981, to cite some of the earliest works on this subject). Generally speaking, overeducation is defined as the situation of individuals that have an educational level higher than the one required for the job that they occupy (Kucel, 2011; McGuinness, 2006). That implies that some individuals are unable to make (full) use of their knowledge, skills and/or educational credentials on the labour market. The incidence of the phenomenon varies between countries, periods and measurements, but it is generally said to affect one quarter to one third of the labour force in OECD countries (Quintini, 2011b). Overeducation is therefore far from marginal. Moreover, the literature has shown its strong negative effects at the individual level, at the firm level and on the economy as a whole (McGuinness, 2006), calling for a better understanding of the phenomenon and the levers to be used to reduce or

---

<sup>5</sup>OECD (2023), Adult education level (indicator). doi: [10.1787/36bce3fe-en](https://doi.org/10.1787/36bce3fe-en) (Accessed on 08 June 2023).

limit it.

The purpose of this thesis is to explore the sources and consequences of overeducation, especially in terms of mobility, whether geographical or professional. More precisely, it investigates how the labour market geography affects overeducation and the role of individual spatial mobility, as well as the impact of educational mismatch on future professional trajectories. The remainder of this introduction will expose some relevant background information, which will set out the framework of this thesis. First, I will outline theories about educational choices and why people engage in education, discuss the existing literature about the returns to education, and show how overeducation can be related to spatial and professional mobility. Then, I will review previous empirical evidence about the incidence, drivers and consequences of overeducation.

## Educational choices and returns to education

### Why do people engage in education?

In labour economics, two dominant frameworks aim to explain the decision to engage in education: the Human Capital Theory (Becker, 1964) and the signaling theory (Spence, 1973). Although these are rival frameworks, they have in common that the main reason individuals participate in (higher) education is that it should improve their future labor market outcomes.

In Human Capital Theory (Becker, 1964), individuals' productivity depends on the skills, knowledge and abilities that they accumulate. This set of characteristics is what is called "human capital". In this framework, education (alongside with on-the-job training, lifelong learning, etc.) is a way to acquire those skills and knowledge, in order to enhance work productivity. Higher human capital is supposed to be associated with higher earnings (Lemieux, 2006; Mincer, 1974) and better employment (Bishop, 1994). Indeed, individuals decide to study only if the benefits exceed the costs: the returns to education should be positive.

On the contrary, in the signaling theory (Spence, 1973), future productivity is not affected by schooling, but is an intrinsic characteristic of workers. Education is not a way to accumulate skills or knowledge, but to provide employers information about the innate ability of individuals. Educational credentials can therefore be used by employers as a "filter" to select the most capable workers (Arrow, 1973).

Building on this, Thurow (1975) describes education not only as a signaling tool, but as a sorting criterion. In his Job Competition Model, individuals wait for jobs in a queue, in which the first ones are allocated the best jobs (in the original model, the jobs with higher earnings). Education is thus a way to acquire and maintain a better position in the queue. Both signaling and sorting are based on the underlying assumption that future productivity is negatively correlated to the cost of the signal: education allows individuals with higher ability to differentiate themselves from low ability workers because they support cheaper educational costs.

Both Human Capital Theory and signaling theory describe education as an investment, which depends on individuals' innate ability, expected earnings (Lemieux, 2006; Mincer, 1974) and direct monetary costs (Fuller et al., 1982), but is also affected by workers' own beliefs (Koch et al., 2015), especially in terms of self-perceived academic capability (Chevalier et al., 2009). However, there can be other motives to engage in education than investment: individuals' choice to study can be related to consumption (Alstadsæter, 2011; Lazear, 1977) or social norms (Coleman, 1961) motives.

Some individuals decide to study because they are interested in the subject of the courses, they enjoy learning or they are attracted by student life: they consume education rather than investing in it (Gullason, 1989; Lazear, 1977). These individuals have a positive willingness to pay for non-pecuniary and non work-related (and often more immediate<sup>6</sup>) returns to education (Alstadsæter et al., 2008): they accept lower future earnings to acquire a specific type of education. This can be a particular major, a particular type of college, a particular location or specific amenities of the school (Alstadsæter, 2011). Especially, Sellami et al. (2020) showed that individuals who engage in higher education for consumption motives choose different fields of study from those of people motivated by investment. In addition, Alstadsæter et al. (2008) found that higher income taxation and lower tuition fees tend to promote consumption over investment, as they reduce the returns to education in terms of earnings.

Educational choices can also be influenced by social norms (Coleman, 1961). Indeed, there are utility gains associated with conforming to the expectations and behaviours of the social group to which one belongs (Akerlof & Kranton, 2000, 2002). This can translate in children graduating in the same field as their parents, students

---

<sup>6</sup>See in particular Gullason (1989) on how young American men decided to continue schooling to avoid being drafted for the Vietnam War in the 1960s.

choosing the same major as their friends, etc., in order to “fit in”. Deviating from these choices can lead to stigmatisation and (social) losses that are not necessarily counterbalanced by better professional positions in the end. Some individuals may thus value other advantages over work-related returns.

## Returns to education and educational inflation

Monetary returns to education are usually estimated through Mincerian wage equations (Lemieux, 2006; Mincer, 1974). In their basic version, earnings are function of schooling (measured in years of education) and professional experience. Since the 1970s, some refinements has been proposed to improve the estimates, such as the inclusion of control variables (notably individual ability) or more complex functional forms (Heckman et al., 2006). However, the key element remains the same: the coefficient associated to education corresponds to its returns and should be positive. Non-monetary work-related returns to education are also supposed to be positive: additional years of education should be associated with a lower risk of unemployment (Kiefer, 1985; Nickell, 1979). Finally, in the workplace, education is in theory associated with jobs with higher responsibilities, managerial tasks and more complex activities. Therefore, access to more demanding labour market positions can eventually be seen as a return to education.

There is cause for concern if returns to education are lower than expected: if the earnings premium associated with a higher degree is not sufficient to compensate the cost of obtaining it; if education does not provide the expected protection against unemployment; if individuals cannot find jobs that match their education. Overeducation falls into the latter. Unfortunately, the literature shows that, whether in terms of educational (mis)match or wages, returns to education have been declining over the last decades.

Recent studies show that monetary returns to education are decreasing with time (see e.g. Argan & Gary-Bobo, 2023; Emmons et al., 2019; Ichino et al., 2022, respectively for France, the US and the UK): having a higher education degree provides a lower wage premium nowadays than thirty years ago. Argan et al. (2023) show that, although there is still a clear hierarchy of degrees in terms of returns to education, the highest credentials are associated with lower real wages in the 2010s than at end of the 1990s in France (around 7.2% lower for Masters’ degrees, around 6% lower for Business and engineering school graduates): degrees are “devaluating”.



Some studies argue that the declining value of degrees results from a “race between education and technology” (Autor et al., 2020; Goldin & Katz, 2008), related to skill-biased technological change (Acemoglu, 1999; Autor et al., 2006): if the demand for skilled labour grows faster than the supply (due in particular to technical change), technology “takes the lead” and wage premiums increase; if the supply of skilled labour exceeds the demand, education “wins the race”, but its returns decrease. This leads us to reconsider the benefits of the educational expansion observed in most developed countries over the last decades.

Because of the strong increase in the number of graduates, degrees are said to lose some of their value on the labour market. By analogy with price inflation, where increasing the amount of money circulating in the economy reduces the amount of goods each monetary unit allows to acquire, this phenomenon is often called “educational inflation” or “credential inflation” (Collins, 2002; Duru-Bellat, 2006). Duru-Bellat (2006) distinguishes between the intrinsic value, i.e. the amount and quality of knowledge and skills that individuals acquired by obtaining the degree, and the instrumental value of degrees, i.e. the relative value of the degree compared to others, especially used by recruiters to classify candidates<sup>7</sup>. The former therefore refers to Human Capital Theory and the latter to the signaling theory.

From the Human Capital point of view, a decrease in the earnings originates from a lower productivity. This suggests the knowledge and skills acquired through education may no longer be as relevant or sufficient as they once were. If education is not able to provide “enough” or “the right” skills, one can question its quality. Indeed, there is previous evidence about a higher quality of education increasing monetary returns (Card & Krueger, 1992; Long, 2010) and improving matching status (Robst, 1995). Conversely, the decrease in earnings and the rise of overeducation could therefore be the consequence of a drop in the quality of schooling over time. Results by Chevalier (2003) somehow support this idea. Indeed, he showed that, among overeducated workers, some individuals have a higher degree than what is theoretically required, but consider having the adequate amount of skills: they are only “apparently overeducated”. This suggests an inefficiency of the educational system to provide the needed skills, which individuals compensate for by studying longer. However, some individuals are “genuinely overeducated”: they both have a

---

<sup>7</sup>It is worth mentioning that Collins (2002), when he writes about “credential inflation”, only discusses the instrumental value aspect. On the contrary, Duru-Bellat (2006) takes into account both values when writing about “educational inflation”. Nevertheless, both terminologies are used interchangeably in a number of works to designate one phenomenon or the other.

higher degree than required and consider that their skills are not fully used in their job. In this case, education seems able to provide the right skills and knowledge, but not to be valued accordingly on the labour market. Therefore, the human capital approach of the decreasing returns to education over time, while not to be dismissed, appears not to fully explain the phenomenon.

Another part of the explanation may lie in questioning the signaling role of education. The underlying idea behind “credential inflation” (Collins, 2002) is that, if everybody has a degree, it cannot be used to distinguish individuals with more ability from the others anymore. In this view, education would have become a poor signal for recruiters. Ingram and Neumann (2006) showed that educational expansion is associated with an increase of the skill heterogeneity among individuals with the same degree. Therefore, in order to balance information asymmetry, employers may reduce offered wages or hire more educated workers as an insurance for a potential lack of skills, even though Argan et al. (2023) showed that, with equivalent degrees, graduates are on average more able nowadays than twenty years ago. Moreover, there is a strong heterogeneity in the returns of degrees coming from different higher education institutions that, on paper, offer the same curriculum (Robst, 1995), which further increases the asymmetry of information between candidates and recruiters.

Everything discussed so far in this section focused on education as an investment, investment which would ultimately not pay off on the labour market. If education is pursued with consumption motives, one can wonder whether lower work-related returns (and in particular overeducation) are really an issue. Indeed, those individuals value other types of returns and have a positive willingness to pay for it (Alstadsæter, 2011). Therefore, overeducation might seem a tolerable strategy, with the lower returns being more accepted by the individuals. However, results by Sellami et al. (2020) do not only confirm the negative effect on wages, but also evidence a greater penalty in terms of job satisfaction than for other overeducated workers. This thus challenges findings by Alstadsæter et al. (2008), who showed that individuals motivated by educational consumption traded “money for joy”.

The question of the occupational returns to education is to some extent less important when it comes to educational choices linked to social norms, as fields related to this motive (especially Health and Education) are generally associated with low risks of overeducation (Rossen et al., 2019; Sellami et al., 2020). Besides, Sellami et al. (2020) showed that the concerned individuals, when overeducated,

suffer only small wage penalties and no significant loss of job satisfaction, suggesting that the utility gains associated to social conformism indeed compensate the eventual lower work-related returns (Akerlof & Kranton, 2000).

Finally, questioning whether overeducation is a problem in cases where education was pursued for motives other than investment still rely on the assumption that individuals make their choices rationally, with perfect information about the future consequences of these decisions. However, educational choices are actually made in an uncertain environment (Altonji, 1993) in terms of the future structure of the labour market and economic climate. Educational choices are also likely to be driven by biased beliefs and to be time-inconsistent: a choice made when 16 to 18 years old may bring immediate satisfaction, but be perceived as detrimental by the same individual when older (Koch et al., 2015).

## Overeducation, matching frictions and mobility

The previous sections described how overeducation can be seen as a lower return to education than expected, and how it can be related to the initial educational choices of the individuals and strong educational expansion over the last decades. Therefore, the large increase in educational attainment is often presented as one of the main factors in the emergence of overeducation (Groot & Maassen van den Brink, 2000; Hartog, 2000; Verhaest & van der Velden, 2013). To deal with this issue, one public policy approach could be to constraint individuals' educational choices, notably using financial incentives and barriers (Alstadsæter et al., 2008), and potentially reduce the number of graduates, although that may not be desirable from a social point of view (Alstadsæter, 2011), especially since education has positive externalities on other social outcomes, such as the reduction of criminality, the improvement of health or the increase of civic participation (Lochner, 2011). Another option could lie in further standardisation of educational curricula, to lower the skill heterogeneity between graduates (Chevalier & Lindley, 2009) and the quality disparities between higher education institutions (Robst, 1995). Finally, providing young people better information about work environments and the structure of the labour market could help to reduce educational mismatch (Quintini, 2011b).

However, this approach of overeducation (and how to tackle it) from the school system's perspective leaves us with two questions. Firstly, some studies moderate the effect of educational expansion on the rising incidence of overeducation (Croce

& Ghignoni, 2012) and evidence effects of the labour market structure (Davia et al., 2017), especially related to the demand in skilled labour (Ghignoni & Verashchagina, 2014), macroeconomic conditions (Croce & Ghignoni, 2012) or institutional context (Mendes de Oliveira et al., 2000). Secondly, policies aiming to transform the educational system may reduce the risk of overeducation for future graduates who have not yet left school, but do not bring any solution for individuals who are already on the labour market. Therefore, education-related policies, while useful and necessary, are not sufficient to completely solve the overeducation issue.

On this basis, the aim of this thesis is to address overeducation from the labour market rather than the educational system perspective. This leads me to explore the phenomenon, not (only) as a negative output of the resource creation, but as an allocation problem of the existing resource (the resource being educated labour). This approach is supported by the fact that the allocation of workers into jobs is rarely optimal, with unemployment and overeducation occurring alongside unfilled vacancies (Cedefop, 2018). On this matter, Dolado et al. (2009) or Gautier (2002) show that overeducation can result from search and matching frictions (see also the very recent work of Navarini & Verhaest, 2023).

Some stylised facts illustrate these search and matching issues. The *Enquête Besoins en Main-d'Œuvre* (BMO) from the French public employment agency *Pôle Emploi* surveys recruitment needs by occupation and business sectors, especially to evaluate hiring difficulties. Looking at the 2023 BMO survey<sup>8</sup>, employers report troubles to hire even in fields generally associated with a high incidence of overeducation, such as Economics, Business and Law (Sellami et al., 2020), Humanities (Barone & Ortiz, 2011) or Natural sciences (Rossen et al., 2019). It is the case, respectively, for legal professionals (66% of the recruitments considered as difficult), accounting occupations (63%) or bank employees (56%); for translators (61%) or socio-cultural animation professionals (59%); or for laboratory staff (52%), for instance. To summarise, it thus appears that, in some sectors, we observe both individuals unable to find suitable jobs and employers unable to find suitable candidates, suggesting that firms and job seekers are not able to meet each other.

Part of the explanation may reside in a discrepancy between the skills that workers have to offer and the ones employers are looking for (Chevalier & Lindley, 2009; Chevalier, 2003). This leads back to what is discussed in the previous sections,

---

<sup>8</sup>Pôle Emploi (2023), *Enquête Besoins en Main-d'Œuvre* (BMO). <https://statistiques.pole-emploi.org/bmo/bmo?le=0&pp=2023&css=1> (Accessed on 20 June 2023).

especially about aiming at a better consistency between education and occupations. As suggested by Dolado et al. (2009), another explanation may be related to a lack of mobility of the labour force. This is the angle from which the subject will be addressed in this thesis: how does mobility, whether professional or geographical, interact with overeducation? On the one hand, this involves exploring to what extent and in which context mobility can reduce overeducation. On the other hand, it involves determining whether mismatched workers are able of being mobile or not.

### **Overeducation and professional mobility**

Regarding professional (im)mobility, the starting point is the job search theory (McCall, 1970; Stigler, 1962). In this framework, employers have imperfect knowledge of the quality of workers, and job seekers have imperfect knowledge of the wages, employment stability, etc., offered by the firm. Therefore, a more efficient allocation of the labour force derives from better information. Building on this, Jovanovic (1979), in his theory of turnover, explains that a way to acquire this information is repeated job search: only matched workers stay in their jobs, while mismatched ones quit their consecutive positions until they find an adequate one, which they achieve by accumulating more information with time.

With regard to overeducation, this is what would predict the career mobility theory (Sicherman & Galor, 1990), where mismatched jobs at the beginning of the career can help accessing better positions later on. In Dolado et al. (2009), mismatched workers engage in on-the-job search in order to find a better match. All of this rely on the assumption that workers are able to move from one job to another, and that this mobility is systematically upward. Gangl (2003) however showed that job mobility can be limited by employment protection legislation (EPL), which tend to be high in most European countries (OECD, 2020). Moreover, fixed-term contracts, less impacted by EPL and allowing more recurrent job search, do not necessarily allow access better labour market positions, because of a stigma effect associated with this type of contract (Filomena & Picchio, 2021; Svalund & Berglund, 2018). Empirical evidence about overeducation points towards a negative effect of mismatch on future positions, with individuals being trapped in overeducated jobs for a long time (see e.g. Baert et al., 2013; Meroni & Vera-Toscano, 2017; Scherer, 2004; Voßemer & Schuck, 2016) and encountering difficulties in accessing more adequate jobs. On this matter, Baert and Verhaest (2019) show that there is a stigma effect of overeducation, and Rubb (2003) suggests that on-the-job training is necessary for

workers to exit overeducation.

### Overeducation and spatial mobility

The links between geographical (or spatial) mobility and opportunities on the labour market are explored since the seminal paper of Kain (1968), first to introduce the “spatial mismatch” hypothesis: there are negative effects associated with living in areas located far from job opportunities<sup>9</sup>. Therefore, less mobile individuals are more likely to experience poor labour market outcomes. Regarding overeducation, this may be amplified by the non-random spatial distribution of jobs: low skilled jobs are more evenly distributed geographically, while higher skilled jobs tend to be available only in specific locations (P. M. Blau & Duncan, 1967; Détang-Dessendre, 1999). Following the assumption that married women limit their job search to the local labour market on which their husbands work, Frank (1978) evidenced that this spatial constraint led to a higher risk of overeducation and lower wages.

A more general framework explaining the potential trade-off between spatial mobility and job match is proposed by Simpson (1992). On a local labour market with no adequately matched job available, job seekers have three options : (1) to stay unemployed, with the expectation of finding a matched job later; (2) to accept a job for which they are overeducated; (3) to geographically broaden their job search in order to find a matched job located farther away. Empirical evidence show that more spatially mobile individuals face lower risks of overeducation (see e.g. Büchel & van Ham, 2003; Hensen et al., 2009; Venhorst & Cörvers, 2018). This is true both for commuting (Di Paolo et al., 2017) and residential migration (Iammarino & Marinelli, 2015; Jauhiainen, 2011). However, the negative effects depend on the distance of the migration: only migrations over a sufficient distance reduce the risk of mismatch (Jauhiainen, 2011), by providing access to another local/regional labour market. There are also strong disparities in the reduction of mismatch depending on both the departure and arrival regions (Devillanova, 2013; Iammarino & Marinelli, 2015), suggesting effects of the local context (Tselios, 2013).

However, spatial mobility involves important costs, both financial (transportation costs in the case of commuting, relocation costs for residential migration) and psychological, which implies that the decision to move is not trivial. Especially,

---

<sup>9</sup>In the original paper, Kain (1968) focused on male black workers. This hypothesis has since been extended to a more general population, although empirical evidence confirms that some demographic groups are more affected than others by spatial mismatch (see Gobillon & Selod, 2021, for a recent review about this subject).

individuals may be reluctant to move if they are not assured to compensate the mobility costs (Greenwood, 1997).

## Overeducation: incidence, consequences, drivers

The previous section gave some insight into the potential mechanisms behind the overeducation phenomenon as well as some theoretical background. This section completes the picture by providing an overview of the empirical literature on overeducation, its drivers and consequences.

It is hard to draw a clear and unique picture of the incidence of overeducation. Quintini (2011b) mentioned a share of overeducated workers around 25% in OECD countries, while pointing out wide variations between countries, periods and measurement approaches. Most of the evidence however agree on an increase of the phenomenon over the past decades (see e.g. Di Paola & Moullet, 2018; Hartog, 2000; Kucel, 2011), although a few studies rather showed a high, but stable, incidence of overeducation (Groot & Maassen van den Brink, 2000; McGuinness, 2006).

There are three main approaches to measure overeducation. Objective measures of overeducation are based on an *ex ante* definition of what should be the “normal” educational level(s) to hold a particular job (Affichard, 1981; Rumberger, 1981). Statistical measures are based on the distribution of education in the population, and rely on the mean (Verdugo & Verdugo, 1989), mode (Davia et al., 2017; Mendes de Oliveira et al., 2000) or median education in each occupation. Subjective measures are based on workers’ self-assessment about their job situation (Duncan & Hoffman, 1981). Surveying previous evidence, Capsada-Munsech (2019) showed that statistical measures lead to the lowest incidence, around 15% on average; objective measures show rates around 20%; subjective measures are associated with both the highest average incidence (30%) and the highest dispersion.

Quintini (2011a) summarised studies from the 1990s and 2000s and highlighted profiles of countries regarding the incidence of overeducation: on average, rates are higher in Southern Europe (Spain, Portugal), they are more moderate in Anglo-Saxon countries (US, UK) and Western Europe (Germany, Belgium). Surprisingly, the highest and the lowest incidence are both reported for Scandinavian countries, respectively Sweden and Finland. Results by McGuinness et al. (2018) somewhat confirm these profiles, providing estimates for additional countries. Overeducation

is higher in Italy (24%), Greece (28%) and Cyprus (31%), moderate in France (17%) and Austria (19%) and significantly lower in Eastern Europe (9% in Slovenia, 11% in Bulgaria). Results for Scandinavia are less spread than in Quintini (2011a), with all countries experiencing a moderate incidence, similar to what is observed in Western Europe. Another interesting result in McGuinness et al. (2018) is the convergence of overeducation between countries, mostly because of a rising incidence in countries with originally low overeducation rates while mismatch is stable in countries with higher rates.

Historically, the literature about the consequences of overeducation first focused on its effects on wages. Duncan and Hoffman (1981) extended the Mincerian equations (Mincer, 1974) to explore the differentiated returns to surplus (and deficit) education. Hartog (2000) referred to this framework as the ORU (Over-, Required and Under-Education) function. Estimates of the ORU function provide strong evidence of a positive return to overeducation years, but smaller than the return to required years (see e.g. Duncan & Hoffman, 1981; McGuinness, 2006; Mendes de Oliveira et al., 2000; Sloane et al., 1999): if overeducated workers earn more than individuals in the same job with a lower degree, they earn less than workers with the same degree in a matched job. At the individual level, overeducation is also associated with lower satisfaction and well-being, both at work (Allen & van der Velden, 2001; Tsang et al., 1991) and in life in general (Piper, 2015), in comparison with workers with the same educational level but in a matched job.

At the firm level, overeducation is associated with a higher turnover (Alba-Ramírez, 1993; Sloane et al., 1999; Tsang et al., 1991) and losses of productivity (Mendes de Oliveira et al., 2000; Tsang et al., 1991), which may have a negative impact on the economy as a whole (McGuinness, 2006; Tsang, 1987). However, Kampelmann et al. (2020) showed that overeducation can in some cases be associated with higher profitability for the companies, especially as overeducated workers are less productive than workers with the same degree in an adequate job, but more productive than less educated workers doing the same job (Büchel, 2002).

There is a abundant literature on how overeducation vary depending on the individual characteristics of the workers, but also depending on characteristics of their job environment, such as firm characteristics, economic context, and education and labour market institutions.

The estimated incidence of overeducation is generally slightly higher for women



than for men, but this difference is mainly due to spatial constraints (Frank, 1978), marital status (McGoldrick & Robst, 1996) or the presence of children (Büchel & van Ham, 2003), rather a pure gender effect. Focusing on the public sector, Di Paola and Moullet (2012) showed that women were statistically more overeducated, but subjectively less likely to consider themselves as such, which can illustrate a higher willingness to pay for the advantages of public jobs. There is also evidence of a higher overeducation risk among individuals with immigrant background (Lindley, 2009), lower socio-economic background (Barone & Ortiz, 2011), disability (Büchel & van Ham, 2003) and past unemployment spells (Baert & Verhaest, 2019).

Young workers are more likely to experience overeducation (Dolton & Vignoles, 2000; Forgeot & Gautié, 1997). However, more than an effect of age, this is related to the distance from the entry on the labour market, since the age at the time of graduation has very limited impact on mismatch (Baert et al., 2013; Meroni & Vera-Toscano, 2017). One explanation is the lack of professional experience and a potential lack of skills of young graduates (Chevalier & Lindley, 2009; Chevalier, 2003). Therefore, overeducation should decrease with the time spent on the labour market (Rubb, 2003; Sicherman & Galor, 1990), although evidence about it is mixed (Baert et al., 2013; Meroni & Vera-Toscano, 2017), as mentioned earlier.

Overeducation is also influenced by variables related to the education of the individuals. Theoretically, the educational level has an ambiguous effect: on the one hand, having a higher degree should be related to higher skills and access to better jobs; on the other hand, higher degrees are above the standard for more occupations, increasing the risk to observe a mismatch. Most of the empirical evidence however points towards a negative effect of higher degrees on overeducation (see e.g. Di Paola & Moullet, 2018; Halaby, 1994; Meroni & Vera-Toscano, 2017). The risk of overeducation also varies by field of study (Barone & Ortiz, 2011; Rossen et al., 2019; Sellami et al., 2020), with Natural Sciences, Humanities and Economics, Business and Law being generally associated with higher rates of overeducation. Finally, Robst (1995) found that mismatch is affected by the quality of the university.

Overeducation tends to be higher in smaller firms (Ghignoni & Verashchagina, 2014; Guironnet, 2006), where hiring a worker above the required level can be a way to limit the risk of a lack of skills, which would be more costly than in larger companies. Regarding the business sectors, the risk of mismatch is generally higher in Agriculture and Sales (Nauze-Fichet & Tomasini, 2002), and in the private compared to the public sector (Di Paola & Moullet, 2009).

A more recent strand of the overeducation literature focuses on its macro-levels drivers (see e.g. Croce & Ghignoni, 2012; Davia et al., 2017), providing a different insight about the factors leading to overeducation. However, these drivers strongly vary depending on the country and institutional context (Ghignoni & Verashchagina, 2014), resulting in contradictory and inconsistent results across studies. Guironnet (2006) evidenced a positive effect of regional unemployment on the risk of overeducation, while Davia et al. (2017) found overeducation to be lower in areas with higher unemployment. Büchel and van Ham (2003) do not find any significant effect of regional unemployment on overeducation. Various results are also reported concerning EPL. Di Pietro (2002) found a stronger EPL to be associated with a higher risk of mismatch, related to a shortage of job opportunities; on the contrary, Davia et al. (2017) found a negative effect, which they associate with employers willing to achieve better matches because of higher costs of separation. Verhaest and van der Velden (2013) found the effect of EPL to be insignificant. Davia et al. (2017) also showed that overeducation is related to the imbalance between supply and demand of skilled labour (the “race between education and technology”, Goldin & Katz, 2008), suggesting that policies aimed at stimulating the demand can contribute to the reduction of mismatch (Ghignoni & Verashchagina, 2014). Regarding the global economic context, Croce and Ghignoni (2012) found that overeducation responds to cyclical fluctuations, increasing during recessions and decreasing in more favourable periods. However, Tselios (2013) found the risk of mismatch to be higher in more dynamic regions because of an excess of competition for jobs.

## Road map of the thesis

This thesis makes empirical contributions relative to the links between overeducation and mobility, whether spatial or professional. More specifically, it aims to assess the extent to which mobility can alleviate overeducation, as well as the issues and limitations that this might encounter. From the geographical point of view, this leads to explore local and regional labour markets disparities, and the links between territorial and professional inadequacy. From the professional point of view, this thesis focuses on the trap effect of mismatch and the ability of individuals to exit from unsuitable job positions.

This thesis is divided in four chapters<sup>10</sup>, in which I investigate overeducation, on the one hand from the perspective of the geography of local labour markets (chapter 2) and individual spatial mobility (chapter 3); on the other hand, in terms of occupational mobility and career trajectories (chapter 4). Prior to these analyses, chapter 1 addresses methodological issues related to the measurement of overeducation. The following paragraphs detail the content of each chapter.

The first chapter, entitled *Measuring overeducation using skills requirements*, differs from what will follow in that it is not an empirical analysis but a methodological work about overeducation and its measures. It does, however, lay the foundations on technical points that will be useful throughout the three other chapters. In particular, it describes the various alternatives available for measuring the phenomenon, as well as the advantages, drawbacks and implications of each of them.

My main contribution in this chapter is to propose a new measure of overeducation, called *Skill Analysis*, based on the skills content of educational levels and occupations. This work is motivated by two main facts. Firstly, it is interesting to look at the content of jobs in terms of tasks and skills required to understand how degrees relate to them, especially in a context of modernisation of the education system that places skills at the heart of the equation (European Commission, 2017). Secondly, measures based on the content of jobs require very frequent updates in order to keep pace with the changes in the labour market structure and not become obsolete (van der Meer, 2006), otherwise the incidence of the phenomenon will be estimated following a standard that is no longer relevant. However, these updates are costly, so objective measures, although conceptually better (Verhaest & Omeij, 2006), are not widely available.

In order to build this measure of overeducation based on skills requirements, I mobilise data from the Occupational Information Network (O\*NET) of the US Department of Labor and information about learning outcomes from the European Qualifications Framework (EQF) defined by the European Commission. The O\*NET database contains detailed information about the content of jobs, from the activities and tasks performed to the knowledge and skills required, working conditions and future prospects for each occupation. The EQF provides information about the qualifications acquired in the different educational systems and labour markets in

---

<sup>10</sup>Each chapter of the thesis is an independent research article. This explains the repetition of some elements between chapters (notably contextual elements and literature) and the use of the terms “article” and “paper”. Chapter 3 is co-authored with Florent Sari, which explains the use of the pronoun “we”.

Europe. It defines levels of qualifications on a common base for each member country, with a description of the skills and knowledge associated to each level, as well as the related degrees and professional certifications. I first associate skill scores to each EQF level, that I compare to the skill scores of O\*NET occupations, in order to determine for each one the EQF requirement. I then link this classification to the International Standard Classification of Occupations (ISCO-08) to define the theoretical education-skills requirements for each ISCO-08 occupation.

This work results in an objective measure of overeducation, which I can keep up-to-date through the procedure described in the paper. I rely on the definition of precise occupations to deal with the aggregation problem that arises with the statistical alternatives, i.e. differences in educational requirements within larger occupational groups. I compare this measure to existing ones using data from the French Labour Force Survey. I find that a significant share of workers is statistically but not objectively overeducated, and only moderate correlations between approaches, which suggests that different measures are related to different phenomena.

In the second chapter, entitled *Local labour markets and spatial determinants of overeducation*, I analyse overeducation from the perspective of local labour markets. Starting from the fact that individuals are primarily looking for jobs around their residential location, I explore how the geography and local context of labour markets affect overeducation, focusing on three aspects. Firstly, this chapter investigates the role of local labour market characteristics in explaining overeducation, with a distinction between urban and rural areas. Secondly, it assesses the spatial dependence of local labour markets regarding overeducation, similarly to what has previously been evidenced for unemployment (Patacchini & Zenou, 2007). Thirdly, it documents the disparities across gender regarding the spatial dimension of overeducation (Frank, 1978).

To explore this matter, I mobilise data from the French population census, that allow me to calculate overeducation rates at the scale of the French employment areas (statistical divisions designed to reflect local labour markets) and provide detailed socio-economic information about a large number of individuals, as well as data from the Directory of Enterprises and Institutions about the demography of firms. I mobilise spatial econometrics models in order to take spatial dependence into account and to determine how neighbouring labour markets influence each other. In addition, I use spatial regimes models to take spatial heterogeneity into account and to explore how macro-level drivers vary between urban and rural areas.

Several results emerge from this analysis. Firstly, I find a strong impact of local labour market characteristics on overeducation, with variations between urban and rural areas. The structure of employment by sector plays a significant role only in rural areas, while the imbalance between the supply and demand for educated work is the main driver in urban areas. This suggests that overeducation is related to the (lack of) variety of job opportunities in rural areas and to an excessive competition for jobs in cities. Secondly, my results confirm the influence of neighbouring areas on educational mismatch only for men, but not for women. More generally, the local labour market context has heterogeneous effects across gender. Women appear to be more affected by the context of the labour market, but at a more local scale than men.

The third chapter, entitled *Spatial mobility and overeducation of young workers*, is directly related to the previous one: given that there is strong spatial disparities in the incidence and drivers of overeducation, one can wonder if workers can relocate to avoid being mismatched. This is what would predict the theory of Simpson (1992), who claimed that spatial flexibility was a way to escape both unemployment and overeducation. Besides, this intuition is supported by the lower effect of the geography of labour markets on young people highlighted in the previous chapter: young workers may be less affected by the local context in a given area because they are more mobile. Therefore, this study aims to assess whether individual spatial mobility reduces the risk of being overeducated or not. Moreover, the geographical disparities are once again addressed, but in a different way, with a focus on residential migration to Paris. Indeed, the French labour market is characterised by the central role of its capital city, which alone accounts for 20% of the country's labour force. This centralism also translates into higher skills and wages on average (Combes et al., 2015). Finally, we explore the differentiated effects of residential migration depending on the educational level of individuals, as the distribution of skilled jobs is not random (Détang-Dessendre, 1999) and relocation behaviours differ according to individuals' level of schooling (Combes et al., 2012).

Using data from the *Génération 2010* survey from Céreq<sup>11</sup>, we estimate the effect of residential migration on the probability of being overeducated three years after graduation. This study is confronted with two identification issues: (1) a sample selection problem, as mismatch can only be observed for individuals in employment and (2) an endogeneity bias in the decision of migration, because of unobservable

---

<sup>11</sup>Centre d'études et de recherches sur les qualifications.

factors affecting both mobility and overeducation and/or reverse causality. To deal with sample selection, we rely on the Heckman two-step correction (Heckman, 1979). Endogeneity is addressed using an instrumental variables approach. Alternatively, we use a triprobit specification to estimate simultaneously the employment, migration and mismatch equations to check the robustness of our results. Our findings are consistent throughout the different specifications. In addition, the data allow us to use several definitions of the phenomenon of overeducation. Many studies only exploit one definition due to data availability, which has often been presented as a limitation because of the divergent results obtained through different approaches. In this work, we rely on both statistical and subjective definitions of mismatch, in order to explore the differentiated effects respectively on the fact of being overeducated and on the feeling of it.

Our contribution in this study is threefold. Firstly, our results confirm that individual spatial mobility reduces the risk of being overeducated (around 5 percentage points on average). Moreover, this negative effect is significantly stronger for migration to Paris than for migration between other regions (five to six times higher). Secondly, we show that, while residential migration reduces the risk of overeducation for more educated workers, it has no significant effect for short-cycle university graduates. This suggests that jobs corresponding to longer university education are available only in specific areas, hence the importance of being mobile to find a matched job. Thirdly, we find a negative effect of migration on mismatch for both the statistical and the subjective definitions, but its magnitude varies between the approaches. In addition, we evidence once again differences between educational levels. In particular, the most educated workers appear to be statistically less mismatched, but not to self-assess as such.

The last chapter, *Educational mismatch and early career trajectories*, extends the analysis to horizontal mismatch, i.e. individuals working in a field different from the one they studied, in addition to overeducation, and moves its focus from spatial to professional mobility. According to the career mobility theory (Sicherman & Galor, 1990), being overeducated at the beginning of the career should be a “stepping stone” to better professional positions. The empirical literature however challenges this, showing a “trap effect” of overeducation (see e.g. Baert et al., 2013; Meroni & Vera-Toscano, 2017), young workers remaining mismatched for a long time if they enter the labour market in such jobs.

In this chapter, I try to assess whether taking a mismatched job at the beginning

of the career has negative effects on future labour market outcomes or not. In particular, I investigate the “stepping stone” or “trap” nature of such jobs, not only for overeducation, but also in the case of horizontal mismatch. If they are trapped in mismatched positions, I also want to determine if they can transit from a type of mismatch to another. Finally, I explore the effects of being, not only overeducated or horizontally mismatched, but also of being mismatched in both dimensions at the same time. Indeed, individuals fully mismatched are the furthest from their education, they might therefore suffer stronger detrimental effects than individuals mismatched in only one dimension.

I use data from the *Génération 2017* survey from Céreq to investigate these questions. Using a dynamic treatment framework based on propensity score matching (Sianesi, 2004, 2008), I estimate the treatment effects of being mismatched in the first job, in comparison to stay longer in unemployment, on labour market outcomes of individuals, observed three years after graduation. Three types of mismatches are explored: overeducation (vertical mismatch), horizontal mismatch and the combination of both. Such a strategy allows to account for heterogeneous times of treatment and graduation-to-job durations, which are likely to influence both the likelihood of treatment and its effects. Indeed, it allows to compare treated (mismatched) individuals, not to never treated ones, but to individuals not yet treated: individuals who decide to remain unemployed at a given point in time might eventually accept to be mismatched later.

The important finding of this chapter is that mismatch, either vertical, horizontal or in both dimensions, is a “trap”: accepting a mismatched first job rather than staying longer in unemployment leads to a higher probability to be in a mismatched job three years after graduation. Besides, I find only small evidence that workers transit from one type of mismatch to another: most of them are trapped in the same type of mismatch throughout the time window. However, accepting to be mismatched in the first job appears to have positive effects on future employment, especially if the first job is taken early after graduation. Therefore, there still can be a trade-off regarding the quality of the job to limit future risks of unemployment. Overall, I find that different mismatches have different effects on labour markets outcomes. Nevertheless, young graduates should not accept jobs for which they are mismatched in both dimensions, as they have very limited positive effects on future labour market outcomes.





# Chapter 1

## Measuring overeducation using skills requirements

Overeducation can be measured through several approaches, each one leading to different results as to its incidence. This paper aims to build a new and up-to-date measure of overeducation, called *Skill Analysis* measure, based on the content of jobs in terms of tasks and skills required to perform a job. To this end, I mobilise data from the Occupational Information Network (O\*NET) and information about learning outcomes from the European Qualifications Framework. I then compare the new measure of overeducation to usual statistical measures using the French Labour Force Survey. Results confirm that objective and statistical measures do not apprehend the exact same matter, showing only moderate correlations between approaches.

**JEL classification:** I21, I23, J24

**Keywords:** overeducation, educational mismatch, skills requirements, measurement issues

---

I would like to thank participants in the LEMNA Doctoral Research Day 2022 for their comments on earlier versions of this chapter.

# Mesurer la sur-éducation à partir des compétences requises

La sur-éducation peut être mesurée via diverses approches, pouvant mener à des résultats parfois très différents. Ce chapitre vise à construire une nouvelle mesure actualisée de la sur-éducation, basée sur le contenu des emplois en termes de tâches à effectuer et de compétences requises. Pour ce faire, je mobilise les données d'O\*NET (*Occupational Information Network*) du Département du Travail des États-Unis, ainsi que les résultats d'apprentissage (*learning outcomes*) associés à chaque niveau d'études par le Cadre Européen des Certifications (CEC). Par la suite, je compare cette nouvelle mesure à ses alternatives statistiques à partir des données de l'Enquête Emploi. Les résultats confirment que les mesures objectives et statistiques n'appréhendent pas le même phénomène, révélant des corrélations relativement modérées entre les approches.

**Mots-clés :** sur-éducation, inadéquation emploi-diplôme, compétences, problèmes de mesure

## 1.1 Introduction

From the labour market perspective, education is not always used to its full potential, as some individuals study longer than they actually need to do their job: they are “overeducated” (Kucel, 2011; McGuinness, 2006). According to the European skill and jobs survey (ESJS), about 30% of the adult population of the European Union is mismatched in some way (Cedefop, 2018).

If the increase of overeducation is a widely shared finding in the literature, its incidence is however still debated, especially since it can be measured through several approaches. Capsada-Munsech (2019) show that, depending on the measure, overeducation rates range from 12% to 38% using the same data. The author points that it has long not been of concern because the differences in measures do not substantially affect the returns of overeducation (McGuinness, 2006). However, recent studies discussed the drivers of the phenomenon (Croce & Ghignoni, 2012; Davia et al., 2017; Ghignoni & Verashchagina, 2014), using overeducation as a dependent variable and raising the question of what is exactly measured and how.

In this paper, “overeducation”<sup>1</sup> is addressed from the labour market point of view, as an inadequacy between the formal educational level of the individuals and the one required for their job<sup>2</sup>. Overeducation can be seen as an educational investment that did not pay off. There is an opportunity cost to study, because of the direct cost of education and the absence of wages during education years (Becker, 1964). Thus, if additional years of education do not lead to a higher professional position and higher wages, there is a risk that the opportunity cost will never be compensated. This is consistent with both the Human Capital Theory (Becker, 1964) and job-market signalling (Spence, 1973). In the former, education allows individuals to accumulate skills and knowledge that will be used on the labour market; in the latter, educational credentials indicate the level of ability of individuals, that allows them to access better (paid) jobs. In both cases, education is supposed to improve labour market outcomes.

---

<sup>1</sup>Overeducation has to be distinguished from other closely related but different concepts: overskilling and horizontal mismatch. Overskilling occurs when the skills of an individual are not fully used in the job she or he occupies. How these skills were acquired (formal education or professional training) is not taken into account, by contrast with overeducation (Capsada-Munsech, 2017). Allen and van der Velden (2001) showed that educational mismatch and skill mismatch, although closely related, do not systematically overlap. Horizontal job-education mismatch occurs when individuals do not work in their field of study.

<sup>2</sup>In this regard, the word “overschooling” might seem less ambiguous and a good alternative to the term “overeducation”. However, although it has sometimes been used, it remains quite rare in the literature.

Overeducation can be measured through several approaches, each one leading to different results as to its incidence. Statistical measures are based on the distribution of education in the population. Subjective measures are based on workers' assessment about their job situation. Objective measures of overeducation are based on an *ex ante* definition of what should be the “normal” (in the sense of “conforming to a standard”)<sup>3</sup> educational level(s) to hold a particular job. Objective measures are often presented as better measures than their statistical or subjective alternatives (Capsada-Munsech, 2019; Verhaest & Omeij, 2006). However, they also raise some issues. First, they require to be updated regularly to remain relevant. Due to the cost of keeping them consistent with the evolutions of the labour market structures, these measures are missing in several countries, or are based on old occupational dictionaries, which raises concerns regarding the accuracy of the measure (van der Meer, 2006). Furthermore, they may not be the most suitable for international comparison because of the differences in education and jobs between countries<sup>4</sup>.

The aim of this paper is to propose an objective measure dealing with these issues, built on the skills requirements in each occupation and the learning outcomes of each educational level in Europe, henceforth *Skill Analysis* measure. To this end, I mobilise data from the Occupational Information Network (O\*NET) of the US Department of Labor, which provide detailed information about the skills used in each occupation, and information about learning outcomes (that is the skills and knowledge provided by each educational level) from the European Qualifications Framework (EQF).

The *Skill Analysis* measure is based on the skills required to do a particular job, the requirements being defined *ex ante* by experts, and those provided by degrees. I first match EQF educational levels to occupations based on the skills scores drawn from O\*NET data and then I match these educational requirements to ISCO-08<sup>5</sup> occupations using the correspondence with O\*NET. This measure of overeducation relies on the definition of precise occupations to deal with the aggregation problem that arises with the statistical alternatives, *i.e.* the averaging of different educational requirements within larger occupational groups. Moreover, I define two versions of the requirements (upper requirement and lower requirement) to take into account

---

<sup>3</sup>This kind of measure is sometimes referred to as “normative measure” in earlier works.

<sup>4</sup>However, in the last decade, there has been an attempt to normalise education systems in Europe, which theoretically reduces differences in educational requirements between those countries.

<sup>5</sup>International Standard Classification of Occupations 2008.

the variety of jobs in a same occupation.

After building this measure, I compare it to existing measures of overeducation, in particular statistical measures. The upper *Skill Analysis* measure leads to the lowest overeducation incidence of all approaches, while the lower *Skill Analysis* shows an incidence similar to what is observed with statistical alternatives. The analysis of the correlations and overlaps between the *Skill Analysis* and the statistical measures shows that the various approaches do not take into account the same individuals, and tend to apprehend slightly different phenomenon.

This paper contributes to the literature about the measurement of overeducation, proposing a new measure that deals with the limitations of standard approaches. Other alternative measures have been proposed, using wages (Nauze-Fichet & Tomasini, 2002) or combining several types of measures (Chevalier, 2003). The *Skill Analysis* measure differs from these approaches by computing educational requirements based on large data about skills and a detailed nomenclature of occupations.

This paper is structured as follows. Section 1.2 exposes the existing measures of overeducation. Section 1.3 describes the construction of the *Skill Analysis* measure. Section 1.4 compares the new measure to existing measures, focusing in particular on young workers. Section 1.5 concludes.

## 1.2 Measuring overeducation

Back in the 1970s and 1980s, the first papers about overeducation relied on an objective measure (Rumberger, 1981). Objective measures are based on a correspondence grid between education and occupations, built *ex ante* by experts according to the content of jobs and degrees. Verhaest and Omey (2006) refer to this measure as *Job Analysis*<sup>6</sup>. Given that the educational norm is defined *a priori*, the objective measure has the advantage not to be affected by the actual distribution of education in the population, which is likely to bias the estimated incidence of overeducation. If this approach has the advantage of being more rigorous, it is however costly. Indeed, it implies the construction and the regular update of a correspondence between education and occupations, in order to keep up with labour market changes

---

<sup>6</sup>Since it is built by looking at the task content of the jobs, to some extent, this type of measure brings together overeducation and overskilling. However, it must be noted that it does not take into account skills that could have been obtained through professional experience.

and modifications of the content of degrees and jobs.

Statistical measures (also called empirical measures) define “normal” educational levels using the statistical distribution of education in each occupation. These measures can be built from the deviation from the mean (Croce & Ghignoni, 2012; Verdugo & Verdugo, 1989) or from the median, using the mode of the distribution (Davia et al., 2017; Ghignoni & Verashchagina, 2014) or the analysis of contingency tables between jobs and degrees (Forgeot & Gautié, 1997; Nauze-Fichet & Tomasini, 2002). They are easier to use because the required variables to build the measure (the educational level and the occupation of individuals) are available in a large number of datasets. However, as they are based on statistical distributions, they depend on the structure of education in the population studied. In particular, if many workers in an occupation are employed below their educational level, their situation will appear as “normal”, although they are objectively overeducated. Verhaest and Omey (2006) refer to these measures as *Realised Matches*. They illustrate the role of education as a signal sent to recruiters. However, as with the objective measure, these approaches are sensitive to the evolution of the structure of jobs and education across time: if educational requirements increase, overeducated workers at a point in time may no longer be mismatched in subsequent periods, without any actual changes in their situation.

Subjective measures are based on the workers’ evaluation of their own job situation. (Verhaest & Omey, 2006) refer to these measures as *Worker Self-Assessments*. Two different approaches can be distinguished, depending on the question asked to respondents. The first approach (direct approach) consists in asking directly to workers if they think they are overeducated. This strategy leads to important biases. First, a similar situation can be interpreted differently by different respondents (one worker declaring being overeducated while another in the same job with the same education states not being overeducated). Second, some individuals may choose (deliberately or not) to give a biased answer to understate an overeducation situation (Capsada-Munsech, 2019). Third, overeducation can be mixed up with other job characteristics, such as wage or job satisfaction (Capsada-Munsech, 2019). The second approach (indirect approach) consists in asking workers which degree is required to get or to do their job, referring to signalling (Spence, 1973) or to human capital (Becker, 1964) respectively. While the indirect subjective measure is less biased than the direct approach, individuals may still overstate the importance of

their own degree (Capsada-Munsech, 2019). Both approaches rely on unusual data (the answer to the question “Which degree do you need to do/get your job?”), but subjective measures allow to explore how overeducation situations are experienced by workers (Di Paola & Moullet, 2009, 2012), by contrast with objective or statistical measures.

Finally, Chevalier (2003) combines subjective workers’ assessment on job satisfaction and an objective definition of educational requirements to build an alternative measure of overeducation, aiming to account for the skill heterogeneity of the workers. The author defines as “genuinely overeducated” individuals who have education in excess of what is required for their job and are dissatisfied at work, and as “apparently overeducated” individuals who have education in excess but are satisfied in their job. Using this methodology, he shows that skills play a role in the risk of overeducation<sup>7</sup>.

Because of the variety of approaches to measure overeducation, several studies have explored their reliability and performance. Verhaest and Omey (2006), on Belgian data, compare five measures of overeducation: one *Job Analysis* measure, two *Self-Assessment measures* (depending on the degree needed to do or to get the job) and two *Realised Matches* measures (based on the mean and the mode of the distribution). Using statistical tests (F-tests and LR-tests), they show that the objective measure is better than the others. Indeed, the *ex ante* classification is not impacted by sample characteristics or workers’ subjectivity, and should be preferred whenever possible. In a meta-analysis of papers about overeducation from the 1970s to the 1990s, Groot and Maassen van den Brink (2000) find that objective and subjective measures lead to similar overeducation rates. Only the statistical measures showed slightly lower results. More recently, Capsada-Munsech (2019) compare two objective, two statistical and one subjective measures using REFLEX/HEGESCO<sup>8</sup> and LFS data for various European countries, and found low correlations between the different types of approaches. The author also underlines the importance of the thresholds above which individuals are overeducated. Finally, for national level analyses, objective measures are the most appropriate, or statistical measures if there is no up-to-date job-education correspondence. For cross-countries comparisons, subjective measures are more suitable because the other approaches raise the question of

---

<sup>7</sup>Meroni and Vera-Toscano (2017) find that both types of overeducation lead to the same detrimental effects on career paths.

<sup>8</sup>Research into Employment and professional/Higher Education as a Generator of Strategic Competences.

the comparability of skill and job structures between countries (Capsada-Munsech, 2019).

However, it appears that the measures do not necessarily substitute for each other. Indeed, some individuals are overeducated according to one measure, but are not according to another. This is particularly striking between subjective and objective measures: around 30% of the individuals are overeducated in only one of these two dimensions (Capsada-Munsech, 2019; Di Paola & Moullet, 2012). Thus, using several measures conjointly can bring light on different aspects of overeducation, but the choice of the measure is often driven by data availability (Capsada-Munsech, 2019).

## 1.3 Building a measure of mismatch based on skills requirements: the *Skill Analysis* measure

### 1.3.1 The European Qualifications Framework (EQF)

In order to build an up-to-date objective measure of overeducation, I mobilise information from the European Qualifications Framework (EQF) from the European Commission. The European Qualifications Framework, adopted in 2008 and revised in 2017<sup>9</sup>. It provides information about the qualifications acquired in the different educational systems and labour markets in Europe. The EQF is part of the Bologna Process, which seeks to build a common base for higher education systems in Europe. The aim of this process is to facilitate employability, as well as student and professional mobility.

The EQF defines eight levels of qualifications<sup>10</sup> on a common basis for each member country, with a description of the skills (cognitive use of logical thinking or practical use of methods and tools), knowledge (theoretical learnings) and responsibilities (ability to apply knowledge and skills autonomously) related to each level (European Commission, 2017). Degrees and professional certifications are associated with EQF levels based on the knowledge and skills they provide. EQF levels,

---

<sup>9</sup>See the Council Recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning (European Commission, 2017).

<sup>10</sup>EQF levels 1 and 2 are associated with basic knowledge and skills and correspond respectively to Kindergarten and Primary School in terms of education. They are not levels that can be valued *per se* on the labour market; thus, they are not used in this work.



the corresponding degrees and the corresponding learning outcomes are described in Table 1.7 in Appendix 1A.

For the construction of the *Skill Analysis* measure, I gather EQF levels 7 and 8 (Master's degrees and PhD, respectively) for two main reasons. First, based on the descriptions of EQF levels, it is hard to define different scores for levels 7 and 8 in terms of importance and level. Second, in most overeducation analyses, PhDs are associated with Master's degrees because there are too few PhD graduates to have a reliable sample size. Thus, it should not be of concern not to take EQF level 8 explicitly into account, even if I will not be able to consider PhD graduates working in EQF 7 occupations as overeducated.

### 1.3.2 The O\*NET database

I also rely on the O\*NET (Occupational Information Network) database which is developed by the US Department of Labor and contains detailed information about the content of jobs. It aims at allowing job seekers and recruiters to better understand the labour market, in order to help the former to find the job and the latter to find the workforce that they need. For each of the 873 occupations in the dataset, the O\*NET database provides information about activities and tasks performed, knowledge and skills required, working conditions, experience and education required, and evolution of wages and hirings. To determine whether individuals are overeducated or not, I rely on the information about the use of skills in each occupation.

35 skills are listed in the O\*NET data, grouped into seven categories: Content, Process, Social Skills, Technical Skills, Systems Skills, Resource Management Skills and Complex Problem Solving. Based on the content of occupations, each skill for each of the O\*NET occupations is rated by labour analysts on the ground of its importance (1 to 5) and of its required level to perform an occupation (1 to 7)<sup>11</sup>. Ratings are then standardised on a scale ranging from 0 to 100<sup>12</sup>. A score of 100 in terms of importance corresponds to an extremely important skill for the occupation, and a score of 0 to an unimportant skill. A score of 100 in terms of level corresponds to requiring an expert level, and a score of 0 to an unneeded skill. Even though the

---

<sup>11</sup>The analysts' selection process, as well as a more detailed overview of the notation process, are provided in Fleisher and Tsacoumis (2012).

<sup>12</sup>Standardisation is given by the equation:  $S = \frac{O-L}{H-L} \times 100$ , with  $S$  the standardised score,  $O$  the original score given by the analysts,  $L$  the lowest possible score on the rating scale and  $H$  the highest possible score on the rating scale. For instance, the importance rating scale goes from 1 to 5, so  $L = 1$  and  $H = 5$ . Thus, a standardised score  $S \geq 50$  corresponds to an original score from 3 to 5.

O\*NET database and the EQF come from different countries, the skills required to perform the tasks should not vary substantially for similar jobs, provided that there is no significant technology gap between those countries<sup>13</sup>.

I keep 20 skills out of the 35 skills available. Some skills are excluded because they are not linked to secondary or tertiary education, but to basic knowledge which are not provided by the corresponding degrees. These are the skills of the Content skill category of O\*NET<sup>14</sup>. The skill Programming is excluded because it relates to only particular fields (rather than levels) of education and occupations. Finally, I exclude most of the Social Skills and Resource Management Skills<sup>15</sup> because there is no information about them in the EQF. The 20 skills that I use in this work are reported in Table 1.8 in Appendix 1B, as well as the description provided by O\*NET for each of them.

### 1.3.3 Matching skills, degrees and occupations

Using O\*NET data and the EQF description of learning outcomes, I aim to define the theoretical education-skills requirements for each ISCO-08 occupation (observed at the 4-digits level). To this end, I proceed in two main steps: I first map each O\*NET occupation to its EQF educational requirement. Second, I match each O\*NET occupation to the corresponding ISCO occupation.

#### Matching O\*NET occupations to EQF levels

I compute the educational requirements for each O\*NET occupation according to the following process:

1. Using the indications given in the EQF, I first define whether a particular skill is provided by each educational level or not. If it is the case, I define ranges of scores (using the O\*NET standardised scores) that are provided by each EQF level. For every skill, I consider that standardised scores lower than 25

---

<sup>13</sup>Alternatively, I could rely directly on the information about the degrees required reported in the O\*NET database. However, the labour market and education structures are different in the United States than in Europe. Since these institutions strongly depend on political decisions and thus are likely to vary between countries (especially when they do not share a common educational framework), required educational levels in the USA may not be the ones required in Europe.

<sup>14</sup>The *Content* skill category includes Active Listening, Mathematics, Reading Comprehension, Science, Speaking and Writing.

<sup>15</sup>In *Social Skills*, I only keep Instructing and exclude Coordination, Negotiation, Persuasion, Service Orientation and Social Perceptiveness. In *Resource Management Skills*, I only keep Management of Personnel Resources and exclude Management of Financial Resources, Management of Material Resources and Time Management.

in terms of importance and lower than 14 in terms of level are not provided by any educational level<sup>16</sup>. The correspondence between EQF levels and skill scores is reported in Table 1.1.

2. I then retrieve from the O\*NET database the scores of level and importance of each skill in each occupation. These data are available on O\*NET Online. As for the level scores, some skills are considered “Not relevant” for some occupations because the related tasks do not require them. I affect a zero value to these occupation-skill couples, leading them not to be linked to any EQF level. I also affect a zero value when skills are stated “Not available” in the O\*NET data, i.e. when scores have low precision<sup>17</sup>.
3. Using the matches realised in the two previous steps, I define the associated EQF level for each skill in each occupation. For some occupation-skill couples, there are several degrees which can bring the required skill level. In this case, all concerned EQF levels are associated with the occupation-skill couple.
4. Finally, I compute the requirements for each skill in each occupation to determine which EQF level brings the larger amount of skills for a particular occupation. At this stage, 768 occupations (88%) are linked to only one particular EQF level, giving the educational requirements directly. For the remaining 105 occupations (12%), several EQF levels (two levels for most of the occupations, three levels for seven occupations) provide the same amount of skills. For these occupations, I keep the two required levels and consider an “upper requirement” and a “lower requirement”<sup>18</sup>.

---

<sup>16</sup>Standardised scores lower than 25 in terms of importance and lower than 14 in terms of level correspond to the minimum scores in the respective original rating scales.

<sup>17</sup>O\*NET recommends suppressing these low precision scores when conducting economic analyses. By affecting a zero value, I ensure that they will not be linked to any EQF level, and thus will not be used in the analysis.

<sup>18</sup>For the seven occupations with three competing EQF levels, I consider the highest of the three as the “upper requirement” and the lowest as the “lower requirement”.

Table 1.1: Learning outcomes expressed in skill scores by EQF level

| Skills                            | EQF 7-8                        | EQF 6  | EQF 5   | EQF 4  | EQF 3                               |
|-----------------------------------|--------------------------------|--|---|--|-------------------------------------|
| Active Learning                   | $I \geq 50$ ; $L \geq 57$      | $I \geq 50$ ; $57 > L \geq 43$                                   | $50 > I \geq 25$ ; $57 > L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Critical Thinking                 | $I \geq 25$ ; $L \geq 57$      | $I \geq 25$ ; $57 > L \geq 43$                                   | $I \geq 25$ ; $57 > L \geq 43$  | $I \geq 25$ ; $43 > L \geq 14$                                   | -                                   |
| Learning Strategies               | $I \geq 50$ ; $L \geq 43$      | $50 > I \geq 25$ ; $L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Monitoring                        | $L \geq 43$                    | $L \geq 43$  | $L \geq 43$   | $43 > L \geq 14$   | -                                   |
| Instructing                       | $I \geq 50$ ; $L \geq 57$      | $I \geq 50$ ; $57 > L \geq 43$                                   | $50 > I \geq 25$ ; $L \geq 14$<br>$I \geq 50$ ; $43 > L \geq 14$      | -  | -                                   |
| Equipment Maintenance             | -                              | -  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $29 > L \geq 14$ |
| Equipment Selection               | -                              | -  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $29 > L \geq 14$ |
| Installation                      | -                              | -  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $29 > L \geq 14$ |
| Operation and Control             | -                              | -  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 29$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $29 > L \geq 14$ |
| Operations Analysis               | $I \geq 50$ ; $L \geq 57$      | $I \geq 50$ ; $57 > L \geq 29$                                   | $I \geq 50$ ; $43 > L \geq 14$  | $50 > I \geq 25$ ; $L \geq 14$                                   | -                                   |
| Operations Monitoring             | $50 > I \geq 25$ ; $L \geq 57$ | $50 > I \geq 25$ ; $L \geq 43$                                   | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $43 > L \geq 14$                              | $50 > I \geq 25$ ; $43 > L \geq 14$ |
| Quality Control Analysis          | $I \geq 50$ ; $L \geq 43$      | $I \geq 50$ ; $L \geq 43$  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 14$<br>$I \geq 50$ ; $43 > L \geq 14$ | -                                   |
| Repairing                         | -                              | -  | $I \geq 50$ ; $L \geq 43$   | $50 > I \geq 25$ ; $L \geq 14$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $L \geq 14$      |
| Technology Design                 | $I \geq 50$ ; $L \geq 43$      | $50 > I \geq 25$ ; $L \geq 14$<br>$I \geq 50$ ; $43 > L \geq 14$ | $50 > I \geq 25$ ; $L \geq 14$<br>$I \geq 50$ ; $43 > L \geq 14$      | -  | -                                   |
| Troubleshooting                   | $50 > I \geq 25$ ; $L \geq 57$ | $I \geq 50$ ; $L \geq 43$  | $I \geq 50$ ; $L \geq 43$   | $I \geq 25$ ; $43 > L \geq 14$                                   | $I \geq 25$ ; $43 > L \geq 14$      |
| Judgment and Decision Making      | $I \geq 50$ ; $L \geq 43$      | $I \geq 50$ ; $43 > L \geq 14$                                   | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Systems Analysis                  | $I \geq 50$ ; $L \geq 43$      | $I \geq 50$ ; $43 > L \geq 14$                                   | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Systems Evaluation                | $I \geq 50$ ; $L \geq 43$      | $I \geq 50$ ; $43 > L \geq 14$                                   | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Management of Personnel Resources | $I \geq 50$ ; $L \geq 43$      | $I \geq 50$ ; $43 > L \geq 14$                                   | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |
| Complex Problem Solving           | $I \geq 50$ ; $L \geq 57$      | $I \geq 50$ ; $57 > L \geq 14$                                   | $50 > I \geq 25$ ; $L \geq 29$  | $50 > I \geq 25$ ; $29 > L \geq 14$                              | -                                   |

Source: Classification based on European Commission (2017), Annex II and O\*NET

Note:  $I$  denotes the skill score of importance,  $L$  denotes the skill score of level. “-” denotes that the skill is not provided by the EQF level.

Among the 768 occupations with unambiguous educational requirements, about 35% require EQF 7 or 8 level, 6% require EQF 6 level, 43% require EQF 5 level and 16% require EQF 4 level. None of these occupations are linked to EQF level 3. As for the 105 other occupations, about 80% have two consecutive EQF requirements (mostly EQF 6 and 5 or EQF 7 and 6).

### Matching O\*NET occupations to ISCO-08 occupations

The second step of the matching process is to map O\*NET occupations into ISCO-08 codes to obtain the EQF requirements for each ISCO occupation. Following Le Barbanchon and Rizzotti (2020), I use the Standard Occupational Classification (SOC) as a pivotal classification. I proceed as follows:

1. The correspondence between ISCO-08 and SOC provided by the International Labour Organization (ILO) uses the SOC 2010 classification, while O\*NET is based on the SOC 2018 classification. Using the crosswalk between SOC2010 and SOC2018 provided by the US Bureau of Labor Statistics, I match ISCO-08 to SOC2018. There are 441 ISCO codes for around 850 SOC codes<sup>19</sup>, matched “many-to-many”: several SOC codes correspond to the same ISCO code, but conversely, in some cases, several ISCO codes correspond to the same SOC code. After the deletion of duplicates (occupations with the same code in ISCO-08 and SOC2018, but not in SOC2010), I obtain a database of 1147 SOC-ISCO matches.
2. The O\*NET classification is based on SOC2018, the six first digits of O\*NET codes being exactly the corresponding SOC2018 code. Some O\*NET occupations are more precisely defined than in SOC, resulting in two additional numbers. Therefore, one SOC code can correspond to several O\*NET occupations, but the opposite does not exist. Using the SOC codes, I match the O\*NET occupations to ISCO codes. The matching is once again “many-to-many”: several O\*NET codes correspond to the same ISCO code, and conversely, several ISCO codes correspond to the same O\*NET code. Before any cleaning of the dataset, it contains 1312 O\*NET-ISCO matches.
3. Some ISCO codes do not have any corresponding O\*NET code. Many of these codes are armed forces occupations or political occupations, which are generally not included in analyses about overeducation. Thus, I exclude ISCO

---

<sup>19</sup>More precisely, there are 840 codes in SOC2010 and 867 in SOC2018.

categories of Armed Forces (ISCO 0) and related occupations (firefighters, police, and prison guards), political positions (ISCO 111), as well as religious occupations (ISCO 2636 and 3413). After these exclusions, 425 ISCO occupations remain, among which 18 do not have any correspondence.

4. Most ISCO codes correspond to several O\*NET occupations. However, for 235 ISCO codes, all occupations share the same educational requirements. For 45 additional ISCO codes, all EQF requirements are the same except where requirements are missing<sup>20</sup>. For these ISCO codes, which represent 66% of the ISCO occupations, I directly compute the EQF requirement based on the non-missing data.
5. When educational requirements are not the same for all occupations in the same ISCO code (126 ISCO codes), I keep the highest and lowest EQF levels and compute them to the “upper requirement” and the “lower requirement” respectively.
6. For the 18 remaining codes without corresponding O\*NET occupations, and thus no EQF level, I affect them the educational requirements of the ISCO occupations sharing the same first three digits (as in step 4) or, if all requirements are not the same, I compute the “upper requirement” and the “lower requirement” (as in step 5).

For about half of the ISCO codes (223), the upper and lower EQF requirements are the same. For the other half, using the upper requirement is likely to lead to an underestimation of overeducation, while using the lower requirement may lead to an overestimation. Indeed, the upper requirement measure is designed not to wrongly consider as overeducated individuals that are actually matched, at the expense of not taking into account a few mismatched workers. On the contrary, the lower requirement measure ensures that all overeducated workers are considered as such, at the cost of considering as overeducated some workers that are actually matched in their job. Both definitions of the requirements will be compared to the statistical measures in the next Section. Using these requirements, the *Skill Analysis* considers individuals as overeducated if their actual EQF level is higher than the (upper or lower) requirement in their particular ISCO-08 occupation. The upper requirement

---

<sup>20</sup>Most codes where requirements are missing correspond in SOC2018 to “All other” or “Not elsewhere classified” categories.

and lower requirement *Skill Analysis* measures are respectively referred to as “Upper SA” and “Lower SA” in the following<sup>21</sup>.

### 1.3.4 Advantages and drawbacks of the *Skill Analysis* measure

Using the skills requirements provided by O\*NET, the *Skill Analysis* measure associates ISCO-08 occupations to EQF levels. The international dimension of these classifications should allow to use this measure in cross-countries analyses of overeducation (within Europe). Another advantage of this measure is to use occupations rather than occupational groups. Indeed, educational requirements are more likely to vary between occupations in the same group than in precisely defined occupations (Halaby, 1994). From this perspective, this approach addresses the aggregation problem, which is one of the biggest issues of overeducation measures (especially statistical measures, for which aggregation of occupations is often required to ensure a sufficient sample size in each occupational group). Therefore, the approach that I propose measures overeducation more accurately than its statistical alternatives.

However, this measure still has limitations. First, it links skill levels for occupations and degrees without taking into account professional experience, which can bring and improve skills through on-the-job training, increasing the level of skills of individuals compared to that provided by education. Second, as it relies on the current definition of educational levels, degrees may not correspond exactly to the EQF levels under consideration for individuals who graduated a long time ago. Therefore, the *Skill Analysis* measure is more appropriate for young workers, who are anyway the most affected by overeducation (Baert et al., 2013; Dolton & Vignoles, 2000). Third, for some occupations, several educational levels appear as “normal”, raising the question of which one is the most appropriate. This drawback is partially managed by making a difference between a lower and an upper educational requirement for these occupations. Fourth, as for all objective and statistical overeducation measures, it is based on the skills that educational levels should theoretically (according to the EQF) bring to graduates, without information on which skills they actually provide nor on the heterogeneity between degrees at the same level.

---

<sup>21</sup>The correspondence table between occupations and educational levels according to the *Skill Analysis* measure can be provided upon request.

## 1.4 Empirical comparison of overeducation measures

### 1.4.1 Data and statistical measures

In order to explore differences in overeducation measures, I use data from the French Labour Force Survey (LFS) conducted by the French National Institute of Statistics for year 2021<sup>22</sup>. The LFS is a large household survey conducted in European countries and built to ensure comparability between countries, allowing both short-term and structural analyses of the labour market. This dataset provides detailed information on individual socioeconomic (gender, age, educational level, etc.) and job characteristics (wage, occupation, hours worked, etc.). I am then able to compute individuals' EQF levels and link them to the requirements in their particular ISCO-08 occupation.

I exclude from my sample Armed forces, Political and religious occupations, as well as self-employed workers. The sample includes 116,512 employed individuals aged 15 to 64 years old, and 20,058 are younger than 30. Among young workers, 48% have a higher education degree: around 16% graduated at each tertiary EQF level (EQF levels 5, 6 and 7-8). The other half of the sample has at most a high school diploma.

Similarly to (Capsada-Munsech, 2019), I compute three statistical overeducation measures in order to compare them to the *Skill Analysis* measures (SA)<sup>23</sup>. All measures to be compared are summarised and described in Table 1.2. For statistical measures, I compute overeducation rates at the ISCO-08 2-digits level. Indeed, I am not able to use the 4-digits level due to the small number of observations in each code at this level, not allowing to calculate the statistics accurately. The first statistical measure (“Mode”) is based on the mode of the distribution: I consider an individual as overeducated if her or his EQF level is higher than the modal EQF level in her or his particular ISCO occupation. Using ISCO codes at the 2-digits level implies the aggregation of several occupations that are sometimes very different and thus have various educational requirements. Therefore, considering only one level as the “normal” requirement does not allow to take into account the variety

<sup>22</sup>At the time of writing, LFS data are not available for more recent years than 2021. Earlier versions of this paper mobilised data from LFS 2018 and 2019, which provided similar results.

<sup>23</sup>Unfortunately, the LFS data do not allow me to compare them to subjective approaches, as there is no information about workers' assessment about the level required to do (or to get) their job or their feeling of being overeducated.



of the actual requirements. Instead, I compute a second measure (“2-Mode”) that considers as “normal” the two most attained educational levels in the occupation (rather than only one in the previous measure). The third measure (“Q3”) is based on the deviation from the median of the distribution. The idea behind this measure is that there is an educational mismatch for the degrees farthest from the median. Using this indicator, I consider that individuals are overeducated if their EQF level is higher than the third quartile of the distribution in their ISCO-08 2-digit code.

Table 1.2: Overeducation measures

| <b>Measure</b>  | <b>Type</b> | <b>Overeducated if...</b>  | <b>Aggregation</b> |
|-----------------|-------------|--|--------------------|
| <b>Upper SA</b> | Objective   | educational level above the highest “normal” level as defined in Section 4       | ISCO 4-digits      |
| <b>Lower SA</b> | Objective   | educational level above the lowest “normal” level as defined in Section 4        | ISCO 4-digits      |
| <b>Mode</b>     | Statistical | educational level above the most common in the occupation                        | ISCO 2-digits      |
| <b>2-Mode</b>   | Statistical | educational level above the highest of the two most common in the occupation     | ISCO 2-digits      |
| <b>Q3</b>       | Statistical | educational level above the third quartile of the distribution in the occupation | ISCO 2-digits      |

Note: SA stands for *Skill Analysis* measure.

## 1.4.2 Overeducation incidence across measures

Table 1.3 shows the differences in overeducation rates depending on the type of measure under consideration. As expected, there are large disparities in overeducation rates between measures. For the whole sample, the incidence of overeducation range from 6.86% (upper *Skill Analysis*-SA) to 26.44% (Mode). Unsurprisingly, the upper requirement objective measure presents the lowest rates regardless of the sample that I consider. It is due to the fact that it is, by design, the most parsimonious of the measures that I explore. On the contrary, the Mode measure leads to the highest overeducation rates, because of the restrictive definition of the “normal” educational requirements.

Regarding the gender differences in overeducation rates, women are more often mismatched than men, whatever measure is used. However, it must be noted that the gap between men and women is larger using the 2-Mode measure than the alternatives. Given that the 2-Mode measure adds a second “normal” educational level for each occupation, which is for about half of the occupations higher than

the first one, this may illustrate the tendency for women to overinvest in education relative to men, in order to compensate for their higher risk of being discriminated on the labour market (Becker et al., 2010). Overall, the 2-Mode measure shows overeducation rates similar to the lower SA (except for men), while the Q3 measure provides lower ones.

Table 1.3: Overeducation rates by measure

| Measure                            | Upper SA | Lower SA | Mode   | 2-Mode | Q3     |
|------------------------------------|----------|----------|--------|--------|--------|
| Whole sample<br>( $N = 116512$ )   | 6.86%    | 17.80%   | 26.44% | 15.40% | 10.40% |
| 15-29 year-olds<br>( $N = 20058$ ) | 10.43%   | 23.39%   | 41.79% | 22.63% | 16.55% |
| Men<br>( $N = 55591$ )             | 5.98%    | 17.75%   | 24.73% | 13.69% | 9.81%  |
| Women<br>( $N = 60921$ )           | 7.67%    | 17.86%   | 27.99% | 16.97% | 10.94% |

Source: Author's calculation based on French LFS.

Note: SA stands for *Skill Analysis* measure.

Because, by design, the *Skill Analysis* measure is more accurate for young workers (see Section 1.3.4.), Table 1.4 displays the overeducation rates for each measure, with detailed rates reported by occupations and EQF levels. Overall, overeducation is higher for young workers, regardless of the measure. The upper SA leads to lower overeducation incidence than statistical measures, while the lower SA provides overeducation rates similar to the ones observed with the 2-Mode. For several ISCO 1 and 2 occupations, I observe overeducation rates equal to zero, especially using statistical indicators. The first plausible explanation is the aggregation problem: while *Skill Analysis* measures are computed using precise occupations, statistical ones are computed on the ISCO 2-digits occupational groups to ensure a sufficient sample size. Thus, the variation in educational requirements between occupations in the same ISCO code may be hidden behind this aggregation, which is not the case for the *Skill Analysis* measures. The other (but not exclusive of the first) explanation is the specific bias of statistical measures, which comes from the design of these measures itself, and can be expressed as “if everybody in an occupation is overeducated, nobody is”. In that case, the results reflect hiring standards that are higher than the real need for skills in the occupation. As I focus on young workers, this questions the educational system, which appears not to provide the skills required on the labour market, at least for the concerned occupations.

Table 1.4: Overeducation rates by measure - 15-29 y-o

| Measure   | Upper SA | Lower SA | Mode   | 2-Mode | Q3     |
|---|----------|----------|--------|--------|--------|
| All occupations and EQF levels (N = 20058)  | 10.43%   | 23.39%   | 41.79% | 22.63% | 16.55% |
| All occupations, EQF levels 4 to 8 (N = 15114)  | 13.84%   | 31.03%   | 55.43% | 30.02% | 21.95% |
| Men (N = 10339)   | 8.86%    | 22.09%   | 40.52% | 21.17% | 15.04% |
| Women (N = 9719)  | 12.10%   | 24.76%   | 43.10% | 24.16% | 18.13% |
| <b>Occupations</b>  |          |          |        |        |        |
| 12 Administrative and Commercial Managers (N = 137)   | 0        | 11.68%   | 0      | 0      | 0      |
| 13 Production and Specialized Services Managers (N = 92)                                      | 0        | 1.09%    | 0      | 0      | 0      |
| 14 Hospitality, Retail and Other Services Managers (N = 90)                                   | 3.33%    | 6.67%    | 52.22% | 27.78% | 27.78% |
| 21 Science and Engineering Professionals (N = 756)  | 8.73%    | 19.84%   | 0      | 0      | 0      |
| 22 Health Professionals (N = 288)   | 0        | 2.43%    | 0      | 0      | 0      |
| 23 Teaching Professionals (N = 529)   | 0.76%    | 13.99%   | 71.08% | 0      | 0      |
| 24 Business and Administration Professionals (N = 1195)                                       | 7.03%    | 60.25%   | 0      | 0      | 0      |
| 25 Information and Communications Technology Professionals (N = 565)                          | 40.53%   | 76.99%   | 0      | 0      | 0      |
| 26 Legal, Social and Cultural Professionals (N = 418)   | 1.91%    | 29.90%   | 0      | 0      | 0      |
| 31 Science and Engineering Associate Professionals (N = 567)                                  | 6.70%    | 40.56%   | 32.63% | 32.63% | 32.63% |
| 32 Health Associate Professionals (N = 823)   | 4.62%    | 15.31%   | 56.62% | 1.46%  | 1.46%  |
| 33 Business and Administration Associate Professionals (N = 1239)                             | 9.69%    | 41.00%   | 44.71% | 44.71% | 18.72% |
| 34 Legal, Social, Cultural and Related Associate Professionals (N = 926)                      | 13.61%   | 22.46%   | 36.61% | 36.61% | 23.00% |
| 35 Information and Communications Technicians (N = 180)                                       | 10.56%   | 39.44%   | 32.78% | 32.78% | 32.78% |
| 41 General and Keyboard Clerks (N = 608)  | 35.53%   | 35.53%   | 59.21% | 59.21% | 35.53% |
| 42 Customer Services Clerks (N = 414)   | 26.09%   | 29.71%   | 44.20% | 44.20% | 28.99% |
| 43 Numerical and Material Recording Clerks (N = 577)  | 23.92%   | 38.47%   | 56.67% | 33.10% | 33.10% |
| 44 Other Clerical Support Workers (N = 264)   | 48.11%   | 50.38%   | 60.61% | 60.61% | 48.11% |
| 51 Personal Services Workers (N = 1116)   | 7.17%    | 14.78%   | 62.81% | 18.10% | 18.10% |
| 52 Sales Workers (N = 1663)   | 13.11%   | 21.71%   | 25.02% | 25.02% | 25.02% |
| 53 Personal Care Workers (N = 897)  | 1.90%    | 5.25%    | 54.85% | 9.03%  | 9.03%  |
| 54 Protective Services Workers (N = 146)  | 0.68%    | 10.96%   | 59.59% | 16.44% | 16.44% |
| 61 Market-oriented Skilled Agricultural Workers (N = 442)                                     | 4.30%    | 17.42%   | 55.20% | 20.81% | 20.81% |
| 62 Market-oriented Skilled Forestry, Fishery and Hunting Workers (N = 38)                     | 10.53%   | 15.79%   | 57.89% | 15.79% | 15.79% |
| 71 Building and Related Trades Workers (excluding Electricians) (N = 850)                     | 3.06%    | 4.35%    | 39.53% | 6.59%  | 6.59%  |
| 72 Metal, Machinery and Related Trades Workers (N = 841)                                      | 1.43%    | 11.77%   | 64.21% | 21.28% | 21.28% |
| 73 Handicraft and Printing Workers (N = 72)   | 6.94%    | 12.50%   | 70.83% | 16.67% | 16.67% |
| 74 Electrical and Electronic Trades Workers (N = 302)   | 0.66%    | 6.95%    | 69.87% | 25.50% | 25.50% |
| 75 Food Processing, Woodworking, Garment and Other Craft and Related Trades Workers (N = 553) | 12.30%   | 12.84%   | 44.67% | 18.44% | 18.44% |
| 81 Stationary Plant and Machine Operators (N = 455)   | 2.86%    | 4.84%    | 53.85% | 12.31% | 12.31% |
| 82 Assemblers (N = 134)   | 2.99%    | 8.21%    | 54.48% | 16.42% | 16.42% |
| 83 Drivers and Mobile Plant Operators (N = 679)   | 7.81%    | 9.72%    | 52.43% | 52.43% | 10.60% |
| 91 Cleaners and Helpers (N = 500)   | 7.00%    | 7.00%    | 52.80% | 52.80% | 52.80% |
| 92 Agricultural, Forestry and Fishery Labourers (N = 204)                                     | 3.92%    | 17.16%   | 57.35% | 57.35% | 17.16% |
| 93 Labourers in Mining, Construction, Manufacturing and Transport (N = 1118)                  | 13.86%   | 17.26%   | 63.15% | 17.26% | 17.26% |
| 94 Food Preparation Assistants (N = 280)  | 12.86%   | 12.86%   | 60.36% | 60.36% | 12.86% |
| 96 Refuse Workers and Other Elementary Workers (N = 90)                                       | 13.33%   | 13.33%   | 50.00% | 50.00% | 13.33% |
| <b>EQF levels<sup>a</sup></b>   |          |          |        |        |        |
| Levels 7-8 (N = 2981)   | 26.78%   | 63.35%   | 39.11% | 25.67% | 25.67% |
| Level 6 (N = 3312)  | 37.71%   | 72.18%   | 72.39% | 56.89% | 45.90% |
| Level 5 (N = 3289)  | 8.71%    | 29.19%   | 56.69% | 50.91% | 37.53% |
| Level 4 (N = 7926)  | 0        | 0.08%    | 54.31% | 11.42% | 3.43%  |

Source: Author's calculation based on French IFS.

Note: SA stands for *Skill Analysis* measure.<sup>a</sup> Individuals without any degree (N = 1178) cannot be overeducated. By design, individuals with EQF level 3 (N = 3766) cannot be overeducated either.

Regardless of the measure, EQF level 6 graduates are more often overeducated than others. In France, this level corresponds to degrees which are designed to give access to longer studies, but not to enter the labour market directly<sup>24</sup>. Therefore, this finding is in line with Goux and Maurin (1994), who found that years of education without formal certification had lower returns on the labour market than degree-validated educational credentials.

Table 1.5: Correspondence between *Skill Analysis* and statistical measures - 15-29 y-o

|                 | Mode              |                  | 2-Mode            |                  | Q3                |                  |
|-----------------|-------------------|------------------|-------------------|------------------|-------------------|------------------|
|                 | Matched           | Overeducated     | Matched           | Overeducated     | Matched           | Overeducated     |
| <b>Upper SA</b> |                   |                  |                   |                  |                   |                  |
| Matched         | 11279<br>(56.26%) | 6677<br>(33.31%) | 15090<br>(75.27%) | 2866<br>(14.30%) | 16246<br>(81.04%) | 1710<br>(8.53%)  |
| Overeducated    | 391<br>(1.95%)    | 1701<br>(8.48%)  | 421<br>(2.10%)    | 1671<br>(8.33%)  | 485<br>(2.42%)    | 1607<br>(8.02%)  |
| <b>Lower SA</b> |                   |                  |                   |                  |                   |                  |
| Matched         | 10072<br>(50.24%) | 5286<br>(26.37%) | 13768<br>(68.68%) | 1590<br>(7.93%)  | 14694<br>(73.29%) | 664<br>(3.31%)   |
| Overeducated    | 1598<br>(7.97%)   | 3092<br>(15.42%) | 1743<br>(8.69%)   | 2947<br>(14.70%) | 2037<br>(10.16%)  | 2653<br>(13.23%) |

Source: Author's calculation based on French LFS.

Note: SA stands for *Skill Analysis* measure.

The computation of overeducation rates by ISCO 2-digits codes reveals strong disparities between occupations (Table 1.4). While some occupations hire less than 5% of overeducated workers, others employ between a third and a half (and, in rare cases, more than a half) of their workforce below their educational level. Most of these occupations with higher overeducation are intermediate professions (associate professionals and clerks), regardless of the measure. As predicted by the *Job Competition Model* (Thurow, 1975) and *assignment models* (Sattinger, 1993), this suggests an overinvestment in education in order to gain access to better positions in more saturated sectors. Indeed, most of the concerned occupations are in the services, law and administration sectors. This is also in line with Sellami et al. (2020), who find that graduates in Economics, Business and Law were among the most at risk of being overeducated.

By design, the upper SA measure systematically underestimates overeducation and the lower SA measure systematically overestimates it. Therefore, the Q3 measure, which reveals an intermediate incidence of the phenomenon, appears to be

<sup>24</sup>EQF level 6 corresponds in France to the degrees *Licence* and *Maîtrise*, which are intermediate steps to Master's degree. The exception is the *Licence professionnelle*, which has a vocational purpose. However, this particular degree represents less than 5% of the students in 2021 (French Ministry of Higher Education, *MESRI-SIES*).

more appropriate than the alternative statistical measures. However, while the estimated incidence seems accurate, the Q3 might not consider exactly the right individuals as overeducated, since statistical measures take larger occupational groups into account. This may raise some concerns when investigating the drivers and/or consequences of the phenomenon. When looking at the overlaps between these measures (Table 1.5), 10% to 15% of the workers are mismatched in only one of the two dimensions. The correspondence is as expected: crossing the upper SA and the Q3 measures, 8.53% of the workers are statistically but not objectively overeducated, against 2.42% that are objectively but not statistically overeducated. I find an opposite correspondence between the Q3 and the lower SA measures: 3.31% of the workers are overeducated statistically but not objectively, and 10.16% objectively but not statistically. The Q3 measure also presents the highest correlations with the *Skill Analysis* measures (Table 1.6), even though the coefficients remain moderate (between 0.55 and 0.60), which tends to confirm a better, although limited, reliability of the Q3 measure (compared to other statistical measures) as an alternative to objective approaches.

Table 1.6: Correlation between overeducation measures - 15-29 y-o

|          | Upper SA | Lower SA | Mode   | 2-Mode | Q3 |
|----------|----------|----------|--------|--------|----|
| Upper SA | 1        |          |        |        |    |
| Lower SA | 0.6177   | 1        |        |        |    |
| Mode     | 0.2735   | 0.2704   | 1      |        |    |
| 2-Mode   | 0.4670   | 0.5310   | 0.6383 | 1      |    |
| Q3       | 0.5536   | 0.5952   | 0.5255 | 0.8233 | 1  |

Source: Author's calculation based on French LFS.

Note: SA stands for *Skill Analysis* measure.

In line with Capsada-Munsech (2019) and Verhaest and Omey (2006), correlations tend to be higher when considering measures of the same type, *i.e.* between the three statistical measures on the one hand and between the two versions of the objective measure on the other hand, as reported in Table 1.6. On the contrary, correlations between objective and statistical measures are lower, except for the Q3 measure. This suggests that these different measures do not reveal the same phenomenon, but rather different dimensions of mismatch.

## 1.5 Conclusion

The aim of this work was to develop a new objective measure of overeducation in Europe, based on skills requirements in occupations. Due to the cost of keeping them consistent with the evolutions of the labour market structures, objective measures are missing in several countries, or are based on old occupational dictionaries, which raises concern regarding the accuracy of the measure (van der Meer, 2006). In this respect, I propose an up-to-date measure of overeducation, similar to what O\*NET offers for the US but based on the recommendations of the European Commission (2017), where the educational level(s) needed for each occupation is defined based on the skills required to do the job. Moreover, the use of precise occupations rather than larger occupational groups deals with the aggregation problem that arises with the statistical alternatives.

The empirical comparison, using data from the French Labour Force Survey, tends to confirm that different measures are related to different matters. The different approaches only partially overlap: I find that a significant share of workers is statistically but not objectively overeducated. The opposite is also true for some workers, but to a lesser extent. The results suggest that in the absence of an adequate objective measure a measure based on the median and quartiles of the distribution is a better alternative than other statistical indicators, with higher correlations and overlaps with the *Skill Analysis* measure. Overall, the upper *Skill Analysis* measure lead to a lower overeducation incidence, while the lower *Skill Analysis* is closer to statistical measures.

## Appendices

### Appendix 1A: Description of EQF levels

Table 1.7: EQF levels with corresponding degrees and learning outcomes

| EQF level | Degrees   | Knowledge, skills and responsibilities   |
|-----------|---|--|
| Level 8   | Ph.D.   | Mobilise the most advanced knowledge, skills and techniques to solve critical and complex problems<br>Extend and redefine existing knowledge and professional practice<br>Demonstrate substantial authority, innovation, autonomy and professional integrity<br>Develop new ideas or processes at the forefront of work or study contexts  |
| Level 7   | Master's degree, Engineering degree, Graduate education (5-year higher education) | Highly specialised knowledge<br>Critical awareness of knowledge issues in a field and at the interface between different fields<br>Specialised problem-solving skills in order to develop new knowledge and procedures<br>Manage and transform work or study contexts that are complex, unpredictable and require new strategic approaches<br>Take responsibility for contributing to professional knowledge and practice<br>Review the strategic performance of teams |
| Level 6   | Bachelor's degree, Undergraduate education (3- to 4-year higher education)        | Advanced knowledge of a field of work and study<br>Critical understanding of theories and principles<br>Advanced skills required to solve complex and unpredictable problems in a specialised field of work or study<br>Manage complex activities or projects<br>Take responsibility for decision-making in unpredictable contexts<br>Manage professional development of individuals and groups  |
| Level 5   | Vocational university degree, Undergraduate education (2-year higher education)   | Comprehensive specialised knowledge within a field of work or study, and awareness of the boundaries of this knowledge<br>Comprehensive range of cognitive and practical skills required to develop solutions to abstract problems<br>Exercise management and supervision in contexts of work or study with unpredictable change<br>Review and develop performance of self and others  |
| Level 4   | 3-year high school degree (general or vocational)                                 | Knowledge in broad contexts within a field of work or study<br>Cognitive and practical skills required generate solutions to specific problems in a field work or study<br>Self-management within the guidelines of contexts that are usually predictable<br>Supervise the routine work of others<br>Take some responsibility for the evaluation and improvement of work and study activities  |
| Level 3   | 2-year high school degree (vocational)  | Knowledge of facts, principles processes and general concepts in a field of work and study<br>Cognitive and practical skills to accomplish tasks and solve problems by selecting and applying basic methods and tools<br>Take responsibility for completion of tasks<br>Adapt own behaviour to circumstances in solving problems   |

Source: European Commission (2017), Annex II

## Appendix 1B: Description of O\*NET skills

Table 1.8: O\*NET skills

| Skills                                | Description  |
|---------------------------------------|--|
| <i>Process</i>                        |  |
| Active Learning                       | Understanding the implications of new information for both current and future problem-solving and decision-making                                      |
| Critical Thinking                     | Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions, or approaches to problems                    |
| Learning Strategies                   | Selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things                   |
| Monitoring                            | Monitoring/Assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action                       |
| <i>Social Skills</i>                  |  |
| Instructing                           | Teaching others how to do something  |
| <i>Technical Skills</i>               |  |
| Equipment Maintenance                 | Performing routine maintenance on equipment and determining when and what kind of maintenance is needed  |
| Equipment Selection                   | Determining the kind of tools and equipment needed to do a job   |
| Installation                          | Installing equipment, machines, wiring, or programs to meet specifications   |
| Operation and Control                 | Controlling operations of equipment or systems   |
| Operations Analysis                   | Analyzing needs and product requirements to create a design  |
| Operations Monitoring                 | Watching gauges, dials, or other indicators to make sure a machine is working properly   |
| Quality Control Analysis              | Conducting tests and inspections of products, services, or processes to evaluate quality or performance  |
| Repairing                             | Repairing machines or systems using the needed tools   |
| Technology Design                     | Generating or adapting equipment and technology to serve user needs  |
| Troubleshooting                       | Determining causes of operating errors and deciding what to do about it  |
| <i>Systems Skills</i>                 |  |
| Judgment and Decision Making          | Considering the relative costs and benefits of potential actions to choose the most appropriate one  |
| Systems Analysis                      | Determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes                               |
| Systems Evaluation                    | Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system |
| <i>Resource Management Skills</i>     |  |
| Management of Personnel Resources     | Motivating, developing, and directing people as they work, identifying the best people for the job   |
| <i>Complex Problem Solving Skills</i> |  |
| Complex Problem Solving               | Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions                                 |

Source: O\*NET Online Documentation



## Chapter 2

# Local labour markets and spatial determinants of overeducation

This chapter aims to explore the geographical dimension of overeducation and the influence of the local economic context. To this end, I mobilise explore spatial econometrics and spatial regimes models using data from the French population census for years 2009, 2013 and 2017. I find that the local labour market context impacts overeducation, but that there is strong differences between urban and rural areas. Interactions between neighbouring labour markets have limited effects on overeducation, affecting only men. More generally, women appears to be more impacted by the labour market context than men, but at a more local level.

**JEL classification:** C33, J24, R23

**Keywords:** overeducation, educational mismatch, local labour markets, spatial heterogeneity, spatial econometrics models

---

I would like to thank participants in the 19th International Workshop on Spatial Econometrics and Statistics, the 17th annual conference of the research federation CNRS TEPP, the 38th *Journées de Microéconomie Appliquée* and the LEMNA and NZWRI seminars for their comments on earlier versions of this chapter.

# Dimension locale du marché du travail et déterminants spatiaux de la sur-éducation

Ce chapitre vise à explorer la dimension géographique de la suréducation et l'influence du contexte économique local. Pour ce faire, je mobilise des modèles d'économétrie spatiale et de régimes spatiaux sur les données du recensement de population pour les années 2009, 2013 et 2017. Je montre que le contexte du marché du travail local a un impact sur la sur-éducation, mais qu'il existe de fortes différences entre les zones urbaines et rurales. Les interactions entre marchés du travail voisins ont des effets limités sur la suréducation, affectant principalement les hommes. Plus généralement, les femmes semblent être plus touchées par le contexte du marché du travail que les hommes, mais à un niveau plus local.

**Mots-clés :** sur-éducation, inadéquation emploi-diplôme, marchés du travail locaux, hétérogénéité spatiale, modèles d'économétrie spatiale

## 2.1 Introduction

Some workers have an educational level higher than required for their job: they are “overeducated” (Kucel, 2011; McGuinness, 2006). One popular explanation for overeducation is the large increase in educational attainment over the last decades in most advanced countries (Groot & Maassen van den Brink, 2000; Verhaest & van der Velden, 2013). Another explanation (not exclusive to the first one) is that overeducation results from search and matching frictions (Dolado et al., 2009; Gautier, 2002; Navarini & Verhaest, 2023). From this perspective, this questions the impact of labour market institutions, context and geography on educational mismatch (Davia et al., 2017).

While in the general labour economics framework, individuals are searching for jobs on a global market, most workers actually have limited spatial flexibility, and search for jobs around their residential locations, on local labour markets (Manning & Petrongolo, 2017). This implies some trade-offs regarding their labour market position. Simpson (1992) explains that, if a job seeker is on a local labour market with no suitable jobs available, she or he has three options. The first is not to participate in the market, *i.e.* to stay unemployed. The second is to accept a job for which she or he is overeducated. The third is to geographically broaden her or his job search. On this last point, empirical evidence confirms that individual spatial mobility helps reducing overeducation (see e.g. Büchel & van Ham, 2003; Venhorst & Cörvers, 2018). However, this effect strongly varies depending on the departure and arrival regions (Devillanova, 2013; Iammarino & Marinelli, 2015).

This paper contributes to the literature by addressing the geographical dimension of overeducation and exploring the influence of the local economic context. This investigation has a threefold objective. First, I want to determine the extent to which the local labour market context influences overeducation. While some papers study spatial variations in overeducation rates (Croce & Ghignoni, 2012; Davia et al., 2017; Ghignoni & Verashchagina, 2014), they focus on disparities between countries or large regions. However, Manning and Petrongolo (2017) show that labour markets are actually much more local, with a strong discouraging effect of distance on job search. I address this issue by conducting the analysis at the scale of the French employment areas<sup>1</sup>. Moreover, I explore the heterogeneous effects

---

<sup>1</sup>Employment areas are defined by the French National Institute of Statistics and Economic Studies (Insee) as areas where most workers both live and work, and where firms can find most of their workforce. They are especially designed to study local labour markets. See Section 2.3.1 for a detailed description of this geographical division.

of the local context between urban and rural areas, as higher skilled jobs are not evenly distributed between areas (Détang-Dessendre, 1999), and workers tend to be attracted by cities (Murphy et al., 1991).

Second, similar to what has been evidenced for unemployment (Patacchini & Zenou, 2007), I want to verify if a local labour market's situation influences overeducation on neighbouring markets. According to Simpson (1992), overeducation and unemployment are alternative situations, meaning that a similar mechanism might lie behind both these outcomes. Besides, Manning and Petrongolo (2017) show that local labour markets are not precisely defined divisions, but overlapping areas, and that shocks impact labour markets at a much broader scale than their own area. Individual behaviours, especially the fact that individuals “flee” less dynamic labour markets (Langella & Manning, 2022) and that spatial mobility decreases overeducation (Büchel & van Ham, 2003; Venhorst & Cörvers, 2018), might also affect the interactions between areas.

Third, I explore whether these geographical factors have the same impacts on overeducation across genders. The seminal paper of Frank (1978) regarding the spatial dimension of overeducation explains that women face a higher risk of overeducation because of their limited spatial mobility. Although the mechanisms behind this are somewhat dated<sup>2</sup>, some studies show disparities in overeducation between men and women, mostly due to familial constraints (Büchel & Battu, 2003; Büchel & van Ham, 2003). If this is true, we can expect the local context to affect more strongly women than men.

To explore these matters, I estimate spatial econometrics and spatial regimes models using data from the French population census, that allow me to calculate overeducation rates at the scale of the French employment areas, for years 2009, 2013 and 2017. Three main results emerge from this analysis. First, I find that local characteristics influence overeducation, but that the effects vary between urban and rural areas. Several factors, such as the structure of employment by sector or the proximity with a foreign country, play a significant role only in rural areas, while the imbalance between supply and demand of skilled labour is the main driver in urban areas. Second, evidence about the influence of neighbouring areas on educational mismatch is mixed. While it appears to have an effect in 2009, it seems not to be the case anymore in 2013 and 2017. However, distinguishing between genders, I show

---

<sup>2</sup>According to this “theory of differential overqualification”, in a couple, it is the husband who looks for a job on the global labour market, whereas the wife looks for a job only on the local labour market of the area where her partner has obtained work.

that there is spatial dependence for men overeducation, but not for women. This leads to the third result: the local labour market context has heterogeneous effects across gender. Women overeducation is impacted by the labour market structure at the local level, but is less affected by neighbouring areas, while I observe the opposite for men. This somehow supports the idea of a “differential overeducation” (Frank, 1978) across gender.

Section 2.2 provides a review of previous works about macro-level drivers of overeducation and the influence of spatial variables. Section 2.3 describes the data and variables. Section 2.4 explains the empirical strategy and spatial econometrics models. Section 2.5 presents and discusses the results. Section 2.6 concludes.

## 2.2 Overeducation, local and spatial drivers

### 2.2.1 Literature review

For a long time, overeducation was only studied at the individual level, with limited interest in its macro-level drivers. Still, a few papers study the effects of some area level variables, such as regional unemployment (Büchel & van Ham, 2003; Croce & Ghignoni, 2012; Guironnet, 2006), economic dynamism (Croce & Ghignoni, 2012; Tselios, 2013) or investment (Ghignoni & Verashchagina, 2014; McGuinness et al., 2018), but find contradictory results<sup>3</sup>. As for unemployment, Guironnet (2006) find a positive effect, which he associates to an indicator of poor health of the labour market, while Davia et al. (2017) evidence a negative effect, that could reflect the trade-off at the individual level between overeducation and unemployment. Büchel and van Ham (2003) do not find any significant effect. McGuinness et al. (2018) show that investment in R&D does not affect overeducation, while Ghignoni and Verashchagina (2014) present it as a lever to reduce mismatch. In the time dimension, Croce and Ghignoni (2012) find that overeducation is lower in more dynamic periods; in the space dimension, Tselios (2013) show that mismatch is higher in dynamic regions because competition between workers is higher. In a cross-countries analysis, Ghignoni and Verashchagina (2014) explain that those disparities come from differences of institutional and economic contexts between countries.

---

<sup>3</sup>There is also a substantial literature about the effect of the employment legislation protection (EPL) on overeducation (see e.g. Davia et al., 2017; Di Pietro, 2002; Verhaest & van der Velden, 2013). However, since my focus is on local disparities at the sub-national level, I will not detail this literature that applies mostly for cross-countries comparisons.

A widespread explanation for spatial differences in overeducation rates is related to educational attainment. The rise of mismatch is often associated to the large increase in the educational level in advanced countries in the last fifty years (see e.g. Groot & Maassen van den Brink, 2000; Hartog, 2000; Verhaest & van der Velden, 2013). Davia et al. (2017) and Croce and Ghignoni (2012) show that a higher number of graduates does not *per se* generate overeducation, but that an excess supply of skilled labour (relative to the demand) in an area explains disparities in the incidence of overeducation between regions. Indeed, educated workers are not evenly distributed across the spatial dimension (Murphy et al., 1991), neither are jobs (Détang-Dessendre, 1999). Verhaest and van der Velden (2013) show that the distribution of graduates in terms of fields of study also impacts overeducation. Overall, Ghignoni and Verashchagina (2014) show that supply factors are more important in lagging countries, while public policy levers related to the demand are more relevant in more advanced countries.

In addition to these macro-level drivers, the geography of labour markets influences overeducation through the individual behaviours of workers. Many studies show that spatial mobility is a way to avoid overeducation (see e.g. Büchel & van Ham, 2003; Jauhiainen, 2011; Venhorst & Cörvers, 2018), but there is also evidence of large disparities in the effect depending on the regions (Devillanova, 2013; Iammarino & Marinelli, 2015). In particular, the effect is stronger for mobility towards important economic centers (Fouquet & Sari, 2023), which offer a larger pool of job opportunities (Combes et al., 2015). The decision to move is not trivial, as mobility involves important psychological and financial costs (Greenwood, 1997). Still, young workers are more prone to mobility than other age groups (Marinescu & Rathelot, 2018), although Bernela and Bonnal (2022) show that post-graduation mobility, while increasing over time, remains quite rare. Higher skilled workers also tend to be more mobile (Combes et al., 2012). Theoretically, this higher mobility of skilled and young workers has an ambiguous effect on how the local labour market context influences overeducation. On the one hand, mobile individuals might be less affected by the local characteristics of an area, as they can relocate in others; on the other hand, their mobility impacts supply-side characteristics of the area, that themselves might affect overeducation. The expected results might be even more inconclusive since the location decision can be related to other reasons than labour market ones, especially familial reasons (Bernela & Bonnal, 2022) or natural amenities (Power, 2005; Roback, 1982).

Extending the focus beyond overeducation, there is evidence of effects of the local context on the labour market, especially unemployment. In a meta-analysis, Elhorst (2003) show an inverse relationship between unemployment and educational attainment. He also evidences significant effects of the industry mix, although the sign remains ambiguous. Rios (2017) find positive effects of economic dynamism (growth of GDP) and sectoral diversification on the reduction of unemployment. Cracolici et al. (2007) find no significant effects of demographic variables and amenities on unemployment, but show a spatial dependence of unemployment rates between neighbouring areas. This is also what predicts the model of Patacchini and Zenou (2007), the main reason for these spatial spillovers being the mobility of workers between areas. To some extent, Elhorst (2003) confirms this, showing a positive effect of both commuting and residential migration on unemployment rates.

## 2.2.2 Hypotheses

This paper aims to assess the link between local labour markets, their geography and overeducation. Four hypotheses are to be tested.

*H1: Overeducation is impacted by the local labour market context.*

There is evidence of the impact of macro-level drivers on overeducation at the scale of countries or large regions (see e.g. Croce & Ghignoni, 2012; Davia et al., 2017; Tselios, 2013). However, labour markets are generally smaller and limited at a local level (Manning & Petrongolo, 2017). I thus verify if the local characteristics also affect overeducation at the infra-national scale of employment areas.

*H2: Overeducation on a local labour market is affected by neighbouring labour markets.*

Patacchini and Zenou (2007) evidence a strong spatial dependence in unemployment rates. Since overeducation and unemployment are alternative situation (Simpson, 1992), the aim is to determine if a similar dependence can be observed for overeducation. Besides, local labour markets are not precisely defined geographical divisions, but overlapping areas (Manning & Petrongolo, 2017), which might also imply interactions between areas as defined in the data.

*H3: Area level drivers of overeducation vary between urban and rural areas.*

There are large disparities between urban and rural areas both in terms of educational attainment (Rodríguez-Pose & Tselios, 2011) and of skilled job opportunities (Détang-Dessendre, 1999). Urban areas also attract more workers (Murphy et al., 1991). Moreover, overeducation is affected by the size of the labour market, which

is also related to urbanisation (Büchel & van Ham, 2003). Then, one can expect overeducation to be driven by different local factors depending on the degree of urbanisation of the area.

*H4: The local context and the geography of labour markets does not affect men and women overeducation in the same way.*

According to Frank (1978), labour markets are more local for women than for men, who have less difficulties to move to jobs located farther. The local context of the labour market might therefore have a stronger effect on women than on men overeducation.

## 2.3 Data and variables

### 2.3.1 Data

I explore overeducation at the scale of local labour markets in order to investigate the role of local characteristics and spatial interactions between neighbouring areas. To this end, I rebuild a large set of variables at the French employment areas level. These geographical divisions are defined by Insee as areas where most workers both live and work, and where firms can find most of their workforce. These areas are defined from commuting flows between municipalities, each municipality being included in only one employment area and indivisible, even when large. Employment areas are composed of neighbouring municipalities; while their population cannot be less than 5000, it should not be too large, ensuring their relevance to the study of local labour markets<sup>4</sup>. I focus on the 297 employment areas of metropolitan France<sup>5</sup>, using the 2010 zoning, for years 2009, 2013 and 2017.

My main source of data is the *Professional activity* database of the population census conducted by Insee. This dataset provides individual information on education, labour market situation and demographic characteristics for around 7.5 millions workers for each year. I use these data to compute the overeducation rates in each employment area, as well as local labour market characteristics such as employment by business sector and indicators of skilled labour supply and demand. In the *Professional activity* database, individuals are localised at the employment area

---

<sup>4</sup>See Appendix 2A for a cartography of French employment areas.

<sup>5</sup>I exclude the seven employment areas of Corsica because of its insular nature.



of their place of work rather than their place of residence, I thus calculate the share of individuals working in the area but living in another to account for this.

I also use the Directory of Enterprises and Institutions data (REE), which provides city-based information on the demography of private firms. From it, I calculate the number of microfirms (with less than 10 workers) at the beginning of the year and the number of new firms created in the last year in each employment area. All variables are described in Table 2.1.

Table 2.1: Description of the variables

| Variables  | Sample          | Source                                  |
|--|-----------------|---|
| Overeducation rate   | 20-39 year-olds | Population census (activity database)   |
| % of the workers who studied outside the area              | 20-39 year-olds | Population census (activity database)   |
| % of 20-39 year-olds in the labour force                   | 15-64 year-olds | Population census (activity database)   |
| % of workers living outside their area of work             | 15-64 year-olds | Population census (activity database)   |
| % of employment in the agricultural sector                 | 15-64 year-olds | Population census (activity database)   |
| % of employment in the industrial sector                   | 15-64 year-olds | Population census (activity database)   |
| % of employment in the construction sector                 | 15-64 year-olds | Population census (activity database)   |
| % of employment in the sales sector                        | 15-64 year-olds | Population census (activity database)   |
| % of employment in the services sector                     | 15-64 year-olds | Population census (activity database)   |
| % of employment in the education, health and public sector | 15-64 year-olds | Population census (activity database)   |
| Supply-demand imbalance of tertiary graduates              | 15-64 year-olds | Population census (activity database)   |
| % of temporary contracts                                   | 15-64 year-olds | Population census (activity database)   |
| Number of microfirms for 100000 inhab. on January 1st      | Private firms   | REE                                     |
| Number of firms created in the last year for 100000 inhab. | Private firms   | REE                                     |
| Long-term unemployment rates (1990)                        | 15-64 year-olds | Population census (city-localised data) |
| Area shares a border with a foreign country                |                 | Employment area data                    |

Source: French population census and REE

## 2.3.2 Variables

### Overeducation rates

In this work, I focus on overeducation of young workers, aged between 20 and 39 years old, for two main reasons. First, overeducation is known to particularly affect young workers at the beginning of their career (Baert et al., 2013; Meroni & Vera-Toscano, 2017) and to decrease after some time on the labour market with the accumulation of professional experience (Rubb, 2003). Second, the choice of localisation on the individuals across the country is not random. I am able to control for this using information about whether the workers graduated in the area or not. However, the drawback of this proxy based on the place of education is to be relevant mostly for younger workers. That is why I focus on overeducation of young workers in the first half of their professional career.

Local overeducation rates are calculated using the French population census, which indicates whether workers have a tertiary degree or not. An individual is here considered as overeducated if she or he has a higher education degree but works in

an occupation not requiring one. Overeducation rates are then calculated as the ratio between the number of overeducated workers and the number of employed individuals in the area. This imperfect measure leads to underestimating overeducation: for example, individuals with an ISCED<sup>6</sup> 7 degree (typically, a Master's degree) but working in an occupation requiring ISCED 6 (Bachelor's degree) are not identified as overeducated. This could bias results, especially if adding these unobserved overeducated workers changed the geography of overeducation, which must be kept in mind in the interpretation of the results.

To determine whether an occupation requires a higher education degree, I rely on the *Professional families* classification of the French Ministry of Labour (Dares)<sup>7</sup>. For each of the 87 professional families<sup>8</sup>, information about required qualifications is available from the Operational Directory of Occupations and Jobs (ROME) of the French state employment agency *Pôle Emploi*. There are nine levels of qualification in ROME, the three highest (Technicians, Supervisors, Managers and engineers) can be considered to require a tertiary degree. For some of the professional families, classification is straightforward because only these three levels are involved. For families with a wider range of required qualifications, I consider that a higher education degree is required if more than half of the workers in the occupation have one.

## Explanatory variables

The choice of the explanatory variables is driven by the theoretical background of overeducation. Human Capital Theory (Becker, 1964) stresses the importance of worker characteristics and institutional factors<sup>9</sup>. The Job Competition Model (Thurow, 1975) underlines the importance of job characteristics in explaining educational mismatches. Assignment models (Sattinger, 1993) can be seen as an intermediate framework between Human Capital Theory and the Job Competition Model. These models predict that job allocation is the result of an assignment

---

<sup>6</sup>International Standard Classification of Education.

<sup>7</sup>The *Professional families* classification is built by linking socio-professional categories (PCS) and qualifications required as defined in the Operational Directory of Occupations and Jobs (ROME). *Professional families* are particularly appropriate to study employment in terms of qualifications, opportunities, mobilities and supply-demand relationships.

<sup>8</sup>The final total is 85 families, after excluding Armed forces and Religious occupations.

<sup>9</sup>In Human Capital Theory, there should not be any underutilisation of human capital. Thus, overeducation should not exist at the equilibrium. However, McGuinness (2006) indicated that, as a short-run phenomenon, overeducation remains consistent with this framework, in particular because of institutional factors that for instance prevent firms from adjusting their production processes to the new characteristics of the labour supply.

problem and depends on both the distribution of workers and the distribution of jobs. McGuinness (2006) indicated that this framework is entirely consistent with overeducation, and I therefore include information on both the distribution of workers and the distribution of firms.

As for the distribution of workers, I include the share of 20-39 year-olds in the active population, drawn from the population census, which corresponds to the importance of the population of interest in the labour force of the area. The choice of localisation of the individuals is not random, as it is impacted by characteristics of the area (Roback, 1982). Moreover, young workers tend to live closer to cities Murphy et al. (1991). To control for this, I include the share of individuals who studied in a different area than the one they work in. To take into account the spatial flexibility of workers (Büchel & van Ham, 2003; Jauhainen, 2011), I include the share of individuals living outside the employment area of their workplace, *i.e.* who commute from another area to work.

Since overeducation affects some sectors more than others (Nauze-Fichet & Tomasini, 2002), I also include the distribution of employment among the different sectors. Besides, Elhorst (2003) show that the industry mix of an area affect the local labour market. As for the distribution of firms, I also include the log of the number of microfirms (with less than 10 employees), since the probability of being overeducated has been found to decrease with size of firm (Ghignoni & Verashchagina, 2014; Guironnet, 2006). One possible explanation is that small firms are less likely to provide on-the-job training, in which case hiring an overeducated worker reduces the risk of choosing a candidate who lacks the required skills. Moreover, recruitment processes are more formalised in larger firms, reducing the risk of mismatch. Thus, overeducation is likely to be more of a problem in areas where firms are smaller. In addition, I include the log of the number of firms created in the last year, as an indicator of the local economic dynamism of the areas.

Finally, I include variables addressing the local labour market context. To control for the effect of increased educational attainment (Groot & Maassen van den Brink, 2000), I use an indicator of the imbalance between supply and demand for tertiary-level graduates in the active population. This variable is the ratio between the share of tertiary graduates in the labour force (proxy for the skilled labour supply) and the share of individuals in ISCO 1, 2 and 3 occupations in the labour force (proxy for the skilled labour demand). If this ratio is greater than one, there are too many

graduates for the skilled jobs available; if this ratio is lower than one, not all jobs can be filled by the area's skilled labour force. In addition, I control for the share of temporary employment (Croce & Ghignoni, 2012; Ghignoni & Verashchagina, 2014). Indeed, according to the theory of turnover (Jovanovic, 1979), mismatches are a consequence of imperfect information. Repeated job seeking is therefore a way to improve matches, by obtaining more and better information. Thus, short-term contracts should help reduce overeducation. Finally, I include a dummy variable equal to one if the employment area shares a border with a foreign country. This allows to control for a potential excess supply of workers living in neighbouring countries but working in France, that would overcrowd the concerned local labour markets and generate overeducation. The mechanism can also work in the opposite way (French inhabitants working in a foreign country). The expected sign of the effect is therefore positive in the former case, and negative in the latter.

### 2.3.3 Descriptive statistics

Table 2.2 presents the main characteristics of the sample for the years 2009, 2013 and 2017. Overeducation rates increase over the period, from 12.8% in 2009 to 15.5% in 2017 on average. The rise is higher at the top (+5.7 points for the maximum value) than at the bottom of the distribution (+1 point for the minimum value). Overeducation rates are higher for women than for men, increase over the period are similar. The proportion that the 20-39 year-olds represent in the whole labour force decrease between 2009 and 2017. The share of workers who relocated in the area since the end of their studies grow by around 5 points over the period, and the share of workers living outside the employment area of their workplace also slightly increase. This possibly indicates that workers broaden their job search area to find a better job.

The distribution of employment among the different sectors vary slightly between 2009 and 2017, the share of the services and public sectors increasing at the expense of the industrial and construction sectors. Overall, services, sales, education, health and the public sector represent about 70% of employment on average. There is also a significant rise in the number of new firms, likely reflecting a more favourable economic climate at the end of the period than at the beginning. The share of temporary contracts appears stable over the period, around 13%.

On average, the indicator of imbalance between the skilled labour supply and demand is lower than one, meaning that there are more jobs requiring educated labour

than workers to occupy them. However, this ratio increase sharply over the period, reaching almost one in 2017. This average hides large disparities between employment areas, with the ratio ranging from about 0.7 to more than 1.15. This may indicate a spatial mismatch between jobs and workers, highlighting the importance of investigating the geographical dimension of overeducation.

Table 2.2: Sample characteristics

| Variables  | 2009  |           | 2013  |           | 2017  |           |
|--|-------|-----------|-------|-----------|-------|-----------|
|  | Mean  | Std. Dev. | Mean  | Std. Dev. | Mean  | Std. Dev. |
| <i>Dependent variables</i>                                   |       |           |       |           |       |           |
| Overeducation rate   | 0.128 | 0.001     | 0.143 | 0.001     | 0.155 | 0.001     |
| Women overeducation rate                                     | 0.159 | 0.002     | 0.172 | 0.002     | 0.183 | 0.002     |
| Men overeducation rate                                       | 0.099 | 0.001     | 0.114 | 0.001     | 0.128 | 0.001     |
| <i>Distribution of workers</i>                               |       |           |       |           |       |           |
| % of the workers who studied outside the area                | 0.316 | 0.009     | 0.333 | 0.010     | 0.361 | 0.011     |
| % of 20-39 year-olds in the labour force                     | 0.398 | 0.001     | 0.371 | 0.001     | 0.360 | 0.001     |
| % of workers living outside their area of work               | 0.197 | 0.006     | 0.209 | 0.006     | 0.224 | 0.006     |
| <i>Employment by sector</i>                                  |       |           |       |           |       |           |
| % of employment in the agricultural sector                   | 0.021 | 0.001     | 0.022 | 0.001     | 0.021 | 0.001     |
| % of employment in the industrial sector                     | 0.186 | 0.004     | 0.174 | 0.004     | 0.168 | 0.004     |
| % of employment in the construction sector                   | 0.066 | 0.001     | 0.063 | 0.001     | 0.058 | 0.001     |
| % of employment in the sales sector                          | 0.211 | 0.002     | 0.213 | 0.002     | 0.213 | 0.002     |
| % of employment in the services sector                       | 0.168 | 0.002     | 0.170 | 0.003     | 0.173 | 0.003     |
| % of employment in the educ., health and public sector       | 0.348 | 0.003     | 0.359 | 0.003     | 0.367 | 0.003     |
| <i>Local characteristics</i>                                 |       |           |       |           |       |           |
| Supply-demand imbalance of tertiary graduates                | 0.856 | 0.004     | 0.908 | 0.004     | 0.964 | 0.004     |
| % of temporary contracts                                     | 0.129 | 0.001     | 0.128 | 0.002     | 0.132 | 0.001     |
| Number of microfirms (for 100000 inhab.)                     | 13282 | 4771      | 15575 | 5441      | 19279 | 6789      |
| Number of firms created in the last year (for 100000 inhab.) | 1216  | 463       | 1912  | 661       | 2120  | 787       |
| Area shares a border with a foreign country                  | 0.148 | 0.021     | 0.148 | 0.021     | 0.148 | 0.021     |
| Observations   | 297   |           | 297   |           | 297   |           |

Note: Raw values are reported for the number of microfirms and newly established firms rather than the log.

Source: French population census and REE

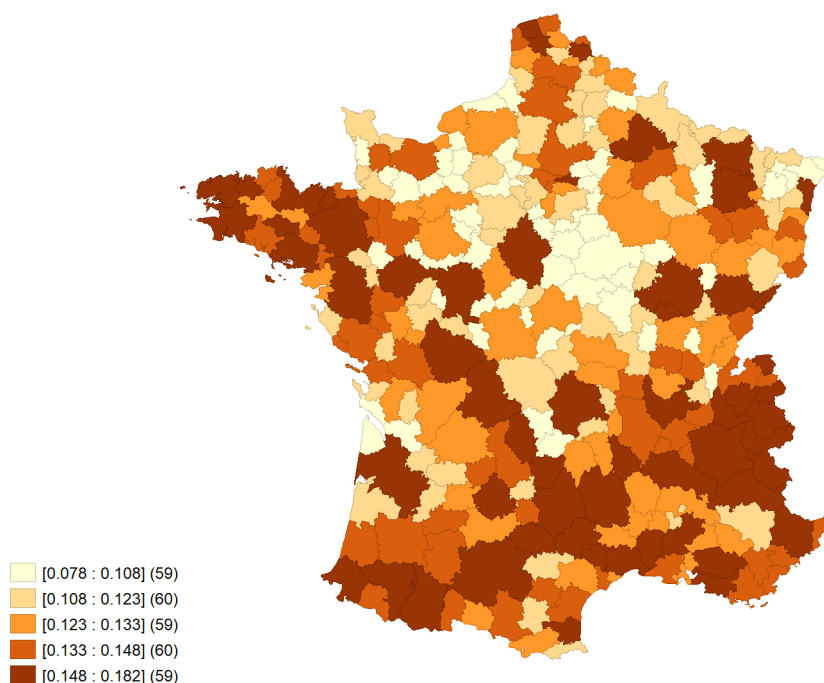
## 2.4 Empirical strategy

### 2.4.1 Spatial interactions

In order to address hypothesis  $H2$ , I explore the geography and disparities of overeducation rates between employment areas. Figures 2.1 to 2.3 display overeducation rates by employment area, respectively for years 2009, 2013 and 2017. The first thing to note is that over-education seems to be more prevalent in the southern and western parts of the country, regardless of the year. Moreover, several clusters of areas show similar overeducation rates, with high values gathered in the south and low values in the center. While the spatial clusters become less apparent with time, patterns remain similar at each date. This suggests that neighbouring observations

are not independent, but spatially autocorrelated, which violates a necessary assumption of standard models and leads to biased estimators if this spatial structure is ignored (Cliff & Ord, 1973, 1981). Spatial autocorrelation is positive (negative) if very similar (very different) values of the same variable are observed in neighbouring areas. To confirm the presence of autocorrelation, I calculate Moran's  $I$  coefficients (Moran, 1950)<sup>10</sup>. Table 2.3 presents the values of the Moran's  $I$  for years 2009, 2013 and 2017. These are always very significant, regardless of the spatial weight matrix<sup>11</sup>.

Figure 2.1: Overeducation rates by French employment area - 2009



Note: Areas in pale yellow have the lowest overeducation rates, areas in dark red have the highest overeducation rates.

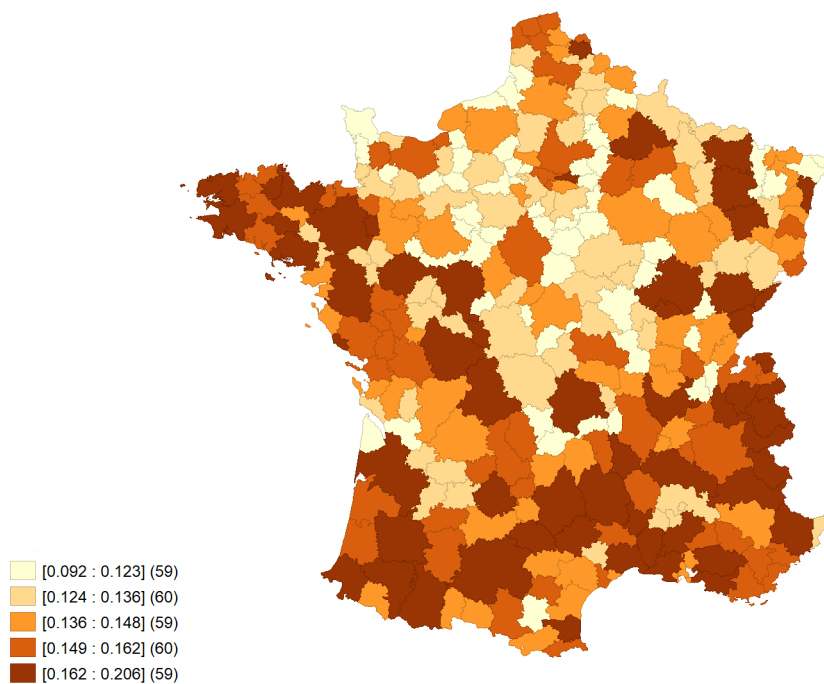
<sup>10</sup>Moran's  $I$  is defined as follows:

$$I_W = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}; i \neq j$$

where  $n$  is the number of observations,  $y_i$  and  $y_j$  are the values of the variable in area  $i$  and  $j$  respectively and  $\bar{y}$  the mean value of that same variable, and  $w_{ij}$  the element of the spatial weight matrix that defines the proximity between area  $i$  and area  $j$ .

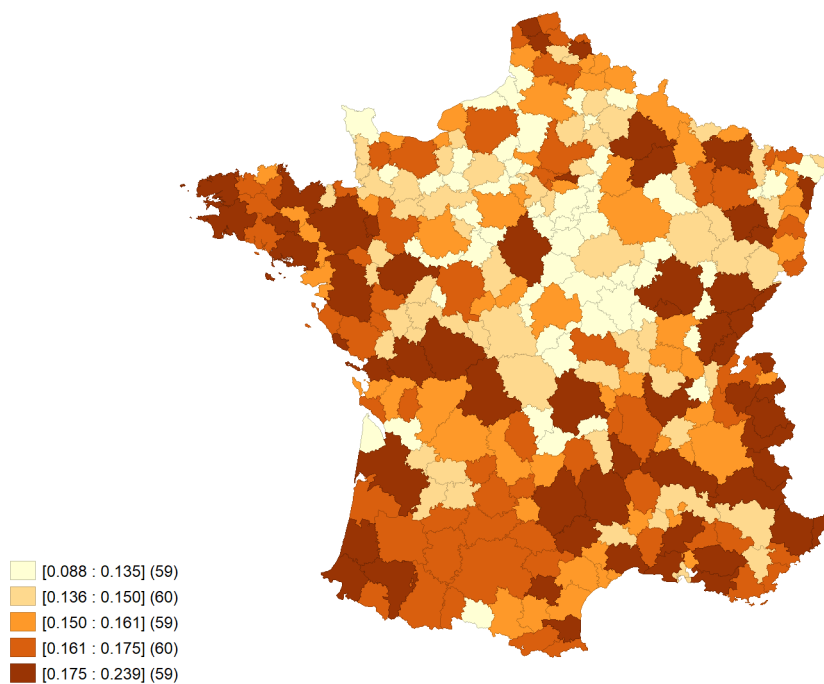
<sup>11</sup>I use contiguity matrices (first- and second-order), k-nearest neighbours (from 3 to 7 neighbours) matrices and an inverse distance matrix (with a threshold at 100km, and a zero weight for areas located farther away).

Figure 2.2: Overeducation rates by French employment area - 2013



*Note:* Areas in pale yellow have the lowest overeducation rates, areas in dark red have the highest overeducation rates.

Figure 2.3: Overeducation rates by French employment area - 2017



*Note:* Areas in pale yellow have the lowest overeducation rates, areas in dark red have the highest overeducation rates.

Table 2.3: Spatial autocorrelation of overeducation rates (Moran's  $I$ )

| Spatial weight matrix        | 2009        |            | 2013        |            | 2017        |            |
|------------------------------|-------------|------------|-------------|------------|-------------|------------|
|                              | Moran's $I$ | $p$ -value | Moran's $I$ | $p$ -value | Moran's $I$ | $p$ -value |
| First-order contiguity       | 0.343***    | 0.002      | 0.310***    | 0.002      | 0.298***    | 0.002      |
| Second-order contiguity      | 0.277***    | 0.002      | 0.227***    | 0.002      | 0.225***    | 0.002      |
| 3 nearest neighbours (3NN)   | 0.384***    | 0.002      | 0.367***    | 0.002      | 0.347***    | 0.002      |
| 4 nearest neighbours (4NN)   | 0.365***    | 0.002      | 0.331***    | 0.002      | 0.322***    | 0.002      |
| 5 nearest neighbours (5NN)   | 0.360***    | 0.002      | 0.323***    | 0.002      | 0.318***    | 0.002      |
| 6 nearest neighbours (6NN)   | 0.341***    | 0.002      | 0.315***    | 0.002      | 0.317***    | 0.002      |
| 7 nearest neighbours (7NN)   | 0.339***    | 0.002      | 0.303***    | 0.002      | 0.305***    | 0.002      |
| Inverse distance (max 100km) | 0.304***    | 0.002      | 0.262***    | 0.002      | 0.266***    | 0.002      |
| Number of observations       | 297         |            | 297         |            | 297         |            |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source : Author's calculation based on French population census and REE

Following Manski (1993), the full spatial model, including all forms of spatial lags, can be written as follows:

$$\begin{aligned}
 Y &= \rho WY + X\beta + WX\theta + u \\
 u &= \lambda Wu + \epsilon
 \end{aligned}
 \tag{2.1}$$

where  $W$  is the spatial weight matrix. This model is not identifiable in this form; to estimate the spatial model, at least one of the spatial lags ( $\rho$ ,  $\theta$  and  $\lambda$ ) has to be constraint to zero. To determine which type of model to use, I follow the specification procedure proposed by Debarsy and Ertur (2010), based on Lagrange multiplier tests (LM-tests)<sup>12</sup>, in order to identify the presence of spatial autocorrelation on the dependent variable (Spatial AutoRegressive model, SAR) or in the error term (Spatial Error Model, SEM). Results of the specification tests are reported in Table 2.4, for various spatial weight matrix designs.

Despite the autocorrelation observed in Table 2.3, the inclusion of a spatial lag is not obvious when looking at the LM-tests. The inclusion of a lagged error term is excluded at all dates. As for the spatial lag of the dependent variable, there are strong variations depending on the year. While it seems important to take this spatial dimension into account for year 2009, it appears not to be necessary for 2013

<sup>12</sup>These LM-tests are score tests based on the gradient of the likelihood function. When testing the SAR model (respectively the SEM model), the null hypothesis is that the spatial coefficient  $\rho$  ( $\lambda$ ) is equal to zero. When the null hypothesis is rejected, the spatial structure of the data needs to be taken into account.



Table 2.4: Specification tests - LM-tests for inclusion of spatial lags

| Spatial weight matrix              | Spatial lag | <i>p</i> -value | Spatial error | <i>p</i> -value |
|------------------------------------|-------------|-----------------|---------------|-----------------|
| <b>Year 2009</b> ( <i>N</i> = 297) |             |                 |               |                 |
| First-order contiguity             | 8.096***    | 0.004           | 0.071         | 0.790           |
| Second-order contiguity            | 19.24***    | 0.000           | 0.266         | 0.606           |
| 3 nearest neighbours (3NN)         | 4.104**     | 0.043           | 0.242         | 0.623           |
| 4 nearest neighbours (4NN)         | 6.290**     | 0.012           | 0.287         | 0.592           |
| 5 nearest neighbours (5NN)         | 8.777***    | 0.003           | 0.031         | 0.861           |
| 6 nearest neighbours (6NN)         | 6.532**     | 0.011           | 0.474         | 0.490           |
| 7 nearest neighbours (7NN)         | 8.152***    | 0.004           | 0.074         | 0.785           |
| Inverse distance (max 100km)       | 8.826***    | 0.003           | 0.078         | 0.781           |
| <b>Year 2013</b> ( <i>N</i> = 297) |             |                 |               |                 |
| First-order contiguity             | 0.381       | 0.537           | 0.268         | 0.605           |
| Second-order contiguity            | 1.887       | 0.170           | 1.316         | 0.251           |
| 3 nearest neighbours (3NN)         | 0.291       | 0.590           | 0.345         | 0.557           |
| 4 nearest neighbours (4NN)         | 0.082       | 0.774           | 1.158         | 0.282           |
| 5 nearest neighbours (5NN)         | 0.287       | 0.592           | 1.345         | 0.246           |
| 6 nearest neighbours (6NN)         | 0.085       | 0.771           | 2.911*        | 0.088           |
| 7 nearest neighbours (7NN)         | 0.377       | 0.539           | 0.849         | 0.357           |
| Inverse distance (max 100km)       | 0.159       | 0.690           | 0.309         | 0.579           |
| <b>Year 2017</b> ( <i>N</i> = 297) |             |                 |               |                 |
| First-order contiguity             | 0.000       | 0.997           | 1.100         | 0.294           |
| Second-order contiguity            | 0.123       | 0.726           | 0.119         | 0.730           |
| 3 nearest neighbours (3NN)         | 0.021       | 0.885           | 1.035         | 0.309           |
| 4 nearest neighbours (4NN)         | 0.119       | 0.730           | 1.471         | 0.225           |
| 5 nearest neighbours (5NN)         | 0.025       | 0.875           | 0.805         | 0.370           |
| 6 nearest neighbours (6NN)         | 0.003       | 0.956           | 1.601         | 0.206           |
| 7 nearest neighbours (7NN)         | 0.258       | 0.611           | 0.829         | 0.363           |
| Inverse distance (max 100km)       | 0.232       | 0.630           | 0.614         | 0.433           |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.  
Source : Author's calculation based on French population census and REE

and 2017. Given these results, I estimate a SAR model<sup>13</sup>, but it is expected that the spatial lag will be significant in the estimation results only for 2009. This model can be written as follows:

$$Y = \rho WY + X\beta + u \quad (2.2)$$

with  $W$  the spatial weight matrix. In the SAR model, coefficients cannot be interpreted directly. Such a model makes it possible to distinguish between direct effects, *i.e.* the average impact of a variable observed in a given area on overeducation in the same area, and indirect effects, *i.e.* the average impact on overeducation in one area of the explanatory variables observed in other areas (LeSage & Pace, 2009). The total effect, which sums direct and indirect effects, is the average impact of a

<sup>13</sup>Alternatively, I could estimate a Spatial Durbin Model (SDM), which includes spatial lags of the covariates in addition to the lag of the dependent variable. This type of model is more general and includes both the SAR and SEM models by transformation, which makes it more robust to misspecifications (Floch & Le Saout, 2018). However, including spatial lags of covariates creates a problem of discontinuous space using spatial regimes (see Section 2.4.2), as it leads to consider as neighbours only areas in the same regime. Moreover, lags of the covariates are never significant in the SDM signification (see Appendix 2C), which supports the use of the SAR model.

regressor on the dependent variable (see Appendix 2B for the calculation of these different effects).

The main results are obtained using a first-order contiguity matrix, *i.e.* areas are considered as neighbours if they share a border. This simple and common matrix generally gives reliable estimates, better reveals the true spatial structure of the data and performs better on average than the others (Stakhovych & Bijmolt, 2009). Moreover, LeSage and Pace (2009) show that results are generally not significantly impacted by the choice of spatial weight matrix.

### 2.4.2 Spatial heterogeneity

There are spatial disparities in terms of skilled labour supply (Rodríguez-Pose & Tselios, 2011) and demand (Détang-Dessendre, 1999) between urban and rural areas. There is also evidence of heterogeneous effects of spatial mobility on overeducation depending on the destination of the migration (Iammarino & Marinelli, 2015), especially migration towards large economic centers (Fouquet & Sari, 2023). This suggests that local characteristics have heterogeneous effects on overeducation depending on the urbanisation of the area (hypothesis *H3*).

I account for spatial heterogeneity using a spatial regimes model (Le Gallo, 2004). The inclusion of spatial regimes allows the estimated effects of the explanatory variables to be different depending on the regime. The first regime corresponds to urban areas and the second regime to rural areas<sup>14</sup>. Following Le Gallo (2004), the model can be written as follows:

$$Y = D_U X \beta_U + D_R X \beta_R + u \quad (2.3)$$

with  $D_U$  and  $D_R$  the dummy variables denoting the two spatial regimes, respectively equal to 1 if urban area (0 otherwise) and equal to 1 if rural area (0 otherwise). I also estimate a SAR spatial regimes model, including both spatial interactions and spatial heterogeneity, which can be written as follows:

$$Y = \rho W Y + D_U X \beta_U + D_R X \beta_R + u \quad (2.4)$$

---

<sup>14</sup>Urban and rural areas are defined following the typology from Insee, which is based on population density. “Dense” and “intermediate density” areas are considered as urban, while “low density” and “very low density” areas are considered as rural. More details about this typology are provided in D’Alessandro et al. (2021).

where  $WY$  do not vary by regime, which allow to account for interactions between urban and rural areas.

### 2.4.3 Endogeneity issue

I control for the choice of localisation of the individuals by including the share of workers who work in the area but studied in another. However, this choice is likely to be endogenous: individuals might have chosen to come in the area to have better labour market prospects, especially to find more suitable job opportunities. To address this issue, I use an instrumental variables approach. The models without spatial interactions are estimated using two-stage least squared (2SLS), the SAR models are estimated using generalized spatial two-stage least squares (GS2SLS) as proposed by Kelejian and Prucha (1998).

The instrument used in this work is the lagged unemployment rate in the area (unemployment rate in 1990), rebuilt from the municipality-level database of the population census. The validity of the instrument is based on the assumption that the long-term unemployment rate is a good predictor of the attractiveness of an area, but that overeducation in 2009, 2013 and 2017 is uncorrelated with unemployment in 1990. Such an assumption can be criticised, especially since local unemployment tend to be persistent. However, such an instrument has been used in the literature (Croce & Ghignoni, 2015, see e.g.) and provides consistent results. Besides, the instrument is statistically validated by the weak instrument test of Olea and Pflueger (2013). Results for the first-stage of the models are presented in Appendix 2D.

## 2.5 Results

### 2.5.1 Effects of the local context

Estimation results for the baseline and SAR models are presented in Table 2.5. The effects of most variables are similar for all years. However, as expected from the specification tests (Table 2.4), the spatial lag of the dependent variable (SAR model) is significant only for year 2009. Hypothesis  $H2$  is therefore only partially verified. The positive sign of the coefficient implies that overeducation is higher when it is also high in neighbouring areas. In a SAR model, estimated coefficients do not directly give the impacts of the variables on overeducation. Direct, indirect

Table 2.5: Estimation results

| Variables   | OLS       |       | SAR       |       |
|---|-----------|-------|-----------|-------|
|   | Coeff.    | SE    | Coeff.    | SE    |
| <b>Year 2009</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | 0.035     | 0.040 | 0.060     | 0.039 |
| % of the workers who studied outside the area     | 0.030**   | 0.014 | 0.041***  | 0.014 |
| % of 20-39 year-olds in the labour force          | -0.074*   | 0.038 | -0.049    | 0.037 |
| % of workers living outside their area of work    | -0.027*** | 0.008 | -0.033*** | 0.007 |
| % of empl. in the agricultural sector             | -0.127*** | 0.031 | -0.121*** | 0.030 |
| % of empl. in the industrial sector               | -0.113*** | 0.019 | -0.128*** | 0.018 |
| % of empl. in the construction sector             | -0.026    | 0.060 | -0.079    | 0.059 |
| % of empl. in the sales sector                    | -0.076**  | 0.030 | -0.079*** | 0.028 |
| % of empl. in the educ., health and public sector | -0.097*** | 0.021 | -0.102*** | 0.020 |
| Supply-demand imbalance of tertiary graduates     | 0.300***  | 0.016 | 0.275***  | 0.016 |
| % of temporary contracts                          | -0.029    | 0.029 | -0.063**  | 0.029 |
| Number of microfirms                              | -0.002    | 0.005 | -0.006    | 0.005 |
| Number of firms created in the last year          | -0.006    | 0.004 | -0.006    | 0.004 |
| Area shares a border with a foreign country       | 0.005**   | 0.002 | 0.005**   | 0.002 |
| Spatial lag of $Y$                                |           |       | 0.223***  | 0.054 |
| <b>Year 2013</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | -0.021    | 0.043 | -0.017    | 0.042 |
| % of the workers who studied outside the area     | 0.029**   | 0.011 | 0.032***  | 0.011 |
| % of 20-39 year-olds in the labour force          | -0.028    | 0.039 | -0.017    | 0.039 |
| % of workers living outside their area of work    | -0.027*** | 0.007 | -0.029*** | 0.007 |
| % of empl. in the agricultural sector             | -0.111*** | 0.030 | -0.111*** | 0.029 |
| % of empl. in the industrial sector               | -0.087*** | 0.019 | -0.090*** | 0.019 |
| % of empl. in the construction sector             | -0.026    | 0.061 | -0.046    | 0.061 |
| % of empl. in the sales sector                    | -0.048    | 0.030 | -0.047    | 0.029 |
| % of empl. in the educ., health and public sector | -0.075*** | 0.020 | -0.074*** | 0.020 |
| Supply-demand imbalance of tertiary graduates     | 0.289***  | 0.015 | 0.280***  | 0.015 |
| % of temporary contracts                          | -0.051*   | 0.029 | -0.065**  | 0.030 |
| Number of microfirms                              | 0.009*    | 0.005 | 0.008     | 0.005 |
| Number of firms created in the last year          | -0.016*** | 0.005 | -0.016*** | 0.005 |
| Area shares a border with a foreign country       | 0.005***  | 0.002 | 0.005***  | 0.002 |
| Spatial lag of $Y$                                |           |       | 0.089     | 0.059 |
| <b>Year 2017</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | -0.103**  | 0.046 | -0.102**  | 0.045 |
| % of the workers who studied outside the area     | 0.056***  | 0.017 | 0.059***  | 0.017 |
| % of 20-39 year-olds in the labour force          | -0.073*   | 0.038 | -0.063*   | 0.038 |
| % of workers living outside their area of work    | -0.022*** | 0.007 | -0.023*** | 0.007 |
| % of empl. in the agricultural sector             | -0.098*** | 0.032 | -0.099*** | 0.031 |
| % of empl. in the industrial sector               | -0.078*** | 0.022 | -0.082*** | 0.022 |
| % of empl. in the construction sector             | -0.042    | 0.086 | -0.065    | 0.087 |
| % of empl. in the sales sector                    | -0.075**  | 0.035 | -0.076**  | 0.035 |
| % of empl. in the educ., health and public sector | -0.054*** | 0.020 | -0.055*** | 0.020 |
| Supply-demand imbalance of tertiary graduates     | 0.322***  | 0.018 | 0.316***  | 0.018 |
| % of temporary contracts                          | -0.090*** | 0.033 | -0.100*** | 0.034 |
| Number of microfirms                              | 0.015*    | 0.007 | 0.015**   | 0.007 |
| Number of firms created in the last year          | -0.017*** | 0.006 | -0.017*** | 0.006 |
| Area shares a border with a foreign country       | 0.012***  | 0.003 | 0.012***  | 0.002 |
| Spatial lag of $Y$                                |           |       | 0.069     | 0.060 |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source : Author's calculation based on French population census and REE

and total effects for this model are reported in Table 2.6 for year 2009<sup>15</sup>.

Overall, I find significant effects of several variables on overeducation at the area level, confirming hypothesis *H1*. There is a strong direct effect of the imbalance between skilled labour supply and demand, confirming that the lack of qualified jobs in an area (relative to the number of tertiary graduates) leads to higher overeducation rates (Davia et al., 2017). I find a significant positive effect of the share of

<sup>15</sup>For sake of clarity, those effects are not reported for years 2013 and 2017. Indeed, since the spatial lag is not significant, the direct effect is almost equal to the coefficient, and the indirect effect is close to zero and never significant.

workers who studied elsewhere. This suggests that areas which attract workers are associated with higher overeducation. I also find a positive effect of being a border area. Those effects suggest some overcrowding mechanism in these regions: areas attract workers who were not originally from there (or at least who did not study there) or come from the neighbouring foreign country, but the demand of skilled labour is not keeping pace with this additional supply.

Table 2.6: Direct, indirect and total effects - SAR model

| Variables   | Direct    | Indirect  | Total     |
|---|-----------|-----------|-----------|
| <b>Year 2009</b> ( $N = 297$ )                    |           |           |           |
| % of the workers who studied outside the area     | 0.041***  | 0.011**   | 0.053***  |
| % of 20-39 year-olds in the labour force          | -0.049    | -0.014    | -0.063    |
| % of workers living outside their area of work    | -0.034*** | -0.009**  | -0.043*** |
| % of empl. in the agricultural sector             | -0.122*** | -0.034*** | -0.156*** |
| % of empl. in the industrial sector               | -0.129*** | -0.035*** | -0.164*** |
| % of empl. in the construction sector             | -0.080    | -0.022    | -0.102    |
| % of empl. in the sales sector                    | -0.080*** | -0.022**  | -0.102*** |
| % of empl. in the educ., health and public sector | -0.104*** | -0.028*** | -0.132*** |
| Supply-demand imbalance of tertiary graduates     | 0.278***  | 0.076***  | 0.355***  |
| % of temporary contracts                          | -0.063**  | -0.017    | -0.080**  |
| Number of microfirms                              | -0.006    | -0.002    | -0.007    |
| Number of firms created in the last year          | -0.006    | -0.002    | -0.007    |
| Area shares a border with a foreign country       | 0.006***  | 0.002**   | 0.007***  |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source : Author's calculation based on French population census and REE

I also find a negative direct effect of the number of new firms created in the previous year in 2013 and 2017. The absence of significant effect in 2009 is likely to be related to the global economic context, with low firm creations during the economic crisis. This positive coefficient might be the sign that economic dynamism in an area reduces overeducation. This could also mean that some workers decide to launch their business to avoid being overeducated, as suggested by Conroy and Watson (2023), although the absence of relevant information on this subject at the individual level leaves this explanation at the level of a hypothesis.

I find effects of the structure of employment across sectors. Especially, a higher share of any sector (except construction) in comparison with the services sector is associated with lower overeducation rates. The indirect effects (Table 2.6), *i.e.* the effects of the variables in an area on overeducation in neighbouring areas, are similar to the direct effects in terms of sign and significance, but their magnitude is three to four times lower. This confirms that the situation on the labour market in an area affects overeducation at a broader scale (for year 2009), but that the effect decreases with distance (Manning & Petrongolo, 2017).

## 2.5.2 Heterogeneity between urban and rural areas

Table 2.7 displays the results of the spatial regimes models, with and without inclusion of spatial autocorrelation. Direct, indirect and total effects for the SAR spatial regimes model are reported in Table 2.8. There is a strong spatial heterogeneity between rural and urban areas, confirming hypothesis *H3*. It has to be noted that the spatial lag of the dependent variable is now significant for the three years, although only at the 10% level for years 2013 and 2017. Still, most indirect effects are significant only in 2009.

The results observed for rural areas are similar to the ones obtained without taking spatial heterogeneity into account, while only a few variables seem to affect overeducation in urban areas. Spatial regimes models confirm the positive direct effect of the excess supply of tertiary graduates in both rural and urban areas, with limited difference in terms of magnitude. However, it appears to be the main effect in urban areas, while many other drivers are evidenced in rural ones. The structure of employment by sector have significant effects in rural areas, as well as the share of young workers in the labour force and the proximity with a foreign country. Overall, this implies that overeducation is driven mostly by the competition between graduates in urban areas, while the local context also have an impact in rural areas. In particular, the effects of the structure of employment by sector suggest that increasing the variety of job opportunities in rural areas might help reduce overeducation.

## 2.5.3 Heterogeneity by gender

To study the heterogeneous effects of the local context on overeducation across gender, I estimate the spatial regimes SAR model separately for men and women. Direct, indirect and total effects by gender are reported in Table 2.9<sup>16</sup>. Overall, results for both gender are similar in terms of sign and significance of the direct effects, and confirm that the local context has a greater influence on rural areas than on urban ones. Still, results for men and women differ on three main points.

Firstly, the magnitude of the effects vary quite strongly between gender. Most significant variables have a stronger effect on women than on men (generally from 1.5 to 2 times higher). Secondly, while the spatial lag of the dependent variable

<sup>16</sup>Estimation results are available in Appendix 2.12. For sake of clarity, I only report the results of the spatial regimes SAR model, which is the most complete. Besides, the inclusion of spatial regimes and of the spatially lagged dependent variable does not affect substantially the effects observed with the other models (full tables for these models available upon request).

Table 2.7: Estimation results - Spatial regimes models

| Variables   | Spatial regimes |                |           |       | SAR + SR  |                  |           |       |
|---|-----------------|----------------|-----------|-------|-----------|------------------|-----------|-------|
|   | Urban           |                | Rural     |       | Urban     |                  | Rural     |       |
|   | Coeff.          | SE             | Coeff.    | SE    | Coeff.    | SE               | Coeff.    | SE    |
| <b>Year 2009</b> ( $N = 297$ )                    |                 |                |           |       |           |                  |           |       |
| Constant  | -0.011          | 0.043          |           |       | 0.016     | 0.041            |           |       |
| % of the workers who studied outside the area     | -0.002          | 0.020          | 0.064***  | 0.022 | 0.012     | 0.019            | 0.070***  | 0.020 |
| % of 20-39 year-olds in the labour force          | -0.090          | 0.059          | -0.025    | 0.044 | -0.058    | 0.056            | -0.006    | 0.041 |
| % of workers living outside their area of work    | -0.008          | 0.012          | -0.041*** | 0.011 | -0.015    | 0.012            | -0.044*** | 0.010 |
| % of empl. in the agricultural sector             | -0.283**        | 0.127          | -0.115*** | 0.039 | -0.210*   | 0.120            | -0.121*** | 0.036 |
| % of empl. in the industrial sector               | -0.036          | 0.030          | -0.161*** | 0.029 | -0.060**  | 0.029            | -0.176*** | 0.028 |
| % of empl. in the construction sector             | 0.234*          | 0.123          | -0.200**  | 0.084 | 0.205*    | 0.115            | -0.255*** | 0.080 |
| % of empl. in the sales sector                    | 0.010           | 0.044          | -0.178*** | 0.045 | -0.005    | 0.042            | -0.177*** | 0.042 |
| % of empl. in the educ., health and public sector | -0.050*         | 0.029          | -0.147*** | 0.031 | -0.061**  | 0.028            | -0.153*** | 0.029 |
| Supply-demand imbalance of tertiary graduates     | 0.319***        | 0.031          | 0.323***  | 0.021 | 0.299***  | 0.029            | 0.297***  | 0.021 |
| % of temporary contracts                          | -0.038          | 0.063          | -0.055    | 0.036 | -0.085    | 0.060            | -0.085**  | 0.034 |
| Number of microfirms                              | -0.005          | 0.008          | 0.007     | 0.007 | -0.008    | 0.008            | 0.005     | 0.006 |
| Number of firms created in the last year          | -0.004          | 0.008          | -0.010*   | 0.005 | -0.005    | 0.008            | -0.010**  | 0.005 |
| Area shares a border with a foreign country       | 0.000           | 0.003          | 0.007**   | 0.003 | 0.001     | 0.003            | 0.008***  | 0.003 |
| Spatial lag of $Y$                                |                 |                |           |       | 0.197***  | 0.054            |           |       |
| Chow test for structural instability              |                 | 1.65* (0.073)  |           |       |           | 20.91* (0.075)   |           |       |
| <b>Year 2013</b> ( $N = 297$ )                    |                 |                |           |       |           |                  |           |       |
| Constant  | -0.035          | 0.047          |           |       | -0.032    | 0.044            |           |       |
| % of the workers who studied outside the area     | 0.009           | 0.016          | 0.050***  | 0.018 | 0.013     | 0.015            | 0.053***  | 0.017 |
| % of 20-39 year-olds in the labour force          | -0.030          | 0.060          | 0.006     | 0.047 | -0.011    | 0.058            | 0.020     | 0.045 |
| % of workers living outside their area of work    | -0.019          | 0.011          | -0.035*** | 0.010 | -0.020*   | 0.011            | -0.036*** | 0.009 |
| % of empl. in the agricultural sector             | -0.010          | 0.132          | -0.122*** | 0.037 | 0.023     | 0.127            | -0.125*** | 0.035 |
| % of empl. in the industrial sector               | -0.042          | 0.032          | -0.122*** | 0.029 | -0.046    | 0.030            | -0.126*** | 0.027 |
| % of empl. in the construction sector             | 0.118           | 0.121          | -0.142*   | 0.083 | 0.114     | 0.115            | -0.176**  | 0.081 |
| % of empl. in the sales sector                    | 0.012           | 0.049          | -0.123*** | 0.046 | 0.014     | 0.046            | -0.123*** | 0.043 |
| % of empl. in the educ., health and public sector | -0.066**        | 0.032          | -0.105*** | 0.029 | -0.066**  | 0.030            | -0.105*** | 0.028 |
| Supply-demand imbalance of tertiary graduates     | 0.290***        | 0.028          | 0.299***  | 0.018 | 0.280***  | 0.027            | 0.287***  | 0.019 |
| % of temporary contracts                          | -0.081          | 0.065          | -0.069*   | 0.035 | -0.109*   | 0.064            | -0.086**  | 0.035 |
| Number of microfirms                              | 0.004           | 0.008          | 0.018**   | 0.007 | 0.003     | 0.008            | 0.018**   | 0.007 |
| Number of firms created in the last year          | -0.011          | 0.009          | -0.024*** | 0.006 | -0.012    | 0.009            | -0.024*** | 0.006 |
| Area shares a border with a foreign country       | 0.002           | 0.003          | 0.009***  | 0.003 | 0.002     | 0.003            | 0.009***  | 0.003 |
| Spatial lag of $Y$                                |                 |                |           |       | 0.109*    | 0.058            |           |       |
| Chow test for structural instability              |                 | 0.88 (0.572)   |           |       |           | 14.06 (0.369)    |           |       |
| <b>Year 2017</b> ( $N = 297$ )                    |                 |                |           |       |           |                  |           |       |
| Constant  | -0.141***       | 0.049          |           |       | -0.139*** | 0.047            |           |       |
| % of the workers who studied outside the area     | 0.039*          | 0.023          | 0.046*    | 0.025 | 0.044**   | 0.022            | 0.052**   | 0.024 |
| % of 20-39 year-olds in the labour force          | 0.099           | 0.065          | -0.105**  | 0.045 | 0.124**   | 0.063            | -0.090**  | 0.043 |
| % of workers living outside their area of work    | -0.032***       | 0.012          | -0.016    | 0.010 | -0.034*** | 0.011            | -0.018*   | 0.009 |
| % of empl. in the agricultural sector             | 0.012           | 0.143          | -0.118*** | 0.038 | 0.048     | 0.137            | -0.122*** | 0.036 |
| % of empl. in the industrial sector               | -0.036          | 0.038          | -0.091*** | 0.034 | -0.048    | 0.036            | -0.099*** | 0.033 |
| % of empl. in the construction sector             | 0.009           | 0.156          | -0.101    | 0.116 | -0.021    | 0.149            | -0.146    | 0.113 |
| % of empl. in the sales sector                    | -0.008          | 0.054          | -0.096*   | 0.054 | -0.007    | 0.052            | -0.099*   | 0.051 |
| % of empl. in the educ., health and public sector | -0.060*         | 0.031          | -0.069**  | 0.031 | -0.063**  | 0.030            | -0.072**  | 0.030 |
| Supply-demand imbalance of tertiary graduates     | 0.296***        | 0.035          | 0.321***  | 0.022 | 0.282***  | 0.034            | 0.313***  | 0.021 |
| % of temporary contracts                          | -0.129*         | 0.069          | -0.111*** | 0.040 | -0.149**  | 0.066            | -0.127*** | 0.039 |
| Number of microfirms                              | 0.006           | 0.010          | 0.032***  | 0.009 | 0.006     | 0.009            | 0.031***  | 0.009 |
| Number of firms created in the last year          | -0.007          | 0.010          | -0.030*** | 0.008 | -0.007    | 0.010            | -0.030*** | 0.008 |
| Area shares a border with a foreign country       | 0.005           | 0.004          | 0.015***  | 0.004 | 0.005     | 0.003            | 0.016***  | 0.004 |
| Spatial lag of $Y$                                |                 |                |           |       | 0.109*    | 0.057            |           |       |
| Chow test for structural instability              |                 | 1.96** (0.024) |           |       |           | 30.05*** (0.005) |           |       |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Note: The coefficients associated to the constant and the spatial lag of  $Y$  is common to both regimes.

Source: Author's calculation based on French population census and REE

is insignificant for women, it is highly significant for men, meaning that they are affected by a larger labour market. The fact that women appear to be more affected by the labour market context than men, but at a more local level, tend to support the idea of a differential overeducation (Frank, 1978). Thirdly, some variables have an impact only on one gender. On the one hand, the proximity with a foreign country has a significant on women, which suggests that they are more likely to be pushed

Table 2.8: Direct, indirect and total effects - Spatial regimes SAR model

| Variables   | Urban areas |          |           | Rural areas |          |           |
|---|-------------|----------|-----------|-------------|----------|-----------|
|   | Direct      | Indirect | Total     | Direct      | Indirect | Total     |
| <b>Year 2009</b> ( $N = 297$ )                    |             |          |           |             |          |           |
| % of the workers who studied outside the area     | 0.012       | 0.003    | 0.015     | 0.071***    | 0.017**  | 0.087***  |
| % of 20-39 year-olds in the labour force          | -0.059      | -0.014   | -0.073    | -0.006      | -0.001   | -0.007    |
| % of workers living outside their area of work    | -0.016      | -0.004   | -0.019    | -0.044***   | -0.010** | -0.055*** |
| % of empl. in the agricultural sector             | -0.212*     | -0.050   | -0.262*   | -0.122***   | -0.029** | -0.151*** |
| % of empl. in the industrial sector               | -0.060**    | -0.014   | -0.074**  | -0.177***   | -0.042** | -0.219*** |
| % of empl. in the construction sector             | 0.207*      | -0.049   | 0.255*    | -0.257***   | -0.061** | -0.318*** |
| % of empl. in the sales sector                    | -0.005      | -0.001   | -0.006    | -0.178***   | -0.042** | -0.220*** |
| % of empl. in the educ., health and public sector | -0.062**    | -0.015*  | -0.076**  | -0.154***   | -0.036** | -0.191*** |
| Supply-demand imbalance of tertiary graduates     | 0.301***    | 0.071*** | 0.372***  | 0.299***    | 0.070*** | 0.369***  |
| % of temporary contracts                          | -0.086      | -0.020   | -0.106    | -0.086**    | -0.020*  | -0.106**  |
| Number of microfirms                              | -0.008      | -0.002   | -0.010    | 0.005       | 0.001    | 0.006     |
| Number of firms created in the last year          | -0.005      | -0.001   | -0.006    | -0.010**    | -0.002*  | -0.012**  |
| Area shares a border with a foreign country       | 0.001       | 0.000    | 0.001     | 0.008***    | 0.002*   | 0.010***  |
| <b>Year 2013</b> ( $N = 297$ )                    |             |          |           |             |          |           |
| % of the workers who studied outside the area     | 0.013       | 0.002    | 0.014     | 0.053***    | 0.006    | 0.060***  |
| % of 20-39 year-olds in the labour force          | -0.011      | -0.001   | -0.012    | 0.021       | 0.002    | 0.023     |
| % of workers living outside their area of work    | -0.021*     | -0.002   | -0.023*   | -0.036***   | -0.004   | -0.040*** |
| % of empl. in the agricultural sector             | 0.023       | 0.003    | 0.026     | -0.126***   | -0.015   | -0.141*** |
| % of empl. in the industrial sector               | -0.046      | -0.006   | -0.051    | -0.126***   | -0.015   | -0.141*** |
| % of empl. in the construction sector             | 0.114       | 0.014    | 0.128     | -0.177**    | -0.021   | -0.198**  |
| % of empl. in the sales sector                    | 0.014       | 0.002    | 0.016     | -0.123***   | -0.015   | -0.138*** |
| % of empl. in the educ., health and public sector | -0.066**    | -0.008   | -0.074**  | -0.105***   | -0.013   | -0.117*** |
| Supply-demand imbalance of tertiary graduates     | 0.281***    | 0.034*   | 0.314***  | 0.288***    | 0.035*   | 0.322***  |
| % of temporary contracts                          | -0.109*     | -0.013   | -0.122*   | -0.086**    | -0.010   | -0.097**  |
| Number of microfirms                              | 0.003       | 0.000    | 0.004     | 0.018**     | 0.002    | 0.020**   |
| Number of firms created in the last year          | -0.012      | -0.001   | -0.013    | -0.024***   | -0.003   | -0.027*** |
| Area shares a border with a foreign country       | 0.002       | 0.000    | 0.002     | 0.009***    | 0.001    | 0.010***  |
| <b>Year 2017</b> ( $N = 297$ )                    |             |          |           |             |          |           |
| % of the workers who studied outside the area     | 0.044**     | 0.005    | 0.050*    | 0.052**     | 0.006    | 0.058**   |
| % of 20-39 year-olds in the labour force          | 0.125**     | 0.015    | 0.140*    | -0.090**    | -0.011   | -0.101**  |
| % of workers living outside their area of work    | -0.034***   | -0.004   | -0.038*** | -0.018*     | -0.002   | -0.020*   |
| % of empl. in the agricultural sector             | 0.048       | 0.006    | 0.054     | -0.123***   | -0.015   | -0.137*** |
| % of empl. in the industrial sector               | -0.048      | -0.006   | -0.054    | -0.100***   | -0.012   | -0.111*** |
| % of empl. in the construction sector             | -0.021      | -0.002   | -0.023    | -0.146      | -0.017   | -0.164    |
| % of empl. in the sales sector                    | -0.007      | -0.001   | -0.008    | -0.099*     | -0.012   | -0.111*   |
| % of empl. in the educ., health and public sector | -0.063**    | -0.008   | -0.071**  | -0.073**    | -0.009   | -0.081**  |
| Supply-demand imbalance of tertiary graduates     | 0.283***    | 0.034*   | 0.317***  | 0.314***    | 0.037*   | 0.351***  |
| % of temporary contracts                          | -0.149**    | -0.018   | -0.167**  | -0.127***   | -0.015   | -0.142*** |
| Number of microfirms                              | 0.006       | 0.001    | 0.007     | 0.031***    | 0.004    | 0.035***  |
| Number of firms created in the last year          | -0.008      | -0.001   | -0.008    | -0.030***   | -0.004   | -0.034*** |
| Area shares a border with a foreign country       | 0.005       | 0.001    | 0.006     | 0.016***    | 0.002    | 0.018***  |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculation based on French population census and REE

down into less qualified jobs if there is an added supply of cross-border workers. On the other hand, I find a negative effect of the share of temporary contracts for men, which support the idea that repeated job search can be a way to improve matches (Jovanovic, 1979) for this demographic group.



Table 2.9: Direct, indirect and total effects - Spatial regimes SAR model - by gender

| Variables   | Women       |          |             |           | Men         |           |             |          |          |           |          |           |
|---|-------------|----------|-------------|-----------|-------------|-----------|-------------|----------|----------|-----------|----------|-----------|
|   | Urban areas |          | Rural areas |           | Urban areas |           | Rural areas |          |          |           |          |           |
|   | Direct      | Indirect | Total       | Direct    | Indirect    | Total     | Direct      | Indirect | Total    |           |          |           |
| <b>Year 2009 (N = 297)</b>                        |             |          |             |           |             |           |             |          |          |           |          |           |
| % of the workers who studied outside the area     | -0.003      | -0.001   | -0.003      | 0.085***  | 0.020*      | 0.105***  | 0.015       | 0.007    | 0.022    | 0.046**   | 0.020*   | 0.065**   |
| % of 20-39 year-olds in the labour force          | -0.230***   | -0.054** | -0.283***   | -0.066    | -0.016      | -0.082    | 0.084       | 0.036    | 0.120    | 0.046**   | 0.020    | 0.066     |
| % of workers living outside their area of work    | -0.006      | -0.001   | -0.007      | -0.047*** | -0.011*     | -0.058*** | -0.019      | -0.008   | -0.028   | -0.041**  | -0.018** | -0.059*** |
| % of empl. in the agricultural sector             | -0.300      | -0.070   | -0.370      | -0.125*** | -0.029      | -0.154*** | -0.149      | -0.064   | -0.213   | -0.129*** | -0.055** | -0.184*** |
| % of empl. in the industrial sector               | -0.086*     | -0.020   | -0.106*     | -0.210*** | -0.049**    | -0.260*** | -0.200      | -0.009   | -0.208   | -0.130*** | -0.056** | -0.186*** |
| % of empl. in the construction sector             | 0.391**     | 0.091*   | 0.482**     | -0.218*   | -0.051      | -0.268*   | 0.102       | 0.044    | 0.146    | -0.244*** | -0.105** | -0.349*** |
| % of empl. in the sales sector                    | -0.053      | -0.012   | -0.066      | -0.170*** | -0.040*     | -0.210**  | 0.054       | 0.023    | 0.077    | -0.167*** | -0.072** | -0.239*** |
| % of empl. in the educ., health and public sector | -0.132***   | -0.031*  | -0.162***   | -0.227*** | -0.053**    | -0.281*** | -0.004      | -0.002   | -0.006   | -0.092**  | -0.040** | -0.132*** |
| Supply-demand imbalance of tertiary graduates     | 0.374***    | 0.087**  | 0.461***    | 0.350***  | 0.082***    | 0.431***  | 0.218***    | 0.094*** | 0.312*** | 0.230***  | 0.099*** | 0.329***  |
| % of temporary contracts                          | -0.053      | -0.012   | -0.065      | -0.018    | -0.004      | -0.023    | -0.095      | -0.041   | -0.136   | -0.125*** | -0.054** | -0.179*** |
| Number of microfirms                              | -0.016      | -0.004   | -0.019      | -0.006    | -0.001      | -0.007    | -0.003      | -0.001   | -0.004   | 0.010     | 0.005    | 0.015     |
| Number of firms created in the last year          | -0.005      | -0.001   | -0.006      | -0.012    | -0.003      | -0.015    | -0.007      | -0.003   | -0.010   | -0.008    | -0.011   | -0.015    |
| Area shares a border with a foreign country       | -0.002      | -0.000   | -0.002      | 0.012***  | 0.003*      | 0.014**   | 0.001       | 0.001    | 0.002    | 0.004     | 0.002    | 0.005     |
| <b>Year 2013 (N = 297)</b>                        |             |          |             |           |             |           |             |          |          |           |          |           |
| % of the workers who studied outside the area     | -0.003      | 0.000    | -0.004      | 0.072***  | -0.001      | 0.071***  | 0.019       | 0.010    | 0.029    | 0.035*    | 0.018    | 0.053     |
| % of 20-39 year-olds in the labour force          | -0.131      | 0.002    | -0.130      | -0.003    | 0.000       | -0.003    | 0.101       | 0.051    | 0.152    | 0.067     | 0.034    | 0.101     |
| % of workers living outside their area of work    | -0.010      | 0.000    | -0.010      | -0.042*** | 0.001       | -0.042*** | -0.029**    | -0.015*  | -0.044** | -0.033*** | -0.017** | -0.050*** |
| % of empl. in the agricultural sector             | -0.141      | 0.002    | -0.139      | -0.163*** | 0.002       | -0.161*** | 0.146       | 0.074    | 0.220    | -0.082*   | -0.042   | -0.124*   |
| % of empl. in the industrial sector               | -0.046      | 0.001    | -0.045      | -0.152*** | 0.002       | -0.150*** | -0.022      | -0.011   | -0.033   | -0.087**  | -0.045*  | -0.132**  |
| % of empl. in the construction sector             | -0.332**    | -0.004   | -0.328**    | -0.097    | 0.001       | -0.096    | -0.031      | -0.016   | -0.047   | -0.235**  | -0.120*  | -0.354**  |
| % of empl. in the sales sector                    | 0.004       | 0.000    | 0.004       | -0.120*   | 0.002       | -0.119*   | 0.043       | 0.022    | 0.065    | -0.109**  | -0.055*  | -0.164**  |
| % of empl. in the educ., health and public sector | -0.134***   | 0.002    | -0.132***   | -0.174*** | 0.002       | -0.172*** | -0.007      | -0.004   | -0.011   | -0.055    | -0.028   | -0.082    |
| Supply-demand imbalance of tertiary graduates     | 0.345***    | -0.005   | 0.341***    | 0.323***  | 0.004       | 0.319***  | 0.210***    | 0.107*** | 0.316*** | 0.245***  | 0.125*** | 0.370***  |
| % of temporary contracts                          | 0.003       | -0.000   | 0.003       | -0.014    | 0.000       | -0.014    | -0.212***   | -0.108*  | -0.320** | -0.137*** | -0.070** | -0.206*** |
| Number of microfirms                              | -0.001      | 0.000    | -0.001      | 0.016     | -0.000      | 0.016     | 0.006       | 0.003    | 0.009    | 0.014     | 0.007    | 0.021     |
| Number of firms created in the last year          | -0.012      | 0.000    | -0.012      | -0.030*** | 0.000       | -0.029*** | -0.012      | -0.006   | -0.018   | -0.016**  | -0.008*  | -0.024**  |
| Area shares a border with a foreign country       | -0.003      | 0.000    | -0.003      | 0.014***  | -0.000      | 0.014***  | 0.004       | 0.002    | 0.006    | 0.005     | 0.002    | 0.007     |
| <b>Year 2017 (N = 297)</b>                        |             |          |             |           |             |           |             |          |          |           |          |           |
| % of the workers who studied outside the area     | 0.037       | 0.001    | 0.039       | 0.034     | 0.001       | 0.035     | 0.048       | 0.019    | 0.067    | 0.063*    | 0.024    | 0.087*    |
| % of 20-39 year-olds in the labour force          | 0.042       | 0.001    | 0.044       | -0.108*   | -0.004      | -0.112*   | 0.226***    | 0.088*   | 0.314**  | -0.069    | -0.027   | -0.096    |
| % of workers living outside their area of work    | -0.024      | -0.001   | -0.025      | -0.014    | -0.001      | -0.014    | -0.044***   | -0.017*  | -0.062** | -0.022*   | -0.009   | -0.031*   |
| % of empl. in the agricultural sector             | -0.200      | -0.008   | -0.208      | -0.158*** | -0.006      | -0.165*** | 0.285       | 0.111    | 0.396    | -0.092**  | -0.036   | -0.128*   |
| % of empl. in the industrial sector               | -0.051      | -0.002   | -0.053      | -0.098**  | -0.004      | -0.102**  | -0.033      | -0.013   | -0.046   | -0.089**  | -0.035   | -0.124*   |
| % of empl. in the construction sector             | 0.190       | 0.007    | 0.198       | 0.018     | 0.001       | 0.019     | -0.183      | -0.071   | -0.254   | -0.294**  | -0.114   | -0.409*   |
| % of empl. in the sales sector                    | -0.078      | -0.003   | -0.081      | -0.068    | -0.003      | -0.070    | 0.072       | 0.028    | 0.100    | -0.126*   | -0.049   | -0.175*   |
| % of empl. in the educ., health and public sector | -0.139***   | -0.005   | -0.137***   | -0.107*** | -0.004      | -0.111**  | -0.007      | -0.002   | -0.010   | -0.059    | -0.023   | -0.081    |
| Supply-demand imbalance of tertiary graduates     | 0.366***    | 0.014    | 0.380***    | 0.329***  | 0.013       | 0.341***  | 0.185***    | 0.072**  | 0.256*** | 0.283***  | 0.110*** | 0.393***  |
| % of temporary contracts                          | -0.079      | -0.003   | -0.082      | -0.043    | -0.002      | -0.044    | -0.215**    | -0.084*  | -0.299** | -0.188*** | -0.073** | -0.261**  |
| Number of microfirms                              | 0.000       | 0.000    | 0.000       | 0.035***  | 0.001       | 0.037     | 0.012       | 0.005    | 0.017    | 0.030**   | 0.012*   | 0.041**   |
| Number of firms created in the last year          | -0.004      | -0.000   | -0.004      | -0.038*** | -0.001      | -0.039*** | -0.013      | -0.005   | -0.019   | -0.026**  | -0.010*  | -0.036**  |
| Area shares a border with a foreign country       | 0.001       | 0.001    | 0.001       | 0.016***  | 0.001       | 0.017***  | 0.007       | 0.003    | 0.009    | 0.015***  | 0.006**  | 0.020***  |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.  
Source: Author's calculation based on French population census and REE

## 2.6 Conclusion

The aim of this work was to study the local determinants of overeducation and the disparities between local overeducation rates in metropolitan France in the 2010s. To this end, I regress overeducation rates on a large set of local characteristics at the scale of French employment areas. To take the spatial structure of the data into account and to explore how neighbouring labour markets can impact each other, I estimate spatial econometrics models that account for spatial dependency, and spatial regimes models that account for the spatial heterogeneity between urban and rural areas.

The results are as follows. I evidence the effect of local characteristics on overeducation, with strong differences between urban and rural areas. While the imbalance between supply and demand for educated work, and thus the competition between workers, is the main driver in urban areas, the local context as a whole appears to have an effect in rural areas. Especially, the industry mix plays a significant role, which questions the variety of job opportunities in these areas. Overall, the situation in neighbouring areas seems to have only limited effects, but affect significantly men overeducation. More generally, the local labour market context has heterogeneous effects across gender: women are more impacted by the labour market context than men, but at a more local level.

From a public policy point of view, these results plead for differentiated policies depending on the characteristics of the workers, but also depending on the area. Overeducation is lower in rural areas when there is less specialisation in terms of sectors. Therefore, increasing the variety of job opportunities in these areas might help reduce overeducation. Moreover, this would also lower the incentive for mismatched workers from rural areas to go to cities to find a suitable job, which overcrowd already highly competitive labour markets.

## Appendices

### Appendix 2A: Cartography of French employment areas

Figure 2.4: French employment areas - 2010 zoning



## Appendix 2B: Direct, indirect and total effects in the SAR model

The Spatial AutoRegressive (SAR) model that includes a spatially lagged dependent variable  $WY$  can be written as follows:

$$Y = \rho WY + X\beta + u$$

As the model includes  $WY$ , the coefficients  $\beta$  associated with independent variables cannot be interpreted directly. Indeed, while a change in  $X$  in an area directly impacts  $Y$  in this particular area, it also impacts  $Y$  in the other areas through  $W$ . Following LeSage and Pace (2009), to identify direct and indirect effects, I rewrite the model as follows:

$$Y = (I - \rho W)^{-1} X\beta (I - \rho W)^{-1} u$$

$$Y = \sum_{k=1}^N \beta_k X_k + (I - \rho W)^{-1} u$$

where  $k$  denotes the  $k$ th explanatory variable.

The direct, indirect and total effects of the  $k$ th explanatory variable are then given by the matrix of partial derivatives of  $E(Y)$  with respect to the  $k$ th variable in areas 1 to  $N$ :

$$\left[ \frac{\partial E(Y)}{\partial x_{1k}} \dots \frac{\partial E(Y)}{\partial x_{Nk}} \right] = (I - \rho W)^{-1} \beta_k$$

Each diagonal element of this matrix is the direct effect for an area of a change in variable  $X_k$  in this particular area. Thus, the average direct effect corresponds to the mean diagonal element of this matrix, *i.e.* the average of the diagonal terms. The other elements of this matrix are the indirect effects of a change in  $X_k$  in a given area on  $Y$  in the other areas. Thus, the average indirect effect is given by the mean row sum of the non-diagonal elements. The average total effect is the sum of the average direct effect and the indirect effect.

## Appendix 2C: Estimations results with an SDM model

Table 2.10: Estimation results - SDM model

| Variables   | X         |       | WX        |       |
|---|-----------|-------|-----------|-------|
|   | Coeff.    | SE    | Coeff.    | SE    |
| <b>Year 2009</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | -0.066    | 0.082 |           |       |
| % of the workers who studied outside the area     | 0.084***  | 0.024 | -0.079**  | 0.033 |
| % of 20-39 year-olds in the labour force          | 0.010     | 0.043 | -0.028    | 0.082 |
| % of workers living outside their area of work    | -0.038*** | 0.010 | 0.040**   | 0.017 |
| % of empl. in the agricultural sector             | -0.055    | 0.035 | -0.023    | 0.081 |
| % of empl. in the industrial sector               | -0.138*** | 0.022 | 0.130***  | 0.041 |
| % of empl. in the construction sector             | -0.181**  | 0.075 | 0.257**   | 0.117 |
| % of empl. in the sales sector                    | 0.098***  | 0.031 | 0.121**   | 0.059 |
| % of empl. in the educ., health and public sector | -0.106*** | 0.024 | 0.098**   | 0.042 |
| Supply-demand imbalance of tertiary graduates     | 0.278***  | 0.021 | -0.142**  | 0.068 |
| % of temporary contracts                          | -0.185*** | 0.039 | 0.261***  | 0.066 |
| Number of microfirms                              | -0.009*   | 0.005 | 0.011     | 0.010 |
| Number of firms created in the last year          | -0.008*   | 0.004 | 0.004     | 0.008 |
| Area shares a border with a foreign country       | 0.012***  | 0.024 | -0.013**  | 0.005 |
| Spatial lag of Y                                  |           |       | 0.572***  | 0.195 |
| <b>Year 2013</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | -0.067    | 0.077 |           |       |
| % of the workers who studied outside the area     | 0.067***  | 0.020 | -0.045    | 0.028 |
| % of 20-39 year-olds in the labour force          | 0.046     | 0.049 | -0.047    | 0.081 |
| % of workers living outside their area of work    | -0.031*** | 0.008 | 0.026     | 0.016 |
| % of empl. in the agricultural sector             | -0.087*** | 0.034 | 0.015     | 0.080 |
| % of empl. in the industrial sector               | -0.081*** | 0.021 | 0.042     | 0.046 |
| % of empl. in the construction sector             | -0.140*   | 0.076 | 0.182     | 0.126 |
| % of empl. in the sales sector                    | -0.061*   | 0.033 | 0.039     | 0.066 |
| % of empl. in the educ., health and public sector | -0.067*** | 0.022 | 0.050     | 0.045 |
| Supply-demand imbalance of tertiary graduates     | 0.288***  | 0.018 | -0.086    | 0.083 |
| % of temporary contracts                          | -0.172*** | 0.041 | 0.220***  | 0.070 |
| Number of microfirms                              | 0.013*    | 0.007 | -0.013    | 0.011 |
| Number of firms created in the last year          | -0.025*** | 0.006 | 0.022**   | 0.010 |
| Area shares a border with a foreign country       | 0.015***  | 0.003 | -0.016*** | 0.005 |
| Spatial lag of Y                                  |           |       | 0.385     | 0.252 |
| <b>Year 2017</b> ( $N = 297$ )                    |           |       |           |       |
| Constant  | -0.016    | 0.086 |           |       |
| % of the workers who studied outside the area     | 0.087***  | 0.031 | -0.073    | 0.046 |
| % of 20-39 year-olds in the labour force          | -0.032    | 0.051 | -0.046    | 0.082 |
| % of workers living outside their area of work    | -0.024*** | 0.009 | 0.021     | 0.017 |
| % of empl. in the agricultural sector             | -0.094**  | 0.036 | 0.039     | 0.089 |
| % of empl. in the industrial sector               | -0.086*** | 0.029 | 0.059     | 0.052 |
| % of empl. in the construction sector             | -0.187    | 0.124 | 0.388**   | 0.192 |
| % of empl. in the sales sector                    | -0.101**  | 0.047 | 0.102     | 0.081 |
| % of empl. in the educ., health and public sector | -0.052**  | 0.026 | 0.034     | 0.046 |
| Supply-demand imbalance of tertiary graduates     | 0.330***  | 0.023 | -0.180*   | 0.106 |
| % of temporary contracts                          | -0.184*** | 0.049 | 0.208**   | 0.084 |
| Number of microfirms                              | 0.026***  | 0.009 | -0.029**  | 0.013 |
| Number of firms created in the last year          | -0.026*** | 0.008 | 0.026**   | 0.012 |
| Area shares a border with a foreign country       | 0.015***  | 0.004 | -0.009    | 0.007 |
| Spatial lag of Y                                  |           |       | 0.510*    | 0.285 |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculation based on French population census and REE

## Appendix 2D: First-stage regression for 2SLS and GS2SLS

Table 2.11: First stage regression results

| Variables   | Coeff.    | SE    |
|---|-----------|-------|
| <b>Year 2009</b> ( $N = 297$ )                    |           |       |
| Constant  | 0.686     | 0.583 |
| % of 20-39 year-olds in the labour force          | -1.208**  | 0.535 |
| % of workers living outside their area of work    | 0.259***  | 0.092 |
| % of empl. in the agricultural sector             | -0.733*   | 0.413 |
| % of empl. in the industrial sector               | 0.333     | 0.249 |
| % of empl. in the construction sector             | 1.932**   | 0.743 |
| % of empl. in the sales sector                    | 0.698*    | 0.383 |
| % of empl. in the educ., health and public sector | 0.451     | 0.281 |
| Supply-demand imbalance of tertiary graduates     | -0.680*** | 0.174 |
| % of temporary contracts                          | 1.010**   | 0.397 |
| Number of microfirms                              | -0.003    | 0.073 |
| Number of firms created in the last year          | 0.034     | 0.061 |
| Area shares a border with a foreign country       | -0.108*** | 0.024 |
| Long-term unemployment                            | -1.674*** | 0.340 |
| Weak identification test stat                     | 24.182    |       |
| <b>Year 2013</b> ( $N = 297$ )                    |           |       |
| Constant  | 1.402**   | 0.583 |
| % of 20-39 year-olds in the labour force          | -1.744*** | 0.516 |
| % of workers living outside their area of work    | 0.108     | 0.093 |
| % of empl. in the agricultural sector             | -0.206    | 0.407 |
| % of empl. in the industrial sector               | 0.172     | 0.263 |
| % of empl. in the construction sector             | 2.291***  | 0.733 |
| % of empl. in the sales sector                    | 0.935**   | 0.378 |
| % of empl. in the educ., health and public sector | 0.305     | 0.280 |
| Supply-demand imbalance of tertiary graduates     | -0.635*** | 0.171 |
| % of temporary contracts                          | 1.324***  | 0.373 |
| Number of microfirms                              | -0.223*** | 0.069 |
| Number of firms created in the last year          | 0.246***  | 0.060 |
| Area shares a border with a foreign country       | -0.110*** | 0.025 |
| Long-term unemployment                            | -2.183*** | 0.345 |
| Weak identification test stat                     | 39.965    |       |
| <b>Year 2017</b> ( $N = 297$ )                    |           |       |
| Constant  | 1.286**   | 0.628 |
| % of 20-39 year-olds in the labour force          | -1.089**  | 0.542 |
| % of workers living outside their area of work    | 0.115     | 0.100 |
| % of empl. in the agricultural sector             | -0.162    | 0.468 |
| % of empl. in the industrial sector               | 0.514*    | 0.290 |
| % of empl. in the construction sector             | 2.944***  | 0.920 |
| % of empl. in the sales sector                    | 1.181***  | 0.415 |
| % of empl. in the educ., health and public sector | 0.454     | 0.291 |
| Supply-demand imbalance of tertiary graduates     | -0.689*** | 0.189 |
| % of temporary contracts                          | 1.193***  | 0.420 |
| Number of microfirms                              | -0.185**  | 0.088 |
| Number of firms created in the last year          | 0.169**   | 0.078 |
| Area shares a border with a foreign country       | -0.111*** | 0.029 |
| Long-term unemployment                            | -1.537*** | 0.396 |
| Weak identification test stat                     | 15.097    |       |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculation based on French population census and REE

## Appendix 2E: Spatial regimes SAR model - Estimations results by gender

Table 2.12: Estimation results by gender - SAR spatial regimes models

| Variables   | Women     |        |           |       | Men       |          |           |       |
|---|-----------|--------|-----------|-------|-----------|----------|-----------|-------|
|   | Urban     |        | Rural     |       | Urban     |          | Rural     |       |
|   | Coeff.    | SE     | Coeff.    | SE    | Coeff.    | SE       | Coeff.    | SE    |
| <b>Year 2009</b> ( $N = 297$ )                    |           |        |           |       |           |          |           |       |
| Constant  | 0.146**   | 0.061  |           |       | -0.070    | 0.044    |           |       |
| % of the workers who studied outside the area     | -0.003    | 0.029  | 0.084***  | 0.031 | 0.015     | 0.020    | 0.045**   | 0.022 |
| % of 20-39 year-olds in the labour force          | -0.228*** | 0.085  | -0.066    | 0.063 | 0.082     | 0.059    | 0.045     | 0.044 |
| % of workers living outside their area of work    | -0.006    | 0.018  | -0.047*** | 0.015 | -0.019    | 0.013    | -0.040*** | 0.011 |
| % of empl. in the agricultural sector             | -0.297    | 0.184  | -0.124*   | 0.055 | -0.146    | 0.127    | -0.126*** | 0.038 |
| % of empl. in the industrial sector               | -0.085*   | 0.044  | -0.209*** | 0.042 | -0.019    | 0.031    | -0.127*** | 0.029 |
| % of empl. in the construction sector             | 0.388**   | 0.175  | -0.216*   | 0.122 | 0.100     | 0.121    | -0.239*** | 0.084 |
| % of empl. in the sales sector                    | -0.053    | 0.063  | -0.169*** | 0.064 | 0.052     | 0.044    | -0.163*** | 0.045 |
| % of empl. in the educ., health and public sector | -0.131*** | 0.042  | -0.226*** | 0.044 | -0.004    | 0.030    | -0.090*** | 0.031 |
| Supply-demand imbalance of tertiary graduates     | 0.371***  | 0.045  | 0.347***  | 0.031 | 0.214***  | 0.031    | 0.225***  | 0.022 |
| % of temporary contracts                          | -0.052    | 0.092  | -0.018    | 0.053 | -0.093    | 0.064    | -0.123*** | 0.036 |
| Number of microfirms                              | -0.015    | 0.012  | -0.006    | 0.010 | -0.003    | 0.008    | 0.010     | 0.007 |
| Number of firms created in the last year          | -0.005    | 0.012  | -0.012    | 0.007 | -0.003    | 0.008    | -0.008    | 0.005 |
| Area shares a border with a foreign country       | -0.002    | 0.005  | 0.012***  | 0.004 | 0.001     | 0.003    | 0.004     | 0.003 |
| Spatial lag of Y                                  | 0.196***  | 0.065  |           |       | 0.316***  | 0.068    |           |       |
| Chow test for structural instability              |           | 14.43  | (0.344)   |       |           | 29.89*** | (0.005)   |       |
| <b>Year 2013</b> ( $N = 297$ )                    |           |        |           |       |           |          |           |       |
| Constant  | 0.057     | 0.065  |           |       | -0.095*   | 0.053    |           |       |
| % of the workers who studied outside the area     | -0.004    | 0.022  | 0.072***  | 0.024 | 0.019     | 0.018    | 0.034*    | 0.020 |
| % of 20-39 year-olds in the labour force          | -0.131    | 0.084  | -0.003    | 0.065 | 0.098     | 0.070    | 0.065     | 0.054 |
| % of workers living outside their area of work    | -0.010    | 0.016  | -0.042*** | 0.014 | -0.028**  | 0.013    | -0.032*** | 0.011 |
| % of empl. in the agricultural sector             | -0.141    | 0.186  | -0.163*** | 0.051 | 0.142     | 0.151    | -0.080*   | 0.042 |
| % of empl. in the industrial sector               | -0.046    | 0.044  | -0.152*** | 0.040 | -0.021    | 0.036    | -0.085*** | 0.033 |
| % of empl. in the construction sector             | 0.332**   | 0.167  | -0.097    | 0.118 | -0.030    | 0.138    | -0.228**  | 0.097 |
| % of empl. in the sales sector                    | 0.004     | 0.068  | -0.120*   | 0.063 | 0.042     | 0.056    | -0.105**  | 0.052 |
| % of empl. in the educ., health and public sector | -0.134*** | 0.044  | -0.174*** | 0.041 | -0.007    | 0.036    | -0.053    | 0.033 |
| Supply-demand imbalance of tertiary graduates     | 0.345***  | 0.040  | 0.323***  | 0.027 | 0.204***  | 0.033    | 0.238***  | 0.022 |
| % of temporary contracts                          | 0.003     | 0.094  | -0.014    | 0.051 | -0.206*** | 0.075    | -0.133*** | 0.041 |
| Number of microfirms                              | -0.001    | 0.011  | 0.016     | 0.010 | 0.006     | 0.009    | 0.013     | 0.008 |
| Number of firms created in the last year          | -0.012    | 0.012  | -0.030*** | 0.009 | -0.012    | 0.010    | -0.016**  | 0.007 |
| Area shares a border with a foreign country       | -0.003    | 0.004  | 0.014***  | 0.004 | 0.004     | 0.003    | 0.004     | 0.003 |
| Spatial lag of Y                                  | -0.013    | 0.073  |           |       | 0.356***  | 0.075    |           |       |
| Chow test for structural instability              |           | 16.11  | (0.243)   |       |           | 9.75     | (0.714)   |       |
| <b>Year 2017</b> ( $N = 297$ )                    |           |        |           |       |           |          |           |       |
| Constant  | -0.100    | 0.068  |           |       | -0.165*** | 0.062    |           |       |
| % of the workers who studied outside the area     | 0.037     | 0.032  | 0.034     | 0.035 | 0.047     | 0.030    | 0.062*    | 0.032 |
| % of 20-39 year-olds in the labour force          | 0.042     | 0.091  | -0.107*   | 0.063 | 0.222***  | 0.082    | -0.068    | 0.057 |
| % of workers living outside their area of work    | -0.024    | 0.017  | -0.014    | 0.014 | -0.044*** | 0.015    | -0.022*   | 0.012 |
| % of empl. in the agricultural sector             | -0.200    | 0.199  | -0.158*** | 0.053 | 0.280     | 0.180    | -0.091*   | 0.048 |
| % of empl. in the industrial sector               | -0.051    | 0.052  | -0.098**  | 0.047 | -0.033    | 0.049    | -0.087**  | 0.043 |
| % of empl. in the construction sector             | 0.190     | 0.216  | 0.018     | 0.164 | -0.180    | 0.197    | -0.289*   | 0.148 |
| % of empl. in the sales sector                    | -0.078    | 0.075  | -0.068    | 0.074 | 0.071     | 0.068    | -0.124*   | 0.067 |
| % of empl. in the educ., health and public sector | -0.132*** | 0.043  | -0.107**  | 0.043 | -0.007    | 0.039    | -0.058    | 0.039 |
| Supply-demand imbalance of tertiary graduates     | 0.366***  | 0.049  | 0.329***  | 0.031 | 0.181***  | 0.045    | 0.277***  | 0.028 |
| % of temporary contracts                          | -0.079    | 0.097  | -0.043    | 0.058 | -0.211**  | 0.086    | -0.184*** | 0.051 |
| Number of microfirms                              | 0.000     | 0.013  | 0.035***  | 0.013 | 0.012     | 0.012    | 0.029**   | 0.012 |
| Number of firms created in the last year          | -0.004    | 0.014  | -0.038*** | 0.011 | -0.013    | 0.013    | -0.025**  | 0.010 |
| Area shares a border with a foreign country       | 0.001     | 0.005  | 0.016***  | 0.005 | 0.007     | 0.005    | 0.014***  | 0.005 |
| Spatial lag of Y                                  | 0.037     | 0.073  |           |       | 0.293***  | 0.079    |           |       |
| Chow test for structural instability              |           | 22.26* | (0.051)   |       |           | 22.60**  | (0.047)   |       |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Note: The coefficients associated to the constant and the spatial lag of Y is common to both regimes.

Source: Author's calculation based on French population census and REE





## Chapter 3

# Spatial mobility and overeducation of young workers

This chapter explores the influence of spatial mobility on the risk of overeducation of French young workers. Mobilising a survey following a cohort of young graduates entering the labour market in 2010 until 2013, our results reveal that interregional migration decreases the risk of (statistical and subjective) overeducation. We also evidence that migration to an economic centre (the Paris region) has an even stronger negative effect, and that more educated workers benefit more from spatial mobility. These results are robust to controlling for self-selection and the endogeneity of migration, as well as to various specifications of the model.

**JEL classification:** C35, J24, R23

**Keywords:** overeducation, educational mismatch, spatial mobility, residential migration, instrumental variables approach

---

This chapter reviews a work co-written with Florent Sari entitled *Spatial mobility and overeducation of young workers: New evidence from France*, accepted for publication in *Papers in Regional Science* (Fouquet & Sari, 2023). I would like to thank participants in the 20th International Workshop on Spatial Econometrics and Statistics, the 70th Congress of the French Economic Association (AFSE), the 63rd New Zealand Association of Economists Annual (NZAE) Conference, the Winter School on Public Policy Evaluation (ETEPP) 2022 and the ERUDITE and GAINS seminars for their comments on earlier versions of this chapter.

# Mobilité spatiale et sur-éducation des jeunes diplômés

Ce chapitre vise à étudier l'impact de la mobilité spatiale sur le risque de sur-éducation des jeunes travailleurs en France. À partir des données de l'enquête *Génération 2010* du Céreq, qui suit jusqu'en 2013 une cohorte d'individus ayant quitté le système éducatif en 2010, nous montrons que la migration résidentielle inter-régionale permet de réduire le risque de sur-éducation (à la fois statistique et subjective). Nos résultats montrent un effet négatif plus important de la migration vers un centre économique (la région parisienne). Par ailleurs, les plus diplômés bénéficient plus fortement de la mobilité spatiale. Ces résultats sont robustes au contrôle d'un effet de sélection dans l'emploi et de l'endogénéité de la décision de mobilité, ainsi qu'à diverses spécifications du modèle.

**Mots-clés :** sur-éducation, inadéquation emploi-diplôme, mobilité spatiale, migration résidentielle, variables instrumentales

## 3.1 Introduction

Most job seekers are searching for jobs on local labour markets, around their residential location. However, smaller markets often mean less suitable job opportunities. If there is no adequate job available in the area, an individual has three options (Simpson, 1992). The first is to remain unemployed. The second is to accept a job requiring less education than she or he actually has, *i.e.* become overeducated. To avoid both unemployment and overeducation, the third option is to search for jobs located farther from them. The economic literature has indeed shown that spatial mobility tends to reduce the risk of unemployment as it permits to access to new areas with better job opportunities (see, for instance, the works of Détang-Dessendre, 1999; Fahr & Sunde, 2006; Langella & Manning, 2022; Pissarides & McMaster, 1990; Rupert & Wasmer, 2012). A recent work by Schmutz et al. (2021) confirms that this is especially true for the most educated workers. While the links between spatial or geographical mobility and unemployment have been the subject of numerous empirical analyses, it is not the case concerning the links with overeducation. This phenomenon, which particularly concerns the most educated young workers, remains somewhat under-studied.

In France, as in many OECD countries, young graduates entering the labour market experience important issues in their early professional careers. Firstly, they face a higher risk of unemployment. In 2021, 19.6% of the 15-24 years old in the labour force were unemployed<sup>1</sup>. Secondly, they also face a higher risk of overeducation (Dolton & Vignoles, 2000). However, being overeducated at the beginning of the career have a detrimental impact on future labour market outcomes (Baert & Verhaest, 2019; Baert et al., 2013), due to the negative signal of overeducation and the decline of skills over time. Then, it is essential for young workers to access a matched job as soon as possible after their graduation. In this context, broadening their job search area may be a strategy for these young workers to avoid overeducation and its scarring effects. In the French case, although young graduates generally tend to stay in the same region for higher education and employment, their share has declined over the past few decades (Bernela & Bonnal, 2022), the spatial mobility allowing them to increase chances to find suitable and better-paid jobs (Lemistre & Moreau, 2009).

In this work, we contribute to the existing literature by investigating the conse-

---

<sup>1</sup>Source: OECD data.

quences of spatial mobility for young workers in their early careers. In particular, we examine the links between geographical mobility (between graduation and the job currently occupied) and the risk of overeducation, the latter being defined as the fact of occupying a job requiring an educational level lower than one's own. Previous works have indeed shown that expanding the job search area could be an important factor in avoiding overeducation in OECD countries (see e.g. Büchel & van Ham, 2003; Devillanova, 2013; Hensen et al., 2009; Romaní et al., 2016, respectively for Germany, Italy, the Netherlands and Spain). However, to our knowledge, such a question has not been studied yet in the French context. There is previous evidence of a negative effect of residential migration on the risk of unemployment, especially for highly educated workers (Détang-Dessendre, 1999; Détang-Dessendre et al., 2004). Still, Bernela and Bonnal (2022) showed that French young workers were not spatially mobile, with less than 30% relocating after graduation, even though the most educated are more likely to move, especially to bigger cities with more labour market prospects (Détang-Dessendre et al., 2004). Therefore, we document the differentiated effects of spatial mobility on overeducation depending on the destination of the residential migration: in addition to the analysis of interregional migrations nationwide, we make a specific focus on migrations to the Paris region. Although Paris is not the only economic centre in France, it accounts for 20% of the country's labour force and concentrates a large pool of job opportunities as well as higher wages (Combes et al., 2015). Similar geographical structures can be observed in other countries (see e.g. Andersson et al., 2014; Faggian & McCann, 2009; Venhorst et al., 2010, respectively for Sweden, the UK and the Netherlands). Previous studies showed a negative effect of migration towards economic centres on overeducation (Iammarino & Marinelli, 2015), but overall, the effects of residential migration remain dependent on the characteristics of the destination region (DeviIlanova, 2013; Jauhiainen, 2011). However, Venhorst et al. (2010) evidenced significant migration flows between peripheral regions in The Netherlands, which remained attractive even for highly educated workers.

In order to examine the relationship between spatial mobility and the risk of overeducation, we use the *Génération 2010* survey from the Céreq (Centre d'études et de recherche sur les qualifications, French Center for Studies and Research on Qualifications), allowing us to follow a cohort of French young people who left the educational system in 2010 until 2013. Because the survey provides information on residential mobility, it is used to check if a regional migration since graduation is a

way to avoid overeducation. The empirical design retained to analyse the relationship between migration and overeducation combines three econometric models. The first one is to account for the sample selection problem, as overeducation is observable only for individuals that actually work. The second one is for the endogeneity of migration decision. Indeed, there may be some unobserved factors simultaneously determining migration and overeducation, or some reverse causality, as overeducation might affect the migration decision. For these reasons, we run instrumental variables estimations, in which the migration decision is instrumented by migrations and stays abroad observed during past education. The last equation is for testing the relationship between the migration decision and the risk of overeducation on the labour market. This econometric strategy should allow us to better identify the causal effect of spatial mobility (or migration) on overeducation.

Our approach is original for at least three reasons. Firstly, if there are already some works investigating the role of migration and the implications for education-job matching issues, there is none for France. It represents a particularly interesting study case as the Paris region polarises a significant proportion of jobs and workers. In this context, a migration to Paris may have differentiated effects compared to migration to other regions. We try to disentangle these questions by examining all migrations between regions and those only related to the Paris region. Secondly, our work is based on a representative sample of the stock of both short-cycle higher education and long-cycle higher education graduates. By studying these different populations, we want to check if migration is more favourable for the most educated profiles. Combes et al. (2012) have indeed shown that the relocation of workers between French local labour markets may differ according to their skills. Thirdly, our data enable us to use two measures of overeducation, a statistical measure and a subjective one. More specifically, the statistical measure is based on statistical characteristics of the distribution of education for each occupation, while the subjective one is based on workers' self-assessments about the skills required to do their job. We can therefore explore the actual situation of overeducation and the feeling of being overeducated, respectively.

The main results are as follows. First, we find that having migrated (*i.e.* changed of region) between the last degree obtained and the job currently held tends to decrease the probability of being overeducated. It is confirmed for both of our measures of overeducation, the statistical and the subjective ones. Considering self-selection into employment and the endogeneity of migration decision does not

change conclusions for the statistical measure, but makes the effect observed for the subjective one significant. In addition, when we focus on the migration to Paris, we evidence an even stronger negative effect on the risk of overeducation, confirming the advantage of such a region in terms of job opportunities. Finally, we show that the benefits of migration may differ according to the educational level of young workers. In fact, no effect is found for individuals with a 2-year degree, while it is particularly “advantageous” for those with a 3- or 4-year degree. In comparison, individuals with at least a 5-year degree also decrease their risk of being overeducated with spatial mobility, but the effect is smaller than for the previous group.

Section 3.2 offers a description of the theoretical links between migration decision and overeducation and a literature review of the previous empirical works exploring this question. Section 3.3 presents the data and variables retained for our analysis. Section 3.4 provides some descriptive statistics of our study sample. Section 3.5 describes the estimation strategy and Section 3.6 displays the results of our estimations. Section 3.7 concludes.

## 3.2 Spatial mobility and overeducation

### 3.2.1 Theoretical links

Individuals working in occupations for which they have more education than what is required are considered overeducated. This phenomenon is known to particularly affect young workers entering the labour market (Dolton & Vignoles, 2000). Theoretically, the career mobility theory (Sicherman & Galor, 1990) exposes overeducation as a temporary phenomenon, where workers accept a job requiring less education than they actually have in order to improve future outcomes on the labour market. Baert et al. (2013) challenge this conclusion by showing that being overeducated at the beginning of the career is detrimental to future labour market positions. Moreover, in the short term, overeducated workers have lower wages (Verdugo & Verdugo, 1989) and a lower satisfaction at work (Tsang et al., 1991) than individuals with the same level of education but working in a matched occupation. Therefore, extending their geographic mobility may be a strategy for these individuals to avoid overeducation and its scarring effects, as in the theory of Simpson (1992).

Concretely, it is expected that the risk of overeducation is higher for the work-

ers who are restricted to a smaller local labour market. In such places, the lower density of suitable job opportunities is likely to hinder the job search process and the matching quality between workers and job vacancies. As a consequence, spatial mobility can be seen as a way of avoiding overeducation because it provides access to more job opportunities. Some previous works have indeed found that geographic mobility positively influences workers' labour market positions (Büchel & Battu, 2003; Hensen et al., 2009; van Ham, 2001).

In this work, we start from this assumption, and we hypothesise that geographically mobile graduates have a lower probability of being overeducated compared to those who are less mobile. With our data, we are able to see if French young workers have changed their region of residence between the last degree obtained and the moment when they are interrogated for the survey (three years later). We believe that those who make the decision to move between the two dates are less likely to be overeducated. The mobility decision is indeed based on arbitrage between the costs and benefits associated with it. Because the costs of mobility may be high (due to financial and psychological costs), the job occupied must offer good working conditions. That is to say, it has to be matched to the worker's skills, and the wage offered should partially or fully offset the costs associated with the mobility decision. To the extent that job opportunities are richer in the Paris region than in other regions, we assume that the probability of finding a well suitable job is higher in this region, and then that the risk of overeducation should be lower.

In addition, it has been shown that spatial mobilities or migrations are more frequent for the most skilled people (Combes et al., 2012). Starting from this fact, if we consider different groups of young workers with different behaviours related to migration (*i.e.* the highly skilled workers who have a high propensity to migrate and the skilled workers who do it less frequently), some differentiated effects on the probability of being overeducated might be observed. We can indeed consider that the expected benefits of migration are higher for those who are less likely to adopt such behaviour. Indeed, if the decision to migrate has been taken, it is theoretically to obtain a job corresponding to the skills offered. Concerning the workers for whom these mobilities are frequent, the expected benefits could be lower. We try to disentangle these questions by considering the short- and long-cycle higher education graduates separately.

Finally, the links between spatial mobility and overeducation are likely to be different between men and women. Frank (1978) argues that, in a couple, the

husband is looking for a job on the global labour market, whereas the wife is looking only in the local labour market where her partner has obtained a job. Then, we can expect that women in a couple are more likely to be overeducated due to a smaller job search area. Büchel and Battu (2003) confirm this hypothesis as they show that the probability of being overeducated is higher for married women living in rural areas. The size of the local labour market or access to other labour markets through some mobilities can help them to find a suitable job. However, when marital status is taken into account, Büchel and van Ham (2003) provide evidence that women experience a penalty in their chances of finding a job, but are not more prone to overeducation than men. The authors also argue that gender differences in household roles may be important in influencing women to accept job offers closer to home. If there are children in the household, women are more likely to be restricted in their spatial mobility. On this point, empirical evidence is mixed: some authors do not find any significant effect on overeducation for women having children (Devillanova, 2013; Ramos & Sanromá, 2013), while some others observe a positive one (Büchel & van Ham, 2003), or even a negative one (Jauhiainen, 2011).

### 3.2.2 Literature review

In recent decades, several studies have examined the connections between spatial mobility and the risk of overeducation, with a primary focus on local labour markets in OECD countries, particularly concerning young individuals with higher education. However, two main issues need to be addressed to establish a clear causal relationship between the decision to move and overeducation. The first one is a possible “selection bias” since overeducation is observable only among graduates who are currently employed. The second one is the endogeneity of the mobility decision. There might be unobserved characteristics influencing both mobility behaviour and the likelihood of being overeducated. While the pioneering studies may not have adequately considered these biases, more recent ones have made attempts to address them. However, identifying suitable instrumental variables to tackle these issues remains challenging.

Büchel and van Ham (2003) were among the first to highlight the role of regional labour market size (determined by the ability to search far away from the place of residence) as a potential explanatory variable of overeducation. Based on German data, the authors find that this risk decreases with individual spatial mobility and



increases with the time needed to travel to a large agglomeration. They confirm that the size of the labour market is an important factor in avoiding overeducation. However, a limitation arises when the authors take into account the self-selection into employment through a bivariate probit model. In doing so, they fail to consider the endogenous relationship between the mobility decision and overeducation, which could potentially bias their results. Using Dutch data, Hensen et al. (2009) investigate the relationship between spatial mobility and education-job mismatch. Although they do not address the two problems mentioned earlier, they found that being more mobile reduces vertical educational mismatches (*i.e.* when the worker's level of education is higher than what is required for his/her employment) but increases horizontal educational mismatches (*i.e.* when the worker's field of education is different from the mode field for his/her occupation). In other words, the risk of being overeducated is lower for spatially flexible workers, but they are more often employed in jobs outside their field of study.

Jauhiainen (2011) also investigates the influence of mobility on the risk of overeducation in Finnish regional labour markets. Using a probit model that controls for selection bias (the Heckman two-step method), she finds an ambiguous effect of spatial mobility on the probability of being overeducated. Specifically, the study reveals that moving to another region has a negative impact on the likelihood of being overeducated, while moving within the same region has a positive effect. This finding is consistent with the idea of limited spatial flexibility increasing overeducation risk. Employing a similar methodology, Ramos and Sanromá (2013) examine the influence of local characteristics and spatial mobility on overeducation in Spanish local labour markets. Their results were consistent with previous findings, demonstrating that the size of local labour markets and the opportunity to expand job search areas (via commuting) are relevant factors that reduce the risk of overeducation.

Devillanova (2013) goes a step further by proposing an analysis that simultaneously considers self-selection and the endogeneity of migration. The results of this study indicate a negative correlation between short-distance mobility and the probability of Italians being overeducated, but they do not reveal a clear relationship for internal migration. As a result, the author concludes that the links between migration and overeducation remain uncertain, emphasizing the need for further research to better identify the underlying effects. Other researchers have also focused on the Italian local labour markets. For instance, Croce and Ghignoni (2015) tend to confirm the relevance of spatial mobility as an explanatory factor

for the individual risk of educational mismatch. In line with Devillanova (2013), they control for self-selection into employment and test for the endogeneity of spatial mobility. However, their results differ somewhat from the previous findings, as they reveal a negative effect for both commuting time (short-distance mobility) and migration. Interestingly, the effects differ based on educational level: commuting is favourable for upper-secondary graduates, while migration is more beneficial for university graduates.

Iammarino and Marinelli (2015) study the determinants of education-job mismatches among recent university graduates in Italy, with a particular focus on the role of inter-regional migration. By controlling for both the endogenous relationship between migration and employment and the self-selection bias, they find that while migration at the national level decreases the probability of overeducation, significant differences emerge when considering the sub-national dimension. In the French context, our goal is to identify similar relationships by distinguishing between Paris and other regions concerning migration. Moving on to another study, Meliciani and Radicchia (2016) specifically investigate the role of informal recruitment channels on the risk of being overeducated, while also considering its impact on spatial mobility. They estimate different equations allowing them to control for selection into employment and a likely endogeneity bias for mobility decisions. The main finding indicates that using an informal channel has a positive effect on the risk of overeducation and a negative effect on migration. Another important result is that migration may reduce the risk of overeducation, although this finding is only applicable to specific geographical areas of the Italian territory, consistent with the work of Iammarino and Marinelli (2015).

Some recent papers explore spatial mobility at different levels. For instance, Di Paolo et al. (2017) examine mobility at a very local level, focusing on the effect of job accessibility on job-education mismatch in the metropolitan area of Barcelona. Their findings, obtained from a joint model for car ownership, employment selectivity and mismatch, confirm that having access to private vehicles reduces the risk of job-education mismatch. Additionally, they reveal that public transport job accessibility also directly influences this mismatch problem. Venhorst and Cörvers (2018) estimate the impact of internal migration on job-match quality for university and college graduates in the Netherlands. Using an instrumental variables approach to control for self-selection biases among migrants, they find a positive effect of migration on the likelihood of achieving a vertical education match and giving positive sub-

jective evaluations of job-match quality. However, no significant effect is evidenced on the likelihood of achieving a horizontal match. Most recently, Ghosh and Grassi (2020) evaluated the effects of international mobility on the education-job mismatch of Italian PhD graduates. By controlling for self-selection into cross-border mobility, they highlight that migration to foreign countries significantly reduces the risk of overeducation.

The various studies reveal contrasting effects depending on the scale of spatial mobility considered. The effects of migration on the risk of overeducation are generally significant for long distances within a country or for international migrations. At the country level, migrations to some specific regions may be favourable, while they may be unfavourable for others. These effects also depend on individual characteristics. For instance, migration is likely to be more beneficial for certain levels of education. This research contributes to the existing literature by examining, in the context of France, the relationships between inter-regional mobility and job-education mismatch, and by distinguishing individuals based on their education levels.

## 3.3 Data and variables

### 3.3.1 Data

This work is based on data from the *Génération 2010* survey conducted by Céreq in spring 2013. The survey captures the entry into and trajectories during the first year in the French labour market of 33547 individuals who left school for the first time between October 2009 and October 2010. Random sampling was employed to ensure the representativeness of the data at the regional level and for each educational level. The data provide socio-demographic information about individuals (gender, age, cohabitation status) and detailed information about their education (highest degree obtained, field of study). Furthermore, the survey also includes information about the successive labour market positions (occupation, firm size, business sector) of these individuals from their graduation to the time they were surveyed.

The focus of this study is on higher education graduates. Workers with lower education are less likely to be overeducated, and in some cases, it is not even possible to be overeducated for the lowest degrees. The sample is divided into three levels

of education based on the number of years of study after the baccalaureate: the 2-year university graduates, the 3- and 4-year university graduates, and the long-cycle higher education graduates who have studied for five years or more. Self-employed workers are excluded from the sample. After these restrictions, our final sample consists of 15865 individuals who were surveyed between 28 and 45 months after their graduation. The study focuses on the labour market position at the time of the survey because it is the only time point for which information about the place of residence is available, which is needed to determine spatial mobility. Additionally, some other variables, such as cohabitation status, are only observed at the time of the survey.

Focusing on higher education graduates in the first three years after graduation is not a trivial choice. It ensures to have a fairly homogenous sample, especially in terms of professional experience, which may have a negative effect on overeducation through job changes after some time (Rubb, 2003). Moreover, all individuals graduated in a short time window, they thus experienced similar economic conditions when entering the labour market (Croce & Ghignoni, 2012). Because they face higher risks of overeducation (Dolton & Vignoles, 2000) and less favourable labour market prospects in general (Bell & Blanchflower, 2011; Blanchflower & Freeman, 2000), young workers also are of particular interest for targeted public policies. Furthermore, even though observing migration only in the first years after graduation has some limitations (Greenwood, 1997), policies to support mobility may be more adapted for younger workers, since they have lower mobility costs (less likely to have children, more often tenants etc.). Besides, in the French context, young students do not necessarily have to move to access higher education: almost all fields of study are available in all regions, with universities accessible with only short mobilities. Since the propensity to relocate is related to past migration (DaVanzo & Morrison, 1981), young workers entering the labour market may not engage in spatial mobility without incentives. The findings of Bernela and Bonnal (2022) somehow confirm this idea, showing that migration between graduation and employment was increasing over time but remained quite rare. Nevertheless, previous works showed that higher skilled workers were more likely to relocate (Combes et al., 2012; Greenwood, 1997), especially for job-related reasons (Détang-Dessendre et al., 2004). For this specific demographic group, residential migration might also be a necessity since highly educated jobs are not available everywhere, but rather tend to be concentrated in specific areas (Détang-Dessendre, 1999).

### 3.3.2 Variables

#### Dependent variable: the risk of overeducation

In this work, we are interested in the risk of overeducation among young graduates. The existing literature offers three main approaches to measuring this phenomenon:

1. *Objective measure*: This method defines what should be considered the “normal” match between education and occupations, based on the skills required for each occupation (Rumberger, 1981).
2. *Subjective measures*: This method is based on workers’ self-assessments regarding the degree they believe is required to do or to get their job (Duncan & Hoffman, 1981).
3. *Statistical measures*: It relies on statistical characteristics of the education distribution for each occupation, such as the mean (Verdugo & Verdugo, 1989) or the mode (Croce & Ghignoni, 2012; Davia et al., 2017).

Our dataset enables us to employ two different measures of mismatch. We can use the subjective and the statistical ones. Therefore, we can explore the feeling of being overeducated and the actual situation of overeducation, respectively.

Regarding the subjective measure, respondents are asked in the *Génération* survey if they are employed above, below or at their skill level<sup>2</sup>. However, it is important to note that the answer to this question may reflect skill mismatch, *i.e.* overskilling, rather than educational mismatch, *i.e.* overeducation. Overskilling refers to the underutilization of the skills possessed by workers, while overeducation refers to the difference between the level of schooling they have acquired and the level required for the job they currently occupy. It has been shown by Allen and van der Velden (2001) that these two dimensions of mismatch do not overlap systematically. Therefore, it becomes quite complicated to precisely determine which dimension respondents consider when answering the survey. To avoid any confusion, in this study, we will refer to subjective mismatch. Individuals will be considered subjectively mismatched if they indicate that they are employed below their level of skills. In this case, our dependent variable takes the value 1 and 0 otherwise.

---

<sup>2</sup>The following question is asked: “About this job, would you say you are used: At your skill level (1); Below your skill level (2); Above your skill level (3)?”.

There is no ambiguity between overeducation and overskilling when using the statistical measure. We mobilize a bi-modal measure, where individuals are considered statistically overeducated if their educational level is above the higher of the two most common levels required for their specific occupation. By retaining two modes instead of one, we account for the wide range of educational levels in consideration and the need to ensure a sufficiently large sample size within occupational groups. In order to determine what degrees are considered “normal” for each occupation, we do not limit ourselves to the sample of young workers. This decision is made to avoid potential alterations in the distribution of education by occupation since younger workers are more likely to experience overeducation compared to older workers. Therefore, we mobilize the 2013 French Labour Force Survey (*Enquête Emploi*) from Insee (Institut National de la Statistique et des Études Économiques), which corresponds to the year when respondents were surveyed for *Génération 2010*. This dataset provides information on workers of all age groups, including educational levels and occupations at a detailed level, enabling us to establish a precise correspondence between them. The classifications of educational and occupational levels, as well as the aforementioned education-occupation correspondence, are reported in Appendix 3A.

### **Independent variables**

The *Génération 2010* survey provides us a large set of explanatory variables about individuals’ socio-economic characteristics, education, and labour market positions. To explain the risk of educational mismatch, we gather various information about young graduates and the jobs they occupy.

Regarding individual characteristics, we take into account gender and age at the time of graduation. In addition, we create binary variables to indicate whether individuals live with a partner or with their parents (the reference being “living alone”) and whether they have children. Having children is also associated with the fact of being a woman. It is a way to check to what extent household configuration matters. In particular, women who are primarily concerned with their children may restrict their spatial horizon of job search and be more often overeducated. To analyse the effects of education and field of study, we create dummy variables. It is worth noting that graduating in certain fields may lead to a higher risk of mismatches compared to others (Meroni & Vera-Toscano, 2017). As for the educational level, the impact on overeducation remains theoretically ambiguous. On the one hand,

having a higher degree may indicate higher skills and knowledge, offering some level of protection against overeducation. On the other hand, higher degrees may exceed the standard requirements for a larger proportion of occupations, resulting in more frequent instances of mismatches.

Regarding the labour market position of the individuals, we incorporate information about the employer and job characteristics. Firstly, we include dummies for the sector of the firm since there may be more instances of overeducation in some sectors compared to others (Nauze-Fichet & Tomasini, 2002). Secondly, we introduce dummies to account for full-time work and permanent jobs. Young workers might be less selective when considering temporary jobs compared to permanent positions. Consequently, they may accept a job for which they are overeducated if they perceive it as a transitional position. This is what would predict the career mobility theory (Sicherman & Galor, 1990). We also include the time spent (in months) between graduation and the first job held for each individual. Because it provides information on the duration of job search, it can be argued that a longer period of time is more likely to lead to a suitable job.

In this work, we argue that one of the main factors determining the risk of overeducation is the spatial mobility of graduates. By examining the regions where young people have obtained their last degree and where they actually live at the time of the survey, we are able to determine if they have known spatial mobility. Concretely, we create a dummy variable taking the value 1 if the graduate has changed of region between his or her last degree and the region where she or he lives three years later, and 0 otherwise. First of all, we consider all inter-regional migrations, whatever the region of arrival of the migrants. Then, we focus on migrations to Paris region exclusively. Because it is a region that polarises a large part of job opportunities and workers and that concentrates high levels of wages and qualifications (Combes et al., 2015), we believe that it must be distinguished from other regions. We retain an administrative definition of French regions<sup>3</sup>. Its geography is presented in Appendix 3B.

---

<sup>3</sup>We use the definition of the regions at the time of the survey, that is to say, before a French territorial reform implemented in 2015. This definition corresponds to the NUTS 2 geographical divisions from Eurostat.

### 3.4 Descriptive statistics

We provide descriptive statistics about our sample in Table 3.1. First, we observe that there are slightly more women (56%) than men in our sample. About half of the individuals live with a partner, and more than 30% live alone. The rest of the individuals still live with their parents at the time of the survey. A large majority of individuals in the sample does not have children. It is not surprising as it represents young people who have recently completed their studies. We do not consider the number of children because the vast majority of households have only one child (approximately 80% of the total).

Concerning educational levels, 3- and 4-year degrees are under-represented in comparison with both shorter and longer tertiary degrees. This can be explained by the purpose of these degrees: 3- and 4-year degrees are transitory levels, designed to give access to longer studies. On the contrary, 2- and 5-year degrees are supposed to allow direct integration into the labour market. All fields of study are represented in the same proportions, with the exception of agriculture (under-represented) and health (over-represented).

Three years after graduation, 81.7% of the individuals were employed. Services and public sector workers represent a large share of our sample (more than 75%). Concerning the job characteristics, a majority of the individuals are working full-time (more than 85%), but it must be noted that about a third of our sample have temporary contracts. It is a rather high proportion of temporary contracts for workers on the French labour market<sup>4</sup>. However, young people are generally confronted to this kind of jobs due to their lack of experience when they leave the school system. The average wage is lower for workers confronted to overeducation. On the one hand, it can be due to the fact that they do not have a job corresponding to their level of qualification and are therefore paid less than if they did. On the other hand, the lower wage could explain the subjective feeling of being mismatched. If residential migration is not a predominant behaviour, it is not marginal either. Indeed, 38.4% of the workers have moved since graduation, among which about a quarter to Paris. When migrations are observed, they appear to be significant, with an average distance of around 350 km. However, French regions cover vast territories, and moving from one region to another often involves long distances. Despite this

---

<sup>4</sup>According to the French National Institute of Statistics and Economic Studies (Insee), in 2013, 86.5% of workers had a permanent contract.



Table 3.1: Descriptive statistics

|                                   | Whole sample |       | Employed |       | Migrant |       | Stat. overed. |       | Subj. mismatch |        |
|-----------------------------------|--------------|-------|----------|-------|---------|-------|---------------|-------|----------------|--------|
|                                   | Mean         | S.E.  | Mean     | S.E.  | Mean    | S.E.  | Mean          | S.E.  | Mean           | S.E.   |
| Employed                          | 0.817        | 0.003 |          |       | 0.834   | 0.004 |               |       |                |        |
| Migration                         | 0.375        | 0.003 | 0.383    | 0.004 |         |       | 0.343         | 0.007 | 0.361          | 0.008  |
| Migration distance (in km)        |              |       | 352.1    | 2.678 | 353.2   | 2.433 | 335.2         | 5.170 | 349.3          | 5.364  |
| Migration to Paris                | 0.089        | 0.002 | 0.095    | 0.003 | 0.238   | 0.005 | 0.071         | 0.004 | 0.094          | 0.005  |
| Statistically overeducated        |              |       | 0.275    | 0.004 | 0.205   | 0.005 |               |       | 0.483          | 0.009  |
| Subjectively mismatched           |              |       | 0.251    | 0.004 | 0.197   | 0.005 | 0.440         | 0.008 |                |        |
| <i>Individual characteristics</i> |              |       |          |       |         |       |               |       |                |        |
| Woman                             | 0.559        | 0.003 | 0.561    | 0.004 | 0.538   | 0.006 | 0.534         | 0.008 | 0.515          | 0.008  |
| Age                               | 24.29        | 0.024 | 24.28    | 0.025 | 24.38   | 0.036 | 23.49         | 0.044 | 24.35          | 0.053  |
| Having children                   | 0.134        | 0.002 | 0.131    | 0.003 | 0.116   | 0.004 | 0.093         | 0.004 | 0.135          | 0.006  |
| <i>Residential status</i>         |              |       |          |       |         |       |               |       |                |        |
| Living alone                      | 0.338        | 0.003 | 0.342    | 0.004 | 0.418   | 0.006 | 0.326         | 0.008 | 0.323          | 0.008  |
| Living with a partner             | 0.492        | 0.004 | 0.515    | 0.004 | 0.491   | 0.006 | 0.435         | 0.008 | 0.484          | 0.008  |
| Living with their parents         | 0.169        | 0.003 | 0.142    | 0.003 | 0.090   | 0.004 | 0.239         | 0.007 | 0.192          | 0.007  |
| <i>Educational level</i>          |              |       |          |       |         |       |               |       |                |        |
| 2-year degree                     | 0.362        | 0.004 | 0.378    | 0.004 | 0.276   | 0.006 | 0.252         | 0.007 | 0.288          | 0.007  |
| 3- or 4- year degree              | 0.223        | 0.003 | 0.204    | 0.003 | 0.210   | 0.005 | 0.396         | 0.008 | 0.285          | 0.007  |
| 5-year degree or more             | 0.415        | 0.003 | 0.417    | 0.004 | 0.513   | 0.006 | 0.351         | 0.008 | 0.427          | 0.008  |
| <i>Field of study</i>             |              |       |          |       |         |       |               |       |                |        |
| Education                         | 0.004        | 0.001 | 0.004    | 0.001 | 0.003   | 0.001 | 0.003         | 0.001 | 0.004          | 0.001  |
| Humanities                        | 0.122        | 0.002 | 0.106    | 0.003 | 0.105   | 0.004 | 0.141         | 0.006 | 0.156          | 0.006  |
| Economics, business and law       | 0.166        | 0.003 | 0.160    | 0.003 | 0.140   | 0.005 | 0.255         | 0.007 | 0.206          | 0.007  |
| Science                           | 0.159        | 0.002 | 0.160    | 0.003 | 0.197   | 0.005 | 0.119         | 0.005 | 0.158          | 0.006  |
| Industry                          | 0.150        | 0.003 | 0.157    | 0.003 | 0.176   | 0.005 | 0.200         | 0.007 | 0.172          | 0.006  |
| Agriculture                       | 0.024        | 0.001 | 0.022    | 0.001 | 0.030   | 0.002 | 0.045         | 0.003 | 0.025          | 0.003  |
| Health                            | 0.223        | 0.003 | 0.246    | 0.003 | 0.200   | 0.005 | 0.034         | 0.003 | 0.099          | 0.005  |
| Services                          | 0.149        | 0.003 | 0.142    | 0.003 | 0.147   | 0.005 | 0.198         | 0.007 | 0.177          | 0.007  |
| <i>Business sector</i>            |              |       |          |       |         |       |               |       |                |        |
| Public sector                     |              |       | 0.339    | 0.004 |         |       | 0.170         | 0.006 | 0.281          | 0.007  |
| Agriculture                       |              |       | 0.007    | 0.001 |         |       | 0.019         | 0.002 | 0.006          | 0.001  |
| Sales                             |              |       | 0.099    | 0.003 |         |       | 0.212         | 0.007 | 0.171          | 0.007  |
| Construction                      |              |       | 0.024    | 0.001 |         |       | 0.024         | 0.003 | 0.022          | 0.003  |
| Industry                          |              |       | 0.095    | 0.003 |         |       | 0.136         | 0.006 | 0.117          | 0.006  |
| Services                          |              |       | 0.434    | 0.004 |         |       | 0.427         | 0.008 | 0.401          | 0.008  |
| <i>Job characteristics</i>        |              |       |          |       |         |       |               |       |                |        |
| Wage                              |              |       | 1835.8   | 6.034 |         |       | 1546.5        | 9.631 | 1641.5         | 10.958 |
| Working full-time                 |              |       | 0.863    | 0.003 |         |       | 0.834         | 0.006 | 0.843          | 0.006  |
| Permanent job                     |              |       | 0.695    | 0.004 |         |       | 0.658         | 0.007 | 0.679          | 0.008  |
| Time to get first job (in months) |              |       | 2.3      | 0.040 |         |       | 3.2           | 0.091 | 3.1            | 0.093  |
| <i>N</i>                          | 15 865       |       | 12 968   |       | 5 961   |       | 3 570         |       | 3 254          |        |

Lecture: Migration distance is measured in kilometers between the centroids of the departure and arrival regions (equal to 0 if there is no interregional migration). The average value reported here is calculated only for people who have migrated in each subsample. Average wage is monthly.

Source: Authors' calculation based on *Génération 2010* data.

observation, it is worth noting that the average distance observed mainly reflects movements between neighbouring regions rather than between distant regions.

There are slightly more individuals who are statistically overeducated (27.5%) than those subjectively mismatched (25.1%). However, when we cross the two situations (Table 3.2), we observe that approximately 30% of the workers are mismatched in only one of the dimensions (around 13% are only subjectively mismatched, and more than 15% are only statistically overeducated). This confirms that these two situations do not systematically overlap, either due to the difference between overeducation and overskilling or because there is a distinction between the feeling of being overeducated and the fact of actually being mismatched. The situation is

Table 3.2: Correspondence between statistical overeducation and subjective mismatch

|  |              | Subjective mismatch |                  |
|--|--------------|---------------------|------------------|
|  |              | Matched             | Overeducated     |
| <b><i>Employed</i></b> ( $N = 12968$ ) |              |                     |                  |
| <b>Statistical overeducation</b>       | Matched      | 7 715<br>(59.5%)    | 1 683<br>(13.0%) |
|  | Overeducated | 1 999<br>(15.4%)    | 1 571<br>(12.1%) |
| <b><i>Migrant</i></b> ( $N = 5961$ )   |              |                     |                  |
| <b>Statistical overeducation</b>       | Matched      | 4080<br>(68.4%)     | 656<br>(11.0%)   |
|  | Overeducated | 706<br>(11.8%)      | 519<br>(8.7%)    |

Source: Author's calculation based on *Génération 2010* data.

similar among individuals who have migrated, although the share of mismatched workers (either in one or both dimensions) is lower for this sub-population. While it seems that there may be a relationship between migration and overeducation, an econometric analysis is necessary to establish whether or not there is a causal relationship.

### 3.5 Estimation strategy

This work examines the links between migration (or spatial mobility) and overeducation for French young workers. However, the existence of a causal effect must be assessed in a regression framework controlling for possible sample selection bias and endogeneity of migration decisions.

Firstly, the estimated results may suffer from a sample selection problem since educational mismatch is observable only for graduates who are actually employed. Analysing the risk of overeducation while restricting the sample to young workers could lead to biased results. This bias may occur if the probability of being mismatched differs between individuals who decide to work and those who decide not to. For instance, some young people may choose inactivity or unemployment to avoid mismatch, and thus, those least likely to accept a job could be the most likely to be overeducated. On the other hand, some young people may prefer to be mismatched rather than be unemployed. In this case, the aversion to the risk of unemployment is likely to increase the likelihood of accepting a mismatched job. Whatever the strategy retained, it indicates that some unobserved characteristics influencing the probability of being employed could also be related to the risk of mismatch.

We deal with this selection bias by using the standard two-step Heckman correction procedure (Heckman, 1979). Technically, the first step involves specifying an employment selection equation, which uses a probit model to explain the probability of being employed based on individuals' socio-economic characteristics. The equation is written as follows:

$$E = \beta_0 + \beta_1 X + \beta_2 Z + \epsilon \quad (3.1)$$

$$\text{Where: } E = \begin{cases} 1 & \text{if } E^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

In Equation 3.1,  $E$  represents the employment status (being employed or not), and  $X$  is a vector of various socio-economic characteristics that are likely to influence the probability of employment for young people. We consider age, gender and family status of the young people, along with variables for their educational level and field of study. Additionally, we include dummies for regions of residence to account for any unobserved heterogeneity between territories or local labour markets<sup>5</sup>. Identifying the risk of overeducation with the presence of a sample selection bias requires at least one selection variable ( $Z$ ) that is correlated with access to employment but unrelated to this risk. Finding such variables, also known as exclusion restrictions, can be challenging when working with observational data.

We retain health status as an exclusion restriction for the employment equation. In particular, we know if an individual is affected or not by a long-term health problem or disability. In fact, 7.4% of our sample is concerned, although we cannot identify whether it is a disability or not. We assume that experiencing a long-lasting health problem is likely to affect the likelihood of finding a job but should not have a direct influence on overeducation once we control for demographic characteristics, job-specific features and local variables. A large part of the literature in health economics has explored and demonstrated the links between health status and employment (see e.g. Barnay, 2020; Vornholt et al., 2018, for recent reviews). For instance, several studies have shown that individuals with disabilities are often underemployed in the workforce. Regarding mental health as a health status, the relationship may be bidirectional (Banerjee et al., 2017). In other words, health status can influence employment positions, and conversely, being employed (or not) can impact health status. However, having a disability restricts the bidirec-

---

<sup>5</sup>*i.e.* the French administrative regions. See Appendix 3B for more details.

tional causal effect, as it can be attributed to accidents, genetic conditions or even congenital malformations (rather than employment status). Additionally, because individuals in our sample declare being affected by a long-term health problem, we can reasonably assume that this health status precedes the time of the survey. This ensures that health status indeed determines the probability of employment. On the other hand, to the best of our knowledge, there is no empirical evidence of a direct link between health status and overeducation. While being mismatched could potentially lead to a deterioration in mental health, using only long-term affections should mitigate the relevance of the phenomenon. In addition, more than 80% of the respondents suffering from a health condition report that it preceded their entry on the labour market.

Once we have taken into account the sample selection problem, we may consider the effect of migration on the risk of overeducation by estimating the following linear probability model<sup>6</sup>:

$$O = \beta_0 + \beta_1 X + \beta_2 \text{Mobility} + \epsilon \quad (3.2)$$

$$\text{Where: } O = \begin{cases} 1 & \text{if } O^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

In Equation 3.2,  $O$  represents the variable of being overeducated or not. It is important to note that overeducation is measured in two ways, using statistical and subjective measures. In both cases, the vector of observable characteristics  $X$  does not vary and includes the same set of variables as defined for Equation 3.1, plus some job-related characteristics such as the time needed to find the first job, the type of contract (permanent or not, full-time or not) and the activity sector (public, agricultural, industrial, construction, services etc.). Additionally, we consider two types of spatial mobility. In one case, mobility corresponds to migration from one region to another between graduation and the current job occupied. In the other case, it corresponds to migration to Paris exclusively.

A further econometric issue could arise from the endogeneity of the migration choice. It is likely that there is a potential correlation between overeducation and

---

<sup>6</sup>Some issues can arise from the use of linear probability models. Especially, they can predict probabilities outside the  $[0; 1]$  interval. Appendix 3C shows the distribution of predicted overeducation in our sample. In the case of our statistical measure of overeducation, only a small number of predicted probabilities are lower than 0, and none of them are higher than 1. In the case of our subjective measure, predicted probabilities are almost all between 0 and 1. Still, to check the robustness of our results, we estimate a triprobit model, which provides results similar to those obtained with linear probability models (see Section 3.6.3.).

some unobserved factors that influence the decision to migrate or relocate to get a job. For example, risk-averse individuals might accept jobs that do not perfectly match their skills to avoid the costs and risks associated with migration. This implies that, in Equation 3.2, the spatial mobility decision is likely to be linked to unobservable individual characteristics that are correlated with the error term, resulting in biased estimates. To address this problem, we retain the instrumental variables method.

Previous studies addressing this endogeneity problem have used the lagged unemployment rate in the area from which the individual moved (Croce & Ghignoni, 2015). However, this approach has been criticized due to the questionable assumption that the lagged unemployment rate in a particular area has no influence on current overeducation in the same area, given the high persistence of local unemployment. As an alternative, some authors have used housing tenure as an instrument (Devillanova, 2013; Meliciani & Radicchia, 2016). The argument is that housing tenure should be correlated with migration since renters can move more easily than homeowners, but it is not directly correlated with the degree of overeducation. Since we lack information on housing tenure status, we follow the approach used by Venhorst and Cörvers (2018) and employ information on past mobility as instrumental variables ( $Z$ ).

First, we define a dummy variable that takes the value 1 if young people have migrated between the time they obtained their baccalaureate and their last degree (*i.e.* if they changed their region of residence during or for university studies), and 0 otherwise. This instrument, by capturing information on spatial mobility during education, is likely to explain recent mobility (DaVanzo & Morrison, 1981) without being directly connected to the actual risk of overeducation. However, one can argue that it is not totally convincing as past mobility might be correlated to unobserved characteristics ( $u$ ) determining simultaneously recent mobility and risk of overeducation. It might be a certain ability and/or motivation of some individuals to succeed in their professional careers. For this reason, we retain a second instrumental variable providing information on whether or not young people have spent time abroad during their academic careers. We only consider mandatory stays abroad because they do not result from a direct voluntary choice of individuals concerned and are less likely to be correlated with omitted variables mentioned above. We then have:  $Cov(Z, u) = 0$  and  $Cov(Z, Mobility) \neq 0$ . Our variable thus satisfies the two conditions of a good instrument (Wooldridge, 2002).

Equation 3.2 is estimated using the two-stage least squares (2SLS) procedure. We employ a linear probability model to explain the migration decision, incorporating the same set of variables as before, plus our instrumental variables. Appendix 3D also presents the results for the first stage of the 2SLS, confirming that past migration and mandatory stays abroad are significantly and positively associated with recent migrations. Additionally, we implement the weak instrument test of Olea and Pflueger (2013) to ensure the validity of our instruments. Our F-test values, which exceed the critical value reported by Stock and Yogo (2005), indicate that we have strong instruments.

Finally, we estimate the probability of being overeducated using three different econometric strategies. Firstly, it is estimated without any control for the selection into employment or the endogeneity of migration decision. Secondly, we control only for the selection bias. Thirdly, we control for both biases. In addition, we first estimate baseline models where individuals' educational level is distinguished through dummy variables to check whether they face a different risk of overeducation. Then, we proceed by performing separate estimations depending on the educational level of the graduates.

## 3.6 Results

### 3.6.1 Effects of migration on overeducation

The results of our models are presented in Table 3.3 (statistical overeducation) and Table 3.4 (subjective mismatch). The first column shows the baseline model without any correction for selection and endogeneity biases. The second and third columns present the results with correction for selection bias (column 2) and both selection and endogeneity biases (column 3), respectively. We will focus our discussion on the last column as it represents the most complete and robust model. On the whole, we find different effects for several variables depending on the type of mismatch we consider.

First, the coefficients and significance for the inverse Mills ratio demonstrate the necessity to take into account the selection into employment in our analysis of the determinants of overeducation. In line with Devillanova (2013), it confirms the intuition that the probability of being overeducated is positively correlated with

the probability of being unemployed. Overall, Tables 3.3 and 3.4 also reveal that self-selection has no influence on the coefficients of interest.

Concerning individual characteristics, we find that women are statistically more likely to be overeducated than men. This could be related to some difficulties in finding a “good” job or potential discrimination experienced during the job search process. Consequently, when women are hired, they might be assigned to less skilled positions. However, there is no significant gender difference for subjective mismatch, indicating that men and women have similar perceptions of the mismatch between their skills and job requirements. The age at the time of graduation has a negative effect on statistical overeducation, but a small positive effect on subjective mismatch. In one case, this can be attributed to the fact that with age, individuals gain experience and skills that align better with job requirements, reducing the risk of overeducation. In the other case, the positive effect on subjective mismatch may suggest that older individuals are more critical of their jobs. Having children has no robust effect on overeducation, whatever the measure retained. However, being a woman with children reveals some negative effects depending on the estimation strategy retained. Although counterintuitive, this finding is in accordance with those obtained by Jauhiainen (2011). Regarding living with parents, we find a positive effect on statistical overeducation, although not always significant. Following Chevalier (2003), a significant effect might reflect lower abilities or a lower motivation that affect both the decision to live with parents and the labour market positions. Surprisingly, we do not observe any significant effect for young people living with a partner on statistical overeducation, but a positive and significant one on subjective mismatch. In this case, it could be due to the fact that individuals accept less favourable employment conditions when it comes to following their spouse.

We observe that the educational level has a strong effect on statistical overeducation, but not necessarily on subjective mismatch. Compared to young people with a 2-year degree, those with more years of education are more likely to be overeducated. This is quite normal as it is less common to have a job below his education level if the latter is not very high. However, the risk of overeducation is highest for 3 or 4-year degree graduates. This might be due to the fact that these degrees represent transitional levels designed to give access to further studies. There is limited demand for these educational levels in the French labour market. This is especially

Table 3.3: Linear probability models for statistical overeducation

|   | OLS                  | Heckman              | Heckman + IV         |
|---|----------------------|----------------------|----------------------|
| Constant                                      | 0.730***<br>(0.043)  | 0.669***<br>(0.055)  | 0.711***<br>(0.125)  |
| Migration                                     | -0.038***<br>(0.009) | -0.038***<br>(0.009) | -0.072***<br>(0.017) |
| <i>Demographic characteristics</i>            |                      |                      |                      |
| Woman   | 0.062***<br>(0.007)  | 0.062***<br>(0.007)  | 0.062***<br>(0.008)  |
| Age   | -0.016***<br>(0.002) | -0.016***<br>(0.002) | -0.017***<br>(0.002) |
| Having children                               | 0.029<br>(0.021)     | 0.033<br>(0.021)     | 0.031*<br>(0.018)    |
| Woman × Children                              | -0.019<br>(0.025)    | -0.055**<br>(0.022)  | -0.055*<br>(0.031)   |
| <i>Residential status (Ref: Living alone)</i> |                      |                      |                      |
| Living with a partner                         | -0.001<br>(0.007)    | 0.012<br>(0.011)     | 0.010<br>(0.011)     |
| Living with their parents                     | 0.108***<br>(0.013)  | 0.059*<br>(0.031)    | 0.052<br>(0.033)     |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                      |                      |
| 3- or 4- year degree                          | 0.248***<br>(0.020)  | 0.228***<br>(0.023)  | 0.233***<br>(0.018)  |
| 5-year degree or more                         | 0.063***<br>(0.018)  | 0.054***<br>(0.016)  | 0.063***<br>(0.013)  |
| <i>Field of study (Ref: Services)</i>         |                      |                      |                      |
| Education                                     | -0.099<br>(0.068)    | -0.057<br>(0.064)    | -0.056<br>(0.059)    |
| Humanities                                    | -0.013<br>(0.024)    | -0.032<br>(0.028)    | -0.034*<br>(0.019)   |
| Economics, business and law                   | 0.058***<br>(0.014)  | 0.060***<br>(0.014)  | 0.059***<br>(0.013)  |
| Science                                       | -0.091***<br>(0.018) | -0.081***<br>(0.016) | -0.080***<br>(0.014) |
| Industry                                      | -0.024<br>(0.016)    | 0.002<br>(0.016)     | 0.004<br>(0.020)     |
| Agriculture                                   | 0.106***<br>(0.030)  | 0.094***<br>(0.032)  | 0.098***<br>(0.027)  |
| Health  | -0.215***<br>(0.020) | -0.182***<br>(0.020) | -0.180***<br>(0.025) |
| <i>Business sector (Ref: Services)</i>        |                      |                      |                      |
| Public sector                                 | -0.078***<br>(0.010) | -0.078***<br>(0.010) | -0.077***<br>(0.008) |
| Agriculture                                   | 0.312***<br>(0.052)  | 0.311***<br>(0.052)  | 0.312***<br>(0.042)  |
| Sales   | 0.220***<br>(0.012)  | 0.221***<br>(0.013)  | 0.220***<br>(0.012)  |
| Construction                                  | 0.066*<br>(0.038)    | 0.067*<br>(0.038)    | 0.066***<br>(0.023)  |
| Industry                                      | 0.074***<br>(0.019)  | 0.075***<br>(0.019)  | 0.075***<br>(0.013)  |
| <i>Job characteristics</i>                    |                      |                      |                      |
| Working full-time                             | -0.050***<br>(0.012) | -0.050***<br>(0.012) | -0.049***<br>(0.011) |
| Permanent job                                 | -0.050***<br>(0.009) | -0.049***<br>(0.009) | -0.050***<br>(0.008) |
| Duration between graduation and first job     | 0.007***<br>(0.001)  | 0.006***<br>(0.001)  | 0.006***<br>(0.001)  |
| Inverse Mills ratio                           |                      | 0.232*<br>(0.121)    | 0.230*<br>(0.140)    |
| Region fixed effects                          | Yes                  | Yes                  | Yes                  |
| R-squared                                     | 0.251                | 0.251                | 0.250                |
| N   | 12 968               | 12 962               | 12 962               |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.

true for 4-year degrees, which may not be sufficient for jobs requiring a Master's degree, making them more susceptible to being preferred for positions that require



a lower level of education, such as those asking for a Bachelor's degree. Moreover, we find a negative effect of graduating in health-related fields on both dimensions of mismatch. While the field of study does not seem to affect subjective mismatch, our results show that graduates in Economics, Business and Law, and Agriculture face a higher risk of statistical overeducation, whereas graduates in Science are significantly less likely to be overeducated.

When we examine the job characteristics, we observe that the public sector seems to be the only one protected from overeducation. This could be attributed to the fact that most public service positions are allocated through formal entrance examinations, which reduces the likelihood of being hired for a job that does not match one's education. However, individuals in the public sector do not feel subjectively less mismatched. On the other hand, the Agriculture and Construction sectors have the highest risk of overeducation. These sectors often involve manual work, and such jobs typically do not require high levels of qualifications from workers.

The estimated coefficient for the variable "working full-time" indicates that this situation is associated with a lower risk of statistical overeducation. One possible explanation is that young workers are more selective when choosing a full-time job and are less likely to accept being overeducated for such positions. On the other hand, the negative sign associated with permanent contracts in Table 3.3 supports this view and is in line with the findings of Baert and Verhaest (2019), who argued that overeducation might be deemed acceptable only for temporary positions. We also observe that the time to get the first job is positively correlated with the risk of overeducation. Young people who have spent the most time searching for a job may decide to take a lower-quality job rather than continue to do so, especially to avoid the more detrimental effects of unemployment on future labour market positions (Baert & Verhaest, 2019). Although this result contradicts the predictions of the job search models, it is confirmed in the different models and for the different measures of over-education. The positive sign might therefore illustrate an effect of professional experience in reducing overeducation (Rubb, 2003). If we consider that overeducation at the beginning of the career may be related to a lack of skills (Chevalier, 2003), then taking a job earlier after graduation allows the individuals to accumulate more experience and potential on-the-job training to compensate.

Our main focus is on the effect of migration (or spatial mobility) on overeducation. The results are presented in Tables 3.3 and 3.4. Overall, we find a significant

Table 3.4: Linear probability models for subjective mismatch

|   | OLS                  | Heckman              | Heckman + IV         |
|---|----------------------|----------------------|----------------------|
| Constant                                      | 0.175***<br>(0.058)  | 0.053<br>(0.067)     | 0.092<br>(0.136)     |
| Migration                                     | -0.015<br>(0.010)    | -0.015<br>(0.010)    | -0.046***<br>(0.019) |
| <i>Demographic characteristics</i>            |                      |                      |                      |
| Woman   | 0.010<br>(0.008)     | 0.011<br>(0.008)     | 0.011<br>(0.009)     |
| Age   | 0.005***<br>(0.001)  | 0.004***<br>(0.001)  | 0.003**<br>(0.002)   |
| Having children                               | 0.024<br>(0.017)     | 0.038<br>(0.018)     | 0.027<br>(0.019)     |
| Woman × Children                              | -0.001<br>(0.023)    | -0.073*<br>(0.037)   | -0.073**<br>(0.033)  |
| <i>Residential status (Ref: Living alone)</i> |                      |                      |                      |
| Living with a partner                         | 0.009<br>(0.008)     | 0.034***<br>(0.012)  | 0.032***<br>(0.012)  |
| Living with their parents                     | 0.070***<br>(0.019)  | -0.027<br>(0.038)    | -0.033<br>(0.035)    |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                      |                      |
| 3- or 4- year degree                          | 0.036**<br>(0.014)   | -0.005<br>(0.016)    | -0.000<br>(0.019)    |
| 5-year degree or more                         | -0.023<br>(0.014)    | -0.040***<br>(0.013) | -0.032***<br>(0.014) |
| <i>Field of study (Ref: Services)</i>         |                      |                      |                      |
| Education                                     | -0.066<br>(0.039)    | 0.017<br>(0.040)     | 0.018<br>(0.063)     |
| Humanities                                    | 0.047*<br>(0.024)    | 0.009<br>(0.031)     | 0.007<br>(0.014)     |
| Economics, business and law                   | 0.004<br>(0.014)     | 0.008<br>(0.014)     | 0.007<br>(0.014)     |
| Science                                       | -0.039**<br>(0.015)  | -0.021<br>(0.015)    | -0.020<br>(0.015)    |
| Industry                                      | -0.025<br>(0.024)    | 0.025<br>(0.027)     | 0.027<br>(0.022)     |
| Agriculture                                   | -0.032<br>(0.033)    | -0.054<br>(0.037)    | -0.050*<br>(0.029)   |
| Health  | -0.180***<br>(0.015) | -0.116***<br>(0.024) | -0.114***<br>(0.027) |
| <i>Business sector (Ref: Services)</i>        |                      |                      |                      |
| Public sector                                 | 0.003<br>(0.009)     | 0.003<br>(0.009)     | 0.004<br>(0.009)     |
| Agriculture                                   | -0.042<br>(0.041)    | -0.044<br>(0.041)    | -0.042<br>(0.045)    |
| Sales   | 0.159***<br>(0.013)  | 0.159***<br>(0.014)  | 0.158***<br>(0.013)  |
| Construction                                  | -0.021<br>(0.035)    | -0.019<br>(0.035)    | -0.020<br>(0.025)    |
| Industry                                      | 0.059**<br>(0.024)   | 0.060**<br>(0.024)   | 0.060**<br>(0.014)   |
| <i>Job characteristics</i>                    |                      |                      |                      |
| Working full-time                             | -0.046***<br>(0.013) | -0.046***<br>(0.013) | -0.045***<br>(0.012) |
| Permanent job                                 | -0.004<br>(0.012)    | -0.003<br>(0.012)    | -0.004<br>(0.009)    |
| Duration between graduation and first job     | 0.007***<br>(0.001)  | 0.007***<br>(0.001)  | 0.007***<br>(0.001)  |
| Inverse Mills ratio                           |                      | 0.456<br>(0.136)     | 0.455<br>(0.156)     |
| Region fixed effects                          | Yes                  | Yes                  | Yes                  |
| R-squared                                     | 0.074                | 0.074                | 0.073                |
| N   | 12 968               | 12 962               | 12 962               |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.

and negative effect of migration on both dimensions of mismatch, although the effect is stronger on statistical overeducation than on subjective mismatch. This suggests

that increasing the job search radius and looking for opportunities on a national (rather than local) level can be a good strategy for young workers to avoid overeducation. Interestingly, this finding partially contradicts the work of Devillanova (2013), who does not find any significant effect of migration on the risk of overeducation for Italian workers when controlling for selection and endogeneity biases. However, the author finds that commuting improves the match between education and job. Similarly, Meliciani and Radicchia (2016) observe no effect of migration on overeducation, except for specific geographical areas in Italy. On the other hand, our findings are in line with the studies conducted by Jauhainen (2011) on Finnish data and Ramos and Sanromá (2013) on Spanish data, which both observe that moving to another region reduces the risk of overeducation for workers. Also on Spanish data, Romaní et al. (2016) confirm the link between spatial mobility and overeducation only for some specific groups as the highly educated, while Croce and Ghignoni (2015) show a negative impact of commuting on the risk of educational mismatch for Italian upper-secondary graduates, as well as a negative impact of migration for university graduates. These results support the need for an analysis that differentiates between education levels, as is done in Table 3.6 below.

### 3.6.2 Differentiation by migration and educational level

Tables 3.5 and 3.6 present the effects of our four variables of interest, differentiating between migration to all regions and migration to Paris. Table 3.5 provides results for the whole population, while Table 3.6 provides results for different educational levels<sup>7</sup>.

Results of our various estimated models reveal that the coefficients for migration to Paris are approximately four to six times higher than for migration to all regions. This finding suggests that the Paris region is particularly advantageous for young people seeking jobs that match their skill levels. The higher job density and various job opportunities in Paris likely enable individuals to search more effectively for well-suited positions. Additionally, previous studies like Combes et al. (2012) have demonstrated that wages in the Paris region are 24% higher than in the rest of France, and a significant part of this wage premium is attributed to the higher skills of the Parisian workforce. The successful matching of skills demanded and

---

<sup>7</sup>For sake of clarity and because the effects of the other variables are similar to those obtained in Tables 3.3 and 3.4, we only report the coefficients for migration in the table. Full tables are available upon request.

Table 3.5: Linear probability models - Effects of migration

| (a) Statistical overeducation |                |           |            |                     |           |            |
|-------------------------------|----------------|-----------|------------|---------------------|-----------|------------|
|                               | All migrations |           |            | Migrations to Paris |           |            |
|                               | OLS            | Heck.     | Heck. + IV | OLS                 | Heck.     | Heck. + IV |
| <b>Coefficient</b>            | -0.038***      | -0.038*** | -0.072***  | -0.023***           | -0.023*** | -0.572***  |
| <b>Standard error</b>         | (0.009)        | (0.009)   | (0.017)    | (0.006)             | (0.006)   | (0.099)    |
| <b>Ind. characteristics</b>   | Yes            | Yes       | Yes        | Yes                 | Yes       | Yes        |
| <b>Job characteristics</b>    | Yes            | Yes       | Yes        | Yes                 | Yes       | Yes        |
| <b>Region fixed effects</b>   | Yes            | Yes       | Yes        | Yes                 | Yes       | Yes        |
| <b>R-squared</b>              | 0.251          | 0.251     | 0.250      | 0.250               | 0.250     | 0.161      |
| <i>N</i>                      | 12 968         | 12 962    | 12 962     | 12 968              | 12 962    | 12 962     |

| (b) Subjective mismatch     |                |         |            |                     |         |            |
|-----------------------------|----------------|---------|------------|---------------------|---------|------------|
|                             | All migrations |         |            | Migrations to Paris |         |            |
|                             | OLS            | Heck.   | Heck. + IV | OLS                 | Heck.   | Heck. + IV |
| <b>Coefficient</b>          | -0.015         | -0.015  | -0.046**   | 0.007               | 0.006   | -0.247**   |
| <b>Standard error</b>       | (0.010)        | (0.010) | (0.019)    | (0.004)             | (0.004) | (0.102)    |
| <b>Ind. characteristics</b> | Yes            | Yes     | Yes        | Yes                 | Yes     | Yes        |
| <b>Job characteristics</b>  | Yes            | Yes     | Yes        | Yes                 | Yes     | Yes        |
| <b>Region fixed effects</b> | Yes            | Yes     | Yes        | Yes                 | Yes     | Yes        |
| <b>R-squared</b>            | 0.074          | 0.074   | 0.073      | 0.074               | 0.074   | 0.054      |
| <i>N</i>                    | 12 968         | 12 962  | 12 962     | 12 968              | 12 962  | 12 962     |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.

offered by employers in Paris may contribute to these higher wages. Furthermore, considering that highly qualified individuals are less likely to be overeducated, and that Paris attracts and concentrates a large number of highly skilled individuals, it is quite logical to observe such a strong effect of migration to Paris. Other studies, such as Büchel and van Ham (2003), Jauhainen (2011), and Ramos and Sanromá (2013), also have similar conclusions, showing that residing in larger labour markets is beneficial as it reduces the probability of overeducation. More generally, previous research has highlighted that the effects of migration may vary depending on the destination. For instance, studies like Meliciani and Radicchia (2016) or Iammarino and Marinelli (2015) have found contrasting effects for Italian workers migrating to the north or south of the country.

Finally, Table 3.6 presents the effects of spatial mobility on overeducation, considering educational level and migration destination. We show that the negative effect observed for the whole sample varies significantly depending on the educational level. Specifically, individuals with at least a 5-year degree are less likely to be statistically overeducated when they migrate. However, we do not find any significant effect of residential migration on subjective mismatch. This may indicate that spatial mobility allows workers to find better-matched jobs, but the costs associated with migration lead them to underestimate its benefits, resulting in self-assessed mismatch. Venhorst and Cörvers (2018) also find that the effects of spatial mobility

Table 3.6: Linear probability models - Effects of migration by educational level

| (a) Statistical overeducation |                |                   |                   |                     |                   |                   |
|-------------------------------|----------------|-------------------|-------------------|---------------------|-------------------|-------------------|
| Degree                        | All migrations |                   |                   | Migrations to Paris |                   |                   |
|                               | 2-year         | 3-year<br>or more | 5-year<br>or more | 2-year              | 3-year<br>or more | 5-year<br>or more |
|                               | OLS            | Heck.             | Heck. + IV        | OLS                 | Heck.             | Heck. + IV        |
| <b>Coefficient</b>            | -0.023         | -0.102***         | -0.102***         | -0.432              | -0.601***         | -0.461***         |
| <b>Standard error</b>         | (0.021)        | (0.026)           | (0.032)           | (0.301)             | (0.117)           | (0.121)           |
| <b>Ind. characteristics</b>   | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>Job characteristics</b>    | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>Region fixed effects</b>   | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>R-squared</b>              | 0.330          | 0.211             | 0.162             | 0.297               | 0.099             | 0.066             |
| <i>N</i>                      | 4 903          | 8 059             | 5 410             | 4 903               | 8 059             | 5 410             |

| (b) Subjective mismatch     |                |                   |                   |                     |                   |                   |
|-----------------------------|----------------|-------------------|-------------------|---------------------|-------------------|-------------------|
| Degree                      | All migrations |                   |                   | Migrations to Paris |                   |                   |
|                             | 2-year         | 3-year<br>or more | 5-year<br>or more | 2-year              | 3-year<br>or more | 5-year<br>or more |
|                             | OLS            | Heck.             | Heck. + IV        | OLS                 | Heck.             | Heck. + IV        |
| <b>Coefficient</b>          | -0.007         | -0.062**          | -0.049            | -0.127              | -0.248**          | -0.228*           |
| <b>Standard error</b>       | (0.024)        | (0.028)           | (0.035)           | (0.340)             | (0.117)           | (0.129)           |
| <b>Ind. characteristics</b> | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>Job characteristics</b>  | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>Region fixed effects</b> | Yes            | Yes               | Yes               | Yes                 | Yes               | Yes               |
| <b>R-squared</b>            | 0.109          | 0.048             | 0.046             | 0.107               | 0.023             | 0.021             |
| <i>N</i>                    | 4 903          | 8 059             | 5 410             | 4 903               | 8 059             | 5 410             |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region. Estimated coefficients are obtained with the model controlling for selection and endogeneity biases (Heckman procedure + Instrumental variables).

Source: Authors' calculations based on *Génération 2010* data.

on subjective evaluations may differ depending on the educational level. For college graduates, they observe negative effects of migration on subjective job-match measures. Conversely, they find a positive effect of migration on subjective evaluations of job-match quality for university graduates. While our analysis confirms some differentiated effects, it contradicts their conclusion as we do not find any significant effects of migration for the most educated or the least educated individuals. We only report some negative effects on the risk of subjective mismatch for 3-year university graduates. These differences in results may be attributed to the contrasting economic and institutional contexts in the Netherlands and France.

In addition, our findings suggest that 2-year university graduates do not seem to benefit from migration in any dimension of mismatch. On the contrary, we observe significant negative effects of spatial mobility (both national migration and migration to Paris) for individuals with intermediate tertiary education (3- and 4-year degrees). This result is consistent with Croce and Ghignoni (2015), who find that moving (by relocation) decreases the risk of a wrong match for more educated graduates but has little or no effect on the least educated ones. Only commuting time appears to be helpful for this latter category. Similarly, in the Spanish case, Romání et al. (2016) evidence that migration has no effect on overeducation, except

for male workers with a university degree. On Dutch data, Venhorst and Cörvers (2018) find a positive effect of internal migration on the likelihood of obtaining a horizontal match for both college and university graduates. However, the authors considered movers on both short and long distances. In fact, the mean distance moved was between 20 km and 30 km, according to the population studied. It is, therefore, difficult to compare these mobilities with inter-regional migrations that take place over several hundred kilometres (as is the case for France, Italy or Spain).

Overall, we find that the results by educational level are qualitatively similar, whether considering only migrations to Paris or all inter-regional migrations: spatial mobility mainly benefits more educated workers. However, interestingly enough, we find that subjective mismatch is reduced only by migrations to Paris for the 5-year university graduates (although the effect is significant only at the 10% level), but not by other residential migrations. This might be the sign that the most educated value some advantages specifically associated with living in the capital city. Regarding statistical overeducation, on the contrary, we find similar coefficients for the intermediate and highest tertiary education graduates for general migration, while the effect is relatively lower for the most educated looking at migration to Paris (but still much higher than the effect of any inter-regional migration). Because overeducation can be related to a lack of job opportunities, it may seem logical that this problem mainly concerns the highly educated. Indeed, they may need to be close to large metropolitan areas (like Paris) in order to find suitable jobs. On the contrary, jobs corresponding to shorter university education may be less concentrated in specific regions, reducing the importance of being mobile to find a matched job. For the less educated, increasing daily commutes might then be sufficient to avoid being overeducated.

### 3.6.3 Robustness checks

#### Model specification

In our work, migrations are defined by a dummy variable indicating whether or not young people have changed of region at the end of their studies. The main limitation of our approach is that moving from a region to a neighbouring region and moving to a region far away are considered similar. In some cases, it is a move of a hundred kilometres or so, while in others, it could be much greater. Then, we propose to compute the migration distance in kilometres when a change of region

is observed<sup>8</sup>. Because we do not have detailed information for the initial place of residence and the current place at the time of the survey, we cannot use more precise distances. Results of estimated models with this alternative measure for migrations are summarized in Table 3.7.

Considering distances, rather than the simple fact of moving or not, yields slightly different results. We still highlight a negative effect of migration on the risk of overeducation: The greater the distance, the lower the risk. Moving to a distant place would only be considered for a job that corresponds to the young people's skill level. It is a clear observation when we focus on the statistical measure. In addition, we confirm the fact that the most educated are the first to benefit from such mobility. In line with our previous models, estimations based on the subjective measure of overeducation reveal some differences depending on the educational level. We find that migrations are helpful for young people with at least a 3-year university degree. However, if the effect for the 5-year university graduates is now significant, it is only at the 10% level. Therefore, it is important to consider this result with caution<sup>9</sup>.

Table 3.7: Linear probability models - Effects of migration based on (log) distance (measured in km)

| (a) Statistical overeducation |           |           |                   |                   |
|-------------------------------|-----------|-----------|-------------------|-------------------|
| Degree                        | All       | 2-year    | 3-year<br>or more | 5-year<br>or more |
| <b>Coefficient</b>            | -0.013*** | -0.011*** | -0.013***         | -0.017***         |
| <b>Standard error</b>         | (0.003)   | (0.004)   | (0.004)           | (0.005)           |
| <b>Ind. characteristics</b>   | Yes       | Yes       | Yes               | Yes               |
| <b>Job characteristics</b>    | Yes       | Yes       | Yes               | Yes               |
| <b>Region fixed effects</b>   | Yes       | Yes       | Yes               | Yes               |
| <b>R-squared</b>              | 0.250     | 0.330     | 0.213             | 0.162             |
| <i>N</i>                      | 12 962    | 4 892     | 8 059             | 5 410             |

| (b) Subjective mismatch     |          |         |                   |                   |
|-----------------------------|----------|---------|-------------------|-------------------|
| Degree                      | All      | 2-year  | 3-year<br>or more | 5-year<br>or more |
| <b>Coefficient</b>          | -0.007** | 0.003   | -0.011**          | -0.010*           |
| <b>Standard error</b>       | (0.003)  | (0.004) | (0.005)           | (0.006)           |
| <b>Ind. characteristics</b> | Yes      | Yes     | Yes               | Yes               |
| <b>Job characteristics</b>  | Yes      | Yes     | Yes               | Yes               |
| <b>Region fixed effects</b> | Yes      | Yes     | Yes               | Yes               |
| <b>R-squared</b>            | 0.073    | 0.109   | 0.047             | 0.045             |
| <i>N</i>                    | 12 962   | 4 892   | 8 059             | 5 410             |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region. Estimated coefficients are obtained with the model controlling for selection and endogeneity biases (Heckman procedure + Instrumental variables).

Source: Authors' calculations based on *Génération 2010* data.

<sup>8</sup>The distance is an Euclidian distance between the centroids of the region.

<sup>9</sup>Verifying the non-linearity of migration distances could be a relevant question. However, the need to calculate the logarithm of this variable (due to the distribution of this variable expressed in kilometres) does not allow us to do it in our model.

When examining subjective mismatch, one can consider that the offered wage for the job is an important omitted variable. Individuals are more likely to feel mismatched when they are offered low-paying jobs rather than well-paid ones. This could introduce bias if the wage variable is correlated with other explanatory variables in our model. However, due to limited information on wages in the *Génération 2010* survey, we chose not to include it in our baseline model. For robustness checks, estimations including wages are still presented in Appendix 3E. Interestingly, while including wages has no effect on the significance of the effect of migration on statistical overeducation, it does have an impact on subjective mismatch. It emphasizes the need for caution when interpreting results from estimations that do not control for wages offered. The divergence in results between the two types of mismatch may illustrate both the actual differences they measure and potential biases in subjective approaches. It is possible that respondents mistake educational or skill mismatch with other factors, such as lower wages, as suggested by Hartog (2000).

### **Estimation strategy and dependent variable**

In our main analysis, we estimate our models in two steps, using a Heckman procedure (Heckman, 1979) to control for selection, then using 2SLS and instrumental variables in linear probability models to control for the endogeneity of migration. Two main limitations may result from such a specification. Firstly, the use of linear probability models can lead to predicted probabilities outside the  $[0; 1]$  interval. Secondly, employment and migration can result from a joint decision and thus be correlated. In order to take both these issues into account, we use a triprobit model to estimate the three equations (employment, migration and mismatch) simultaneously. Tables 3.14 and 3.15 (in Appendix 3E) presents the estimated coefficients of the model. As these coefficients do not represent marginal effects, their orders of magnitude cannot be discussed.

With this alternative estimation strategy, we still observe that health status has a significant impact on employment probability (column 1). Concretely, having a long-lasting health problem remains negatively associated with the likelihood of finding a job. Additionally, most of our independent variables show the expected signs. Specifically, women with children are less likely to be employed. This could be attributed to the challenges of balancing family and professional life, especially when the child is at an early age. Living with parents also has a negative effect on employment probability, which may be due to young people having fewer financial



constraints in this living arrangement. We also find that obtaining a higher diploma decreases the chances of being employed at the time of the survey. This could indicate a more selective job search, with individuals seeking well-paid positions or roles that align better with their personal expectations. It is important to consider that the field of study may significantly influence the chances of finding a job, with graduates from health or education studies being the most advantaged.

Concerning migration (column 2), the decision seems to be linked to some demographic characteristics such as age or having children. People living alone are more likely to migrate as they are more autonomous in their decision-making. The most educated (5-year degree or more) are also the most inclined to migrate, which is consistent with the assumption that their job search is longer and more intense. Moreover, this equation confirms the necessity to consider past migration to explain recent moving. Whether it is a voluntary inter-regional migration or a mandatory stay abroad during studies, a positive and significant influence is shown.

Results relating to the equation for risk of overeducation (column 3) are in straight line with our linear probability models (see Table 3.3). In particular, we confirm that migration (or spatial mobility) helps to reduce the chances of being mismatched. The effect, significant at the 1% level, is then robust to the estimation method retained. Most interesting, the parameters  $\rho$  represent the correlation coefficient between the errors of each of the three probits. As the parameter  $\rho_{23}$  is statistically significantly different from zero at the 5% level, the two probits (for migration and overeducation) have to be estimated jointly. In this case, running simple probits may give biased results. It also confirms the fact that the migration decision is endogenous to the risk of overeducation. On the contrary, the parameter  $\rho_{13}$  associated with probits for employment and overeducation shows that we cannot reject the null hypothesis that the errors are uncorrelated. Therefore, the two equations could be estimated separately. This observation is not surprising as the inverse Mills ratio was only significant at the 10% level in Table 3.3. This does not mean that these equations are uncorrelated, but it could indicate that the data is consistent with no selection.

Finally, even if we propose two different measures of overeducation (the statistical and the subjective ones), they appear to be complementary. To ensure that our results are not related to the measure retained, we use an alternative measure of statistical overeducation, based on the median and quartiles of the distribution. We consider that there is an educational mismatch for the levels farthest from the

median. Therefore, an individual is considered overeducated if his or her level of schooling is higher than the third quartile of the distribution of education in his or her particular occupation. Results for our main estimated models are presented in Appendix 3E. We find a negative and significant effect in the same order of magnitude as the one observed for our first statistical measure. The effects of most of the other independent variables are in line with those previously obtained.

### 3.7 Conclusion

The aim of this study was to examine the impact of spatial mobility on the risk of overeducation among young workers. Specifically, we focused on how residential migration affects their likelihood of experiencing job mismatches, with a particular emphasis on migration to Paris and the educational attainment of individuals. For this analysis, we mobilized data from the *Génération 2010* survey conducted by Céreq, which allowed us to explore both statistical overeducation and subjective mismatch. To address potential biases, we employed the Heckman procedure to account for self-selection into employment and retained an instrumental variables approach to address the endogeneity of migration. Our findings remained robust even after considering these two potential issues.

We show that individuals who change their region of residence between graduation and their current job (observed three years later) have a lower likelihood of being overeducated, based on both the statistical and subjective definitions of mismatch. We find that this negative effect is particularly pronounced for those who migrate to Paris, confirming the attractiveness of this area with its abundance of job opportunities. We also find that spatial mobility does not provide an advantage for all young workers. Specifically, it does not significantly reduce the risk of overeducation for 2-year university graduates. On the other hand, 3- and 4-year university graduates experience a reduction in their risk of overeducation after migration.

In terms of public policy recommendations, these findings support measures or programs aimed at improving workers' mobility. Encouraging or promoting inter-regional migrations of young people appears to be an efficient strategy, as it is associated with a lower risk of job mismatch. It is particularly relevant for this category as the costs associated with migrations might be limited compared to some others. Young people are indeed often renters, single, and without children (or with

young children), making residential relocation more feasible for them. However, for other categories, higher financial and psychological costs might discourage mobility.

It is also important to consider that individuals' migrations are often motivated by the desire to be closer to job opportunities. Therefore, it would be beneficial to develop regions or territories that are suffering from a lack of such opportunities. By increasing the attractiveness of these regions (for firms or economic activities), it becomes easier to facilitate better job matching on the local labour markets. This recommendation is in line with the literature on agglomeration effects developed by Duranton and Puga (2004), as well as recent research such as the one by Berlingieri (2019). The author demonstrates that workers in large cities are less likely to be overqualified for their jobs and more likely to work in the field for which they were trained. Besides, Venhorst et al. (2010) showed that migration between peripheral regions (as opposed to economic centres) is not marginal and does not concern only the least able graduates, but might not be exclusively related to work-related reasons (see also Bernela & Bonnal, 2022). This implies that, in economies showing this centre-periphery geography of labour markets, these attractiveness policies may be needed to ensure a correct match between jobs and workers. This question, in the French context, should nevertheless be examined in order to convince of the necessity of such a policy. It must be the issue for future research on qualification and educational mismatches.

## Appendices

### Appendix 3A: Statistical overeducation measure

Table 3.8: Correspondence of educational levels between classifications

| Educational levels                | Isced                 | RNCP       |
|-----------------------------------|-----------------------|------------|
| 5-year tertiary degree and higher | Isced 7 & 8           | RNCP 7 & 8 |
| 3- or 4-year tertiary degree      | Isced 6               | RNCP 6     |
| 2-year tertiary degree            | Isced 5               | RNCP 5     |
| General high school               | Isced 3 (General) & 4 | RNCP 4     |
| Vocational high school            | Isced 3 (Vocational)  | RNCP 4     |
| Short vocational high school      | Isced 3 (Vocational)  | RNCP 3     |
| No degree                         | Isced 0, 1 & 2        | -          |

Isced: International standard classification of education

RNCP: *Répertoire national des certifications professionnelles*, French classification of education and degrees

Table 3.9: Correspondence of occupational levels between classifications

| PCS     | Isco                  | Occupations  |
|---------|-----------------------|--|
| 32      | Isco 2 except 24 & 25 | Science, health, teaching and cultural professionals |
| 36      | Isco 1, 24 & 25       | Managers & Business and administration professionals |
| 41      | Isco 32               | Health and teaching associate professionals          |
| 46      | Isco 33, 34 & 35      | Business and administration associate professionals  |
| 47 & 48 | Isco 31               | Technicians  |
| 51      | Isco 4* & 5*          | Public sector workers                                |
| 54 & 55 | Isco 4 & 52           | Clerical support and sales workers                   |
| 56      | Isco 51, 53 & 54      | Personal service workers                             |
| 61      | Isco 7 & 8            | Blue-collar skilled workers                          |
| 66      | Isco 9                | Blue-collar unskilled workers                        |
| 69      | Isco 6                | Agricultural workers                                 |

Isco: International standard classification of occupations

\* Occupations in PCS 51 are the similar to those in PCS 54, 55 & 56, but in the public sector.

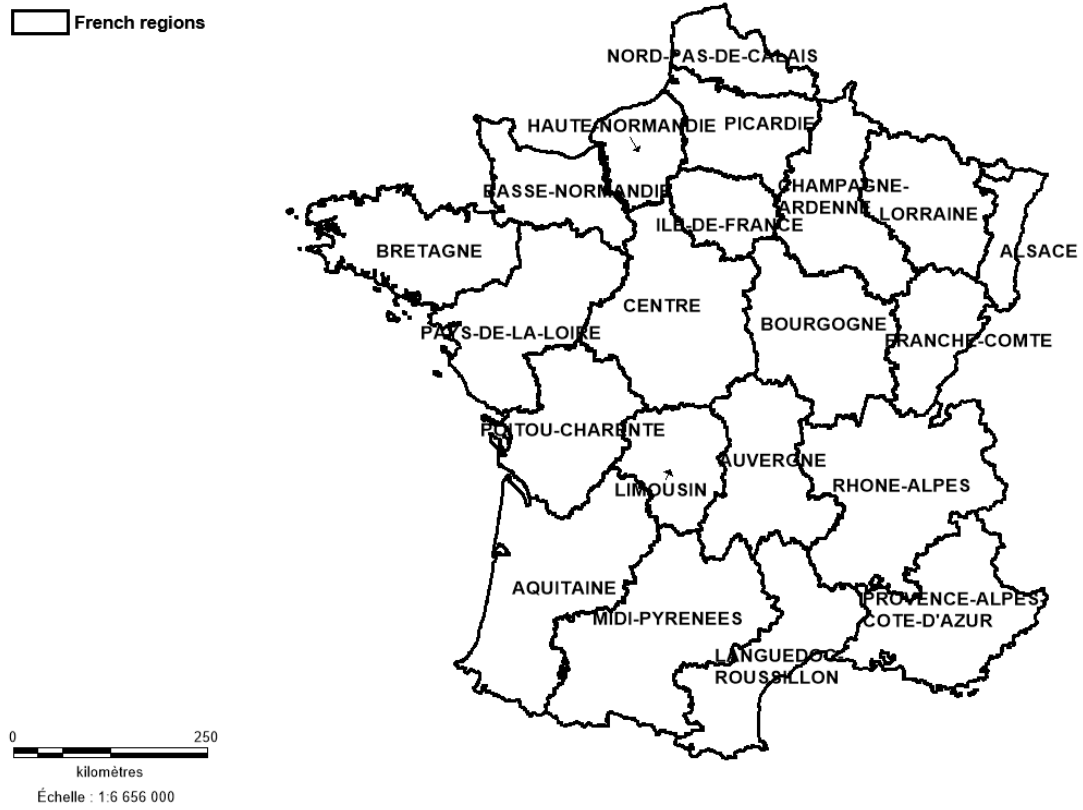
Table 3.10: Occupation-degree correspondence grid - Statistical overeducation

|  | 32 | 36 | 41 | 46 | 47-48 | 51 | 54-55 | 56 | 61 | 66 | 69 |
|--|----|----|----|----|-------|----|-------|----|----|----|----|
| <b>5-year tertiary degree and higher</b> | N  | N  | OE | OE | OE    | OE | OE    | OE | OE | OE | OE |
| <b>3- and 4-year tertiary degree</b>     | N  | N  | N  | OE | OE    | OE | OE    | OE | OE | OE | OE |
| <b>2-year tertiary degree</b>            | N  | N  | N  | N  | N     | OE | OE    | OE | OE | OE | OE |
| <b>General high school</b>               | N  | N  | N  | N  | N     | OE | OE    | OE | OE | OE | OE |
| <b>Vocational high school</b>            | N  | N  | N  | N  | N     | OE | OE    | OE | OE | OE | OE |
| <b>Short vocational high school</b>      | N  | N  | N  | N  | N     | N  | N     | N  | N  | N  | N  |
| <b>No degree</b>                         | N  | N  | N  | N  | N     | N  | N     | N  | N  | N  | N  |

OE stands for overeducation, N for normal. Undereducation is labelled here as normal.

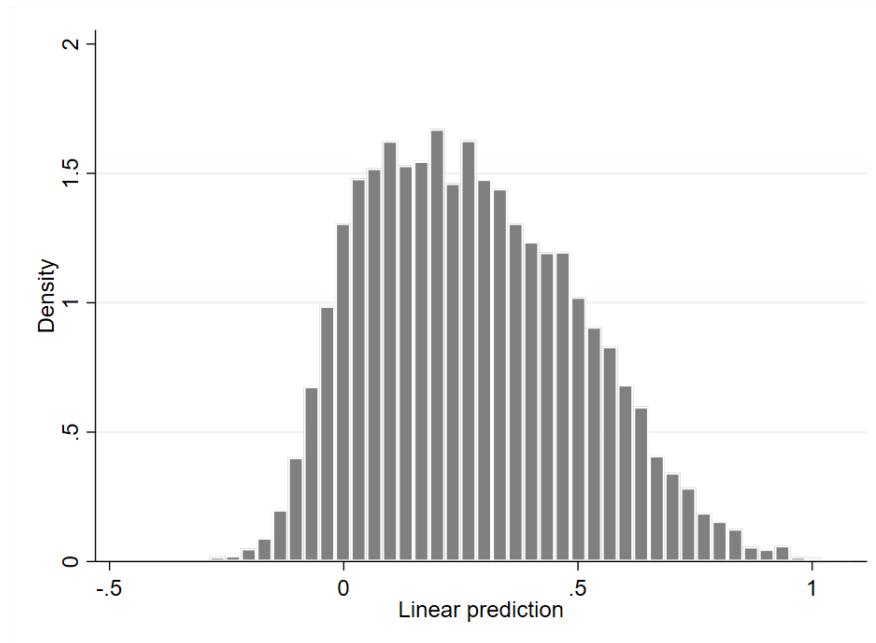
## Appendix 3B: Geography of French regions

Figure 3.1: French regions - NUTS2 areas



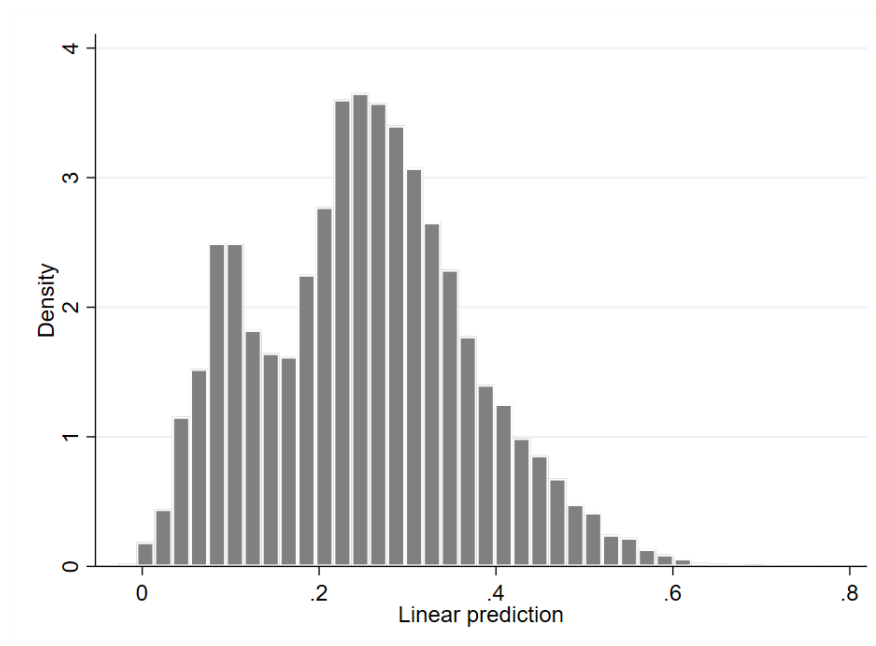
## Appendix 3C: Distribution of predicted mismatch within the sample

Figure 3.2: Predicted probabilities - Statistical overeducation



Source: Authors' calculations based on *Génération 2010* data.

Figure 3.3: Predicted probabilities - Subjective mismatch



Source: Authors' calculations based on *Génération 2010* data.

## Appendix 3D: Selection equation and first stage regression for 2SLS

Table 3.11: Employment equation and IV estimation (first stage)

|   | Employment           | Migration<br>IV (first stage) |
|---|----------------------|-------------------------------|
| Constant                                      | 1.101***<br>(0.107)  | 0.687***<br>(0.080)           |
| <i>Demographic characteristics</i>            |                      |                               |
| Woman   | 0.012<br>(0.025)     | -0.005<br>(0.007)             |
| Age   | -0.005<br>(0.004)    | -0.018***<br>(0.003)          |
| Having children                               | 0.022<br>(0.057)     | -0.057***<br>(0.011)          |
| Woman × Children                              | -0.420***<br>(0.046) | 0.017<br>(0.015)              |
| <i>Residential status (Ref: Living alone)</i> |                      |                               |
| Living with a partner                         | 0.159***<br>(0.027)  | -0.049***<br>(0.008)          |
| Living with their parents                     | -0.492***<br>(0.047) | -0.135***<br>(0.027)          |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                               |
| 3- or 4- year degree                          | -0.221***<br>(0.043) | 0.077***<br>(0.019)           |
| 5-year degree or more                         | -0.105**<br>(0.049)  | 0.133***<br>(0.019)           |
| <i>Field of study (Ref: Services)</i>         |                      |                               |
| Education                                     | 0.526*<br>(0.279)    | 0.082<br>(0.057)              |
| Humanities                                    | -0.175***<br>(0.035) | -0.002<br>(0.012)             |
| Economics, business and law                   | 0.018<br>(0.041)     | -0.002<br>(0.012)             |
| Science                                       | 0.106**<br>(0.044)   | 0.045***<br>(0.015)           |
| Industry                                      | 0.296***<br>(0.047)  | 0.042***<br>(0.017)           |
| Agriculture                                   | -0.109<br>(0.089)    | 0.075**<br>(0.030)            |
| Health  | 0.421***<br>(0.081)  | 0.081***<br>(0.016)           |
| Long-lasting health problems                  | -0.131***<br>(0.045) |                               |
| Migration during studies                      |                      | 0.413***<br>(0.052)           |
| Mandatory study abroad                        |                      | 0.094**<br>(0.036)            |
| Region fixed effects                          | Yes                  | Yes                           |
| R-squared                                     | 0.063                | 0.263                         |
| N   | 15 865               | 15 865                        |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.

## Appendix 3E: Alternative estimations and robustness checks

Table 3.12: Linear probability models for statistical overeducation - Alternative measure

|   | OLS                  | Heckman              | Heckman + IV         |
|---|----------------------|----------------------|----------------------|
| Constant                                      | 0.512***<br>(0.035)  | 0.357***<br>(0.038)  | 0.393***<br>(0.111)  |
| Migration                                     | -0.025**<br>(0.009)  | -0.025**<br>(0.009)  | -0.053***<br>(0.015) |
| <i>Demographic characteristics</i>            |                      |                      |                      |
| Woman   | 0.037***<br>(0.006)  | 0.038***<br>(0.006)  | 0.038***<br>(0.007)  |
| Age   | -0.012***<br>(0.001) | -0.013***<br>(0.001) | -0.014***<br>(0.001) |
| Having children                               | 0.008<br>(0.006)     | 0.015<br>(0.019)     | 0.014<br>(0.016)     |
| Woman × Children                              | 0.006<br>(0.023)     | -0.085***<br>(0.023) | -0.085***<br>(0.027) |
| <i>Residential status (Ref: Living alone)</i> |                      |                      |                      |
| Living with a partner                         | -0.007<br>(0.009)    | 0.024*<br>(0.013)    | 0.022**<br>(0.010)   |
| Living with their parents                     | 0.076***<br>(0.011)  | -0.047<br>(0.031)    | -0.052*<br>(0.029)   |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                      |                      |
| 3- or 4- year degree                          | 0.118***<br>(0.011)  | 0.066***<br>(0.014)  | 0.070***<br>(0.016)  |
| 5-year degree or more                         | 0.228***<br>(0.016)  | 0.207***<br>(0.014)  | 0.214***<br>(0.012)  |
| <i>Field of study (Ref: Services)</i>         |                      |                      |                      |
| Education                                     | 0.021<br>(0.050)     | 0.126**<br>(0.056)   | 0.127**<br>(0.052)   |
| Humanities                                    | 0.073***<br>(0.014)  | 0.025<br>(0.018)     | 0.024<br>(0.017)     |
| Economics, business and law                   | -0.006<br>(0.014)    | -0.002<br>(0.013)    | -0.003<br>(0.011)    |
| Science                                       | -0.069***<br>(0.013) | -0.045***<br>(0.012) | -0.044***<br>(0.013) |
| Industry                                      | -0.040***<br>(0.014) | 0.043**<br>(0.019)   | 0.026<br>(0.018)     |
| Agriculture                                   | 0.139***<br>(0.035)  | 0.110***<br>(0.035)  | 0.114***<br>(0.024)  |
| Health  | -0.066***<br>(0.012) | 0.015<br>(0.017)     | 0.017<br>(0.022)     |
| <i>Business sector (Ref: Services)</i>        |                      |                      |                      |
| Public sector                                 | -0.045***<br>(0.010) | -0.045***<br>(0.010) | -0.044***<br>(0.007) |
| Agriculture                                   | 0.399***<br>(0.066)  | 0.397***<br>(0.066)  | 0.398***<br>(0.037)  |
| Sales   | 0.168***<br>(0.011)  | 0.168***<br>(0.011)  | 0.168***<br>(0.011)  |
| Construction                                  | 0.048*<br>(0.027)    | 0.049*<br>(0.027)    | 0.049***<br>(0.021)  |
| Industry                                      | 0.042**<br>(0.019)   | 0.043**<br>(0.019)   | 0.043***<br>(0.011)  |
| <i>Job characteristics</i>                    |                      |                      |                      |
| Working full-time                             | -0.077***<br>(0.014) | -0.077***<br>(0.014) | -0.076***<br>(0.010) |
| Permanent job                                 | 0.055***<br>(0.010)  | -0.054***<br>(0.010) | -0.054***<br>(0.007) |
| Duration between graduation and first job     | 0.006***<br>(0.001)  | 0.006***<br>(0.001)  | 0.006***<br>(0.001)  |
| Inverse Mills ratio                           |                      | 0.580***<br>(0.121)  | 0.578***<br>(0.128)  |
| Region fixed effects                          | Yes                  | Yes                  | Yes                  |
| R-squared                                     | 0.160                | 0.161                | 0.160                |
| N   | 12 968               | 12 962               | 12 962               |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.



Table 3.13: Linear probability models including wage (mismatch equations)

|   | Statistical<br>overeducation | Subjective<br>mismatch |
|---|------------------------------|------------------------|
| Constant                                      | 2.655***<br>(0.151)          | 1.778***<br>(0.166)    |
| Migration                                     | -0.062***<br>(0.018)         | -0.033*<br>(0.019)     |
| <i>Demographic characteristics</i>            |                              |                        |
| Woman   | 0.042***<br>(0.008)          | -0.018**<br>(0.009)    |
| Age   | -0.014***<br>(0.002)         | 0.007***<br>(0.002)    |
| Having children                               | 0.036**<br>(0.018)           | 0.007***<br>(0.002)    |
| Woman × Children                              | -0.047<br>(0.031)            | -0.074**<br>(0.034)    |
| <i>Residential status (Ref: Living alone)</i> |                              |                        |
| Living with a partner                         | 0.012<br>(0.011)             | 0.036***<br>(0.012)    |
| Living with their parents                     | 0.049<br>(0.033)             | -0.045<br>(0.036)      |
| <i>Educational level (Ref: 2-year degree)</i> |                              |                        |
| 3- or 4- year degree                          | 0.252***<br>(0.018)          | 0.016<br>(0.019)       |
| 5-year degree or more                         | 0.128***<br>(0.014)          | 0.030**<br>(0.015)     |
| <i>Field of study (Ref: Services)</i>         |                              |                        |
| Education                                     | -0.116**<br>(0.059)          | -0.026<br>(0.064)      |
| Humanities                                    | -0.055***<br>(0.019)         | -0.009<br>(0.021)      |
| Economics, business and law                   | 0.071***<br>(0.013)          | 0.018<br>(0.014)       |
| Science                                       | -0.069***<br>(0.014)         | -0.014<br>(0.016)      |
| Industry                                      | 0.008<br>(0.021)             | 0.036<br>(0.023)       |
| Agriculture                                   | 0.096***<br>(0.027)          | -0.054*<br>(0.029)     |
| Health  | -0.139***<br>(0.025)         | -0.069**<br>(0.027)    |
| <i>Business sector (Ref: Services)</i>        |                              |                        |
| Public sector                                 | -0.079***<br>(0.008)         | -0.021**<br>(0.009)    |
| Agriculture                                   | 0.284***<br>(0.044)          | -0.034<br>(0.048)      |
| Sales   | 0.210***<br>(0.012)          | 0.140***<br>(0.014)    |
| Construction                                  | 0.086***<br>(0.023)          | -0.010<br>(0.025)      |
| Industry                                      | 0.089***<br>(0.013)          | 0.065***<br>(0.014)    |
| <i>Job characteristics</i>                    |                              |                        |
| Wage (ln)                                     | -0.298***<br>(0.013)         | -0.237***<br>(0.014)   |
| Working full-time                             | 0.090***<br>(0.014)          | -0.033**<br>(0.016)    |
| Permanent job                                 | -0.019**<br>(0.008)          | -0.010<br>(0.009)      |
| Duration between graduation and first job     | 0.004***<br>(0.001)          | 0.004***<br>(0.001)    |
| Inverse Mills ratio                           | 0.298***<br>(0.071)          | 0.431***<br>(0.160)    |
| Region fixed effects                          | Yes                          | Yes                    |
| R-squared                                     | 0.250                        | 0.160                  |
| N   | 12 337                       | 12 337                 |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region. Estimated coefficients are obtained with the model controlling for selection and endogeneity biases (Heckman procedure + Instrumental variables).

Source: Authors' calculations based on *Génération 2010* data.

Table 3.14: Triprobit model for employment, migration and overeducation

|   | (1) Employment       | (2) Migration        | (3) Overeducation    |
|---|----------------------|----------------------|----------------------|
| Constant                                      | 1.111***<br>(0.112)  | 6.753***<br>(0.261)  | 1.172***<br>(0.242)  |
| Migration                                     |                      |                      | -0.274***<br>(0.069) |
| <i>Demographic characteristics</i>            |                      |                      |                      |
| Woman   | 0.011<br>(0.025)     | -0.021<br>(0.023)    | 0.216***<br>(0.024)  |
| Age   | -0.005<br>(0.004)    | -0.056***<br>(0.008) | -0.066***<br>(0.007) |
| Having children                               | 0.022<br>(0.057)     | -0.179***<br>(0.035) | 0.106<br>(0.087)     |
| Woman × Children                              | -0.419***<br>(0.046) | 0.043<br>(0.055)     | -0.049<br>(0.105)    |
| <i>Residential status (Ref: Living alone)</i> |                      |                      |                      |
| Living with a partner                         | 0.159***<br>(0.027)  | -0.151***<br>(0.022) | -0.007<br>(0.034)    |
| Living with their parents                     | -0.492***<br>(0.047) | -0.563***<br>(0.091) | 0.344***<br>(0.039)  |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                      |                      |
| 3- or 4- year degree                          | -0.222***<br>(0.044) | 0.271***<br>(0.064)  | 0.816***<br>(0.034)  |
| 5-year degree or more                         | -0.104**<br>(0.048)  | 0.442***<br>(0.058)  | 0.342***<br>(0.052)  |
| <i>Field of study (Ref: Services)</i>         |                      |                      |                      |
| Education                                     | 0.526*<br>(0.278)    | 0.268<br>(0.186)     | -0.298<br>(0.232)    |
| Humanities                                    | -0.175***<br>(0.035) | -0.001<br>(0.059)    | -0.021<br>(0.076)    |
| Economics, business and law                   | 0.018<br>(0.041)     | -0.007<br>(0.042)    | 0.163***<br>(0.040)  |
| Science                                       | 0.106**<br>(0.044)   | 0.147***<br>(0.050)  | -0.316***<br>(0.048) |
| Industry                                      | 0.296***<br>(0.047)  | 0.134**<br>(0.058)   | -0.078*<br>(0.041)   |
| Agriculture                                   | -0.109<br>(0.089)    | 0.273**<br>(0.108)   | 0.329***<br>(0.092)  |
| Health  | 0.421***<br>(0.081)  | 0.273***<br>(0.054)  | -1.132***<br>(0.083) |
| <i>Business sector (Ref: Services)</i>        |                      |                      |                      |
| Public sector                                 |                      |                      | -0.370***<br>(0.042) |
| Agriculture                                   |                      |                      | 0.847***<br>(0.164)  |
| Sales   |                      |                      | 0.626***<br>(0.041)  |
| Construction                                  |                      |                      | 0.193*<br>(0.115)    |
| Industry                                      |                      |                      | 0.202***<br>(0.055)  |
| <i>Job characteristics</i>                    |                      |                      |                      |
| Working full-time                             |                      |                      | -0.218***<br>(0.043) |
| Permanent job                                 |                      |                      | -0.203***<br>(0.037) |
| Duration between graduation and first job     |                      |                      | 0.020***<br>(0.003)  |
| Long-lasting health problems                  | -0.128***<br>(0.045) |                      |                      |
| Migration during studies                      |                      | 1.186***<br>(0.148)  |                      |
| Mandatory study abroad                        |                      | 0.302***<br>(0.102)  |                      |
| Region fixed effects                          | Yes                  | Yes                  | Yes                  |
| $\rho_{13}$                                   |                      | -0.071 (0.113)       |                      |
| $\rho_{23}$                                   |                      | 0.096** (0.047)      |                      |
| $\rho_{12}$                                   |                      | 0.023 (0.020)        |                      |
| Log-pseudolikelihood                          |                      | -21096.84            |                      |
| $N$   |                      | 15 864               |                      |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.

Table 3.15: Triprobit model for employment, migration and subjective mismatch

|   | (1) Employment       | (2) Migration        | (3) Mismatch         |
|---|----------------------|----------------------|----------------------|
| Constant                                      | 1.128***<br>(0.102)  | 6.773***<br>(0.269)  | -0.881***<br>(0.153) |
| Migration                                     |                      |                      | -0.143***<br>(0.044) |
| <i>Demographic characteristics</i>            |                      |                      |                      |
| Woman   | 0.014<br>(0.025)     | -0.021<br>(0.023)    | 0.027<br>(0.026)     |
| Age   | -0.005<br>(0.004)    | -0.056***<br>(0.008) | 0.014***<br>(0.005)  |
| Having children                               | 0.020<br>(0.055)     | -0.179***<br>(0.035) | 0.071<br>(0.050)     |
| Woman × Children                              | -0.418***<br>(0.046) | 0.044<br>(0.055)     | -0.063<br>(0.083)    |
| <i>Residential status (Ref: Living alone)</i> |                      |                      |                      |
| Living with a partner                         | 0.156***<br>(0.027)  | -0.151***<br>(0.022) | 0.044<br>(0.030)     |
| Living with their parents                     | -0.494***<br>(0.046) | -0.563***<br>(0.091) | 0.084<br>(0.093)     |
| <i>Educational level (Ref: 2-year degree)</i> |                      |                      |                      |
| 3- or 4- year degree                          | -0.222***<br>(0.044) | 0.271***<br>(0.064)  | 0.077*<br>(0.030)    |
| 5-year degree or more                         | -0.111**<br>(0.049)  | 0.442***<br>(0.059)  | -0.052<br>(0.041)    |
| <i>Field of study (Ref: Services)</i>         |                      |                      |                      |
| Education                                     | 0.524*<br>(0.273)    | 0.273<br>(0.182)     | -0.109<br>(0.123)    |
| Humanities                                    | -0.173***<br>(0.035) | -0.001<br>(0.059)    | 0.073<br>(0.089)     |
| Economics, business and law                   | 0.020<br>(0.041)     | -0.008<br>(0.042)    | 0.012<br>(0.040)     |
| Science                                       | 0.106**<br>(0.044)   | 0.147***<br>(0.050)  | -0.098**<br>(0.041)  |
| Industry                                      | 0.296***<br>(0.047)  | 0.135***<br>(0.057)  | -0.013<br>(0.064)    |
| Agriculture                                   | -0.106<br>(0.088)    | 0.273**<br>(0.107)   | 0.090<br>(0.094)     |
| Health  | 0.419***<br>(0.081)  | 0.273***<br>(0.054)  | -0.607***<br>(0.056) |
| <i>Business sector (Ref: Services)</i>        |                      |                      |                      |
| Public sector                                 |                      |                      | -0.006<br>(0.029)    |
| Agriculture                                   |                      |                      | -0.134<br>(0.126)    |
| Sales   |                      |                      | 0.425***<br>(0.037)  |
| Construction                                  |                      |                      | -0.059<br>(0.117)    |
| Industry                                      |                      |                      | 0.171***<br>(0.064)  |
| <i>Job characteristics</i>                    |                      |                      |                      |
| Working full-time                             |                      |                      | -0.143***<br>(0.038) |
| Permanent job                                 |                      |                      | -0.010<br>(0.036)    |
| Duration between graduation and first job     |                      |                      | 0.018***<br>(0.003)  |
| Long-lasting health problems                  | -0.153***<br>(0.045) |                      |                      |
| Migration during studies                      |                      | 1.187***<br>(0.147)  |                      |
| Mandatory study abroad                        |                      | 0.293***<br>(0.102)  |                      |
| Region fixed effects                          | Yes                  | Yes                  | Yes                  |
| $\rho_{13}$                                   |                      | 0.511* (0.286)       |                      |
| $\rho_{23}$                                   |                      | 0.075** (0.038)      |                      |
| $\rho_{12}$                                   |                      | 0.024 (0.019)        |                      |
| Log-pseudolikelihood                          |                      | -22097.66            |                      |
| $N$   |                      | 15 864               |                      |

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level. Standard errors are clustered by region.

Source: Authors' calculations based on *Génération 2010* data.



## Chapter 4

# Educational mismatch and early career trajectories

This chapter investigates the “stepping stone” effect of vertical, horizontal and full mismatch at the beginning of the career. Using dynamic propensity score matching in a multiple treatment framework, I find that mismatch in the first job, either vertical or horizontal, leads to a higher probability of still being mismatched after three years in comparison with longer unemployment. While vertical and horizontal mismatch can improve future labour market outcomes (employment and earnings), especially if the first job is taken early after graduation, full mismatch has very limited positive effects on outcomes, and stronger detrimental effects on the transition to a matched job.

**JEL Classification:** I21, J24, J62

**Keywords:** educational mismatch, overeducation, horizontal mismatch, school-to-work transitions, dynamic matching

---

This chapter has been written as part of the exploitation group of the *Génération 2017* data from Céreq (*Centre d'études et de recherches sur les qualifications*). I thank the Céreq for providing anticipated access to the data.

# Inadéquation emploi-diplôme et trajectoires professionnelles en début de carrière

Cet article étudie l'effet de « tremplin » des inadéquations verticale, horizontale et complète en début de carrière. Utilisant un modèle de *matching* dynamique sur le score de propension, appliqué à des traitements multiples, je montre que l'inadéquation dans le premier emploi, qu'elle soit verticale ou horizontale, augmente la probabilité de ne pas être dans un emploi adéquat après trois ans, relativement au fait de rester plus longtemps au chômage. Alors que les inadéquations verticale et horizontale permettent dans une certaine mesure d'améliorer l'emploi et les salaires futurs, en particulier si le premier emploi est accepté rapidement après l'obtention du diplôme, l'inadéquation complète a des effets positifs très limités sur ces variables et des effets négatifs plus importants sur la transition vers un emploi adéquat.

**Mots-clés :** inadéquation emploi-diplôme, sur-éducation, inadéquation horizontale, transitions éducation-marché du travail, *matching* dynamique

## 4.1 Introduction

At the beginning of their career, young workers face limited labour market prospects (Bell & Blanchflower, 2011; Blanchflower & Freeman, 2000). They are more often unemployed (23.5% for the 15-24 year-olds against 9.4% for the whole labour force in France in 2017) and in temporary jobs (57.8% for the 15-24 year-olds against 16.9% for the whole labour force in France in 2017) than older workers<sup>1</sup>. To avoid the scarring effects of unemployment (Arulampalam, 2001) and temporary employment (Filomena & Picchio, 2021), young workers may be tempted to accept any job early after graduation, leading to potential mismatches between their education and their labour market position. These mismatches can be either vertical (overeducation), *i.e.* the educational level of the individuals is higher than the one required for their job (Kucel, 2011; McGuinness, 2006), or horizontal, *i.e.* individuals work in fields that are different from the one they studied (Robst, 2008; Somers et al., 2019). There is large evidence of negative effects of both overeducation (Hartog, 2000; Rubb, 2003; Verdugo & Verdugo, 1989) and horizontal mismatch (Nordin et al., 2010; Robst, 2007a, 2007b) on labour market outcomes (especially wages) in comparison with individuals with the same education but working in a matched job.

Despite these detrimental effects, according to the career mobility theory, being overeducated at the beginning of the career should contribute to a shorter and better access to higher level positions (Sicherman & Galor, 1990). In particular, these authors show that the most educated workers in a given occupation were more likely to evolve to higher labour market positions. However, a number of empirical studies show that overeducation is more of a “trap” than a “stepping stone” for young graduates (Baert et al., 2013; Büchel & Mertens, 2004; Meroni & Vera-Toscano, 2017; Scherer, 2004). It has been shown that entering the labour market as overeducated leads to remain mismatched for a long time before accessing a suitable job (Büchel & Mertens, 2004; Scherer, 2004). It has been argued that overeducated workers may be individuals with lower abilities, which could be the reason of the trap rather than the overeducation situation itself (Büchel & Mertens, 2004). To deal with this potential issue, Baert et al. (2013) use a Timing of Events approach to control for selection on unobservables. Their results are consistent to the previous ones: overeducation delays the transition to an adequate job. Meroni and Vera-

---

<sup>1</sup>OECD (2023), Unemployment rate by age group (indicator). doi: 10.1787/997c8750-en; Unemployment rate (indicator). doi: 10.1787/52570002-en; Temporary employment (indicator). doi: 10.1787/75589b8a-en (Accessed on 06 September 2023)

Toscano (2017) find similar results using dynamic treatment approaches. They show that accepting a first job for which they are overeducated leads to stay mismatched in subsequent jobs, at least during the five years following graduation. Still, many young workers accept such a job in order to enter the labour market more quickly and to avoid unemployment, as it has larger scarring effects than overeducation on their career, especially if mismatched only in temporary jobs (Baert & Verhaest, 2019).

Evidence about the stepping stone or detrimental effects of horizontal mismatch on future occupational mobility is more limited. Wolbers (2003) find that horizontally mismatched workers have a lower occupational status than their matched counterparts, but that they more often look for another job and engage in on-the-job training in order to improve their match. However, nothing is said about the actual impact on future positions. On a slightly different matter, Meroni and Vera-Toscano (2017) show that horizontal mismatch amplifies the “trap” effect of overeducation.

The existence of a “stepping stone effect” for young people entering the job market has not exclusively been studied for mismatch, but also related to temporary contracts or unemployment. Empirical evidence is mixed. Some studies find a positive effect of temporary employment on future labour market outcomes (Cockx & Picchio, 2012; Ichino et al., 2008), while others evidence a “dead end effect” (Alba-Ramírez, 1998; Amuedo-Dorantes, 2000). Some also show an ambiguous effect: for instance, de Graaf-Zijl et al. (2011) show that temporary jobs reduce the duration to find a permanent job in comparison with unemployment, but find no significant effect on the probability of finding a permanent job itself. In a meta-analysis, Filomena and Picchio (2021) show that “stepping stone effects” are less likely to occur when economic conditions are less favorable. As for unemployment, some works show that unemployed job search gives access to better paid jobs than on-the-job search (Kahn & Low, 1982), while others evidence the opposite (D. M. Blau & Robins, 1990).

I contribute to this literature by investigating the “stepping stone effect” of mismatch (Sicherman & Galor, 1990), not only in the case of overeducation, but also in the case of horizontal and full (both vertical and horizontal) mismatches. From a theoretical point of view, horizontal mismatch has an ambiguous effect. On the one hand, it might demonstrate a failure to find a job related to the degree obtained, and thus lower abilities. Indeed, firms first hire the most competent workers in their particular field. On the other hand, it suggests an ability to adapt



that could be valued in the workplace and facilitate job-to-job mobility. In the latter case, repeated job search could improve the quality of the match (Jovanovic, 1979) and allow access to better positions on the labour market. As for full mismatch, we can expect strong negative effects on future occupational positions, as individuals mismatched both horizontally and vertically are the furthest from the skills and knowledge acquired through education. Indeed, detrimental effects of mismatch are generally higher when associated to a lack of (relevant) skills (Chevalier, 2003). Nevertheless, empirical evidence related to full mismatch is scarce. Verhaest et al. (2017) show that differences regarding the drivers of full mismatch in comparison with other mismatches, but did not document their consequences on labour market outcomes.

In addition, I want to explore the effect of taking a mismatched job at the beginning of the career on future labour market outcomes (employment and earnings). Some studies explore the effects of overeducation on further employment, with inconclusive results. On the one hand, Baert and Verhaest (2019) show that overeducation leads to smaller detrimental effects than unemployment on further employment, although it increases the risk in comparison of taking a suitable job. Rubb (2003) also find that past overeducation came with a relative protection against part-time employment in subsequent positions. On the other hand, Meroni and Vera-Toscano (2017) show that being overeducated does not increase the probability of having a job later on, in comparison with remaining longer unemployed. I contribute to this literature by exploring the effect, not only of vertical mismatch, but also of horizontal and full mismatch on further employment, which has to my knowledge not been explored yet. Finally, I investigate the effects of these mismatches on earnings at the beginning of the career. On the one hand, accepting a job earlier when entering the labour market increases work experience, and thus should be associated with higher wages. On the other hand, mismatched workers earn lower wages than individuals with the same education but working in a matched job (Hartog, 2000; Nordin et al., 2010; Robst, 2007a; Verdugo & Verdugo, 1989). I thus aim to verify if there is a significant premium on further earnings to accept a mismatched job rather than staying longer in unemployment.

To investigate these questions, I use data from the labour market integration survey *Génération 2017* from the French Institute of studies and research on qualifications (Céreq<sup>2</sup>), which allows me to follow young workers during three years after

---

<sup>2</sup>*Centre d'études et de recherches sur les qualifications.*

graduation. I use a dynamic treatment framework based on propensity score matching (Abadie & Imbens, 2016; Rosenbaum & Rubin, 1983) to estimate the effects of being mismatched in the first job, rather than staying longer in unemployment, on the labour market position and outcomes after three years. I use a multiple treatment framework (Lechner, 2001), where treatments correspond to the mismatch status in the first job. In this regard, my approach is similar to the one used in Voßemer and Schuck (2016) and Meroni and Vera-Toscano (2017), although the treatments considered are different. Such an approach allows me to account for the differences in the timing of treatment (Sianesi, 2004, 2008), as some individuals take their first job earlier than others after graduation. Indeed, the decision to accept (or not) a job earlier might be correlated with the risk of mismatch: individuals who prefer to stay unemployed at a given point in time (possibly not to be mismatched) may eventually accept to be mismatched to find a job (Meroni & Vera-Toscano, 2017).

The main finding coming from this analysis is that accepting a first job for which individuals are mismatched, either vertically, horizontally or in both dimensions, rather than staying longer unemployed results in a higher risk of being mismatched three years after graduation. Therefore, mismatch, regardless of its type, is not a “stepping stone” to better positions, but a trap for young graduates entering the labour market. The heterogeneity analysis reveal that the more educated suffer the highest detrimental effects on future match, there is no significant difference in the “trap effect” between men and women.

If taken early (but not necessarily immediately) after graduation, mismatched positions have positive effects on future labour market outcomes (employment and earnings), which can explain why some young workers accept such jobs despite the risk of being trapped in it. If taken more than six months after graduation, however, mismatched jobs do not lead to better prospects than longer unemployment. In any case, young graduates should avoid jobs for which they are fully mismatched, as it has very limited positive effects on future labour market outcomes in comparison with longer unemployment, and even has negative effects if taken too late after graduation.

Section 4.2 describes the data and the construction of the mismatch statuses variables, and provides some descriptive statistics. Section 4.3 discusses the treatment framework and the estimation strategy. Section 4.4 provides and comments

the results. Section 4.5 concludes.

## 4.2 Data and variables

### 4.2.1 Sample

This work is based on the most recent iteration of the *Génération* survey from Céreq. Created at the beginning of the 1990s, this survey aims to inform on school-to-work transitions and the labour market entry of French young people. By interviewing these individuals at different times, it enables to follow their early career paths during their first years on the labour market (3 years for the first interrogation, 5 to 10 years using subsequent interrogations, depending on the generation). The *Génération 2017* (the eighth generation to be surveyed) has been interrogated for the first time in 2020, and will be interrogated again in 2023 (6 years after their entry on the labour market).

The *Génération 2017* survey documents the entry and the trajectories on the labour market of 25164 individuals who finished their studies between October 2016 and December 2017, and interrogated around three years after graduation. This data provides some individual information about these workers (such as gender, age, cohabitation status, number of children), and detailed information about the highest degree obtained and the field of study. It also provides information on past education, as well as familial and geographical background. Moreover, a monthly calendar allows to observe the successive labour market positions of these individuals between their graduation and the time at which they were interrogated. Especially, I am able to determine the exact date (month and year) of their first job, and thus the duration since graduation.

This paper focuses on the educational mismatch of tertiary graduates and the evolution of their mismatch status between their first job and their position at the time of the survey (around three years after graduation). I exclude from my sample the individuals who, at one of these dates at least, occupy political positions, religious occupations or are business owners with more than 10 employees. Following Meroni and Vera-Toscano (2017) and Verhaest and van der Velden (2013), are also excluded individuals who never found a job on the whole period and those who do not report their occupation. Finally, the sample is composed of 13,056 individuals, among which 10,630 are used for dynamic matching (by design, individuals who

find a matched job right after graduation are never in the treated nor in the control group).

In addition to the *Génération 2017* survey, I mobilise other datasets to define the different educational mismatches. As for vertical mismatch (overeducation), I mobilise the 2017 French Labour Force Survey (LFS) from the French National Institute of Statistics and Economic Studies (Insee) to define the required educational levels for each occupation on the whole population. Indeed, since young workers are more likely to be overeducated than older workers, observing the distribution of education by occupation only for their demographic group may bias the definition of overeducation. The LFS dataset provides detailed information on socioeconomic, job and educational characteristics for almost 200,000 employed individuals for year 2017, ensuring a sufficient sample size in each occupation to compute the needed statistics.

As for horizontal educational mismatch, I determine the correspondence between occupations and fields of education using the NSF classification<sup>3</sup> from Insee. This nomenclature defines 93 fields of education (3-digits level), divided in four large groups: “Disciplinary fields”, “Technical-professional fields of production”, “Technical-professional fields of services” and “Personal development fields”. This detailed classification allows me to link precisely occupations to the associated fields of study.

## 4.2.2 Mismatches

### Overeducation

Overeducation (or vertical mismatch) occurs when individuals’ educational level is higher than the one which is *a priori* required to do their job (Kucel, 2011; McGuinness, 2006). In this paper, I will rely on a statistical approach to measure overeducation<sup>4</sup>, based on the modes of the distribution of education in each occupation (Davia et al., 2017; Ghignoni & Verashchagina, 2014; Mendes de Oliveira

<sup>3</sup>*Nomenclature des spécialités de formation*, Educational specialties classification.

<sup>4</sup>Alternatively, overeducation could be measured objectively (Rumberger, 1981) using an *ex ante* correspondence that matches occupations to the most adequate educational levels; or subjectively (Duncan & Hoffman, 1981) using workers’ self-assessments (see e.g. Capsada-Munsech, 2017; Verhaest & Omey, 2006, for further discussion of these different measurement approaches). However, the structure of the data does not allow me to use these approaches, as there is no indication of the theoretical adequate levels for each occupation, needed for the objective measure. As for the subjective approach, the dataset includes a worker’s assessment about whether she or he is employed above, below or at her or his level of skills, but this information is available only for the job occupied at the time of the survey, and not for the first job, preventing me to exploit this approach.

et al., 2000). Given that I rely on large occupational groups (2-digits PCS<sup>5</sup>) to ensure a sufficient sample size, I use two modes rather than only one to take the variety of jobs in a same group into account. Using this definition, individuals are considered as statistically overeducated if their educational level is above the higher of the two most common levels for their particular occupation (see Appendix 4A for more details on the measure of overeducation).

### Horizontal mismatch

Horizontal mismatch occurs when individuals work in a field that does not match the one of their studies (Robst, 2008; Somers et al., 2019). Based on the NSF classification, I define which fields of study are associated with each of the 311 occupations (4-digits PCS)<sup>6</sup>. Most occupational titles refer directly to particular educational specialties, although several fields can be linked to the same occupation<sup>7</sup>. Conversely, most educational fields are linked to several occupations. When individuals work in occupations that are not linked to their particular field of study, they are considered as horizontally mismatched<sup>8</sup>.

### 4.2.3 Descriptive statistics

Table 4.1 provides descriptive statistics about the sample, depending on the matching status of the individuals in their first job. The first striking fact is that, in the first job, more than two third of the graduates are mismatched in some way. More than 40% are overeducated and more than a half are horizontally mismatched, with around 25% of the workers being mismatched in both dimensions. Among overeducated workers, 60% are also horizontally mismatched; among horizontally mismatched workers, around half are also vertically mismatched.

Mismatched workers tend to wait longer before taking their first job, although the gap is small: the average duration between graduation and first is lower than

<sup>5</sup>*Professions et Catégories Socio-professionnelles* 2020, French occupational classification

<sup>6</sup>Due to its large size, the fields of education-occupations matching table is not reported here, but can be provided upon request.

<sup>7</sup>Some educational fields are labelled as “pluridisciplinary” in the classification. These fields are systematically considered as matched to an occupation if another educational field of the same group is also matched to this occupation.

<sup>8</sup>Teachers and researchers occupations (9 socio-economic categories) require particular attention. They can be linked to various educational fields depending on their specialty, on which the *Génération 2017* does not provide any information. I match them to all fields in the “Disciplinary fields” group, as well as the “Education, training” field. This should avoid these individuals to be wrongly considered as mismatched. However, it has to be noted that teachers (researchers) who graduated in a particular field but teaching (doing research) in another are considered as matched.

three months in any case. Overeducated workers are slightly younger than the others (regardless whether they are only overeducated or also mismatched horizontally). Probably related to that, they also are more likely to still live with their parents. There are not much gender disparities between types of mismatch, except for horizontal mismatch only, where women are underrepresented. Horizontally mismatched workers also appear to be slightly more spatially mobile than other workers.

Table 4.1: Descriptive statistics - By mismatch status in the first job

|  | Matched |       | Only vertical mismatch |       | Only horizontal mismatch |       | Both mismatches |       |
|--|---------|-------|------------------------|-------|--------------------------|-------|-----------------|-------|
|  | Mean    | S.E.  | Mean                   | S.E.  | Mean                     | S.E.  | Mean            | S.E.  |
| Duration from grad. to first job (in months) | 1.772   | 0.051 | 1.983                  | 0.075 | 2.166                    | 0.061 | 2.764           | 0.072 |
| <i>Individual characteristics</i>            |         |       |                        |       |                          |       |                 |       |
| Women  | 0.549   | 0.008 | 0.520                  | 0.011 | 0.475                    | 0.009 | 0.539           | 0.009 |
| Age (in years)                               | 27.15   | 0.041 | 26.03                  | 0.046 | 27.06                    | 0.043 | 26.17           | 0.041 |
| Has moved since graduation                   | 0.488   | 0.008 | 0.472                  | 0.011 | 0.542                    | 0.009 | 0.460           | 0.009 |
| <i>Cohabitation status</i>                   |         |       |                        |       |                          |       |                 |       |
| Living alone                                 | 0.412   | 0.008 | 0.344                  | 0.010 | 0.451                    | 0.009 | 0.341           | 0.008 |
| Living with a partner                        | 0.278   | 0.007 | 0.233                  | 0.009 | 0.258                    | 0.008 | 0.209           | 0.007 |
| Living with their parents                    | 0.309   | 0.007 | 0.422                  | 0.011 | 0.289                    | 0.008 | 0.448           | 0.009 |
| <i>Past variables</i>                        |         |       |                        |       |                          |       |                 |       |
| Job in parallel of studies                   | 0.337   | 0.007 | 0.297                  | 0.010 | 0.346                    | 0.008 | 0.367           | 0.008 |
| Higher education scholarship                 | 0.289   | 0.007 | 0.296                  | 0.010 | 0.318                    | 0.008 | 0.386           | 0.009 |
| Repeated a year                              | 0.060   | 0.004 | 0.072                  | 0.006 | 0.045                    | 0.004 | 0.083           | 0.005 |
| High school with honours                     | 0.373   | 0.008 | 0.257                  | 0.010 | 0.410                    | 0.009 | 0.254           | 0.008 |
| <i>Educational level</i>                     |         |       |                        |       |                          |       |                 |       |
| 2-year degree                                | 0.186   | 0.006 | 0.209                  | 0.009 | 0.200                    | 0.007 | 0.237           | 0.007 |
| 3- or 4- year degree                         | 0.217   | 0.007 | 0.383                  | 0.011 | 0.104                    | 0.005 | 0.403           | 0.009 |
| 5-year degree or more                        | 0.597   | 0.008 | 0.408                  | 0.011 | 0.696                    | 0.008 | 0.360           | 0.008 |
| <i>Field of study</i>                        |         |       |                        |       |                          |       |                 |       |
| Humanities                                   | 0.050   | 0.003 | 0.036                  | 0.004 | 0.156                    | 0.006 | 0.172           | 0.007 |
| Economics, business and law                  | 0.313   | 0.007 | 0.512                  | 0.011 | 0.220                    | 0.007 | 0.235           | 0.007 |
| Science                                      | 0.068   | 0.004 | 0.036                  | 0.004 | 0.097                    | 0.005 | 0.081           | 0.005 |
| Industry                                     | 0.107   | 0.005 | 0.156                  | 0.008 | 0.295                    | 0.008 | 0.215           | 0.007 |
| Agriculture                                  | 0.023   | 0.002 | 0.065                  | 0.005 | 0.045                    | 0.003 | 0.056           | 0.004 |
| Health                                       | 0.215   | 0.006 | 0.024                  | 0.003 | 0.035                    | 0.003 | 0.029           | 0.003 |
| Services                                     | 0.224   | 0.007 | 0.172                  | 0.008 | 0.152                    | 0.006 | 0.212           | 0.007 |
| Observations                                 | 3,999   |       | 2,072                  |       | 3,344                    |       | 3,266           |       |

Source: Author's calculation based on *Génération 2017* data.

Statistics about past variables show interesting patterns. Vertical mismatch seem to be associated with slightly lower abilities (more likely to have repeated a year in primary or secondary school and less likely to have graduated high school with honours), while we observe the opposite for horizontal mismatch. In this regard, individuals mismatched in both dimensions appear to be closer to overeducated workers. Fully mismatched individuals are also more likely to have benefited from a scholarship during higher education, which illustrate a less favourable family background.

In terms of educational level, 2-year degrees account for around 20% of the workers in all mismatch statuses. We observe stronger disparities for longer tertiary education. 3- and 4- year degrees are overrepresented (relative to 5-year and longer

degrees) among overeducated workers. This might be related to the purpose of these levels in the French higher education system, where they are intermediate steps to longer programs rather than degrees giving direct access to the labour market. In line with Goux and Maurin (1994), that could indicate lower returns and a lower value of these educational credentials on the labour market. By contrast, 5-year degrees are overrepresented among horizontally mismatched workers, which might indicate a transferability of the skills acquired through this type of education between sectors.

There are large disparities between mismatches in terms of field of study. Unsurprisingly, graduating in Health appears to be a protection against all kinds of mismatch. On the contrary, graduates in Agriculture are overrepresented among mismatched workers. Graduates in Humanities, Science and Industry are overrepresented among horizontally mismatched workers (whether they are overeducated or not), while Economics, business and law graduates are strongly overrepresented among vertically mismatched individuals but underrepresented in horizontal mismatch.

## 4.3 Empirical strategy

### 4.3.1 Treatment framework

The aim of this paper is to study the impact of being vertically and/or horizontally mismatched in the first job after graduation on future labour market positions. To this end, I work with a multiple treatment framework (Lechner, 2001). In their first job, individuals can be in five mutually exclusive states: (1) unemployed, (2) in a matched job, (3) only overeducated, (4) only horizontally mismatched and (5) both vertically and horizontally mismatched. I aim to compare the impact of different mismatches on future labour market outcomes. I consider the pairwise comparisons of three mismatch states (only overeducated, only horizontally mismatched and both vertically and horizontally mismatched) with longer unemployment<sup>9</sup>.

All individuals in my sample do not enter their first job at the same time, and thus are not treated at the same time. As underlined by Meroni and Vera-Toscano (2017), individuals who prefer to remain unemployed at a given point in time may accept to be mismatched later in order to get a job. Therefore, I use a dynamic treatment framework, that allows me to estimate heterogeneous treatment effects

---

<sup>9</sup>As a benchmark, I also estimate the effect of being mismatched in any dimension (without distinction between vertical, horizontal and full mismatch) on future labour market outcomes.

depending on the time of treatment (Sianesi, 2004, 2008), here the duration since graduation. Individuals enter the sample at the time of graduation and remain in it until they are treated, *i.e.* they take their first job. I consider three different times of treatment<sup>10</sup>:  $T_0$  when individuals take their first job less than one month after graduation;  $T_1$  when they take it between one and six months after graduation;  $T_2$  between six and eighteen months after graduation<sup>11</sup>.

In each time window  $t = \{T_0; T_1; T_2\}$ , the effect of receiving treatment  $k$  (mismatch) rather than treatment  $k'$  (unemployment) can be rewritten as:

$$E(Y_t^k - Y_t^{k'} | D = k) = E(Y_t^k | D = k) - E(Y_t^{k'} | D = k) \quad (4.1)$$

with  $D$  the indicator of treatment. However,  $E(Y_t^{k'} | D = k)$  is not observed in the data, as it is the outcome that would have been experienced by individuals receiving treatment  $k'$  if they had received treatment  $k$ . I use propensity score matching (Abadie & Imbens, 2016; Rosenbaum & Rubin, 1983) to compare individuals with the same duration since graduation, with the only difference between them being one received treatment  $k$  and the other received treatment  $k'$ .

Table 4.2: Sample size by treatment status and time of treatment

| <b>Treatment status</b>      | $T_0$  | $T_1$ | $T_2$ |
|------------------------------|--------|-------|-------|
| Unemployed                   | 5,968  | 1,814 | 375   |
| Employed                     |        |       |       |
| Matched                      | 2,426  | 1,234 | 339   |
| Only overeducated            | 1,217  | 645   | 210   |
| Only horizontally mismatched | 1,821  | 1,144 | 379   |
| Both mismatches              | 1,624  | 1,131 | 511   |
| Observations                 | 13,056 | 5,968 | 1,814 |

Source: Author's calculations based on *Génération 2017* data.

Practically, I redefine the treatment and control groups and estimate a specific matching estimator for each time window  $t = \{T_0; T_1; T_2\}$ . In each time window, the control group is composed of the individuals not yet treated, *i.e.* still unemployed at time  $t$ , and the treatment group of individuals receiving treatment  $k$  at time  $t$ . After being treated, individuals leave the sample: individuals treated (entering their first job) at time  $T_0$  are matched to individuals receiving the treatment at time  $T_1$

<sup>10</sup>I am not able to use monthly information to run the analysis due to sample size issues.

<sup>11</sup>Most individuals (97%) in the sample found their job in the first 18 months after graduation. I do not consider individuals taking a first job after this duration because the sample size become too small to provide reliable treated and control groups.



or later, and are excluded of the sample for subsequent dates; individuals treated at time  $T_1$  are matched to individuals receiving the treatment at time  $T_2$  or later, and are excluded of the sample for subsequent dates; and so on. The share of individuals in each treatment at each date is shown in Table 4.2.

### 4.3.2 Conditional independence assumption (CIA)

This dynamic propensity score matching framework relies on a conditional independence assumption (CIA), *i.e.* that conditional on the independent variables  $X$  and for a given unemployment duration since graduation, both  $k$ -treated and  $k'$ -treated individuals would end up with the same outcome if they received the same treatment. Therefore, the vector of characteristics  $X$  must be as complete as possible. The richness of the data allows me to take into account a large set of regressors including individual socio-demographic characteristics and educational variables<sup>12</sup>.

As for the individual characteristics, in addition to gender and age (observed at the time of survey), I take into account their cohabitation status at the time of treatment. Living with a partner may imply joined decisions (about the geographical location, the type of job, etc.) in comparison with individuals living alone, which can affect the labour market positions. Living with parents might entail lower financial pressure, and thus lower incentives to find a (matched) job. I also include information on both the educational level (three categories) and the field of study (seven categories) of the individuals (Ortiz & Kucel, 2008; Rossen et al., 2019).

The credibility of the CIA and the quality of the matching are improved by conditioning on past variables (Caliendo et al., 2017). Therefore, I include information on past education and background. I include a dummy variable equal to one if graduates had a student job in parallel of their studies, for at least three reasons. Firstly, it can reveal the motivation of the graduate; secondly, the professional experience acquired in the job could facilitate the access to (matched) jobs after graduation; and thirdly, the necessity of taking a job while studying could illustrate a less advantageous family background. I also take into account the benefit of a higher education scholarship based on social criteria to control for family background. Besides, identification issues may arise if individuals who accept a mismatched job more quickly after graduation act this way because of lower abilities. Because the *Génération*

---

<sup>12</sup>Characteristics of the job are likely to affect the decision to take a mismatched job. However, since these information can only be observed for those who are currently employed, such variables cannot be used when comparing mismatch to unemployment.

2017 dataset do not provide a measurement of skills, I proxy individual abilities using information on past education (Barone & Ortiz, 2011), in the form of two dummy variables: the first equal to one if the individual repeated a school year in primary or secondary school; the second equal to one if the individual graduated high school (*i.e.* obtained her or his baccalaureate) with honours or highest honours.

The risk of mismatch is also related to geographical variables. At the individual level, residential migration can be a strategy to get closer to job opportunities (Langella & Manning, 2022; Rupert & Wasmer, 2012) and to avoid overeducation (Büchel & van Ham, 2003; Iammarino & Marinelli, 2015), while it tends to increase horizontal mismatch (Hensen et al., 2009). I include a binary variable equal to one if the employment area<sup>13</sup> of residency at the time of the first job is different from the one at the time graduation. The regional labour market context also impacts the risk of overeducation (Büchel & van Ham, 2003; Tselios, 2013). To control for it, I include fixed effects for the region of residency at the time of graduation.

Finally, Croce and Ghignoni (2012) showed that overeducation is impacted by macroeconomic fluctuations. In particular, they found that graduating during an economic crisis leads to a higher risk of mismatch. Even though the time window of graduation in my sample is relatively short (around one year), I include fixed effects for the month of graduation to ensure to capture these potential variations of the economic context. Moreover, these fixed effects can capture the seasonal variations of economic activity, which can also influence the entry on the labour market.

### 4.3.3 Matching specification and estimation

The main specification is estimated using Kernel matching with a bandwidth of 0.025 and imposing exact matching on gender, educational level and field of study. This matching estimator provides a good balance of the variables between treatment and control groups, for all the different treatments and all time windows (see Appendix 4B). To check for robustness, other matching strategies are mobilised (Kernel matching with a bandwidth of 0.05, 5 nearest neighbours matching and inverse probability weighting), which provide similar results (see Appendix 4C), although the quality of the matching is lower.

I estimate the average treatment effects on the treated (ATT) of the three dif-

---

<sup>13</sup>French employment areas are defined by Insee as areas where most workers both live and work, and where firms can find most of their workforce, based on commuting flows between municipalities. They are especially designed to study local labour markets.

ferent treatments on four different outcomes. The first outcome, estimated on the whole sample, is (1) the probability of being employed at the time of the survey. The other outcomes are estimated only for individuals who are employed at the time of the survey. I estimate the average effect on (2) the probability of being in a permanent job, (3) the probability of working full-time, (4) earnings at the time of the survey<sup>14</sup>. To explore the “stepping stone” effect of mismatch, I also estimate the effects on the probabilities of being in each of the matching statuses at the time of the survey, conditionally on being employed.

## 4.4 Results

### 4.4.1 Effects on employment and job characteristics

Table 4.3 displays the average treatment effects on the treated on the labour market outcomes of young graduates, observed around three years after graduation. Column 1 shows the ATT on the probability of being employed three years after graduation. For  $T_0$  and  $T_1$ , *i.e.* for first jobs taken within six months after graduation, I find positive effects of mismatch: taking a mismatch job rather than staying unemployed reduces the probability of being unemployed three years after of 5 to 6 percentage points (between 25% and 30%). The effect is higher for horizontal mismatch (7,5 pp.), and lower for full mismatch (3 to 5 pp.). Moreover, the effect tends to be slightly higher for jobs taken one to six months after graduation than for job taken immediately. Taking a longer time to find a job might illustrate a higher selectivity of job offers and the will to find a better job which allows accumulation of more and/or better professional experience, consistently with the theory of turnover (Jovanovic, 1979). In  $T_2$ , however, the effect is no more significant, suggesting that mismatch offers a protection against future unemployment only if the job is taken early after graduation. It is even worse for full mismatch, for which a job taken too late leads to a lower probability of being employed after three years.

Effects on the probability of having a permanent job (column 2) and the probability of working full-time (column 3) three years after graduation are qualitatively similar to the ones observed for employment. When taken early (in the six months following graduation), mismatched jobs lead to less precarious job positions after

---

<sup>14</sup>Monthly net wages for salaried workers, monthly net income for self-employed individuals.

Table 4.3: Treatment effects on labour market outcomes

| Treatment                | Time of treatment | Outcomes (at the time of survey) |                     |                     |                       |
|--------------------------|-------------------|----------------------------------|---------------------|---------------------|-----------------------|
|                          |                   | (1) Employed                     | (2) Permanent job   | (3) Work full-time  | (4) Earnings          |
| Any mismatch             | $T_0$             | 0.052***<br>(0.008)              | 0.084***<br>(0.009) | 0.036***<br>(0.008) | 138.72***<br>(16.28)  |
|                          | $T_1$             | 0.059***<br>(0.014)              | 0.086***<br>(0.018) | 0.077***<br>(0.015) | 153.23***<br>(24.81)  |
|                          | $T_2$             | -0.025<br>(0.017)                | -0.045**<br>(0.023) | -0.034*<br>(0.020)  | -85.90***<br>(31.29)  |
| Only vertical mismatch   | $T_0$             | 0.050***<br>(0.013)              | 0.099***<br>(0.015) | 0.034**<br>(0.014)  | 88.24***<br>(24.54)   |
|                          | $T_1$             | 0.056**<br>(0.022)               | 0.117***<br>(0.027) | 0.104***<br>(0.025) | 144.72***<br>(39.33)  |
|                          | $T_2$             | 0.009<br>(0.031)                 | 0.014<br>(0.041)    | -0.060<br>(0.037)   | -199.00***<br>(52.26) |
| Only horizontal mismatch | $T_0$             | 0.075***<br>(0.009)              | 0.093***<br>(0.011) | 0.043***<br>(0.010) | 250.44***<br>(25.40)  |
|                          | $T_1$             | 0.075***<br>(0.016)              | 0.077***<br>(0.021) | 0.067***<br>(0.018) | 239.56***<br>(36.40)  |
|                          | $T_2$             | 0.020<br>(0.024)                 | -0.065*<br>(0.033)  | -0.017<br>(0.027)   | 39.27<br>(50.61)      |
| Both mismatches          | $T_0$             | 0.031***<br>(0.012)              | 0.059***<br>(0.015) | 0.020<br>(0.013)    | 7.35<br>(18.79)       |
|                          | $T_1$             | 0.048**<br>(0.019)               | 0.079***<br>(0.025) | 0.051**<br>(0.022)  | 80.77***<br>(29.50)   |
|                          | $T_2$             | -0.078***<br>(0.026)             | -0.053<br>(0.034)   | -0.070**<br>(0.029) | -160.85***<br>(37.39) |

Note: This table displays the ATT on the different outcomes, observed at the time of the survey. Estimations are realised using propensity score, using Kernel matching with a bandwidth of 0.025, exact matching on gender, educational level and field of study, and imposing common support. The effect on the probability of employment (1) is calculated on the whole sample. Effects on outcomes (2) to (4) are calculated considering only individuals employed at the time of the survey.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculations based on *Génération 2017* data.

three years, but when taken later, the observed effect is negative, although barely significant in most cases. Effects are higher for vertically mismatched workers than for horizontally mismatched ones. This is somehow consistent with results by Wolbers (2003), who find that horizontally mismatched workers are also more likely to engage in on-the-job search in order to find a more adequate position (Jovanovic, 1979), which can imply a preference for temporary contracts over permanent ones, at least until a matched job is found.

Overall, accepting to be mismatched in the first job rather than waiting longer in unemployment appears to have positive effects on future employment and job characteristics, but only if the first job is taken early enough after graduation. In any case, being mismatched both vertically and horizontally seems to bring limited benefits compared to longer unemployment.

#### 4.4.2 Effects on earnings

Column 4 in Table 4.3 shows the ATT on monthly earnings of individuals three years after their graduation. Because they are monthly rather than hourly earnings, they do not take into account the full-time or part-time character of the job. Therefore,

I will remain careful in interpreting the effects on earnings<sup>15</sup>.

There is a wage premium for accepting a mismatched job early, while taking it later (six to eighteen months after graduation) has a negative effect on future earnings. This can illustrate that taking a job earlier allows to accumulate more work experience, which is reflected on the earnings observed later. Another explanation is that workers taking a job later may suffer from the scarring effects of both unemployment (Van Belle et al., 2018) and overeducation (Baert & Verhaest, 2019), and are thus perceived as less productive. Similar to what is observed for employment, the premium is higher when the first job is taken after a short job search duration rather than immediately, at least for overeducated workers.

Horizontal mismatch seem more beneficial for earnings than other types of mismatch, in comparison with longer unemployment. The effect is higher and stable for  $T_0$  and  $T_1$ , around 250€ against 150€ at most for other mismatches (7% to 12% higher than individuals who stay longer in unemployment). For  $T_2$ , while other mismatches have significant negative effects, horizontal mismatch is not associated with either a significant wage penalty or premium. This kind of mismatch do not seem to affect earnings trajectories at the beginning of the career (in addition, Table ?? in Appendix 4D shows only a small difference in earnings between matched and horizontally mismatched workers in the current job). This suggests that some skills may not be specific to a particular field and thus may be transferable to others, limiting the lack of productivity of horizontally mismatched workers and resulting in lower wage penalties than other types of mismatch (Nordin et al., 2010; Robst, 2007b).

### 4.4.3 Effects on matching status

In order to explore whether mismatched jobs at the beginning of the career can be a “stepping stone” towards better labour market positions or a “trap” where young workers are stuck for a long time, the effects of the different treatments on the matching status in the current job are reported in Table 4.4. I find a negative and significant effect on the probability to be in a matched job three years after graduation, regardless of the time of treatment and of the type of mismatch (column 4). This adds to previous evidence of the detrimental effects of overeducation on

---

<sup>15</sup>It has to be noted that more than 80% of the workers have a full-time job at the time of the survey, with small differences in this proportion depending on their matching status (see Table 4.11 in Appendix 4D). That is why I attempt to explore the effects on earnings, even though their measure is imperfect.

the following of the career (Baert et al., 2013; Büchel & Mertens, 2004; Meroni & Vera-Toscano, 2017; Scherer, 2004) and confirms that a similar effect occurs for horizontal mismatch. The detrimental effect is higher if the first job is taken six to eighteen months after graduation than earlier. Similar to the results in Table 4.3, taking a job too quickly after graduation is not a good strategy either, as the negative effect is higher for  $T_0$  than for  $T_1$ .

The negative effects of horizontal mismatch are higher than the ones of vertical or full mismatch. However, given the results obtained for earnings, individuals working in the “wrong” field of study may have less incentive to search for a matched job than other mismatched workers. Individuals mismatched in both dimensions in their first job suffer stronger penalties on the probability to be in a matched job than only overeducated ones (especially when the first job is taken early), although they do not have the wage incentives of only horizontally mismatched workers. This supports the idea that the individuals occupying jobs the furthest from their education suffer from the higher detrimental effects.

Table 4.4: Treatment effects on matching status

| Treatment                | Time of treatment | Outcomes (at the time of survey) |                      |                      |                      |
|--------------------------|-------------------|----------------------------------|----------------------|----------------------|----------------------|
|                          |                   | (1) Ver. mismatch                | (2) Hor. mismatch    | (3) Full mismatch    | (4) Match            |
| Any mismatch             | $T_0$             | 0.043***<br>(0.010)              | 0.098***<br>(0.011)  | 0.019*<br>(0.011)    | -0.159***<br>(0.009) |
|                          | $T_1$             | 0.031**<br>(0.015)               | 0.112***<br>(0.019)  | -0.013<br>(0.018)    | -0.130***<br>(0.015) |
|                          | $T_2$             | 0.044**<br>(0.018)               | 0.091***<br>(0.022)  | 0.080***<br>(0.023)  | -0.215***<br>(0.018) |
| Only vertical mismatch   | $T_0$             | 0.443***<br>(0.020)              | -0.160***<br>(0.011) | -0.192***<br>(0.016) | -0.091***<br>(0.014) |
|                          | $T_1$             | 0.413***<br>(0.031)              | -0.167***<br>(0.024) | -0.204***<br>(0.026) | -0.042*<br>(0.023)   |
|                          | $T_2$             | 0.472***<br>(0.042)              | -0.150***<br>(0.028) | -0.120***<br>(0.035) | -0.203***<br>(0.033) |
| Only horizontal mismatch | $T_0$             | -0.092***<br>(0.006)             | 0.447***<br>(0.013)  | -0.130***<br>(0.009) | -0.225***<br>(0.011) |
|                          | $T_1$             | -0.088***<br>(0.012)             | 0.467***<br>(0.023)  | -0.183***<br>(0.017) | -0.197***<br>(0.020) |
|                          | $T_2$             | -0.090***<br>(0.015)             | 0.446***<br>(0.030)  | -0.125***<br>(0.022) | -0.231***<br>(0.024) |
| Both mismatches          | $T_0$             | -0.117***<br>(0.012)             | -0.114***<br>(0.013) | 0.376***<br>(0.017)  | -0.145***<br>(0.011) |
|                          | $T_1$             | -0.075***<br>(0.018)             | -0.135***<br>(0.023) | 0.301***<br>(0.027)  | -0.092***<br>(0.018) |
|                          | $T_2$             | -0.044*<br>(0.024)               | -0.131***<br>(0.026) | 0.381***<br>(0.033)  | -0.205***<br>(0.025) |

Note: This table displays the ATT on mismatch status, observed at the time of the survey. Estimations are realised using propensity score, using Kernel matching with a bandwidth of 0.025, exact matching on gender, educational level and field of study, and imposing common support. Effects on mismatch statuses are calculated considering only individuals employed at the time of the survey.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author’s calculations based on *Génération 2017* data.

In addition to the negative effect on the probability of being in a matched job three years after graduation, I find strong positive effects on the probability to be

in the same mismatch in the current job and in the first job. One can argue that, because the time of observation is not that distant from graduation (around three years), individuals could still be in their first job at the time of the survey. If that may explain part of the trap effect, it must be noted that more than half of the sample changed of job between the first and current position. This share is even higher for vertical mismatch (63%) and full mismatch (70%). This suggests that some individuals are “trapped” in mismatched labour market positions (Baert et al., 2013) even when they are looking for other jobs.

The magnitude of this effect does not vary much with the timing of the first job. In comparison with individuals remaining longer in unemployment, being overeducated in the first job increases by more than 40 percentage points the risk of still being overeducated in the current job. I find similar results for horizontal mismatch, while the ATT of being fully mismatched in the first job is lower (between 0.30 and 0.40).

The negative effects of full mismatch in the first job on the probability of the other mismatches in the current job are also smaller compared to the other effects of one mismatch on another. This may indicate that individuals mismatched in both dimensions in their first job would accept a job for which they are only overeducated or only horizontally mismatch to get closer to their education, rather than looking only for a completely matched job.

Following the same idea, I find that the negative effect of being horizontally mismatched on the probability of overeducation is lower than the opposite. This suggests that some individuals who obtain a first job in the “wrong” field accept a job for which they are overeducated in order to transit back to their field of study, while the opposite is less likely. This may seem surprising because horizontal mismatch tend to be related to smaller penalties on job satisfaction (Allen & van der Velden, 2001; Somers et al., 2019) and earnings than vertical mismatch. However, it is also easier for workers to fit a lower skilled job in their field of study than a higher skilled job in another field. Moreover, some individuals’ academic choices are motivated by educational consumption more than investment (Sellami et al., 2020), which may result in a preference for the subject of the work over the returns of education on the labour market.

#### 4.4.4 Heterogeneity analysis

Tables 4.5 and 4.6 display the heterogeneous effects of mismatch<sup>16</sup> in the first job across genders and educational levels, respectively on labour market outcomes and matching status three years after graduation. These results confirm that taking a mismatched job too late (more than six months) after graduation does not lead to better outcomes than longer unemployment, regardless of the demographic group under consideration.

Table 4.5: Treatment effects on labour market outcomes - Any mismatch

| Treatment                 | Time of treatment | Outcomes (at the time of survey) |                     |                      |                       |
|---------------------------|-------------------|----------------------------------|---------------------|----------------------|-----------------------|
|                           |                   | (1) Employed                     | (2) Permanent job   | (3) Work full-time   | (4) Earnings          |
| Women                     | $T_0$             | 0.040***<br>(0.011)              | 0.088***<br>(0.013) | 0.043***<br>(0.012)  | 109.43***<br>(21.54)  |
|                           | $T_1$             | 0.052***<br>(0.018)              | 0.115***<br>(0.025) | 0.107***<br>(0.023)  | 152.45***<br>(33.61)  |
|                           | $T_2$             | -0.036<br>(0.026)                | -0.042<br>(0.035)   | -0.024<br>(0.029)    | -144.49***<br>(42.90) |
| Men                       | $T_0$             | 0.059***<br>(0.011)              | 0.077***<br>(0.014) | 0.026**<br>(0.011)   | 160.71***<br>(24.75)  |
|                           | $T_1$             | 0.046**<br>(0.019)               | 0.039*<br>(0.023)   | 0.025<br>(0.019)     | 153.58***<br>(38.06)  |
|                           | $T_2$             | 0.007<br>(0.027)                 | -0.052*<br>(0.032)  | -0.052*<br>(0.028)   | -52.33<br>(48.43)     |
| 2-year graduates          | $T_0$             | 0.053***<br>(0.007)              | 0.094***<br>(0.023) | 0.045**<br>(0.021)   | 100.77***<br>(23.72)  |
|                           | $T_1$             | 0.064***<br>(0.013)              | 0.073*<br>(0.042)   | 0.071*<br>(0.040)    | 171.73***<br>(45.41)  |
|                           | $T_2$             | -0.027<br>(0.017)                | -0.066<br>(0.054)   | -0.176***<br>(0.049) | -141.68***<br>(49.87) |
| 3- and 4-year graduates   | $T_0$             | 0.071***<br>(0.016)              | 0.086***<br>(0.021) | 0.031<br>(0.019)     | 106.74<br>(24.72)     |
|                           | $T_1$             | 0.047<br>(0.030)                 | 0.135***<br>(0.043) | 0.139***<br>(0.043)  | 182.18***<br>(47.60)  |
|                           | $T_2$             | -0.030<br>(0.042)                | -0.105*<br>(0.057)  | -0.072<br>(0.052)    | -38.48<br>(61.04)     |
| 5-year and more graduates | $T_0$             | 0.041***<br>(0.009)              | 0.076***<br>(0.011) | 0.032***<br>(0.010)  | 146.70***<br>(23.79)  |
|                           | $T_1$             | 0.044**<br>(0.017)               | 0.067***<br>(0.022) | 0.065***<br>(0.018)  | 168.65***<br>(34.02)  |
|                           | $T_2$             | 0.005<br>(0.024)                 | -0.054*<br>(0.029)  | -0.007<br>(0.021)    | -94.07**<br>(41.91)   |

Note: This table displays the ATT on the different outcomes, observed at the time of the survey. Estimations are realised using propensity score, using Kernel matching with a bandwidth of 0.025, exact matching on gender, educational level and field of study, and imposing common support. The effect on the probability of employment (1) is calculated on the whole sample. Effects on outcomes (2) to (4) are calculated considering only individuals employed at the time of the survey.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculations based on *Génération 2017* data.

Results by gender show that women benefit more from mismatch in terms of protection against precarious employment (part-time and temporary jobs) than men. By contrast, accepting a mismatched job rather than staying longer in employment

<sup>16</sup>Results of the heterogeneity analysis are presented only for the treatment “Any mismatch”. Overall, results and interpretations for the other treatments appear to be similar, but the sample size for each group is too small to provide reliable estimates when distinguishing by type of mismatch. Those results can be provided upon request.



leads to a higher earnings increase for men than for women. Mismatched jobs taken between six and eighteen months are even associated to negative effects on wages for women (the coefficient is not significant for men). Effects on the probability of being in a matched job after three years (Table 4.6, column 4) are similar for both genders. The effect on the risk of being horizontally mismatch (Table 4.6, column 2) is slightly higher for women than for men, and I find a significant effect on the probability of vertical mismatch (Table 4.6, column 1) only for men. This might indicate a selection of men in overeducated jobs and of women in jobs that do not match their field of study, which may also reflect a selection into the fields of education. Indeed, women tend to be more present in Humanities, which is overrepresented in horizontal mismatch (see Table 4.1).

Table 4.6: Treatment effects on matching status - Any mismatch

| Treatment                 | Time of treatment | Outcomes (at the time of survey) |                     |                     |                      |
|---------------------------|-------------------|----------------------------------|---------------------|---------------------|----------------------|
|                           |                   | (1) Ver. mismatch                | (2) Hor. mismatch   | (3) Full mismatch   | (4) Match            |
| Women                     | $T_0$             | 0.023<br>(0.015)                 | 0.113***<br>(0.015) | 0.024<br>(0.015)    | -0.159***<br>(0.013) |
|                           | $T_1$             | 0.016<br>(0.024)                 | 0.124***<br>(0.025) | -0.013<br>(0.025)   | -0.127***<br>(0.021) |
|                           | $T_2$             | 0.055*<br>(0.028)                | 0.062**<br>(0.031)  | 0.092***<br>(0.034) | -0.209***<br>(0.029) |
| Men                       | $T_0$             | 0.070***<br>(0.014)              | 0.094***<br>(0.016) | -0.001<br>(0.017)   | -0.163***<br>(0.012) |
|                           | $T_1$             | 0.047**<br>(0.019)               | 0.105***<br>(0.028) | -0.002<br>(0.027)   | -0.149***<br>(0.024) |
|                           | $T_2$             | 0.034<br>(0.026)                 | 0.114***<br>(0.037) | 0.074**<br>(0.034)  | -0.222***<br>(0.026) |
| 2-year graduates          | $T_0$             | 0.088***<br>(0.021)              | 0.023<br>(0.026)    | 0.063***<br>(0.024) | -0.174***<br>(0.021) |
|                           | $T_1$             | 0.037<br>(0.033)                 | 0.044<br>(0.045)    | -0.009<br>(0.042)   | -0.072**<br>(0.034)  |
|                           | $T_2$             | 0.025<br>(0.039)                 | 0.011<br>(0.056)    | 0.142***<br>(0.054) | -0.178***<br>(0.050) |
| 3- and 4-year graduates   | $T_0$             | 0.040<br>(0.024)                 | 0.007<br>(0.016)    | 0.013<br>(0.026)    | -0.060***<br>(0.013) |
|                           | $T_1$             | 0.034<br>(0.045)                 | 0.015<br>(0.031)    | -0.012<br>(0.047)   | -0.038*<br>(0.021)   |
|                           | $T_2$             | 0.003<br>(0.058)                 | 0.074<br>(0.047)    | 0.020<br>(0.064)    | -0.097***<br>(0.027) |
| 5-year and more graduates | $T_0$             | 0.030***<br>(0.011)              | 0.192***<br>(0.016) | -0.005<br>(0.012)   | -0.217***<br>(0.013) |
|                           | $T_1$             | 0.009<br>(0.019)                 | 0.177***<br>(0.026) | -0.006<br>(0.020)   | -0.180***<br>(0.021) |
|                           | $T_2$             | 0.057**<br>(0.024)               | 0.104***<br>(0.033) | 0.085***<br>(0.027) | -0.247***<br>(0.025) |

Note: This table displays the ATT on mismatch status, observed at the time of the survey. Estimations are realised using propensity score, using Kernel matching with a bandwidth of 0.025, exact matching on gender, educational level and field of study, and imposing common support. Effects on mismatch statuses are calculated considering only individuals employed at the time of the survey.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculations based on *Génération 2017* data.

There are larger disparities depending on the educational level than depending on gender. For intermediate graduates (3- and 4-year degrees), accepting a mismatched job does not provide much benefits on labour market outcomes in comparison with

longer unemployment. The effect on the probability of employment (Table 4.5, column 1) is significant only if the first job is taken right after graduation, and I find positive effects of earnings and the probability of working full-time only for jobs taken between one and six months after graduation. Mismatch however seems to be associated to a greater protection against temporary employment (Table 4.5, column 2). For both shorter and longer degrees, early mismatched jobs are associated with positive effects on outcomes in comparison with longer unemployment. Effects are relatively higher on employment outcomes and lower on earnings for 2-year graduates, while I find the opposite for 5-year graduates. Overall, this suggests a lower value of the intermediate university degrees on the labour market.

I find negative effects on future match for all educational levels (Table 4.6, column 4), but their magnitude varies considerably. The detrimental effects are smaller for intermediate graduates, moderate for short-cycle graduates and the highest for long-cycle graduates. For 3- and 4-year graduates, mismatch in the first job does not seem to increase the risk of a particular type of mismatch later on. To some extent, 2-year graduates appear to be more at risk to be fully mismatched after three years if they accept to be mismatched in the first job. The highest detrimental effects for 5-year graduates are observed on the probability of horizontal mismatch.

## 4.5 Conclusion

The aim of this work was to study the impact of educational mismatch in the first job after graduation on the following of the career. I especially wanted to explore whether there is a “stepping stone” effect of mismatch at the beginning of the career or not. To this end, I use a dynamic treatment framework, that allows me to explore the effects of several types of mismatch (overeducation, horizontal mismatch and full mismatches) on the labour market position and outcomes observed three years after graduation, and to take into account the differences in the timing of the first job.

I show that mismatch, in any dimension, is not a stepping stone but a trap for young workers at the beginning of their career. Indeed, individuals who accept a first job for which they are mismatched are more likely to still be mismatched three years after graduation than if they had stayed longer unemployed. Moreover, I find only small evidence that workers transit from one type of mismatch to another: most of them are trapped in the same type of mismatch in their first and current

job. The magnitude of these effects do not vary much with the elapsed duration between graduation and the first job.

I find a positive (although small) effect of taking a mismatched job earlier on future employment. Therefore, there might be a trade-off between waiting longer in unemployment to find a matched job but with a higher risk of being unemployed after three years, and taking a mismatched job which protects against future unemployment but leads to be trapped in the mismatch, with the detrimental effects associated. Nevertheless, this effect is significant only in the six months following graduation, meaning that taking a mismatched job too late after graduation do not improve future labour market prospects in comparison with longer unemployment.

While horizontal mismatch appears to have higher benefits in terms of employment and earnings, overeducation provides more protection against precarious employment (temporary and part-time jobs) in comparison with longer unemployment. Full mismatch have very limited benefits on labour market outcomes, and even negative effects if the first job is taken too late. This suggests that, the further their job is from their education, the lower are individuals' returns from education, both in terms of protection against unemployment and in terms of earnings. Therefore, it seems better for young graduates to remain longer in unemployment than to accept a position where they will be mismatched in both dimensions.

The heterogeneity analysis do not show much differences in the trap effect between men and women, with significant negative effects on future match regardless of the timing of the first job. However, results suggest disparities in the type of mismatch: men appear to end up more frequently in overeducated jobs, while women have a higher probability to work outside their field of study. In terms of educational level, long-cycle graduates have the most to lose to accept to be mismatched in their first job, as they suffer the highest detrimental effects on the probability of being matched after three years.

From a public policy point of view, the trap effect of mismatch implies that policies encouraging workers to accept any job (even inconsistent with their education) in order to get out of unemployment faster might have adverse effects, by locking them into mismatched jobs that have detrimental consequences on their future career (Baert et al., 2013; Voßemer & Schuck, 2016). Results suggest that the trap effect is stronger for horizontal mismatch than for overeducation, but horizontal mismatch is also associated with better labour market outcomes. Full mismatch is the worst situation for young graduates entering the labour market, this situa-

tion being barely better than longer unemployment. Given these disparities, further research is needed to understand the mechanisms of the trap effects, which might vary according to the type of mismatch. It would also be interesting to check, with a longer time horizon, whether these effects eventually disappear or whether they persist in the longer term.

## Appendices

### Appendix 4A: Statistical measure of overeducation

Table 4.7: Description of occupational codes - PCS 2020

| <b>Code</b> | <b>Occupations</b>   |
|-------------|--|
| 10          | Farming, forestry, fishing and aquaculture business owners                     |
| 21          | Craftspersons  |
| 22          | Retailers, shopkeepers   |
| 31          | Liberal professions  |
| 33          | Administrative and technical civil service managers                            |
| 34          | Teachers and higher scientific occupations                                     |
| 35          | Information, arts and culture occupations                                      |
| 37          | Administrative and sales managers  |
| 38          | Engineers and technical managers   |
| 42          | Primary school teachers, professional and sports trainers                      |
| 43          | Health and social associate professionals                                      |
| 45          | Intermediate occupations of the public sector                                  |
| 46          | Business and administration associate professionals                            |
| 47          | Technicians  |
| 48          | Foremen  |
| 52          | Public service administrative employees, service agents and health auxiliaries |
| 53          | Police, army, firefighters and private security workers                        |
| 54          | Clerical support workers   |
| 55          | Sales workers  |
| 56          | Personal service workers   |
| 62          | Skilled industrial workers   |
| 63          | Skilled craft workers  |
| 64          | Transport vehicle drivers, delivery drivers,                                   |
| 65          | Equipment drivers, forklift truck operators, warehouse workers                 |
| 67          | Unskilled industrial workers   |
| 68          | Unskilled craft workers  |
| 69          | Agricultural workers   |

Table 4.8: Occupation-degree correspondence grid - Statistical overeducation

|                                   |           |           |           |           |           |           |           |           |           |
|-----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                                   | <b>10</b> | <b>21</b> | <b>22</b> | <b>31</b> | <b>33</b> | <b>34</b> | <b>35</b> | <b>37</b> | <b>38</b> |
| 5-year tertiary degree and higher | OE        | OE        | OE        | N         | N         | N         | N         | N         | N         |
| 3- and 4-year tertiary degree     | OE        | OE        | OE        | N         | N         | N         | N         | N         | N         |
| 2-year tertiary degree            | OE        | OE        | OE        | N         | N         | N         | N         | N         | N         |
| General high school               | OE        | OE        | OE        | N         | N         | N         | N         | N         | N         |
| Vocational high school            | N         | N         | OE        | N         | N         | N         | N         | N         | N         |
| Short vocational high school      | N         | N         | N         | N         | N         | N         | N         | N         | N         |

|                                   |           |           |           |           |           |           |           |           |           |
|-----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                                   | <b>42</b> | <b>43</b> | <b>45</b> | <b>46</b> | <b>47</b> | <b>48</b> | <b>52</b> | <b>53</b> | <b>54</b> |
| 5-year tertiary degree and higher | N         | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| 3- and 4-year tertiary degree     | N         | N         | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| 2-year tertiary degree            | N         | N         | N         | N         | N         | OE        | OE        | OE        | N         |
| General high school               | N         | N         | N         | N         | N         | OE        | OE        | OE        | N         |
| Vocational high school            | N         | N         | N         | N         | N         | N         | OE        | N         | N         |
| Short vocational high school      | N         | N         | N         | N         | N         | N         | N         | N         | N         |

|                                   |           |           |           |           |           |           |           |           |           |
|-----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                                   | <b>55</b> | <b>56</b> | <b>62</b> | <b>63</b> | <b>64</b> | <b>65</b> | <b>67</b> | <b>68</b> | <b>69</b> |
| 5-year tertiary degree and higher | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| 3- and 4-year tertiary degree     | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| 2-year tertiary degree            | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| General high school               | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| Vocational high school            | N         | OE        | OE        | OE        | OE        | OE        | OE        | OE        | OE        |
| Short vocational high school      | N         | N         | N         | N         | N         | N         | N         | N         | N         |

OE stands for overeducation, N for normal. Undereducation is labelled here as normal.

## Appendix 4B: Quality of the matching

Figure 4.1: Distribution of the propensity score on the matched and unmatched sample -  $T_0$

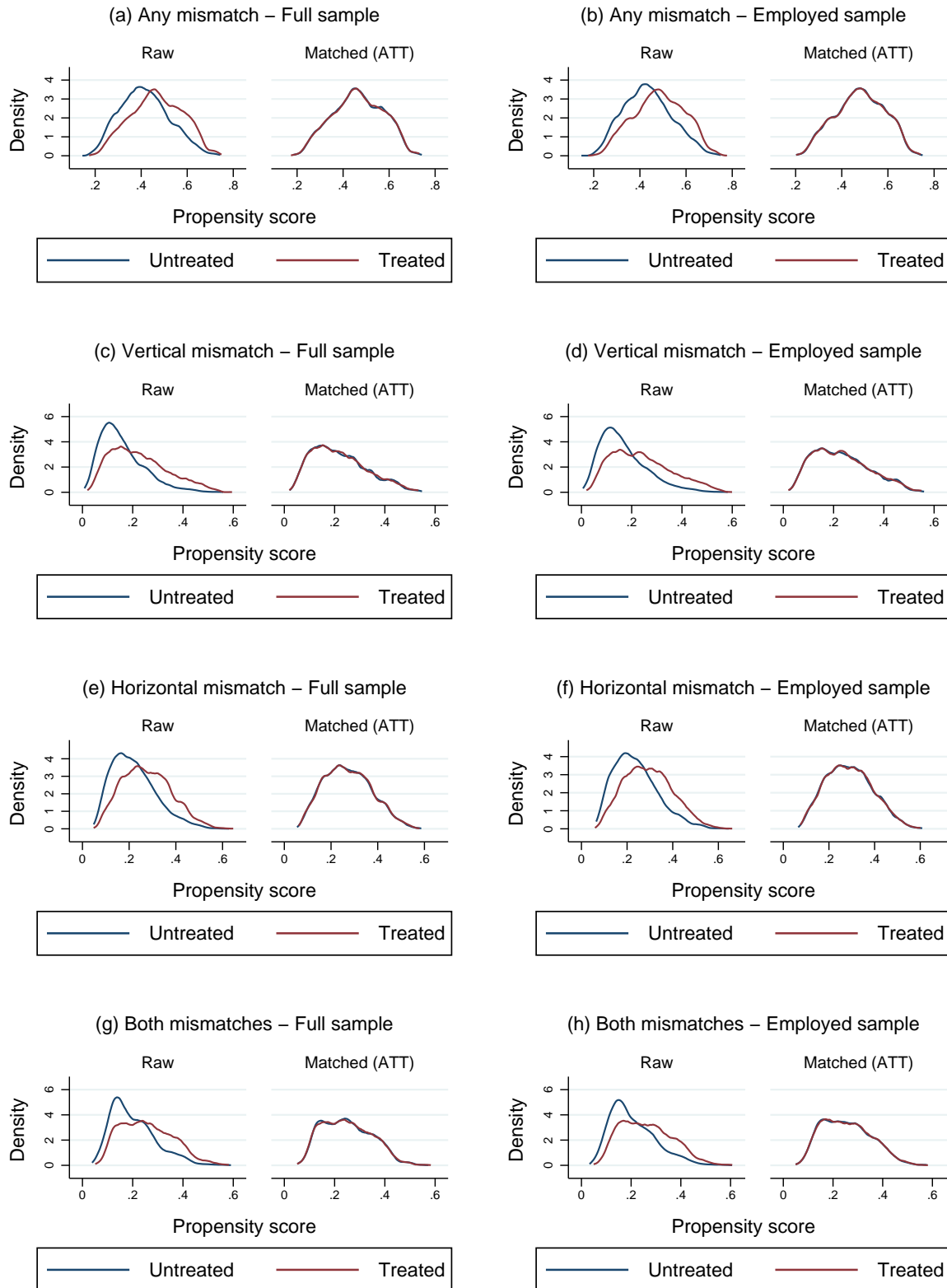


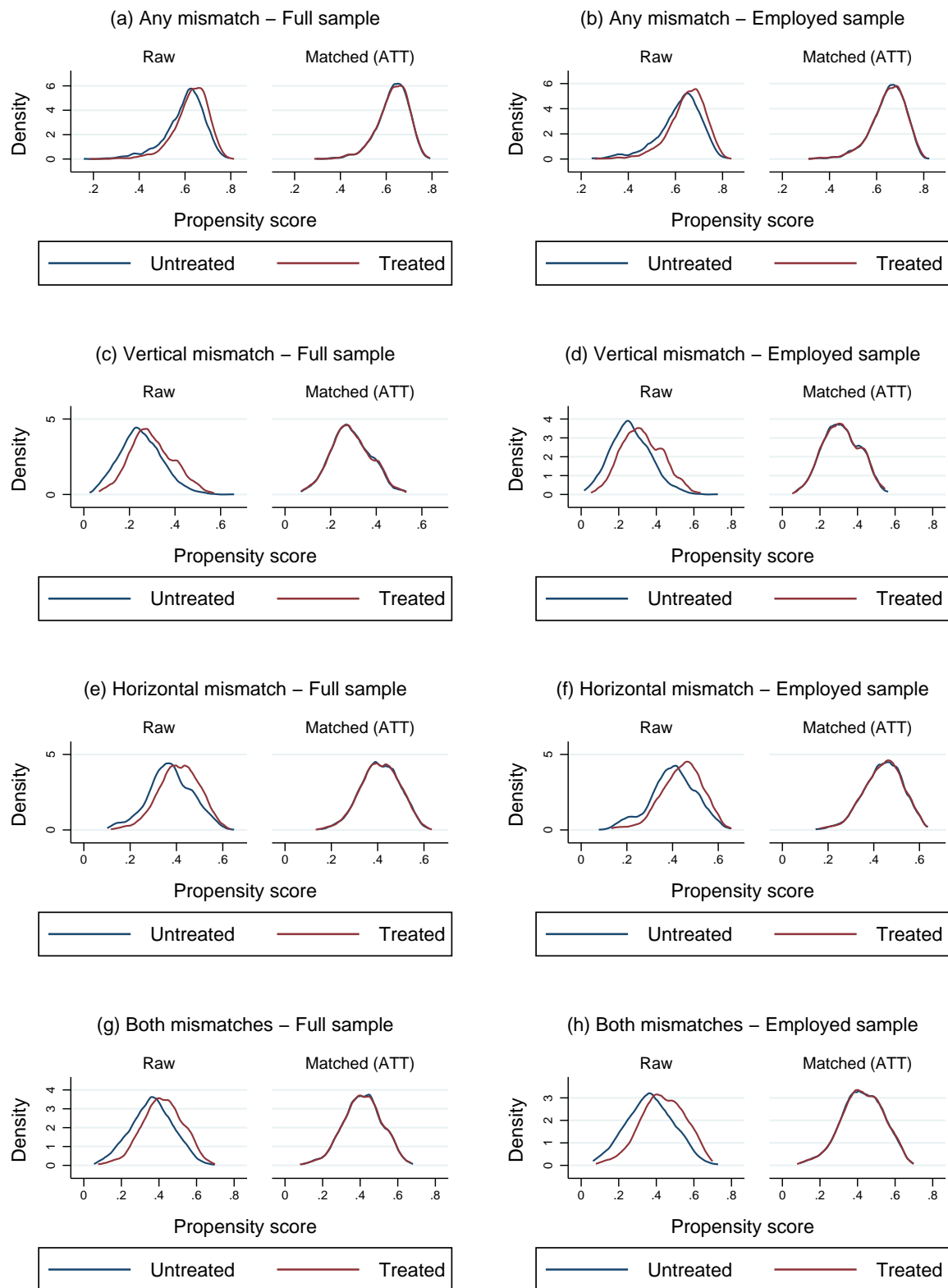
Figure 4.2: Distribution of the propensity score on the matched and unmatched sample -  $T_1$ 



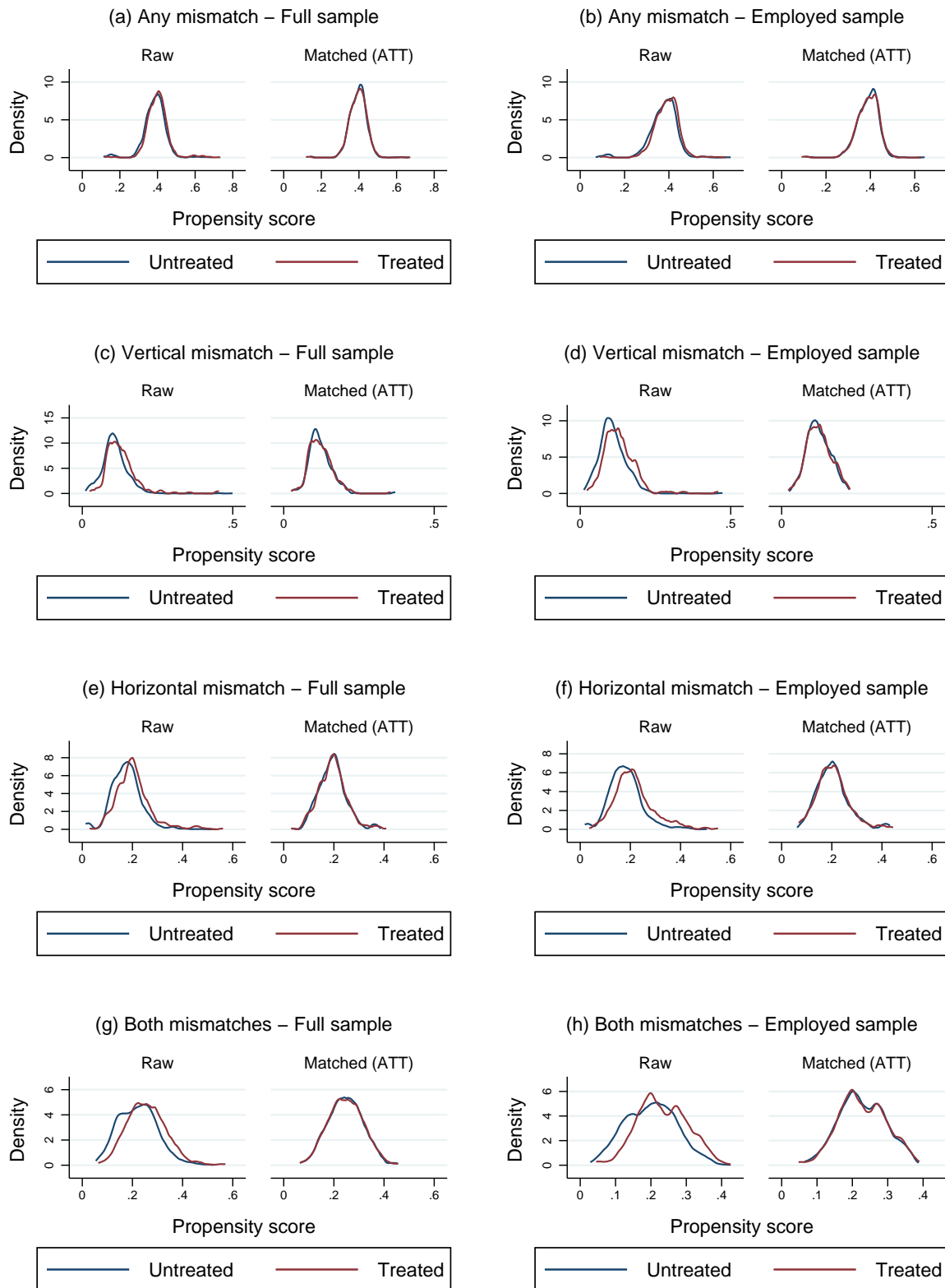
Figure 4.3: Distribution of the propensity score on the matched and unmatched sample -  $T_2$ 

Figure 4.4: Covariate balance on the matched and unmatched sample -  $T_0$

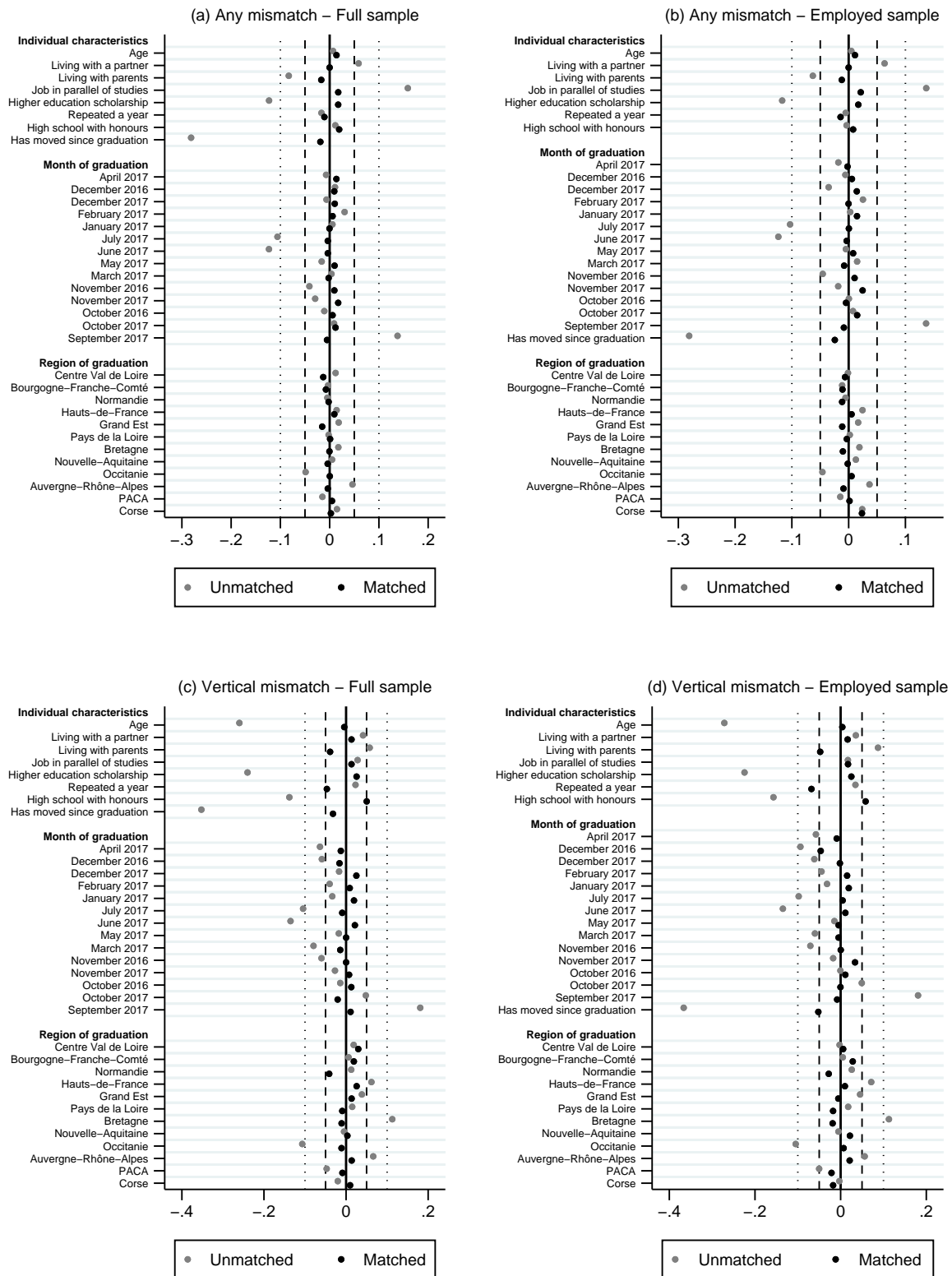


Figure 4.4: Covariate balance on the matched and unmatched sample -  $T_0$  (continued)

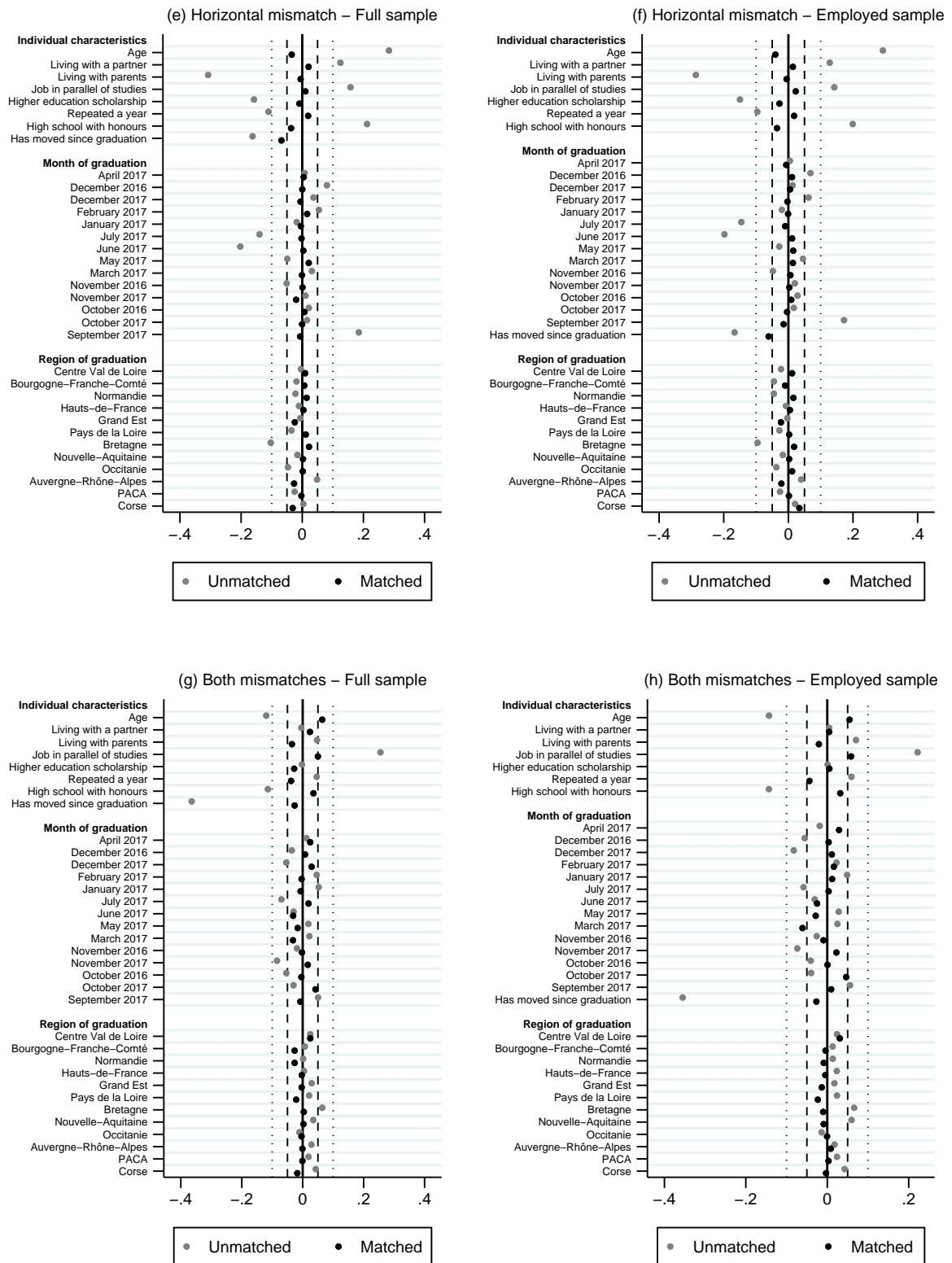


Figure 4.5: Covariate balance on the matched and unmatched sample -  $T_1$

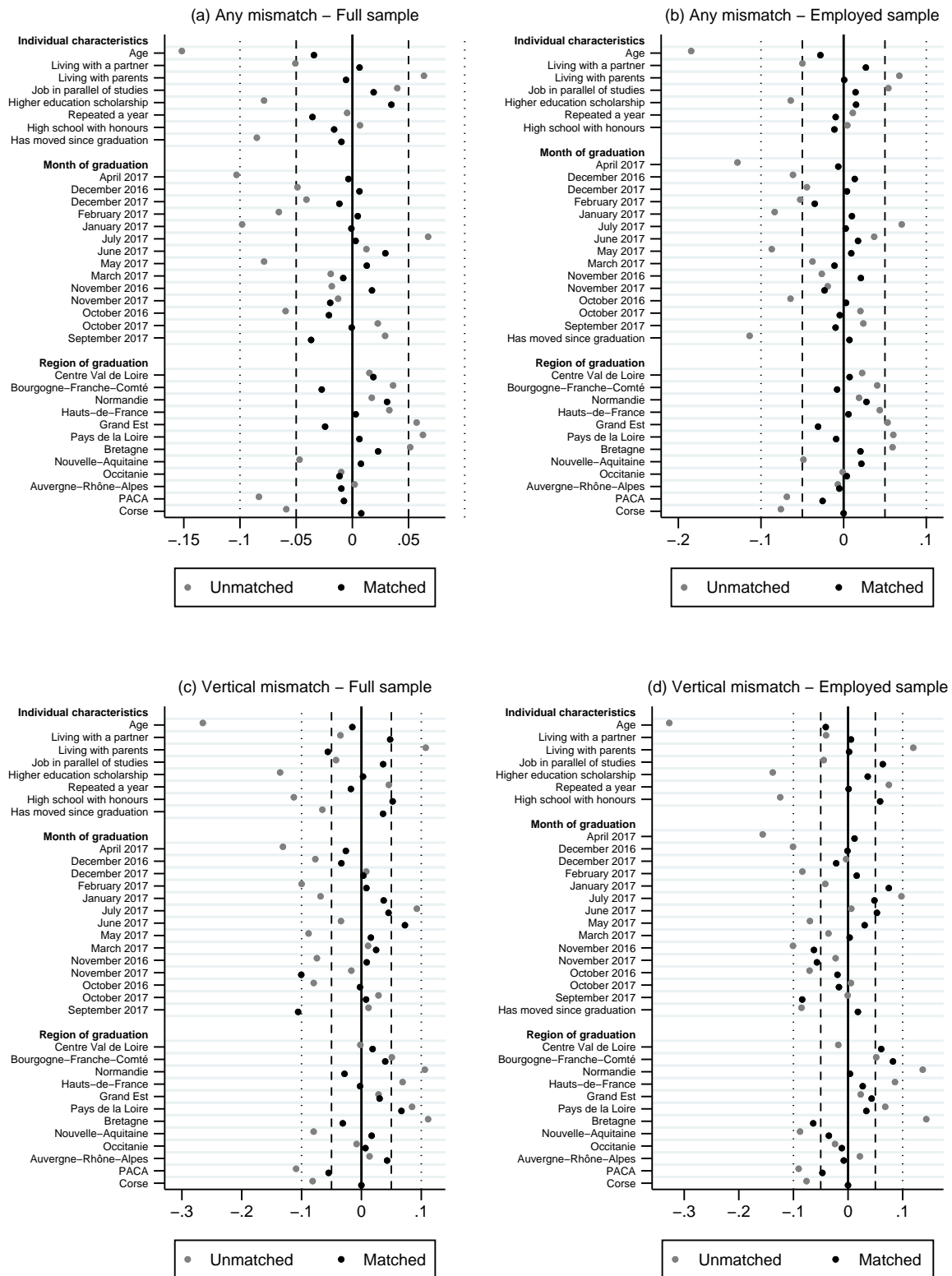


Figure 4.5: Covariate balance on the matched and unmatched sample -  $T_1$  (continued)

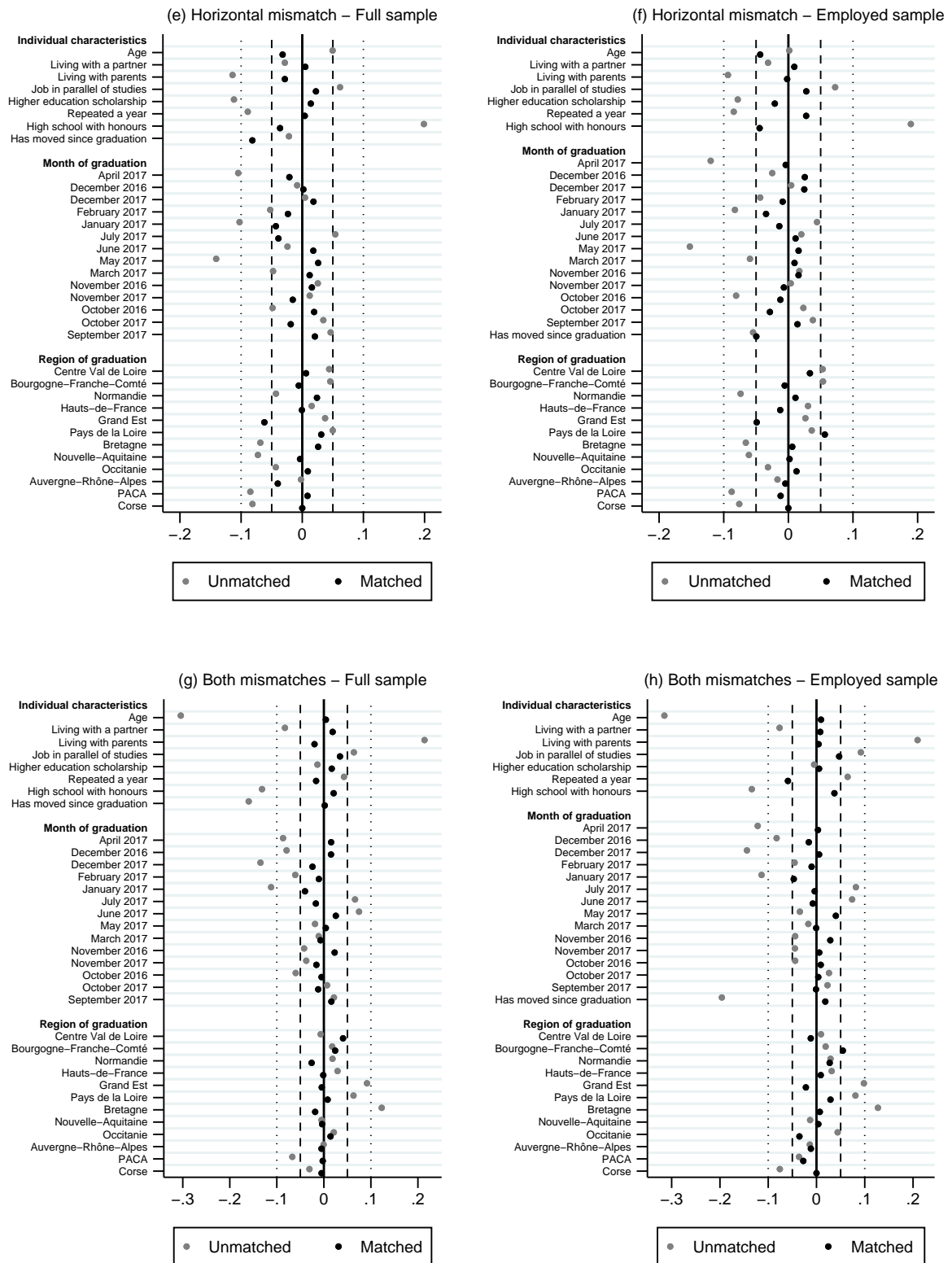


Figure 4.6: Covariate balance on the matched and unmatched sample -  $T_2$

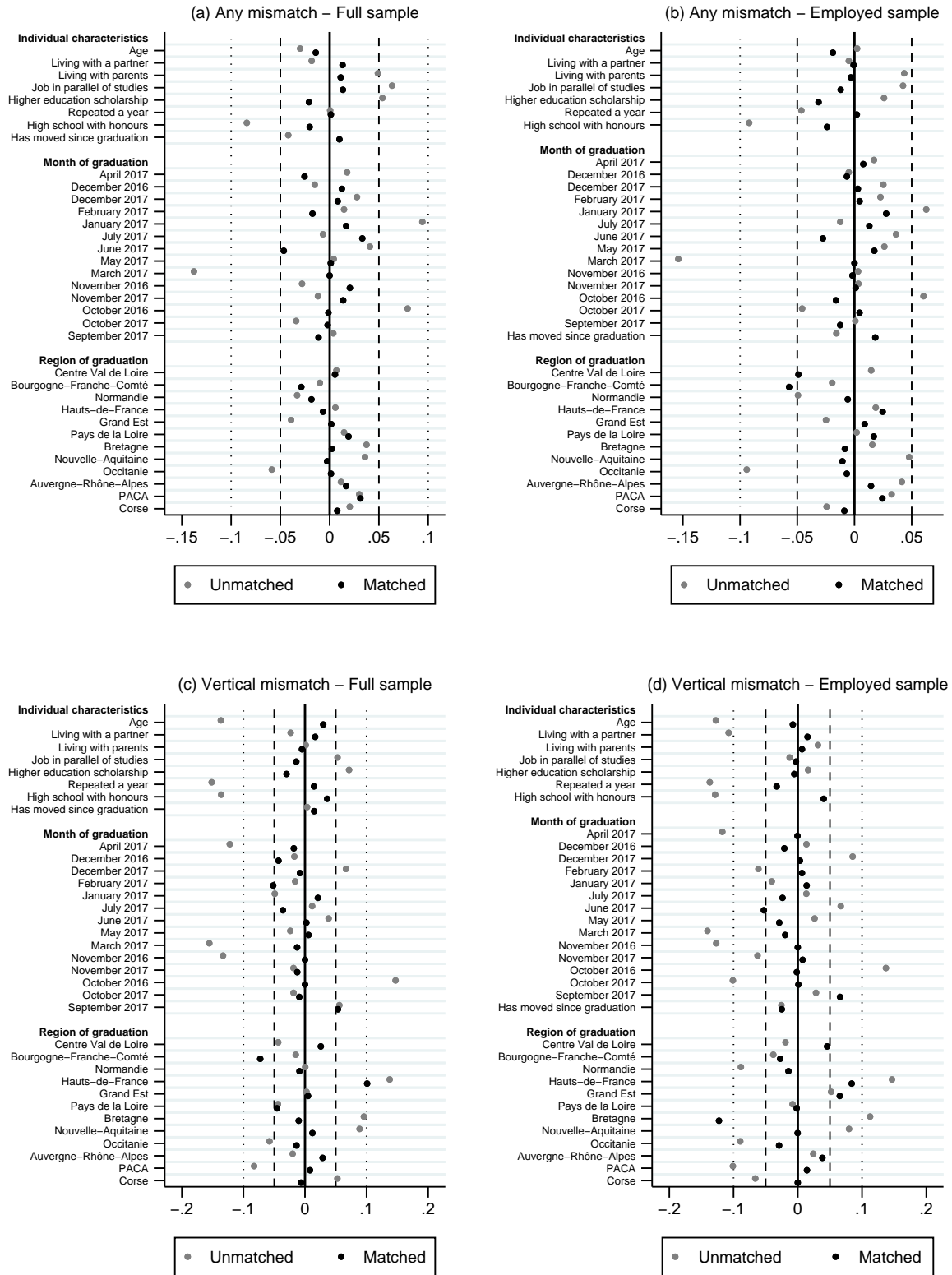
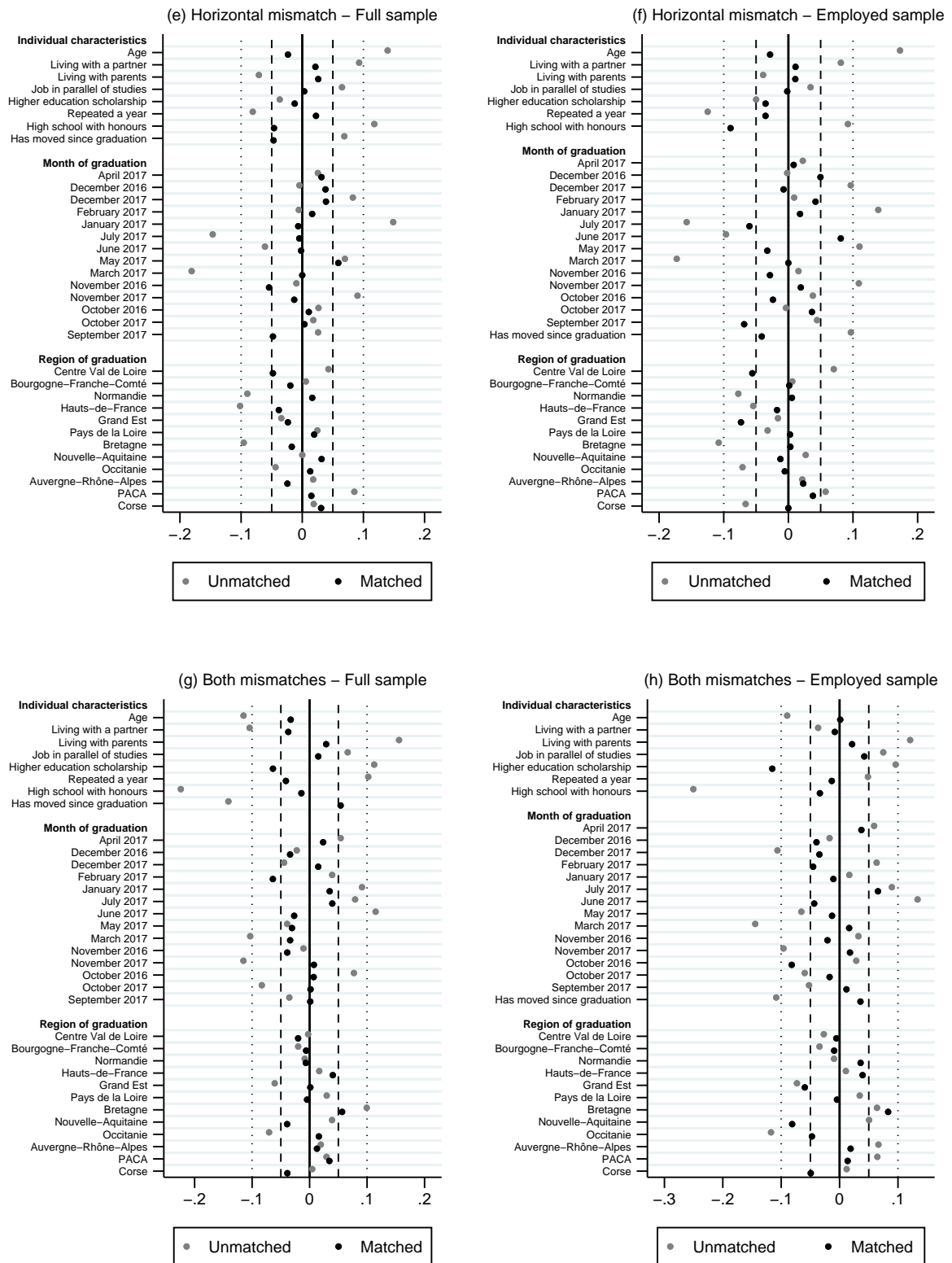


Figure 4.6: Covariate balance on the matched and unmatched sample -  $T_2$  (continued)



## Appendix 4C: Robustness checks

Table 4.9: Treatment effects on labour market outcomes - Any mismatch

| Matching strategy                    | Time of treatment | Outcomes (at the time of survey) |                     |                      |                      |
|--------------------------------------|-------------------|----------------------------------|---------------------|----------------------|----------------------|
|                                      |                   | (1) Employed                     | (2) Permanent job   | (3) Work full-time   | (4) Earnings         |
| Kernel matching<br>bandwidth = 0.025 | $T_0$             | 0.052***<br>(0.008)              | 0.084***<br>(0.009) | 0.036***<br>(0.008)  | 138.72***<br>(16.28) |
|                                      | $T_1$             | 0.059***<br>(0.014)              | 0.086***<br>(0.018) | 0.077***<br>(0.015)  | 153.23***<br>(24.81) |
|                                      | $T_2$             | -0.025<br>(0.017)                | -0.045**<br>(0.023) | -0.034*<br>(0.020)   | -85.90***<br>(31.29) |
| Kernel matching<br>bandwidth = 0.05  | $T_0$             | 0.053***<br>(0.008)              | 0.085***<br>(0.009) | 0.036***<br>(0.008)  | 135.97***<br>(16.16) |
|                                      | $T_1$             | 0.061***<br>(0.013)              | 0.091***<br>(0.017) | 0.078***<br>(0.015)  | 159.71***<br>(24.19) |
|                                      | $T_2$             | -0.027<br>(0.017)                | -0.055**<br>(0.022) | -0.039**<br>(0.019)  | -92.68***<br>(29.92) |
| 5 nearest neighbours<br>matching     | $T_0$             | 0.048***<br>(0.008)              | 0.087***<br>(0.010) | 0.036***<br>(0.009)  | 136.92***<br>(16.77) |
|                                      | $T_1$             | 0.067***<br>(0.014)              | 0.097***<br>(0.018) | 0.077***<br>(0.016)  | 158.11***<br>(25.28) |
|                                      | $T_2$             | -0.027<br>(0.019)                | -0.051**<br>(0.023) | -0.041**<br>(0.020)  | -97.99***<br>(32.59) |
| Inverse probability<br>weighting     | $T_0$             | 0.054***<br>(0.007)              | 0.090***<br>(0.009) | 0.036***<br>(0.008)  | 127.21***<br>(15.71) |
|                                      | $T_1$             | 0.062***<br>(0.013)              | 0.096***<br>(0.016) | 0.069***<br>(0.014)  | 146.62***<br>(23.36) |
|                                      | $T_2$             | -0.027<br>(0.017)                | -0.047**<br>(0.021) | -0.048***<br>(0.018) | -93.86***<br>(28.70) |

Note: This table displays the ATT of being mismatched (in any dimension) in the first job on the different outcomes, observed at the time of the survey. The effect on the probability of employment (1) is calculated on the whole sample. Effects on outcomes (2) to (4) are calculated considering only individuals in employment.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculations based on *Génération 2017* data.

Table 4.10: Treatment effects on matching status - Any mismatch

| Treatment                            | Time of treatment | Outcomes (at the time of survey) |                     |                     |                      |
|--------------------------------------|-------------------|----------------------------------|---------------------|---------------------|----------------------|
|                                      |                   | (1) Ver. mismatch                | (2) Hor. mismatch   | (3) Full mismatch   | (4) Match            |
| Kernel matching<br>bandwidth = 0.025 | $T_0$             | 0.043***<br>(0.010)              | 0.098***<br>(0.011) | 0.019*<br>(0.011)   | -0.159***<br>(0.009) |
|                                      | $T_1$             | 0.031**<br>(0.015)               | 0.112***<br>(0.019) | -0.013<br>(0.018)   | -0.130***<br>(0.015) |
|                                      | $T_2$             | 0.044**<br>(0.018)               | 0.091***<br>(0.022) | 0.080***<br>(0.023) | -0.215***<br>(0.018) |
| Kernel matching<br>bandwidth = 0.05  | $T_0$             | 0.042***<br>(0.010)              | 0.099***<br>(0.011) | 0.020*<br>(0.011)   | -0.161***<br>(0.009) |
|                                      | $T_1$             | 0.033**<br>(0.014)               | 0.112***<br>(0.018) | -0.011<br>(0.017)   | -0.133***<br>(0.015) |
|                                      | $T_2$             | 0.039**<br>(0.018)               | 0.089***<br>(0.022) | 0.076***<br>(0.022) | -0.204***<br>(0.017) |
| 5 nearest neighbours<br>matching     | $T_0$             | 0.040***<br>(0.011)              | 0.099***<br>(0.012) | 0.019***<br>(0.011) | -0.158***<br>(0.009) |
|                                      | $T_1$             | 0.032**<br>(0.015)               | 0.110***<br>(0.019) | -0.011<br>(0.018)   | -0.131***<br>(0.015) |
|                                      | $T_2$             | 0.037**<br>(0.019)               | 0.083***<br>(0.024) | 0.087***<br>(0.023) | -0.206***<br>(0.018) |
| Inverse probability<br>weighting     | $T_0$             | 0.060***<br>(0.009)              | 0.095***<br>(0.011) | 0.027***<br>(0.010) | -0.182***<br>(0.008) |
|                                      | $T_1$             | 0.039***<br>(0.013)              | 0.111***<br>(0.017) | -0.015<br>(0.016)   | -0.135***<br>(0.014) |
|                                      | $T_2$             | 0.041**<br>(0.017)               | 0.086***<br>(0.021) | 0.092***<br>(0.021) | -0.218***<br>(0.016) |

Note: This table displays the ATT of being mismatched (in any dimension) in the first job on the different outcomes, observed at the time of the survey. Effects calculated considering only individuals in employment.

\*\*\* indicates significance at the 1% level, \*\* at the 5% level and \* at the 10% level.

Source: Author's calculations based on *Génération 2017* data.



## Appendix 4D: Additional descriptive statistics

Table 4.11: Descriptive statistics - Outcomes

| <b>Matching status<br/>in the current job</b> | Matched | Only vertical<br>mismatch | Only horizontal<br>mismatch | Both<br>mismatches |
|---|---------|---------------------------|-----------------------------|--------------------|
| Number of observations                        | 3,881   | 1,722                     | 3,345                       | 2,358              |
| Proportion                                    | 34.3%   | 15.2%                     | 29.6%                       | 20.9%              |
| Share of permanent jobs                       | 84.3%   | 83.4%                     | 81.2%                       | 75.8%              |
| Share of full-time jobs                       | 88.1%   | 85.8%                     | 88.5%                       | 80.5%              |
| Average earnings (in €)                       | 2,311   | 1,818                     | 2,210                       | 1,688              |

Source: Author's calculations based on *Génération 2017* data.



# Conclusion

An extensive literature show that overeducation, far from marginal, is high and/or increasing in many advanced countries. More than a quarter of the labour force is affected by mismatch in OECD countries, with detrimental effects both for individuals on wages, career paths or job satisfaction, and the economy as the whole on global productivity.

The aim of this thesis was to assess the extent to which mobility, whether geographical or occupational, interacts with mismatch, in order to identify public policy levers that could help counter this phenomenon. More specifically, I focused on the context of local labour markets and their geography (chapter 2), the role of individual spatial mobility (chapter 3) and the influence of mismatch on occupational mobility and career trajectories (chapter 4). On another note, chapter 1 documents overeducation measures.

Several key elements emerge from this work. First, the methodological work in chapter 1 results in a new measure of overeducation, which accounts for the skill content of the jobs. The comparison between measures tend to confirm that the various approaches apprehend different dimensions of the phenomenon. Second, geographical mobility appears to have some strong links with mismatch. Results in chapter 2 suggest that, in rural areas, the lack of job opportunities forces educated workers to move if they want to avoid overeducation. In urban areas, overeducation seem to be more driven by the competition between workers. Chapter 3 somehow confirms these findings, showing that individual spatial mobility is a way to avoid mismatch. This is especially true if workers are ready to move to large economic centers. Third, although sometimes presented as a “stepping stone” towards better professional positions, mismatch (either horizontal or vertical) rather appear to have detrimental effects on future occupational mobility (chapter 4).

Finally, some elements mentioned in this thesis still raise questions that may pave the way for future research.

Although not the main focus of this work, one side result stands out clearly in three of the four chapters: Bachelor's degrees are associated with a higher risk of overeducation and greater detrimental effects. Given that this level is supposed to sanction the first cycle of university (European Commission, 2017), this claims for further investigation of the potential disparities between the qualifications provided by the education system and those sought by employers. In particular, it could be important to compare Bachelor's degrees to shorter vocational degrees, which seem to be preferred on the labour market.

Chapters 2 and 3 evidence the influence of the local context and the importance of being mobile to avoid overeducation, for the main reason that skilled jobs appear not to be evenly distributed across areas. However, the recent spread of teleworking might be a game-changer, as it should allow a physical (or spatial) disconnection between workers and firms and reduce the importance of spatial mismatch. It thus will be interesting to explore the extent to which remote working could be an alternative lever to alleviate educational mismatch.

Chapter 4 evidences a “trap effect” of mismatch for young workers entering the labour market. However, the analysis conducted here allows to observe this effect only in the first years following graduation. It would therefore be of great interest to explore whether those workers eventually access a matched job in the longer term, and if so, at which cost.

# Bibliography

- Abadie, A., & Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2), 781–807. <https://doi.org/10.3982/ECTA11293>
- Acemoglu, D. (1999). Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review*, 89(5), 1259–1278. <https://doi.org/10.1257/aer.89.5.1259>
- Affichard, J. (1981). Quels emplois après l'école : La valeur des titres scolaires depuis 1973. *Économie et Statistique*, 134(1), 7–26. <https://doi.org/10.3406/estat.1981.4497>
- Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, 115(3), 715–753. <https://doi.org/10.1162/003355300554881>
- Akerlof, G. A., & Kranton, R. E. (2002). Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature*, 40(4), 1167–1201. <https://doi.org/10.1257/002205102762203585>
- Alba-Ramírez, A. (1993). Mismatch in the Spanish labor market: Overeducation? *Journal of Human Resources*, 28(2), 259–278. <https://doi.org/10.2307/146203>
- Alba-Ramírez, A. (1998). How temporary is temporary employment in Spain? *Journal of Labor Research*, 19(4), 695–710. <https://doi.org/10.1007/s12122-998-1056-5>
- Allen, J., & van der Velden, R. (2001). Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job search. *Oxford economic papers*, 53(3), 434–452. <https://doi.org/10.1093/oep/53.3.434>
- Alstadsæter, A. (2011). Measuring the consumption value of higher education. *CESifo Economic Studies*, 57(3), 458–479. <https://doi.org/10.1093/cesifo/ifq009>
- Alstadsæter, A., Kolm, A.-S., & Larsen, B. (2008). Money or joy: The choice of educational type. *European Journal of Political Economy*, 24(1), 107–122. <https://doi.org/10.1016/j.ejpoleco.2007.04.004>
- Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1, Part 1), 48–83. <https://doi.org/10.1086/298317>
- Amuedo-Dorantes, C. (2000). Work transitions into and out of involuntary temporary employment in a segmented market: Evidence from Spain. *ILR Review*, 53(2), 309–325. <https://doi.org/10.1177/001979390005300207>
- Andersson, M., Klaesson, J., & Larsson, J. P. (2014). The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies? *Papers in Regional Science*, 93(4), 727–747. <https://doi.org/10.1111/pirs.12025>

- Argan, D., & Gary-Bobo, R. (2023). Les diplômés français se sont-ils dévalorisés ? Les rendements de l'éducation et de l'expérience en France, 1992-2017. *Revue Économique*, 74(2), 95–137. <https://doi.org/10.3917/reco.742.0095>
- Argan, D., Gary-Bobo, R., & Goussé, M. (2023). Is there a devaluation of degrees? Unobserved heterogeneity in returns to education and early experience. *DT 2304*, CEPREMAP. <http://www.cepremap.fr/depot/docweb/docweb2304.pdf>
- Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics*, 2(3), 193–216. [https://doi.org/10.1016/0047-2727\(73\)90013-3](https://doi.org/10.1016/0047-2727(73)90013-3)
- Arulampalam, W. (2001). Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, 111(475), 585–606. <https://doi.org/10.1111/1468-0297.00664>
- Attali, J. (1998). *Pour un modèle européen d'enseignement supérieur* [rédigé par P. Brandys, G. Charpak, S. Feneuille, A. Kahn, J. Kristeva, M. Lazdunski, M.-E. Leclerc, N. Le Douarin, C. Lewiner, C. Marchello Nizia, F. Mer, J. Monod, R. Pellat, A. Touraine ; rapporteurs : O. Brigaud, F. Mion]. Rapport à M. le Ministre de l'Éducation Nationale, de la Recherche et de la Technologie. <https://www.vie-publique.fr/rapport/26165-pour-un-modele-europeen-denseignement-superieur-rapport-m-le-minis>
- Autor, D. H., Goldin, C. D., & Katz, L. F. (2020). Extending the Race between Education and Technology. *AEA Papers and Proceedings*, 110, 347–351. <https://doi.org/10.1257/pandp.20201061>
- Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the US labor market. *American Economic Review*, 96(2), 189–194. <https://doi.org/10.1257/000282806777212620>
- Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: Stepping stone or trap? *Labour Economics*, 25, 123–140. <https://doi.org/10.1016/j.labeco.2013.04.013>
- Baert, S., & Verhaest, D. (2019). Unemployment or overeducation: Which is a worse signal to employers? *De Economist*, 167(1), 1–21. <https://doi.org/10.1007/s10645-018-9330-2>
- Banerjee, S., Chatterji, P., & Lahiri, K. (2017). Effects of psychiatric disorders on labor market outcomes: A latent variable approach using multiple clinical indicators. *Health Economics*, 26(2), 184–205. <https://doi.org/10.1002/hec.3286>
- Barnay, T. (2020). Interactions of Work and Health: An Economic Perspective. In T. Theorell (Ed.), *Handbook of Socioeconomic Determinants of Occupational Health: From Macro-level to Micro-level Evidence* (pp. 37–52). Springer. [https://doi.org/10.1007/978-3-030-31438-5\\_30](https://doi.org/10.1007/978-3-030-31438-5_30)
- Barone, C., & Ortiz, L. (2011). Overeducation among European University Graduates: A comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education*, 61, 325–337. <https://doi.org/10.1007/s10734-010-9380-0>
- Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.
- Becker, G. S., Hubbard, W. H. J., & Murphy, K. M. (2010). Explaining the worldwide boom in higher education of women. *Journal of Human Capital*, 4(3), 203–241. <https://doi.org/10.1086/657914>

- 
- Bell, D. N. F., & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, 27(2), 241–267. <https://doi.org/10.1093/oxrep/grr011>
- Berlingieri, F. (2019). Local labor market size and qualification mismatch. *Journal of Economic Geography*, 19(6), 1261–1286. <https://doi.org/10.1093/jeg/lby045>
- Bernela, B., & Bonnal, L. (2022). Mobilités géographiques et relation formation-emploi: Une analyse longitudinale de l'enseignement supérieur français. *Revue Économique*, 73(5), pp. 695–734. <https://doi.org/10.3917/reco.735.0695>
- Bishop, J. (1994). Schooling, learning and worker productivity. In R. Asplund (Ed.), *Human capital creation in an economic perspective* (pp. 13–67). [https://doi.org/10.1007/978-3-642-99776-1\\_2](https://doi.org/10.1007/978-3-642-99776-1_2)
- Blanchflower, D. G., & Freeman, R. B. (2000). The declining economic status of young workers in OECD countries. In D. G. Blanchflower & R. B. Freeman (Eds.), *Youth Employment and Joblessness in Advanced Countries* (pp. 19–56). University of Chicago Press. <http://www.nber.org/chapters/c6801>
- Blau, D. M., & Robins, P. K. (1990). Job search outcomes for the employed and unemployed. *Journal of Political Economy*, 98(3), 637–655. <https://www.jstor.org/stable/2937703>
- Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. New York: John Wiley and Sons.
- Büchel, F., & Battu, H. (2003). The theory of differential overqualification: Does it work? *Scottish Journal of Political Economy*, 50(1), 1–16. <https://doi.org/10.1111/1467-9485.00251>
- Büchel, F., & Mertens, A. (2004). Overeducation, undereducation, and the theory of career mobility. *Applied Economics*, 36(8), 803–816. <https://doi.org/10.1080/0003684042000229532>
- Büchel, F., & van Ham, M. (2003). Overeducation, regional labor markets, and spatial flexibility. *Journal of Urban Economics*, 53(3), 482–493. [https://doi.org/10.1016/S0094-1190\(03\)00008-1](https://doi.org/10.1016/S0094-1190(03)00008-1)
- Büchel, F. (2002). The effects of overeducation on productivity in Germany — the firms' viewpoint. *Economics of Education Review*, 21(3), 263–275. [https://doi.org/10.1016/S0272-7757\(01\)00020-6](https://doi.org/10.1016/S0272-7757(01)00020-6)
- Caliendo, M., Mahlstedt, R., & Mitnik, O. A. (2017). Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. *Labour Economics*, 46, 14–25. <https://doi.org/10.1016/j.labeco.2017.02.001>
- Capsada-Munsech, Q. (2017). Overeducation: Concept, theories, and empirical evidence. *Sociology Compass*, 11(10), e12518. <https://doi.org/10.1111/soc4.12518>
- Capsada-Munsech, Q. (2019). Measuring overeducation: Incidence, correlation and overlaps across indicators and countries. *Social Indicators Research*, 145(1), 279–301. <https://doi.org/10.1007/s11205-019-02112-0>
- Card, D., & 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), 1–40. <https://doi.org/10.1086/261805>
- Cedefop. (2018). *Insights into skill shortages and skill mismatch: Learning from Cedefop's European skills and jobs survey*. <http://hdl.voced.edu.au/10707/448822>

- Chevalier, A., Gibbons, S., Thorpe, A., Snell, M., & Hoskins, S. (2009). Students' academic self-perception. *Economics of Education Review*, 28(6), 716–727. <https://doi.org/10.1016/j.econedurev.2009.06.007>
- Chevalier, A., & Lindley, J. (2009). Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(2), 307–337. <https://doi.org/10.1111/j.1467-985X.2008.00578.x>
- Chevalier, A. (2003). Measuring over-education. *Economica*, 70(279), 509–531. <https://doi.org/10.1111/1468-0335.t01-1-00296>
- Cliff, A. D., & Ord, J. K. (1973). *Spatial Autocorrelation*. London: Pion.
- Cliff, A. D., & Ord, J. K. (1981). *Spatial processes, Models and Applications*. London: Pion.
- Cockx, B., & Picchio, M. (2012). Are short-lived jobs stepping stones to long-lasting jobs? *Oxford Bulletin of Economics and Statistics*, 74(5), 646–675. <https://doi.org/10.1111/j.1468-0084.2011.00668.x>
- Coleman, J. S. (1961). *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*. New York: Free Press.
- Collins, R. (2002). Credential inflation and the future of universities. In S. Brint (Ed.), *The future of the city of intellect: The changing American university* (pp. 23–46). Stanford University Press.
- Combes, P.-P., Duranton, G., & Gobillon, L. (2015). Workers and wages in the Greater Paris region. *Revue Économique*, 66(2), 317–350. <https://doi.org/10.3917/reco.pr2.0039>
- Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2012). Sorting and local wage and skill distributions in France [Special issue 'In honor of Jacques Thisse']. *Regional Science and Urban Economics*, 42(6), 913–930. <https://doi.org/10.1016/j.regsciurbeco.2012.11.003>
- Conroy, T., & Watson, P. (2023). Overeducation, natural amenities, and entrepreneurship. *Small Business Economics*, 1–21. <https://doi.org/10.1007/s11187-023-00734-w>
- Coombs, P. H., & Manzoor, A. (1974). *Attacking rural poverty: How nonformal education can help* [A Research Report for the World Bank Prepared by the International Council for Educational Development]. Johns Hopkins University Press, Baltimore. <https://eric.ed.gov/?id=ED118340>
- Cracolici, M. F., Cuffaro, M., & Nijkamp, P. (2007). Geographical distribution of unemployment: An analysis of provincial differences in Italy. *Growth and Change*, 38(4), 649–670. <https://doi.org/10.1111/j.1468-2257.2007.00391.x>
- Croce, G., & Ghignoni, E. (2012). Demand and supply of skilled labour and overeducation in Europe: A country-level analysis. *Comparative Economic Studies*, 54(2), 413–439. <https://doi.org/10.1057/ces.2012.12>
- Croce, G., & Ghignoni, E. (2015). Educational mismatch and spatial flexibility in Italian local labour markets. *Education Economics*, 23(1), 25–46. <https://doi.org/10.1080/09645292.2012.754121>
- D'Alessandro, C., Levy, D., & Regnier, T. (2021). Une nouvelle définition du rural pour mieux rendre compte des réalités des territoires et de leurs transformations. In *La France et ses territoires*, Insee References. INSEE. <https://www.insee.fr/fr/statistiques/5039991?sommaire=5040030>



- DaVanzo, J. S., & Morrison, P. A. (1981). Return and other sequences of migration in the United States. *Demography*, 18(1), 85–101. <https://doi.org/10.2307/2061051>
- Davia, M. A., McGuinness, S., & O’Connell, P. J. (2017). Determinants of regional differences in rates of overeducation in Europe. *Social science research*, 63, 67–80. <https://doi.org/10.1016/j.ssresearch.2016.09.009>
- Debarsy, N., & Ertur, C. (2010). Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 40(6), 453–470. <https://doi.org/10.1016/j.regsciurbeco.2010.06.001>
- de Graaf-Zijl, M., van den Berg, G. J., & Heyma, A. (2011). Stepping stones for the unemployed: the effect of temporary jobs on the duration until (regular) work. *Journal of Population Economics*, 24, 107–139. <https://doi.org/10.1007/s00148-009-0287-y>
- Détang-Dessendre, C. (1999). Reciprocal link between exit from unemployment and geographical mobility. *Environment and Planning A: Economy and Space*, 31(8), 1417–1431. <https://doi.org/10.1068/a311417>
- Détang-Dessendre, C., Drapier, C., & Jayet, H. (2004). The impact of migration on wages: Empirical evidence from French youth. *Journal of Regional Science*, 44(4), 661–691. <https://doi.org/10.1111/j.0022-4146.2004.00353.x>
- Devillanova, C. (2013). Over-education and spatial flexibility: New evidence from Italian survey data. *Papers in Regional Science*, 92(3), 445–464. <https://doi.org/10.1111/j.1435-5957.2012.00434.x>
- Di Paola, V., & Moullet, S. (2009). Femmes et fonction publique. Un risque calculé de déclassement ? *Travail et emploi*, (120), 47–61. <https://doi.org/10.4000/travailemploi.1655>
- Di Paola, V., & Moullet, S. (2012). Peut-on en finir avec le déclassement ? : Évolution du déclassement dans la fonction publique en début de carrière. *Formation emploi. Revue française de sciences sociales*, (118), 83–101. <https://doi.org/10.4000/formationemploi.3610>
- Di Paola, V., & Moullet, S. (2018). Le déclassement : un phénomène enraciné. In *20 ans d’insertion professionnelle des jeunes : entre permanences et évolutions* (pp. 79–83). Céreq Essentiels. <https://www.cereq.fr/20-ans-dinsertion-professionnelle-des-jeunes-entre-permanences-et-evolutions>
- Di Paolo, A., Matas, A., & Raymond, J. L. (2017). Job accessibility and job-education mismatch in the metropolitan area of Barcelona. *Papers in Regional Science*, 96(S1), S91–S112. <https://doi.org/10.1111/pirs.12179>
- Di Pietro, G. (2002). Technological change, labor markets, and ‘low-skill, low-technology traps’. *Technological Forecasting and Social Change*, 69(9), 885–895. [https://doi.org/10.1016/S0040-1625\(01\)00182-2](https://doi.org/10.1016/S0040-1625(01)00182-2)
- Dolado, J. J., Jansen, M., & Jimeno, J. F. (2009). On-the-job search in a matching model with heterogeneous jobs and workers. *The Economic Journal*, 119(534), 200–228. <https://doi.org/10.1111/j.1468-0297.2008.02210.x>
- Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, 19(2), 179–198. [https://doi.org/10.1016/S0272-7757\(97\)00036-8](https://doi.org/10.1016/S0272-7757(97)00036-8)

- Duncan, G. J., & Hoffman, S. D. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75–86. [https://doi.org/10.1016/0272-7757\(81\)90028-5](https://doi.org/10.1016/0272-7757(81)90028-5)
- Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-foundations of Urban Agglomeration Economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics, Vol. 4 - cities and Geography* (pp. 2063–2117, Vol. 4). Elsevier. [https://doi.org/10.1016/S1574-0080\(04\)80005-1](https://doi.org/10.1016/S1574-0080(04)80005-1)
- Duru-Bellat, M. (2006). *L'inflation scolaire: Les désillusions de la méritocratie*. Paris, Le Seuil, collection La République des idées.
- Elhorst, J. P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of Economic Surveys*, 17(5), 709–748. <https://doi.org/10.1046/j.1467-6419.2003.00211.x>
- Emmons, W. R., Kent, A. H., & Ricketts, L. R. (2019). Is College Still Worth It? The New Calculus of Falling Returns. *Federal Reserve Bank of St. Louis Review*, 101(4), 297–329. <https://doi.org/10.20955/r.101.297-329>
- European Commission. (2017). Council Recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning [2017/C 189/03]. [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32017H0615\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32017H0615(01))
- Faggian, A., & McCann, P. (2009). Human capital, graduate migration and innovation in British regions. *Cambridge Journal of Economics*, 33(2), 317–333. <https://doi.org/10.1093/cje/ben042>
- Fahr, R., & Sunde, U. (2006). Spatial mobility and competition for jobs: Some theory and evidence for Western Germany. *Regional Science and Urban Economics*, 36(6), 803–825. <https://doi.org/10.1016/j.regsciurbeco.2006.05.001>
- Filomena, M., & Picchio, M. (2021). Are temporary jobs stepping stones or dead ends? A meta-analytical review of the literature. *IZA Discussion Paper No. 14367*. <https://doi.org/10.2139/ssrn.3851042>
- Fleisher, M. S., & Tsacoumis, S. (2012). O\*NET analyst occupational skills ratings: Procedures update. In *Raleigh, NC: National Center for O\*NET Development*. [https://www.onetcenter.org/reports/AOSkills\\_ProcUpdate.html](https://www.onetcenter.org/reports/AOSkills_ProcUpdate.html)
- Floch, J.-M., & Le Saout, R. (2018). Économétrie spatiale : Modèles courants. In *Manuel d'analyse spatiale, Insee Méthodes 131*. INSEE. <https://www.insee.fr/fr/information/3635442>
- Forgeot, G., & Gautié, J. (1997). Insertion professionnelle des jeunes et processus de déclassement. *Économie et Statistique*, 304(1), 53–74. <https://doi.org/10.3406/estat.1997.2556>
- Fouquet, F., & Sari, F. (2023). Spatial mobility and overeducation of young workers: New evidence from France. *Papers in Regional Science*. <https://doi.org/10.1111/pirs.12755>
- Frank, R. H. (1978). Why women earn less: The theory and estimation of differential overqualification. *American Economic Review*, 68(3), 360–373. <https://www.jstor.org/stable/1805267>
- Freeman, R. B. (1976). *The Overeducated American*. New York: Academic Press.

- 
- Fuller, W. C., Manski, C. F., & Wise, D. A. (1982). New Evidence on the Economic Determinants of Postsecondary Schooling Choices. *Journal of Human Resources*, 17(4), 477–498. <https://doi.org/10.2307/145612>
- Gangl, M. (2003). The only way is up? Employment protection and job mobility among recent entrants to European labour markets. *European Sociological Review*, 19(5), 429–449. <https://doi.org/10.1093/esr/19.5.429>
- Gautier, P. A. (2002). Unemployment and Search Externalities in a Model with Heterogeneous Jobs and Workers. *Economica*, 69(273), 21–40. <https://doi.org/10.1111/1468-0335.00269>
- Ghignoni, E., & Verashchagina, A. (2014). Educational qualifications mismatch in Europe. Is it demand or supply driven? *Journal of Comparative Economics*, 42(3), 670–692. <https://doi.org/10.1016/j.jce.2013.06.006>
- Ghosh, S., & Grassi, E. (2020). Overeducation and overskilling in the early careers of PhD graduates: Does international migration reduce labour market mismatch? *Papers in Regional Science*, 99(4), 915–944. <https://doi.org/10.1111/pirs.12509>
- Gobillon, L., & Selod, H. (2021). Spatial mismatch, poverty, and vulnerable populations. In M. M. Fischer & P. Nijkamp (Eds.), *Handbook of Regional Science* (pp. 573–588). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-662-60723-7\\_7#DOI](https://doi.org/10.1007/978-3-662-60723-7_7#DOI)
- Goldin, C. D., & Katz, L. F. (2008). *The Race between Education and Technology*. Harvard University Press.
- Goux, D., & Maurin, E. (1994). Education, expérience et salaire. *Économie & Prévision*, 116(5), 155–178. <https://doi.org/10.3406/ecop.1994.5704>
- Greenwood, M. J. (1997). Internal migration in developed countries. In M. R. Rosenzweig & O. Stark (Eds.), *Handbook of Population and Family Economics, Vol. 1 - Part B* (pp. 647–720). Elsevier. [https://doi.org/10.1016/S1574-003X\(97\)80004-9](https://doi.org/10.1016/S1574-003X(97)80004-9)
- Groot, W., & Maassen van den Brink, H. (2000). Overeducation in the labour market: A meta-analysis. *Economics of Education Review*, 21(2), 149–158. [https://doi.org/10.1016/S0272-7757\(99\)00057-6](https://doi.org/10.1016/S0272-7757(99)00057-6)
- Guironnet, J.-P. (2006). La suréducation en France : vers une dévalorisation des diplômes du supérieur? *Économie appliquée*, 59(1), 93–120.
- Gullason, E. T. (1989). The consumption value of schooling: An empirical estimate of one aspect. *Journal of Human Resources*, 287–298. <https://doi.org/10.2307/145857>
- Halaby, C. N. (1994). Overeducation and Skill Mismatch. *Sociology of Education*, 67(1), 47–59. <https://doi.org/10.2307/2112749>
- Hanushek, E. A., & Woessmann, L. (2020). Education, knowledge capital, and economic growth. In S. Bradley & C. Green (Eds.), *The Economics of Education: A Comprehensive Overview, Second Edition* (pp. 171–182). <https://doi.org/10.1016/B978-0-12-815391-8.00014-8>
- Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131–147. [https://doi.org/10.1016/S0272-7757\(99\)00050-3](https://doi.org/10.1016/S0272-7757(99)00050-3)
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 47(1), 153–161. <https://doi.org/10.2307/1912352>
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. Hanushek & F. Welch (Eds.),

- Handbook of the Economics of Education, Vol. 1* (pp. 307–458). Amsterdam: Elsevier Science. [https://doi.org/10.1016/S1574-0692\(06\)01007-5](https://doi.org/10.1016/S1574-0692(06)01007-5)
- Hensen, M. M., De Vries, M. R., & Cörvers, F. (2009). The role of geographic mobility in reducing education-job mismatches in the Netherlands. *Papers in Regional Science*, 88(3), 667–682. <https://doi.org/10.1111/j.1435-5957.2008.00189.x>
- Iammarino, S., & Marinelli, E. (2015). Education-job (mis)match and interregional migration: Italian university graduates' transition to work. *Regional Studies*, 49(5), 866–882. <https://doi.org/10.1080/00343404.2014.965135>
- Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, 23(3), 305–327. <https://doi.org/10.1002/jae.998>
- Ichino, A., Rustichini, A., & Zanella, G. (2022). College education, intelligence, and disadvantage: Policy lessons from the UK in 1960-2004. *CEPR Discussion Paper No. DP17284*. <https://ssrn.com/abstract=4121468>
- Ingram, B. F., & Neumann, G. R. (2006). The returns to skill. *Labour Economics*, 13(1), 35–59. <https://doi.org/10.1016/j.labeco.2004.04.005>
- Jauhainen, S. (2011). Overeducation in the Finnish regional labour markets. *Papers in Regional Science*, 90(3), 573–588. <https://doi.org/10.1111/j.1435-5957.2010.00334.x>
- Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5, Part 1), 972–990. <https://doi.org/10.1086/260808>
- Kahn, L. M., & Low, S. A. (1982). The relative effects of employed and unemployed job search. *The Review of Economics and Statistics*, 234–241. <https://doi.org/10.2307/1924302>
- Kain, J. F. (1968). Housing Segregation, Negro Employment, and Metropolitan Decentralization. *The Quarterly Journal of Economics*, 82(2), 175–197. <https://doi.org/10.2307/1885893>
- Kampelmann, S., Mahy, B., Rycx, F., & Vermeulen, G. (2020). Over-, Required, and Undereducation: Consequences on the Bottom Lines of Firms. *Labour*, 34(1), 80–112. <https://doi.org/10.1111/labr.12152>
- Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17, 99–121. <https://doi.org/10.1023/A:1007707430416>
- Kiefer, N. M. (1985). Evidence on the role of education in labor turnover. *Journal of Human Resources*, 20(3), 445–452. <https://doi.org/10.2307/145895>
- Koch, A., Nafziger, J., & Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115, 3–17. <https://doi.org/10.1016/j.jebo.2014.09.005>
- Kucel, A. (2011). Literature survey of the incidence of over-education: A sociological approach. *Revista Española de Investigaciones Sociológicas*, 1(134), 125–142. <https://www.jstor.org/stable/41304938>
- La Belle, T. J. (1982). Formal, nonformal and informal education: A holistic perspective on lifelong learning. *International Review of Education*, 28, 159–175. <https://doi.org/10.1007/BF00598444>

- 
- Langella, M., & Manning, A. (2022). Residential mobility and unemployment in the UK. *Labour Economics*, 75, 102104. <https://doi.org/10.1016/j.labeco.2021.102104>
- Lazear, E. (1977). Education: Consumption or Production? *Journal of Political Economy*, 85(3), 569–597. <https://doi.org/10.1086/260584>
- Le Barbanchon, T., & Rizzotti, N. (2020). *The task content of French jobs* [Available at SSRN 3653262]. <https://doi.org/10.2139/ssrn.3653262>
- Le Gallo, J. (2004). Hétérogénéité spatiale, principes et méthodes. *Économie & prévision*, 162(1), 151–172. <https://doi.org/10.3917/ecop.162.0151>
- Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In *Econometric evaluation of labour market policies* (pp. 43–58). Physica-Verlag HD. [https://doi.org/10.1007/978-3-642-57615-7\\_3](https://doi.org/10.1007/978-3-642-57615-7_3)
- Lemieux, T. (2006). The “Mincer equation” thirty years after *Schooling, Experience, and Earnings*. In S. Grossbard (Ed.), *Jacob Mincer a pioneer of modern labor economics* (pp. 127–145). Springer, Boston. [https://doi.org/10.1007/0-387-29175-X\\_11](https://doi.org/10.1007/0-387-29175-X_11)
- Lemistre, P., & Moreau, N. (2009). Spatial mobility and returns to education: Some evidence from a sample of French youth. *Journal of Regional Science*, 49(1), 149–176. <https://doi.org/10.1111/j.1467-9787.2008.00574.x>
- LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press.
- Lindley, J. (2009). The over-education of UK immigrants and minority ethnic groups: Evidence from the Labour Force Survey. *Economics of Education Review*, 28(1), 80–89. <https://doi.org/10.1016/j.econedurev.2007.11.003>
- Lochner, L. J. (2011). Nonproduction benefits of education: Crime, health, and good citizenship. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education, Vol. 4* (pp. 183–282). Amsterdam: Elsevier Science. <https://doi.org/10.1016/B978-0-444-53444-6.00002-X>
- Long, M. C. (2010). Changes in the returns to education and college quality. *Economics of Education review*, 29(3), 338–347. <https://doi.org/10.1016/j.econedurev.2009.10.005>
- Manning, A., & Petrongolo, B. (2017). How local are labor markets? Evidence from a spatial job search model. *American Economic Review*, 107(10), 2877–2907. <https://doi.org/10.1257/aer.20131026>
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), 531–542. <https://doi.org/10.2307/2298123>
- Marinescu, I., & Rathelot, R. (2018). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics*, 10(3), 42–70. <https://doi.org/10.1257/mac.20160312>
- McCall, J. J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics*, 84(1), 113–126. <https://doi.org/10.2307/1879403>
- McGoldrick, K., & Robst, J. (1996). Gender differences in overeducation: A test of the theory of differential overqualification. *American Economic Review*, 86(2), 280–284. <https://www.jstor.org/stable/2118137>
- McGuinness, S. (2006). Overeducation in the Labour Market. *Journal of Economic Surveys*, 20(3), 387–418. <https://doi.org/10.1111/j.0950-0804.2006.00284.x>

- McGuinness, S., Bergin, A., & Whelan, A. (2018). Overeducation in Europe: trends, convergence, and drivers. *Oxford Economic Papers*, 70(4), 994–1015. <https://doi.org/10.1093/oeqpgy022>
- Meliciani, V., & Radicchia, D. (2016). Informal networks, spatial mobility and overeducation in the Italian labour market. *The Annals of Regional Science*, 56, 513–535. <https://doi.org/10.1007/s00168-016-0752-y>
- Mendes de Oliveira, M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review*, 19(2), 199–206. [https://doi.org/10.1016/S0272-7757\(99\)00020-5](https://doi.org/10.1016/S0272-7757(99)00020-5)
- Meroni, E. C., & Vera-Toscano, E. (2017). The persistence of overeducation among recent graduates. *Labour Economics*, 48, 120–143. <https://doi.org/10.1016/j.labeco.2017.07.002>
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.
- Moran, P. A. (1950). A test for serial independence of residuals. *Biometrika*, 37(1/2), 178–181. <https://doi.org/10.2307/2332162>
- Murphy, K. M., Shleifer, A., & Vishny, R. W. (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics*, 106(2), 503–530. <https://doi.org/10.2307/2937945>
- Nauze-Fichet, E., & Tomasini, M. (2002). Diplôme et insertion sur le marché du travail : Approches socioprofessionnelle et salariale du déclassement suivi d'un commentaire de Saïd Hanchane et Eric Verdier. *Économie et Statistique*, 354(1), 21–44. <https://doi.org/10.3406/estat.2002.7383>
- Navarini, L., & Verhaest, D. (2023). Educational Attainment, Overeducation and Wages: Evidence from a Dynamic Model [Unpublished]. <https://lorenzonavarini.github.io/JM/Overeducation.pdf>
- Nickell, S. (1979). Education and lifetime patterns of unemployment. *Journal of Political Economy*, 87(5, Part 2), S117–S131. <https://doi.org/10.1086/260825>
- Nordin, M., Persson, I., & Rooth, D. O. (2010). Education-occupation mismatch: Is there an income penalty? *Economics of Education Review*, 29(6), 1047–1059. <https://doi.org/10.1016/j.econedurev.2010.05.005>
- OECD. (2020). *OECD Employment Outlook 2020: Worker Security and the COVID-19 Crisis*. Organisation for Economic Co-operation and Development Publishing, Paris. <https://doi.org/10.1787/1686c758-en>
- OECD. (2022). *Education at a Glance 2022: OECD indicators*. Organisation for Economic Co-operation and Development Publishing, Paris. <https://doi.org/10.1787/3197152b-en>
- Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358–369. <https://doi.org/10.1080/00401706.2013.806694>
- Ortiz, L., & Kucel, A. (2008). Do fields of study matter for over-education? The cases of Spain and Germany. *International Journal of Comparative Sociology*, 49(4-5), 305–327. <https://doi.org/10.1177/0020715208093079>
- Patacchini, E., & Zenou, Y. (2007). Spatial dependence in local unemployment rates. *Journal of Economic Geography*, 7(2), 169–191. <https://doi.org/10.1093/jeg/lbm001>

- 
- Piper, A. (2015). Heaven knows I'm miserable now: overeducation and reduced life satisfaction. *Education Economics*, 23(6), 677–692. <https://doi.org/10.1080/09645292.2013.870981>
- Pissarides, C. A., & McMaster, I. (1990). Regional migration, wages and unemployment: Empirical evidence and implications for policy. *Oxford Economic Papers*, 42(4), 812–831. <https://doi.org/10.1093/oxfordjournals.oep.a041980>
- Power, T. M. (2005). The supply and demand for natural amenities: An overview of theory and concepts. In G. P. Green, S. C. Deller, & D. W. Marcouiller (Eds.), *Amenities and rural development: Theory, methods and public policy* (pp. 63–77). Edward Elgar Publishing Limited Northampton, MA.
- Quintini, G. (2011a). Over-qualified or under-skilled: A review of existing literature. *OECD Social, Employment and Migration Working Papers*, (121). <https://doi.org/10.1787/5kg58j9d7b6d-en>
- Quintini, G. (2011b). Right for the job: Over-qualified or under-skilled? *OECD Social, Employment and Migration Working Papers*, (120). <https://doi.org/10.1787/5kg59fcz3tkd-en>
- Ramos, R., & Sanromá, E. (2013). Overeducation and local labour markets in Spain. *Tijdschrift voor Economische en Sociale Geografie*, 104(3), 278–291. <https://doi.org/10.1111/j.1467-9663.2012.00752.x>
- Rios, V. (2017). What drives unemployment disparities in European regions? a dynamic spatial panel approach. *Regional Studies*, 51(11), 1599–1611. <https://doi.org/10.1080/00343404.2016.1216094>
- Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6), 1257–1278. <https://www.jstor.org/stable/1830947>
- Robst, J. (1995). College quality and overeducation. *Economics of Education Review*, 14(3), 221–228. [https://doi.org/10.1016/0272-7757\(94\)00054-A](https://doi.org/10.1016/0272-7757(94)00054-A)
- Robst, J. (2007a). Education and job match: The relatedness of college major and work. *Economics of Education Review*, 26(4), 397–407. <https://doi.org/10.1016/j.econedurev.2006.08.003>
- Robst, J. (2007b). Education, college major, and job match: Gender differences in reasons for mismatch. *Education Economics*, 15(2), 159–175. <https://doi.org/10.1080/09645290701263070>
- Robst, J. (2008). Overeducation and college major: Expanding the definition of mismatch between schooling and jobs. *The Manchester School*, 76(4), 349–368. <https://doi.org/10.1111/j.1467-9957.2008.01064.x>
- Rodríguez-Pose, A., & Tselios, V. (2011). Mapping the European regional educational distribution. *European Urban and Regional Studies*, 18(4), 358–374. <https://doi.org/10.1177/0969776411399345>
- Romaní, J., Casado-Díaz, J. M., & Lillo-Banuls, A. (2016). On the links between spatial variables and overeducation. *Applied Economics Letters*, 23(9), 652–655. <https://doi.org/10.1080/13504851.2015.1095996>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rossen, A., Boll, C., & Wolf, A. (2019). Patterns of overeducation in Europe: The role of field of study. *IZA Journal of Labor Policy*, 9(1). <https://doi.org/10.2478/izajolp-2019-0003>

- Rubb, S. (2003). Overeducation: a short or long run phenomenon for individuals? *Economics of Education Review*, 22(4), 389–394. [https://doi.org/10.1016/S0272-7757\(02\)00052-3](https://doi.org/10.1016/S0272-7757(02)00052-3)
- Rumberger, R. W. (1981). The rising incidence of overeducation in the US labor market. *Economics of Education Review*, 1(3), 293–314. [https://doi.org/10.1016/0272-7757\(81\)90001-7](https://doi.org/10.1016/0272-7757(81)90001-7)
- Rupert, P., & Wasmer, E. (2012). Housing and the labor market: Time to move and aggregate unemployment. *Journal of Monetary Economics*, 59(1), 24–36. <https://doi.org/10.1016/j.jmoneco.2011.10.008>
- Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31, 831–880. <http://www.jstor.org/stable/2728516>
- Scherer, S. (2004). Stepping-stones or traps? The consequences of labour market entry positions on future careers in West Germany, Great Britain and Italy. *Work, employment and society*, 18(2), 369–394. <https://doi.org/10.1177/09500172004042774>
- Schmutz, B., Sidibé, M., & Vidal-Naquet, E. (2021). Why are low-skilled workers less mobile? The role of mobility costs and spatial frictions. *Annals of Economics and Statistics*, (142), 283–304. <https://doi.org/10.15609/annaeconstat2009.142.0283>
- Sellami, S., Verhaest, D., Nonneman, W., & Van Trier, W. (2020). Education as investment, consumption or adapting to social norm: Implications for educational mismatch among graduates. *Education Economics*, 28(1), 26–45. <https://doi.org/10.1080/09645292.2019.1680955>
- Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programs in the 1990s. *Review of Economics and Statistics*, 86(1), 133–155. <https://doi.org/10.1162/003465304323023723>
- Sianesi, B. (2008). Differential effects of active labour market programs for the unemployed. *Labour Economics*, 15(3), 370–399. <https://doi.org/10.1016/j.labeco.2007.04.004>
- Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, 98(1), 169–192. <https://doi.org/10.1086/261674>
- Simpson, W. (1992). *Urban structure and the labour market. Worker mobility, commuting and underemployment in cities*. Clarendon Press, Oxford.
- Sloane, P., Battu, H., & Seaman, P. (1999). Overeducation, undereducation and the British labour market. *Applied Economics*, 31, 1437–1453. <https://doi.org/10.1080/000368499323319>
- Somers, M. A., Cabus, S. J., Groot, W., & Maassen van den Brink, H. (2019). Horizontal mismatch between employment and field of education: Evidence from a systematic literature review. *Journal of Economic Surveys*, 33(2), 567–603. <https://doi.org/10.1111/joes.12271>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Stakhovych, S., & Bijmolt, T. H. (2009). Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2), 389–408. <https://doi.org/10.1111/j.1435-5957.2008.00213.x>
- Stigler, G. J. (1962). Information in the Labor Market. *Journal of Political Economy*, 70(5, Part 2), 94–105. <https://doi.org/10.1086/258727>
- Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews (Ed.), *Identification and inference for econometric models* (pp. 80–108). Cambridge University Press. <https://doi.org/10.3386/t0284>



- 
- Svalund, J., & Berglund, T. (2018). Fixed-term employment in Norway and Sweden: A pathway to labour market marginalization? *European Journal of Industrial Relations*, *24*(3), 261–277. <https://doi.org/10.1177/0959680117740777>
- Thurow, L. C. (1975). *Generating inequality*. Basic Books: New York.
- Tsang, M. C. (1987). The impact of underutilization of education on productivity: A case study of the US Bell companies. *Economics of Education Review*, *6*(3), 239–254. [https://doi.org/10.1016/0272-7757\(87\)90003-3](https://doi.org/10.1016/0272-7757(87)90003-3)
- Tsang, M. C., Rumberger, R. W., & Levin, H. M. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations*, *30*(2), 209–228. <https://doi.org/10.1111/j.1468-232X.1991.tb00786.x>
- Tselios, V. (2013). Overeducation and externalities in the EU: The combined moderating influence of migration and gender. *Environment and Planning C: Government and Policy*, *31*(2), 193–221. <https://doi.org/10.1068/c1207r>
- Van Belle, E., Di Stasio, V., Caers, R., De Couck, M., & Baert, S. (2018). Why are employers put off by long spells of unemployment? *European Sociological Review*, *34*(6), 694–710. <https://doi.org/10.1093/esr/jcy039>
- van der Meer, P. H. (2006). The validity of two education requirement measures. *Economics of Education Review*, *25*(2), 211–219. <https://doi.org/10.1016/j.econedurev.2005.02.004>
- van Ham, M. (2001). Workplace mobility and occupational achievement. *International Journal of Population Geography*, *7*(4), 295–306. <https://doi.org/10.1002/ijpg.225>
- Venhorst, V. A., & Cörvers, F. (2018). Entry into working life: Internal migration and the job match quality of higher-educated graduates. *Journal of Regional Science*, *58*(1), 116–140. <https://doi.org/10.1111/jors.12347>
- Venhorst, V. A., van Dijk, J., & van Wissen, L. (2010). Do the best graduates leave the peripheral areas of the Netherlands? *Tijdschrift voor economische en sociale geografie*, *101*(5), 521–537. <https://doi.org/10.1111/j.1467-9663.2010.00629.x>
- Verdugo, R. R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, 629–643. <https://doi.org/10.2307/145998>
- Verhaest, D., & Omeij, E. (2006). Discriminating between alternative measures of over-education. *Applied Economics*, *38*(18), 2113–2120. <https://doi.org/10.1080/00036840500427387>
- Verhaest, D., Sellami, S., & van der Velden, R. (2017). L'inadéquation horizontale et verticale entre compétences et emploi selon le pays et le domaine d'études. *Revue internationale du Travail*, *156*(1), 1–26. <https://doi.org/10.1111/ilrf.12043>
- Verhaest, D., & van der Velden, R. (2013). Cross-country differences in graduate overeducation. *European Sociological Review*, *29*(3), 642–653. <https://doi.org/10.1093/esr/jcs044>
- Vornholt, K., Villotti, P., Muschalla, B., Bauer, J., Colella, A., Zijlstra, F., Van Ruitenbeek, G., Uitdewilligen, S., & Corbière, M. (2018). Disability and employment - overview and highlights. *European Journal of Work and Organizational Psychology*, *27*(1), 40–55. <https://doi.org/10.1080/1359432X.2017.1387536>

- Voßemer, J., & Schuck, B. (2016). Better overeducated than unemployed? The short- and long-term effects of an overeducated labour market re-entry. *European Sociological Review*, 32(2), 251–265. <https://doi.org/10.1093/esr/jcv093>
- Wolbers, M. H. J. (2003). Job mismatches and their labour-market effects among school-leavers in Europe. *European Sociological Review*, 19(3), 249–266. <https://doi.org/10.1093/esr/19.3.249>
- Wooldridge, J. M. (2002). *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning: Mason, Ohio.

**Titre :** Sur-éducation et trajectoires sur le marché du travail

**Mots-clés :** sur-éducation ; inadéquation emploi-diplôme ; mobilité spatiale ; mobilité professionnelle ; transitions éducation-emploi

**Résumé :** La sur-éducation correspond à la situation d'individus dont le niveau d'éducation est a priori supérieur à celui requis pour l'emploi qu'ils occupent. Touchant entre un quart et un tiers des travailleurs des pays de l'OCDE, la sur-éducation a des effets négatifs tant au niveau individuel que collectif.

Cette thèse comporte des contributions empiriques relatives aux liens entre la sur-éducation et la mobilité, qu'elle soit spatiale ou professionnelle. Plus précisément, elle vise à évaluer dans quelle mesure la mobilité peut réduire la sur-éducation, ainsi que les limites associées. D'un point de vue géographique, cela conduit à explorer les disparités locales et régionales des marchés du travail et les liens entre inadéquation territoriale et professionnelle. Sur le plan professionnel, cette thèse cherche à

explorer l'effet de « trappe » du *mismatch* et sur la capacité des individus à sortir d'un emploi inadapté.

À travers plusieurs analyses économétriques, cette thèse étudie la sur-éducation, d'une part du point de vue de la géographie des marchés du travail locaux (chapitre 2) et de la mobilité spatiale (chapitre 3) ; d'autre part, en termes de mobilité professionnelle et de trajectoires de carrière (chapitre 4). En amont de ces analyses, le chapitre 1 aborde des questions méthodologiques liées à la mesure de la sur-éducation. En outre, le dernier chapitre explore une définition plus large de l'inadéquation emploi-diplôme, en abordant l'inadéquation horizontale, c'est-à-dire la situation d'individus qui ne travaillent pas dans le domaine de leur diplôme.

**Title:** Overeducation and labour market trajectories

**Keywords:** overeducation; educational mismatch; spatial mobility; occupational mobility; school-to-work transitions

**Abstract:** Overeducation is defined as the situation of individuals having a level of education *a priori* higher than the one required for the job they occupy. Affecting between one quarter and one third of the workers in OECD countries, overeducation has detrimental effects both at the individual and collective levels.

This thesis makes empirical contributions relative to the links between overeducation and mobility, whether spatial or professional. More specifically, it aims to assess the extent to which mobility can alleviate overeducation, as well as the issues and limitations that this might encounter. From the geographical point of view, this leads to explore local and regional labour markets disparities, and the links between territorial and professional inadequacy. From the

professional point of view, this thesis focuses on the trap effect of mismatch and the ability of individuals to exit from unsuitable job positions. Through several econometric analyses, this thesis investigates overeducation, on the one hand, from the perspective of the geography of local labour markets (chapter 2) and spatial mobility (chapter 3); on the other hand, in terms of occupational mobility and career trajectories (chapter 4). Prior to these analyses, chapter 1 addresses methodological issues related to the measurement of overeducation. Besides, the last chapter considers a larger definition of job-education mismatches, addressing horizontal mismatch, *i.e.* the situation of individuals not working in their field of study.