



HAL
open science

Parenthood, norms and public policies : Three essays in gender economics

Adèle Lemoine

► **To cite this version:**

Adèle Lemoine. Parenthood, norms and public policies : Three essays in gender economics. Economics and Finance. Université Paris sciences et lettres, 2023. English. NNT : 2023UPSLD037 . tel-04405504

HAL Id: tel-04405504

<https://theses.hal.science/tel-04405504>

Submitted on 19 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.



THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à Université Paris-Dauphine

Parenthood, Norms and Public Policies

Three Essays in Gender Economics

Soutenue par

Adèle LEMOINE

Le 20 Décembre 2023

École doctorale n°ED 543

SDOSE

Spécialité

Sciences économiques

Composition du jury :

Sarah FLÈCHE Chargée de recherche, Centre National de Recherche Scientifique Université Paris 1 Panthéon-Sorbonne	<i>Rapporteure</i>
Jürgen MAURER Professeur, Université de Lausanne	<i>Président du jury</i>
Peter EIBICH Professeur, Université Paris Dauphine - PSL	<i>Examineur</i>
Libertad GONZÁLEZ Professeure associée, Université Pompeu Fabra	<i>Examinatrice</i>
Marion LETURCQ Chargée de recherche, Institut National d'Etudes Démographiques	<i>Examinatrice</i>
Éric BONSANG Professeur, Université Paris Dauphine - PSL	<i>Directeur de thèse</i>

L'Université Paris Dauphine - PSL n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Remerciements

Il y a quelques années, je débutais cette thèse sans savoir exactement ce qui m’attendait sur son chemin. Aujourd’hui, j’arrive au terme de cette aventure et mesure la chance qui m’a été donnée de la vivre. Avec ces quelques mots, je souhaiterais remercier toutes les personnes qui m’ont accompagnée durant ce périple.

Tout d’abord, vous ne seriez pas en train de lire cette thèse si je n’avais pas reçu le soutien de mon directeur de thèse. Eric, je ne te remercierai jamais assez de m’avoir accordé ta confiance en acceptant d’encadrer ce travail et d’être co-auteur d’un des chapitres de cette thèse. Merci pour ce que tu m’as transmis durant toutes ces années avec un juste mélange d’exigence et de bienveillance. Je n’oublie pas ton sens de l’humour qui a toujours été le bienvenu dans nos échanges, ni tes mots de réconfort dans les moments de doute. Je suis honorée d’avoir parcouru ce chemin sous ton aile.

Je remercie Sarah Flèche et Jürgen Maurer d’avoir accepté d’être les rapporteur·e-s de cette thèse. Our discussions during the pre-defense led me to the completion of this work. Many thanks for your careful reading and your support. Mes remerciements s’adressent également aux examinateur·rice-s de cette thèse, Peter Eibich, Libertad González et Marion Leturcq. Peter, thanks for taking time to answer my questions at the lab and for sharing your insights on German data. Marion, merci pour nos échanges récents, toujours très constructifs et emplis de bienveillance. Libertad, we did not have the chance to discuss about research yet, but I am sure your comments will greatly improve this work.

Toute ma gratitude s’adresse également à Julien Bergeot, co-auteur d’un des chapitres de cette thèse et jeune chercheur dont les qualités ne sont plus à prouver. Je te remercie sincèrement pour tous les bons conseils dispensés avant même le début de la thèse et pendant ton passage à Dauphine. J’ai beaucoup appris de notre collaboration et je t’en remercie sincèrement. Merci de m’avoir accueillie en Italie dans le cadre de notre travail commun et de m’avoir fait découvrir Venise. I do not forget Agar Brugiavini and Giacomo Pasini who kindly hosted me at Ca’Foscari.

Je souhaiterais remercier chaleureusement les membres de mon comité de suivi de thèse, Eve Caroli et Brigitte Dormont, pour leurs retours constructifs sur mes travaux ainsi que pour leurs encouragements. Eve, je te remercie du fond du coeur pour tes précieux conseils qui m’ont guidée dans les moments de doute. J’ai eu une chance inouïe de croiser ton chemin. Brigitte, merci pour ton accompagnement tout au long de la thèse et pour ton soutien envers les Walking Docs !

En parallèle des travaux que je présente aujourd’hui dans cette thèse, j’ai eu la chance de travailler avec Florence Jusot dont j’admire les qualités tant humaines qu’académiques. Florence, ta bienveillance et ton rire communicatif ont rendu cette collaboration riche et lumineuse. En tant que professeure d’abord puis comme co-auteure ensuite, tu as su me transmettre ton goût pour la recherche sur les questions d’inégalités de santé. Je t’en suis infiniment reconnaissante.

J'en profite pour remercier les membres du projet EquiRAC qui ont accompagné nos recherches.

Cette thèse a été menée au Laboratoire d'Economie de Dauphine (LEDA) dont je remercie les membres avec qui j'ai pu échanger pendant ces quatre années. Merci en particulier à Elise Huillery, Eve Caroli, Farid Toubal et Anne Epaulard d'avoir, à tour de rôle, dirigé le programme doctoral et à Sidartha Gordon de donner les clés aux doctorant·e-s pour l'après-thèse. Le LEDA ne pourrait pas exister sans l'aide précieuse d'Eva Annane, de Cécile Billois-Fontalavie et de Vanina Youmbi. Votre bonne humeur va me manquer. Merci également à Karène Alarich-Val qui faisait partie de l'équipe administrative lors de mes premières années de thèse. Je souhaite également remercier Daniel Herrera-Araujo et Romain Plassard de m'avoir offert ma première expérience en tant que co-organisatrice du séminaire LEDA. Daniel, je te suis particulièrement reconnaissante d'avoir sans cesse porté attention à l'avancement de ma thèse et de m'avoir initiée à l'économie expérimentale que j'espère mettre en pratique un jour.

Le LEDA ne serait tout simplement le même sans ses (plus ou moins) nouvelles recrues : Romain Gâté, Mathilde Godard, Loïc Henry, Laura Khoury, Elsa Perdrix, Claire Rimbaud et Arthur Thomas. Mathilde, merci beaucoup pour le temps que tu m'as accordé pour discuter de mes travaux. Arthur, Claire, Laura, Loïc et Romain, merci pour vos sourires et les moments partagés autour d'une bière ou d'un café. Elsa, merci infiniment pour tes précieux conseils, tes relectures, les pédibus et ta joie de vivre qui égaye le labo.

Bien sûr, je n'oublie pas les personnes qui m'ont mis le pied à l'étrier, notamment Anne-Laure Samson qui m'a permis d'intégrer le Master d'Economie de la santé à Dauphine et m'a ouvert la voie vers mon premier stage de recherche avec Anne Laferrère, que j'en profite pour remercier chaleureusement. Mes remerciements s'adressent également à Clémentine Garrouste qui a encadré mon tout premier projet de recherche lors du mémoire de M1. Enfin, je porte une attention toute particulière à remercier l'équipe française de l'enquête SHARE qui a accompagné mes premiers pas dans le monde de la recherche. Merci à Benjamin et Louis de m'avoir accueillie et d'avoir rendu mon intégration si facile. Constance, merci d'avoir été ma première acolyte de stage. Justine, merci pour ta générosité et ton amitié. Courage, le bout du tunnel est bientôt là pour toi aussi ! Thomas, je te remercie d'avoir toujours été là pour moi, dans les bons moments comme dans les plus difficiles. Merci également à tou·te·s les chercheur·e-s qui m'ont conseillée et que j'ai toujours plaisir à croiser en conférence ou ailleurs : Elena Bassoli, Damien Bricard, Paul Dourgnon, Roméo Fontaine, Sandrine Juin, Hélène Le Forner, Anne-Kim Ristori, Lise Rochaix, Julie Tréguier, Quitterie Roquebert et Jérôme Wittwer.

Cette thèse n'aurait évidemment pas eu la même saveur sans les doctorant·e-s du LEDA. Arthur, je te remercie pour tous ces moments passés ensemble depuis le Master, les éclats de rire et ton réconfort dans les moments compliqués. Cécile, merci d'avoir toujours été à mes côtés (en C104G, en soirée et en voyage !). Ta présence a beaucoup compté pour moi. Benoît, fidèle compagnon de thèse, d'aventure sportive et d'apéro. Je te remercie pour ton amitié. Cette thèse a été écrite en grande partie dans le bureau C104G où j'y ai trouvé de fantastiques co-bureaux. Cédric, pilier du C104G et fournisseur officiel de joie au bureau, en vélo comme en soirée ! L'aventure

n'aurait pas été aussi belle sans toi. Alexis, happy-culteur des doctorant·e·s et compagnon de télé-travail-café. Côme, qui a tout de suite mis le feu dans le bureau (et beaucoup de sucre dans nos veines), Chloé mon acolyte du vendredi, Coline ma copine nordiste, Lucille mon soutien de fin de thèse et Anatole, la relève du C104G. Merci aussi à Léopold et Inès de m'avoir ouvert le P202 dans les moments de solitude au 1er étage et pour tous les bons moments passés ensemble. Mathilde L., merci d'avoir insufflé la création des Walking Docs et d'avoir rendu ce défi possible. Julien D., qui aurait cru qu'on se retrouverait en thèse ensemble lorsqu'on était en L2 ? Merci enfin à tou·te·s les autres : Amélie, Anthonin, Antonio, Eléonore, Fatou, Maeva, Malte, Matteo, Pauline, Paul-Emile, Paul B., Raphaël et Zélie. J'ai également une pensée pour les ancien·ne·s doctorant·e·s qui ont montré la voie : Alexis D., Christian, Doriane et Sandra (les ancien·ne·s du LEGOS), Aimée, Anne et Noémie M. (les un peu moins anciennes du LEGOS) et aussi Charlotte, Emy, Eugénie, Leslie, Morgan et Noémie C. Yohan, merci pour ton aide précieuse, autant en économétrie que pour les procédures administratives. Paul C, merci de m'avoir légué (en toute objectivité) le meilleur bureau du C104G ! A tou·te·s les autres dont j'ai pu croiser la route : Merci à vous. A tou·te·s les nouv·eaux·elles arrivant·e·s que je n'ai eu le temps de mieux connaître : Courage ! Le chemin est sinueux, mais il est aussi plein de bonnes surprises.

Lorsque la thèse était une source de stress, j'ai eu la chance de pouvoir compter sur Fans de terre pour m'évader. Merci Bénédicte et Monika de m'avoir initiée à la céramique et de continuer à me transmettre votre passion. Merci également aux céramistes du jeudi soir pour votre bonne humeur, vos créations inspirantes et vos *"Alors ça y est, tu as rendu ta thèse ?"*.

Je tiens à remercier mes ami·e·s sur lesquel·le·s je peux compter, certain·e·s depuis plus de 10 ans. Aude, merci pour ton indéfectible soutien, ton indulgence et pour ta force que j'ai toujours admirée. Chaïma, mon petit rayon de soleil sans qui la vie serait bien plus terne. Courage, toi aussi tu seras bientôt Docteure (la vraie, avec un stéthoscope). Simon B., merci pour ton amitié qui ne s'étiole pas malgré les années qui passent, parfois sans se voir. Antoine, merci de m'avoir supportée pendant toute la thèse et d'avoir toujours répondu présent. Ton amitié m'est très chère. Merci également à Ginie et Nico pour les bons moments partagés, toujours autour de bons petits plats !

Enfin, cette thèse n'aurait jamais pu voir le jour sans mes proches dont je chéris la présence. Merci à mes frères adorés, les personnes les plus fortes que je connaisse. Simon, merci d'avoir ouvert la voie et d'avoir toujours été là pour protéger tes petits frère et soeur. Paul, merci pour tes encouragements, tes conseils et tes pâtisseries réconfortantes. Merci également à José et Céleste d'apporter la joie dont une famille a besoin (et toujours plein de bonnes choses à manger !). Je remercie ma mère Christine de m'avoir permis de poursuivre mes études à Paris, Sylvie, Philippe, Hugo et Frédérique pour leur soutien. Bien sûr, je ne serais pas en train d'écrire ces mots sans toi, Greg. Je te remercie d'avoir toujours cru en moi et de m'avoir rendu la vie si facile à la maison (alors que j'écrivais sur l'inégale répartition des tâches au sein du couple...oups). C'est grâce à ta patience et ta tendresse que j'ai pu arriver au bout de ce travail.

A vous tou·te·s : Merci.

Résumé

Au cours du XXe siècle, la place des femmes dans l'économie occidentale a connu de profonds changements. Pourtant, d'importants écarts liés au genre persistent aujourd'hui sur le marché du travail et d'autant plus dans la sphère domestique. Cette thèse propose un éclairage sur ces inégalités avec trois analyses empiriques basées sur des données d'enquête.

Le premier chapitre s'intéresse à la contribution de la parentalité aux inégalités femme-homme en matière de répartition du temps et de bien-être en Allemagne. Dans un premier temps, l'évolution de ces caractéristiques parentales est analysée autour de la première naissance à l'aide d'une *étude d'évènement* qui exploite la variation du *timing* de la naissance. Les résultats montrent que la naissance du premier enfant constitue un facteur aggravant de l'inégale répartition des tâches au sein du couple. Après la première naissance, les femmes réallouent du temps de travail sur le marché au profit de tâches non-rémunérées alors que l'allocation du temps des pères reste pratiquement inchangée. En explorant l'effet de la parentalité sur le bien-être parental, ce chapitre montre que les femmes déclarent une baisse de satisfaction avec les tâches non-rémunérées moindre que celle des hommes alors qu'elles en portent majoritairement la charge. Ce résultat illustre ainsi un ancrage profond des normes de genre au sein du couple. Dans un deuxième temps, l'effet du nombre d'enfants est estimé à l'aide d'une *approche par variable instrumentale*. Les résultats suggèrent qu'au-delà de l'entrée dans la parentalité, le nombre d'enfants intensifie l'inégale répartition des tâches. Cette analyse montre également que les femmes d'au moins trois enfants déclarent davantage d'insatisfaction avec leur temps de loisir relativement à celles qui ont moins d'enfants alors que le temps de loisir reste significativement inchangé à l'arrivée d'un enfant supplémentaire, ce qui suggère une qualité dégradée du temps de loisir des femmes de familles nombreuses. La contribution de cette étude réside en deux points principaux. Tout d'abord, elle explore le rôle de la parentalité dans la répartition du travail non-rémunéré qui est souvent négligé au profit de l'analyse sur le travail rémunéré et ses caractéristiques. Pourtant, le travail non-rémunéré représente une part importante de la production totale qui est en majeure partie fournie par des femmes. Ensuite, elle questionne l'effet de la parentalité sur la satisfaction relative à ces différentes tâches, mettant ainsi en exergue la persistance des normes de genre au sein du couple.

Le deuxième chapitre est un travail commun avec Julien Bergeot. Dans cette étude, nous évaluons l'effet de l'entrée à l'école pré-élémentaire sur l'offre de travail des mères dans une perspective comparative entre pays européens. En raison du caractère non-aléatoire de la décision d'inscription à l'école pré-élémentaire par les mères, nous exploitons la relation discontinue entre l'éligibilité à l'école et l'année de naissance de l'enfant afin de prédire l'inscription à l'aide d'une *régression par discontinuité*. A l'aide de données d'enquête européennes harmonisées, nous montrons que les programmes éducatifs précoces représentent un levier efficace au travail des mères dans les pays où les normes de genre conservatrices dominent ainsi que dans les pays où la part

des dépenses publiques allouée à l'offre de modes de garde est plus faible. Nos résultats suggèrent la nécessité d'une offre publique de modes de garde accessible et abordable afin d'annihiler la contribution de la parentalité aux inégalités femme-homme sur le marché du travail, d'autant plus dans les pays où la participation des femmes au marché du travail est plus faible et l'offre de mode de garde intervenant avant les programmes éducatifs est moins développée.

Le troisième chapitre, co-écrit avec Eric Bonsang, s'intéresse à la transmission inter-générationnelle des normes de genre et son implication à long terme dans les écarts de santé cognitive entre les femmes et les hommes d'âge avancé. Les troubles cognitifs représentent l'un des principaux défis des sociétés vieillissantes, il est donc essentiel de comprendre les déterminants du fonctionnement cognitif des femmes et des hommes. Nous recourons à l'*approche épidémiologique* qui permet d'identifier la contribution des normes sociales en gardant le cadre institutionnel fixe. Concrètement, nous utilisons la variation des normes de genre entre pays de naissance des parents parmi des individus vivant dans un pays de résidence donné et dont au moins un parent est né dans un autre pays. Nous constatons que, parmi les individus âgés d'au moins 50 ans interrogés dans le pays de résidence, le fonctionnement cognitif des femmes diminue relativement à celui des hommes lorsque les normes de genre sont plus conservatrices dans le pays de naissance parental. L'exploration des mécanismes potentiels montre que cette relation s'explique au moins partiellement par des parcours professionnels différents entre les femmes et les hommes. Enfin, nous montrons que cet écart cognitif pourrait avoir d'autres implications pour la santé, puisque les femmes âgées de plus de 65 ans déclarent plus de difficultés dans les activités de la vie quotidienne que leurs homologues masculins lorsqu'elles sont exposées à des normes de genre plus conservatrices. Nos résultats constituent un argument supplémentaire à la lutte contre les stéréotypes de genre qui, en plus d'empêcher la convergence des femmes et des hommes dans le paysage économique, contribuent à dégrader l'état de santé des femmes.

Summary

Over the course of the 20th century, women's role in the Western economy underwent profound changes. However, major gender gaps persist on the labour market and even more so in the domestic sphere. This thesis sheds light on these inequalities with three empirical analyses based on survey data.

The first chapter explores the contribution of parenthood to gender inequalities in time-allocation and well-being in Germany. First, the evolution of these parental characteristics is analysed around the first birth using an *event study* analysis that exploits variations in birth *timing*. Results show that the first child birth emphasizes the unequal division of tasks within couples. After the first birth, women reallocate working time on the market to unpaid tasks, while fathers' allocation of time remains virtually unchanged. By exploring the effect of parenthood on parental well-being, this chapter shows that women report a lower drop in satisfaction with unpaid work than men despite the fact that they bear the major load of domestic work. This result suggests that gender norms are deeply integrated by individuals. Secondly, the effect of the number of children is estimated with an *instrumental variable approach*. Results suggest that, beyond the onset of parenthood, the number of children intensifies pre-existing inequalities in time-allocation at the couple level. This analysis also shows that women with at least three children report greater dissatisfaction with leisure time than those with fewer children, while the amount of leisure time remains significantly unchanged with the arrival of an additional child. This finding suggests a deteriorated quality of leisure time for women with large families. The main contribution of this study is twofold. First, it explores the role of parenthood in the distribution of unpaid work within couples, which is often neglected in favour of paid work and its characteristics in empirical analyses. Yet unpaid work accounts for a significant proportion of total production, most of which is provided by women. Second, it focuses on the effect of parenthood on satisfaction with these different tasks, highlighting the persistence of gender norms within couples.

The second chapter is a joint work with Julien Bergeot. In this study, we examine the effect of pre-school enrollment on mothers' labour supply in a comparative perspective between European countries. Given that mothers' pre-school enrollment decisions are unlikely to be random, we exploit the discontinuous relationship between pre-school eligibility and the child's year of birth to predict enrollment using a *fuzzy regression on discontinuity design*. Using harmonised European survey data, we show that pre-school programs represent an effective lever to mothers' labour force in countries where conservative gender norms dominate as well as in countries where the share of public spending allocated to the provision of early childcare is lower. Our results support the need for accessible and affordable public childcare provision in order to counteract the contribution of parenthood to gender inequalities on the labour market, especially in countries where women's labour supply is lower and early childcare provision is poorer.

The third chapter, co-written with Eric Bonsang, focuses on the inter-generational transmission of gender norms and its long-term impact on cognitive health differences between women and men in old age. Cognitive impairment represents one of the major challenges of ageing societies. Thus, it is necessary to target the determinants of women's and men's cognitive functioning. We use an *epidemiological approach* to identify the contribution of social norms keeping the institutional framework fixed. Specifically, we use the variation in gender norms between parents' countries of birth among individuals living in a given country of residence and who have at least one parent born in another country. We find that, among individuals aged at least 50 in a given country of residence, women's cognitive scores decrease compared to men's when gender norms are more conservative in the parental country of birth. Exploration of potential mechanisms shows that this relationship can be at least partly attributed to gender differences in career paths. Finally, we show that the cognitive gap could have further implications for health since women aged over 65 report more difficulties with activities of daily living than their male counterparts when they are exposed to more conservative gender norms. Our results provide an additional argument for fighting against gender stereotypes which, in addition to preventing the convergence of women and men in the economic landscape, contribute to deteriorate women's health.

Table des matières

Remerciements	vi
Résumé	ix
Summary	xiii
Introduction	2
1 Gender economic convergence : features and driving forces	6
1.1 The rise in female labour force	6
1.2 Public policies as tools to empower women	7
1.3 Unpaid work : a similar convergence ?	8
2 Convergence is not equality	9
2.1 Gender gaps remain on the labour market	9
2.1.1 Horizontal segregation	10
2.1.2 Vertical segregation	12
2.2 Men and women still allocate time differently	12
2.2.1 The child(ren) penalty	13
2.2.2 When gender norms stick	15
3 This dissertation	17
3.1 Contributions	17
3.2 Data and methods	19
Bibliography	22
1 Parenthood, Time-Allocation and Well-being	34
1 Introduction	40
2 The German context	44
3 Methodology	46
3.1 The first birth penalty	46

3.2	The effect of family size	48
3.2.1	Standard linear regression	49
3.2.2	Fixed effects model	50
3.2.3	Instrumental variable strategy	50
4	Data	54
4.1	Sample restrictions	54
4.2	Description of variables	55
5	Results	57
5.1	The first birth penalty	57
5.2	Family size effects	62
5.2.1	Main results	62
5.2.2	Internal validity	65
5.2.3	Heterogeneous results	67
5.2.4	Additional results	69
6	Conclusion	72
	Appendix	81
A	Identification tests	81
B	Descriptive statistics	84
C	First birth penalties	86
E	Internal validity	89
E.1	Robustness checks	89
E.2	Exclusion restriction test	90
F	Further results	91
G	Complier analysis	93
G.1	Methodology	93
G.2	Results	94
2	Pre-school Enrollment and Mothers' Labour Supply : Evidence from Europe	96
1	Introduction	102
2	Institutional framework	104
2.1	Early childcare provision	104
2.2	Pre-school system	105
3	Methodology	108
4	Data	114
4.1	Sample restriction	114
4.2	Description of variables	115
5	Estimation results	118
5.1	Average effects	118

5.1.1	Mothers' labour supply	118
5.1.2	Childcare use	121
5.2	Heterogeneous effects with respect to individual characteristics	123
5.2.1	Level of education	123
5.2.2	Partnership status	124
5.2.3	Birth order	125
5.3	Heterogeneous effects with respect to country level characteristics	125
5.3.1	Gender norms	125
5.3.2	Early childcare characteristics	127
5.3.3	Analysis by country	130
6	Conclusion	130
	Appendix	137
A	Childcare systems in Europe	137
B	Descriptive statistics by country	139
C	Identification tests	141
D	Robustness checks	143
E	Heterogeneous effects with respect to individual characteristics	145
E.1	Methodology	145
E.2	Results	146
F	Heterogeneous effects with respect to country characteristics	152
F.1	Methodology	152
F.2	Measures at the country level	153
F.2.1	Gender norms	153
F.2.2	Family benefits	153
F.2.3	Net childcare costs	154
F.3	Results	158
G	Complier analysis	166
G.1	Methodology	166
G.2	Results	167
H	Results by country	168
3	Inherited Gender Norms and the Gender Gap in Cognitive Functioning	170
1	Introduction	176
2	Empirical strategy	178
2.1	Baseline model	178
2.2	Heterogeneity with respect to parental place of birth	179
3	Data	180
3.1	The analytical sample	180

3.2	Measures of cognitive functioning	181
3.3	Covariates and mediating variables	185
3.4	Gender Norms Index	185
3.5	Graphical description	187
4	Results	189
4.1	Main results	189
4.2	Robustness checks	191
4.2.1	Alternative proxies for gender norms	191
4.2.2	Additional controls	193
4.2.3	Additional sample restrictions	194
4.2.4	Exploring alternative explanations	194
4.3	Mediating factors	195
4.3.1	Literature	195
4.3.2	Analyses	197
4.4	Further findings	197
5	Conclusion	198
	Bibliography	204
	Appendix	205
A	Descriptive statistics	205
B	Main results by sub-group	207
B.1	Methodology	207
B.2	Results	208
C	Alternative gender norms proxies	210
C.1	Description of alternative measures	210
C.2	Results	211
D	Additional control variables	214
D.1	Individual level controls	214
D.2	Country level controls	215
E	Sample restrictions	218
F	Further results	219
	Conclusion	223
	List of Tables	231
	List of Figures	234

Introduction

Table of contents

1	Gender economic convergence : features and driving forces	6
1.1	The rise in female labour force	6
1.2	Public policies as tools to empower women	7
1.3	Unpaid work : a similar convergence ?	8
2	Convergence is not equality	9
2.1	Gender gaps remain on the labour market	9
2.1.1	Horizontal segregation	10
2.1.2	Vertical segregation	12
2.2	Men and women still allocate time differently	12
2.2.1	The child(ren) penalty	13
2.2.2	When gender norms stick	15
3	This dissertation	17
3.1	Contributions	17
3.2	Data and methods	19
	Bibliography	22

1 Gender economic convergence : features and driving forces

The convergence of economic gender roles that occurred over the twentieth century is one of the most striking transformations of the Western world economy. Women's empowerment started from the beginning of the century, leading to profound changes in the way women are considered in the economic framework. Ever since, economists have been documenting this converging trend and investigating the factors leading to what Nobel Prize winner Claudia Goldin calls the *Quiet revolution*.

1.1 The rise in female labour force

One of the main features of this convergence is the rise in female labour supply over time. In the United States, female labour force participation increased from less than 20% to about 35% between 1900 and 1950 and more than doubled from 1950 to the beginning of the twenty-first century to reach 80% in 2000 [Goldin, 2006], with similar trends in other OECD countries [Olivetti and Petrongolo, 2016]. At the same time, the labour market became less gender-segregated with a higher proportion of women in occupations that were traditionally male-dominated. These modifications contributed to reducing the gender gap in earnings over time [Blau and Kahn, 2000].

Several driving forces have been highlighted in the literature. Among them, structural changes in the economy have been at play. Since women were initially more represented than men in the service sector, the shift of labour demand towards more tertiary occupations translated into a higher representation of women in the labour force in several OECD countries [Olivetti and Petrongolo, 2016; Ngai and Petrongolo, 2017; Petrongolo and Ronchi, 2020]. At the same time, the arrival of technological substitutes to domestic production on the market helped women outsourcing housework and saving time for paid work [Coen-Pirani et al., 2010] and the decrease in market prices for home appliances went along with a higher female labour market participation [Cavalcanti and Tavares, 2008].

Moreover, changes in expectations about women's role in the economy contributed to female empowerment. Because more women were drawn into the workforce during the Second World War, notably due to male mobilisation [Goldin and Olivetti, 2013] and casualties [Brodeur and

Kattan, 2022], the societal view of women’s role in the economy evolved and contributed to change gender role attitudes from one generation to the next. In that respect, [Fernández et al., 2004] provide evidence that the increase in women’s workforce due to the war changed the view their sons have about working women which, in turn, had an effect on female labour supply in the next generation.

1.2 Public policies as tools to empower women

Another important driver of female empowerment is access to birth control methods for young women. Initiated from the 1960s, policies aiming at lowering legal age for access to contraceptive pills contributed to postponing age at marriage and age at first birth. With a higher opportunity to control birth timing, education investment costs have been reduced for women [Goldin and Katz, 2002], female labour force has increased and the gender wage gap has shrunk in the United States [Bailey et al., 2012]. Abortion legalisation also played a substantial role in giving women more control over their fertility which, in turn, fostered their economic opportunities [Myers, 2017].

Second, public childcare policies emerged by the end of the century to give mothers the opportunity to better reconcile work and family duties. A vast literature has evaluated the effect of childcare supply expansion for children in their first years of life but results diverge according to the context. In the 1990s, reforms that increased the number of available day-care slots fostered maternal labour supply in Canada [Baker et al., 2008; Lefebvre and Merrigan, 2008]. A similar conclusion is put forward in Belgium [Dujardin et al., 2018], in Germany [Müller and Wrohlich, 2020], in Norway [Andresen and Havnes, 2019] and in Switzerland [Ravazzini, 2018] in the 2000s. In Italy, childcare expansion has been shown particularly efficient in provinces with low initial levels of public childcare provision [Brilli et al., 2016]. In some other cases, such reforms did not boost female labour supply but crowded out informal childcare arrangements [Havnes and Mogstad, 2011] or more costly formal childcare solutions instead [Pora et al., 2020]. In addition to childcare provision expansion, reducing childcare costs has also been highlighted as a driving force. In Germany, the introduction of free day-care slots increased maternal labour supply [Huebener et al., 2020]. This policy particularly benefited mothers with low socio-economic status [Busse and Gathmann, 2020], promoting equality of opportunities among mothers with

different backgrounds. This positive result is somewhat contrasted by other studies that find null or only modest effects [Lundin et al., 2008; Bettendorf et al., 2015], or positive effects only among mothers of large families [Givord and Marbot, 2015] when reducing or abolishing day-care fees.

Then, pre-school programs have been implemented to provide pre-elementary instruction for all, irrespective of parental socio-economic background. At first glance, pre-school programs are not expected to lever maternal labour force if childcare for very young children help mothers resuming work earlier. They nonetheless have been shown effective in boosting maternal labour supply in the United States [Gelbach, 2002], in Argentina [Berlinski et al., 2011], in Germany [Bauernschuster and Schlotter, 2015], in Spain [Nollenberger and Rodríguez-Planas, 2015] and in Italy [Carta and Rizzica, 2018]. In some contexts, pre-school enrollment levers employment among single mothers only, implying that pre-school helps mothers with less childcare options to go back on the labour market [Cascio, 2009; Goux and Maurin, 2010; Fitzpatrick, 2012]. The positive effect of these programs can be explained by the fact that enrollment to public pre-school is free while it is usually not the case for early childcare arrangements such as collective day-care or child-minders. Nevertheless, a remaining problem that arises for women with school-age children is that school schedules and parents' office hours do not always match. If mothers have only few alternatives for after-school care, the gap between schedules might prevent mothers from working full time. In that respect, school schedule extensions in the 2010s have allowed mothers to work more hours, providing smaller gender gaps in pay [Duchini and Van Effenterre, 2022; Berthelon et al., 2023].

1.3 Unpaid work : a similar convergence ?

It is also important to consider the domestic sphere when dealing with economic gender gaps. This point has been much less explored in the literature, and yet, unpaid work plays a tremendous role in the economic landscape [Stiglitz et al., 2009]. Omitting domestic production from the scope would therefore lead to undervalue gender inequalities in the economy. Furthermore, the gender gap in unpaid work is a major challenge worldwide and is part of the United Nations' sustainable development goals to attain gender equality [United Nations, 2022]. A few studies have explored the evolution of the gender gap in unpaid work over time and shown that women

have reallocated time from home production to the labour market and that men have increased their involvement in domestic tasks [Gimenez-Nadal and Sevilla, 2012], therefore resulting in a decreasing gender gap between the 1980s and the 2010s [D’Albis et al., 2016]. A recent literature shows that the introduction of shared parental leave has been successful in increasing fathers’ involvement [Schober, 2014; Tamm, 2019]. Looking closer at domestic time-allocation trends, it nonetheless appears that men have increased the amount of time they spend with children while time allocated to housework remains virtually stable [Champagne et al., 2015]. All in all, the fact remains that gender convergence in unpaid work has been far less dramatic than the one observed for paid work and that women still bear the major part of domestic work.

2 Convergence is not equality

Nowadays, gender discrepancies remain on two aspects of the economy. First, the gender convergence observed on the labour market started to level off as soon as in the 1990s despite great progress over several decades, leaving the prospect of reaching gender equality fading away. Second, the modest convergence in unpaid work has failed to compensate the meteoric rise in female labour force. A major problem stemming from those two stylized facts is that women, and especially mothers, meet the reality of the *second shift* since they work almost as much as men (but for a lower pay) while being the main provider of domestic tasks in the household [Bertrand, 2011].

2.1 Gender gaps remain on the labour market

In 2017, women still earned less than men with large cross-country variations in gender income differentials. The gender gap ranged from 5% to 20% depending on countries, with the United States on the top of gender income inequalities [Petrongolo and Ronchi, 2020]. Not only that average income differences exist between men and women, but they also vary along the earnings distribution. The gender gap is larger at the upper tail of the distribution which illustrates the existence of a *glass ceiling* [Albrecht et al., 2003; Fortin et al., 2017; Piketty et al., 2018; Garbinti et al., 2018] while *sticky floor* effects, i.e. larger gender gaps at the bottom of the distribution, are also demonstrated in some countries [Arulampalam et al., 2007]. In France, Bozio et al. [2014] estimate that about 30% of the gender gap in pay is attributed to occupational

segregation.¹ Two different types of gender occupational segregation have been put forward in the literature : *horizontal segregation*, which is that women are over-represented in occupations with lower potential income and *vertical segregation* or the fact that women are less represented in higher-paid positions in a given sector.

2.1.1 Horizontal segregation

This first type of segregation stems from the fact that men and women make different choices in terms of study fields earlier in their life. Despite the fact that young women now outperform their male counterparts in high school, are more likely to attend college [Fortin et al., 2015] and have higher chances to graduate than men once enrolled [Goldin et al., 2006; OCDE, 2021], they are still under-represented in fields with higher potential incomes. In the literature, this phenomenon is referred as the *leaky pipeline* to illustrate that fewer women choose to major in these fields as education level increases. Concretely, it has been shown that women are less likely than men to complete a tertiary degree in Science, Technology, Engineering and Mathematics fields [Bertrand, 2020].

With an extensive literature review, Azmat and Petrongolo [2014] provide various answers to this phenomenon. On one hand, there is evidence that women are less willing to compete than men or perform less well in competitive settings [Croson and Gneezy, 2009]. Landaud et al. [2020] illustrate the gender differential in willingness to compete with observational data in France. They exploit the discontinuous relationship between secondary-school grade scores and admission to selective high-schools to show that female students are less likely to choose sciences as a major field of study when enrolled in a more selective high school. However, this is not the case for male students. A similar conclusion is drawn from an experiment among high-school students in the Netherlands [Buser et al., 2014]. Concretely, the authors elicit willingness to compete by allowing students to choose between a competitive or a non-competitive payment scheme in a setting where they are asked to perform an arithmetic task. Using this experimentally drawn measure of taste for competition, they show that 23% of the gender gap in major choice is attributable to competitiveness. In addition, Ors et al. [2013] find evidence of lower relative female performances in competitive academic examinations relative to men while it is

¹More precisely, [Bozio et al., 2014] find that the gender gap in monthly wage amounts to 24.5% among which 7.3% are attributable to occupational segregation.

the reverse for national non-competitive exams. Their findings are based real-life results to entrance examinations for selective business schools as opposed to national examination results at the end of high school in France. [Flory et al. \[2015\]](#) also find that women turn away from competition in a professional framework by running a field experiment in the United States in which they randomize job advertisements into different compensation regimes, i.e., either non-competitive (a fixed-wage compensation) or competition-driven (based on individual or team performance) for the same job and show that female job-applicants are less likely than men to choose competitive-based compensations.

Another potential explanation includes the fact that female students under-perform in mathematics when they are exposed to more conservative gender norms, which supports the idea that gender differences in scientific skills are socially constructed. In this vein, [Guiso et al. \[2008\]](#) find that there is a male advantage in mathematics in countries with a high level of gender inequality but not in the most gender-equal countries. [Nollenberger et al. \[2016\]](#) and [Rodríguez-Planas and Nollenberger \[2018\]](#) show that conservative gender norms transmitted across generations increase the gender gap in mathematics performance while [Dossi et al. \[2021\]](#) use parental preference for sons to show that gender stereotypes have a negative impact on girls' performances at school. Gender stereotypes held by teachers are also at stake in student achievement and study field choice. In Italy, [Carlana \[2019\]](#) concludes that girls under-perform in mathematics when they are exposed to gender-biased teachers but no such effect is found for boys. The author uses the Gender-Science Implicit Association Test, a tool developed by psychologists and widely used by economists to measure teachers' expectations with achievement in mathematics and reading by female and male students. Teachers are defined as "gender-biased" when they implicitly associate boys with mathematics and girls with reading. [Lavy and Sand \[2018\]](#) find that teachers' gender-bias encourages boys to choose advanced mathematics courses in high-school and discourages girls to do so. Using a grading reform in Swedish universities, [Jansson and Tyrefors \[2022\]](#) bring a concrete example of *leaky pipeline* in economics by showing that female students' mid-term grades are higher when they are marked anonymously.

2.1.2 Vertical segregation

In France, [Gobillon et al. \[2015\]](#) report that women have a lower access to high-paid jobs compared to low-paid ones in the private sector. In the United States, 75% of the gender gap in pay is explained by a lower representation of women in managerial positions or by the fact that they manage smaller companies [[Bertrand and Hallock, 2001](#)]. Two main explanations can be drawn from the literature. First, *bargaining effects* seem to be at stake. Based on matched worker-firm data in Portugal, [Card et al. \[2016\]](#) advocate that women receive less premiums than men in a given workplace, which is attributed to lower bargaining behaviours among women. In addition, women are more likely to conduct tasks with low promotion-potential than men (e.g. writing reports, attending committees, organising events...) [[Babcock et al., 2017](#)]. Interestingly, this gender differential stems from different channels : women volunteer more, are more often asked to volunteer and are more likely to accept to run such tasks than men, revealing that gender stereotypes also have an influence on women's progress within the workplace.

Discriminatory behaviours have also been highlighted as a driver of vertical segregation. First, sex-biased hiring has been put forward by [Goldin and Rouse \[2000\]](#) using a natural experiment based on the change in recruitment practices among symphony orchestras in the 1970s. They find that the shift to *blind* auditions, that guarantee the candidate's anonymity, increases women's probability to pass the first selection round by 50%. Field experiments are also used in this literature, such as the one conducted by [Kübler et al. \[2018\]](#). By giving fictitious CVs to human resources managers, they show that women receive poorer evaluations than men for a similar qualification level. Discrimination against women is even larger for applications on male-dominated occupations. In contrast, [Bertrand and Mullainathan \[2004\]](#) find evidence of racial and social discrimination in hiring, but not on the grounds of the applicant's sex in the United States. In addition to biased hiring, discrimination in the workplace has also been shown detrimental to women's career aspirations which, in turn, influences the gender gap in promotion among lawyers in the United States [[Azmat et al., 2020](#)].

2.2 Men and women still allocate time differently

While the dramatic change in women's role on the labour market is an undeniable fact, a puzzle remains : Why do women still bear the major load of unpaid work in the household despite

their empowerment ? A number of answers has been brought by the literature in economics. Among them, parenthood has been shown to have asymmetrical effects between women's and men's labour market outcomes. A few recent studies have also explored the contribution of parenthood on gender gaps in time-allocation, which includes unpaid tasks. Another strand of literature highlights that the persistence of traditional gender norms prevents further gender convergence in the domestic sphere.

2.2.1 The child(ren) penalty

There is an abundant and recent research investigating the role played by parenthood in economic gender gaps. Based upon [Angelov et al. \[2016\]](#)'s methodology perfected by [Kleven et al. \[2019a\]](#) a few years later, these studies exploit variations in the first birth timing in an event-study design to assess the evolution of gender gaps in labour market outcomes around the first birth. These replications are informative in showing that despite variations in the magnitude of estimates, the *child penalty*, i.e. the change in a parental outcome value following the first birth relative to the pre-birth period, is still a concern nowadays and this is the case in many countries. All studies show that *i*) only women's labour market outcome values drop after the first birth while fathers' trajectories remain virtually unchanged and *ii*) the penalty reduces with distance to the first birth but does not disappear entirely. Several years after the first birth, the penalty in earnings is about 25% in Denmark [[Bütikofer et al., 2018](#)], Sweden [[Angelov et al., 2016](#); [Kleven et al., 2019a,b](#)] and Norway [[Andresen and Nix, 2022](#)]. It ranges between 30% and 40% in the United States [[Kleven et al., 2019a](#); [Cortés and Pan, 2020](#)], in France [[Meurs and Pora, 2019](#)] and in Spain [[De Quinto et al., 2021](#)], and reaches 50% in the Netherlands [[Rabaté and Rellstab, 2022](#); [Artmann et al., 2022](#)] and in Austria and even 60% in Germany [[Kleven et al., 2019a](#)]. Some of these studies disentangle the underlying channels and show that the penalty is mainly attributed to labour supply, both at the extensive and intensive margins.

To date, only one paper considers this question with respect to time-allocation instead of only focusing on labour market outcomes. Using Mexican data, [Aguilar-Gomez et al. \[2019\]](#) find an increase in time spent on unpaid tasks that is about 3 times higher for women than for men shortly after the birth and that persists with time. This results has two implications. First, the penalty is also a concern in the domestic sphere and should not be disregarded in empirical

works as unpaid work is an integral part of the economy. Second, the magnitude is even larger with respect to unpaid work, which means that the *child penalty* is by far undervalued when focusing only on labour market outcomes.

To add to this, the penalty borne by mothers is not confined to the first child. Several studies have provided evidence that the number of children also plays a role in gender gaps. Using multiple births [Rosenzweig and Wolpin, 1980; Bronars and Grogger, 1994; Jacobsen et al., 1999; Black et al., 2005], infertility shocks [Aguero and Marks, 2008; Markussen and Strøm, 2022], fertility treatments [Lundborg et al., 2017] or the sex-composition in the sibship [Angrist and Evans, 1998; Cruces and Galiani, 2007; Cools et al., 2017] to instrument the number of children that parents have, all studies converge towards a negative effect of family size on maternal labour supply. Angrist and Evans [1998] conclude on a negative effect of the number of children on mothers' labour supply and no effect for fathers in the United States. Cruces and Galiani [2007] replicate this method on Mexican and Argentinian data and find similar results. In Norway, Cools et al. [2017] show that having more children reduces mothers' opportunities on the labour market such as being employed by higher paying firms and being at the top of the wage distribution in their workplace. Once again, research examining the effect of family size on gender gaps in time-allocation is very scarce. To the best of my knowledge, it is summed up to one paper. Taking the sex-composition among the first two children as an instrument, Frenette [2011] shows that mothers decrease the time spent on the labour market and reallocate it to unpaid tasks such as housework and childcare when the number of children increases while fathers do not change their time-allocation from the exception of a modest increase in childcare time.

A last and important point in this literature is the fact that parenthood still contributes to gender gaps on many aspects of the economy nowadays, which constitutes a potential answer to the fact that gender equality has not been reached yet. In a second part of their paper, Kleven et al. [2019a] use a wage decomposition method² and delve into the residual part of gender differentials by integrating the contribution of parenthood to income inequality in a dynamic perspective. Their investigations show that the gender gap in earnings is largely attributed to

²Methodology introduced by Oaxaca [1973] and Blinder [1973] which consists in decomposing gender income inequality between *explained* and *unexplained* components. The explained part relates to gender differences in observed characteristics associated to the gender gap while the unexplained (or residual) part refers to gender differences in returns to these characteristics.

parenthood. Between 1980 and 2010, the contribution of children relative to other factors has increased from 40% to 80% which is due to the fact that residual gender inequality that is not attributed to children reduces over time while the contribution of parenthood remains stable.

2.2.2 When gender norms stick

In an effort to better understand individual behaviours, economists have started to consider cultural factors as potential drivers from the 2000s. Before that, such factors were virtually absent of research in economics, considering that *culture* was too difficult to define and measure in order to be properly used in econometric models. Guiso et al. [2006] define it as "*customary beliefs and values that ethnic, religious and social groups transmit fairly unchanged from generation to generation*" and review the contemporary literature that assesses the relationship between social norms and several economic outcomes such as saving behaviours, trust or preference for redistribution. As part of culture, gender norms have been shown to play a substantial role in the persistence of gender gaps in the economy.

Empirically, disentangling norms from institutions is not an easy task because of the two-way causal relationship that links them together [Alesina and Giuliano, 2015]. In that respect, the German divide illustrates how institutional shocks can affect social norms. In 1949, the world entered the Cold War that would divide Europe for 45 years between the Western bloc and the Eastern bloc. From each side of the *iron curtain*, economic and political regimes were radically opposed. In East Germany, workforce was lacking due to the post-war mass exodus. As a result, the German Democratic Republic rapidly implemented work-family compatible measures with the aim to draw women into the workforce as much as men. Among them, public childcare provision and flexible working hours for women were aimed at bringing mothers on the labour market. On the other side of the border, the *breadwinner model*³ was still widespread and childcare supply scarcity would hamper women's labour supply [Bauernschuster and Rainer, 2012]. After reunification, economic behaviours are shaped differently between the East and the West, in particular with regard to gender role attitudes. Even nowadays, women from East Germany are less likely to hold traditional gender role attitudes towards work than in the West

³A *single earner model* in which the man earns the total household income or its major part and the woman takes over the unpaid work in an heterosexual couple.

[Campa and Serafinelli, 2019], they are not likely to *do gender*⁴ [Lippmann et al., 2020] and the gender gap in mathematics is smaller in the East [Lippmann and Senik, 2018]. In particular, Lippmann et al. [2020] show that, in West Germany, women who switch from earning less to earning more than their partners compensate by increasing their amount of unpaid work or even withdraw from the labour market while this is not the case for women from the East. The main message to be drawn from this literature is that gender gaps are socially constructed and that public policies not only affect short-term economic behaviours but also social norms on the long run.

In order to isolate the contribution of norms in individual behaviours, the *epidemiological approach* has been developed in the literature. Concretely, the methodology proposed by the authors consists in focusing on a sample of second-generation immigrants, i.e. individuals who are born in a country with at least one parent born in another country, and exploiting variations in gender norms across parental countries of birth while keeping the institutional environment fixed. In particular, Fernández and Fogli [2009] use female labour force participation and fertility rates in the country of birth of second-generation immigrants' parents as proxies for gender norms and show that these measures predict women's workforce and fertility decisions in the next generation. With the same approach, Blau et al. [2013] confirm the inter-generational transmission of gender norms and go a step further in showing that transmission is more powerful through mothers' country of birth and when both parents come from the same country. This approach has been used with alternative measures of gender norms and a wide range of outcomes such as school performances, intimate partner violence or competitive behaviours.⁵ More recently, this approach has been used to assess the contribution of gender norms to the division of unpaid tasks between men and women [Blau et al., 2020]. The authors conclude that male second-generation immigrants' time allocated to domestic tasks depends on whether norms converge towards more gender equality in the parental country of birth. This results suggest that inter-generational transmission is also at stake when it comes to unpaid work, therefore explaining at least part of the persistence of gender gaps in time-allocation over time.

Sticky gender gaps in paid and unpaid work also seem to reside in the fact that women tend to *do gender* when the couple derives from the male *breadwinner model* [Bittman et al., 2003; Bertrand

⁴Wording used by the authors to describe socially constructed behaviours [West and Zimmerman, 1987].

⁵This literature is further detailed in the third chapter of this dissertation that builds on this methodology.

et al., 2015]. Interestingly, there is a sharp drop in the distribution of female contribution to household income at exactly 50%, suggesting that women are averse to earning more than their male partners. In turn, this conservative view of gender role attitudes affects the division of domestic work at the couple level in the sense that women compensate a higher relative income by providing more unpaid work in order to stick to gender identity norms [Akerlof and Kranton, 2000]. These results suggest that the perspective of reaching perfect equality in market and non-market work cannot be considered if gender norms do not evolve. This phenomenon is all the more problematic that it negatively affects women's well-being [Flèche et al., 2020]. Engaging men to involve more in domestic work should therefore be considered as a priority by policy makers.

3 This dissertation

In a context of stalled gender economic convergence, this thesis aims at bringing additional insights on causes, consequences and solutions to gender inequalities with three empirical works lying at the intersection between gender, family and health economics.

3.1 Contributions

In Chapter 1, *Parenthood, Time-Allocation and Well-being*, I explore how parenthood affects gender gaps in time-allocation and well-being in Germany. While the role of parenthood in labour market gender gaps has widely been explored in the literature, there is only very scarce evidence on time-allocation behaviours. Yet, market work is only one part of economic production. Taking a step towards filling this gap, this chapter examines the evolution of mothers' and fathers' time-allocation between paid work, unpaid work and unproductive activities, also called *leisure*, around the first birth in a first part, and with respect to the number of children in a second part. Results concerning time allocated to paid work around the first birth confirm previous findings of the literature according to which fathers' trajectory remains unchanged while women's paid work time sharply reduces after the birth and never entirely converges back. I observe that fathers get involved in unpaid tasks following the birth of the first child, but to a much lower extent compared to mothers. Interestingly, while the gender gap in paid work penalty gets lower with distance to birth for both one-child and several-children

mothers, the gap in unpaid work is greater the further from the first birth for mothers with additional children. This result suggests that even if they resume work, mothers keep the major part of the unpaid work load, consistent with the *second-shift* problem. Both parents reduce their leisure time, again with a larger penalty for women, especially if they have several children. I also show a drop in satisfaction with housework for both parents but with a greater magnitude among men. If this result may come as a surprise, it actually sheds light on the fact that gender identity norms are deeply rooted. In a second part, I analyse whether parents change their time-allocation when they have more children. Intra-household time-allocation modifications with respect to additional children are not straightforward considering that such decisions might be taken at entry into parenthood and remain stable afterwards, irrespective of further fertility decisions. Actually, results show that the number of children emphasizes pre-existing gender gaps in time-allocation. Interestingly, mothers declare a deteriorated satisfaction with respect to leisure time while keeping leisure time constant when they have more children, which could stem from a lower quality of leisure time for women with larger families.

Chapter 2, *Pre-school Enrollment and Mothers' Labour Supply : Evidence from Europe*, is a joint work with Julien Bergeot. Starting from the fact that pre-school programs provide very different effects on maternal labour supply depending on the country of observation, we wondered what could explain such variability in the literature. To this aim, we evaluate the effect of pre-school enrollment in a cross-country comparison perspective with harmonized data in ten European countries. The contribution of this paper resides in the fact that, to date, there is no research exploring country-level heterogeneity in the effect of pre-school on mothers' labour supply. We first show that, pooling all types of individuals and countries together, pre-school has no significant effect on mothers' workforce. We then turn to heterogeneity analyses at the individual and at the country levels respectively. At the individual level, further explorations show that pre-school enrollment positively affects the intensive margin of labour supply among the most educated mothers and among mothers who do not have younger children. Country-level analyses reveal that pre-school is a leverage for maternal labour supply in countries with strong gender stereotypes, where female labour force is lower, and in countries with a low investment in early childcare, where mothers have fewer formal childcare solutions to resume work before their child enrolls in a pre-school program.

Chapter 3, *Inherited Gender Norms and the Gender Gap in Cognitive Functioning* is co-authored with Eric Bonsang. In line with the literature that assesses the contribution of gender norms to gender gaps in various economic outcomes, we investigate whether gender norms contribute to the gender gap in cognitive health among individuals aged 50 at least. Drawing on the medical literature on cognitive development factors over the life course, our hypothesis is that a gender gap in cognitive functioning might occur at older ages if men and women are exposed differently to such cognitive development factors during their lives due to conservative gender norms. Because cognitive impairment is one of the major challenges for aging societies, understanding the causes of women’s and men’s cognitive functioning in later ages is pivotal. To this aim, we build on the literature that exploits the channel of inter-generational transmission of social norms to assess the relationship between gender norms and the cognitive gender gap later in life. We find that women’s cognitive skills decrease relatively to men’s when gender norms transmitted *via* parents are more conservative. Explorations of potential mechanisms show that the type of occupation is likely to be a driver. Lastly, we show that the cognitive gender gap might have further health implications since 65+ aged women declare more difficulties with activities of daily living than their male counterparts when they are exposed to more conservative gender norms.

3.2 Data and methods

For each chapter, a different empirical strategy is implemented and various sources of data are mobilised. In the next paragraphs, I give more details on the empirical strategies and data sets used in the three chapters.

In Chapter 1, I explore two related but different research questions. For this reason, two different methodologies are used. In the first part in which I explore how gender gaps in time-allocation evolve when couples enter into parenthood, I use an *event-study analysis* around the first birth. This methodology exploits variations in the timing of the first birth to yield the effect of entry into parenthood on parents’ time-allocation and well-being, provided that the first birth timing is exogenous to unobserved determinants of parental outcomes.⁶ In the second part, I evaluate the effect of family size with an *instrumental variable strategy* (IV). Under specific

⁶This assumption is further discussed in Chapter 1.

assumptions⁷, this strategy provides the Local Average Treatment Effect (LATE) of family size on parent's time-allocation and well-being, i.e. the effect among individuals who comply to the instrument. Building on previous literature, I use the sex-composition of the first two children to predict fertility choices. Analyses in Chapter 1 are based on the German Socio-Economic Panel (GSOEP), a longitudinal survey that collects information at the household and the individual levels. I use observations between 1991 and 2018 since information was collected only in former West Germany before reunification. The population of interest gathers parents who live in the same household in order to focus on the intra-household division of tasks and I specifically focus on heterosexual couples to target gender discrepancies at the household level with regard to time-allocation and well-being.

Chapter 2 relies on a *fuzzy regression on discontinuity design* (RDD) which exploits the discontinuous relationship between eligibility to pre-school and the child's year of birth to predict pre-school enrollment because mothers' enrollment decisions are unlikely to be random, preventing us from estimating any causal effect of pre-school enrollment on maternal labour supply absent of exogenous variation in enrollment. Identification with a fuzzy RDD approach relies on the assumptions of a standard IV strategy with independence requiring absence of manipulation of the child's date of birth and continuity of potential outcomes at the eligibility threshold. Information on mothers' labour supply as well as their children's characteristics are drawn from the European Union Statistics on Income and Living Conditions data set (EU-SILC), harmonised data collected between 2004 and 2020 in several European Union member states to which we have been granted access by EUROSTAT. For the cross-country perspective, we use data from the OECD to measure public spending in in-cash and in-kind family benefits and the net cost of early childcare for each country. Gender stereotypes are proxied by measures drawn from the Integrated Values Surveys (IVS). By design, we focus our analyses on working age mothers who have at least one pre-school age child in a bandwidth of one year around the pre-school eligibility threshold in each country.

For Chapter 3, we build on the *epidemiological approach* in order to explore the contribution of inter-generational transmission of gender norms in the gender gap in cognitive functioning. This methodology consists in isolating social norms by holding the institutional background

⁷Relevance, Independence, Exclusion and Monotonicity [Angrist et al., 1996].

fixed using a sample of second-generation immigrants who are exposed to the same institutional environment but differ with respect to gender norms transmitted by their parents. We construct a sample of second-generation immigrants from the Survey of Health, Ageing and Retirement in Europe (SHARE) for which we observe cognitive functioning measures among individuals aged at least 50. This sample of interest is matched with other sources of data that contain information on gender norms at the parental country-of-birth level. Gender norms are measured using prevailing opinions on gender role attitudes for each corresponding parental birth country in the IVS. Alternatively, we also use the Gender Gap Index provided by the World Economic Forum (WEF) as measure of gender norms and we further check the robustness of our results by adding country-level controls made available by the World Bank.

Bibliography

- Aguero, J. M. and Marks, M. S. (2008). Motherhood and female labor force participation: Evidence from infertility shocks. *American Economic Review*, 98(2):500–504. [14, 42]
- Aguilar-Gomez, S., Arceo-Gomez, E., and De la Cruz Toledo, E. (2019). Inside the black box of child penalties: Unpaid work and household structure. *Available at SSRN 3497089*. [13, 42]
- Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753. [17, 41, 70]
- Albrecht, J., Björklund, A., and Vroman, S. (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21(1):145–177. [9]
- Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944. [15, 176]
- Andresen, M. E. and Havnes, T. (2019). Child care, parental labor supply and tax revenue. *Labour Economics*, 61:101762. [7, 102]
- Andresen, M. E. and Nix, E. (2022). What causes the child penalty? Evidence from adopting and same-sex couples. *Journal of Labor Economics*, 40(4):971–1004. [13, 42]
- Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579. [13, 42]
- Angrist, J. D. and Evans, W. N. (1998). Children and their parents’ labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3):450–477. [14, 41, 42, 43, 51, 53, 57, 66, 67]
- Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455. [20, 51]
- Artmann, E., Oosterbeek, H., and van der Klaauw, B. (2022). Household specialization and the child penalty in the Netherlands. *Labour Economics*, 78:102221. [13, 42]
- Arulampalam, W., Booth, A. L., and Bryan, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *IIR Review*, 60(2):163–186. [9]

- Azmat, G., Cuñat, V., and Henry, E. (2020). Gender promotion gaps: Career aspirations and workplace discrimination. *Available at SSRN 3518420*. [12]
- Azmat, G. and Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, 30:32–40. [10]
- Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, 107(3):714–747. [12]
- Bailey, M. J., Hershbein, B., and Miller, A. R. (2012). The opt-in revolution? Contraception and the gender gap in wages. *American Economic Journal: Applied Economics*, 4(3):225–54. [7]
- Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745. [7, 102]
- Bauernschuster, S. and Rainer, H. (2012). Political regimes and the family: How sex-role attitudes continue to differ in reunified Germany. *Journal of Population Economics*, 25(1):5–27. [15, 44]
- Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers’ labor supply : Evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16. [8, 44, 103]
- Berlinski, S., Galiani, S., and Mc Ewan, P. J. (2011). Preschool and maternal labor market outcomes: Evidence from a regression discontinuity design. *Economic Development and Cultural Change*, 59(2):313–344. [8, 103]
- Berthelon, M., Kruger, D., and Oyarzún, M. (2023). School schedules and mothers’ employment: Evidence from an education reform. *Review of Economics of the Household*, 21(1):131–171. [8]
- Bertrand, M. (2011). New perspectives on gender. In *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier. [9]
- Bertrand, M. (2020). Gender in the twenty-first century. In *AEA Papers and Proceedings*, volume 110, pages 1–24. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. [10]

- Bertrand, M. and Hallock, K. F. (2001). The gender gap in top corporate jobs. *ILR Review*, 55(1):3–21. [12]
- Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614. [16, 43]
- Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4):991–1013. [12]
- Bettendorf, L. J., Jongen, E. L., and Muller, P. (2015). Childcare subsidies and labour supply : Evidence from a large Dutch reform. *Labour Economics*, 36:112–123. [8, 129]
- Bittman, M., England, P., Sayer, L., Folbre, N., and Matheson, G. (2003). When does gender trump money? Bargaining and time in household work. *American Journal of Sociology*, 109(1):186–214. [16, 43]
- Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children’s education. *The Quarterly Journal of Economics*, 120(2):669–700. [14, 42]
- Blau, F. D. and Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4):75–100. [6, 40]
- Blau, F. D., Kahn, L. M., Comey, M., Eng, A., Meyerhofer, P., and Willén, A. (2020). Culture and gender allocation of tasks: Source country characteristics and the division of non-market work among US immigrants. *Review of Economics of the Household*, 18:907–958. [16]
- Blau, F. D., Kahn, L. M., Liu, A. Y.-H., and Papps, K. L. (2013). The transmission of women’s fertility, human capital and work orientation across immigrant generations. *Journal of Population Economics*, 26:405–435. [16, 103, 177, 179, 199]
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, pages 436–455. [14]
- Bozio, A., Dormont, B., and García-Peñalosa, C. (2014). Reducing wage inequalities between men and women. *Notes du Conseil d’Analyse Economique*, 17(7):1–12. [9, 10]
- Brilli, Y., Del Boca, D., and Pronzato, C. D. (2016). Does child care availability play a role in

- maternal employment and children’s development? Evidence from Italy. *Review of Economics of the Household*, 14:27–51. [7, 102]
- Brodeur, A. and Kattan, L. (2022). World War II, the baby boom and employment: County-level evidence. *Journal of Labor Economics*, 40(2):437–471. [6]
- Bronars, S. G. and Grogger, J. (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *American Economic Review*, pages 1141–1156. [14, 42]
- Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447. [10]
- Busse, A. and Gathmann, C. (2020). Free daycare policies, family choices and child development. *Journal of Economic Behavior & Organization*, 179:240–260. [7, 103]
- Bütikofer, A., Jensen, S., and Salvanes, K. G. (2018). The role of parenthood on the gender gap among top earners. *European Economic Review*, 109:103–123. [13, 42]
- Campa, P. and Serafinelli, M. (2019). Politico-economic regimes and attitudes: Female workers under State socialism. *Review of Economics and Statistics*, 101(2):233–248. [16]
- Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, sorting and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633–686. [12]
- Carlana, M. (2019). Implicit stereotypes: Evidence from teachers’ gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224. [11]
- Carta, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children’s outcomes: Evidence from Italy. *Journal of Public Economics*, 158:79–102. [8, 103]
- Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources*, 44(1):140–170. [8, 103, 124]
- Cavalcanti, T. V. and Tavares, J. (2008). Assessing the “engines of liberation”: Home appliances and female labor force participation. *The Review of Economics and Statistics*, 90(1):81–88. [6]
- Champagne, C., Pailhé, A., and Solaz, A. (2015). Le temps domestique et parental des hommes et des femmes: Quels facteurs d’évolutions en 25 ans? *Économie et Statistique*, 478(1):209–

242. [9, 43]

- Coen-Pirani, D., León, A., and Lugauer, S. (2010). The effect of household appliances on female labor force participation: Evidence from microdata. *Labour Economics*, 17(3):503–513. [6]
- Cools, S., Markussen, S., and Strøm, M. (2017). Children and careers: How family size affects parents' labor market outcomes in the long run. *Demography*, 54(5):1773–1793. [14, 43, 52]
- Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. Technical report, National Bureau of Economic Research. [13, 42]
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474. [10]
- Cruces, G. and Galiani, S. (2007). Fertility and female labor supply in latin america: New causal evidence. *Labour Economics*, 14(3):565–573. [14, 43]
- De Quinto, A., Hospido, L., and Sanz, C. (2021). The child penalty: Evidence from Spain. *SERIEs*, 12:585–606. [13, 42]
- Dossi, G., Figlio, D., Giuliano, P., and Sapienza, P. (2021). Born in the family: Preferences for boys and the gender gap in math. *Journal of Economic Behavior & Organization*, 183:175–188. [11, 200]
- Duchini, E. and Van Effenterre, C. (2022). School schedule and the gender pay gap. *Journal of Human Resources*, pages 0121–11431R2. [8]
- Dujardin, C., Fonder, M., and Lejeune, B. (2018). Does formal child care availability for 0-3 year olds boost mothers' employment rate? Panel data based evidence from Belgium. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (129):103–126. [7, 102]
- D'Albis, H., Bonnet, C., Navaux, J., Pelletan, J., and Solaz, A. (2016). Travail rémunéré et travail domestique: Une évaluation monétaire de la contribution des femmes et des hommes à l'activité économique depuis 30 ans. *Revue de l'OFCE*, (5):101–130. [9, 43]
- Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177. [16, 103, 125, 177, 186, 191, 199]
- Fernández, R., Fogli, A., and Olivetti, C. (2004). Mothers and sons: Preference formation and

- female labor force dynamics. *The Quarterly Journal of Economics*, 119(4):1249–1299. [7]
- Fitzpatrick, M. D. (2012). Revising our thinking about the relationship between maternal labor supply and preschool. *Journal of Human Resources*, 47(3):583–612. [8, 103]
- Flèche, S., Lepinteur, A., and Powdthavee, N. (2020). Gender norms, fairness and relative working hours within households. *Labour Economics*, 65:101866. [17, 43]
- Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1):122–155. [11]
- Fortin, N. M., Bell, B., and Böhm, M. (2017). Top earnings inequality and the gender pay gap: Canada, Sweden and the United Kingdom. *Labour Economics*, 47:107–123. [9]
- Fortin, N. M., Oreopoulos, P., and Phipps, S. (2015). Leaving boys behind: Gender disparities in high academic achievement. *Journal of Human Resources*, 50(3):549–579. [10]
- Frenette, M. (2011). How does the stork delegate work? Childbearing and the gender division of paid and unpaid labour. *Journal of Population Economics*, 24(3):895–910. [14, 43]
- Garbinti, B., Goupille-Lebret, J., and Piketty, T. (2018). Income inequality in France, 1900–2014: Evidence from distributional national accounts (DINA). *Journal of Public Economics*, 162:63–77. [9]
- Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, 92(1):307–322. [8, 103]
- Gimenez-Nadal, J. I. and Sevilla, A. (2012). Trends in time allocation: A cross-country analysis. *European Economic Review*, 56(6):1338–1359. [9, 43]
- Givord, P. and Marbot, C. (2015). Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies. *Labour Economics*, 36:99–111. [8, 103]
- Gobillon, L., Meurs, D., and Roux, S. (2015). Estimating gender differences in access to jobs. *Journal of Labor Economics*, 33(2):317–363. [12]
- Goldin, C. (2006). The quiet revolution that transformed women’s employment, education and family. *American Economic Review*, 96(2):1–21. [6]

- Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women’s career and marriage decisions. *Journal of Political Economy*, 110(4):730–770. [7]
- Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4):133–156. [10]
- Goldin, C. and Olivetti, C. (2013). Shocking labor supply: A reassessment of the role of World War II on women’s labor supply. *American Economic Review*, 103(3):257–262. [6]
- Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of “blind” auditions on female musicians. *American Economic Review*, 90(4):715–741. [12]
- Goux, D. and Maurin, E. (2010). Public school availability for two-year olds and mothers’ labour supply. *Labour Economics*, 17(6):951–962. [8, 103, 124]
- Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender and math. *Science*, 320(5880):1164–1165. [11, 185, 192]
- Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48. [15]
- Havnes, T. and Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics*, 95(11-12):1455–1465. [7, 102]
- Huebener, M., Pape, A., and Spiess, C. K. (2020). Parental labour supply responses to the abolition of day care fees. *Journal of Economic Behavior & Organization*, 180:510–543. [7, 103, 129]
- Jacobsen, J. P., Pearce III, J. W., and Rosenbloom, J. L. (1999). The effects of childbearing on married women’s labor supply and earnings: Using twin births as a natural experiment. *Journal of Human Resources*, pages 449–474. [14, 42]
- Jansson, J. and Tyrefors, B. (2022). Grading bias and the leaky pipeline in economics: Evidence from Stockholm University. *Labour Economics*, 78:102212. [11]
- Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, volume 109, pages 122–126. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. [13, 14, 42, 46]

- Kleven, H., Landais, C., and Sogaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209. [13, 42, 46]
- Kübler, D., Schmid, J., and Stüber, R. (2018). Gender discrimination in hiring across occupations: A nationally-representative vignette study. *Labour Economics*, 55:215–229. [12]
- Landaud, F., Ly, S. T., and Maurin, É. (2020). Competitive schools and the gender gap in the choice of field of study. *Journal of Human Resources*, 55(1):278–308. [10]
- Lavy, V. and Sand, E. (2018). On the origins of gender gaps in human capital: Short-and long-term consequences of teachers’ biases. *Journal of Public Economics*, 167:263–279. [11]
- Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, 26(3):519–548. [7, 102]
- Lippmann, Q., Georgieff, A., and Senik, C. (2020). Undoing gender with institutions: Lessons from the German division and reunification. *The Economic Journal*, 130(629):1445–1470. [16, 43, 44]
- Lippmann, Q. and Senik, C. (2018). Math, girls and socialism. *Journal of Comparative Economics*, 46(3):874–888. [16]
- Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review*, 107(6):1611–37. [14, 42]
- Lundin, D., Mörk, E., and Öckert, B. (2008). How far can reduced childcare prices push female labour supply? *Labour Economics*, 15(4):647–659. [8]
- Markussen, S. and Strøm, M. (2022). Children and labor market outcomes: Separating the effects of the first three children. *Journal of Population Economics*, 35(1):135–167. [14, 42]
- Meurs, D. and Pora, P. (2019). Gender equality on the labour market in France: A slow convergence hampered by motherhood. *Economie et Statistique*, 510(1):109–130. [13, 42]
- Müller, K.-U. and Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare. *Labour Economics*, 62:101776. [7, 44, 102]
- Myers, C. K. (2017). The power of abortion policy: Reexamining the effects of young women’s

- access to reproductive control. *Journal of Political Economy*, 125(6):2178–2224. [7]
- Ngai, L. R. and Petrongolo, B. (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, 9(4):1–44. [6]
- Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36:124–136. [8, 103]
- Nollenberger, N., Rodríguez-Planas, N., and Sevilla, A. (2016). The math gender gap: The role of culture. *American Economic Review*, 106(5):257–261. [11, 177, 192, 200]
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, pages 693–709. [14]
- OCDE (2021). *Education at a glance 2021*. [10]
- Olivetti, C. and Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8:405–434. [6, 40]
- Ors, E., Palomino, F., and Peyrache, E. (2013). Performance gender gap: Does competition matter? *Journal of Labor Economics*, 31(3):443–499. [10]
- Petrongolo, B. and Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64:101819. [6, 9]
- Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the United States. *The Quarterly Journal of Economics*, 133(2):553–609. [9]
- Pora, P. et al. (2020). *Keep working and spend less?: Collective childcare and parental earnings in France*. INSEE, Institut national de la statistique et des études économiques. [7, 102]
- Rabaté, S. and Rellstab, S. (2022). What determines the child penalty in the Netherlands? The role of policy and norms. *De Economist*, 170(2):195–229. [13, 42]
- Ravazzini, L. (2018). Childcare and maternal part-time employment: A natural experiment using Swiss cantons. *Swiss Journal of Economics and Statistics*, 154:1–16. [7, 102]
- Rodríguez-Planas, N. and Nollenberger, N. (2018). Let the girls learn! It is not only about math... it’s about gender social norms. *Economics of Education Review*, 62:230–253. [11, 177, 179, 186, 192, 200]

- Rosenzweig, M. R. and Wolpin, K. I. (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: Journal of the Econometric Society*, pages 227–240. [14, 42]
- Schober, P. S. (2014). Parental leave and domestic work of mothers and fathers: A longitudinal study of two reforms in West Germany. *Journal of Social Policy*, 43(2):351–372. [9, 45, 225]
- Stiglitz, J. E., Sen, A., Fitoussi, J.-P., et al. (2009). Report by the commission on the measurement of economic performance and social progress. [8]
- Tamm, M. (2019). Fathers’ parental leave-taking, childcare involvement and labor market participation. *Labour Economics*, 59:184–197. [9, 45, 225]
- United Nations (2022). The Sustainable Development Goals: Report 2022. [8]
- West, C. and Zimmerman, D. H. (1987). Doing gender. *Gender & Society*, 1(2):125–151. [16]

Chapter 1

Parenthood, Time-Allocation and Well-being

This chapter has greatly benefited from discussions with Dan Anderberg, Julien Bergeot, Eric Bonsang, Eve Caroli, Brigitte Dormont, Peter Eibich, Sarah Flèche, Gabrielle Gambuli, Aimée Kingsada, Camille Landais, Hélène Le Forner, Marion Leturcq, Jürgen Maurer, Elsa Perdrix and Agathe Simon. I am grateful to participants of the 2022 Annual Meeting of the Society for the Economics of the Household, the 35th Annual Conference of the European Society for Population Economics, the 38th Journées de Microéconomie Appliquée, the RHUL-Dauphine PhD Workshop, the 20th Augustin Cournot Doctoral Days, the 22nd LAGV Public Economics Conference, the 71st Annual Congress of the French Economics Association and the 2nd Dauphine-PSE Workshop for their comments on earlier versions of this chapter.

Table of contents

1	Introduction	40
2	The German context	44
3	Methodology	46
	3.1 The first birth penalty	46
	3.2 The effect of family size	48
4	Data	54
	4.1 Sample restrictions	54
	4.2 Description of variables	55
5	Results	57
	5.1 The first birth penalty	57
	5.2 Family size effects	62
6	Conclusion	72
	Appendix	81
A	Identification tests	81
B	Descriptive statistics	84
C	First birth penalties	86
E	Internal validity	89
	E.1 Robustness checks	89
	E.2 Exclusion restriction test	90
F	Further results	91
G	Complier analysis	93
	G.1 Methodology	93
	G.2 Results	94

Abstract

This paper investigates how parenthood affects gender inequalities in time-allocation and well-being in Germany. Using an event-study analysis around the first birth, this paper shows that mothers reallocate time from paid work to unpaid work following the first birth while fathers' paid work time remains unchanged. Fathers' time spent on unpaid tasks increases, but to a much lower extent compared to mothers. The gender gap in paid work time gets lower with distance to birth while the gap in unpaid work gets larger for mothers with additional children, consistent with the *second-shift* problem. I also show that the drop in satisfaction with housework following the birth is greater among men, suggesting that gender identity norms are deeply rooted. Since the number of children is also expected to play a role, I analyse whether parents change their time-allocation when family size increases using an instrumental variable approach. Results show that the number of children emphasizes pre-existing gender gaps in time-allocation. Interestingly, mothers declare a deteriorated satisfaction with respect to leisure time while keeping leisure time constant, which could stem from a lower quality of leisure time for women with larger families.

1 Introduction

Gender gaps on the labour market have narrowed over the past century [Blau and Kahn, 2000; Goldin, 2014; Olivetti and Petrongolo, 2016] but the convergence is plateauing since the 1990s, leading to persistent gender gaps nowadays. In that respect, parenthood is likely to play a role. Several descriptive studies report a negative association between motherhood and wages in the United States [Waldfogel, 1997; Budig and England, 2001; Anderson et al., 2002] while no significant relationship can be shown with respect to fathers [Loughran and Zissimopoulos, 2009]. On another hand, Lundberg and Rose [2000] provide evidence of substitution mechanisms between parents, with mothers reducing their working hours and, *de facto*, their income, while an increase for fathers is observed¹. In Norway, Cools and Strøm [2016] find wage penalties for fathers, though much smaller than for mothers.

The aim of this study is to explore how parenthood affects gender differences in time-allocation at the couple level in Germany. In a first part, I focus on entry into parenthood relying on an event-study analysis around the first birth. Provided that birth timing is exogenous to time-allocation², this approach exploits variations in the first birth timing to identify the effect of entry into parenthood, also called the *child penalty* in the literature. Event-study estimates show that mothers' time allocated to the labour market is reduced by more than 80% one year after the birth relative to one year before while their time spent on unpaid tasks is multiplied by almost 4. Fathers' time allocated to paid work remains virtually unchanged and their time spent on unpaid tasks is higher by 10% after the birth relative to one year before. Mothers' penalty in time spent on paid work narrows with time but is still higher than 30% for women with one child and reaches 60% for those with more children 10 years after the first birth. Among women, the penalty in domestic work decreases with distance to the first birth for one-child mothers but keeps on increasing for women with additional children, suggesting that the number of children might also be at stake in time-allocation gender gaps. Finally, leisure time is reduced for both

¹Note that Waldfogel [1997]; Budig and England [2001]; Anderson et al. [2002]; Loughran and Zissimopoulos [2009] all use fixed effect models on the National Longitudinal Survey of Youth (United States), but with different survey cohorts. Lundberg and Rose [2000] use the same approach on North American data from the Panel Study of Income Dynamics.

²This hypothesis is further discussed in Section 3.1.

men and women after the first birth with a larger amplitude among women, leading to a gender gap in the birth penalty of about 15% one year after the birth that slightly recovers over time, especially among one-child parents. I also investigate changes in subjective well-being with paid work, housework and leisure around the first birth. Satisfaction with paid work remains virtually smooth around the event for both men and women and satisfaction with leisure decreases for both parents in a similar (and modest) proportion. However, the decrease in satisfaction with housework is of greater magnitude among men despite the fact that the amount of unpaid work they take on after the birth is much lower compared to mothers'. This suggests that gender role attitudes are integrated by individuals, consistent with the identity formation theory [[Akerlof and Kranton, 2000](#)].

The second part of this paper explores whether the number of children has an effect on the gender gap in time-allocation. I investigate this question using an instrumental variable strategy in order to correct for the endogenous relationship between time-allocation and fertility decisions. Following [Angrist and Evans \[1998\]](#), family size is instrumented by the sex-composition of the first two children, thus identifying the Local Average Treatment Effect (LATE) of family size on parental time-allocation.³ I show that mothers reallocate time from the labour market to unpaid tasks while fathers' time allocation remains unchanged when family size increases. Effects are driven by the least educated mothers and mothers whose first-born children are older when an additional child enters the sibship. This paper also explores the effect of family size on parents' well-being. While fathers' well-being is not significantly affected by a higher number of children, mothers declare less satisfaction with respect to leisure while keeping leisure time unchanged, which might reflect a deteriorated leisure time quality. On the other hand, mothers do not declare less satisfaction with unpaid tasks despite the fact that they take on the major part of domestic work within the couple, in line with the results I find for entry into parenthood.

This paper contributes to three different strands of the literature. First, it relates to studies that use the event-study approach around the first birth. All these studies converge towards the same results: the first birth is only detrimental to women on the labour market and the gender gap gets smaller with time but never turns null. They also reveal cross-country variations in the size of penalties several years after the first birth. Long-run penalties⁴ are the lowest

³Under the standard IV assumptions that are further detailed in Section [3.2.3](#)

⁴In the literature, the *long-run penalty* is defined as the change in a given labour market outcome for women

in Denmark [Bütikofer et al., 2018], Sweden [Angelov et al., 2016; Kleven et al., 2019a,b] and Norway [Andresen and Nix, 2022], where mothers’ earnings fall behind men’s by around 25%. It ranges from 30% to 40% in the United States [Kleven et al., 2019a; Cortés and Pan, 2020], in France [Meurs and Pora, 2019] and in Spain [De Quinto et al., 2021] and reaches 50% in the Netherlands [Rabaté and Rellstab, 2022; Artmann et al., 2022] and in Austria [Kleven et al., 2019a]. The child penalty is particularly high in Germany where mothers’ earnings are reduced by 80% compared to fathers one year after the birth and by 60% ten years after, which is in line with the fact that conservative gender role attitudes are still widespread in Germany [Kleven et al., 2019a]. Indeed, the division of home-production between men and women remains highly unequal. Nowadays, women still earn 20% less than men, have a probability to work part-time that is almost 5 times higher and spend 52% more time than men on home-production and care⁵, and with even more pronounced gaps among couples with children [Federal Ministry for Family Affairs and Youth, 2020]. This paper adds to this literature investigating how parents reallocate time after the first birth between paid work, unpaid work and leisure. To my knowledge, the work of Aguilar-Gomez et al. [2019] is, to date, the only available study investigating the question of unpaid work with this approach. The authors show that the change in time spent on unpaid tasks is about 3 times higher for women than for men shortly after the birth. They nonetheless point out the fact that their data does not allow to investigate any changes in unproductive time, or also called *leisure*, nor any potential implication on maternal mental load. This paper aims at filling the gap in exploring both changes in leisure time and well-being. Second, it adds to the literature that examines the effect of the number of children on maternal labour supply. Several instrumental variables have emerged in this literature, such as multiple births [Rosenzweig and Wolpin, 1980; Bronars and Grogger, 1994; Jacobsen et al., 1999; Black et al., 2005], infertility shocks [Aguero and Marks, 2008; Markussen and Strøm, 2022], fertility treatments [Lundborg et al., 2017] and, finally, the sex-composition in the sibship [Angrist and Evans, 1998] on which this paper builds on.⁶ This instrument is based on the documented parental preference for a sex-mixed sibship [Ben-Porath and Welch, 1976]. Because parents of

compared to men 5 to 10 years after the birth relative to the pre-birth period (usually 1 year before birth).

⁵In the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth’s report, “care” refers to both childcare and informal help provided to adults within or outside the household.

⁶The other instruments could not be implemented using GSOEP data because of information unavailability or lack of statistical power.

same-sex first-borns are more likely to have more than two children than other parents and that sex at birth is randomly assigned, the sex-composition of the first two children appears as a good candidate for causal identification. Angrist and Evans [1998] conclude on a negative effect of the number of children on mothers' labour supply and no effect for fathers in the United States. Cruces and Galiani [2007] replicate this method on Mexican and Argentinian data and find similar results. In Norway, Cools et al. [2017] show that having more children reduces mothers' opportunities on the labour market such as being employed by higher paying firms and being at the top of the wage distribution in their workplace. To the best of my knowledge, only one paper has investigated the effect of family size on the division of tasks between mothers and fathers in the previous literature. Using Canadian data, Frenette [2011] shows that mothers decrease the time spent on the labour market and reallocate it to unpaid tasks such as housework and childcare when the number of children increases. However, fathers do not change their working time, neither the time they spend on housework. They spend more time on childcare, though the increase is two times lower than for mothers. The contribution of my paper resides in several points. First, it provides further information on time-allocation, disentangling different types of domestic work (i.e. housework, errands and gardening/repairs) and of leisure (i.e. sleep, hobbies and other free-time). Second, the effect of family size on subjective well-being is estimated. Third, results are likely to differ compared to Canada due to the high prevalence of traditional gender role attitudes and low fertility in Germany. This paper therefore adds an element of comparison of the family size effect on time-allocation to the international scope.

Lastly, this paper relates to the literature on the division of tasks within couples. Despite an increase in men's involvement in domestic tasks and childcare over the last decades, women still provide the major part of unpaid tasks [Gimenez-Nadal and Sevilla, 2012; Champagne et al., 2015; D'Albis et al., 2016] and this is especially the case among parents [Baxter et al., 2008; Craig and Mullan, 2010]. The unequal division of unpaid work is not specific to couples following the *male breadwinner model*⁷ since women who earn more than their husbands compensate by providing more domestic work [Bittman et al., 2003; Bertrand et al., 2015; Lippmann et al., 2020]. This literature reveals the importance of gender identity norms in the unequal division of tasks which, in turn, negatively affects women's well-being [Flèche et al., 2020].

⁷A *single earner model* in which the man earns the total household income or its major part and the woman takes over the unpaid work in an heterosexual couple.

The remainder of this paper is organized as follows. Section 2 sheds light on the German context, Section 3 explains the identification strategies, Section 4 describes the data, Section 5 presents the results and Section 6 concludes.

2 The German context

Despite a greater female labour market participation nowadays, the traditional *male breadwinner model* is still widespread in Germany, especially in the West. Such geographical discrepancies are attributed by the literature to fifty years of different political regimes during the German division [Bauernschuster and Rainer, 2012; Lippmann et al., 2020].⁸

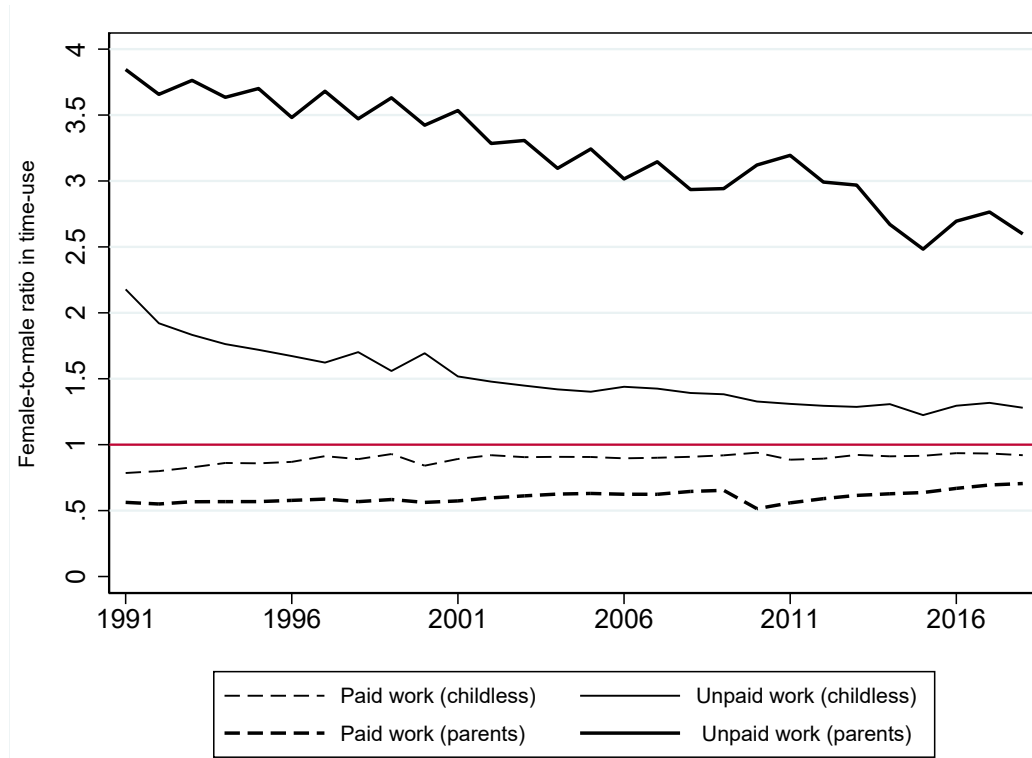
After reunification, family policies were aimed at reducing the East-West gap in public childcare provision. In 1996, a childcare reform was implemented to foster childcare access for children from three to six years old. The reform consisted in providing more highly subsidized half-day care slots for children from three years old until school enrollment and especially targeted households living in West Germany. The reform has been shown successful in rising female employment [Bauernschuster and Schlotter, 2015]. However, public childcare provision for children under three years old remained scarce until the mid 2000s and the private market for childcare was virtually non-existent. In 2002, only 2% of children aged under three were enrolled in day-care in West Germany while this was the case for 35% of them in East Germany. Childcare expansion laws emerged in 2005 and 2008 aiming at providing more accessibility to affordable public child care for children under three years old. From 2013, all children are guaranteed a subsidized day care slot after their first birthday. Müller and Wrohlich [2020] evaluate the effect of childcare expansion on maternal labour market participation and show that it contributed to an increase in part-time maternal employment.

At the same time, family policies were running counter to female employment. As fertility plummeted in East Germany after reunification⁹, public policies were aimed at boosting fertility. Thus, maternity leave gradually expanded until it reached an amount of 36 months of protected leave with low benefits at a flat rate. As expected, these reforms constituted inactivity traps for

⁸More generally, literature shows that the German divide induced differences in individual preferences such as taste for redistribution, public intervention [Alesina and Fuchs-Schündeln, 2007] and trust [Rainer and Siedler, 2009]

⁹This phenomenon has been attributed to the negative shock on labour demand due to the fall of the Soviet bloc [Lippmann, 2018].

Figure 1: Evolution of the time-use gender gap in Germany



Note: Two-way line plot between survey years and the ratio of female-to-male time-use. The red line represents perfect equality between men and women in terms of time-use (i.e. a null gender gap). Unpaid work refers to domestic tasks such as doing housework, running errands and taking care of children (if any).

Source : German Socio-Economic Panel (1991-2018).

Sample: Working age population (15-65 years old).

mothers who withdrew from the labour market after the birth [Dustmann and Schönberg, 2012; Schönberg and Ludsteck, 2014]. In 2007, a reform on parental leave was implemented with a two sided objective: encouraging mothers to resume work earlier after a birth and promoting paternity leave take-up which was very low. This policy consists in a 12-month protected leave for either the mother or the father at a 67% replacement rate of previous net labour earnings. Parents are granted two extra months if both father and mother take a leave. The total 14 months can be freely distributed between both parents. This reform induced an earlier return to work for mothers [Kluge and Tamm, 2013; Geyer et al., 2015], an increase in childcare time for fathers [Schober, 2014] as well as in fathers' housework time [Tamm, 2019].

In addition, unpaid work¹⁰ is still unequally distributed between men and women, especially among parents. Figure 1 displays the evolution of time-use gender gaps from 1991 to 2018 in

¹⁰Unpaid work refers to domestic tasks such as doing housework and caring for children.

Germany.¹¹ For both childless individuals and parents, gender gaps in paid and unpaid work weekly hours reduced through time, though with differences in levels and evolution trends. While the female-to-male ratio in time spent on paid work increased from 78% to 92% for individuals without children, it only increased from 53% to 62% among parents, implying that the gender gap in paid work time among parents in 2018 is still larger than it was among childless individuals in 1991. Concerning unpaid work, the time-use gender gap is larger among parents, irrespective of the survey year considered. In 1991, time spent on unpaid work was twice higher among childless women than among childless men and higher by only 28% in 2018. As far as parents are concerned, the female-to-male ratio is still higher than 2.5 in 2018.

3 Methodology

This section details the estimation strategies used to explore how men and women reallocate time with respect to parenthood. The first section examines how time-allocation evolves around the first child birth among men and women in Section 3.1. Then, I check whether the number of children plays a role in shaping gender gaps in time-allocation in Section 3.2. Results are presented in Section 5.

3.1 The first birth penalty

The first birth penalty, usually called *child penalty* in the literature, is estimated with an event-study approach around the first birth as introduced in Kleven et al. [2019a,b]. This methodology exploits the change in a given parental outcome after the first birth relative to a pre-birth period (usually one year before birth) keeping age and year constant. Only individuals who become parents during the observation period are included in the analysis. By design, event-study estimates are based on variations in timing of birth rather than fertility decisions.

The first birth penalty is estimated in two steps. First, the time-use outcome Y_{ht}^p of parent p , with $p = m$ for mothers and $p = f$ for fathers respectively, in household h in year t is regressed

¹¹The gender gap is computed as the ratio of female-to-male time-use for paid and unpaid work on an average week-day.

on a set of distance-to-event, age and year binary variables such as :

$$Y_{ht}^p = \sum_{d \neq -1} \alpha_d^p \cdot \mathbb{1}[Distance_{ht} = d] + \sum_a \alpha_a^p \cdot \mathbb{1}[Age_{ht}^p = a] + \sum_t \alpha_t^p \cdot \mathbb{1}[Year = t] + \epsilon_{ht}^p \quad (1)$$

Each binary variable is equal to 1 if the expression in the brackets is fulfilled, 0 otherwise. $Distance_{ht}$ takes values of the gap between survey year t and the event year b_h (i.e. the first child's birth year in household h) such as $d = t - b_h$. Following the existing literature, it ranges from -5 (i.e. 5 years before birth) to 10 (i.e. 10 years after birth). $d = -1$ is excluded so that each α_d^p captures the change in the outcome relative to the year before the first child birth. Age_{ht}^p refers to the age of parent p in household h at time t and $Year$ to the survey year. Age and survey year dummies control for life-cycle and time trends respectively as they might affect both birth timing and time-allocation. Each α_d^p is interpreted as the change in time-use for an individual observed d years after the birth (or before the birth if $d \in [-5; -2]$) relative to an individual observed the year prior to birth for a given age and a given year. It therefore captures variations in parental age at first birth.

The second step is the computation of the penalty for each parent at each period. It is defined as the change in time-allocation for parent p at distance-to-event d related to the average time-allocation value among parents-to-be (i.e. individuals observed in the reference period prior to birth) conditional on d . It is such as :

$$P_d^p = \frac{\hat{\alpha}_d^p}{E[\tilde{Y}_{ht}^p | d]} \quad (2)$$

where \tilde{Y}_{ht}^p are the predicted values from the regression described in Equation 1 when omitting the contribution of distance-to-event-dummies. More precisely, they are computed such as $\tilde{Y}_{ht}^p = \hat{Y}_{ht}^p - \hat{\alpha}_d^p$ with \hat{Y}_{ht}^p the predicted outcome values of Equation 1. Then, the gender gap in the first birth penalty is computed as follows for each d :

$$P_d^p = \frac{\hat{\alpha}_d^m - \hat{\alpha}_d^f}{E[\tilde{Y}_{ht}^m | d]} \quad (3)$$

Penalties for men and women are represented in Figure 3 taking paid work, unpaid work and

leisure time as outcomes respectively. The gender gap in first birth penalty for each value of d is presented in Table C1.¹²

It is important to keep in mind that interpreting the first birth penalty as the causal effect of entering into parenthood requires birth timing exogeneity, which is arguable. First, it might not be the case considering that some individuals might enter into parenthood earlier if they anticipate that they will have several children while investing less in their career (i.e. spend less time on paid work) due to lower expected returns associated with parenthood. In this case, the change in time-allocation after the birth relative to a pre-birth period might be overestimated because individuals who are not yet parents might work more than individuals who are already parents for the same age and year of observation. This risk is expected to increase with distance to the first birth because individuals are less likely to be similar when the difference in age at first birth is larger. Moreover, penalties estimated several years after the first birth are more likely to capture changes due to potential additional children. Nevertheless, penalties estimated within a short bandwidth around the event time (e.g. 1 or 2 years after the birth for example) should be less affected by this bias, even if its existence cannot be entirely ruled out.

Second, anticipation effects could arise if birth timing is not independent of expectations on future outcomes. If individuals anticipate their first birth and consequently adapt their time-allocation prior to birth (i.e. change to a job that reconciles work and family duties better), the discontinuity magnitude around the first birth is expected to be lower. If anticipation effects exist, first birth penalty estimates are lower bounds of the effect of entering into parenthood.

Event-study estimates should therefore be interpreted with caution, keeping in mind that the change in time-allocation around the first birth might not capture the effect of the first birth, especially for penalties estimated far from birth. First child penalties are nonetheless informative on gender differences in time-allocation with respect to parenthood.

3.2 The effect of family size

This section implements empirical strategies to examine whether the number of children has an effect on gender gaps in time-allocation. I first consider a regression of time-allocation on family

¹²As a further investigation, the same analysis is run taking well-being measures as outcomes. Penalties are reported in Figure 4 and gender gaps in well-being penalties can be found in Table C2.

size without correcting for endogeneity (Section 3.2.1) and I present the potential biases at stake in this relationship. I then introduce a fixed effects model (Section 3.2.2) and an instrumental variable model (Section 3.2.3) that yield causal estimates under specific assumptions that I describe later in the section.

3.2.1 Standard linear regression

A naive approach to determine how the number of children affects parental time-allocation would rely on a standard linear regression of a time-allocation measure on family size for each parent p , with $p = f$ for fathers and $p = m$ for mothers, of household h at time t such as:

$$Y_{ht}^p = \beta_0^p + \beta_1^p FamSize_{ht} + X_{ht}^m \beta_2^p + X_{ht}^f \beta_3^p + X_{ht}^h \beta_4^p + \theta_t^p + \mu_{ht}^p \quad (4)$$

Y_{ht}^p is the outcome of parent p of household h at time t , $FamSize_{ht}$ the variable characterizing family size in household h at time t , X_{ht}^m , X_{ht}^f and X_{ht}^h control variables vectors at the mother, the father and the household levels respectively, θ_t^p a year specific effect¹³ and μ_{ht}^p the error term. Variables are further detailed in Section 4. Coefficients are estimated using the Ordinary Least Squares (OLS) estimator.

Interpreting β_1^p as the effect of family size on time-allocation of parent p could only be possible with family size deemed exogenous, i.e. orthogonal to the error term. However, this assumption is not likely to hold since unobserved characteristics could be associated with both fertility decisions and time-allocation. At the individual level, *taste for labour* is expected positively correlated to time spent on paid work and negatively to time spent on unpaid work, leisure and to the number of children. Omitting this characteristic would result in a negative bias for paid work and a positive bias for unpaid work and leisure, both for men and women. At the household level, couples following a conservative family scheme might have a higher number of children along with an unequal division of tasks. In this case, mothers (resp. fathers) are likely to provide more (resp. less) unpaid work than paid work. For mothers, the bias would be negative for paid work and positive for unpaid work while it would be the opposite for fathers. Then, estimates could suffer from a systematic measurement error if parents who have more children are more

¹³The year specific effect is denoted by θ_t^p for simplification matters but is defined such as $\theta_t^p = \sum_t \theta_t^p \cdot \mathbb{1}[Year = t]$ with $Year$ the survey year variable and t the values that $Year$ can take.

prone to misreporting their time-allocation (e.g. parents of large families might be more likely to conduct tasks simultaneously, which might lead to more misreporting). In this case, the bias sign depends on whether individuals tend to overestimate or underestimate time spent on a given task when they have more children. This point is further discussed in Section 5.

Consequently, Equation 4 *a priori* fails to identify any causal effect of family size on time-allocation of mothers and fathers. If the presence of a bias is very likely, whether estimates are underestimated or overestimated is undetermined. For this reason, I use a fixed effects model to control for time-invariant characteristics at the individual level (Section 3.2.2) and an instrumental variable strategy (Section 3.2.3) in an attempt to correct for both time-variant and time-invariant unobserved characteristics.

3.2.2 Fixed effects model

I first consider a fixed effects model written as follows :

$$Y_{ht}^p = \gamma_h^p + \gamma_1^p FamSize_{ht} + X_{ht}^{m'} \gamma_2^p + X_{ht}^{f'} \gamma_3^p + X_{ht}' \gamma_4^p + \lambda_t^p + \nu_{ht}^p \quad (5)$$

γ_h^p is the individual fixed effect such as $\gamma_h^p = \gamma_0 + X_h^{m'} \rho_1^p + X_h^{f'} \rho_2^p + X_h' \rho_3^p$ with γ_0 the intercept, $X_h^{m'}$, $X_h^{f'}$ and X_h' vectors of time-invariant control variables at the mother, the father and the household level respectively. As in Equation 4, the model also contains a time component (here, λ_t^p) and time-variant control variable vectors. γ_1^p is computed using the *within* estimator [Angrist and Pischke, 2009].

Since individual fixed effect models only correct for time-invariant characteristics, the within estimator is unlikely to yield a causal estimate if time-variant confounding variables are at stake. In our case, deeming taste for labour, gender role attitudes or misreporting as constant over time would be difficult to claim. For this reason, I rely on an instrumental variable strategy in the next section.

3.2.3 Instrumental variable strategy

Assumptions. Identifying a causal effect with an instrumental variable strategy requires to fulfill the following assumptions : *relevance, independence, exclusion restriction and monotonicity*

[Angrist et al., 1996]. Under these assumptions, the effect of the number of children is yielded by the Wald estimator :

$$\tau_{Wald} = \frac{E[Y_{ht}^p | Z_h = 1] - E[Y_{ht}^p | Z_h = 0]}{E[FamSize_{ht} | Z_h = 1] - E[FamSize_{ht} | Z_h = 0]} \quad (6)$$

where τ_{Wald} is the Local Average Treatment Effect (LATE) of family size on parental time-allocation, i.e. the effect among individuals who react to the instrument or *compliers*.

Considering these conditions, I use the sex-composition among the first two children as an instrument to predict family size [Angrist and Evans, 1998]. Technically, the instrument is a binary variable equal to 1 if the couple has two first children of the same sex (i.e. two sons or two daughters), 0 otherwise (i.e. one son and one daughter) and is denoted as Z_h . Family size, denoted as $FamSize_{ht}$, is a binary variable equal to 1 if parents of household h have more than two children and 0 if they do not have supplementary children at time t . The intuition of this instrument is based on the documented parental preferences with respect to the sex-composition of the sibship. Ben-Porath and Welch [1976] show that parents who have same-sex first-borns have a higher probability to have more than two children than parents with a sex-mixed sibship. In other words, they suggest the existence of a parental *taste for balance* driving the final family size. This instrument appears as a good candidate for identification for several reasons.

Relevance. The first-stage relationship must be large enough in order for the instrumental variable to be relevant, otherwise, it is considered as weak as it increases the relative TSLS bias. Relevance can be checked looking at first-stage results and particularly at the Fisher statistic for the weak instrument test.¹⁴ In the case of a single instrument variable, the threshold suggested by the econometric literature above which the weak instrument hypothesis is rejected is $F > 10$ [Stock et al., 2002]. First-stage results are available in Table 1 and show that the instrument fulfills the relevance assumption in this context as it is strongly correlated to family size. This point is further detailed in Section 5.2.

Independence. This assumption requires the instrumental variable to be uncorrelated to the error term. Unlike relevance, independence cannot directly be tested. Nevertheless, knowing

¹⁴The null hypothesis of this test is such that the instrument is only weakly correlated to the instrumented variable, i.e. the instrument is irrelevant. As standard-errors are corrected for clustering at the individual level in regressions, Kleibergen-Paap F statistics are reported.

that sex at birth is randomly assigned, the sex-composition of the first two children should also be random by construction. Mean comparison tests covariates between parents of same-sex versus mixed-sex first-borns are in line with this hypothesis. They show that both groups do not significantly differ with respect to observables¹⁵ (see Table A1).

Exclusion restriction. The exclusion restriction hypothesis is a more challenging assumption as it requires that the instrumental variable affects the dependent variable only through its relationship with the instrumented variable. Violating this assumption would suggest that the sex-composition of the first two children has an effect on parental time-allocation through another channel than via the number of children. Two potential effects could be at stake in this context.

First, a *direct gender effect* could drive parental time-allocation without transiting through family size. Some studies have shown that fathers have a higher involvement within the household with sons than with daughters. More precisely, girls who grow up with brothers have more involved fathers than girls with sisters [Butcher and Case, 1994]. Fathers are more likely to spend time with their children when the sibship is only constituted of boys compared to a sibship of girls [Mammen, 2011]. In addition, considering that girls contribute more to housework duties than their brothers [Cools et al., 2017] and increase their contribution when the sibship size is larger [Menta and Lepinteur, 2021], direct gender effects might affect parental time-allocation through daughters' investment in housework chores. Though the absence of such an effect is arguable in the light of this literature, I test this hypothesis by performing an over-identification test¹⁶ and show that the IV strategy validity is not compromised by such direct effects. The results of this test are further detailed in Section 5.2.2.

Second, *direct gender-composition effects* could also occur considering a different time-allocation when rearing same-sex children as opposed to a mixed-sex sibship. In households following traditional gender norms, parents of a mixed-sex sibship might spend more time on childcare than other parents if they consider that girls and boys should be treated differently. For example, parents enrolling daughters to different types of extracurricular activities compared to sons

¹⁵Both groups only differ with respect to the share of highly educated fathers, but only at the 10% significance level.

¹⁶The hypothesis is tested using the Hansen-Sargan test. To this aim, two different instruments are used: *two sons* and *two daughters*. *Two sons* (resp. *Two daughters*) is a binary variable equal to 1 when the first two borns are boys (resp. girls), 0 otherwise. Rejecting the null hypothesis casts doubt on the validity of the IV strategy as at least one of both instruments is likely to be correlated with the error term [Wooldridge, 2015].

because of gender stereotypes would induce more commuting time associated with childcare for parents of a mixed-sex sibship than for parents of same-sex children. To rule out the possibility of any *direct gender-composition effect* on parental time allocation, a mean comparison test is conducted on time-allocation variables between parents of same-sex children and parents of a son and a daughter before the birth of any third child. The results of this test show that parents of same-sex first-borns do not allocate time differently compared to parents of a mixed-sex sibship before any third birth (see Table A2).

Monotonicity. It suggests the absence of defiers. In this context, defiers are couples who would choose to have more than two children when the first two borns are of the opposite sex but not when they are of the same sex. Such couples would have a preference for having at least two children of the same-sex in the sibship. However, De Chaisemartin [2017] argues that the IV validity holds under a weaker assumption that tolerates the presence of defiers. He shows that the presence of defiers does not impair the validity of the IV if the ratio between the LATE of compliers and the one estimated for defiers is lower than a bound he estimates from the data used by Angrist and Evans [1998]. Arguing that this upper bound is large and that the LATE of compliers should not dramatically differ from the LATE of defiers in this case, this assumption is expected to hold.

Estimation. The identification strategy relies on replacing, in a second-stage equation (Equation 8), the endogenous variable (here, $FamSize_{ht}$) by its fitted value obtained from a first-stage equation (Equation 7) that regresses this variable on the instrument Z_h such as:

$$FamSize_{ht} = \delta_0 + \delta_1 Z_h + X_{ht}^m \delta_2 + X_{ht}^f \delta_3 + X_{ht}' \delta_4 + \phi_t + v_{ht} \quad (7)$$

$$Y_{ht}^p = \delta_0^p + \delta_1^p \widehat{FamSize}_{ht} + X_{ht}^m \delta_2^p + X_{ht}^f \delta_3^p + X_{ht}' \delta_4^p + \phi_t^p + v_{ht}^p \quad (8)$$

Under the IV strategy assumptions, δ_1^p captures the LATE and is computed by the Two-Stage Least Squares (TSLS) estimator. Similarly to λ_t^p in Equation 4, ϕ_t and ϕ_t^p denote year specific effects.

4 Data

All the analyses are based on the German Socio Economic Panel (GSOEP), a longitudinal survey that collects information on various topics at the household and the individual levels. The survey started in 1984 and the last available wave took place in 2018. Before reunification, the dataset only included households who used to live in West Germany. For this reason, the sample is restricted on observations collected from 1991 onwards. Only parents who live in the same household are part of the analysis in order to focus on the intra-household repartition of tasks. Then, only heterosexual couples are included in the sample because this paper focuses on gender discrepancies with regard to time-allocation. The remainder of this section presents the samples of interest in Section 4.1 and describes the variables in Section 4.2.

4.1 Sample restrictions

Event-study analysis. In addition to restrictions mentioned above, the sample of interest for the event-study framework contains individuals who experience their first birth between the observation period (i.e. 1991-2018). Following the previous literature, the sample is further restricted to individuals who are observed at least once before and once after the reference period (i.e. one year before birth) and who are between 20 and 45 years old at the time of the first birth. The sample contains 10,806 observations and 1,263 households. Table B1 reports average values of pre-birth variables of interest.

Instrumental variable analysis. The nature of the instrumental variable used in the TSLS estimation (i.e. the sex-composition of the first two children) implies to further restrict the sample on parents who have at least two children at the time of interview. Only parents of biological children are included to ensure the random distribution of children's sex. Individuals whose second child is older than 18 years old at the time of interview are excluded as they are more likely to live outside the household. Taking all these restrictions into account, the final sample holds 50,149 observations for 9,737 households. Table B2 describes this sample of interest.

4.2 Description of variables

Time-allocation. Individuals are asked about the number of hours they spend on paid work, unpaid work and leisure in an average week day. *Paid work* refers to time spent working for pay including all types of jobs (main occupation, additional jobs and apprenticeship) as well as commuting time. *Unpaid work* covers domestic tasks such as housework (washing, cooking, cleaning), errands, repairs/gardening and childcare. *Leisure* includes all remaining time so that the total amount of daily hours amounts 24. Main analyses are run using these three items as outcomes. Then, I run further analyses for each sub-item of unpaid work (i.e. housework, errands, repairs/gardening and childcare) and leisure is split between sleeping time, time spent on hobbies and any other free-time activity unspecified in the survey (see Section 5.2.4).

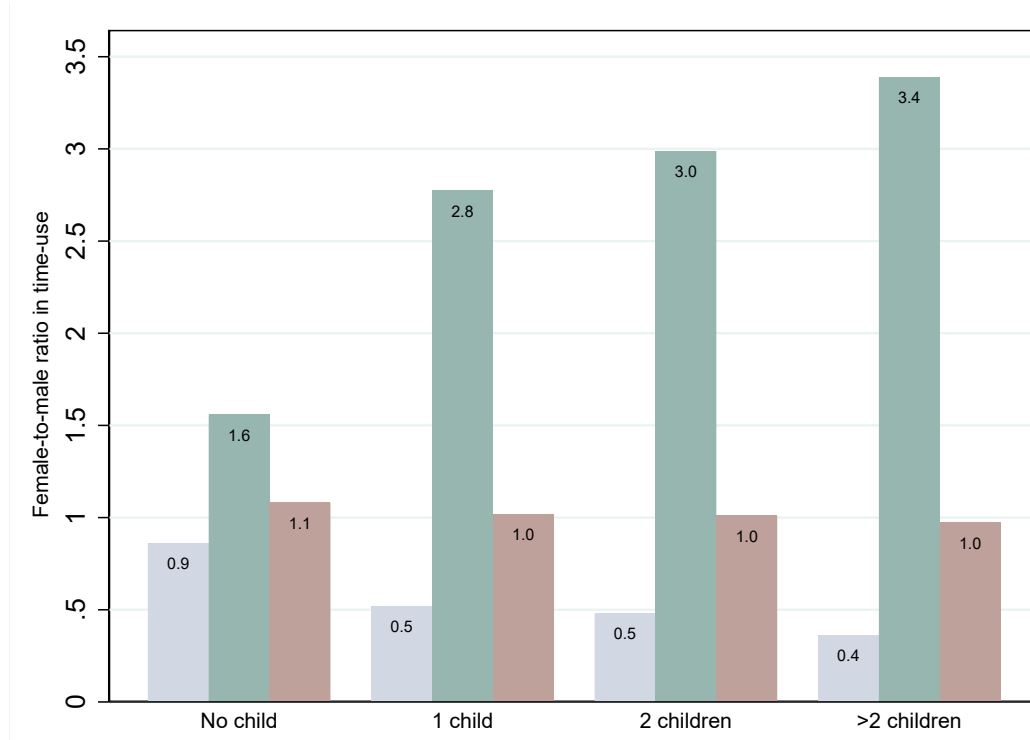
Figure 2 shows how these tasks are divided between men and women among childless couples, couples with one, two and more than two children respectively. The intra-household division of tasks is measured as the female-to-male ratio in daily time used for each of the time items mentioned above. It shows that women spend less time than their male partners on paid work, whether they have children or not, but the ratio is lower among couples with children. Women without children spend 86% of their partner's working time on average while women with one or two children work about half less than them (52% and 48% of male working hours respectively). The female-to-male ratio in paid work drops to 36% when it comes to couples with a larger family size (i.e. at least 3 children). Women take on more unpaid tasks, even among childless couples. While women spend 1.5 more daily time on unpaid activities than their partners, the female-to-male ratio is equal to 2.8 for couples with one child, to 3 for couples with two children while it reaches 3.4 in households with more than two children. Concerning leisure time, the ratio is about 1 in all groups, meaning that both partners allocate the same amount of daily time to "unproductive tasks" irrespective of the number of children they have.

Well-being. Well-being is measured by satisfaction scores declared by individuals on a scale ranging from 0 "*Completely dissatisfied*" to 10 "*Completely satisfied*" concerning paid work, housework, childcare and leisure.¹⁷ Average values are presented in Table B1 for the sample of parents observed around the first birth and in Table B2 for parents of at least two children.

¹⁷Other satisfaction scores are available in the GSOEP but they are not used because they contain high proportions of missing values. Scores that do not provide enough statistical power to run the instrumental variable strategy are excluded from the scope of all analyses.

In both cases, satisfaction scores are relatively high (i.e. around 7) and are virtually the same between men and women.

Figure 2: Intra-household time-allocation by family size



Note: Bar chart of average female-to-male ratio in time-use by number of children. Blue (resp. green, pink) bars display the ratio for time used for paid work (resp. unpaid work, leisure). Averages are computed among individuals subject to sample restrictions as described in the first paragraph of Section 4 without any restriction on the number of children.

Source : German Socio-Economic Panel (1991-2018).

Family size. In the instrumental variable strategy, family size is predicted by a variable defining whether the first two children are of the same-sex or not. Concretely, it is a binary variable equal to 1 if parents have two sons or two daughters as first-borns, 0 otherwise. By construction, family size is a binary variable taking 1 if both parents have more than two children at the time of interview and 0 if they have exactly two children. The same variable is used to measure family size in the OLS and FE regressions for results to be comparable. Conditional on having two children, the probability to have least three children in the sample is almost 36% (see Table B2).

Individual characteristics. Regressions presented in Section 3.2 include vectors of variables at the mother, the father and the household levels. I include the mother's, the father's, the first

born child’s and the second born child’s age respectively.¹⁸ At the time of interview, the first-born child is around 13 years old and the second one is almost 9 on average. Mothers are 39 years old on average at the time of interview and fathers reach 42 (Table B2). Following Angrist and Evans [1998], education levels are included as covariates since they are expected to be correlated to both time-allocation and fertility choices. For each parent, education level is defined as a binary variable taking 1 if the individual has completed the first level of tertiary education at least and 0 for lower education levels.¹⁹ 38% of mothers have completed at least the first stage of tertiary education and 43% of fathers did (Table B2). It is nonetheless important to note that, in the case education is not completed yet when the decision to have an additional child is determined, education could be itself influenced by family size. As a result, education levels could be *bad controls*²⁰ and should not be included in regressions to avoid biasing estimates. As a robustness check, education levels are excluded and results are presented in Section 5.2.2.

5 Results

The results section is organised as follows. First, event study estimates around the first birth are presented for time-allocation and well-being outcomes in Section 5.1. Then, results with respect to the effect family size on time-allocation are presented in Section 5.2.1, the internal validity of results is checked in Section 5.2.2, heterogeneity is explored in Section 5.2.3 and further results are investigated in Section 5.2.4.

5.1 The first birth penalty

First birth penalties in time-allocation (resp. well-being) are represented in Figure 3 (resp. Figure 4) for men (in grey) and women (in blue) separately. Penalties are computed among parents with only one child (dashed lines) and parents with several children (solid lines). Gender gaps in first birth penalties are reported in Table C1 for time-allocation and in Table C2 for well-being outcomes.

Time-allocation. For all time-use outcomes, trajectories before the event (i.e. coefficients

¹⁸In the fixed effects model, these variables are dropped due to multi-collinearity.

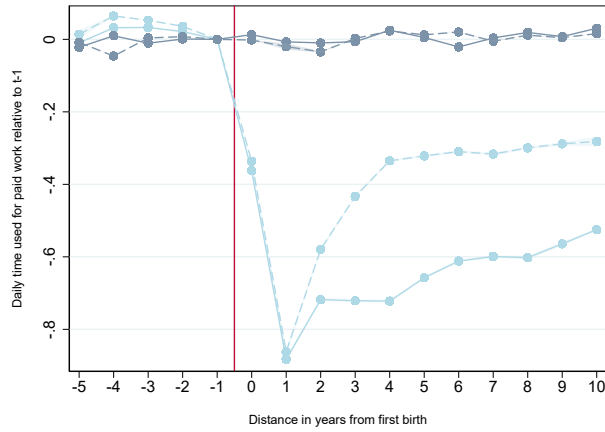
¹⁹Levels of education are based on the International Standard Classification of Education (ISCED), 1997 version.

²⁰According to Angrist and Pischke [2009], *bad controls* are variables that might not be “fixed at the time the regressor of interest is determined”.

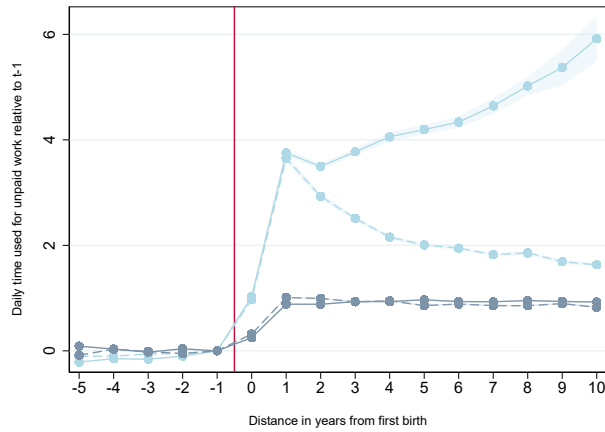
estimated for $-5 < d < -2$ relative to $d = -1$ are virtually smooth and close to zero for both men and women, supporting the absence of any anticipation effect. Figure 3a reports penalties when daily hours of paid work are taken as an outcome and shows a striking gender difference in trends after the birth of the first child. Daily hours spent on paid work are lower by 34% for women with one child and by 38% for women with more children compared to men in the year of birth ($d = 0$) relative to one year before and by 83% (resp. 88%) one year after the birth while the male trajectory is only slightly affected and remains close to zero. In line with previous literature, the penalty gets lower with distance to birth but never converges back to the pre-birth level. Moreover, convergence trends differ with respect to the number of children born over the period. Ten years after the birth, the gender gap in the first birth penalty is about 30% for parents with one child while it reaches 55% when family size is higher.

Figure 3b displays the coefficients for daily hours spent on unpaid tasks. Once again, female and male trends largely differ after the birth. More precisely, time spent on unpaid tasks is lower by about 10% for men one year after the birth while it is almost multiplied by 4 for women compared to one year before birth. The male penalty remains essentially the same irrespective of distance with the first birth and of the number of children born afterwards. As far as women are concerned, the penalty is lower (resp. higher) with distance to birth for mothers of one child (resp. several children). Ten years after the birth, women with one child spend twice their partners' time on unpaid tasks relative to one year before (i.e. the gender gap in penalty of about 1) while the gender gap in penalty reaches 4.6 for couples with several children.

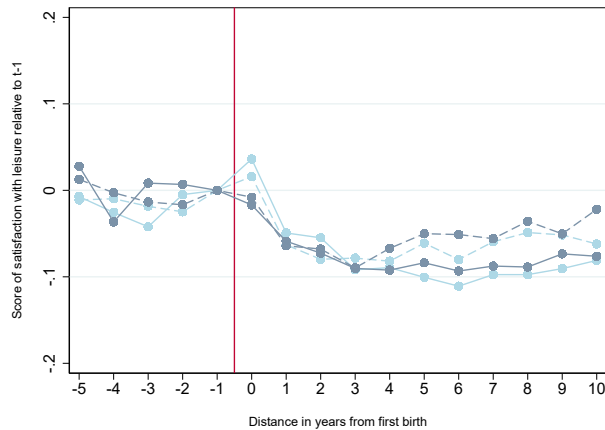
Figure 3: First birth penalties in time-allocation



(a) Time for paid work



(b) Time for unpaid work



(c) Time for leisure

Note : Event-study estimates by distance-to-event d from -5 (i.e. 5 years before the first birth) to 10 (i.e. 10 years after) as described in Section 3.1 with (a) time for paid work, (b) time for unpaid work and (c) time for leisure as an outcome. For each d , the coefficient captures the change in a given time-use outcome d years after the birth relative to one year before the birth as a percentage of the outcome among not-yet parents (i.e. the counterfactual). Male (resp. female) estimates are represented in grey (resp. blue) among individuals who have only one child over the period (dashed lines) and individuals with several children (solid lines). Confidence intervals at the level of 95% are represented by shaded areas.

Source : German Socio-Economic Panel (1991-2018).

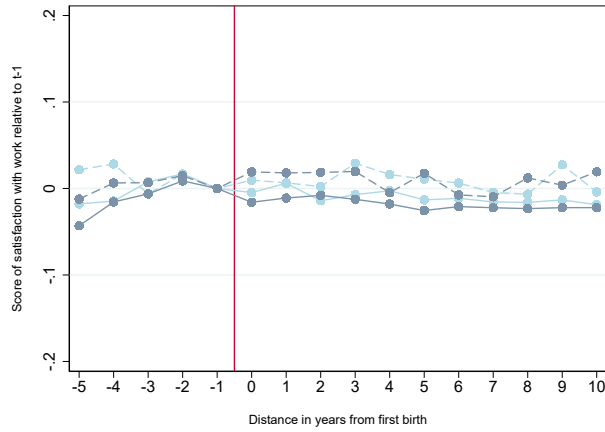
The estimate for leisure time is positive in the year of birth only for mothers, which is consistent with the fact that mothers are likely to be on leave at least part of that year. After the birth, both male and female daily leisure time is lower after the birth compared to one year before, but the penalty is of a higher magnitude among women. More precisely, mothers' leisure time fall behind men's by about 15% of not-yet parents' leisure time one year after the birth compared to one year before. The gender gap in child penalty is lower several years after birth and this is especially the case among one-child parents (see Figure 3c).

For all time-use outcomes, trends differ according to the number of children couples have over the period. This result suggests that the number of children plays a role, either because each additional child increases the child penalty or because of selection effects (e.g. if women with more children have a higher probability to work part-time after the birth, or if they meet more difficulties to be hired or promoted for example). Therefore, differences in trends should not be interpreted as the causal effect of family size on time allocation. This question is further explored with another strategy in the next section.

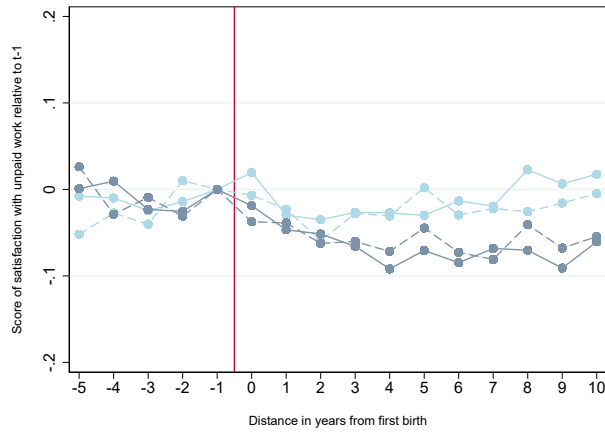
As previously explained in Section 3.1, it is also important to recall that coefficients, especially when estimated far from the birth threshold, should not be interpreted as causal effects since they are also likely to capture differences in time-allocation due to unobserved characteristics linked to birth timing. Coefficients estimated close to birth should nevertheless be less affected by this bias. In any case, penalties should rather be interpreted as the change in time used d years after the birth relative to one year before as a percentage of time used among not-yet parents for a given age and year.

Well-being. Figure 4 reports event-study estimates taking well-being measures as outcomes. With respect to satisfaction with paid work, post-birth coefficients are close to zero for both men and women among one-child as well as several-children couples (Figure 4a).

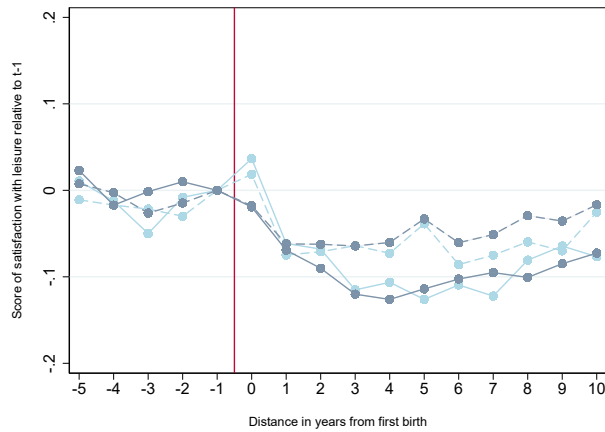
Figure 4: Child penalties in well-being



(a) Satisfaction with paid work



(b) Satisfaction with housework



(c) Satisfaction with leisure

Note : Event-study estimates by distance-to-event d from -5 (i.e. 5 years before the first birth) to 10 (i.e. 10 years after) as described in Section 3.1 with (a) satisfaction with paid work, (b) satisfaction with housework and (c) satisfaction with leisure as an outcome. For each d , the coefficient captures the change in a given time-use outcome d years after the birth relative to one year before the birth as a percentage of the outcome among not-yet parents (i.e. the counterfactual). Male (resp. female) estimates are represented in grey (resp. blue) among individuals who have only one child over the period (dashed lines) and individuals with several children (solid lines). Confidence intervals at the level of 95% are represented by shaded areas.

Source : German Socio-Economic Panel (1991-2018).

Concerning housework, the drop in level of satisfaction is higher among fathers than mothers, resulting in a gender gap in first birth of 3% (resp. 4%) one year after the birth and reaching 8% (resp. 9%) ten years after among one-child parents (resp. parents of several children), which can be surprising considering that the positive change in unpaid work after the first birth is higher for women than for men. This might stem from the fact that men contribute less to unpaid tasks before the birth and do not expect the housework load while women integrate that they are expected to provide the largest part of unpaid work when entering into parenthood (Figure 4b).

The last result relates to satisfaction with leisure. Figure 4c shows that daily time spent on leisure is higher in the year of birth compared to one year before only for women, consistent with previous results found on leisure time. One year after the birth, satisfaction with leisure is lower compared to the pre-birth period and estimates are very similar between men and women for both one-child and several-children couples. Above one year after the birth, estimates are closer to 0 for parents with one child compared to parents with several children, but trends are similar between men and women.

5.2 Family size effects

5.2.1 Main results

First-stage results of the instrumental variable strategy are in line with the previous literature : parents who have two first children of the same sex are more likely to have more than two children than parents who have a son and a daughter by 6.6 percentage points ($p < 0.01$). Thus, the *same-sex* instrument is also a predictor of family size in Germany. Moreover, the F-statistic of the weak instrument test equals 34.065, which is far above the threshold that determines the instrument's relevance (see Table 1). Reduced-form results are also presented in the table and provide evidence of an effect of the instrumental variable on a re-allocation of time from the labour market (-0.186 daily hours, $p < 0.01$) to home-production (+ 0.199 daily hours, $p < 0.05$) for mothers. Fathers' time-allocation remains unchanged.

Table 2 reports second-stage results of the instrumental variable strategy and shows that having three children or more decreases time spent on paid work by 2.811 daily hours ($p < 0.01$) and increases time allocated to unpaid tasks by 3.021 hours per day ($p < 0.01$) for mothers,

Table 1: First-stage and reduced-form estimates

	<i>First-stage</i>	<i>Reduced-form</i>		
	More than two children	Paid work	Unpaid work	Leisure
<i>Mothers</i>				
Same-sex children	0.066*** (0.011)	-0.186*** (0.071)	0.199** (0.080)	-0.014 (0.061)
<i>Fathers</i>				
Same-sex children	0.066*** (0.011)	-0.000 (0.057)	0.015 (0.049)	-0.014 (0.048)
<i>Weak instrument test (F-stat)</i>	34.065			
Number of individuals	9,737	9,737	9,737	9,737
Number of observations	50,149	50,149	50,149	50,149

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01. Source : German Socio-Economic Panel (1991-2018).

representing a decrease in 12% and an increase in 13% of total daily time respectively²¹ while leisure time remains significantly unaffected. Fathers do not significantly reallocate their time-use when they have more children.

As explained earlier in Section 3.2, several sources of bias could exist in the relationship between family size and time-allocation. Consequently, the bias sign is undetermined *a priori*.

When comparing OLS with TSLS estimates, it appears that the bias is positive for paid work and negative for unpaid work for mothers. Interestingly, the bias sign is reversed for fathers : it is negative for paid work and positive for unpaid work. This result suggests that a misreporting bias might be at stake. Since mothers tend to overestimate the time they spend on paid work and underestimate the time spent on unpaid tasks when they have larger family sizes while it is the opposite for men, it could be the case that some parents report their time-allocation differently if they consider that another time-allocation is more socially desirable with respect to the number of children they have (i.e. parents of large families with an unequal division of tasks within the couple might want to provide a more progressive view of their gender roles attitudes to the interviewer). This result is consistent with the fact that women tend to underestimate the amount of time they spend on housework in surveys compared to diary settings while it is the opposite for fathers. Interestingly, these systematic gaps in reported hours increase with the number of children in the household [Kan, 2008], which is in line with the results I put forward.

²¹2.11/24=0.117 and 3.021/24=0.125.

Table 2: Effect of family size on time-allocation

	Paid work	Unpaid work	Leisure
Mothers			
OLS estimate	-1.680*** (0.073)	2.628*** (0.083)	-0.948*** (0.067)
FE estimate	-1.625*** (0.116)	3.386*** (0.160)	-1.761*** (0.133)
TSLS estimate	-2.811*** (1.045)	3.021*** (1.117)	-0.210 (0.920)
<i>Endogeneity test (p-value)</i>	0.272	0.724	0.417
Fathers			
OLS estimate	-0.358*** (0.065)	0.396*** (0.055)	-0.038 (0.054)
FE estimate	0.351*** (0.125)	0.438*** (0.102)	-0.789*** (0.117)
TSLS estimate	-0.020 (0.863)	0.229 (0.739)	-0.209 (0.733)
<i>Endogeneity test (p-value)</i>	0.694	0.821	0.815
Number of individuals	9,737	9,737	9,737
Number of observations	50,149	50,149	50,149

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source : German Socio-Economic Panel (1991-2018).

When comparing OLS with FE estimates, the bias is negative for paid work and unpaid work and positive for leisure for both parents. The negative bias on paid work is in line with a *taste-for-labour* bias captured by the FE model, suggesting that career-attachment is, at least partly, time-varying. However, the negative bias with respect to unpaid work for both parents provides unclear information on what is at stake. For mothers, the bias goes in the same way and when comparing OLS with TSLS estimates, suggesting that the FE model might also capture a *social desirability bias*. For fathers, the negative bias suggests that the FE model corrects for traditional gender norms that lead fathers to prioritise paid work over unpaid work when the number of children increases. This also suggests that gender norms vary with time.

In order to investigate whether the negative effect on mothers' working time stems from a withdrawal from the labour market or a reduction in hours among mothers who work, I explore the extensive and intensive margins of the effect. Table 3 shows that the negative effect of family size on mother's working time is entirely extensive as having more than two children decreases the maternal employment rate by 31 percentage points while mothers who stay on the labour

Table 3: Effects of family size on paid work: intensive and extensive margins

	Working	Conditional working hours	Working part-time
Mothers			
OLS estimate	-0.199*** (0.009)	-0.764*** (0.071)	0.088*** (0.012)
FE estimate	-0.195*** (0.017)	-0.935*** (0.153)	0.095*** (0.023)
TSLS estimate	-0.310** (0.126)	-1.241 (1.059)	0.199 (0.185)
<i>Endogeneity test (p-value)</i>	0.373	0.650	0.538
<i>Weak instrument test (F-stat)</i>	34.065	24.242	24.242
Number of individuals	9,737	7,478	7,478
Number of observations	50,149	32,140	32,140
Fathers			
OLS estimate	-0.037*** (0.005)	-0.008 (0.039)	0.012*** (0.004)
FE estimate	0.024** (0.012)	0.178*** (0.065)	-0.002 (0.009)
TSLS estimate	-0.001 (0.072)	-0.015 (0.538)	-0.052 (0.055)
<i>Endogeneity test (p-value)</i>	0.618	0.990	0.235
<i>Weak instrument test (F-stat)</i>	34.065	31.754	31.754
Number of individuals	9,737	9,257	9,257
Number of observations	50,149	45,840	45,840

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01. Conditional working time corresponds to working hours among working individuals (i.e. excluding null working time). *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

market do not significantly reduce their amount of working hours. Despite that the intensive margin coefficient is lower in absolute value compared to the average result (-1.241 daily hours), I cannot discard the fact that the lack of significance might be explained by a loss in precision due to the lowered sample size.

5.2.2 Internal validity

Robustness checks. In order to assess the internal validity of the instrumental variable results, several robustness checks are implemented. First, regressions are run excluding education levels from the control variables vectors to ensure that they are not *bad controls*. Results show that excluding education levels provide virtually unchanged estimates (see Table E1).

Given the evidence on persistent differences in gender role attitudes between East and West Germany (see Section 2), family size can be expected to have a different effect across these two regions. Since the traditional *male breadwinner model* is more widespread in West Germany, one can expect couples to divide tasks unequally even before having children. In this case, having an additional child might not play an important role in the time-allocation gender gap in West Germany. To check that results are not entirely driven by individuals with less conservative gender norms, former East German states (i.e. Berlin, Brandenburg, Mecklenburg-Vorpommern, Saxony, Saxony-Anhalt and Thuringia) are excluded from the sample. Results are similar, implying that family size also plays a role among couples who might already have an uneven division of unpaid tasks (see Table E1).²²

Using the sex-composition as an instrument for family size might yield biased estimates if the probability of having a son is not exactly 50% at each birth. In this case, the probability to have same-sex children would be correlated to the sex of each child, occurring a bias due to potential omitted variables. As proposed by Angrist and Evans [1998], controls for the sex of both first and second children are included in the regression to check for potential *direct gender effects* (see Section 3.2) due to the deviation of the probability to have a son from exactly 50% in the data. Coefficients remain virtually unchanged, which mitigates the existence of any direct gender effect (see Table E1)

It is important to keep in mind that the length of maternity leave coverage is high in Germany, which could cause the negative effect of family size on the extensive margin of labour supply to be temporary. I check whether this is the case or not taking the probability to be on leave at the time of interview as an outcome. Table E2 shows that family size does not have any significant effect on the probability to be on parental leave, which supports the idea that mothers withdraw somewhat definitively from the labour market.

Testing the exclusion restriction assumption. In the case the sex-composition of the first two children affects parental time-allocation through another channel than the number of children, the exclusion restriction assumption is unfulfilled. Knowing that *direct gender effects* might be at stake due to findings in the previous literature (see Section 3.2), this assumption could

²²Ideally, one would also want to run the analysis on East Germany only to determine whether the effects are similar or not across Germany. Unfortunately, the smaller sample size in East Germany does not allow to identify any causal effect properly due to a lack of precision (F -stat < 10).

be violated. In order to know whether or not this is the case, I conduct an over-identification test. Testing for over-identification restrictions requires to have more instruments than instrumented variables in the regression. For this purpose, the sex-composition variable is split into two different instruments: a binary variable taking 1 if the first two children are girls, 0 otherwise, and a second variable taking 1 if the first two children are boys, 0 otherwise [Angrist and Evans, 1998]. In this test, rejecting the null hypothesis (i.e. the IV is uncorrelated to the error term) casts doubts on the suitability of at least one of the instruments. Table E3 reports the results of the test for each time-use outcome. The first-stage coefficient is higher when using the *two sons* instrument compared to the *two daughters* instrument but coefficients do not significantly differ from one another as shown by the result of the equality of coefficients test (p-value=0.299). For all time-use variables, and for both parents, the null hypothesis of the over-identification test is not rejected, ruling out any violation of the exclusion restriction assumption.

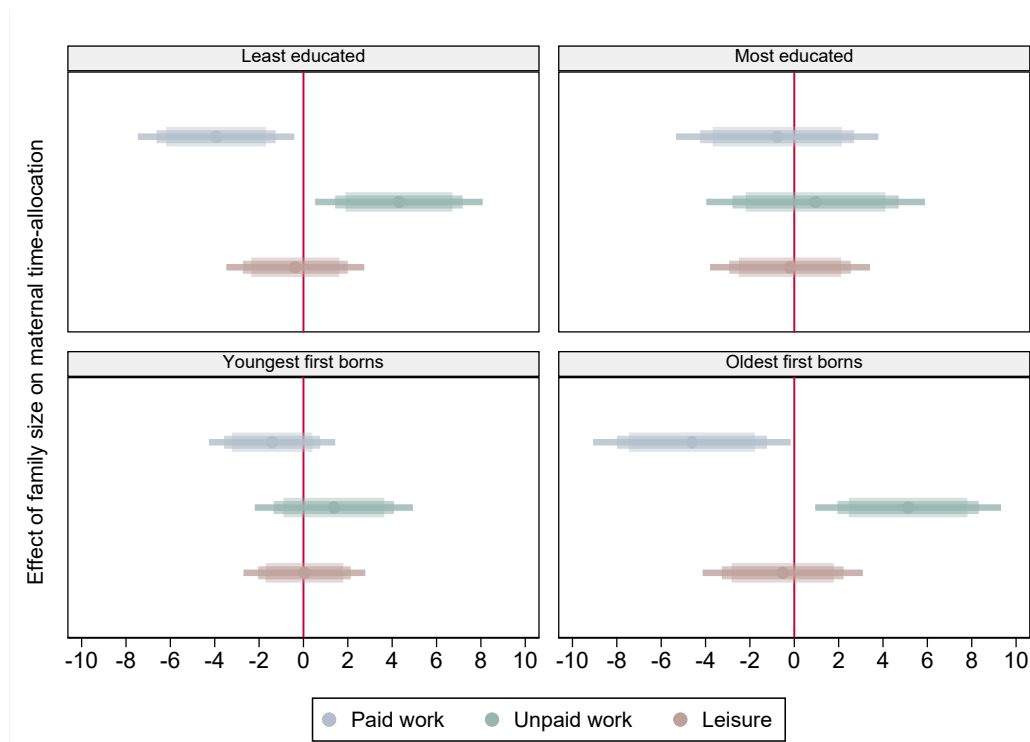
5.2.3 Heterogeneous results

Because the effect of family size on time-allocation might be different according to some individual characteristics, I run heterogeneity analyses. The most appropriate methodology would be to interact the characteristic of interest with the dependent variable in the second-stage and with the instrument in the first-stage to determine whether effects differ significantly across individuals or not. However, adding an interacted instrument requires large samples to provide estimates that are precise enough to be properly interpreted. Here, samples are too small to run this type of heterogeneity analysis. Alternatively, analyses are conducted separately across groups.

Education levels. First, I investigate whether the effect is different or not between the least and the most educated parents of the sample. For this purpose, the sample is divided into two distinct populations: the least (most) educated category defines individuals who completed high school at most (at least).²³ Figure 5 (resp. figure 5) present coefficients with confidence intervals at the level of 95% for each time-use variable for the least versus the most educated mothers (resp. fathers) in the upper part of the figure. Coefficients are displayed in Figure 5 for mothers and in Figure 6 with confidence intervals at the levels of 90, 95 and 99%. The effect of

²³In all cases, the instrument is relevant. The F -stat equals 21.820 (resp. 19.471) for the least educated mothers (resp. fathers) and 12.670 (resp. 15.730) for the most educated mothers (resp. fathers).

Figure 5: Heterogeneous effect of family size on maternal time allocation



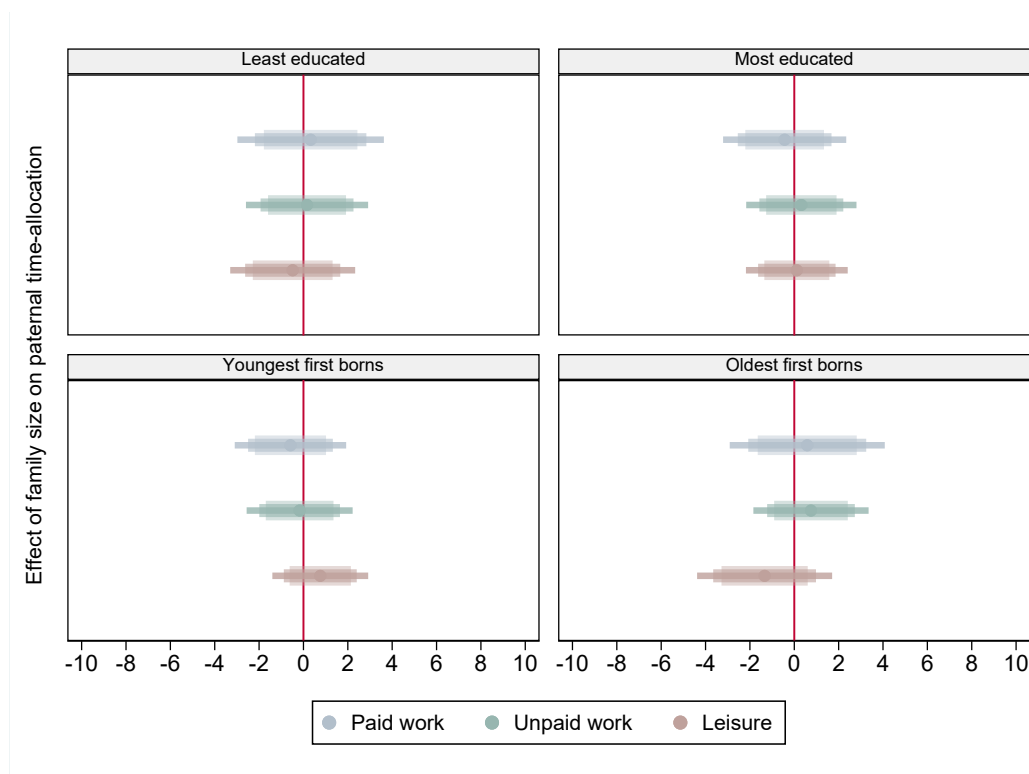
Note: Coefficient plot with confidence intervals at the levels of 90,95, and 99% are represented from transparent to opaque. The least (resp. most) educated group refers to mothers who completed high school at most (resp. who completed at least the first level of tertiary education). The youngest (resp. oldest) first-borns are children aged under (resp. at least) the median age (i.e. 14 years old).

Source : German Socio-Economic Panel (1991-2018).

family size on mothers' re-allocation of time from paid work to unpaid work is entirely driven by the least educated mothers of the sample. For the most educated mothers, coefficients are not significantly different from zero at any conventional level. Fathers do not reallocate time in response to a higher number of children, regardless of their education level.

Age of the first child. Then, heterogeneous results are explored considering the age of the first child. The underlying hypothesis is that mothers who have additional children when their first child is younger might still be out of the labour market when a new child arrives (e.g. on parental leave), leading to lower effects on time re-allocation. On the other hand, older first-borns might be more likely to contribute to domestic tasks, which could result in lower effects too. Parents are split into two groups according to their first child's age at the time of interview. The group of the youngest first-borns are children aged under the median age which, here, is

Figure 6: Heterogeneous effect of family size on paternal time allocation



Note: Coefficient plot with confidence intervals at the levels of 90,95, and 99% are represented from transparent to opaque. The least (resp. most) educated group refers to fathers who completed high school at most (resp. who completed at least the first level of tertiary education). The youngest (resp. oldest) first-borns are children aged under (resp. at least) the median age (i.e. 14 years old). Source : German Socio-Economic Panel (1991-2018).

14 years old while the oldest first-borns are aged at least 14.²⁴ Figure 5 show that results are entirely driven by mothers with an older first-born child when family size increases, in line with the first hypothesis.

5.2.4 Additional results

Is there an effect on well-being ? Considering that taking care of domestic tasks might induce a higher mental load, I explore the potential implications of having more children on well-being using indicators of satisfaction with work, housework, childcare and leisure as outcomes.²⁵ Table 4 shows that mothers' probability to declare a high satisfaction level (i.e. above the median score) with respect to housework is unaffected by family size. This result is in line with previous results presented in Section 5.1 (i.e. a lower decline in satisfaction with housework for

²⁴In all cases, the instrument is relevant. The F -stat equals 28.048 for the youngest first-borns and 17.328 for the oldest ones.

²⁵Other dimensions of satisfaction are available in the survey but with higher proportions of missing values (some questions were introduced in later survey years) or not referring to any time-use dimension.

women compared to men after the first child birth), which suggests that individuals stick to gender identity norms [Akerlof and Kranton, 2000]. Results also show that the probability to declare a level of satisfaction with leisure time above the median score decreases by nearly 32 percentage points ($p < 0.05$) when family size increases. Interestingly, mothers declare lower satisfaction with leisure while the amount of time they spend on leisure activities is not reduced. This could be explained by a negative effect of family size on the quality of leisure time. This could stem from the fact that mothers of larger families might be less likely to have leisure time on their own, which could generate stress when the number of children is high. As far as fathers are concerned, levels of satisfaction remain significantly unchanged.

Table 4: Effect of family size on well-being

	Paid work	Housework	Childcare	Leisure
Mothers				
OLS estimate	0.037*** (0.009)	-0.013 (0.010)	0.030*** (0.011)	-0.041*** (0.009)
FE estimate	-0.006 (0.026)	-0.043** (0.020)	0.004 (0.021)	-0.085*** (0.018)
TSLS estimate	-0.014 (0.130)	-0.185 (0.157)	0.119 (0.181)	-0.323** (0.134)
<i>Endogeneity test (p-value)</i>	0.692	0.260	0.617	0.022
Fathers				
OLS estimate	0.028*** (0.009)	-0.001 (0.010)	0.031*** (0.010)	-0.016* (0.009)
FE estimate	-0.001 (0.025)	-0.007 (0.021)	0.002 (0.021)	-0.102*** (0.017)
TSLS estimate	0.096 (0.126)	-0.009 (0.159)	0.031 (0.173)	-0.016 (0.125)
<i>Endogeneity test (p-value)</i>	0.587	0.957	0.994	0.997
<i>Weak instrument test (F-stat)</i>	25.230	21.218	16.775	31.330
Number of individuals	7,381	7,805	6,393	9,250
Number of observations	31,308	29,567	21,949	43,262

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source : German Socio-Economic Panel (1991-2018).

Disaggregated time-use outcomes. Time-use outcomes are decomposed in order to delve into different time-allocation effects. In the main analysis, leisure represents the daily remaining time after deducing hours allocated to paid and unpaid work (see Section 4). Table F1 reports the results of the instrumental variable strategy splitting leisure time into three sub-items : sleep, hobbies and other free-time. Information on hobbies is available at each wave but sleeping time is only available from 2008 to 2013 and in 2015 and 2017. Therefore, the sample is restricted to couples for which sleeping time is available for both parents. All remaining hours of leisure that are not categorised as sleeping time or time spent on hobbies by individuals are referred as "other free-time". Results show that none of these sub-items of leisure time is significantly affected.²⁶

Then, unpaid work is divided into four different tasks : housework, errands, repairs or gardening and childcare (see Table F2). The analysis reveals that the increase in unpaid work provided by mothers is driven by housework tasks (+1.250 daily hours, $p < 0.01$), childcare (+1.458, $p < 0.1$) and errands (+0.421, $p < 0.05$). The effect on childcare is the largest but also the least precise. This could stem from the fact that domestic work and childcare are tasks that are likely to be conducted simultaneously (e.g. doing housework while children are home). In this case, some mothers might tend to report hours spent on both housework and childcare under the "housework" category or under the "childcare" category only since the questionnaire does not allow to declare simultaneous activities. Another potential explanation could reside in the fact that the effect of an additional child on childcare time is expected to be marginally decreasing.

Intra-household time-allocation. In order to know if these shifts in individual time-allocation change how mothers contribute to the total amount of time spent on a given task at the couple level, I conduct the analysis taking the female-to-couple ratio for each time-use item (i.e. female time related to the sum of female and male time used). This analysis adds additional information because changes at the individual level could exist without modifying the intra-household division of tasks. This could happen if both parents' time spent on a given task shifts in the same way with a similar amplitude or if the division of tasks between parents before the arrival of the additional child is very unequal (e.g. mothers change their time allocated to a task

²⁶All coefficients are not significant for mothers. For fathers, the coefficient when taking time spent on hobbies as the outcome is positive but the magnitude is very low (+0.907) and is only significantly different from 0 at the level of 10%.

while fathers' time is close to zero irrespective of family size). Table F3 reports results taking the female-to-couple ratio as an outcome for each type of task. Results indicate that family size decreases mothers' share in the total time the couple allocates to paid work, which is in line with the fact that mothers significantly reduce their paid work time as opposed to fathers who keep it very high irrespective of the number of children. Interestingly, the change in the intra-household time-allocation ratio for domestic tasks is no significantly different from 0 while mothers do increase their time spent on such tasks. Since fathers do not significantly increase their time spent on domestic tasks, the first hypothesis (i.e. both parents' time spent on a given task shifts in the same way with a similar amplitude) is not realistic. However, the second one is very likely since fathers' time spent on unpaid tasks is very low. As shown in Table B2, fathers' time spent on domestic tasks is not null (2.17 daily hours on average) because all types of domestic tasks are added up. Delving into time allocated to each item of domestic task, it appears that fathers spend less than 1 hour per day doing housework (0.67 daily hours), running errands (0.67 daily hours) and doing repairs (0.83 daily hours).²⁷ As a result, even if mothers increase time spent on housework by 1.250 hours and time spent running errands by 0.421 hours (Table F2), they do not contribute more to the total household time which emphasizes that they bear the major part irrespective of family size. For childcare and leisure time, the intra-household ratio in time-allocation remains unchanged because each parent keeps virtually the same time allocated to these tasks irrespective to the number of children they have.

6 Conclusion

This paper aims at investigating how parents reallocate time with respect to parenthood. I first focus on the evolution of time-allocation of both parents around the first birth using an event-study strategy. For paid work, unpaid work as well as leisure time, results show gender gaps in the first birth penalty in favor of men. Gender gaps are lower for higher distances to birth but never converge to zero. In addition, long-run penalties are higher among couples with several children compared to one-child parents. Surprisingly, the decreased satisfaction with housework after the birth is of greater magnitude among men while the load of unpaid work they take on after the birth is much lower compared to mothers' relative to one year before birth. This result

²⁷Statistics for each item of domestic tasks are not reported in tables.

suggest that gender roles are deeply integrated by individuals.

Then, the paper investigates whether the number of children is also at stake in time-allocation gender gaps using an instrumental variable strategy. Knowing that a parental preference for a sex-mixed sibship has been documented in the previous literature and that sex at birth is randomly assigned, the sex-composition of the first two children is used for instrumentation. Results show that mothers reallocate time spent on the labour market to unpaid tasks such as doing housework, running errands and providing childcare when family size increases. This effect is driven by the least educated mothers. This result can be explained by the fact that mothers with a lower education level are expected to have lower opportunity costs of time due to lower wages compared to other mothers. The re-allocation effect is also driven by mothers whose first-born child is older when family size increases, suggesting that an additional child implies a new career interruption when the mother has time to go back to the labour market after a previous birth compared to mothers of younger first-borns. This paper also explores the effect of family size on parents' satisfaction with housework and leisure and shows that a larger sibship has a negative impact on mothers' satisfaction with leisure. This result is interpreted as a negative effect of family size on leisure time quality as mothers declare less satisfaction with leisure when the number of children increases while they do not reduce the time they spend on leisure.

This paper suffers from several limitations. First, event-study estimates might fail to identify the causal effect of entry into parenthood considering that birth timing is not exogenous. It is nonetheless useful to have in mind that this is a concern especially for coefficients estimated far from birth. In any case, event-study estimates are informative on the description of gender gaps in time-allocation and satisfaction with time-use outcomes around the first birth. Second, the external validity of the instrumental variable strategy is limited. The first reason is that, due to the nature of the instrument (i.e. the sex-composition among the first two children), the analysis is confined to couples who already have two children. Nevertheless, the event-study analysis fills this gap by providing evidence on parents with only one child. The second reason is that the effect is specific to compliers (i.e. parents who have a preference for a mixed-sex sibship). Using the methodology proposed by [Marbach and Hangartner \[2020\]](#), I compare compliers' characteristics with non-compliers'. Figure G1 shows that compliers do not significantly differ

from non-compliers except from two points : complying mothers are younger than never-takers (i.e. those who have more than two children irrespective of sex-composition) by about one year and younger than always-takers (i.e. those who have more than two children irrespective of sex-composition) by about two years, and the compliers' second child is older than never-takers' and younger than always-takers by one year. With regard to these characteristics, the LATE cannot be generalised to the Average Treatment Effect (ATE). Then, further data exploration is limited due to insufficient sample sizes. As a result, some investigations cannot be conducted, such as the evolution of the effect with time or its variation with respect to characteristics at the state level, for example.

Bibliography

- Aguero, J. M. and Marks, M. S. (2008). Motherhood and female labor force participation: Evidence from infertility shocks. *American Economic Review*, 98(2):500–504. [14, 42]
- Aguilar-Gomez, S., Arceo-Gomez, E., and De la Cruz Toledo, E. (2019). Inside the black box of child penalties: Unpaid work and household structure. *Available at SSRN 3497089*. [13, 42]
- Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753. [17, 41, 70]
- Alesina, A. and Fuchs-Schündeln, N. (2007). Good-bye lenin (or not?): The effect of communism on people’s preferences. *American Economic Review*, 97(4):1507–1528. [44]
- Anderson, D. J., Binder, M., and Krause, K. (2002). The motherhood wage penalty: Which mothers pay it and why? *American Economic Review*, 92(2):354–358. [40]
- Andresen, M. E. and Nix, E. (2022). What causes the child penalty? Evidence from adopting and same-sex couples. *Journal of Labor Economics*, 40(4):971–1004. [13, 42]
- Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579. [13, 42]
- Angrist, J. D. and Evans, W. N. (1998). Children and their parents’ labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3):450–477. [14, 41, 42, 43, 51, 53, 57, 66, 67]
- Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455. [20, 51]
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press. [50, 57, 180]
- Artmann, E., Oosterbeek, H., and van der Klaauw, B. (2022). Household specialization and the child penalty in the Netherlands. *Labour Economics*, 78:102221. [13, 42]
- Bauernschuster, S. and Rainer, H. (2012). Political regimes and the family: How sex-role attitudes continue to differ in reunified Germany. *Journal of Population Economics*, 25(1):5–

27. [15, 44]

- Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers' labor supply : Evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16. [8, 44, 103]
- Baxter, J., Hewitt, B., and Haynes, M. (2008). Life course transitions and housework: Marriage, parenthood and time on housework. *Journal of Marriage and Family*, 70(2):259–272. [43]
- Ben-Porath, Y. and Welch, F. (1976). Do sex preferences really matter? *The Quarterly Journal of Economics*, 90(2):285–307. [42, 51]
- Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614. [16, 43]
- Bittman, M., England, P., Sayer, L., Folbre, N., and Matheson, G. (2003). When does gender trump money? Bargaining and time in household work. *American Journal of Sociology*, 109(1):186–214. [16, 43]
- Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2):669–700. [14, 42]
- Blau, F. D. and Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4):75–100. [6, 40]
- Bronars, S. G. and Grogger, J. (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *American Economic Review*, pages 1141–1156. [14, 42]
- Budig, M. J. and England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, pages 204–225. [40]
- Butcher, K. F. and Case, A. (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics*, 109(3):531–563. [52]
- Bütikofer, A., Jensen, S., and Salvanes, K. G. (2018). The role of parenthood on the gender gap among top earners. *European Economic Review*, 109:103–123. [13, 42]
- Champagne, C., Pailhé, A., and Solaz, A. (2015). Le temps domestique et parental des hommes et des femmes: Quels facteurs d'évolutions en 25 ans? *Économie et Statistique*, 478(1):209–242. [9, 43]

- Cools, S., Markussen, S., and Strøm, M. (2017). Children and careers: How family size affects parents' labor market outcomes in the long run. *Demography*, 54(5):1773–1793. [14, 43, 52]
- Cools, S. and Strøm, M. (2016). Parenthood wage penalties in a double income society. *Review of Economics of the Household*, 14(2):391–416. [40]
- Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. Technical report, National Bureau of Economic Research. [13, 42]
- Craig, L. and Mullan, K. (2010). Parenthood, gender and work-family time in the United States, Australia, Italy, France and denmark. *Journal of Marriage and Family*, 72(5):1344–1361. [43]
- Cruces, G. and Galiani, S. (2007). Fertility and female labor supply in latin america: New causal evidence. *Labour Economics*, 14(3):565–573. [14, 43]
- De Chaisemartin, C. (2017). Tolerating defiance? Local average treatment effects without monotonicity. *Quantitative Economics*, 8(2):367–396. [53]
- De Quinto, A., Hospido, L., and Sanz, C. (2021). The child penalty: Evidence from Spain. *SERIEs*, 12:585–606. [13, 42]
- Dustmann, C. and Schönberg, U. (2012). Expansions in maternity leave coverage and children's long-term outcomes. *American Economic Journal: Applied Economics*, 4(3):190–224. [45]
- D'Albis, H., Bonnet, C., Navaux, J., Pelletan, J., and Solaz, A. (2016). Travail rémunéré et travail domestique: Une évaluation monétaire de la contribution des femmes et des hommes à l'activité économique depuis 30 ans. *Revue de l'OFCE*, (5):101–130. [9, 43]
- Federal Ministry for Family Affairs, Senior Citizens, W. and Youth (2020). Who takes care of children, household and the elderly? A dossier on the societal dimension of a private question. [42]
- Flèche, S., Lepinteur, A., and Powdthavee, N. (2020). Gender norms, fairness and relative working hours within households. *Labour Economics*, 65:101866. [17, 43]
- Frenette, M. (2011). How does the stork delegate work? Childbearing and the gender division of paid and unpaid labour. *Journal of Population Economics*, 24(3):895–910. [14, 43]
- Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach.

- Labour Economics*, 36:84–98. [45, 226]
- Gimenez-Nadal, J. I. and Sevilla, A. (2012). Trends in time allocation: A cross-country analysis. *European Economic Review*, 56(6):1338–1359. [9, 43]
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119. [40]
- Jacobsen, J. P., Pearce III, J. W., and Rosenbloom, J. L. (1999). The effects of childbearing on married women’s labor supply and earnings: Using twin births as a natural experiment. *Journal of Human Resources*, pages 449–474. [14, 42]
- Kan, M. Y. (2008). Measuring housework participation: The gap between “stylised” questionnaire estimates and diary-based estimates. *Social Indicators Research*, 86:381–400. [63]
- Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, volume 109, pages 122–126. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. [13, 14, 42, 46]
- Kleven, H., Landais, C., and Søgaaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209. [13, 42, 46]
- Kluge, J. and Tamm, M. (2013). Parental leave regulations, mothers’ labor force attachment and fathers’ childcare involvement: Evidence from a natural experiment. *Journal of Population Economics*, 26(3):983–1005. [45]
- Liepmann, H. (2018). The impact of a negative labor demand shock on fertility : Evidence from the fall of the berlin wall. *Labour Economics*, 54:210–224. [44]
- Lippmann, Q., Georgieff, A., and Senik, C. (2020). Undoing gender with institutions: Lessons from the German division and reunification. *The Economic Journal*, 130(629):1445–1470. [16, 43, 44]
- Loughran, D. S. and Zissimopoulos, J. M. (2009). Why wait? The effect of marriage and childbearing on the wages of men and women. *Journal of Human Resources*, 44(2):326–349. [40]
- Lundberg, S. and Rose, E. (2000). Parenthood and the earnings of married men and women.

- Labour Economics*, 7(6):689–710. [40]
- Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review*, 107(6):1611–37. [14, 42]
- Mammen, K. (2011). Fathers’ time investments in children: Do sons get more? *Journal of Population Economics*, 24(3):839–871. [52]
- Marbach, M. and Hangartner, D. (2020). Profiling compliers and noncompliers for instrumental-variable analysis. *Political Analysis*, 28(3):435–444. [73, 93, 132, 166]
- Markussen, S. and Strøm, M. (2022). Children and labor market outcomes: Separating the effects of the first three children. *Journal of Population Economics*, 35(1):135–167. [14, 42]
- Menta, G. and Lepinteur, A. (2021). Boys don’t cry (or do the dishes): Family size and the housework gender gap. *Journal of Economic Behavior & Organization*, 186:164–188. [52]
- Meurs, D. and Pora, P. (2019). Gender equality on the labour market in France: A slow convergence hampered by motherhood. *Economie et Statistique*, 510(1):109–130. [13, 42]
- Müller, K.-U. and Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare. *Labour Economics*, 62:101776. [7, 44, 102]
- Olivetti, C. and Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8:405–434. [6, 40]
- Rabaté, S. and Rellstab, S. (2022). What determines the child penalty in the Netherlands? The role of policy and norms. *De Economist*, 170(2):195–229. [13, 42]
- Rainer, H. and Siedler, T. (2009). Does democracy foster trust? *Journal of Comparative Economics*, 37(2):251–269. [44]
- Rosenzweig, M. R. and Wolpin, K. I. (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: Journal of the Econometric Society*, pages 227–240. [14, 42]
- Schober, P. S. (2014). Parental leave and domestic work of mothers and fathers: A longitudinal study of two reforms in West Germany. *Journal of Social Policy*, 43(2):351–372. [9, 45, 225]
- Schönberg, U. and Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers’

- labor market outcomes after childbirth. *Journal of Labor Economics*, 32(3):469–505. [45]
- Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529. [51, 111]
- Tamm, M. (2019). Fathers’ parental leave-taking, childcare involvement and labor market participation. *Labour Economics*, 59:184–197. [9, 45, 225]
- Waldfogel, J. (1997). The effect of children on women’s wages. *American Sociological Review*, pages 209–217. [40]
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage Learning. [52]

Appendix

A Identification tests

This section includes elements that support the identification strategy presented in Section 3.2.3.

Table A1: Mean comparison test on observable characteristics between treated and control individuals

	Mixed-sex sibship (1)	Same-sex sibship (2)	Difference (1) - (2)
<i>Mothers</i>			
Age	38.888 (6.323)	38.890 (6.397)	-0.002 (0.144)
Highly educated (%)	0.378	0.387	-0.009 (0.013)
<i>Fathers</i>			
Age	41.677 (6.799)	41.847 (6.950)	-0.170 (0.159)
Highly educated (%)	0.422	0.439	-0.018* (0.013)
<i>Household</i>			
Age of first child	13.283 (5.787)	13.225 (5.723)	0.058 (0.126)
Age of second child	9.351 (5.270)	9.350 (5.206)	0.001 (0.109)
Number of individuals	4,802	4,935	9,737
Number of observations	24,475	25,674	50,149

Note: Columns 1 and 2 report means with standard deviations in parentheses for parents of a mixed-sex sibship and parents of a same-sex sibship respectively. Column 3 reports results of the mean comparison t-test for each variable (i.e. the difference in means with standard errors corrected for clustering at the individual level in parentheses). *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

Table A2: Means comparison test on time-allocation between treated and control individuals before the third birth

	Mixed-sex sibship (1)	Same-sex sibship (2)	Difference (1) - (2)
<i>Mothers</i>			
Paid work	4.356 (3.650)	4.254 (3.662)	0.102 (0.095)
Unpaid work			
– Domestic work	4.675 (2.216)	4.703 (2.244)	-0.028 (0.054)
– Childcare	4.335 (3.850)	4.375 (3.835)	-0.040 (0.094)
Leisure	10.635 (3.812)	10.669 (3.784)	-0.034 (0.085)
<i>Fathers</i>			
Paid work	8.925 (3.074)	8.918 (3.029)	0.007 (0.072)
Unpaid work			
– Domestic work	2.188 (1.800)	2.155 (1.759)	0.033 (0.041)
– Childcare	1.384 (1.655)	1.388 (1.607)	-0.004 (0.039)
Leisure	11.504 (3.043)	11.538 (2.957)	-0.034 (0.065)
Number of individuals	3,347	3,226	6,573
Number of observations	16,953	16,081	33,034

Note: Sample restricted on observations before the birth of any third child. Columns 1 and 2 report means with standard deviations in parentheses for parents of a mixed-sex sibship and parents of a same-sex sibship respectively. Column 3 reports results of the mean comparison t-test for each covariate (i.e. the difference in means with standard errors corrected for clustering at the individual level in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

Table A3: Mean comparison test between parents of two girls and parents of two boys as first-borns

	Two daughters as first-borns (1)	Two sons as first-borns (2)	Difference (1) - (2)
<i>Mothers</i>			
Age	38.887 (6.435)	38.893 (6.360)	-0.006 (0.203)
Highly educated (%)	0.386	0.388	-0.002 (0.018)
<i>Fathers</i>			
Age	41.806 (6.897)	41.886 (6.998)	-0.080 (0.226)
Highly educated (%)	0.437	0.442	-0.005 (0.018)
<i>Household</i>			
Age of first child	13.109 (5.698)	13.334 (5.745)	-0.225* (0.175)
Age of second child	9.292 (5.216)	9.405 (5.197)	-0.113 (0.152)
Number of individuals	2,373	2,562	9,737
Number of observations	12,391	13,283	50,149

Note: Sample restricted on parents of same-sex first borns. Columns 1 and 2 report means with standard deviations in parentheses for parents of two daughters and parents of two sons respectively. Column 3 reports results of the mean comparison t-test for each covariate (i.e. the difference in means with standard errors corrected for clustering at the individual level in parentheses. *p<0.10, **p<0.05, ***p<0.01. Source : German Socio-Economic Panel (1991-2018).

B Descriptive statistics

This section presents descriptive statistics for the populations of interest presented in Section 4.

Table B1: Pre-birth characteristics

	Women	Men
<i>Time-allocation (one year before birth)</i>		
Daily hours allocated to paid work	7.918 (3.429)	8.802 (3.183)
Daily hours allocated to unpaid work	3.156 (1.983)	2.326 (1.792)
Daily hours allocated to leisure	12.926 (3.300)	12.872 (3.302)
<i>Well-being (one year before birth)</i>		
Score of satisfaction with paid work	7.113 (2.027)	7.213 (1.993)
Score of satisfaction with housework	7.057 (1.773)	7.270 (1.861)
Score of satisfaction with leisure	6.867 (2.067)	6.695 (2.122)
<i>Individual characteristics</i>		
Age at first birth	29.527 (4.489)	32.212 (5.128)
Number of individuals	1,263	1,263
Number of observations	10,806	10,806

Note: Sample restricted on individuals observed in a bandwidth of 15 years around the first birth. The first (resp. second) column reports means and standard deviations of female (resp. male) characteristics one year before the birth.

Source : German Socio-Economic Panel (1991-2018).

Table B2: Descriptive statistics among couples with at least two children

	Mothers	Fathers	Household
<i>Time-allocation</i>			
Daily hours allocated to paid work	3.908 (3.627)	8.805 (3.217)	
Daily hours allocated to unpaid work			
– Domestic work	4.880 (2.276)	2.176 (1.814)	
– Childcare	4.696 (3.873)	1.471 (1.705)	
Daily hours allocated to leisure	10.516 (3.847)	11.548 (3.057)	
<i>Well-being</i>			
Score of satisfaction with paid work	7.197 (1.993)	7.123 (1.976)	
Score of satisfaction with domestic work	6.588 (1.857)	6.870 (1.957)	
Score of satisfaction with childcare	7.361 (2.253)	7.233 (2.153)	
Score of satisfaction with leisure	6.339 (2.244)	6.307 (2.202)	
<i>Individual characteristics</i>			
Age	38.889 (6.361)	41.764 (6.877)	
Highly educated (%)	0.383	0.431	
Employment rate (%)	0.641	0.914	
<i>Household characteristics</i>			
Age of first child			13.254 (5.755)
Age of second child			9.351 (5.238)
More than two children (%)			0.358
Same-sex first-borns (%)			0.512
First child is a boy (%)			0.508
Second child is a boy (%)			0.509
Number of individuals	9,737	9,737	9,737
Number of observations	50,149	50,149	50,149

Note: Sample restricted on parents of at least two children. Means with standard deviations in parentheses are reported for characteristics at the mother level (resp. father level, household level) in the first (resp. second, third) column.

Source : German Socio-Economic Panel (1991-2018).

C First birth penalties

This section reports results of the empirical strategy presented in Section [3.1](#).

Table C1: Relative gender gap in first birth penalty : time-allocation

	Daily hours on paid work	Daily hours on unpaid work	Daily hours on leisure
<i>1 child</i>	Penalty	Penalty	Penalty
	95 % CI	95 % CI	95 % CI
0	-0.335 [-0.338 ; -0.331]	0.734 [0.719 ; 0.749]	0.0613 [0.0611 ; 0.0616]
1	-0.834 [-0.844 ; -0.825]	2.910 [2.846 ; 2.974]	-0.1510 [-0.1516 ; -0.1504]
2	-0.542 [-0.547 ; -0.538]	2.208 [2.165 ; 2.251]	-0.1784 [-0.1791 ; -0.1777]
3	-0.436 [-0.440 ; -0.432]	1.850 [1.815 ; 1.885]	-0.1488 [-0.1493 ; -0.1482]
4	-0.361 [-0.364 ; -0.358]	1.504 [1.477 ; 1.531]	-0.1080 [-0.1084 ; -0.1075]
5	-0.336 [-0.339 ; -0.332]	1.417 [1.391 ; 1.443]	-0.1036 [-0.1040 ; -0.1032]
6	-0.332 [-0.336 ; -0.328]	1.347 [1.323 ; 1.370]	-0.1170 [-0.1175 ; -0.1166]
7	-0.311 [-0.315 ; -0.307]	1.256 [1.233 ; 1.280]	-0.1013 [-0.1017 ; -0.1008]
8	-0.312 [-0.316 ; -0.308]	1.295 [1.270 ; 1.320]	-0.1144 [-0.1151 ; 0.1138]
9	-0.293 [-0.298 ; -0.288]	1.128 [1.104 ; 1.152]	-0.0880 [-0.0886 ; -0.0874]
10	-0.299 [-0.314 ; -0.285]	1.124 [1.097 ; 1.152]	-0.0935 [-0.0943 ; -0.0928]
<i>>1 child</i>	Penalty	Penalty	Penalty
	95 % CI	95 % CI	95 % CI
0	-0.376 [-0.379 ; -0.374]	0.840 [0.827 ; 0.854]	0.0607 [0.0606 ; 0.0609]
1	-0.876 [-0.881 ; -0.870]	3.082 [3.031 ; 3.134]	-0.1588 [-0.1591 ; -0.1585]
2	-0.707 [-0.712 ; -0.703]	2.799 [2.753 ; 2.844]	-0.1548 [-0.1551 ; -0.1545]
3	-0.714 [-0.718 ; -0.710]	3.019 [2.964 ; 3.073]	-0.1665 [-0.1667 ; -0.1662]
4	-0.749 [-0.753 ; -0.745]	3.257 [3.184 ; 3.329]	-0.1537 [-0.1539 ; -0.1535]
5	-0.663 [-0.666 ; -0.660]	3.333 [3.261 ; 3.405]	-0.1614 [-0.1616 ; -0.1612]
6	-0.590 [-0.593 ; -0.586]	3.453 [3.364 ; 3.541]	-0.1535 [-0.1537 ; -0.1533]
7	-0.603 [-0.606 ; -0.599]	3.674 [3.556 ; 3.793]	-0.1433 [-0.1435 ; -0.1432]
8	-0.622 [-0.627 ; -0.617]	3.948 [3.812 ; 4.083]	-0.1484 [-0.1486 ; -0.1482]
9	-0.572 [-0.578 ; -0.567]	4.210 [3.940 ; 4.479]	-0.1320 [-0.1322 ; -0.1318]
10	-0.555 [-0.560 ; -0.552]	4.652 [4.319 ; 4.985]	-0.1178 [-0.1180 ; -0.1176]

Note: Sample restricted on parents observed in a bandwidth of 15 years around the first birth.
Source : German Socio-Economic Panel (1991-2018).

Table C2: Relative gender gap in first birth penalty : well-being

	Distance to birth	Daily hours on paid work	Daily hours on unpaid work	Daily hours on leisure
	Penalty	95 % CI	Penalty	95 % CI
<i>1 child</i>				
0	-0.0097	[-0.0097 ; -0.0096]	0.0317	[0.0315 ; 0.0318]
1	-0.0114	[-0.0114 ; -0.0114]	0.0174	[0.0173 ; 0.0175]
2	-0.0164	[-0.0165 ; -0.0164]	0.0063	[0.0062 ; 0.0063]
3	0.0094	[0.0093 ; 0.0094]	0.0355	[0.0354 ; 0.0357]
4	0.0210	[0.0210 ; 0.0211]	0.0450	[0.0448 ; 0.0452]
5	-0.0068	[-0.0069 ; -0.0068]	0.0492	[0.0489 ; 0.0494]
6	0.0137	[0.0136 ; 0.0137]	0.0477	[0.0475 ; 0.0479]
7	0.0089	[0.0038 ; 0.0141]	0.0646	[0.0643 ; 0.0651]
8	-0.0187	[-0.0188 ; -0.0187]	0.0186	[0.0185 ; 0.0187]
9	0.0236	[0.0236 ; 0.0237]	0.0587	[0.0581 ; 0.0592]
10	-0.0233	[-0.0234 ; -0.0232]	0.0548	[0.0542 ; 0.0553]
<i>>1 child</i>				
0	0.0117	[0.0116 ; 0.0117]	0.0391	[0.0389 ; 0.0393]
1	0.0173	[0.0172 ; 0.0173]	0.0193	[0.0192 ; 0.0194]
2	-0.0060	[-0.0060 ; -0.0059]	0.0203	[0.0202 ; 0.0204]
3	0.0057	[0.0056 ; 0.0057]	0.0451	[0.0449 ; 0.0452]
4	0.0157	[0.0156 ; 0.0157]	0.0739	[0.0736 ; 0.0742]
5	0.0127	[0.0126 ; 0.0127]	0.0484	[0.0482 ; 0.0486]
6	0.0098	[0.0098 ; 0.0099]	0.0815	[0.0811 ; 0.0817]
7	0.0070	[0.0069 ; 0.0070]	0.0575	[0.0573 ; 0.0577]
8	0.0074	[0.0074 ; 0.0074]	0.1028	[0.1022 ; 0.1033]
9	0.0091	[0.0091 ; 0.0092]	0.1106	[0.1101 ; 0.1112]
10	0.0035	[0.0035 ; 0.0035]	0.0877	[0.0872 ; 0.0882]
	Penalty	95 % CI	Penalty	95 % CI
	0.0359	[0.0358 ; 0.0360]	0.0555	[0.0554 ; 0.0557]
	-0.0151	[-0.0151 ; -0.0150]	0.0065	[0.0065 ; 0.0066]
	-0.0107	[-0.0107 ; -0.0106]	0.0209	[0.0209 ; 0.0210]
	-0.0019	[-0.0019 ; -0.0018]	0.0032	[0.0031 ; 0.0032]
	-0.0147	[-0.0147 ; -0.0147]	0.0175	[0.0175 ; 0.0176]
	-0.0069	[-0.0070 ; -0.0069]	-0.0143	[-0.0144 ; -0.0143]
	-0.0271	[-0.0272 ; -0.0269]	-0.0091	[-0.0092 ; -0.0091]
	-0.0251	[-0.0252 ; -0.0250]	-0.0293	[-0.0294 ; -0.0292]
	-0.0306	[-0.0308 ; -0.0305]	0.0175	[0.0174 ; 0.0175]
	-0.0349	[-0.0352 ; -0.0348]	0.0182	[0.01814 ; 0.0183]
	-0.0086	[-0.0086 ; -0.0085]	-0.0062	[-0.0063 ; -0.0062]

Note: Sample restricted on parents observed in a bandwidth of 15 years around the first birth.
Source : German Socio-Economic Panel (1991-2018).

E Internal validity

This section presents elements that support the internal validity of the results (Section 5.2.2).

E.1 Robustness checks

Table E1: Robustness checks

	Mothers			Fathers		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Time-allocation</i>						
Daily hours allocated to paid work	-2.632** (1.066)	-2.760*** (1.013)	-2.795*** (1.050)	0.167 (0.868)	-0.016 (0.876)	-0.027 (0.867)
Daily hours allocated to unpaid work	2.872** (1.128)	2.707** (1.144)	2.976*** (1.122)	0.098 (0.741)	0.094 (0.742)	0.211 (0.742)
Daily hours allocated to leisure	-0.241 (0.909)	0.053 (0.978)	-0.181 (0.924)	-0.266 (0.728)	-0.078 (0.741)	-0.184 (0.736)
<i>Weak instrument test (F-stat)</i>	34.826	32.215	33.790	34.826	32.215	33.790
Education level controls are excluded	Yes	No	No	Yes	No	No
East Germany is excluded	No	Yes	No	No	Yes	No
Sex of each first born child is included	No	No	Yes	No	No	Yes
Number of individuals	9,737	9,737	7,915	9,737	9,737	7,915
Number of observations	50,149	50,149	39,495	50,149	50,149	39,495

Note: IV-TSLS estimates excluding education levels as covariates (Columns 1 and 4), excluding East Germany (Columns 2 and 5) and including the sex of each first-born child (Columns 3 and 6). Standard-errors corrected for clustering at the individual level are reported in parentheses. Age controls and survey year specific effects are included in all regressions as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

Table E2: Robustness check : Parental leave

	On parental leave
<i>Mothers</i>	
TSLS estimate	-0.051 (0.065)
<i>Fathers</i>	
TSLS estimate	-0.026 (0.017)
<i>Weak instrument test (F-stat)</i>	27.665
Number of individuals	8,253
Number of observations	36,906

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

E.2 Exclusion restriction test

Table E3: Testing the over-identification assumption : Time-allocation

		<u>More than two children</u>		
Two sons				
OLS estimate		0.074***		
		(0.014)		
Two daughters				
OLS estimate		0.057***		
		(0.014)		
<i>Equality of coefficients test (p-value)</i>		0.299		
<i>Weak instrument test (F-stat)</i>		17.464		
			Paid work	Unpaid work
				Leisure
More than two children				
Mothers				
TOLS estimate		-2.922***	3.351***	-0.428
		(1.030)	(1.106)	(0.909)
<i>Over-identification test (p-value)</i>		0.556	0.107	0.517
Fathers				
TOLS estimate		0.026	0.364	-0.389
		(0.850)	(0.727)	(0.725)
<i>Over-identification test (p-value)</i>		0.770	0.308	0.204
Number of individuals	9,737	9,737	9,737	9,737
Number of observations	50,149	50,149	50,149	50,149

Note: The *same-sex* instrument is split into two different instruments, *two sons* and *two daughters*. Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01.

Source : German Socio-Economic Panel (1991-2018).

F Further results

This section presents additional results that are discussed in Section 5.2.4.

Table F1: Effects of family size on leisure time: sleep or hobbies ?

	Leisure		
	Sleep	Hobbies	Other free time
Mothers			
TSLS estimate	-0.183 (0.372)	0.550 (0.432)	-0.671 (0.912)
Fathers			
TSLS estimate	-0.169 (0.368)	0.907* (0.494)	-0.891 (0.934)
<i>Weak instrument test (F-stat)</i>	22.134	22.134	22.134
Number of individuals	5,318	5,318	5,318
Number of observations	13,611	13,611	13,611

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01. *p<0.10, **p<0.05, ***p<0.01. Source : German Socio-Economic Panel (1991-2018).

Table F2: Effects of family size on unpaid work time

	Unpaid work			
	Housework	Errands	Repairs	Childcare
Mothers				
TSLS estimate	1.250*** (0.461)	0.421** (0.178)	-0.109 (0.193)	1.458* (0.768)
Fathers				
TSLS estimate	-0.079 (0.223)	-0.095 (0.184)	0.171 (0.238)	0.233 (0.399)
<i>Weak instrument test (F-stat)</i>	34.065	34.065	34.065	34.065
Number of individuals	9,737	9,737	9,737	9,737
Number of observations	50,149	50,149	50,149	50,149

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01. *p<0.10, **p<0.05, ***p<0.01. Source : German Socio-Economic Panel (1991-2018).

Table F3: Effect of family size on intra-household time-allocation ratio

	Paid work	Unpaid work		Leisure
		Domestic work	Childcare	
TSLS estimate	-0.196*** (0.075)	0.066 (0.055)	0.014 (0.056)	-0.002 (0.025)
<i>Weak instrument test (F-stat)</i>	31.533	34.394	32.846	34.053
Number of individuals	9,492	9,735	9,056	9,737
Number of observations	48,000	50,109	43,869	50,147

Note: Standard-errors corrected for clustering at the individual level are reported in parentheses. Individual controls and survey year specific effects are included as described in Section 3.2. *p<0.10, **p<0.05, ***p<0.01. Source : German Socio-Economic Panel (1991-2018).

G Complier analysis

G.1 Methodology

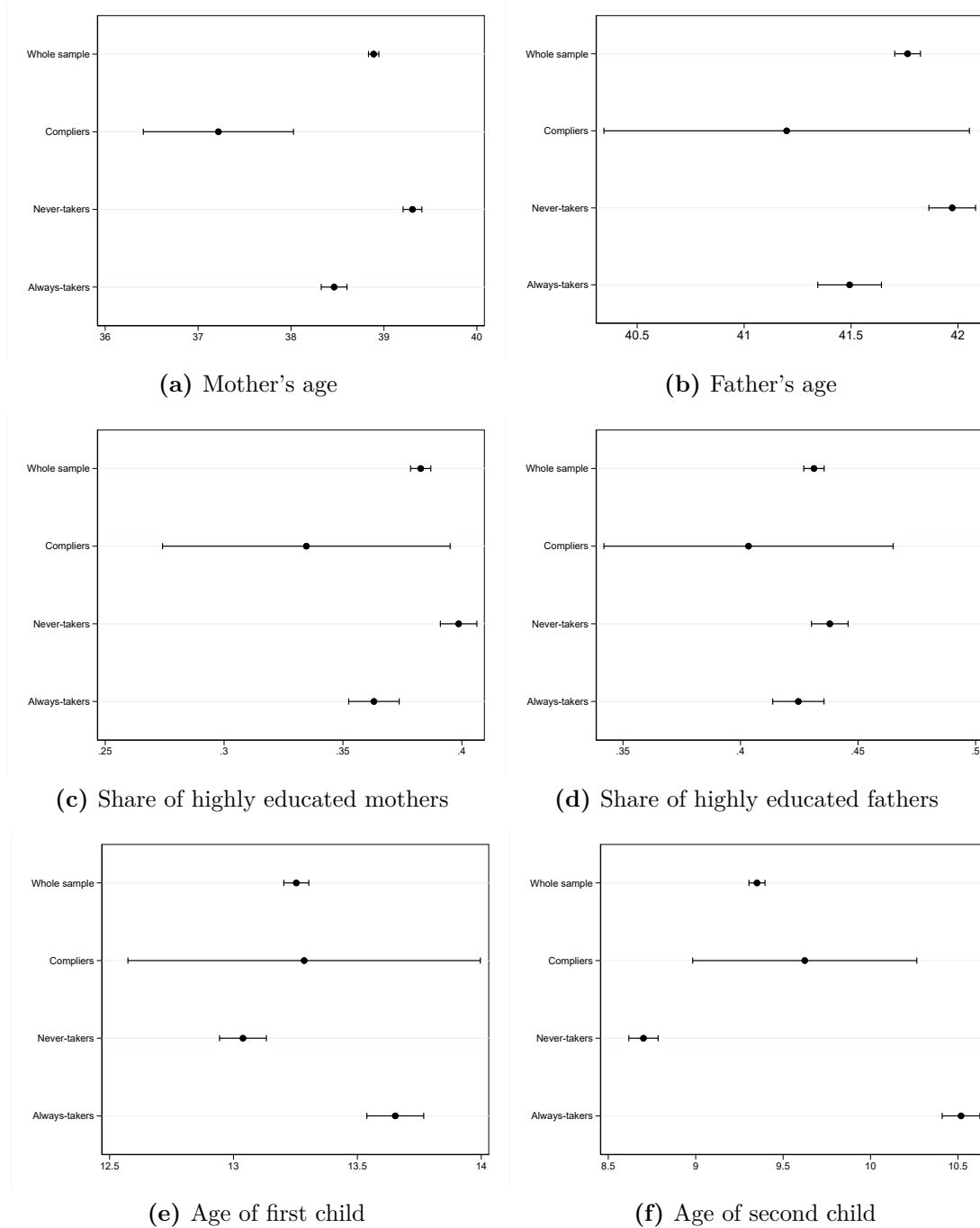
Profiling compliers can be useful to know if the Local Average Treatment Effect (LATE) can be generalized to the Average Treatment Effect (ATE). Under the hypotheses of independence and monotonicity (see Section 3.2.3), the characteristics of compliers and non-compliers can be compared by subtracting the weighted observable characteristic mean of non-compliers (i.e. always-takers and never-takers) to the one computed in the entire sample [Marbach and Hangartner, 2020] such as :

$$\hat{\mu}_{co} = \frac{1}{\hat{\pi}_{co}} * \hat{\mu} - \frac{\hat{\pi}_{nt}}{\hat{\pi}_{co}} * \hat{\mu}_{nt} - \frac{\hat{\pi}_{at}}{\hat{\pi}_{co}} * \hat{\mu}_{at} \quad (9)$$

with $\hat{\mu}_{co}$, $\hat{\mu}$, $\hat{\mu}_{nt}$ and $\hat{\mu}_{at}$ the mean value of a given observable characteristic for compliers, the entire sample, never-takers and always-takers respectively and $\hat{\pi}_{co}$, $\hat{\pi}_{nt}$ and $\hat{\pi}_{at}$ the estimated sample share of compliers, never-takers and always-takers respectively.

G.2 Results

Figure G1: Profile of compliers, never-takers and always-takers



Note : For each observable characteristic, means are reported for the whole sample, compliers, always-takers and never-takers with confidence intervals at the level of 95%.

Source : German Socio-Economic Panel (1991-2018).

Chapter 2

Pre-school Enrollment and Mothers' Labour Supply : Evidence from Europe

co-authored with Julien Bergeot

We would like to thank Eric Bonsang, Eve Caroli, Brigitte Dormont, Sarah Flèche, Mathilde Godard, Arthur Juet, Fanny Landaud, Jürgen Maurer and Fabien Tripier for their valuable comments and suggestions, as well as participants of the Dauphine PhD workshop. This study is based on data from EUROSTAT (EU-SILC, 2004-2020). We are grateful to EUROSTAT for granting us data access and especially to Nicoletta Schweikle-Hilgner for her help with the procedures. *The responsibility for all conclusions drawn from the data lies entirely with the authors.*

Table of contents

1	Introduction	102
2	Institutional framework	104
2.1	Early childcare provision	104
2.2	Pre-school system	105
3	Methodology	108
4	Data	114
4.1	Sample restriction	114
4.2	Description of variables	115
5	Estimation results	118
5.1	Average effects	118
5.2	Heterogeneous effects with respect to individual characteristics	123
5.3	Heterogeneous effects with respect to country level characteristics	125
6	Conclusion	130
	Appendix	137
A	Childcare systems in Europe	137
B	Descriptive statistics by country	139
C	Identification tests	141
D	Robustness checks	143
E	Heterogeneous effects with respect to individual characteristics	145
E.1	Methodology	145
E.2	Results	146
F	Heterogeneous effects with respect to country characteristics	152
F.1	Methodology	152
F.2	Measures at the country level	153
F.3	Results	158
G	Complier analysis	166
G.1	Methodology	166
G.2	Results	167
H	Results by country	168

Abstract

Given that child-related career interruptions are mostly borne by women, childcare provision appears as one of the cornerstones of gender equality on the labour market. In this context, this paper evaluates the effect of pre-school availability on maternal labour supply in a European comparison perspective. We use data from the European Union Statistics on Income and Living Conditions (EU-SILC) and exploit the discontinuous relationship between eligibility to pre-school and the child's year of birth to predict enrollment in a fuzzy regression on discontinuity design. We show that pre-school enrollment does not increase maternal labour supply on average. Nevertheless, further explorations show that the intensive margin is positively affected for the most educated mothers and mothers who do not have younger children. At the country level, pre-school is a leverage for maternal labour market participation in countries with stronger gender stereotypes and in countries with a low investment in early childcare.

1 Introduction

Still today, child-related career interruptions are mostly borne by women. Public childcare provision therefore appears as one of the cornerstones of gender equality on the labour market [OECD, 2012]. In the very first years of a child’s life, usually from 0 to 3 years old¹, parents can benefit from childcare solutions in collective centre-based facilities (i.e. day-care centres) or can opt for home-based childcare provided by professional child-minders. These early childcare solutions are not free of charge, though usually publicly subsidized, and their access is not necessarily guaranteed by legal entitlements [Eurydice, 2019]. When this type of care is not accessible or affordable, parents may have to rely on informal care (e.g. grand-parents) or to reduce their labour supply to care for their child at home. Aimed at older children, usually from 3 years old until elementary schooling age, public pre-school programs have been introduced in order to provide pre-elementary instruction for all children. Pre-school aims at enhancing child development with collective activities improving language, numeracy, artistic and social skills as a preparation for elementary school. As opposed to early childcare solutions (i.e. day-care and child-minders), public pre-school is often provided free of charge.²

The purpose of this paper is to estimate the causal impact of pre-school enrollment on mothers’ labour supply in Europe. The extent to which pre-school enrollment can stimulate maternal labour supply depends on whether mothers manage to resume work during the years before the child enters pre-school. The economic literature has investigated the effect of policies aiming at improving access to affordable early childcare solutions (i.e. childcare provided before pre-school programs, e.g. collective day-care). Studies suggest that providing more early childcare slots enhances mothers’ labour supply [Baker et al., 2008; Lefebvre and Merrigan, 2008; Dujardin et al., 2018; Ravazzini, 2018; Andresen and Havnes, 2019]. Nonetheless, these estimated effects can be rather small or even null [Havnes and Mogstad, 2011; Pora et al., 2020], are heterogeneous in the population [Müller and Wrohlich, 2020] and depend on the initial level of early childcare availability [Brilli et al., 2016]. Giving access to childcare might nevertheless not be enough

¹Eligibility age requirements marginally vary across countries. More details are provided in Section 2.

²In most cases, parents are only charged meals that are highly subsidised.

if slots are not affordable. In that respect, reducing childcare costs can also lever maternal workforce [Givord and Marbot, 2015; Busse and Gathmann, 2020; Huebener et al., 2020].³ Since pre-school is often free and largely supplied, it is expected to increase the labour supply of mothers who have not been able to resume work before because early childcare was not available or affordable enough.

In most countries, eligibility to pre-school enrollment is subject to an age criteria. We use this eligibility rule to predict pre-school enrollment in a fuzzy regression on discontinuity design (RDD). More precisely, we exploit the discontinuous relationship between the child's year of birth and pre-school eligibility to predict enrollment rates using the distance to the eligibility cut-off in quarters of birth. With this design, we estimate the Local Average Treatment Effect (LATE) of pre-school enrollment on mothers' labour supply. We use harmonized survey data collected between 2004 and 2020 in several European Union member states, which is the European Union Statistics on Income and Living Conditions (EU-SILC).

We build on the existing literature exploring the effect of pre-school enrollment on maternal labour supply using discontinuities in eligibility by year of birth to identify a causal effect. Some studies highlight a positive effect of pre-school enrollment on mothers' labour supply in the United States [Gelbach, 2002], in Argentina [Berlinski et al., 2011], in Germany [Bauernschuster and Schlotter, 2015], in Spain [Nollenberger and Rodríguez-Planas, 2015] and in Italy [Carta and Rizzica, 2018]. This literature is nevertheless contrasted by other studies finding positive effects among single mothers only [Cascio, 2009; Goux and Maurin, 2010; Fitzpatrick, 2012].

There is also an economic literature showing the role of gender norms [Antecol, 2000; Fernández and Fogli, 2009; Blau et al., 2013; Bursztyrn et al., 2020] and of public family benefits [Piketty, 2005; González, 2013; Olivetti and Petrongolo, 2017]⁴ on female labour supply. We can expect these contextual factors to shape the relationship between pre-school enrollment of their children and mothers' labour supply. Taking advantage of differences in gender norms and generosity in public benefits between countries, we add to the literature by conducting cross-country comparisons with respect to these characteristics.

³Givord and Marbot [2015] find positive effects of childcare costs reduction only among mothers of large families.

⁴The effect of family benefits largely differ with their nature. Both [Piketty, 2005] and González [2013] find that *in-cash* transfers keep mothers away from the labour market in France and Spain respectively. Among OECD countries, public spending in early childcare (*in-kind* transfers) is positively correlated to female employment [Olivetti and Petrongolo, 2017].

Our results show that, on average, pre-school enrollment does not increase maternal labour supply. Looking at heterogeneous effects with respect to characteristics at the individual level, we find that the intensive margin of labour supply is positively affected for the most educated mothers, who are more likely to adjust their working hours, and decreases it for the least educated ones which suggests that mothers with lower living standards need to work more to pay for formal early childcare when pre-school is not available yet. We also find that pre-school helps mothers who do not have any younger child to go back to the labour market. We explore country-level heterogeneous effects with prevailing gender norms, public spending on family benefits and the net cost of childcare. We find that mothers resume work in response to pre-school enrollment only in countries with the most conservative gender norms, where female labour market participation is lower on average. Regarding the generosity of public benefits, we find different results according to the nature of benefits (in-kind vs in-cash benefits). We do not find any heterogeneous effects when considering generosity in terms of in-cash family benefits. However, mothers living in countries allocating a low share of GDP to in-kind family benefits increase their labour supply at the extensive margin due to pre-school enrollment, while pre-school is not a leverage for maternal labour supply in the most generous countries. Finally, no heterogeneous effect is put forward with regard to the cost of childcare.

The remainder of this paper is organised as follows. We first give details on early childcare provision and pre-school systems in Europe in Section 2. We present the methodology in Section 3 and the data in Section 4. Section 5 reports estimation results and Section 6 concludes.

2 Institutional framework

2.1 Early childcare provision

In Europe, parents are provided several *early childcare* solutions for their children who are not enrolled in school yet. Early childcare refers to childcare solutions for very young children aged from 0 to pre-elementary or elementary school age. It involves childcare-type settings, i.e. collective day-care centres and home-based care by child-minders, but not education-type settings such as pre-schools. Parents can ask for a slot in a day-care centre or at a child-minder's when the child is a few weeks or a few months old, depending on countries (See

Table A1 for further details.). Day-care is the most widespread option and child-minders only represent a small part of public childcare provision. France is an exception since it is the only European country where home-based childcare arrangements are more used than centre-based childcare solutions [Eurydice, 2019]. Only a few European countries grant statutory rights to early childcare, but none of these countries are included in our sample. Such early childcare solutions are not free of charge, but fees are regulated and often subsidized, at least in public childcare infrastructures. Finally, the level of childcare costs largely varies across countries.

Figure 1 displays *effort rates* of net childcare costs in countries that are studied in this paper (except from Cyprus which data is not provided) in 2004 and 2020. This indicator is collected from the Organisation for Economic Co-operation and Development (OECD)⁵. Childcare costs are net of all public childcare benefits and are calculated among parents using childcare facilities. In order to get a measure of effort rate, they are related to the income of a couple with two children with average wages in their own country. Except for Portugal, Greece, Luxembourg and Hungary, net childcare costs effort rates decreased from 2004 to 2020, in line with European Commission's *European Pillar of Social Rights Action Plan* which objective - among others - is to provide equitable access to childcare solutions for all children [European Commission, 2017]. Childcare costs represent the largest part of parents' average income in the Czech Republic in both 2004 (19%) and 2020 (16%) and the lowest in Greece in 2004 (3%) and in Portugal in 2020 (5%).

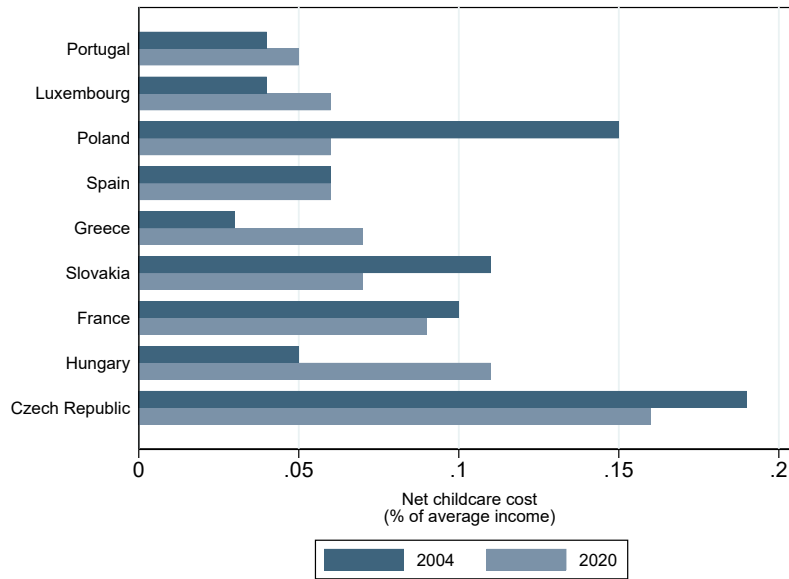
Such costs might prevent mothers from participating to the labour market or working full-time. If these early childcare solutions do not enhance mothers' labour supply due to high costs, it is possible that pre-school enrollment is more effective given that it is free in almost all cases. We present pre-school systems in the following section.

2.2 Pre-school system

Pre-school programs constitute the first stage of educational system. They are subject to education guidelines that some countries include in their legislation [Eurydice, 2019]. The purpose of pre-schools is to prepare children, during one or several years, to elementary school with programs aimed at fostering child development and socialisation. Pre-school programs are taught

⁵For further details, see the OECD website : <https://stats.oecd.org/Index.aspx?DataSetCode=NCC>

Figure 1: Effort rates of net childcare costs by country



Note: Effort rates are defined as the ratio between net childcare costs and average income in each country. It is interpreted as the share of income allocated to childcare costs (net of family benefits).

Sample: Countries included in the EU-SILC selected sample.

Source: Organisation for Economic Co-operation and Development (OECD), 2004 and 2020.

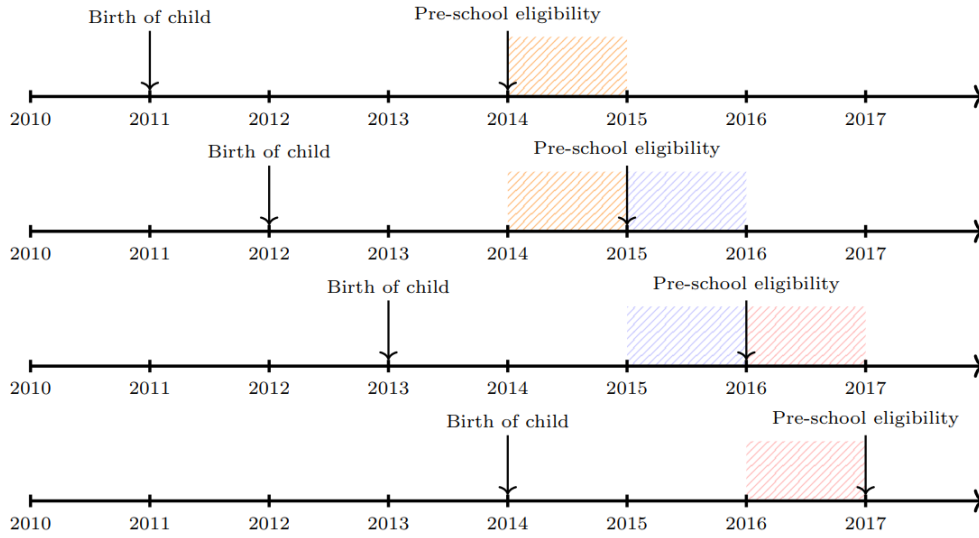
by professional instructors whose level of qualification depends on the country's requirements, but is usually higher than the one required for early childcare professionals.

In some countries, pre-school programs are compulsory, most of the time only the year prior to elementary school entry but some of them implement compulsory pre-school programs from the earliest age.⁶ When not compulsory, it is often made available one or several years earlier for parents willing to enroll their child as soon as possible, usually from the age of 3. In some cases, children are legally entitled to a slot in a pre-school program. Among the countries that are part of this study, this is the case for all of them except for Cyprus, France⁷ and Slovakia. When compulsory, pre-school programs are offered for free by public institutions but parents can also use private options that can charge supplementary fees. Free pre-school education is also provided even when it is not compulsory, except for Cyprus, the Czech Republic and Slovakia in this sample (see Table A1 for more details).

⁶It is the case in Hungary where pre-school is compulsory from the age of 3. It is also the case for France since 2019.

⁷Before 2019. Since then, pre-school is compulsory from the age of 3, and still available from 2 years old under specific conditions.

Figure 2: Pre-school eligibility rule



Note : This diagram illustrates the rule for pre-school eligibility with a 3-year-old threshold. The orange hatched area corresponds to the year in which a child born in 2011 becomes eligible for a pre-school start in September of the same year (September 2014). In the same way, the blue hatched area indicates the year in which a child born one year later (here, in 2012) will be eligible for pre-school. In the second line, if we observe individuals in 2015 in a month prior to school start (from January to August), the child born in 2011 will already be enrolled from the year before (orange area) while the child born in 2012 will be enrolled a few months later but will not be technically enrolled at the time of interview. The same explanation can be brought for the third line (blue and red areas) and so on.

In most cases, pre-school eligibility is granted following a birth-year based principle. The rule is such that any child turning a given age a in a calendar year t (or, in the strictest cases, before school start in September of year t)⁸ can start pre-school in September of that same year t . The age cut-off varies across countries. Among countries included in the study, pre-school is available for children turning 3 years old during the year (or before school start in some cases) in all countries except for Greece (4 years old) and Poland (5 years old before 2015 and 4 years old between 2015 and 2017). In the analysis, we only include countries and years in which pre-school is available from 3 or 4 years old.

Figure 2 represents a timeline illustrating the eligibility rule for pre-school based on the most common setting, the 3-year-old cut-off. For example, a child born in 2011 will turn 3 in 2014 and will be eligible for pre-school in September 2014 if the rule is such that all children turning 3 within the year can enroll.⁹ In 2015, this child is very likely to be enrolled already since she

⁸This difference does not modify our empirical strategy. This point is further detailed in Section 3.

⁹In the strictest cases where the child must turn 3 before school start, the child will be eligible in the following

was eligible as soon as September 2014. On the opposite, a child born in 2012 will be eligible for pre-school in September 2015 since she will only turn 2 in 2014. This means that, in any month of observation prior to September in 2015 (i.e. from January to August 2015), the child born in 2011 has a very high chance of being enrolled while it is less likely for the one born in 2012 since this child will be eligible a few months after the interview but not before.¹⁰ We could even argue that a child born in December 2011 is more likely to be enrolled in pre-school than a child born in January 2012 given the discontinuity generated by the rule based on the calendar year. We use this discontinuity in eligibility rule to predict pre-school enrollment at time t in order to correct for the endogenous relationship between the child’s pre-school enrollment and the mother’s labour supply. The methodology is further detailed in the following section.

3 Methodology

Our main purpose is to identify the causal effect of pre-school enrollment of a child on the mother’s labour supply. This relationship can be represented with the following linear model:

$$Y_{ict} = \alpha + \beta D_{ict} + \epsilon_{ict}$$

where α is an intercept, β is the coefficient of interest to be estimated, Y_{ict} is one of the labour supply outcomes of mother i in country c at time t , D_{ict} is a binary variable indicating whether her child is enrolled in a pre-school ($D_{ict} = 1$) or not ($D_{ict} = 0$) and ϵ_{ict} is an unobserved error term. If we assume that ϵ_{ict} is uncorrelated with D_{ict} , we can estimate the causal effect of pre-school enrollment with a simple regression using the Ordinary Least Squares estimator. But this assumption may not be valid since mothers seeking to enroll their child are likely to differ from other mothers with respect to unobservable characteristics related to labour supply, e.g. career-attachment, family values, or opportunity cost. In order to correct for this endogenous relationship, we exploit the discontinuity between the child’s year of birth and her eligibility to pre-school to predict pre-school enrollment. We consider the youngest age at which children are eligible to pre-school in a given country, implying imperfect compliance of individuals due to

year only, unless earlier starts are possible. See Section 3 for more details on how we exploit this information in our empirical strategy.

¹⁰Except from cses where earlier starts are possible. Once again, see Section 3 for more details.

the fact that pre-school is not compulsory in most cases (see Section 2). For this reason, we use a fuzzy regression on discontinuity design (RDD).

Eligibility rule. Eligibility is based on a date-of-birth criteria. More precisely, there are two types of eligibility rules.¹¹ In a first case, the rule is such that children turning a given age a within the calendar year can qualify for pre-school starting in September of the same year t . These children are born in year $t - a$. This implies that children born in calendar year $t - a - 1$ (i.e. the year before) are eligible as soon as in year $t - 1$ and thus had the possibility to start pre-school in September $t - 1$, while children born in $t - a$ will only qualify for pre-school in September t . For this reason, we expect a higher pre-school enrollment rate for children born in year $t - a - 1$ than for children born a year later in any interview month prior to September of year t (i.e. between January and August of year t).

In a second case, the rule is that children turning a given age a before school start (i.e. September) can qualify for pre-school in September of year t . Here, the difference with the first case is that eligible children are also born in year $t - a$, but necessarily in a month before September (i.e. born between January and the school start day, usually before mid-September). Children born in year $t - a$ after September (i.e. when school starts) have to wait September $t + 1$ to enroll. However, some schools give parents the possibility to enroll early under specific conditions, even if this is not the official rule. As it happens, children born after September (therefore not officially eligible) can be enrolled anyway if there are slots that remain available and if they are deemed ready for pre-school by instructors. Although this practice is usually not officially enshrined in the law, it is common knowledge that some schools actually do it. For this reason, we decide to use the first rule (i.e. the *born-within-the-year* criteria) for all countries since it is not possible to know the extent to which countries are flexible or not on their *born-before-September* rule.

This is important to keep in mind that doing so does not impair the identification of the causal effect we aim to estimate. To be more precise, it implies that, if the *born-before-September* rule is strictly implemented, i.e. without possibility of early starts, pre-school enrollment of children born in year $t - a - 1$ will be lower for children born in the second part of the year (i.e. after September). This is actually what we can see in Figure 3. This implies that the jump between

¹¹See Table A1 for more details on which country follows which rule.

December $t-a$ and January $t-a-1$ is smaller than what we could expect when eligibility is only based as in the first case we described earlier (all children born in $t-a-1$ are eligible). Hence, the main implication is to make our first-stage statistic weaker. One should note that we also observe an upward trend among children who are theoretically not eligible (in year $t-a$). This is because children in their second year of life and born at the very beginning are allowed enter school early if slots are available. This should also make our first-stage weaker. Nonetheless, we show in Table 2 that the first-stage is strong enough (F -stat > 900) for the relevance assumption to be met.

Technically, the eligibility rule is defined as follows. Let us first define q_{ict} , the normalized distance from eligibility for a child i living in country c in year t . In the data, we observe the year and the quarter of birth of each child. We therefore define this variable as the distance from the eligibility threshold in quarters. It is positive when a child is eligible (i.e born in year $t-a-1$) and it is negative when she is not (i.e born in year $t-a$). q_{ict} is equal to 0, 1, 2 and 3 for children born in the fourth, third, second and first quarter of year $t-a-1$, respectively. These children are eligible for pre-school since year $t-1$. On the contrary, this variable is equal to -1 (resp. -2 , -3 , -4) for children born in the first (resp. second, third, fourth) quarter of year $t-a$. These children are not eligible yet. The eligibility status can therefore be defined with the following binary variable, which is equal to 1 if the condition in the brackets is verified and 0 otherwise:

$$E_{ict} = \mathbb{1}[q_{ict} \geq 0] \quad (1)$$

Local Average Treatment Effect. In the case of a fuzzy RDD, the effect of pre-school enrollment on maternal labour supply can be estimated as the ratio of the drop in the labour supply outcome Y and the drop in the probability of pre-school enrollment D at the eligibility cut-off $q = 0$ [Hahn et al., 2001] :

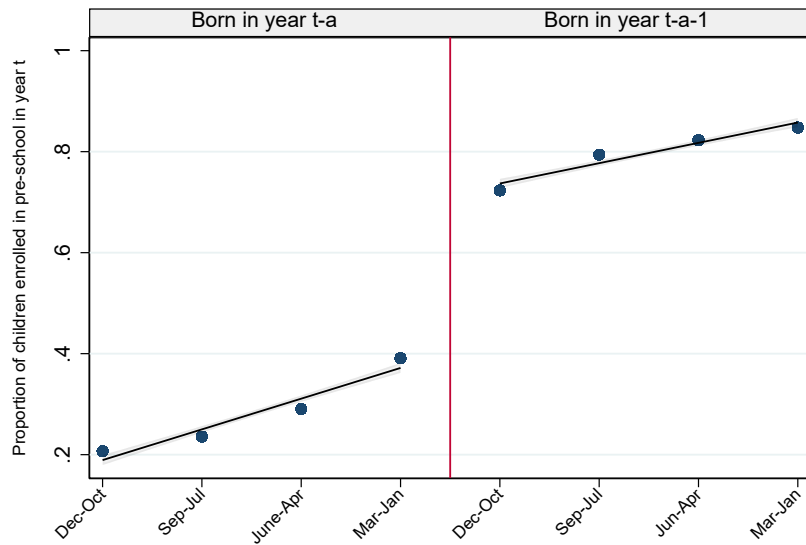
$$\tau_{FRD} = \frac{\lim_{q \rightarrow 0^+} E(Y|q) - \lim_{q \rightarrow 0^-} E(Y|q)}{\lim_{q \rightarrow 0^+} P(D|q) - \lim_{q \rightarrow 0^-} P(D|q)} \quad (2)$$

τ_{FRD} is the local average treatment effect (LATE), i.e. the effect of pre-school enrollment on

labour supply among complying mothers at the neighbourhood of the cutoff point, under the following assumptions.

Relevance. It requires a discontinuity in the probability of pre-school enrollment at the eligibility threshold. In other words, the pre-school eligibility status should be a strong predictor of enrollment. As shown in Figure 3, there is a clear drop in the probability of pre-school enrollment at the cut-off point. This graphical evidence is confirmed by first-stage results in Table 2 : an eligible child, i.e. born in year $t - a - 1$, has a higher probability to be enrolled in pre-school at the time of interview t by 31 to 32 percentage points (ppt) according to the sample considered ($p < 0.01$). In addition, the F-statistic of the Wald test for weak instruments is far above the threshold advocated by the econometric literature [Stock et al., 2002] which confirms the relevance assumption. Further details are given in Section 5.

Figure 3: Discontinuity in the probability of pre-school enrollment at the eligibility threshold



Note : Graphical representation of the fuzzy RDD first-stage regression. Dots are the observed proportion of children enrolled in a pre-school at time t for each quarter of birth around the pre-school eligibility threshold. The cut-off which determines eligibility (i.e. born in year $t-a-1$) is shown by the vertical red line. Linear fitting lines are represented by the solid black lines on each side of the cut-off with confidence intervals at the level of 95% as grey shaded areas.

Independence. The second assumption requires the eligibility status to be locally randomly assigned. In other words, pre-school eligibility should be orthogonal to the error term for the empirical strategy to yield unbiased estimates. With an RDD approach, the independence

assumption requires the *absence of manipulation* of the running variable and the *continuity of potential outcomes* at the cut-off point.

Manipulation could occur in the case mothers willing to go back to work as soon as possible would manipulate birth timing so that the child can enter pre-school a year earlier. In other words, they would give birth in late $t - a - 1$ rather than in early $t - a$. An argument to rule out this possibility is that procreation and timing of birth are not under mothers' perfect control. Despite their anticipation, mothers might not manage to give birth according to their expected timing. As advocated by Lee and Lemieux [2010], "*when individuals have imprecise control over the assignment variable, even if some are especially likely to have values of X near the cutoff, every individual will have approximately the same probability of having an X that is just above (receiving the treatment) or just below (being denied the treatment) the cutoff*". This means that such observations are still randomly distributed around the cut-off. Thus, manipulation (if it exists), should not be a threat to the validity of our identification strategy since it is imperfect. The absence of manipulation can be checked by looking at the local distribution of the running variable, i.e. the number of children born in each quarter around the eligibility threshold. Figure C1 shows the absence of discontinuity in the distribution of births around the threshold.

We also check the continuity of mothers' observable characteristics at the eligibility threshold. As shown in Table 1, *eligible mothers*, i.e. mothers of children born in year $t - a - 1$, significantly differ from *ineligible* ones, i.e. mothers of children born in year $t - a$, with respect to their age, probability to be single and the number of children they have. Column 4 of Table 1 reports the difference in average values between the *eligible* (Column 2) and the *ineligible mothers* (Column 3).¹² It shows that *eligible mothers* are older by almost 10 months¹³ and are more likely to be lone mothers by 1.5 percentage point (ppt). They also have more children on average : their probability to have only one child is lower by 6 ppt while their probability to have two children (resp. more than two children) is higher by 4 ppt (resp. 2 ppt). Such differences are not surprising since all these characteristics are correlated with the child's age. If *eligible* and *ineligible mothers* significantly differ on average with respect to these characteristics, those differences do not impair the RDD validity if the average value of any individual characteristic

¹²From Columns 1 to 3, standard-deviations are presented in parentheses. In Column 4, we report cluster-adjusted standard errors in parentheses. Standard errors are corrected for clustering at the mother's level.

¹³ $9.826 = 0.819 * 12$ months

does not sharply change at the eligibility threshold. We check this assumption taking each of these individual characteristics as the outcome of the RDD estimation and show the absence of discontinuity for all of them (see Figure C2 and Table C1).

Exclusion Restriction. The child’s pre-school eligibility status should affect maternal labour supply only through its effect on pre-school enrollment. This hypothesis would not be verified if, in a given country, the birth year threshold defining pre-school eligibility status also grants eligibility to another program that would influence maternal labour supply, e.g. family allowances. Because the perceived amount of family allowances is observed in the data, we rule out any violation of the exclusion restriction for this reason by confirming the continuity in the probability to receive family allowances and in the conditional amount of family allowances among recipients around the birth year threshold (see Table D3).

Monotonicity (or the absence of defiers). In this case, defying mothers would enroll their child when she is not eligible but not when she is. If enrolling a child before the official cut-off is possible in some cases (see Section 2), such behaviour is not realistic because there is no reason to explain why a mother would want to enroll her child as early as possible (before eligibility), but not after her child has become eligible. Therefore, the monotonicity assumption should hold in our context.

Estimation. τ_{FRD} can be estimated using a Two Stage Least Squares (TSLS) estimation in a fuzzy regression on discontinuity design. We use a triangular Kernel weighted estimation that puts more weight on observations closer to the threshold, which allows to obtain similar estimates as with a local linear regression [Imbens and Lemieux, 2008; Lee and Lemieux, 2010]. In the first stage, we estimate the following regression:

$$D_{ict} = \alpha_0 + \alpha_1 E_{ict} + \alpha_2 f(q_{ict}) + \alpha_3 E_{ict} * f(q_{ict}) + \gamma_c + \delta_t + \mu_{ict} \quad (3)$$

where $f(q_{ict})$ is a linear function of the distance to the eligibility cut-off. We allow its trend to be different from one side of the cut-off to the other including an interaction term between this function and E_{ict} . γ_c and δ_t respectively denote country and year specific effects, while is μ_{ict} an unobserved error term.

Then, we estimate the following second-stage equation using the predicted probabilities of pre-school enrollment, \widehat{D}_{ict} as follows :

$$Y_{ict} = \beta_0 + \beta_1 \widehat{D}_{ict} + \beta_2 f(q_{ict}) + \beta_3 E_{ict} * f(q_{ict}) + \zeta_c + \eta_t + \nu_{ict} \quad (4)$$

where ζ_c and η_t denote country and year specific effects respectively, and ν_{ict} is the unobserved error term. β_1 is the estimated LATE, our coefficient of interest. It is important to note that, due to the longitudinal dimension of the data, mothers are observed several times. Thus, standard-errors are corrected for clustering at the mother's level. Results of this regression model are presented in Section 5.

4 Data

This paper uses data from the European Union Statistics on Income and Living Conditions (EU-SILC) from 2004 to 2020. The survey is harmonized across EU member states and collects information at the household level with a rotating panel design (households are interviewed for a period of four years). It provides information at the household level such as household income, living conditions and family structure. All household members aged at least 16 are eligible for an individual survey including questions on education, labour supply, health and other related topics. Individuals with children are asked about formal and informal childcare use and characteristics.

4.1 Sample restriction

Following the estimation strategy described in Section 3, only mothers with a child born within a bandwidth of one year around the pre-school eligibility threshold are included in the sample. Hence, we focus on mothers with children who are in their year of eligibility (with age a or born in year $t - a$) and those with children who have been eligible the year before (with age $a + 1$, or born in year $t - a - 1$). For instance, if the eligibility rule is such as a child can qualify for pre-school in September of year t if she turns 3 within the year (or before school start in some cases), mothers with a child in her 3rd or 4th year of life are included in the sample. Concretely, we select mothers of children born in one of the four quarters of birth before or after the eligibility

threshold. Because children born in year $t - a$ become eligible for pre-school in September t , it requires to exclude all mothers interviewed after school start in order to avoid a bias towards zero due to individuals considered as ineligible but who would be eligible to pre-school in reality. As we only observe the quarter of interview instead of the precise month, we exclude the last two quarters of interview, i.e. July-August-September and October-November-December in order to be sure to exclude all individuals observed from September. We also proceed to restrictions at the country level. Countries for which the child's quarter of birth is not available or with no clear eligibility age rule for pre-school enrollment at the time of interview are excluded. Taking all these sample restrictions into account, the total sample contains 42,343 observations for 37,546 in 10 countries: Cyprus, Czech Republic, France, Greece, Hungary, Luxembourg, Poland, Portugal, Slovakia and Spain. The first column of Table 1 displays descriptive statistics for the whole sample. Table B1 reports the same statistics for each country separately.

4.2 Description of variables

Labour supply. Several variables are used to measure maternal labour supply. Maternal employment is measured with a binary variable, denoted as *Working*, which is equal to 1 if the mother declares to be working at time t , 0 otherwise. In order to take job search into account, we use a variable *Active*, which is equal to 1 if the mother is active on the labour market (i.e. working or searching for a job) and 0 otherwise. We use two different measures among working mothers to study the intensive margin of labour supply. We use the working hours per week, which is a continuous variable, and a binary variable that captures part-time employment that is equal to 1 if the mother works less than 35 hours per week. In the sample, 58% of mothers are employed and 66% are active, but there is a large cross country variation (see Table B1). Employment and activity rates are the lowest in Czech Republic (33% and 39% respectively) while they are the highest in Portugal (78% and 88% respectively). Working mothers spend 35 hours per week at work on average and 30% of them working part-time. The share of mothers working part-time is high in Luxembourg (46%), while the lowest in Portugal (9%).

Pre-school enrollment. *Pre-school enrollment* is defined as a binary variable equal to 1 if the child is enrolled at the time of interview and 0 otherwise. In the whole sample, 53% of children are enrolled in a pre-school. Enrollment is the highest in Spain, Greece and in France, where

it is higher than 60%, which is in line with the fact that these countries give opportunities for early starts.

Childcare use. Formal and informal childcare are substitutes of mothers', and their use can shape the relationship between labour supply and the child's preschool enrollment. We therefore conduct analyses where formal child care use is the outcome variable. Mothers are asked about their use of both formal and informal childcare arrangements. Formal childcare includes two different childcare solutions : day-care and child-minders. Informal childcare refers to care provided by any person living in the household or not, e.g. grand-parents, friends or neighbours. Mothers are asked how many daily hours her child is cared for in each childcare type. We define daycare, childminder and informal care take-up rates as binary variables taking value 1 if the child is cared for at least 1 hour per day by the given childcare type. In the whole sample, informal childcare take-up (29%) is higher than for daycare (17%) and for childminder care (6%). The share of mothers using informal childcare reaches 43% in Czech Republic while it is the lowest in Spain (12%) (see Table B1). With respect to formal childcare, it is in Luxembourg that mothers have the highest day-care and childminder-care take-up rates (48% and 13% respectively). In line with their high use of informal care, Czech mothers' use of day-care is very low (0.3%).

Individual characteristics. Mothers in the whole sample are 34 years old on average, 6% of them are single and 41% have completed high school at least. The number of children is distributed such as half of mothers of the sample has two children, 28% have only one child and 23% have more than two children. Descriptive statistics by country are reported in Table B1. The baseline model does not control for any of these individual characteristics. Our results are not sensitive to their inclusion (see Tables D1 and D2).

Table 1: Descriptive statistics

	Whole sample	Eligible	Ineligible	Difference
<i>Labour supply</i>				
Employment rate (%)	0.5752	0.6108	0.5412	0.0696*** (0.0051)
Activity rate (%)	0.6643	0.7049	0.6255	0.0793*** (0.0049)
Conditional weekly working hours	35.3690 (9.3897)	35.4436 (9.2651)	35.2885 (9.5222)	0.1551 (0.1278)
Working part-time (%)	0.3065	0.3017	0.3117	-0.0099* (0.0063)
<i>Pre-school</i>				
Pre-school enrollment rate (%)	0.5331	0.7969	0.2809	0.5160*** (0.0044)
<i>Formal childcare</i>				
Daycare take-up (%)	0.1661	0.0896	0.2392	-0.1496*** (0.0037)
Childminder care take-up (%)	0.0565	0.0486	0.0642	-0.0156*** (0.0024)
<i>Informal childcare</i>				
Informal childcare take-up (%)	0.2892	0.2721	0.3055	-0.0334*** (0.0047)
<i>Individual characteristics</i>				
Age	34.4988 (5.1339)	34.9174 (5.1168)	34.0986 (5.1183)	0.8188*** (0.0543)
Single mother (%)	0.0589	0.0652	0.0528	0.0125*** (0.0025)
Completed high school at least (%)	0.4062	0.4038	0.4086	-0.0048 (0.0052)
One child (%)	0.2782	0.2460	0.3090	-0.0631*** (0.0047)
Two children (%)	0.4953	0.5177	0.4740	0.0437*** (0.0053)
More than two children (%)	0.2264	0.2363	0.2170	0.0193*** (0.0044)
Number of individuals	37,546	19,787	20,591	37,546
Number of observations	42,343	20,696	21,647	42,343

Note: Columns 1, 2 and 3 report means with standard deviations in parentheses. Conditional weekly working hours and the probability to work part-time are reported for working mothers only (Nb obs.= 24,356; Nb ind.= 22,223). Column 4 reports results of the mean comparison t-test for each variable (i.e. the difference in means between eligible (Column 2) and ineligible (Column 3) individuals with cluster-adjusted standard-errors at the individual level in parentheses. *p<0.10, **p<0.05, ***p<0.01. Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

5 Estimation results

5.1 Average effects

In this section, we present the estimated effects on mothers' labour supply (Section 5.1.1) and informal childcare use (Section 5.1.2) for the whole sample. Heterogeneous effects are explored at the individual and the country levels in Sections 5.2 and 5.3, respectively.

5.1.1 Mothers' labour supply

Table 2 reports estimation results for maternal labour supply. The upper part of the table contains results for the first-stage estimation, i.e. the effect of the child's pre-school eligibility status on the probability of enrollment. In the lower part of the table, we report the reduced-form estimates, i.e. the direct effect of the child's pre-school eligibility status on the mother's labour supply outcomes, as well as the second-stage estimates, i.e. the effect of pre-school enrollment on maternal labour supply. In order to disentangle the intensive from the extensive margin of the effect of pre-school enrollment on labour supply, the estimation is run on the whole sample of mothers (Columns 1 and 2) and on working mothers (Columns 3 and 4).

First-stage results show that the eligibility rule based on the birth-year threshold is a good predictor of pre-school enrollment for all individuals. In all cases, the F-stat equals 925 in the whole sample and 681 among working mothers. With respect to reduced-form results, none of the coefficients significantly differ from zero at any conventional level, which means that eligibility to pre-school based on the child's birth-year has no significant impact on mothers' labour market outcomes on average. In Figure 4, the absence of discontinuity at the cut-off is clearly visible. Logically, the second-stage coefficient, which is equal to the ratio between the reduced-form and the first-stage estimates, is insignificant. Thus, we do not identify any causal effect of pre-school enrollment on labour supply.

The absence of an average effect could be explained by several potential factors. On one hand, mothers might have resumed work before the child enters pre-school thanks to earlier childcare arrangements. In countries where formal childcare solutions are made available and affordable for parents, pre-school might not be a leverage for maternal labour supply since mothers have returned on the labour market shortly after a birth (see Section 1). On another hand, a low

supply of formal childcare, high costs of childcare or strong preferences for informal childcare could explain such a result. Parents might rely on childcare provided for free by relatives (mostly grand-parents), friends or neighbours. The literature supports that grand-parental childcare provision allows mothers to increase their labour supply [Bratti et al., 2018; Aparicio Fenoll, 2020; Pinto, 2023]. In addition, some mothers might not react to their child’s enrollment to pre-school if they are on maternity or parental leave (in the case they have younger children or anticipate another birth). Our average effect can therefore hide potential heterogeneous effects that we explore later.

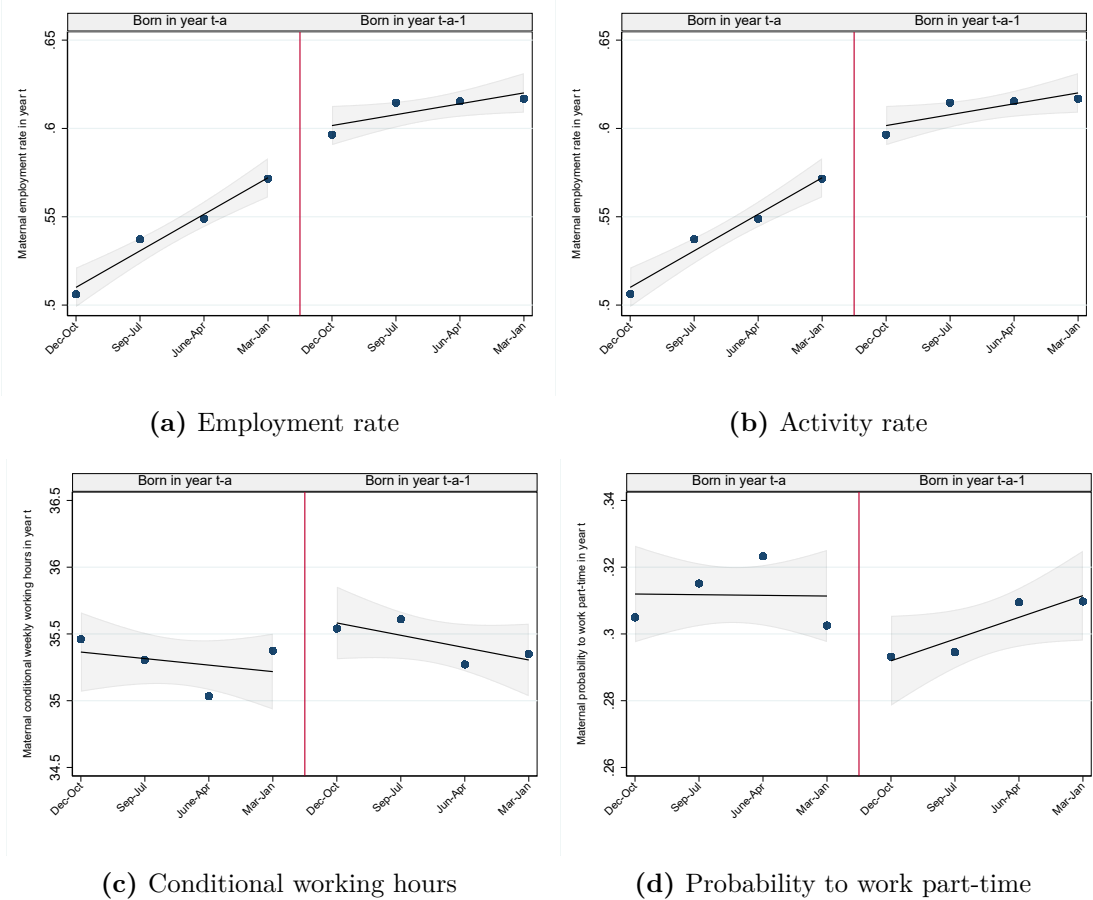
Table 2: Effect of pre-school enrollment on mothers’ labour supply

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school Eligibility	0.3112*** (0.0102)	0.3112*** (0.0102)	0.3225*** (0.0124)	0.3325*** (0.0124)
Reduced-form				
Pre-school Eligibility	0.0127 (0.0123)	0.0164 (0.0116)	0.2960 (0.3006)	-0.0129 (0.0148)
Second-stage				
Pre-school enrollment	0.0409 (0.0394)	0.0527 (0.0370)	0.9177 (0.9318)	0.0401 (0.0458)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	924.924	924.924	681.494	681.494
Number of individuals	33,660	33,660	19,681	19,681
Number of observations	37,349	37,349	21,276	21,276

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Figure 4: Graphical representation of reduced-form estimations : Maternal labour market outcomes



Note : Graphical representation of the fuzzy RDD reduced-form regression. Dots are the observed maternal (a) employment rate, (b) activity rate, (c) average conditional weekly working hours and (d) probability to work part-time at time t for each quarter of birth around the pre-school eligibility threshold. The cut-off which determines eligibility is shown by the vertical red line (born in year $t-a-1$). Linear fitting lines are represented by solid black lines on each side of the cut-off with confidence intervals at the level of 95% as grey shaded areas.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

5.1.2 Childcare use

While pre-school is designed to substitute formal early childcare solutions, substitution effects between pre-school and informal care are less straightforward. On one hand, a decrease in informal childcare use can be expected among mothers who rely on such type of childcare before pre-school age to be able to work. On the other hand, there could be an increase in informal childcare use among mothers resuming work or increasing their working hours due to pre-school enrollment, and previously relying on formal childcare solutions. We could think of mothers with the longest daily working hours could need informal solutions to take care of their child after school for example.

Table 3 displays the estimated effect of pre-school enrollment on informal childcare take-up (Column 1) and informal childcare weekly hours among mothers who use this type of care (Column 2). Results are in line with the first hypothesis. Pre-school enrollment decreases informal childcare take-up by 9 ppt and weekly hours by almost 5 units ($p < 0.05$). Figure 5 graphically supports this finding. The effect on informal childcare is rather strong given that it represents a reduction of 32%¹⁴ at the extensive margin and of 29%¹⁵ at the intensive margin. Nonetheless, the substitution effect is not one for one, indicating that informal caregivers play a role in helping mothers of young children maintaining their work-family balance.

¹⁴ $-0.3164 = -0.0915 / 0.2892$, with 0.2892 the average informal childcare take-up in the whole sample.

¹⁵ $-0.2916 = -4.5945 / 15.7403$, with 15.7403 the average number of hours per week the child spends at an informal caregiver's among mothers who use informal childcare.

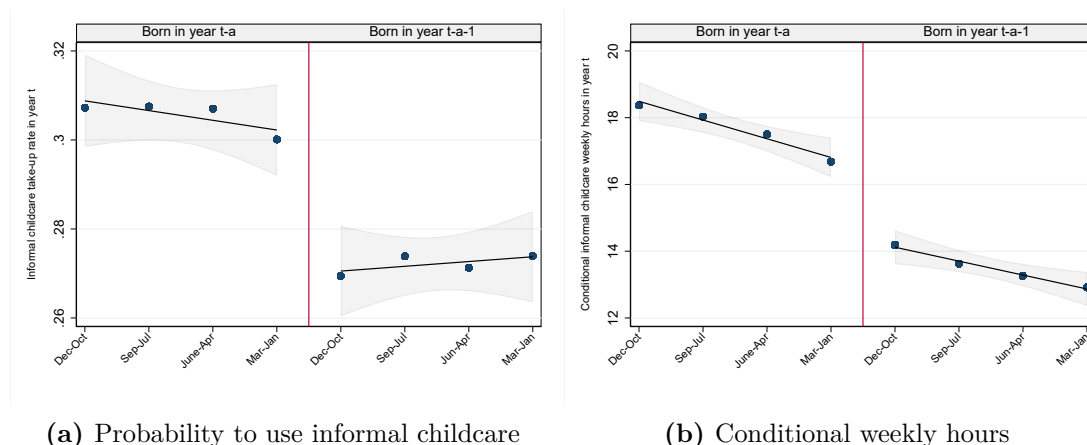
Table 3: Effect of pre-school enrollment on informal childcare use

All mothers		
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school Eligibility	0.3112*** (0.0102)	0.3187*** (0.0209)
	Informal childcare take-up	Conditional childcare weekly hours
	(1)	(2)
Reduced-form		
Pre-school Eligibility	-0.0285** (0.0110)	-1.4644** (0.5848)
Second-stage		
Pre-school enrollment	-0.0915** (0.0355)	-4.5945** (1.7743)
Country specific effect	yes	yes
Year specific effect	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	924.924	231.741
Number of individuals	33,660	10,194
Number of observations	37,349	10,877

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Figure 5: Graphical representation of reduced-form estimations : Informal childcare use



Note : Graphical representation of the fuzzy RDD reduced-form regression. Dots are the observed maternal (a) probability to use informal childcare and (b) conditional informal childcare weekly hours t for each quarter of birth around the pre-school eligibility threshold. The cut-off which determines eligibility is shown by the vertical red line (born in year $t-a-1$). Linear fitting lines are represented by the solid black lines on each side of the cut-off with confidence intervals at the level of 95% as grey shaded areas.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

5.2 Heterogeneous effects with respect to individual characteristics

The effect of pre-school enrollment is likely to differ according to mothers' characteristics. We present the estimation model for heterogeneity analyses at the individual level in Section E.1 as well as corresponding results tables in Section E.2. A summary of heterogeneous results with respect to individual characteristics is available in Table 4.

Table 4: Summary of heterogeneous results at the individual level

	Labour supply				Informal childcare use	
	Working	Active	Working hours	Part-time	Take-up	Cond. hours
<i>Education level</i>						
Least educated	NS	NS	-2.607*	NS	-0.177***	NS
Most educated	NS	NS	+3.898***	-0.124**	NS	NS
<i>Partnership status</i>						
In a partnership	NS	NS	NS	NS	-0.101***	X
Single mother	NS	NS	NS	NS	-0.101***	X
<i>Birth order</i>						
Not the youngest child	NS	NS	NS	NS	NS	NS
Youngest child	+0.084*	NS	NS	NS	NS	NS

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. NS = Not significant. X = coefficient cannot be interpreted (further details in Section 5.2.2).

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

5.2.1 Level of education

First, Table E1 shows that pre-school enrollment reduces the intensive margin of labour supply for the least educated mothers (-2 weekly hours, $p < 0.1$). This result could be explained by an *income effect*. Because income is positively correlated with the level of education, mothers with less education could be constrained to work more hours to pay for early childcare when their child is not eligible for pre-school yet. For example, they might have to work extra hours at their workplace or take a second job to finance day-care or care provided by a childminder in order to stay on the labour market. Then, the relative price of childcare decreases when their child enrolls in a pre-school since it is provided almost free of charge, leading to a reduction in labour supply at the intensive margin. To support this idea, Table E2 shows that pre-school is a substitute to informal childcare at the extensive margin for the least educated mothers only (-18 ppt, $p < 0.01$), implying that these mothers probably also need to rely on informal childcare

solutions before pre-school in order to work enough hours to pay for formal childcare.

This mechanism is less likely to be at stake for mothers with more education because they are more likely to afford formal early childcare. Besides, we find an opposite result for the most educated since pre-school enrollment increases their weekly working hours by about 4 units ($p < 0.01$) and reduces their probability to work part-time by 12 ppt ($p < 0.05$). A potential explanation could be that the most educated women are more likely to have flexible working hours, due to a managerial position, for example. Table E1 shows that if they are more likely to be employed and active in general, they work fewer hours when they are employed, supporting the idea that they might adjust their working hours more easily. Those mothers could take advantage of this flexibility to work less when the child is not enrolled in pre-school yet, knowing that a reduction in working hours would have a more limited impact on their ability to pay for childcare compared to less educated mothers. Then, because attendance is usually much larger in a pre-school than in a day-care or at a childminder's because it is part of the schooling system, mothers might resume their pre-birth amount of working hours due to pre-school enrollment.

5.2.2 Partnership status

Some studies have shown that public pre-school is a leverage for lone mothers' labour supply but not for mothers who declare to be in a relationship [Cascio, 2009; Goux and Maurin, 2010; Fitzpatrick, 2010]. This heterogeneity could be attributed to lower living standards among single mothers compared to two-parent families. For this reason, we explore potential differences in our results by partnership status.

Table E3 shows that the partnership status does not yield any heterogeneous effect of pre-school enrollment on maternal labour supply. Substitution effects between pre-school and informal childcare arrangements are the same with respect to the mother's partnership status at the extensive margin (-10 ppt, $p < 0.01$). In this specific case, the intensive margin of the effect of pre-school on informal childcare use cannot be interpreted because the instrument is too weak, leading to biased estimates (Table E4). This is due to the low share of single mothers in the sample (see Table 1).

5.2.3 Birth order

Another potential source of heterogeneity could come from the child's birth order in the sibship. We could expect mothers to be more likely to resume work due to her child's enrollment in a pre-school if there are no younger children to provide care to at home. If the observed child is not the last of the sibship, the mother might still be on maternity or parental leave due to a more recent birth or might anticipate a following birth and therefore stay out of the labour market. Because birth order depends on the sibship size (i.e. an only child is always the last one), the number of children is controlled for in the regression.

Results suggest that pre-school helps mothers resuming work when their child is the last born, irrespective of the number of children they have (Table E5). More precisely, the employment rate increases by 9 ppt among mothers who do not have younger children when their child enrolls in a pre-school ($p < 0.1$). Birth order does not yield any heterogeneous effect in terms of informal childcare use (Table E6).

5.3 Heterogeneous effects with respect to country level characteristics

The main contribution of this paper is to exploit cross-country differences. To this aim, we investigate potential heterogeneity in our results according to country level characteristics. For each of them, we use the method described in Section F.1. We use measures of gender norms, generosity in terms of family benefits and the net cost of childcare that we describe with more details in Section F.2. Results tables are available in Section F.3. A summary of heterogeneous results with respect to country characteristics is available in Table 5.

5.3.1 Gender norms

The economic literature has investigated the effect of gender norms on female labour supply. Results show that conservative gender role attitudes, that is men perceived as *breadwinners* and women as *caregivers*, hamper women's labour market participation [Antecol, 2000; Fernández and Fogli, 2009; Bursztyn et al., 2020]. Therefore, two types of results can be expected. First, pre-school enrollment could constitute a leverage for women's labour supply in countries with low female employment rates due to traditional gender role attitudes. Pre-school enrollment

Table 5: Summary of heterogeneous results at the country level

	Labour supply				Informal childcare use	
	Working	Active	Working hours	Part-time	Take-up	Cond. hours
<i>Gender norms</i>						
Least conservative	NS	NS	NS	NS	-0.111***	NS
Most conservative	+0.142**	+0.173***	NS	NS	-0.111***	NS
<i>In-cash benefits</i>						
Least generous	NS	NS	NS	NS	-0.092*	-8.882***
Most generous	NS	NS	NS	NS	-0.092*	NS
<i>In-kind benefits</i>						
Least generous	+0.272***	+0.299***	NS	NS	-0.164**	-8.632***
Most generous	NS	NS	NS	NS	-0.164**	-8.632***
<i>Net cost of childcare</i>						
Least costly	NS	NS	NS	NS	-0.104*	-8.192***
Most costly	NS	NS	NS	NS	-0.104*	-8.192***

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. NS = Not significant.

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

would allow these women to provide less hours of care to their child. Second, pre-school might have no effect in such countries if conservative gender norms keep mothers away from the labour market irrespective of public childcare provision.

To answer this question, we construct a measure for gender norms at the country level using data from the Integrated Values Surveys (IVS) which merges the European Values Study (EVS) and the World Value Survey (WVS) files together. These surveys have been used in the literature exploring the relationship between gender norms and economic outcomes [Fortin, 2005; Alesina et al., 2015; Bonsang et al., 2017]. Both of them collect information on individuals' beliefs and values related to several topics in European countries from 1981 onwards. This survey is of particular interest for our study since individuals are asked about their opinions regarding gender role attitudes.

We describe how we construct our Gender Norms Index (GNI) in Section F.2.1. Out of this index, we construct a binary variable equal to 1 for GNI values above the median, and 0 otherwise (see Figure F2 for the distribution of the GNI). We therefore classify Poland, Slovakia, Cyprus, Portugal, Czech Republic and Greece as *the most conservative* countries, and Spain, Hungary, France and Luxembourg as the *least conservative* countries of our sample.

Table F1 shows that mothers living in countries with stronger gender stereotypes work less (-24

ppt) and are less active (-31 ppt), but those who work have longer working hours (+4 weekly hours) and a lower probability to work part-time by 26 ppt than in less conservative countries. We observe a positive effect of pre-school enrollment on maternal labour supply at the extensive margin in the most conservative countries only. More precisely, we estimate an increase in the employment rate by 14 ppt ($p < 0.05$) and in the activity rate by 17 ppt ($p < 0.01$). No significant effect is found at the intensive margin. This result follows the first hypothesis according to which the provision of pre-school helps mothers resuming work or encourages them to search for a job in countries where gender norms are the most conservative and female labour supply is lower on average. We find that pre-school is a substitute to informal childcare at the extensive margin irrespective of the level of conservatism (-11 ppt, $p < 0.01$) (Table F2).

5.3.2 Early childcare characteristics

As explained in Section 2, public pre-school is provided (almost) free of charge with a legal entitlement for all children in most cases, while it is not the case for earlier formal childcare solutions. One could therefore expect pre-school to lever mothers' labour supply, especially in countries where earlier formal childcare solutions are not available enough - because of a shortage in childcare slots - or not affordable enough - because of low levels of subsidies. The purpose of this section is to further investigate the role played by pre-school enrollment in maternal labour supply according to the country's *i*) level of generosity in terms of family benefits and *ii*) the net cost of childcare. To this aim, we collect information from the Organisation for Economic Cooperation and Development (OECD) *Social and Welfare Statistics*.

Family benefits. We use the *Family benefits public spending* indicator to measure countries' level of generosity.¹⁶ The indicator is split into two different parts:

1. *In-cash benefits* perceived by households with children (e.g., family allowances, parental leave coverage...)
2. *In-kind benefits* that ensure the running of public childcare services through direct financing (e.g., construction of infrastructures, staff hiring...) and public childcare support for families (e.g., financial aid for childcare use).

¹⁶OCDE (2023), « Social Expenditure : Aggregated data », OECD Social and Welfare Statistics (base de données), <https://doi.org/10.1787/data-00166-en>.

For each type of family benefits spending, we use the share of Gross Domestic Product (GDP) allocated to such expenditures as a measure for generosity in family benefits. More details on the construction on these variable are given in Section F.2.2. The distribution of family benefits shares over time in each country is provided in Figure F3.

A priori, pre-school enrollment is expected to have a stronger effect on mothers' labour supply, if any, in countries where the share of GDP allocated to in-kind family benefits is low since female employment is positively correlated to public investment in early childcare [Olivetti and Petrongolo, 2017]. When spending on in-kind transfers is high, pre-school substitutes to earlier formal childcare arrangements, implying no labour supply effect since mothers managed to resume work before pre-school start. However, the effect of pre-school availability on maternal labour supply according to countries' generosity in terms of in-cash family benefits is less straightforward. Such benefits have been shown detrimental to female labour market participation as they constitute inactivity traps for mothers [Schöne, 2004; Piketty, 2005; González, 2013]. Thus, pre-school enrollment could help mothers who withdrew from the labour market due to these disincentives going back to the labour market. On another hand, pre-school enrollment would have no effect if mothers are pushed out of the labour market on the long-run by financial transfers that encourage them to stay home to take care of children.

The level of in-cash family benefits does not yield any significant heterogeneous effect of pre-school enrollment on maternal labour supply (Table F3). Pre-school enrollment decreases informal childcare take-up the same way in both country types (-9 ppt, $p < 0.1$) and decreases conditional weekly hours only in the least generous countries (-9 hours, $p < 0.01$) (Table F4). The absence of a substitution effect in countries where the in-cash family benefits share is higher is in line with the fact that mothers use less informal childcare hours in these countries on average (-5 weekly hours, $p < 0.05$).

In countries with a higher level of in-kind family benefits, mothers work more (+20 ppt) and are more active (+21 ppt) on average (see Table F5). Regarding causal effects, pre-school enrollment increases mothers' probability to be employed (resp. to be active) by 27 ppt (resp. 30 ppt) ($p < 0.01$) in the least generous countries. In contrast, pre-school enrollment has no significant effect on maternal labour supply in countries with a higher share of GDP spent on in-kind family benefits. In the most generous countries, pre-school is therefore not a leverage for

maternal labour supply, probably because mothers living in such countries were able to resume work before pre-school. Table F6 shows that the decrease in informal childcare use in response to pre-school enrollment is the same irrespective of the level of in-kind family benefits (-16 ppt in take-up, $p < 0.05$ and -8 weekly hours $p < 0.01$).

Net cost of early childcare. The OECD also provides information on net childcare costs at the country level.¹⁷ Costs are computed among parents using centre-based childcare and are net of all family benefits. Because childcare costs are expected to be higher in countries with higher income levels, we consider net childcare cost as a share of household disposable income in order to correct for disparities in living standards between countries. Further details concerning this variable are provided in Section F.2.3.

Childcare costs are of interest in our case because they are expected to drive mothers' labour supply. When childcare is not affordable, mothers might reduce their working hours or stop participating to the labour market to provide childcare at home. Previous literature showing that reducing childcare costs foster maternal labour supply [Bettendorf et al., 2015; Huebener et al., 2020] supports this hypothesis. Because public pre-school programs are free in almost all cases, one can expect pre-school to lever maternal labour supply in countries where early childcare costs are higher. On another hand, mothers might work more to increase their income in order to afford formal childcare solutions. Thus, the expected result is not straightforward.

Table F7 shows that pre-school does not increase mothers' labour supply irrespective of the weight early childcare costs represent in households' disposable incomes. The absence of leverage in countries where childcare costs are high could somewhat be explained by the positive correlation between maternal labour market participation and net childcare costs effort rates (+9 ppt for employment and +8 ppt for activity, $p < 0.01$). Differences in net childcare cost effort rates between countries do not yield any heterogeneity in substitution effects between pre-school and informal childcare use (Table F8).

¹⁷OCDE (2023), « Benefits and wages : Net childcare cost for parents using childcare », OECD Social and Welfare Statistics (base de données), <https://doi.org/10.1787/b0781729-en>

5.3.3 Analysis by country

We implement an analysis by country to explore which countries drive our findings. Results, displayed in Appendix H, show that there is a strong effect on mothers' employment rate in Slovakia and in the Czech Republic. In the other countries, there is no significant effect on the extensive margin of labour supply and no heterogeneous effects are found for other outcomes. This result on employment rate is in line with our heterogeneous results when pooling all countries together. Indeed, these countries have a very low use of formal childcare (lower than 1%), are among those with the most conservative gender norms and have a rather low share of in-kind family benefits in their public spending. In countries that share similar characteristics (Portugal and Hungary), the effect is also positive but not significant. Overall, this suggests that our heterogeneous results are driven by Slovakia and the Czech Republic. Public childcare provision effort should therefore be concentrated in these countries.

6 Conclusion

The purpose of this paper is to evaluate the impact of a child's pre-school enrollment on her mother's labour supply in a cross-country perspective. The underlying hypothesis is that pre-school might be a leverage for mothers whose access to earlier childcare solutions is limited, due to a lack of availability (i.e., shortage in day-care slots) or affordability (i.e., low willingness to pay, low public subsidies...). Pre-school appears as a convenient alternative since it is provided free of charge with a legal entitlement for eligible children in most cases. Moreover, pre-school programs are part of the education system, which offers an opportunity for child development as a preparation for elementary school. On another hand, pre-school might not have any effect if early childcare solutions help mothers resuming work before pre-school enrollment.

In order to correct for the endogenous relationship between pre-school enrollment decision and mother's labour supply, we make use of the discontinuous relationship between pre-school eligibility and the child's year of birth in a fuzzy regression on discontinuity design. We measure the distance to pre-school eligibility in quarters of birth in a bandwidth of one year around the eligibility cut-off. We use data from the EU-SILC from 2004 to 2020 which allows us to conduct country level heterogeneity analyses.

Pooling all countries together, we do not find any effect of pre-school enrollment on mothers' labour supply. We then explore heterogeneous effects at the individual level. We find that pre-school increases the intensive margin of labour supply for the most educated mothers - who are more likely to adjust their working hours - and decreases it for the least educated ones - which could be explained by an income effect. This result suggests that efforts still need to be done with respect to early childcare provision since mothers with lower living standards need to work more to pay for formal early childcare. We also find that pre-school helps mothers resuming work when the child who is enrolled is the last born in the sibship.

At the country level, we investigate potential differences in the effect of pre-school enrollment according to gender norms, public spending in family benefits and the net cost of childcare. First, we find that pre-school is a leverage for maternal labour supply in the most conservative countries, where female labour market participation is lower on average. Family benefits can be transferred in cash (through family allowances or parental leave coverage for example) and in kind (through direct public investment in childcare services). Since expected mechanisms are different between those two types of transfers, we conduct separate analyses for both of them. We find that in-cash benefits do not yield any heterogeneous effect of pre-school on mothers' labour supply. Following previous studies showing that in-cash family benefits keep mothers out of the labour market, one could expect a positive effect of pre-school enrollment on maternal labour market participation in countries with high levels of in-cash transfers. Our results are not in line with this hypothesis. Our results also indicate that pre-school fosters labour market participation in countries with low levels of in-kind family benefits. This suggests that investment in early childcare provision helps mothers resuming work before their child reaches pre-school age.

Several limitations to our study have to be highlighted. First, we do not observe the child's exact date of birth which would allow us to compare individuals as similar as possible. Nevertheless, we show that standard RDD hypotheses are still fulfilled in our setting, which suggests that the internal validity of our results should not be impaired. Moreover, even if EU-SILC data only provides the child's quarter of birth, it is -to the best of our knowledge- the most appropriate dataset we could use to explore the effect of pre-school on maternal labour supply in a European comparison perspective. Second, the external validity is limited because the effects we estimate

are local. The first reason is that they are estimated only for mothers around the threshold, i.e. mothers of pre-school age children. Results therefore cannot be generalized to women who have children of other ages. However, it is important to keep in mind that pre-school programs always focus on children of the same age range, with only slight differences between countries (between 3 to 5 years old). Therefore, our results could not be generalized to mothers with children of other ages by definition because we focus on pre-school enrollment. The effects are local also because they are estimated among compliers, i.e. mothers who react to their child's eligibility to pre-school in terms of enrollment decision. We borrow the methodology proposed by [Marbach and Hangartner \[2020\]](#) to compare compliers to non-compliers with respect to observable characteristics. Their method is briefly described in Section G and results show that compliers do not statistically differ from non-compliers when never-takers and always-takers are pooled together as non-compliers. Figure G1 nevertheless shows that compliers are younger than always-takers but older than never-takers by about 6 months, they are more (resp. less) educated than never-takers (resp. always-takers), more likely to be single mothers than always-takers and they have less (resp. more) children than never-takers (resp. always-takers). If this profiling analysis suggests that our results are specific to complying mothers, differences therefore remain modest in terms of amplitude.

Bibliography

- Alesina, A., Algan, Y., Cahuc, P., and Giuliano, P. (2015). Family values and the regulation of labor. *Journal of the European Economic Association*, 13(4):599–630. [126]
- Andresen, M. E. and Havnes, T. (2019). Child care, parental labor supply and tax revenue. *Labour Economics*, 61:101762. [7, 102]
- Antecol, H. (2000). An examination of cross-country differences in the gender gap in labor force participation rates. *Labour Economics*, 7(4):409–426. [103, 125, 177, 191, 199]
- Aparicio Fenoll, A. (2020). The uneven impact of women’s retirement on their daughters’ employment. *Review of Economics of the Household*, 18(3):795–821. [119]
- Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745. [7, 102]
- Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers’ labor supply : Evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16. [8, 44, 103]
- Berlinski, S., Galiani, S., and Mc Ewan, P. J. (2011). Preschool and maternal labor market outcomes: Evidence from a regression discontinuity design. *Economic Development and Cultural Change*, 59(2):313–344. [8, 103]
- Bettendorf, L. J., Jongen, E. L., and Muller, P. (2015). Childcare subsidies and labour supply : Evidence from a large Dutch reform. *Labour Economics*, 36:112–123. [8, 129]
- Blau, F. D., Kahn, L. M., Liu, A. Y.-H., and Papps, K. L. (2013). The transmission of women’s fertility, human capital and work orientation across immigrant generations. *Journal of Population Economics*, 26:405–435. [16, 103, 177, 179, 199]
- Bonsang, E., Skirbekk, V., and Staudinger, U. M. (2017). As you sow, so shall you reap: Gender-role attitudes and late-life cognition. *Psychological Science*, 28(9):1201–1213. [126, 176, 185]
- Bratti, M., Frattini, T., and Scervini, F. (2018). Grandparental availability for child care and maternal labor force participation: Pension reform evidence from Italy. *Journal of Population Economics*, 31:1239–1277. [119]

- Brilli, Y., Del Boca, D., and Pronzato, C. D. (2016). Does child care availability play a role in maternal employment and children’s development? Evidence from Italy. *Review of Economics of the Household*, 14:27–51. [7, 102]
- Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029. [103, 125]
- Busse, A. and Gathmann, C. (2020). Free daycare policies, family choices and child development. *Journal of Economic Behavior & Organization*, 179:240–260. [7, 103]
- Carta, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children’s outcomes: Evidence from Italy. *Journal of Public Economics*, 158:79–102. [8, 103]
- Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources*, 44(1):140–170. [8, 103, 124]
- Dujardin, C., Fonder, M., and Lejeune, B. (2018). Does formal child care availability for 0-3 year olds boost mothers’ employment rate? Panel data based evidence from Belgium. *Annals of Economics and Statistics/Annales d’Économie et de Statistique*, (129):103–126. [7, 102]
- European Commission (2017). *European pillar of social rights*. Publications Office. [105]
- Eurydice (2019). Key data on early childhood education and care in europe – 2019 edition. Technical report, European Commission/EACEA/Eurydice. Eurydice Report. Luxembourg: Publications Office of the European Union. [102, 105]
- Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177. [16, 103, 125, 177, 186, 191, 199]
- Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics*, 28(1):51–85. [124]
- Fitzpatrick, M. D. (2012). Revising our thinking about the relationship between maternal labor supply and preschool. *Journal of Human Resources*, 47(3):583–612. [8, 103]
- Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3):416–438. [126, 185]

- Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, 92(1):307–322. [8, 103]
- Givord, P. and Marbot, C. (2015). Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies. *Labour Economics*, 36:99–111. [8, 103]
- González, L. (2013). The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply. *American Economic Journal: Economic Policy*, 5(3):160–188. [103, 128]
- Goux, D. and Maurin, E. (2010). Public school availability for two-year olds and mothers' labour supply. *Labour Economics*, 17(6):951–962. [8, 103, 124]
- Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica*, 69(1):201–209. [110]
- Havnes, T. and Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics*, 95(11-12):1455–1465. [7, 102]
- Huebener, M., Pape, A., and Spiess, C. K. (2020). Parental labour supply responses to the abolition of day care fees. *Journal of Economic Behavior & Organization*, 180:510–543. [7, 103, 129]
- Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635. [113]
- Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in Economics. *Journal of Economic Literature*, 48(2):281–355. [112, 113]
- Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, 26(3):519–548. [7, 102]
- Marbach, M. and Hangartner, D. (2020). Profiling compliers and noncompliers for instrumental-variable analysis. *Political Analysis*, 28(3):435–444. [73, 93, 132, 166]
- Müller, K.-U. and Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare. *Labour Economics*,

- 62:101776. [7, 44, 102]
- Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36:124–136. [8, 103]
- OECD (2012). *Closing the gender gap: Act now*. OECD Publishing Paris. [102]
- Olivetti, C. and Petrongolo, B. (2017). The economic consequences of family policies: Lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, 31(1):205–230. [103, 128]
- Piketty, T. (2005). Impact de l’allocation parentale d’éducation sur l’activité féminine. *Histoires de familles, histoires familiales: Les résultats de l’enquête famille de 1999*, (156):79. [103, 128]
- Pinto, M. F. (2023). Stay at home with grandma, mom is going to work: The impact of grandmothers’ retirement on mothers’ labor decisions. *Economic Development and Cultural Change*, 72(1):283–327. [119]
- Pora, P. et al. (2020). *Keep working and spend less?: Collective childcare and parental earnings in France*. INSEE, Institut national de la statistique et des études économiques. [7, 102]
- Ravazzini, L. (2018). Childcare and maternal part-time employment: A natural experiment using Swiss cantons. *Swiss Journal of Economics and Statistics*, 154:1–16. [7, 102]
- Schøne, P. (2004). Labour supply effects of a cash-for-care subsidy. *Journal of Population Economics*, 17:703–727. [128]
- Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529. [51, 111]

Appendix

A Childcare systems in Europe

Table A1: Early childcare arrangements and pre-school characteristics by country

Country	Early childcare				Pre-school			
	Eligibility	Availability	Free	Legal entitlement	Eligibility	Availability	Free	Legal entitlement
Cyprus	From birth to compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the 1st of September	Half or full-day	No	No
Czech Rep.	From 6 months old until compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the 1st of September. Available from 2 years old but not guaranteed	Full-day	No	Yes
France	From 2 months to compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the 1st of September. Compulsory from 2019. Available from 2 years old but not guaranteed.	Half or full-day	Yes	No
Greece	From 20 weeks to compulsory schooling	Full-day	No	No	Available for children reaching the age of 4 by the end of the year from 2006 to 2019. Compulsory from 2020	Half or full-day	Yes	Yes
Hungary	From 20 weeks to compulsory schooling	Full-day	No	No	Compulsory for children reaching the age of 3 by the 31st of August	Full-day	Free	Compulsory
Luxembourg	From birth until compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the 1st of September since 2009. Compulsory from 4 years old	Max. 8 half-days per week	Yes	Yes
Poland	From 20 weeks to compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the end of the year since 2017. In 2015-2016 : 4 years old, in 2009-2014 : 5 years old.	Half or full day	Yes	Yes
Portugal	From 3 months to compulsory schooling	Full-day	No	No	Available for children reaching the age of 3 before the end of school year can be registered throughout the year	Full-day	Yes	Yes
Slovakia	From 6 months of age until 3 years old	Full-day	No	No	Available for children reaching the age of 3 before the end of school year can be registered throughout the year. If the kindergarten's capacity allows, 2-year-old children can also be admitted.	Half or full-day	No	No
Spain	From 3 months of age until 3 years old	No	No	In some autonomous communities, yes.	Available for children reaching the age of 3 before the end of school year	Half or full-day	Yes	Yes

B Descriptive statistics by country

Table B1: Descriptive statistics by country

	CY	CZ	FR	GR	HU	LU	PL	PT	SK	ES
Labour supply										
Employment rate (%)	0.699	0.332	0.671	0.523	0.432	0.635	0.608	0.775	0.466	0.607
Activity rate (%)	0.786	0.392	0.747	0.663	0.486	0.707	0.645	0.882	0.551	0.738
Conditional weekly working hours	36.094	36.240	33.699	36.083	37.628	32.610	37.778	38.763	38.799	34.090
	(7.812)	(8.351)	(9.342)	(10.707)	(7.319)	(11.497)	(7.984)	(7.207)	(6.191)	(9.748)
Working part-time (%)	0.261	0.255	0.386	0.328	0.184	0.476	0.175	0.091	0.099	0.370
Pre-school										
Pre-school enrollment rate (%)	0.377	0.417	0.605	0.653	0.442	0.245	0.390	0.385	0.464	0.823
Formal childcare										
Daycare take-up (%)	0.306	0.003	0.305	0.015	0.215	0.417	0.116	0.343	0.009	0.021
Childminder care take-up (%)	0.043	0.015	0.109	0.035	0.013	0.148	0.037	0.061	0.007	0.032
Informal childcare										
Informal care take-up (%)	0.402	0.430	0.189	0.415	0.394	0.246	0.373	0.325	0.278	0.124
Individual characteristics										
Age	34.239	33.463	34.272	35.486	33.364	34.323	33.579	35.720	33.514	35.790
	(4.999)	(4.722)	(5.095)	(5.246)	(5.531)	(5.252)	(5.011)	(5.147)	(4.956)	(4.841)
Single mother (%)	0.042	0.091	0.085	0.015	0.078	0.054	0.052	0.077	0.030	0.037
Completed high school at least (%)	0.546	0.292	0.447	0.361	0.333	0.367	0.508	0.370	0.379	0.458
Number of children	2.101	1.930	2.148	1.994	2.137	2.132	2.062	1.824	2.076	1.843
	(0.911)	(0.793)	(0.935)	(0.872)	(1.086)	(0.975)	(0.954)	(0.814)	(1.089)	(0.779)
Number of individuals	2,115	4,364	7,532	4,162	2,625	2,230	2,481	2,472	2,320	6,318
Number of observations	2,385	5,242	9,003	4,800	2,775	3,634	2,656	2,570	2,431	6,847

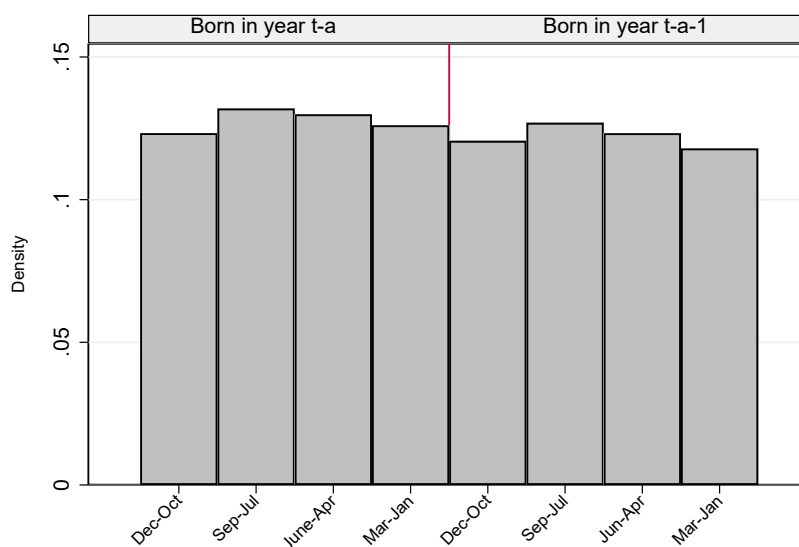
Note: Means with standard deviations in parentheses. CY = Cyprus, CZ = Czech Republic, FR = France, GR = Greece, HU = Hungary, LU = Luxembourg, PL = Poland, PT = Portugal, SK = Slovakia, ES = Spain. This table is commented in Section 4.

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

C Identification tests

This section reports elements that support the identification strategy presented in Section 3.

Figure C1: Distribution of birth quarters



Note : Grey bars represent birth frequencies by quarter of birth around the pre-school eligibility cut-off (born in year t-a-1).

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

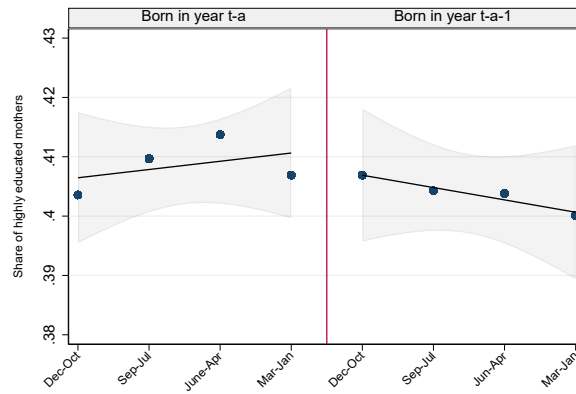
Table C1: Continuity of individual characteristics

	All mothers		
	Highly educated (1)	Single (2)	Number of children (3)
Reduced-form			
Pre-school eligibility	-0.0067 (0.0124)	-0.0067 (0.0059)	0.0159 (0.0243)
Second-stage			
Pre-school enrollment	-0.0215 (0.0399)	-0.0215 (0.0188)	0.0509 (0.0780)
Country specific effects	yes	yes	yes
Year specific effects	yes	yes	yes
Individual controls	no	no	no
Weak instrument test (F-stat)	924.924	924.924	924.924
Number of individuals	33,660	33,660	33,660
Number of observations	37,349	37,349	37,349

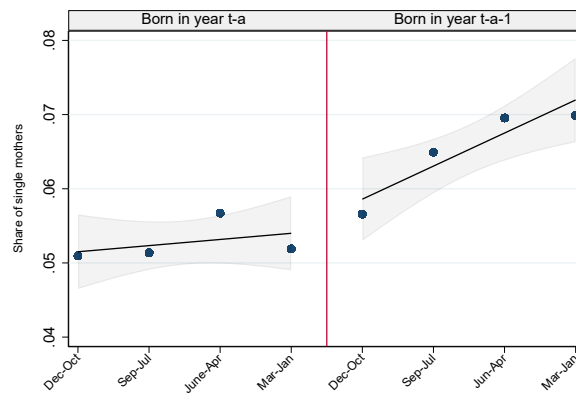
Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

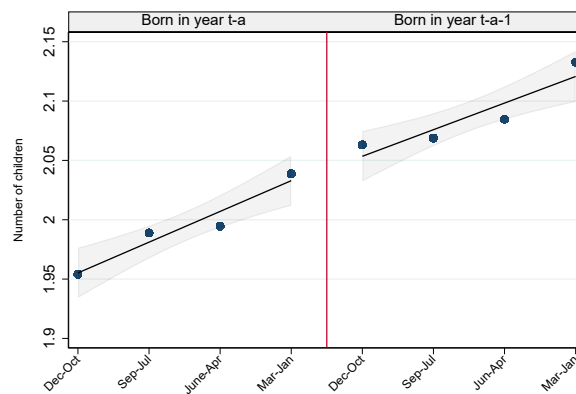
Figure C2: Continuity of individual characteristics at the eligibility threshold



(a) Education level



(b) Probability to be single



(c) Number of children

Note : Graphical representation of the fuzzy RDD reduced-form regression. Dots are the observed maternal (a) probability to be highly educated, (b) probability to be single and (c) average number of children at time t for each quarter of birth around the pre-school eligibility threshold. The cut-off which determines eligibility is shown by the vertical red line (born in year $t-a-1$). Linear fitting lines are represented by the solid black lines on each side of the cut-off with confidence intervals at the level of 95% as grey shaded areas.

Source: European Statistics on Income and Living Conditions (EU-SILC), 2004-2020.

D Robustness checks

Table D1: Effect of pre-school enrollment on mothers' labour supply with individual covariates

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.3106*** (0.0103)	0.3106*** (0.0103)	0.3325*** (0.0123)	0.3325*** (0.0123)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Reduced-form				
Pre-school eligibility	0.0115 (0.0116)	0.0161 (0.0111)	0.2521 (0.2988)	-0.0106 (0.0147)
Second-stage				
Pre-school enrollment	0.0369 (0.0373)	0.0519 (0.0354)	0.7817 (0.9258)	-0.0327 (0.0455)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes
Weak instrument test (F-stat)	929.252	929.252	682.677	682.677
Number of individuals	33,660	33,660	19,681	19,681
Number of observations	37,349	37,349	21,276	21,276

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Covariates are presented in Section 4.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table D2: Effect of pre-school enrollment on informal childcare use with individual covariates

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.3106*** (0.0102)	0.3106*** (0.0102)
	Informal childcare take-up	Formal take-up
	(1)	(2)
Reduced-form		
Pre-school eligibility	-0.0266** (0.0109)	-0.1299*** (0.0088)
Second-stage		
Pre-school enrollment	-0.0856** (0.0352)	-0.4183*** (0.0279)
Country specific effect	yes	yes
Year specific effect	yes	yes
Individual controls	yes	yes
Weak instrument test (F-stat)	929.252	929.252
Number of individuals	33,660	33,660
Number of observations	37,349	37,349

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Covariates are presented in Section 4.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table D3: Placebo test : Family allowances

	All mothers		Working mothers	
	Family allowance reception	Conditional amount	Family allowance reception	Conditional amount
Reduced-form				
Pre-school eligibility	-0.0054 (0.0096)	60.4688 (124.2516)	-0.0108 (0.0133)	66.1604 (162.9668)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Number of individuals	29,980	21,128	18,027	11,741
Number of observations	33,485	23,570	19,577	12,768

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. This analysis is a check for the exclusion restriction assumption (see Section 3).

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

E Heterogeneous effects with respect to individual characteristics

E.1 Methodology

We describe here the methodology used for heterogeneity results presented in Section 5.2. For each source of heterogeneity, we implement the same RDD strategy we use for average results, with additional interaction terms as follows:

$$D_{ict} = \gamma_0 + \gamma_1 E_{ict} + \gamma_2 I_{ict} + \gamma_3 E_{ict} * I_{ict} + \gamma_4 f(q_{ict}) + \gamma_5 E_{ict} * f(q_{ict}) + \gamma_6 f(q_{ict}) * I_{ict} + \gamma_7 E_{ict} * f(q_{ict}) * I_{ict} + \theta_c + \lambda_t + \epsilon_{ict} \quad (5)$$

$$Y_{ict} = \delta_0 + \delta_1 \widehat{D}_{ict} + \delta_2 I_{ict} + \delta_3 \widehat{D}_{ict} * I_{ict} + \delta_4 f(q_{ict}) + \delta_5 E_{ict} * f(q_{ict}) + \delta_6 f(q_{ict}) * I_{ict} + \delta_7 E_{ict} * f(q_{ict}) * I_{ict} + \phi_c + \rho_t + v_{ict} \quad (6)$$

I_{ict} is the binary variable defining the mother's i source of heterogeneity (family size or education) interviewed in country c at time t . In the first-stage equation (Equation 5), this variable is interacted with the eligibility status to pre-school E_{ict} , the linear function of the distance to the eligibility cut-off $f(q_{ict})$ and with the trends on both sides of the cut-off $E_{ict} * f(q_{ict})$ to allow their relationship with pre-school enrollment rate D_{ict} to differ according to the mother's characteristic. θ_c and λ_t denote country and time specific effects and ϵ_{ict} the unobserved error term. In the second-stage equation (Equation 6), I_{ict} is interacted with the fitted values of the variable of interest \widehat{D}_{ict} in order to assess whether the effect of pre-school enrollment on the mother's labour supply outcome Y_{ict} is heterogeneous with respect to the characteristic considered (family size or education). As in the first-stage equation, I_{ict} is also interacted with $f(q_{ict})$ and E_{ict} . Here, country and time specific effects are defined by ϕ_c and ρ_t respectively, and the unobserved error term by v_{ict} .

E.2 Results

We report here the detailed results of the heterogeneity analysis at the individual level presented in Section 5.2.

Table E1: Heterogeneous effects on mothers' labour supply: Education level

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.2983*** (0.0136)	0.2983*** (0.0136)	0.2952*** (0.0182)	0.2952*** (0.0182)
High education	0.0570*** (0.0106)	0.0570*** (0.0106)	0.0155 (0.0135)	0.0155 (0.0135)
Pre-school eligibility x High education	0.0326 (0.0206)	0.0326 (0.0206)	0.0532** (0.0248)	0.0532** (0.0248)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	0.0557 (0.0563)	0.0781 (0.0547)	-2.6071* (1.5672)	0.0584 (0.0725)
High education ($\hat{\delta}_2$)	0.2513*** (0.0374)	0.1994*** (0.0351)	-3.0188*** (1.0634)	0.0836 (0.0510)
Pre-school enrollment x High education ($\hat{\delta}_3$)	-0.0214 (0.0791)	-0.0489 (0.0745)	6.5048*** (2.0310)	-0.1821* (0.0979)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.0343 (0.0525)	0.0292 (0.0476)	3.8976*** (1.2013)	-0.1237** (0.0619)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	298.322	298.322	155.797	155.797
Number of individuals	33,660	33,660	19,681	19,681
Number of observations	37,349	37,349	21,276	21,276

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.1.
Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table E2: Heterogeneous effects on informal childcare use: Education level

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.2983*** (0.0136)	0.2689*** (0.0293)
High education	0.0570*** (0.0106)	0.0164 (0.0192)
Pre-school eligibility x High education	0.0327 (0.0206)	0.1033** (0.0419)
	Informal childcare take-up (1)	Conditional weekly hours (2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1769*** (0.0487)	-3.3604 (3.1992)
High education ($\hat{\delta}_2$)	-0.0281 (0.0365)	0.3592 (1.6202)
Pre-school enrollment x High education ($\hat{\delta}_3$)	0.2012*** (0.0757)	-1.8823 (3.9708)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.0243 (0.0554)	-5.2427** (2.1559)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	298.322	37.146
Number of individuals	33,660	9,950
Number of observations	37,349	10,627

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.1.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table E3: Heterogeneous effects on mothers' labour supply: Partnership status

	All mothers		Working mothers	
	Pre-school enrollment			
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.3115*** (0.0105)	0.3115*** (0.0105)	0.3232*** (0.0127)	0.3232*** (0.0127)
Single mother	-0.0388* (0.0220)	-0.0387* (0.0220)	-0.0508* (0.0301)	-0.0508* (0.0301)
Pre-school eligibility x Single mother	0.0008 (0.0442)	0.0007 (0.0442)	-0.0045 (0.0565)	-0.0045 (0.0565)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	0.0450 (0.0405)	0.0534 (0.0381)	0.7037 (0.9596)	-0.0317 (0.0471)
Single mother ($\hat{\delta}_2$)	-0.0133 (0.0723)	0.0268 (0.0648)	-1.3187 (2.0513)	0.0746 (0.0996)
Pre-school enrollment x Single mother ($\hat{\delta}_3$)	-0.0737 (0.1855)	0.0194 (0.1638)	3.9438 (4.3458)	-0.1729 (0.2159)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.0287 (0.1803)	0.0727 (0.1587)	4.6474 (4.2210)	-0.2045 (0.2099)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	22.133	22.133	13.524	13.524
Number of individuals	33,660	33,660	19,089	19,089
Number of observations	37,349	37,349	20,669	20,669

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table E4: Heterogeneous effects on informal childcare use: Partnership status

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.3115*** (0.0105)	X X
Single mother	-0.0388* (0.0220)	X X
Pre-school eligibility x Single mother	0.0008 (0.0442)	X X
	Informal childcare take-up (1)	Conditional weekly hours (2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1010*** (0.0365)	X X
Single mother ($\hat{\delta}_2$)	-0.0159 (0.0669)	X X
Pre-school enrollment x Single mother ($\hat{\delta}_3$)	0.1994 (0.1690)	X X
$(\hat{\delta}_1 + \hat{\delta}_3)$	0.0984 (0.1644)	X X
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	22.133	2.858
Number of individuals	33,660	10,194
Number of observations	37,349	10,877

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table E5: Heterogeneous effects on mothers' labour supply: Birth order

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.2794*** (0.0213)	0.2794*** (0.0213)	0.2938*** (0.0299)	0.2938*** (0.0299)
Last child	-0.0192 (0.0130)	-0.0192 (0.0130)	-0.0252 (0.0185)	-0.0252 (0.0185)
Pre-school eligibility x Last child	0.0393 (0.0243)	0.0393 (0.0243)	0.0339 (0.0329)	0.0339 (0.0329)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	-0.0870 (0.0905)	-0.0485 (0.0886)	2.9751 (2.6965)	-0.0632 (0.1411)
Last child ($\hat{\delta}_2$)	0.0250 (0.0487)	0.0901* (0.0471)	1.9577 (1.6087)	-0.0371 (0.0829)
Pre-school enrollment x Last child ($\hat{\delta}_3$)	0.1706* (0.1021)	0.1421 (0.0986)	-2.4314 (2.9549)	0.0304 (0.1529)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.0836* (0.0436)	0.0936** (0.0398)	0.5437 (1.0230)	-0.0329 (0.0495)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	N. of children	N. of children	N. of children	N. of children
Weak instrument test (F-stat)	83.161	83.161	35.538	35.538
Number of individuals	33,660	33,660	19,089	19,089
Number of observations	37,349	37,349	20,669	20,669

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.3.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table E6: Heterogeneous effects on informal childcare use: Birth order

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.2922*** (0.0218)	0.3061*** (0.0470)
Last child	-0.0198 (0.0135)	0.0062 (0.0243)
Pre-school eligibility x Last child	0.0364 (0.0248)	0.0148 (0.0525)
	Informal childcare take-up	Conditional weekly hours
	(1)	(2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1134 (0.0786)	-0.9584 (3.9941)
Last child ($\hat{\delta}_2$)	0.0037 (0.0429)	4.1347** (1.8530)
Pre-school enrollment x Last child ($\hat{\delta}_3$)	0.0306 (0.0900)	-4.2754 (4.5590)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.0828** (0.0407)	-5.2338** (2.0300)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	N. of children	N. of children
Weak instrument test (F-stat)	83.161	38.937
Number of individuals	32,580	9,950
Number of observations	36,213	10,627

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.2.3.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

F Heterogeneous effects with respect to country characteristics

F.1 Methodology

We describe here the methodology used for heterogeneity results presented in Section 5.3. For each source of heterogeneity, we implement the same RDD strategy we use for average results with additional interaction terms as follows:

$$\begin{aligned} D_{ict} = & \zeta_0 + \zeta_1 E_{ict} + \zeta_2 C_{ct} + \zeta_3 E_{ict} * C_{ct} \\ & + \zeta_4 f(q_{ict}) + \zeta_5 E_{ict} * f(q_{ict}) + \zeta_6 f(q_{ict}) * C_{ct} + \zeta_7 E_{ict} * f(q_{ict}) * C_{ct} + \pi_c + \tau_t + e_{ict} \end{aligned} \quad (7)$$

$$\begin{aligned} Y_{ict} = & \eta_0 + \eta_1 \hat{D}_{ict} + \eta_2 C_{ct} + \eta_3 \hat{D}_{ict} * C_{ct} \\ & + \eta_4 f(q_{ict}) + \eta_5 E_{ict} * f(q_{ict}) + \eta_6 f(q_{ict}) * C_{ct} + \eta_7 E_{ict} * f(q_{ict}) * C_{ct} + \psi_c + \omega_t + u_{ict} \end{aligned} \quad (8)$$

C_{ct} is the variable defining the characteristic of country c at time t . For the specific case of gender norms, the indicator for the country level characteristic is denoted as C_c since the GNI does not vary with time but with cohorts only. The remainder of the specification is similar to what is done when we explore heterogeneity based on individual characteristics (see Equations 5 and 6).

F.2 Measures at the country level

This section aims at giving more details on the construction of country level indicators that we use as presented in Section 5.3.

F.2.1 Gender norms

Because gender norms are expected to vary across birth cohorts, we select individuals born in the same cohorts in the IVS as in our EU-SILC sample. Individuals are asked whether they agree or not with several conservative statements on gender roles. We focus on two different statements provided by the IVS to compute the GNI:

1. *"When jobs are scarce, men have more right to a job than women"*
2. *"Men should take as much responsibility than women for home and children"*

First of all, we calculate the proportion of individuals who agree or strongly agree (versus disagree or strongly disagree) with Statements 1 and 2, giving two composite gender norms indices which average values are reported for each country in Figure F1. Then, the GNI is calculated as the average between these two measures of gender norms (Figure F2). Heterogeneity results using the GNI are presented in Section 5.3.1.

F.2.2 Family benefits

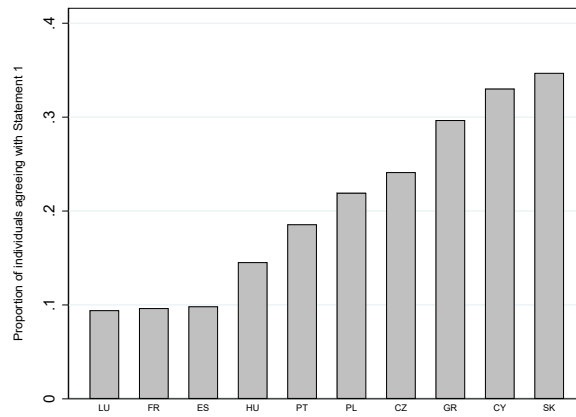
Data from the OECD is merged with EU-SILC data by country and year. As previously done for the GNI, we construct a binary variable taking 1 if the share of GDP spent on family benefits is higher than the median value in the sample (equal to 1.32% for in-cash and 0.57% for in-kind transfers). Figure F3 displays the evolution of public expenditures allocated to in-cash (Figure F3a) and to in-kind transfers (Figure F3b) from 2004 to 2020 for each country of the EU-SILC sample. Note that information on family benefits is not available for Cyprus in the OECD *Social protection* dataset. On the whole time period, Luxembourg is the country which spends the largest part of its GDP on family benefits (in cash) while Spain has the lowest share. With regard to in-kind benefits, France is the most generous over the period and Greece is the country which allocates the lowest share of GDP to such transfers. In Spain, France and Portugal, public expenditures as a share of GDP are higher for in-kind than for in-cash family

benefits. Heterogeneity results using generosity in family benefits are presented in Section [5.3.2](#).

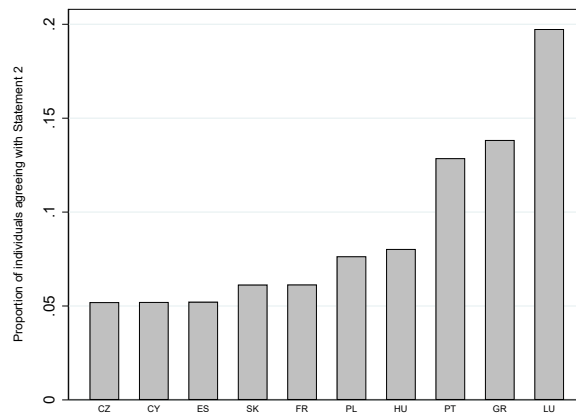
F.2.3 Net childcare costs

In line with what we do for family benefits, we use a binary variable taking 1 if the effort rate of net childcare costs is higher than the median value in the sample (equal to 8.7%) using data from the OECD. Figure [F3c](#) displays the evolution of net childcare costs effort rates for each country of the EU-SILC sample, except for Cyprus, in 2004, 2008, 2012, 2015, 2018, 2019 and 2020. Except from 2015, the effort rate is the highest in Czech Republic. In Greece, Hungary, France, Spain and Portugal, effort rates always remain lower than 10%, irrespective of the year of observation. Heterogeneity results using effort rates of net childcare costs are presented in Section [5.3.2](#).

Figure F1: Composite gender norms indices by country



(a) Proportion agreeing with *"When jobs are scarce, men have more right to a job than women"*

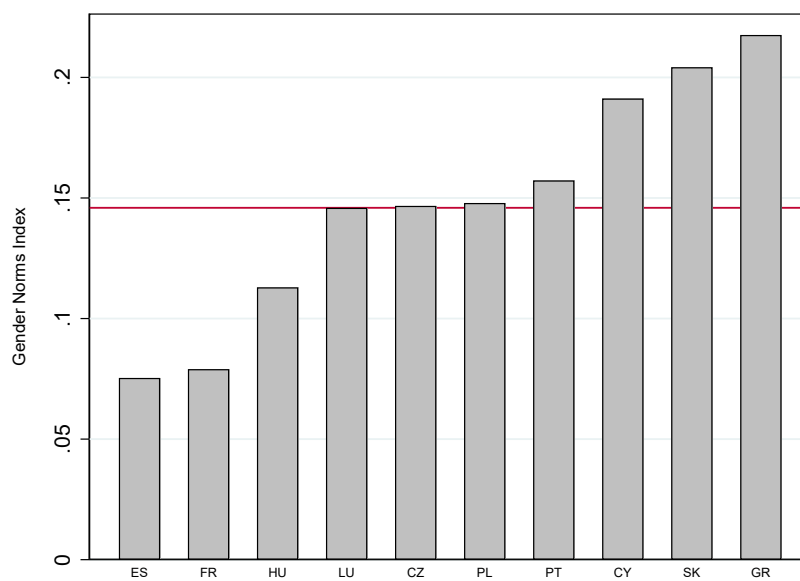


(b) Proportion agreeing more with *"A woman has to have children to be fulfilled"* than with *"A man has to have children to be fulfilled"*

Note : Both measures are used to compute the GNI as detailed in Section F.2.1.

Source : Integrated Values Study (IVS).

Figure F2: Gender norms index by country

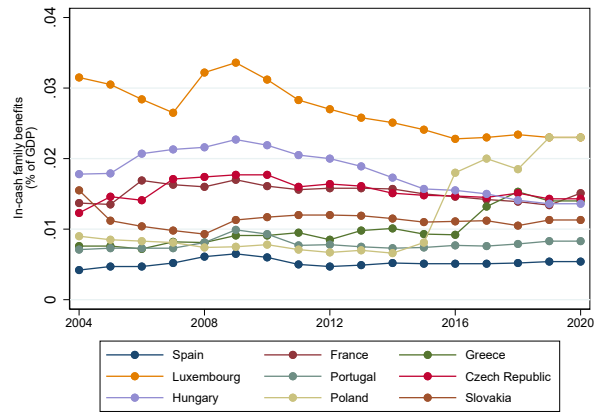


Note : Grey bars represent the average value of the Gender Norms Index by country included in the analysis. The horizontal red line indicates the GNI median value such that countries with an average GNI above the median are categorised as *more conservative* and the ones below the median are defined as *less conservative*. The construction of the GNI is detailed in Section F.2.1 and this figure is commented in Section 5.3.1.

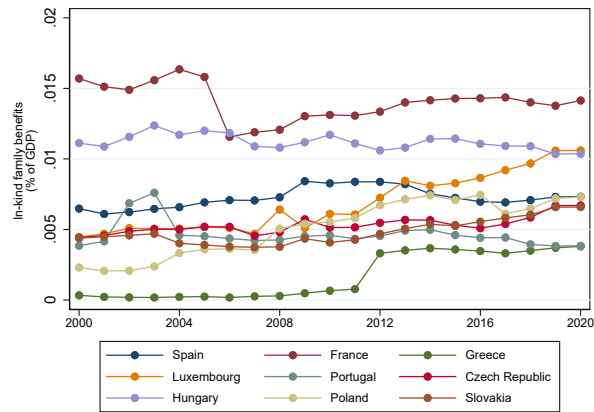
Sample : Countries included in the EU-SILC selected sample.

Source : Integrated Values Study (IVS).

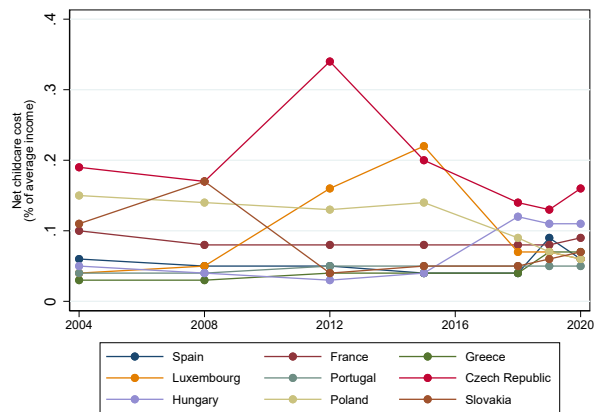
Figure F3: Childcare characteristics by country and year



(a) In-cash family benefits



(b) In-kind family benefits



(c) Net childcare cost

Note : Sub-figure (a) displays the evolution of *in-cash* family benefits spent by each country as a share of gross domestic product from 2004 to 2020. Sub-figure (b) exhibits the evolution of *in-kind* family benefits for the same time period and countries. The construction of these indicators is detailed in Section F.2.2 for family benefits and in Section F.2.3 for net childcare costs. Sample : Countries included in the EU-SILC selected sample.

Source : Organisation for Economic Co-operation and Development (OECD), 2004-2020.

F.3 Results

We report here the detailed results of the heterogeneity analysis at the country level presented in Section 5.3.

Table F1: Heterogeneous effects on mothers' labour supply: Gender norms

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.3255*** (0.0120)	0.3255*** (0.0120)	0.3067*** (0.0145)	0.3067*** (0.0145)
More conservative	-0.3937*** (0.0160)	-0.3937*** (0.0160)	-0.2772*** (0.0216)	-0.2772*** (0.0216)
Pre-school eligibility x More conservative	-0.0320 (0.0211)	-0.0320 (0.0211)	0.0359 (0.0256)	0.0359 (0.0256)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	-0.0349 (0.0516)	-0.0376 (0.0479)	1.0009 (1.3864)	-0.0383 (0.0689)
More conservative ($\hat{\delta}_2$)	-0.2443*** (0.0468)	-0.3086*** (0.0440)	4.4541*** (1.3784)	-0.2600*** (0.0676)
Pre-school enrollment x More conservative ($\hat{\delta}_3$)	0.1764** (0.0789)	0.2104*** (0.0741)	-0.1728 (1.8366)	-0.0045 (0.0900)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.1415** (0.0599)	0.1727*** (0.0567)	0.8282 (1.2050)	-0.0429 (0.0580)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	144.137	144.137	133.764	133.764
Number of individuals	33,660	33,660	19,681	19,681
Number of observations	37,349	37,349	21,276	21,276

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.1.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F2: Heterogeneous effects on informal childcare use: Gender norms

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.3255*** (0.0120)	0.3593*** (0.0295)
More conservative	-0.0394*** (0.0160)	-0.3219*** (0.0341)
Pre-school eligibility x More conservative	-0.0320 (0.0211)	-0.0641 (0.0410)
	Informal childcare take-up	Conditional weekly hours
	(1)	(2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1106*** (0.0409)	-3.7786 (2.3967)
More conservative ($\hat{\delta}_2$)	0.0938** (0.0403)	-1.4541 (3.4513)
Pre-school enrollment x More conservative ($\hat{\delta}_3$)	0.0453 (0.0743)	-1.6998 (1.9837)
$(\hat{\delta}_1 + \hat{\delta}_3)$	-0.0653 (0.0622)	-5.2326** (2.5038)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	144.137	53.823
Number of individuals	33,660	10,194
Number of observations	37,349	10,877

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.1.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F3: Heterogeneous effects on mothers' labour supply: In-cash family benefits

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.3023*** (0.0141)	0.3023*** (0.0141)	0.2882*** (0.0171)	0.2882*** (0.0171)
High share	0.0306** (0.0139)	0.0306** (0.0139)	0.0054 (0.0186)	0.0054 (0.0186)
Pre-school eligibility x High share	-0.0148 (0.0209)	-0.0148 (0.0209)	0.0129 (0.0253)	0.0129 (0.0253)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	0.0673 (0.0596)	0.0772 (0.0546)	1.5861 (1.5271)	-0.0393 (0.0728)
High share ($\hat{\delta}_2$)	0.0266 (0.0421)	0.0259 (0.0397)	0.3853 (1.1479)	-0.0040 (0.0560)
Pre-school enrollment x High share ($\hat{\delta}_3$)	-0.0259 (0.0863)	-0.0230 (0.0812)	-0.7558 (2.1528)	-0.0155 (0.1047)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.0415 (0.0623)	0.0542 (0.0600)	0.8303 (1.5095)	-0.0548 (0.0749)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	234.882	234.882	141.483	141.483
Number of individuals	31,790	31,790	18,339	18,339
Number of observations	35,257	35,257	19,815	19,815

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F4: Heterogeneous effects on informal childcare use: In-cash family benefits

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.3023*** (0.0141)	0.3367*** (0.0324)
High share	0.0306** (0.0139)	0.1087*** (0.0261)
Pre-school eligibility x High share	-0.0148 (0.0209)	-0.0914** (0.0439)
	Informal childcare take-up	Conditional weekly hours
	(1)	(2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.0915* (0.0504)	-8.8823*** (2.7397)
High share ($\hat{\delta}_2$)	0.0734* (0.0382)	-5.0191** (1.9732)
Pre-school enrollment x High share ($\hat{\delta}_3$)	-0.0394 (0.0771)	7.9332* (4.0847)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.1309** (0.0581)	-0.9492 (3.0403)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	234.882	34.800
Number of individuals	31,790	9,394
Number of observations	35,257	10,032

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F5: Heterogeneous effects on mothers' labour supply: In-kind family benefits

	All mothers		Working mothers	
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.2292*** (0.0164)	0.2292*** (0.0164)	0.2356*** (0.0201)	0.2356*** (0.0201)
High share	0.0439*** (0.0133)	0.0439*** (0.0133)	0.0576*** (0.0164)	0.0576*** (0.0164)
Pre-school eligibility x High share	0.1261*** (0.0210)	0.1261*** (0.0210)	0.1065*** (0.0256)	0.1065*** (0.0256)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	0.2720*** (0.0775)	0.2994*** (0.0740)	1.5707 (1.9550)	-0.0977 (0.0897)
High share ($\hat{\delta}_2$)	0.2013*** (0.0425)	0.2070*** (0.0402)	1.7312 (1.1657)	-0.0832 (0.0550)
Pre-school enrollment x High share ($\hat{\delta}_3$)	-0.3433*** (0.0916)	-0.3682*** (0.0868)	-0.4191 (2.2995)	0.0730 (0.1089)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.0712 (0.0497)	-0.0689 (0.0464)	1.1516 (1.2453)	-0.0247 (0.0632)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	95.207	95.207	65.933	65.933
Number of individuals	31,790	31,790	18,339	18,339
Number of observations	35,257	35,257	19,815	19,815

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F6: Heterogeneous effects on informal childcare use: In-kind family benefits

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.2292*** (0.0164)	0.2350*** (0.0300)
High share	0.0439*** (0.0133)	0.0474* (0.0268)
Pre-school eligibility x High share	0.1261*** (0.0210)	0.1201*** (0.0430)
	Informal childcare take-up	Conditional weekly hours
	(1)	(2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1637** (0.0762)	-8.6323*** (3.2693)
High share ($\hat{\delta}_2$)	-0.0041 (0.0395)	-1.2650 (1.5748)
Pre-school enrollment x High share ($\hat{\delta}_3$)	0.0869 (0.0859)	6.7078 (4.1001)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.0768* (0.0407)	-1.9245 (2.5205)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	95.207	30.387
Number of individuals	31,790	9,394
Number of observations	35,257	10,032

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F7: Heterogeneous effects on mothers' labour supply: Net childcare cost effort rates

	All mothers		Working mothers	
	Pre-school enrollment			
	(1)	(2)	(3)	(4)
First-stage				
Pre-school eligibility	0.2537*** (0.0142)	0.2537*** (0.0142)	0.2173*** (0.0175)	0.2173*** (0.0175)
High cost	0.0785*** (0.0135)	0.0785*** (0.0135)	0.0388** (0.0177)	0.0388** (0.0177)
Pre-school eligibility x High cost	0.0802*** (0.0209)	0.0802*** (0.0209)	0.1539*** (0.0249)	0.1539*** (0.0249)
	Working (1)	Active (2)	Working hours (3)	Part-time (4)
Second-stage				
Pre-school enrollment ($\hat{\delta}_1$)	0.0752 (0.0696)	0.0935 (0.0642)	1.6989 (2.0505)	-0.0871 (0.0949)
High cost ($\hat{\delta}_2$)	0.0885*** (0.0278)	0.0800*** (0.0264)	0.5079 (0.8570)	-0.0368 (0.0402)
Pre-school enrollment x High cost ($\hat{\delta}_3$)	-0.0397 (0.0873)	-0.0526 (0.0817)	-0.7604 (2.3397)	0.0647 (0.1113)
($\hat{\delta}_1 + \hat{\delta}_3$)	0.0355 (0.0538)	0.0409 (0.0516)	0.9385 (1.1737)	-0.0224 (0.0603)
Country specific effects	yes	yes	yes	yes
Year specific effects	yes	yes	yes	yes
Individual controls	no	no	no	no
Weak instrument test (F-stat)	158.339	158.339	76.183	76.183
Number of individuals	31,790	31,790	18,339	18,339
Number of observations	35,257	35,257	19,815	19,815

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.
Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Table F8: Heterogeneous effects on informal childcare use: Net childcare cost effort rates

	All mothers	
	Pre-school enrollment	
	(1)	(2)
First-stage		
Pre-school eligibility	0.2537*** (0.0142)	0.2911*** (0.0309)
High cost	0.0681*** (0.0135)	0.1012*** (0.0251)
Pre-school eligibility x High cost	0.0802*** (0.0209)	-0.0147 (0.0438)
	Informal childcare take-up (1)	Conditional weekly hours (2)
Second-stage		
Pre-school enrollment ($\hat{\delta}_1$)	-0.1040* (0.0605)	-8.1924*** (2.8770)
High cost ($\hat{\delta}_2$)	0.0009 (0.0255)	-0.8583 (1.3364)
Pre-school enrollment x High cost ($\hat{\delta}_3$)	-0.0111 (0.0774)	6.3144 (3.9755)
($\hat{\delta}_1 + \hat{\delta}_3$)	-0.1150** (0.0492)	-1.8780 (2.8424)
Country specific effects	yes	yes
Year specific effects	yes	yes
Individual controls	no	no
Weak instrument test (F-stat)	158.339	83.213
Number of individuals	31,790	9,394
Number of observations	35,257	10,032

Note: IV-TSLS estimates computed with a fuzzy RDD. Standard errors corrected for clustering at the individual level in parentheses. *p<0.1, **p<0.05, ***p<0.01. Results are commented in Section 5.3.2.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

G Complier analysis

G.1 Methodology

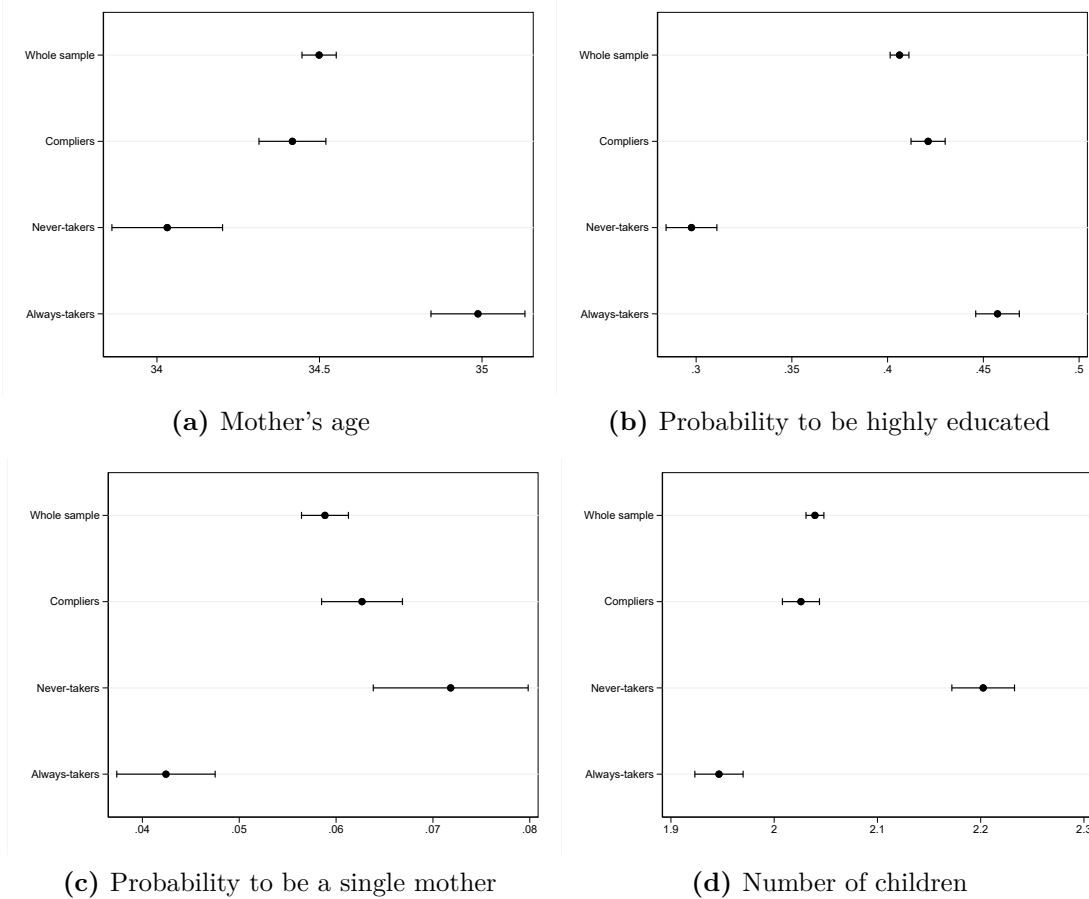
Profiling compliers can be useful to know if the Local Average Treatment Effect (LATE) can be generalized to the Average Treatment Effect (ATE). Under the hypotheses of independance and monotonicity (see Section 3), the characteristics of compliers and non-compliers can be compared by subtracting the weighted observable characteristic mean of non-compliers (i.e. always-takers and never-takers) to the one computed in the entire sample [Marbach and Hangartner, 2020] such as :

$$\hat{\mu}_{co} = \frac{1}{\hat{\pi}_{co}} * \hat{\mu} - \frac{\hat{\pi}_{nt}}{\hat{\pi}_{co}} * \hat{\mu}_{nt} - \frac{\hat{\pi}_{at}}{\hat{\pi}_{co}} * \hat{\mu}_{at} \quad (9)$$

with $\hat{\mu}_{co}$, $\hat{\mu}$, $\hat{\mu}_{nt}$ and $\hat{\mu}_{at}$ the mean value of a given observable characteristic for compliers, the entire sample, never-takers and always-takers respectively and $\hat{\pi}_{co}$, $\hat{\pi}_{nt}$ and $\hat{\pi}_{at}$ the estimated sample share of compliers, never-takers and always-takers respectively.

G.2 Results

Figure G1: Profile of compliers, never-takers and always-takers

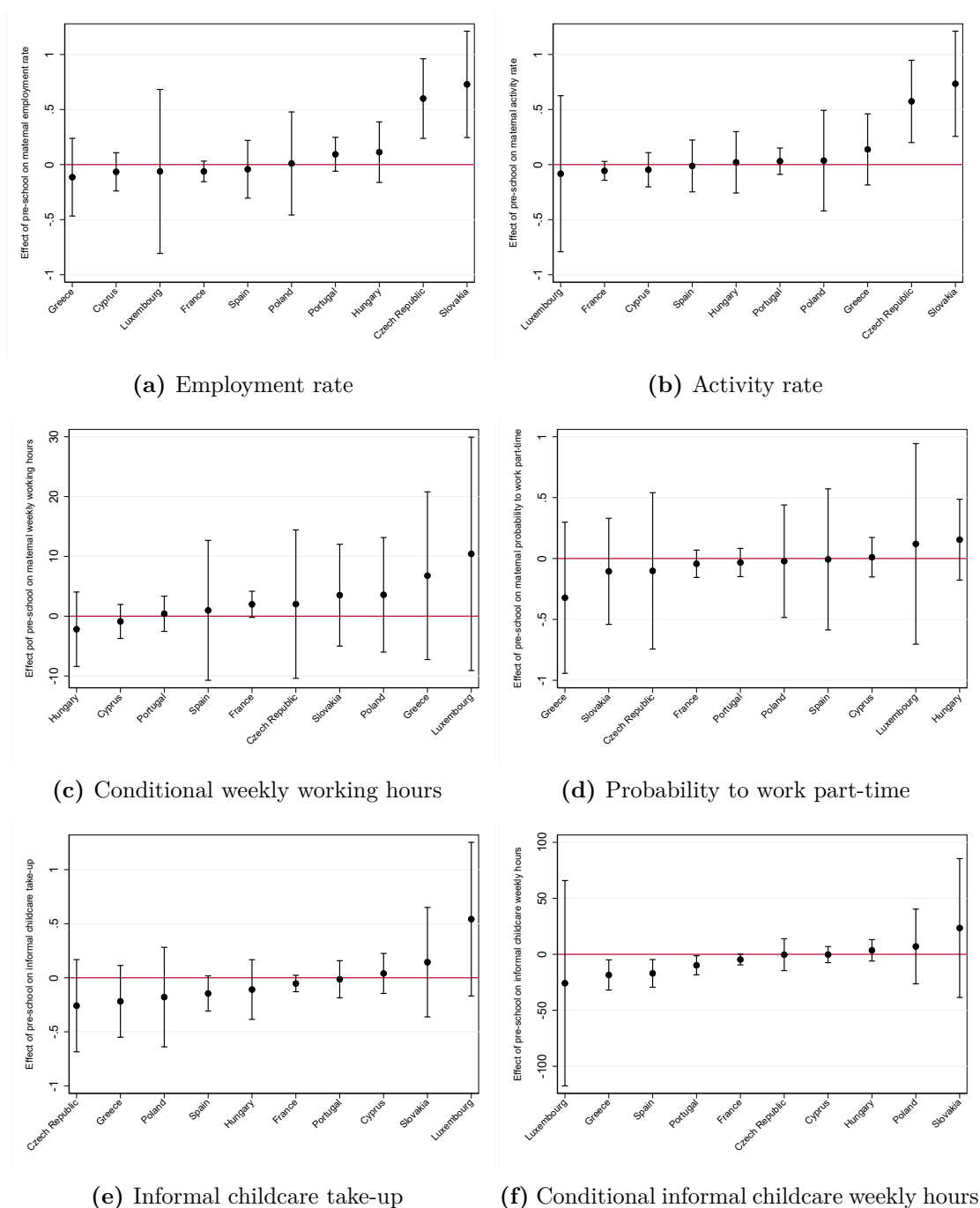


Note : For each observable characteristic, means are reported for the whole sample, compliers, always-takers and never-takers with confidence intervals at the level of 95%.
 Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

H Results by country

In Section 5.1, we provide main results pooling all countries together. Here, we report the results by country.

Figure H1: Results by country



Note : Coefficients with confidence intervals at the level of 95%.

Source : European Union Survey on Income and Living Conditions (EU-SILC), 2004-2020.

Chapter 3

Inherited Gender Norms and the Gender Gap in Cognitive Functioning

co-authored with Eric Bonsang

Our acknowledgments go to Thomas Barnay, Justine Bondoux, Eve Caroli, Brigitte Dormont, Sarah Flèche and Jürgen Maurer for their valuable comments on earlier drafts. We also thank participants of the Dauphine PhD Workshop, the 42nd Journées des Economistes de la Santé Français, the IRES Lunch Seminar (UCLouvain), the LIVES Workshop (University of Lausanne), the Lille Economie Management Seminar (University of Lille) and the Netspar Workshop for their suggestions. This study is based on data from the Survey of Health, Ageing and Retirement in Europe (Waves 1-8). We would like to thank members of the SHARE project board for granting us access to the data and Thomas Renaud for sharing his insights on the data.

Table of contents

1	Introduction	176
2	Empirical strategy	178
2.1	Baseline model	178
2.2	Heterogeneity with respect to parental place of birth	179
3	Data	180
3.1	The analytical sample	180
3.2	Measures of cognitive functioning	181
3.3	Covariates and mediating variables	185
3.4	Gender Norms Index	185
3.5	Graphical description	187
4	Results	189
4.1	Main results	189
4.2	Robustness checks	191
4.3	Mediating factors	195
4.4	Further findings	197
5	Conclusion	198
	Bibliography	204
	Appendix	205
A	Descriptive statistics	205
B	Main results by sub-group	207
B.1	Methodology	207
B.2	Results	208
C	Alternative gender norms proxies	210
C.1	Description of alternative measures	210
C.2	Results	211
D	Additional control variables	214
D.1	Individual level controls	214
D.2	Country level controls	215
E	Sample restrictions	218
F	Further results	219

Abstract

This paper investigates the relationship between gender norms and the cognitive gender gap among older individuals. We use data from the Survey of Health, Ageing and Retirement in Europe (SHARE) and the Integrated Values Surveys (IVS) to examine how gender differences in cognitive functioning among second-generation immigrants are related to gender norms in their parents' country of birth. This approach allows to identify the effect of social norms while holding the institutional background fixed. Our results indicate that more conservative gender norms in the parents' country of birth are associated with lower cognitive test scores for women compared to men. Further explorations suggest that gender differences in the type of occupation partly explain this relationship. This study highlights that policies aiming at promoting gender equality might have important implications for the cognitive health of older women.

1 Introduction

Age-related cognitive decline is a widespread phenomenon that raises a significant challenge for ageing societies. Cognitive impairment is linked to functional disability [McGuire et al., 2006], to a decreased quality of life [Hussenoeder et al., 2020] and to higher healthcare costs [Vecchio et al., 2016]. Older women are particularly affected as they have higher age-specific dementia rates and longevity than men [Prince et al., 2013; Wimo et al., 2013]. In most countries, the proportion of women is about two to one for individuals above 80 years old [United Nations, 2017]. This implies that the potential burden of age-related cognitive impairment is larger for ageing women than ageing men. Therefore, it is crucial to understand what contributes to cognitive functioning differences between women and men as they grow older.

A limited number of studies has investigated the determinants of gender disparities in cognitive functioning during later stages of life. Among these studies, the predominant focus has been on socio-economic factors and economic development [Yount, 2008; Lei et al., 2012; Weber et al., 2014; Maurer, 2011] with comparatively less emphasis on cultural determinants. Yet, gender norms dictate individual behaviors and have been shown detrimental to a wide range of women's economic outcomes [Giuliano, 2020]. There is, however, one study of Bonsang et al. [2017] that analyses the relationship between gender-role attitudes and gender differences in cognitive functioning in later life across countries and cohorts. The authors find that older women perform relatively better than men in countries characterised by more equal gender-role attitudes. By exploiting cross-country variations in gender-role attitudes across cohorts, they show that this association is robust to the inclusion of country and cohort-fixed effects and further analyses suggest that this association is unlikely to be driven by reverse causality.

This paper investigates whether gender norms contribute to the cognitive gender gap among individuals aged at least 50 relying on the *epidemiological approach*. Identifying the effect of gender norms is challenging as social norms and institutions are intertwined [Alesina and Giuliano, 2015]. This methodology overcomes this issue by exploiting variations in social norms while holding the institutional background fixed. This is possible using a sample of second-generation immigrants because they are exposed to the same institutional environment but

have different cultural backgrounds depending on the country of origin of their parents. We use data from the Survey of Health, Ageing and Retirement in Europe (SHARE) that includes cognitive test scores as well as information on individuals' and parents' countries of birth to target second-generation immigrants. Drawing on the existing literature, we proxy gender norms at the parental country-of-birth level (i.e. thereafter country of ancestry) using prevailing gender role attitudes in the Integrated Values Surveys (IVS).

This paper adds to the literature that uses this approach to investigate the contribution of gender norms to female outcomes or gender gaps. [Fernández and Fogli \[2009\]](#) and [Blau et al. \[2013\]](#) show that female labour force participation and fertility outcomes¹ in the country of ancestry drive labour and fertility decisions of female second-generation immigrants in the United States. [Finseraas and Kotsadam \[2017\]](#) find similar results for female labour force participation in Norway. In order to capture the effect on gender differences rather than on female absolute outcome values, [Antecol \[2000\]](#) interacts the gender gap in labour force participation in the country of ancestry with the individual's sex in the second-generation immigrants sample. She shows that the inter-generational transmission of gender norms plays an important role in labour force participation gender gaps in the host country. In this literature, other studies use measures of gender equality in the country of ancestry to proxy gender norms. Among them, [Hauge et al. \[2023\]](#) assess the effect of gender norms on the gender gap in competitive behaviours. The authors conduct an experiment to elicit willingness to compete among second-generation immigrants in Norway and use the World Economic Forum's Gender Gap Index as a measure of gender norms in the country of ancestry. Their study shows that the gender gap in competition decreases in Norway when gender equality is higher in the country of ancestry. Using the same gender equality measures, [González and Rodríguez-Planas \[2020\]](#) show that second-generation immigrants women in the host country are less exposed to intimate partner violence when the country of ancestry is more gender-equal. In this literature, our paper is mostly related to the work of [Nollenberger et al. \[2016\]](#) and [Rodríguez-Planas and Nollenberger \[2018\]](#) who investigate whether gender norms contribute to the gender gap in cognitive performance at school. They suggest that the gender gap in mathematics in the host country reduces with more gender-equal norms in the country of ancestry.

¹[Fernández and Fogli \[2009\]](#) use fertility rates and [Blau et al. \[2013\]](#) the number of children.

Among second-generation immigrants aged 50 and above, we find that women’s cognitive skills decrease relatively to men’s in the country of residence when gender norms are more conservative in the country of birth of both parents. Among individuals who have only one parent born abroad, these relationships are not significantly different from zero at any conventional level, which is in line with previous findings in the literature. We also find that, when the parents come from a country with more conservative gender norms, women have more difficulties than men with some instrumental activities of daily living such as using a telephone and finding the way in an unfamiliar place. Our results are robust to the use of alternative measures of gender norms and to controlling for individual as well as country level characteristics that could confound the relationship between gender norms and the cognitive gender gap. Lastly, explorations of potential mechanisms show that the type of occupation is likely to be a driver of this relationship.

The remainder of this paper is organized as follows. First, the empirical strategy is detailed in Section 2, the data is described in Section 3, results are presented and their robustness is checked in Section 4. Our findings are then discussed in Section 5.

2 Empirical strategy

This section presents the *epidemiological approach* developed in the literature using a sample of second-generation immigrants who have at least one parent born abroad (Section 2.1). In Section 2.2, we allow the relationship of interest to vary according to whether individuals have one or both parents born abroad.

2.1 Baseline model

First, we run the following regression using the Ordinary Least Squares (OLS) estimator among second-generation immigrants :

$$C_{isrta} = \alpha_0 + \alpha_1 Female_i + \alpha_2 Female_i * GNI_a + \gamma_a + \gamma_r + \gamma_{rs} + \delta_t + \rho X'_{it} + \epsilon_{isrta} \quad (1)$$

C_{isrta} is the cognitive score of individual i of sex s living in country of residence r at time t who has at least one parent born in country of ancestry a . α_0 is the intercept, $Female_i$ the

binary variable indicating individual i 's sex equal to 0 for a man and to 1 for a woman, GNI_a the Gender Norms Index at the country of ancestry level such that gender norms are more conservative when the index has a high value, γ_a the country of ancestry specific effect, γ_r the country of residence specific effect, γ_{rs} the country of residence by sex specific effect, δ_t the interview wave specific effect, X'_{it} the vector of individual control variables and ϵ_{isrta} the error term.

α_2 is the coefficient of interest as it captures the effect of traditional gender norms in the country of ancestry on the cognitive gender gap among second-generation immigrants in the country of residence. We standardise both the cognitive test score and the GNI in order to simplify the interpretation of the results. Consequently, the effect is interpreted as the change in $\hat{\alpha}_2$ standard deviation from the average cognitive test score for a woman compared to a man when the GNI increases by 1 standard deviation from the mean. $\hat{\alpha}_2$ is expected negative, suggesting that the female cognitive score decreases relatively to the male score when gender norms are more conservative.

The model includes country of residence and country of ancestry specific effects to control for country characteristics that might be related to cognitive performances the same way for men and women. The country of residence by sex specific effect accounts for country characteristics contributing to the cognitive gender gap that affect men and women differently. Interview wave effects account for time-specific characteristics that might correlate to individuals' cognitive performances. The vector X'_{it} aims at controlling for any individual characteristic associated with cognitive functioning but unrelated to gender norms (i.e. age and age squared).

2.2 Heterogeneity with respect to parental place of birth

Because results are expected to differ according to whether parents come from the same country or not [Blau et al., 2013; Rodríguez-Planas and Nollenberger, 2018], we distinguish second-generation immigrants with both of their parents born in another country from those with only one parent born abroad. Gender norms vary at the level of one country of ancestry. In that respect, we exclude individuals who have two parents born abroad but from different countries. Further details on each second-generation immigrants group are given in Section 3.1.

We construct a binary variable (One_i) equal to 1 if only one parent of individual i is born abroad

and to 0 if both parents are born in the same country. This variable is added to the baseline model and interacted with the binary variable indicating the individual's sex ($Female_i$), with the Gender Norms Index (GNI_a) and with both of them, giving :

$$C_{isrta} = \beta_0 + \beta_1 Female_i + \beta_2 Female_i * GNI_a + \beta_3 One_i + \beta_4 One_i * Female_i + \beta_5 One_i * GNI_a + \beta_6 Female_i * GNI_a * One_i + \theta_a + \theta_r + \theta_{rs} + \lambda_t + \tau X'_{it} + \mu_{isrta} \quad (2)$$

Here, β_2 and β_6 are of particular interest. β_2 represents the change in the cognitive gender gap in the country of residence when the GNI increases in the country of ancestry of both parents while β_6 captures the change for second-generation immigrants with only one parent (either the mother or the father) born in this country compared to those with both parents born in the same country. The same control variables and specific effects as in the baseline model are included.

Because individuals are observed several times, standard-errors are corrected for clustering at the individual level and reported in parentheses in all specifications. In line with previous literature, we also provide standard-errors corrected at the country of ancestry level. However, the number of countries of ancestry is rather low, leading to downward biased standard-errors when correcting for clustering at the country of ancestry level [Angrist and Pischke, 2009]. We therefore do not report them to avoid interpreting misleading results. Instead, standard-errors are corrected for clustering at the country of ancestry level using wild bootstrap for a better inference [Cameron et al., 2008]. P-values of the coefficient's significance test are reported in brackets.

3 Data

3.1 The analytical sample

The sample of interest is drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE), a longitudinal dataset containing demographic, socio-economic and health information on individuals aged 50 and older in Europe from 2004 onwards. Our study exploits data from Wave 1 to Wave 8, providing observations until 2020, except from Wave 3 which was dedicated to a retrospective questionnaire and thus did not include the information we need for the

analysis.

We define second-generation immigrants as individuals who are born in their country of residence and have at least one parent born in another country. As presented in Section 2.2, second-generation immigrants can either have both of their parents born abroad or only one of them. Among second-generation immigrants whose both parents are born in another country, we exclude individuals who have two parents born in different countries since we need a unique measure for gender norms at the country of ancestry level.

Countries of ancestry that contain less than 10 observations or contain only men or only women are excluded. Moreover, we exclude second-generation immigrants from Sweden as it has been found to be an outlier in terms of cognitive gender gap. Taking these restrictions into account, the final sample contains 10,435 second-generation immigrants among which 2,509 have both parents born in the same country and 7,926 have only one parent born in another country (4,229 whose mother is born abroad and 3,697 whose father is born abroad). Sample sizes by country of ancestry are presented in Table A1.

3.2 Measures of cognitive functioning

Cognitive functioning is assessed by a series of tests conducted during a face-to-face interview. These tests are based on the Mini Mental State Examination which aim is to evaluate the individual's general cognitive status [Folstein et al., 1975]. They are used by medical doctors to detect dementia symptoms. This section presents how these tests are conducted by interviewers. Means and standard deviations are presented in Table 1 for the whole second-generation immigrants sample (Column 1), for second-generation immigrants whose both parents are born in the same country (Column 2) and for second-generation immigrants with only one parent born abroad (Column 3). Differences in means (with standard-errors) between both groups are also reported.

Episodic memory. Individuals are asked to memorize 10 common words that the interviewer gives out loud. Episodic memory is measured by asking individuals to list as many of these words they can remember immediately and with a delay. The immediate word recall (IWR) test takes place just after the interviewer reads the 10 words out. The delayed word recall (DWR) test occurs after completing the fluency and numeracy tests that we describe below.

Table 1: Descriptive statistics

	All second-generation immigrants (1)	Both parents are born abroad (2)	Only one parent is born abroad (3)	Difference (2) - (3)
Main outcomes				
Immediate word recall score	5.635 (1.673)	5.525 (1.726)	5.670 (1.654)	-0.145* (0.110)
Delayed word recall score	4.346 (2.139)	4.133 (2.142)	4.413 (2.133)	-0.281* (0.185)
Verbal fluency score	22.148 (7.230)	21.398 (7.447)	22.386 (7.144)	-0.988 (1.158)
Numeracy score	3.591 (0.988)	3.518 (1.005)	3.615 (0.981)	-0.097 (0.104)
Subtraction score	4.467 (0.891)	4.397 (1.005)	4.489 (0.923)	-0.092 (0.104)
Covariates				
Age	65.989 (9.331)	64.814 (9.499)	66.361 (9.247)	-1.547 (1.745)
Woman (%)	0.570 (0.245)	0.564 (0.496)	0.571 (0.495)	-0.007 (0.043)
Mediating variables				
No education (%)	0.010 (0.100)	0.016 (0.124)	0.009 (0.092)	0.007 (0.012)
At least secondary education (%)	0.333 (0.222)	0.294 (0.456)	0.345 (0.475)	-0.052 (0.059)
Never worked (%)	0.025 (0.155)	0.030 (0.171)	0.023 (0.150)	0.007 (0.027)
White collar worker (%)	0.500 (0.500)	0.422 (0.494)	0.524 (0.499)	-0.102* (0.077)
Non-professional activity (%)	0.919 (0.074)	0.865 (0.341)	0.936 (0.244)	-0.071** (0.031)
Number of countries of ancestry	44	28	43	44
Number of observations	10,435	2,509	7,926	10,435

Note: Column 1 (resp. 2,3) reports means with standard deviations in parentheses for all second-generation immigrants (resp. second-generation immigrants with both parents in the same country, second-generation immigrants with only one parent born abroad). The fourth column reports the difference in mean values between the two sub-groups of second-generation immigrants with standard errors corrected for clustering at the country of ancestry level in parentheses. The proportion of white-collar workers is calculated among individuals who ever worked.

Samples : All second-generation immigrants.

Source : Survey of Health, Ageing and Retirement in Europe (Waves 1-8).

Both the IWR and DWR measures range from 0 (i.e. the individual does not remember any of the listed words) to 10. On average, individuals remember 5.6 words immediately and 4.3 with the delay. second-generation immigrants with only one parent born abroad perform better, but the difference is very low (-0.145 words for IWR and -0.281 words for DWR, $p < 0.1$).

Verbal fluency. For this test, individuals have to name as many different animals as possible in 1 minute. The verbal fluency task is a tool used to measure executive functions [Shao et al., 2014]. Executive functions are necessary for the cognitive control of behaviour, i.e. selecting and successfully monitoring behaviours that facilitate the attainment of goals. The verbal fluency

test score ranges from 0 (i.e. the individual does not give any animal name) to 70. On average, individuals provide 22 names in 1 minute.

Numeracy. Numeracy is assessed by asking a few questions that involve simple calculations based on real life situations. Respondents who correctly answer the first question are asked a more difficult one while those who make a mistake are asked an easier one. The first mathematical task is the following:

1) *“If the chance of getting a disease is 10%, how many people out of 1,000 would be expected to get the disease?”.*

If the individual gives a wrong answer, she gets the following question:

2.a) *“In a sale, a shop is selling all items at half price. Before the sale, a sofa costs 300 (local currency). How much will it cost in the sale?”.*

After this question, the numeracy test is stopped, independently of whether the answer is correct or not. If the first question is correctly answered, the following question is asked:

2.b) *“A second hand car dealer is selling a car for 6,000 (local currency). This is two-thirds of what it costs new. How much did the car cost new?”.*

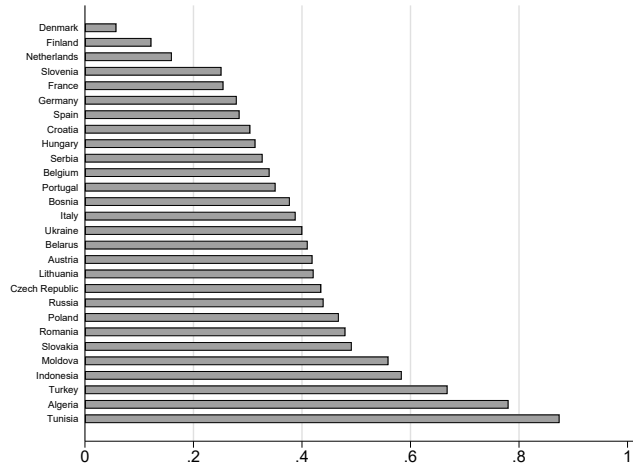
Only if both the first and second numeracy questions are correctly answered, there is a last question:

3) *“Let’s say you have 2,000 (local currency) in a savings account. The account earns 10% interest each year. How much would you have in the account at the end of two years?”.*

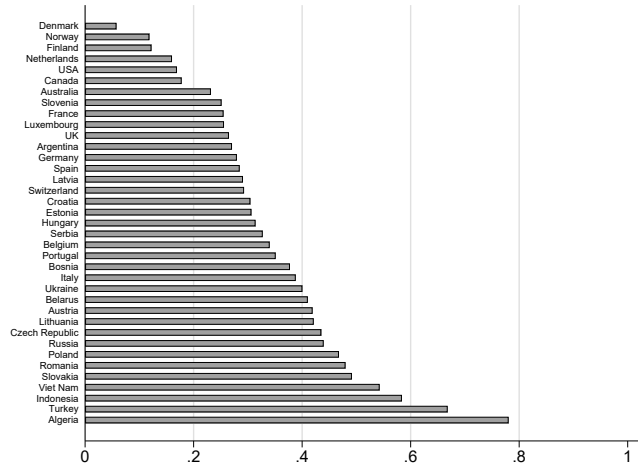
The numeracy score ranges from 1 to 5 and is computed as follows. If answers to question 1) and question 2.a) are wrong, the score equals 1 (i.e. the lowest value) and it equals 2 for individuals who gave the right answer to 2.a). For individuals who correctly answered question 1) but not to 2.b), the score equals 3. If the answer given to question 2.b) was correct, individuals get a score of 4 if they did not manage to find the answer to question 3) and a score of 5 otherwise. In the whole population, this score averages at 3.6.

Serial Subtractions. This test evaluates the individual’s ability to subtract 7 from 100 five consecutive times. The test score ranges from 0 to 5 and averages at 4.5.

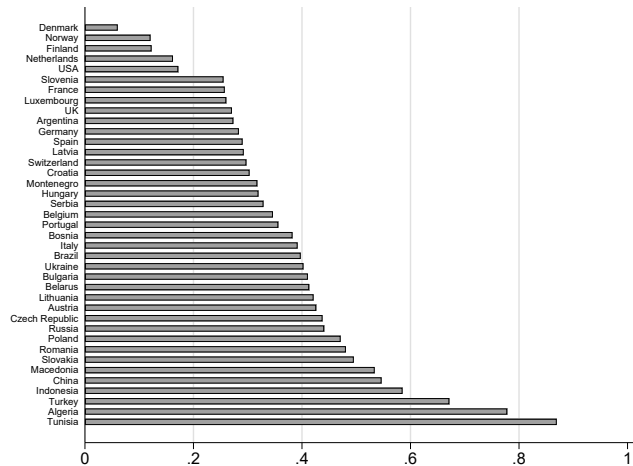
Figure 1: Gender Norms Index by country of ancestry



(a) Both parents are born abroad



(b) Only mother is born abroad



(c) Only father is born abroad

Note : The Gender Norms Index is the proportion of individuals agreeing with the following statement: “When jobs are scarce, men should have more right to a job than women” in the countries of ancestry declared by the population of interest (second-generation immigrants in SHARE).

Source : Integrated Values Surveys, 1981-2020.

For fluency, numeracy and serial subtraction tests, scores do not significantly differ between second-generation immigrants groups. We use the average of these cognitive test scores, denoted as the Cognitive Score Index (CSI), as the main outcome for the graphical description (Section 3.5) and the estimations (Section 4). Given that some cognitive test scores take their values in different ranges, we use their standardised values to compute the CSI.

3.3 Covariates and mediating variables

On average, individuals are 66 years old at the time of interview and 57% of them are women. 1% of individuals has no education (i.e. no or incomplete primary education) and 33% of them completed at least a secondary education level². Among those who have ever worked (97.5%), 50% are *white-collars*³ (or used to be, if retired). second-generation immigrants whose both parents are born in the same country are less likely to have a white-collar occupation by 10 percentage points ($p < 0.1$).

More than 90% of the sample has conducted at least one non-professional activity in the last 12 months among voluntary work, training courses, sports or social activity, religious activity, political activity, reading books and playing games. second-generation immigrants with both parents born in another country are less likely to have done at least one of these activities in the year than other second-generation immigrants by 7 percentage points ($p < 0.05$). Variables related to education, professional and non-professional activity are explored as potential mediating factors in Section 4.3.

3.4 Gender Norms Index

Following previous studies investigating the relationship between gender norms and individual characteristics, we proxy gender norms with the average opinion of individuals regarding gender roles in the Integrated Values Surveys (IVS)⁴ [Fortin, 2005; Guiso et al., 2008; Bonsang et al., 2017; Hauge et al., 2023]. The aim of these surveys is to shed light on beliefs and values on

²Education level is based on the International Standard Classification of Education.

³We categorise *white-collars* following the International Standard Classification of Occupations. *White-collars* include managers (e.g. chief executives, senior officials, administrative and commercial managers...), professionals (e.g. engineers, health professionals, teachers...) technical professionals and clerks (e.g. healthcare assistants, social workers, secretaries...).

⁴The IVS merges the European Values Study (EVS) and the World Values Survey (WVS) files together. See <https://europeanvaluesstudy.eu/methodology-data-documentation/integrated-values-surveys/>

various topics such as politics, social relations, or gender roles on the international scale.

We construct a Gender Norms Index (GNI) which takes its values in IVS countries that match countries of ancestry in SHARE. It is measured as the proportion of individuals agreeing with the following statement: “*When jobs are scarce, men should have more right to a job than women*”. In the surveys, individuals are asked if they agree or disagree with the statement or are indifferent to it. We construct a binary variable taking 1 if the individual agrees, 0 if she disagrees. Thus, the GNI ranges from 0 to 1, 0 meaning that no one agrees with the conservative statement in the given country. We select this statement because it implies for individuals to compare the perceptions they have of both male and female roles in society. On the contrary, using a question such as “*Pre-school child suffers with a working mother*”⁵ does not provide any information about what the individual would answer for a man, which would not allow to elicit gender role attitudes. IVS data are collected from 1981 onwards, which does not allow to proxy gender norm at the time parents used to live in the country. Nonetheless, using contemporaneous measures of gender norms is common in the literature [Fernández and Fogli, 2009; Rodríguez-Planas and Nollenberger, 2018; González and Rodríguez-Planas, 2020] since they capture culture in the country of ancestry around the time we observe second-generation immigrants in the country of residence. For more consistence, the GNI is computed among individuals in the IVS who are born in the same cohorts as second-generation immigrants in SHARE.

Figure 1 displays GNI values in countries of ancestry. Sub-figure 1a (resp. Sub-figure 1b, Sub-figure 1c) reports the average GNI value in each country if both parents are born in the same country (resp. if only the mother is born in this country, if only the father is born in this country). In all cases, the lowest GNI is held by Denmark where 6% of the population agree with giving a priority to men on the labour market when jobs are scarce, while it is the highest in Tunisia (88%) among individuals who have both parents or only their father born abroad and in Algeria (78%) for individuals whose only the mother is born in another country.

⁵This question is also asked in the IVS.

3.5 Graphical description

We first provide a graphical representation of the association between the cognitive gender gap among second-generation immigrants with the same country of ancestry and the GNI in this given country. The Cognitive Gender Gap (CGG) is defined as:

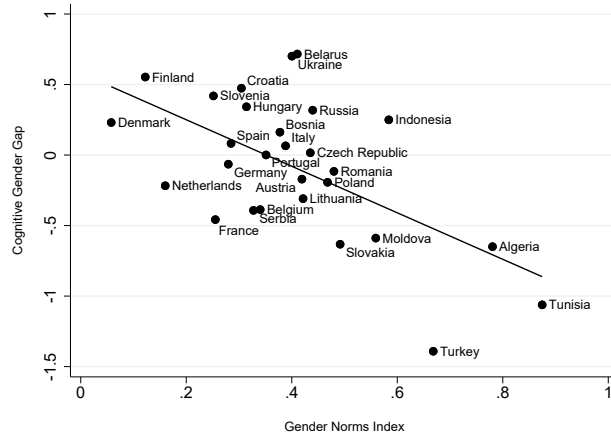
$$CGG_a = CSI_{fa} - CSI_{ma} \quad (3)$$

CGG_a is the difference in Cognitive Score Index⁶ between female and male second-generation immigrants who have at least one parent born in country of ancestry a , CSI_{fa} is the female Cognitive Score Index and CSI_{ma} the male Cognitive Score Index among second-generation immigrants with country of ancestry a . A positive (resp. negative) CGG represents a female (resp. male) advantage in CSI. This relationship is estimated using the OLS estimator without any control variable nor specific effect.

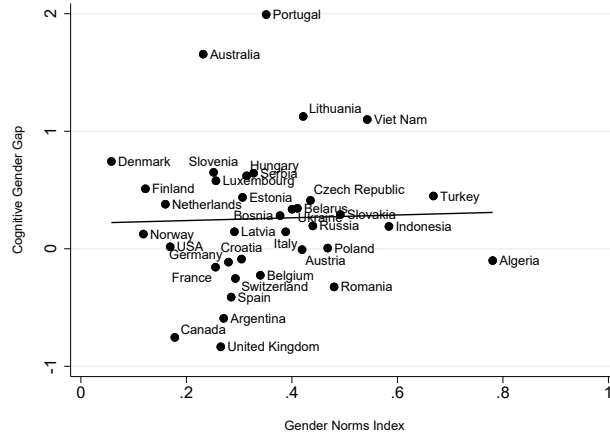
Figure 2 displays this relationship for individuals whose parents are both born abroad and in the same country (Sub-figure 2a), for individuals whose only the mother is born abroad (Sub-figure 2b) and for individuals whose only the father is born abroad (Sub-figure 2c). It shows that the CGG is negatively associated with the GNI among second-generation immigrants who have both parents born in the same country, with a slope of -1.649 ($p < 0.01$). In other words, the cognitive female advantage that exists among second-generation immigrants from a country of ancestry with a low GNI (i.e. less conservative countries) decreases while the GNI increases and even becomes a male advantage in more conservative countries. For second-generation immigrants with only one parent born in country of ancestry a , the relationship is not significant at any conventional level.

⁶This index is described in Section 3.2.

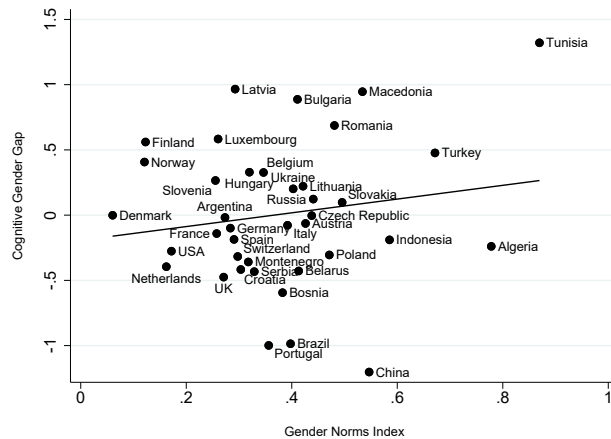
Figure 2: Association between the Gender Norms Index and the Cognitive Gender Gap



(a) Both parents are born abroad



(b) Only mother is born abroad



(c) Only father is born abroad

Note : Scatterplot of the association between the Gender Norms Index and the Cognitive Gender Gap among second-generation immigrants with the best fitting regression line (OLS).

Sample : Second-generation immigrants whose (a) both parents are born abroad and in the same country, (b) only the mother is born abroad and (c) only the father is born abroad.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

4 Results

We report estimation results based on Equation 1 in the first two columns of Table 2. Column 1 (resp. 2) displays results without (resp. with) country of ancestry specific effects. We report estimation results related to Equation 2 in the last two columns of Table 2. As in the baseline model, results are reported without (resp. with) country of ancestry specific effects in Column 3 (resp. 4). We also report results for each second-generation immigrants group separately such as described in Section B.1.

4.1 Main results

In Table 2, Column 1 (resp. 2) reports results of the baseline model without (resp. with) country of ancestry specific effects. In both specifications, there is no significant difference in CSI between men and women among all second-generation immigrants. Column 1 shows that the GNI is not significantly correlated to the CSI for men, but decreases by 0.073 standard deviation for women relatively to men when the GNI increases by 1 standard deviation ($p < 0.05$). Controlling for country of ancestry specific effects (Column 2), the change in the cognitive gender gap has a lower amplitude (-0.063 sd, $p < 0.1$), suggesting that other factors at the country of ancestry level are correlated to cognitive functioning among all second-generation immigrants. Because it is more conservative, the specification that includes country of ancestry specific effects is preferred.

When allowing results to differ according to whether both parents are born abroad or not, we observe that the GNI negatively contributes to the cognitive gender gap among second-generation immigrants whose both parents are born in the same country (Columns 3 and 4). More precisely, the CSI decreases by 0.243 sd ($p < 0.01$) (resp. by 0.237, $p < 0.01$) for women compared to men when the GNI increases by 1 sd among second-generation immigrants whose both parents are from the same country. The effect is entirely driven by second-generation immigrants whose both parents are born in the same country since the effect is higher by 0.200 sd (resp. 0.201) among second-generation immigrants with only one parent born abroad relatively to those with both parents born in the same country ($p < 0.01$). The difference between both coefficients equals -0.043 (resp. -0.036) and is not significantly different from 0 at

Table 2: Estimation results

	Cognitive Score Index			
	(1)	(2)	(3)	(4)
Gender Norms Index	0.023 (0.025)		0.117** (0.052)	
Woman	0.009 (0.109)	0.016 (0.108)	0.277 (0.231)	0.252 (0.245)
Gender Norms Index * Woman	-0.073** (0.037) [0.075]	-0.063* (0.035) [0.100]	-0.243*** (0.093) [0.016]	-0.237*** (0.081) [0.010]
Only one parent is born abroad			0.425 (1.386)	0.430 (1.397)
Only one parent is born abroad * Woman			-0.299 (0.261)	-0.264 (0.271)
Only one parent is born abroad * Gender Norms Index			-0.101* (0.060) [0.190]	-0.121** (0.059) [0.109]
Only one parent is born abroad * Gender Norms Index * Woman			0.200** (0.102) [0.057]	0.201** (0.091) [0.050]
Country of ancestry FE	no	yes	no	yes
Country of residence FE	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes
Age, age squared	yes	yes	yes	yes
Fully interacted model	yes	yes	yes	yes
R^2	0.163	0.172	0.178	0.187
Number of ancestry countries	44	44	44	44
Number of individuals	4,023	4,023	4,023	4,023
Number of observations	10,435	10,435	10,435	10,435

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Regressions are detailed in Section 2. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : All second-generation immigrants.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

any conventional level, implying the absence of effect among those with only one parent born abroad. This result is in line with a higher cultural transmission when both parents share the same culture, as highlighted in the previous literature (see Section 2.2). Remaining analyses therefore focus on individuals with both parents born in the same country.

Since only second-generation immigrants with both parents born abroad drive the results, all the following analyses focus on this sample for the sake of a simpler interpretation. Table B1 reports the results for all cognitive test scores among second-generation immigrants who have both of their parents born abroad.⁷ In order for coefficients to be comparable, we use the standardised values of cognitive test scores as outcomes. Table B1 shows that, among second-generation

⁷Results for second-generation immigrants with only on parent born abroad are reported in Section B.

immigrants whose both parents are born in the same country, the GNI mostly contributes to the gender gap in numeracy test score (-0.184 sd, $p < 0.01$). To a lower extent, gender gaps in DWR (-0.128 sd, $p < 0.05$) and in VF test scores (-0.127 sd, $p < 0.05$) are also affected. The effect is the lowest with respect to IWR test score (-0.098, $p < 0.1$). With a more conservative computation of standard-errors (i.e. bootstrapped standard-errors corrected for clustering at the country of ancestry level), the effect is significant at the level of 5% for episodic memory measures (IWR and DWR).

4.2 Robustness checks

This section aims at checking the robustness of our results. We have shown in Section 4.1 that gender norms in the country of ancestry increase the cognitive gender gap in the country of residence only among second-generation immigrants whose both parents are born in the same country. We thus conduct robustness checks on this specific sample.

4.2.1 Alternative proxies for gender norms

First, we check whether our results hold when using other measures of gender norms. Following previous literature, we use the gender gap in labour force participation (LFP) and the *World Economic Forum's* Gender Gap Index (GGI).⁸ For LFP, we use information from the *World Development Indicators* database.⁹ In both cases, data is made available by the World Bank.

Gender gap in labour force participation. Several studies in the literature we build on use the logarithm (thereafter, log) of female labour force participation (FLFP) in the country of ancestry as a proxy for gender norms [Fernández and Fogli, 2009; Finseraas and Kotsadam, 2017; Hauge et al., 2023]. However, a low FLFP might not be informative on gender norms prevailing in a given country if male labour force participation (MLFP) has the same level. For this reason, we follow Antecol [2000] and use the difference in (log of) LFP between men and women. LFP rates are provided from 1990 to 2021. For each country of ancestry, we compute the gender gap in log(LFP) for each year between 1990 and 2021 and then take the average over the whole time period as a proxy for gender norms in the country of ancestry. As for the GNI, this measure is then standardised with a mean of 0 and standard deviation of 1.

⁸<https://www.weforum.org/reports/global-gender-gap-report-2023/>

⁹<https://databank.worldbank.org/source/world-development-indicators>

Table C2 displays the results with this alternative proxy for gender norms in Column 1. Results show that the CSI decreases by 0.150 sd for women relatively to men when the gender gap in LFP increases by 1 sd in the country of ancestry. The coefficient is slightly lower than the one obtained when using the GNI, but is still significant. We also conduct the analysis for each year of LFP measurement instead of the average value over the time period. Figure C1 shows that results are not sensitive to the year we use.

Gender Gap Index. The GGI is commonly used in the literature we refer to [Guiso et al., 2008; Nollenberger et al., 2016; Rodríguez-Planas and Nollenberger, 2018; González and Rodríguez-Planas, 2020; Hauge et al., 2023]. It measures gender gaps in 4 categories : Economic participation and opportunity, Educational attainment, Political empowerment and Health and survival. More details on each sub-index are given in Section C.1. Each GGI ranges from 0 (equality) to 1 (inequality). In opposition with the GNI, a higher GGI value therefore indicates a lower gender gap. In this case, the coefficient of interest is expected positive. As for the gender gap in $\log(\text{LFP})$, we use the average value of each GGI over the time period of data collection for the main analysis. We then check that results hold for each year of GGI measurement.

Results using each GGI sub-index are reported in Table C2. Interestingly, the cognitive gender gap is only affected by the GGI in economic participation and opportunity. The CSI increases by 0.155 sd for women relatively to men when there is more gender equality in economic opportunities in the country of ancestry. The effect holds irrespective of the year of GGI measurement we use (Figure C2). For other GGI sub-indices, coefficients are positive but insignificantly different from 0. This result is consistent with the fact that the cognitive gender gap increases with the gender gap in LFP and with the conservative gender role attitudes on the labour market (GNI). This might indicate that gender gaps on the labour market drive the relationship between the gender norms and the cognitive gender gap. We further investigate this hypothesis in Section 4.3. However, we cannot discard that the absence of results when using the GGI in education or the GGI in health might be explained by the fact that their means are very close to the upper bound and that their variability is very low (Table C1).

4.2.2 Additional controls

We also want to know if our results are sensitive to controlling for other potential factors that could confound the relationship between gender norms in the country of ancestry and the cognitive gender gap in the country of residence.

Individual level controls. If less (or more) educated parents invest more in their sons' human capital than in their daughters' and that their education level is correlated to the level of gender norms conservatism in their country of birth, omitting parental education levels as well as their interactions with the individual's sex might lead to biased estimates. It is nonetheless important to note that controlling for parents' education levels might also bias the coefficient of interest if they are causally affected by gender norms in the country of ancestry. Keeping this in mind, we conduct the robustness check by adding indicators for each parent's levels of education as well as its interaction with the binary variable that indicates the sex of the individual to the regression and check if results are sensitive to those changes or not.¹⁰

Table D1 presents the results controlling for the mother's (resp. the father's, both parents') education level(s) and its interaction with the sex indicator in Column 1 (resp. 2, 3). Both mother's and father's levels of education are positively correlated to the individual's CSI, but not differently between men and women. The coefficient of interest appears insensitive to the inclusion of these variables and similar to the baseline result.

Country level controls. We also want to check that our results are not explained by economic nor human development at the country of ancestry level. If it were the case, it would cast doubt on the existence of any effect of gender norms transmission on the cognitive gender gap. To this aim, we include the level of Gross Domestic Product per capita per purchasing power (GDP), the Human Development Index (HDI) and the unemployment rate as well as interactions between the variable indicating the sex of the individual and each of these country-level variables. These indicators are provided by the *World Development Indicators* database from 1990 to 2022. As done previously, we take the average value over the whole time period for each indicator and run a sensitivity analysis according to the year of measurement. Table D2 shows that GDP and HDI are highly correlated (0.938) and might therefore capture similar characteristics. Unemployment

¹⁰Parental education levels are defined as previously done for second-generation immigrants' educational levels (see Section 3.3).

being poorly correlated to GDP (-0.131) as well as to HDI (-0.117), including it as a control variable is likely to bring additional information.

Results are presented in Table D3. Column 1 shows that the GDP is negatively correlated with the cognitive gender gap, in line with the literature documenting the negative relationship between gender inequality and economic development [Seguino, 2000]. The coefficient of interest remains significantly different from zero and is even larger compared to the main results, suggesting that omitting GDP in the regression biases the coefficient of interest toward 0. The HDI and the unemployment rate do not significantly contribute to the cognitive gender gap (see Columns 2 and 3). Including the interaction between the HDI and the individual's sex indicator keeps the coefficient of interest significant which amplitude is similar to the one obtained while controlling for the level of GDP interacted with the sex indicator. However, including the interaction term between the unemployment rate and the female indicator leaves the coefficient size virtually unchanged. Column 4 reports coefficients when including all of the 3 interacted indicators and shows that results hold with this specification.

4.2.3 Additional sample restrictions

We want to check that our results are not driven by the largest groups of second-generation immigrants (see Table A1). To this aim, we exclude countries of ancestry containing more than 100 second-generation immigrants in our sample of interest. Table E1 provides the results excluding second-generation immigrants with both parents born in Germany (Column 1), in Italy (Column 2), in Poland (Column 3) and in Russia (Column 4). Results hold in all cases. Coefficients are slightly larger, in line with a larger transmission from parents to children when cultural traits are not dominant in the host country [Bisin and Verdier, 2011].

4.2.4 Exploring alternative explanations

This section checks that the decrease in the cognitive gender gap is not explained by differences in non-cognitive characteristics that might be affected by gender norms. One could think about differences in languages skills if women are less exposed to social contacts than men due to conservative gender norms and that mother tongue is different that the one prevailing in the country of residence. It is not possible to check that in the data, but language skills can somewhat

be proxied by the individual's need for clarification and issues to understand the interviewer's questions. Table F1 report the results taking the probability to ask for clarification (Column 1) and the probability to misunderstand questions (Column 2) as the outcome respectively. Then, considering that female mental health could be negatively affected by traditional gender role attitudes, women who are transmitted such norms could be less willing to answer to the interviewer's questions (Column 3) or more likely to declare depressive symptoms (Column 4). Results show that none of these non-cognitive characteristics are affected differently between men and women due to conservative gender norms, ruling out the possibility that the decrease in the female CSI compared to men covers other explanations.

4.3 Mediating factors

This section investigates potential mediating channels that could be at stake in the relationship between gender norms and the cognitive gender gap. We first expose potential mechanisms highlighted in the literature (Section 4.3.1) and present the results using these mediating factors as outcomes (Section 4.3.2).

4.3.1 Literature

Medical literature has documented the relationship between cognitive performance and exposure to stimulating activities during life. It shows that more educated individuals have a higher cognitive level [Le Carret et al., 2003] and are less exposed to dementia risk later in life [Lövdén et al., 2020]. In economics, gender differences in cognitive functioning in later life have been attributed to differences in education levels between men and women [Maurer, 2011; Angrisani et al., 2020]. Some studies have identified a positive effect of education on cognitive functioning later in life using compulsory schooling reforms in England [Banks and Mazzonna, 2012]. Those findings are corroborated using the same empirical strategy in European countries [Crespo et al., 2014; Schneeweis et al., 2014] and in the United States [Glymour et al., 2008].

Then, cognitive functioning has been shown correlated to lifestyle. Fratiglioni et al. [2004] provide a literature review of medical studies describing a positive association between an active and socially stimulating lifestyle and cognitive performances among older individuals. Adam et al. [2007] corroborate this finding showing that the practice of professional, physical and social

Table 3: Investigation of mediating factors : Education and labour supply

	Both parents are born abroad			
	No education	> High school	Never worked	White collar occupation
Woman	0.011 (0.013)	-0.492*** (0.159)	0.027 (0.056)	-0.042 (0.194)
Gender Norms Index * Woman	0.017 (0.016) [0.495]	-0.036 (0.032) [0.325]	-0.035 (0.025) [0.366]	-0.078** (0.036) [0.033]
Country of ancestry FE	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes
Age, age squared	yes	yes	yes	yes
R^2	0.078	0.149	0.215	0.123
Number of ancestry countries	28	28	28	28
Number of individuals	1,001	1,001	1,001	961
Number of observations	2,509	2,509	2,509	2,433

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

activities is positively associated with cognitive reserve. However, these studies do not allow to conclude on whether these characteristics are causes or consequences of cognitive development as they do not correct for the endogenous relationship between cognitive functioning and the practice of such activities.

Regarding professional activity, several studies in economics have documented the effect of retirement on cognitive functioning. While [Bonsang et al. \[2012\]](#) find a negative effect, suggesting that employment is beneficial for cognitive development, several studies explore the heterogeneous effect of retirement on cognition and find a negative effect for higher skilled workers and a positive one for lower skilled workers [[Coe et al., 2012](#); [Celidoni et al., 2017](#); [Mazzonna and Peracchi, 2017](#)]. This result implies that occupation contributes to cognitive functioning only if is stimulating. Interestingly, [Atalay et al. \[2019\]](#) find a negative effect of retirement on cognitive functioning only for men.

4.3.2 Analyses

These hypothetical mechanisms are explored using the mediating variables described in Section 3.3 as outcomes. We first explore educational attainment and professional activity as potential channels using the probability to have no education, the probability to have completed at least secondary school, the probability to have never worked and the probability to have, or have had, a *white-collar* occupation conditional on having ever worked. The first two columns of Table 3 show that gender differences in educational attainment are not likely to underlie the relationship between gender norms in the country of ancestry and the cognitive gender gap in the country of residence. Women are not more likely to have never worked during their life than men when gender norms are more conservative in the country of ancestry (Column 3), but they are less likely than men to have or have had a *white-collar* occupation if they have ever worked by almost 7.8 percentage points when the GNI increases by 1 sd, representing a decrease of 2% with respect to the mean (Column 4).¹¹ Note that the proportion of individuals without any education as well as the proportion of individuals who have always been out of the labour market are very low, which may contribute to the absence of results (Table 1).

We also explore non-professional activities as potential mediators as active lifestyle has been shown positively correlated to cognitive functioning in the literature. For each activity listed in Section 3.3, we use the probability to have conducted the activity at least once in the past 12 months as the outcome. As observed in Table 4, there is no significant evidence of gender differences in non-professional activity related to gender norms in the country of ancestry.

4.4 Further findings

This section aims at investigating potential implications of our results in terms of health. In particular, we explore whether gender norms contribute to limitations with instrumental activities of daily living (IADL). To this aim, we take as the outcome a binary variable equal to 1 if the individual declares at least one limitation among the following IADLs : bathing, eating, getting in or out of bed, going to the toilet, cooking, shopping, doing housework, finding one's way in an unknown place, using a telephone, managing medication consumption and managing money,

¹¹With g =second-generation immigrants with both parents born in the same country, $\frac{\hat{\alpha}_2^g * GNI_{sd}^g}{WhiteCollar_{mean}^g} = \frac{-0.078 * 0.102}{0.422} = -0.019$. GNI_{sd}^g values are available in Table C1 and $WhiteCollar_{mean}^g$ in Table 1.

Table 4: Investigation of mediating factors : Non-professional activities

	Both parents are born abroad						
	Voluntary work	Training course	Sports or social club	Religious activity	Political activity	Reading books	Playing games
Woman	0.063 (0.127)	0.013 (0.079)	-0.054 (0.134)	0.154* (0.084)	-0.115* (0.069)	-0.004 (0.067)	-0.025 (0.137)
Gender Norms Index * Woman	0.006 (0.022) [0.836]	-0.009 (0.016) [0.597]	-0.002 (0.029) [0.945]	0.024 (0.029) [0.576]	-0.014 (0.020) [0.350]	-0.011 (0.024) [0.719]	0.047 (0.034) [0.231]
Country of ancestry FE	yes	yes	yes	yes yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes yes	yes	yes	yes
Age, age squared	yes	yes	yes	yes yes	yes	yes	yes
R^2	0.096	0.103	0.094	0.111	0.046	0.125	0.092
Number of ancestry countries	28	28	28	28	28	28	28
Number of individuals	997	997	997	609	997	997	997
Number of observations	2,494	2,494	2,494	948	2,494	2,494	2,494

Note: OLS estimates, with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born in the same country.

Source : Survey of Health, Ageing and Retirement in Europe (Waves 1-8).

0 otherwise. We construct a similar binary variable for each IADL item. Table F2 reports the results taking these variables as outcomes and shows that women are more likely to declare at least one limitation than men by 8 ppt ($p < 0.01$) when gender norms are more conservative by 1 sd. This result stems from the fact that women are more likely to declare space orientation problems (i.e. finding one's way) by 3 ppt ($p < 0.05$) and problems with telephone use by about 2 ppt ($p < 0.1$) compared to men.

5 Conclusion

The aim of this study is to explore how gender norms are associated with the gender gap in cognitive functioning among individuals aged 50 and above. Because norms and institutions are entangled, gender differences in cognitive skills are measured among second-generation immigrants in the country of residence and gender norms are proxied at the parental country-of-birth level (also called country of ancestry). Using the OLS estimator, we investigate how the gender gap in cognitive test score changes among second-generation immigrants in the country of residence when gender norms are more conservative in the country of ancestry.

We find that female cognitive skills decrease relatively to male ones when gender norms are more conservative among individuals whose both parents are born in the same country. Our results are robust to using alternative measures for gender norms, to controlling for individual as well as country-of-ancestry level characteristics and to further sample restrictions. In line with previous literature, we do not find any significant relationship for individuals with only one parent born abroad, suggesting that inter-generational transmission is stronger among homogamous couples. We explore the potential underlying mechanisms of this relationship and find that gender differences in the type of occupation are likely to be at stake while differences in education do not appear as an explanation in our case. Moreover, this gender differential in cognitive functioning could result in higher functional disability for women since they declare more limitations with IADLs than men when gender norms are more traditional.

We acknowledge several limitations to our results. We are not able to conduct any further analysis due to the limited number of observations in our samples. Using a fraction of the population such as a sample of second-generation immigrants necessarily conducts to some constraints in terms of empirical analysis. Then, considering selection into migration with respect to gender norms (i.e. if individuals migrate because they do not agree with gender norms in their country of birth) implies that norms prevailing on the country of ancestry might not be transmitted to children. In this case, our results would be underestimated. Then, the empirical strategy we build on implies that our results are specific to the population of second-generation immigrants. Nonetheless, Table A2 shows that they do not significantly differ from individuals with native parents with respect to observable characteristics. Of course, we cannot discard the fact that they might have different characteristics that are not observed in the data. In any case, this strategy is, to our knowledge, the best suited to our research question. Investigating the contribution of gender norms to gender gaps in the general population is left for future research.

Three main conclusions can be drawn from this paper. First, focusing on second-generation immigrants highlights the persistence of traditional gender norms through generations which, in addition to disadvantage women on the labour market [Antecol, 2000; Fernández and Fogli, 2009; Blau et al., 2013; Finseraas and Kotsadam, 2017], in competitive situations [Hauge et al., 2023] and to expose them to intimate partner violence [González and Rodríguez-Planas, 2020],

also contribute to deteriorate their health. Second, conservative gender norms are not only related to the cognitive gender gap early in life [Nollenberger et al., 2016; Rodríguez-Planas and Nollenberger, 2018; Dossi et al., 2021], they might also have long-term implications at older ages. Third, this paper is an additional evidence in the literature that social norms matter to individual economic outcomes, implying that they should not be disregarded by economists.

Bibliography

- Adam, S., Bonsang, É., Germain, S., and Perelman, S. (2007). Retraite, activités non professionnelles et vieillissement cognitif. Une exploration à partir des données de SHARE. *Économie et Statistique*, (403-404). [195]
- Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944. [15, 176]
- Angrisani, M., Lee, J., and Meijer, E. (2020). The gender gap in education and late-life cognition: Evidence from multiple countries and birth cohorts. *The Journal of the Economics of Ageing*, 16:100232. [195]
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. [50, 57, 180]
- Antecol, H. (2000). An examination of cross-country differences in the gender gap in labor force participation rates. *Labour Economics*, 7(4):409–426. [103, 125, 177, 191, 199]
- Atalay, K., Barrett, G. F., and Staneva, A. (2019). The effect of retirement on elderly cognitive functioning. *Journal of Health Economics*, 66:37–53. [196]
- Banks, J. and Mazzonna, F. (2012). The effect of education on old age cognitive abilities: Evidence from a regression discontinuity design. *The Economic Journal*, 122(560):418–448. [195]
- Bisin, A. and Verdier, T. (2011). The economics of cultural transmission and socialization. In *Handbook of Social Economics*, volume 1, pages 339–416. Elsevier. [194]
- Blau, F. D., Kahn, L. M., Liu, A. Y.-H., and Papps, K. L. (2013). The transmission of women's fertility, human capital and work orientation across immigrant generations. *Journal of Population Economics*, 26:405–435. [16, 103, 177, 179, 199]
- Bonsang, E., Adam, S., and Perelman, S. (2012). Does retirement affect cognitive functioning? *Journal of Health Economics*, 31(3):490–501. [196]
- Bonsang, E., Skirbekk, V., and Staudinger, U. M. (2017). As you sow, so shall you reap: Gender-role attitudes and late-life cognition. *Psychological Science*, 28(9):1201–1213. [126,

176, 185]

- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427. [180]
- Celidoni, M., Dal Bianco, C., and Weber, G. (2017). Retirement and cognitive decline. A longitudinal analysis using SHARE data. *Journal of Health Economics*, 56:113–125. [196]
- Coe, N. B., von Gaudecker, H.-M., Lindeboom, M., and Maurer, J. (2012). The effect of retirement on cognitive functioning. *Health Economics*, 21(8):913–927. [196]
- Crespo, L., López-Noval, B., and Mira, P. (2014). Compulsory schooling, education, depression and memory: New evidence from SHARELIFE. *Economics of Education Review*, 43:36–46. [195]
- Dossi, G., Figlio, D., Giuliano, P., and Sapienza, P. (2021). Born in the family: Preferences for boys and the gender gap in math. *Journal of Economic Behavior & Organization*, 183:175–188. [11, 200]
- Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177. [16, 103, 125, 177, 186, 191, 199]
- Finseraas, H. and Kotsadam, A. (2017). Ancestry culture and female employment—an analysis using second-generation siblings. *European Sociological Review*, 33(3):382–392. [177, 191, 199]
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). “Mini-Mental State”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric research*, 12(3):189–198. [181]
- Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3):416–438. [126, 185]
- Fratiglioni, L., Paillard-Borg, S., and Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. *The Lancet Neurology*, 3(6):343–353. [195]
- Giuliano, P. (2020). Gender and culture. *Oxford Review of Economic Policy*, 36(4):944–961. [176]

- Glymour, M. M., Kawachi, I., Jencks, C. S., and Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology & Community Health*, 62(6):532–537. [195]
- González, L. and Rodríguez-Planas, N. (2020). Gender norms and intimate partner violence. *Journal of Economic Behavior & Organization*, 178:223–248. [177, 186, 192, 199]
- Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender and math. *Science*, 320(5880):1164–1165. [11, 185, 192]
- Hauge, K. E., Kotsadam, A., and Riege, A. (2023). Culture and gender differences in willingness to compete. *The Economic Journal*, page uead033. [177, 185, 191, 192, 199]
- Hussenoeder, F. S., Conrad, I., Roehr, S., Fuchs, A., Pentzek, M., Bickel, H., Moesch, E., Weyerer, S., Werle, J., Wiese, B., et al. (2020). Mild cognitive impairment and quality of life in the oldest old: A closer look. *Quality of Life Research*, 29:1675–1683. [176]
- Le Carret, N., Lafont, S., Letenneur, L., Dartigues, J.-F., Mayo, W., and Fabrigoule, C. (2003). The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. *Developmental Neuropsychology*, 23(3):317–337. [195]
- Lei, X., Hu, Y., McArdle, J. J., Smith, J. P., and Zhao, Y. (2012). Gender differences in cognition among older adults in China. *Journal of Human Resources*, 47(4):951–971. [176]
- Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., and Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest*, 21(1):6–41. [195]
- Maurer, J. (2011). Education and male-female differences in later-life cognition: International evidence from Latin America and the Caribbean. *Demography*, 48(3):915–930. [176, 195]
- Mazzonna, F. and Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources*, 52(1):128–151. [196]
- McGuire, L. C., Ford, E. S., and Ajani, U. A. (2006). Cognitive functioning as a predictor of functional disability in later life. *The American Journal of Geriatric Psychiatry*, 14(1):36–42. [176]
- Nollenberger, N., Rodríguez-Planas, N., and Sevilla, A. (2016). The math gender gap: The role

- of culture. *American Economic Review*, 106(5):257–261. [11, 177, 192, 200]
- Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., and Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia*, 9(1):63–75. [176]
- Rodríguez-Planas, N. and Nollenberger, N. (2018). Let the girls learn! It is not only about math... it's about gender social norms. *Economics of Education Review*, 62:230–253. [11, 177, 179, 186, 192, 200]
- Schneeweis, N., Skirbekk, V., and Winter-Ebmer, R. (2014). Does education improve cognitive performance four decades after school completion? *Demography*, 51(2):619–643. [195]
- Seguino, S. (2000). Gender inequality and economic growth: A cross-country analysis. *World Development*, 28(7):1211–1230. [194]
- Shao, Z., Janse, E., Visser, K., and Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5:772. [182]
- United Nations (2017). World population prospects: The 2008 revision. URL: <http://esa.un.org/unpp> (gelesen am 16, 2010. accessed: may 19. [176]
- Vecchio, N., Fitzgerald, J. A., Radford, K., and Fisher, R. (2016). The association between cognitive impairment and community service use patterns in older people living in Australia. *Health & Social Care in the Community*, 24(3):321–333. [176]
- Weber, D., Skirbekk, V., Freund, I., and Herlitz, A. (2014). The changing face of cognitive gender differences in Europe. *Proceedings of the National Academy of Sciences*, 111(32):11673–11678. [176]
- Wimo, A., Jönsson, L., Bond, J., Prince, M., Winblad, B., and International, A. D. (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia*, 9(1):1–11. [176]
- Yount, K. M. (2008). Gender, resources across the life course and cognitive functioning in Egypt. *Demography*, 45(4):907–926. [176]

Appendix

A Descriptive statistics

Table A1: Number of individuals by country of ancestry

Country of birth	Both parents	Only mother	Only father
Algeria	28	28	29
Argentina	0	31	14
Australia	0	10	0
Austria	65	240	145
Belgium	38	160	102
Bosnia	52	14	36
Brazil	0	0	18
Bulgaria	0	0	25
Belarus	30	14	28
China	0	0	15
Canada	0	25	0
Croatia	65	87	97
Czech Republic	19	108	127
Denmark	13	46	88
Estonia	0	15	0
Finland	19	157	45
France	26	397	265
Germany	164	813	457
Hungary	54	52	120
Indonesia	35	30	15
Italy	439	231	379
Latvia	0	39	40
Lithuania	15	25	14
Luxembourg	0	43	22
Macedonia	0	0	10
Moldova	10	0	0
Montenegro	0	0	10
Netherlands	39	116	102
Norway	0	87	46
Poland	188	246	291
Portugal	19	14	22
Romania	49	41	46
Russia	807	566	511
Serbia	15	21	26
Slovakia	44	69	57
Slovenia	89	119	87
Spain	95	34	79
Switzerland	0	76	55
Tunisia	14	0	10
Turkey	27	18	40
Ukraine	51	101	124
United Kingdom	0	51	33
United States of America	0	85	67
Vietnam	0	20	0
Total	2,509	4,229	3,697

Note : Number of individuals by second-generation immigrants group. Countries with less than 10 individuals are excluded (see Section 3.1). As a robustness check, countries of ancestry with more than 100 individuals are excluded (see Section 4.2.3). Source : Survey of Health, Ageing and Retirement in Europe (Waves 1-8).

Table A2: External validity : Mean comparison tests

	Second-generation immigrants (1)	Individuals with native parents (2)	Difference in means (1) - (2)
<i>Main outcomes</i>			
Immediate word recall score	5.635 (1.673)	5.309 (1.779)	0.326 (0.394)
Delayed word recall score	4.346 (2.139)	3.936 (2.175)	0.409 (0.500)
Verbal fluency score	22.148 (7.230)	20.634 (7.763)	1.514 (2.519)
Numeracy score	3.591 (0.988)	3.466 (1.050)	0.125 (0.316)
Subtraction score	4.467 (0.944)	4.300 (1.105)	0.166 (0.291)
<i>Covariates</i>			
Age	65.989 (9.331)	67.705 (9.700)	-1.716* (1.129)
Woman (%)	0.570 (0.495)	0.557 (0.497)	0.012 (0.028)
<i>Mediating variables</i>			
No education (%)	0.010 (0.100)	0.038 (0.191)	-0.027 (0.052)
At least secondary education (%)	0.333 (0.471)	0.272 (0.445)	0.061 (0.100)
Never worked (%)	0.025 (0.155)	0.059 (0.235)	-0.034 (0.058)
White collar worker (%)	0.500 (0.500)	0.476 (0.499)	0.024 (0.090)
Non-professional activity (%)	0.919 (0.272)	0.859 (0.348)	0.061 (0.123)
Number of countries of ancestry	27	44	48
Number of observations	10,435	155,317	165,752

Note: Columns 1 and 2 report means with standard deviations in parentheses for second-generation immigrants and individuals whose parents are natives respectively. The third column reports the difference in mean values between the two groups with standard errors in parentheses. The proportion of white-collar worker is calculated among individuals who ever worked. Samples : All second-generation immigrants (Col. 1) and individuals whose parents are natives (Col. 2). Source : Survey of Health, Ageing and Retirement in Europe (Waves 1-8).

B Main results by sub-group

We report results by group of second-generation immigrants for each cognitive test score and the Cognitive Score Index. Results are commented in Section 4.1.

B.1 Methodology

For each group of second-generation immigrants g (i.e. both parents are born abroad, only the mother is born abroad or only the father is born abroad), we run the following regression (OLS) :

$$C_{isrta}^g = \alpha_0^g + \alpha_1^g Female_i^g + \alpha_2^g Female_i^g * GNI_a^g + \gamma_a^g + \gamma_r^g + \gamma_{rs}^g + \delta_t^g + \rho^g X_{it}^g + \epsilon_{isrta}^g \quad (4)$$

For each g , the specification is similar to the one described in Equation 1. Results are available in Section B.2.

B.2 Results

Table B1: Estimation results - Both parents are born abroad

	Both parents are born abroad					
	Immediate word recall	Delayed word recall	Word fluency	Serial subtraction	Numeracy	Cognitive Score Index
Woman	0.139 (0.221)	0.246 (0.244)	0.209 (0.313)	-0.258 (0.196)	0.020 (0.282)	0.103 (0.244)
Gender Norms Index*Woman	-0.098* (0.054) [0.037]	-0.128** (0.055) [0.028]	-0.127** (0.058) [0.115]	-0.097 (0.068) [0.219]	-0.186** (0.079) [0.104]	-0.184*** (0.063) [0.031]
Country of ancestry FE	yes	yes	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes	yes	yes
R^2	0.192	0.178	0.167	0.095	0.162	0.210
Number of ancestry countries	28	28	28	28	28	28
Number of individuals	1,001	1,001	1,001	1,001	1,001	1,001
Number of observations	2,509	2,509	2,509	2,509	2,509	2,509

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Cognitive tests are described in Section 3.2. Results are interpreted in Section 4.1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

Table B2: Estimation results - Only the mother is born abroad

Only the mother is born abroad						
	Immediate word recall	Delayed word recall	Word fluency	Serial subtraction	Numeracy	Cognitive Score Index
Woman	0.166 (0.178)	0.238 (0.164)	0.061 (0.165)	-0.092 (0.123)	-0.260 (0.199)	0.033 (0.178)
Gender Norms Index*Woman	-0.038 (0.046) [0.443]	0.005 (0.047) [0.913]	-0.028 (0.053) [0.595]	0.002 (0.052) [0.955]	-0.049 (0.072) [0.455]	-0.032 (0.055) [0.520]
Country of ancestry FE	yes	yes	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes	yes	yes
R^2	0.182	0.186	0.195	0.067	0.107	0.206
Number of ancestry countries	37	37	37	37	37	37
Number of individuals	1,616	1,616	1,616	1,616	1,616	1,616
Number of observations	4,229	4,229	4,229	4,229	4,229	4,229

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Cognitive tests are described in Section 3.2. Results are interpreted in Section 4.1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose only the mother is born abroad.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

Table B3: Estimation results - Only the father is born abroad

Only the father is born abroad						
	Immediate word recall	Delayed word recall	Word fluency	Serial subtraction	Numeracy	Cognitive Score Index
Woman	0.118 (0.135)	0.114 (0.157)	0.011 (0.149)	-0.303*** (0.105)	-0.206 (0.176)	-0.077 (0.161)
Gender Norms Index*Woman	-0.057 (0.051) [0.276]	-0.055 (0.053) [0.480]	0.024 (0.055) [0.693]	0.015 (0.050) [0.811]	-0.077 (0.065) [0.184]	-0.044 (0.059) [0.532]
Country of ancestry FE	yes	yes	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes	yes	yes
R^2	0.157	0.172	0.190	0.081	0.149	0.201
Number of ancestry countries	39	39	39	39	39	39
Number of individuals	1,406	1,406	1,406	1,406	1,406	1,406
Number of observations	3,697	3,697	3,697	3,697	3,697	3,697

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Cognitive tests are described in Section 3.2. Results are interpreted in Section 4.1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose only the father is born abroad.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

C Alternative gender norms proxies

C.1 Description of alternative measures

This section provides further information on alternative measures of gender norms that we use in Section 4.2.1. Table C1 reports means and standard-deviations for each of them.

Each GGI sub-index is based on several composite indicators :

- The GGI in economic participation and opportunity ($GGI_{economic}$) is based on the difference in LFP between women and men, the female-to-male ratio of estimated earned income for similar work and the women-to-men ratio among legislators, senior officials, managers and workers.
- The GGI in educational attainment ($GGI_{education}$) on female-to-male literacy rate, female-to-male enrollment rate in primary, secondary and tertiary education respectively.
- The GGI in political empowerment ($GGI_{politics}$) on the women-to-men ratio in parliament seats, the women-to-men ratio in ministerial positions, and the female-to-male ratio in number of years as a head of state.
- The GGI in health and survival (GGI_{health}) on the difference in healthy life expectancy between women and men and the sex ratio at birth.

Table C1: Description of gender norms measures

	GNI	$GG_{log(LFP)}$	$GGI_{politics}$	$GGI_{economic}$	$GGI_{education}$	GGI_{health}
<i>Both parents are born abroad</i>						
Mean	0.400	0.251	0.215	0.688	0.995	0.976
Standard-deviation	0.102	0.210	0.124	0.073	0.008	0.005
<i>Only the mother is born abroad</i>						
Mean	0.331	0.189	0.308	0.707	0.993	0.975
Standard-deviation	0.115	0.088	0.147	0.056	0.010	0.004
<i>Only the father is born abroad</i>						
Mean	0.356	0.198	0.272	0.694	0.994	0.975
Standard-deviation	0.118	0.101	0.138	0.064	0.009	0.005

Sources : Integrated Values Study + World Bank. These statistics are useful to the interpretation of results in Section 4.2.1.

C.2 Results

We present here the results using alternative gender norms proxies as a robustness check. Tables and Figures are commented in Section 4.2.1.

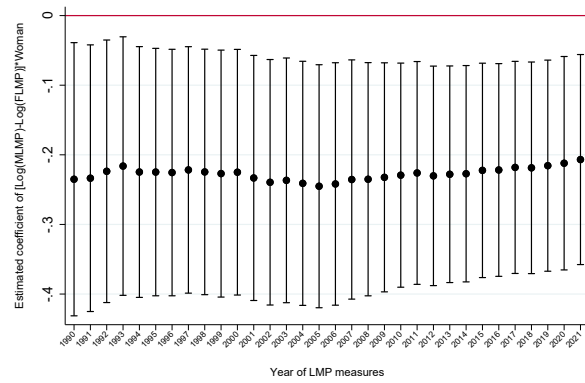
Table C2: Estimation results - Alternative gender norms proxies

	Cognitive Score Index				
	(1)	(2)	(3)	(4)	(5)
Woman	0.205 (0.241)	0.192 (0.238)	0.244 (0.248)	0.242 (0.240)	0.180 (0.238)
$[Log(MLMP) - Log(FLMP)] * \text{Woman}$	-0.150** (0.067) [0.078]				
$GGI_{politics} * \text{Woman}$	0.095 (0.090) [0.268]				
$GGI_{economics} * \text{Woman}$	0.155** (0.075) [0.067]				
$GGI_{education} * \text{Woman}$	0.101 (0.067) [0.320]				
$GGI_{health} * \text{Woman}$	0.095 (0.090) [0.268]				
Country of ancestry FE	yes	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes	yes
R^2	0.208	0.205	0.207	0.206	0.205
Number of ancestry countries	28	28	28	28	28
Number of individuals	1,001	1,001	1,001	1,001	1,001
Number of observations	2,509	2,509	2,509	2,509	2,509

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.
Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + World Bank.

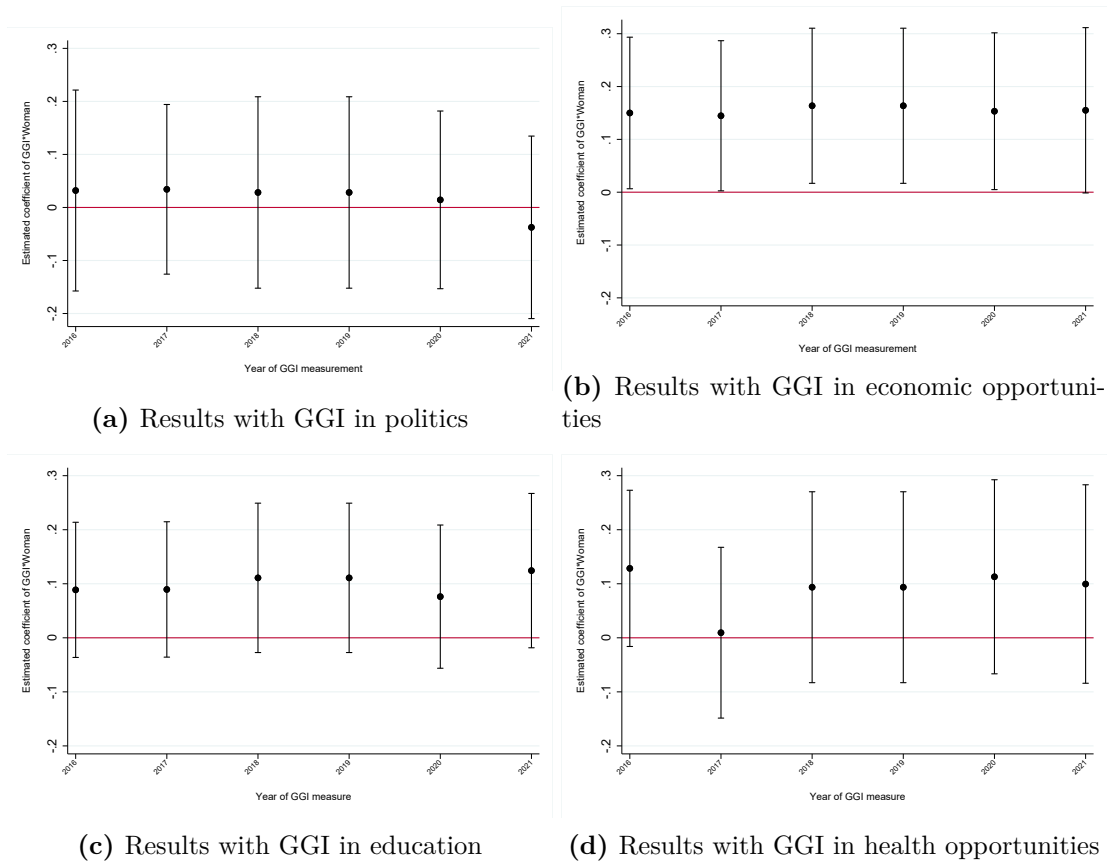
Figure C1: Estimation results by year of employment rates measurement



Note: Estimated coefficients of the interaction between the individual's sex and the gender gap in labour force participation (LFP) rate in the country of ancestry by year of LFP measurement. The gender gap is computed as the difference in $\log(\text{LFP})$ between men and women.

Source : World Bank (1990-2021).

Figure C2: Estimation results by year of GGI measurement



Note: Estimated coefficients of the interaction between the individual's sex and the Gender Gap Index (GGI) by year of GGI measurement. Sub-figure (a) (resp. (b), (c) and (d)) reports the coefficient of the interaction with the GGI in politics (resp. economic opportunities, education and health opportunities). Only years for which data is available for all countries of ancestry are used in the analyses. Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + World Bank (2016-2021).

D Additional control variables

D.1 Individual level controls

Table D1: Robustness check : Controlling for parents' levels of education

	Cognitive Score Index		
	(1)	(2)	(3)
Woman	0.059 (0.248)	0.112 (0.257)	0.075 (0.257)
Gender Norms Index * Woman	-0.176*** (0.063) [0.034]	-0.174*** (0.063) [0.035]	-0.170*** (0.063) [0.038]
Mother's level of education	0.334** (0.158)		0.191 (0.195)
Mother's level of education * Woman	0.003 (0.235)		0.032 (0.256)
Father's level of education		0.287** (0.144)	0.197 (0.176)
Father's level of education * Woman		0.005 (0.192)	0.037 (0.216)
Country of ancestry FE	yes	yes	yes
Country of residence FE	yes	yes	yes
Country of residence by sex FE	yes	yes	yes
Interview wave FE	yes	yes	yes
Age, Age squared	yes	yes	yes
R^2	0.215	0.217	0.219
Number of ancestry countries	28	28	28
Number of individuals	1,001	1,001	1,001
Number of observations	2,509	2,509	2,509

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Column 1 (resp. 2) reports estimation results controlling for maternal (resp. paternal) education level and its interaction with the individual's sex. In Column 3, both parents' levels of education are included. Results are commented in Section 4.2.2. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

D.2 Country level controls

Table D2: Correlations between country of ancestry's characteristics

	GDP per capita	HDI	Unemployment rate
GDP per capita	1		
HDI	0.938	1	
Unemployment rate	-0.131	-0.117	1

Source : World Bank. Correlations are commented in Section [4.2.2](#).

Table D3: Robustness check : Controlling for country of ancestry's characteristics

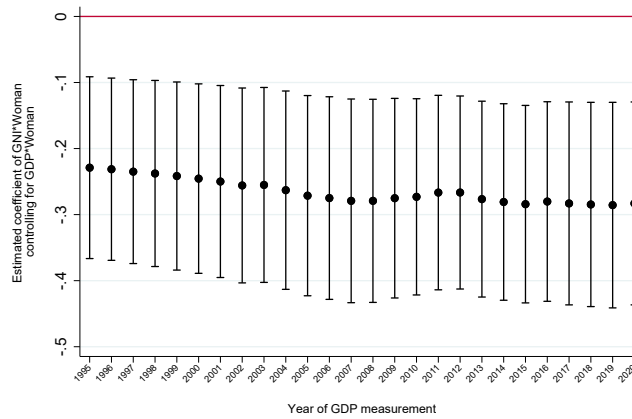
	Cognitive Score Index			
	(1)	(2)	(3)	(4)
Woman	0.168 (0.246)	0.150 (0.243)	0.102 (0.246)	0.199 (0.247)
Gender Norms Index * Woman	-0.272*** (0.077) [0.002]	-0.273*** (0.085) [0.005]	-0.185*** (0.063) [0.024]	-0.251*** (0.088) [0.019]
Gross Domestic Product per capita * Woman	-0.201** (0.098)			-0.361 (0.225)
Human Development Index * Woman		-0.172 (0.106)		0.142 (0.241)
Unemployment rate * Woman			0.005 (0.065)	-0.087 (0.077)
Country of ancestry FE	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes
R^2	0.212	0.211	0.210	0.213
Number of ancestry countries	28	28	28	28
Number of individuals	1,001	1,001	1,001	1,001
Number of observations	2,509	2,509	2,509	2,509

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Column 1 (resp. 2, 3) reports estimation results controlling for the interaction between the individual's sex and the level of Gross Domestic Product per capita (resp. the Human Development Index, the unemployment rate) at the country of ancestry level. In Column 4, all three interactions are included as controls. Results are commented in Section 4.2.2. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

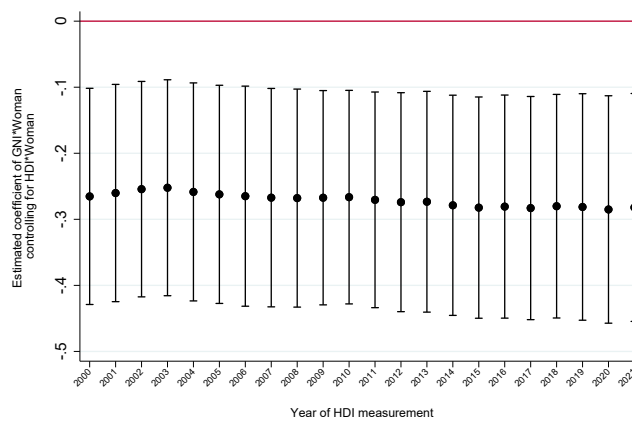
Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

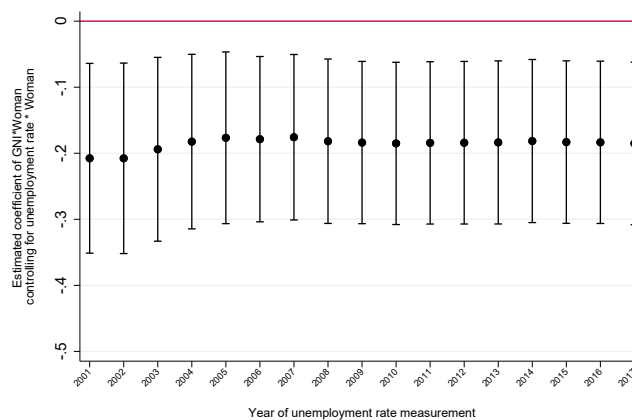
Figure D1: Estimation results by year of country level control measurement



(a) Results controlling for Gross Domestic Product per capita



(b) Results controlling for Human Development Index



(c) Results controlling for unemployment rate

Note : Estimated coefficients of the interaction between the individual's sex and the Gender Norms Index (GNI) controlling for (a) the GDP per capita, (b) the HDI and (c) the unemployment rate in the country of ancestry by year of measurement. Only years for which data is available for all countries of ancestry are used in the analyses. Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + World Bank (1995-2021).

E Sample restrictions

Table E1: Robustness check : Exclusion of largest groups among countries of ancestry

	Cognitive Score Index			
	Without Germany	Without Italy	Without Poland	Without Russia
Woman	0.329 (0.245)	-0.042 (0.285)	0.113 (0.256)	0.148 (0.252)
Gender Norms Index*Woman	-0.199*** (0.063) [0.022]	-0.249*** (0.070) [0.008]	-0.189*** (0.068) [0.086]	-0.213*** (0.075) [0.045]
Country of ancestry FE	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes
R^2	0.215	0.200	0.217	0.293
Number of ancestry countries	27	27	27	27
Number of individuals	942	837	936	692
Number of observations	2,345	2,070	2,321	1,702

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Column 1 (resp. 2,3 and 4) reports estimation results excluding Germany (resp. Italy, Poland and Russia) among countries of ancestry. Results are commented in Section 4.2.3. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

F Further results

Table F1: Robustness check : Check for alternative explanations

	Both parents are born abroad			
	Asks for clarification	Does not understand	Is unwilling to answer	Is depressed
Woman	0.063 (0.087)	0.029 (0.041)	0.028 (0.032)	0.026 (0.034)
Gender Norms Index * Woman	-0.003 (0.019) [0.859]	0.001 (0.010) [0.946]	-0.013 (0.012) [0.392]	-0.009 (0.018) [0.719]
Country of ancestry FE	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes
Age, Age squared	yes	yes	yes	yes
R^2	0.111	0.052	0.075	0.052
Outcome mean	0.195	0.049	0.075	0.128
Number of ancestry countries	28	28	28	28
Number of individuals	1,001	1,001	1,001	993
Number of observations	2,507	2,506	2,506	2,478

Note: OLS estimates with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. Results are commented in Section 4.2.4. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample : Second-generation immigrants whose both parents are born abroad and in the same country.

Sources : Survey of Health, Ageing and Retirement in Europe (Waves 1-8) + Integrated Values Surveys (1981-2020).

Table F2: Investigation health implications : Limitations with Instrumental Activities of Daily Living (IADL)

Both parents are born abroad												
	At least 1 limitation	Bathing	Eating	Getting in or out of bed	Go to toilet	Orientation	Cooking	Shopping	Using a telephone	Managing medication	Doing housework	Managing money
Woman	0.197** (0.090)	0.110* (0.060)	0.052 (0.036)	0.036 (0.027)	0.019 (0.025)	0.097 (0.077)	0.023 (0.040)	0.072* (0.040)	0.037 (0.024)	0.040* (0.022)	0.149** (0.075)	0.048 (0.043)
Gender Norms Index * Woman	0.080*** [0.053] (0.029)	0.016 [0.368] (0.016)	-0.002 [0.691] (0.005)	0.003 [0.851] (0.011)	0.002 [0.768] (0.007)	0.033** [0.068] (0.016)	0.025 [0.249] (0.016)	0.002 [0.879] (0.017)	0.018* [0.118] (0.010)	0.010 [0.407] (0.009)	0.037 [0.276] (0.026)	0.022 [0.253] (0.017)
Country of ancestry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Country of residence FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Country of residence by sex FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Interview wave FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age, age squared	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R ²	0.177	0.140	0.062	0.073	0.077	0.117	0.103	0.128	0.096	0.069	0.130	0.114
Number of ancestry countries	31	31	31	31	31	31	31	31	31	31	31	31
Number of individuals	613	613	613	613	613	613	613	613	613	613	613	613
Number of observations	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591

Note: OLS estimates, with standard errors corrected for clustering at the individual level in parentheses and wild bootstrap test p-values for clustering correction at the country of ancestry level in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Sample : Second-generation immigrants whose both parents are born in the same country aged at least 65.
Source : Survey of Health, Ageing and Retirement in Europe (Waves 1-8).

Conclusion

Despite considerable progress over decades, economic convergence between men and women has levelled off. This dissertation joins the ongoing effort of researchers in bringing empirical support to gender inequalities in the economic landscape as well as pathways to public intervention.

Among the reasons why women are still lagging behind men on the labour market, parenthood is at the forefront. After the first birth, women reallocate time from paid to unpaid work while men's trajectory remains essentially stable, leading to substantial gender gaps in time-allocation that widen with the number of children. These results, brought by the first chapter of this dissertation, call for public intervention to annihilate the contribution of parenthood to gender inequalities. Recently, public policy evaluations of parental leave schemes have provided promising results. In particular, shared parental leave is useful in shortening mothers' career breaks and in increasing their earnings on the long run [Pérvier et al., 2021; Frodermann et al., 2023]. Such settings also serve gender equality in fostering father's involvement in childcare and housework chores [Schober, 2014; Patnaik, 2019; Tamm, 2019] and in changing gender role attitudes among children [Farré et al., 2022]. If these policies are encouraging, it is nonetheless important to note that father's leave take-up remains low in some countries despite implementation of shared leave schemes with incentive settings. As part of future research, understanding fathers' reluctance to parental leave is pivotal. In particular, social stigma that prevents fathers from taking time off to take care of children may be at stake.

As part of public intervention to reduce mothers' career interruptions, childcare provision is of paramount importance. However, if childcare slots are scarce or poorly subsidised, mothers might have to find alternative solutions. Among them, informal childcare provided by relatives, usually grand-parents, is widespread but might not be enough for women to work full-time. Moreover, it requires to live near the informal care provider, when such a person exists. In this context, pre-school programs designed to foster child development before entering the schooling system can be a lever since it is usually provided free of charge as opposed to earlier childcare solutions, such as day-care. The second chapter of this dissertation addresses this question in a cross-country perspective and shows that pre-school programs help mothers resuming work in countries where investment in public early childcare is low and gender stereotypes are strong. This result highlights the need for such countries to increase their public spending in early childcare structures in order for women to be able to go back to paid work before their child is

eligible to pre-school, especially in countries with lower female workforce. For greater results, public early childcare provision and shared parental leave designs can be combined [[Geyer et al., 2015](#)].

In addition to hampering gender economic convergence over the life course, gender stereotypes can be detrimental to women's health at older ages. In the third chapter, we show that traditional gender norms affect women's cognitive functioning relative to men. This relationship can be at least partly explained by differences in type of occupation during life between men and women due to conservative gender norms. Our results have further implications in terms of quality of life since women are more likely to meet difficulties with instrumental activities of daily living than men when exposed to traditional gender norms. Despite the richness of data used for this chapter, low statistical power due to small samples does not allow us to go further in the explorations. In order to delve into the mechanisms at stake in this relationship, it would be worth replicating this analysis on larger samples as part of further research. This chapter informs on the long-run effects of traditional gender norms that go beyond the scope of usual economic outcomes. Overcoming such effects requires to get to the root of gender stereotypes. In that respect, education is an efficient tool to phase traditional gender norms out [[Rivera-Garrido, 2022](#)] and gender role models are key to shaping new aspirations for women [[Bettinger and Long, 2005](#); [Blau et al., 2010](#); [Carrell et al., 2010](#); [Breda et al., 2021](#)] that will hopefully lead to greater convergence in the future.

Bibliography

- Bettinger, E. P. and Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, 95(2):152–157. [226]
- Blau, F. D., Currie, J. M., Croson, R. T. A., and Ginther, D. K. (2010). Can mentoring help female assistant professors? Interim results from a randomized trial. *American Economic Review*, 100(2):348–352. [226]
- Breda, T., Grenet, J., Monnet, M., and Van Effenterre, C. (2021). Do female role models reduce the gender gap in science? Evidence from French high schools. [226]
- Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 125(3):1101–1144. [226]
- Farré, L., Felfe, C., Gonzalez, L., and Schneider, P. (2022). Changing gender norms across generations: Evidence from a paternity leave reform. [225]
- Frodermann, C., Wrohlich, K., and Zucco, A. (2023). Parental leave policy and long-run earnings of mothers. *Labour Economics*, 80:102296. [225]
- Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics*, 36:84–98. [45, 226]
- Patnaik, A. (2019). Reserving time for Daddy: The consequences of fathers’ quotas. *Journal of Labor Economics*, 37(4):1009–1059. [225]
- Pérvier, H., Verdugo, G., et al. (2021). *Can parental leave be shared?* OFCE. [225]
- Rivera-Garrido, N. (2022). Can education reduce traditional gender role attitudes? *Economics of Education Review*, 89:102261. [226]
- Schober, P. S. (2014). Parental leave and domestic work of mothers and fathers: A longitudinal study of two reforms in West Germany. *Journal of Social Policy*, 43(2):351–372. [9, 45, 225]
- Tamm, M. (2019). Fathers’ parental leave-taking, childcare involvement and labor market participation. *Labour Economics*, 59:184–197. [9, 45, 225]

List of Tables

Chapter 1	34
1 First-stage and reduced-form estimates	63
2 Effect of family size on time-allocation	64
3 Effects of family size on paid work: intensive and extensive margins	65
4 Effect of family size on well-being	70
A1 Mean comparison test on observable characteristics between treated and control individuals	81
A2 Means comparison test on time-allocation between treated and control individuals before the third birth	82
A3 Mean comparison test between parents of two girls and parents of two boys as first-borns	83
B1 Pre-birth characteristics	84
B2 Descriptive statistics among couples with at least two children	85
C1 Relative gender gap in first birth penalty : time-allocation	87
C2 Relative gender gap in first birth penalty : well-being	88
E1 Robustness checks	89
E2 Robustness check : Parental leave	89
E3 Testing the over-identification assumption : Time-allocation	90
F1 Effects of family size on leisure time: sleep or hobbies ?	91
F2 Effects of family size on unpaid work time	91

F3	Effect of family size on intra-household time-allocation ratio	92
----	--	----

Chapter 2 **96**

1	Descriptive statistics	117
2	Effect of pre-school enrollment on mothers' labour supply	119
3	Effect of pre-school enrollment on informal childcare use	122
4	Summary of heterogeneous results at the individual level	123
5	Summary of heterogeneous results at the country level	126
A1	Early childcare arrangements and pre-school characteristics by country	138
B1	Descriptive statistics by country	140
C1	Continuity of individual characteristics	141
D1	Effect of pre-school enrollment on mothers' labour supply with individual covariates	143
D2	Effect of pre-school enrollment on informal childcare use with individual covariates	144
D3	Placebo test : Family allowances	144
E1	Heterogeneous effects on mothers' labour supply: Education level	146
E2	Heterogeneous effects on informal childcare use: Education level	147
E3	Heterogeneous effects on mothers' labour supply: Partnership status	148
E4	Heterogeneous effects on informal childcare use: Partnership status	149
E5	Heterogeneous effects on mothers' labour supply: Birth order	150
E6	Heterogeneous effects on informal childcare use: Birth order	151
F1	Heterogeneous effects on mothers' labour supply: Gender norms	158
F2	Heterogeneous effects on informal childcare use: Gender norms	159
F3	Heterogeneous effects on mothers' labour supply: In-cash family benefits	160
F4	Heterogeneous effects on informal childcare use: In-cash family benefits	161
F5	Heterogeneous effects on mothers' labour supply: In-kind family benefits	162
F6	Heterogeneous effects on informal childcare use: In-kind family benefits	163
F7	Heterogeneous effects on mothers' labour supply: Net childcare cost effort rates	164
F8	Heterogeneous effects on informal childcare use: Net childcare cost effort rates	165

Chapter 3 **170**

1	Descriptive statistics	182
2	Estimation results	190
3	Investigation of mediating factors : Education and labour supply	196
4	Investigation of mediating factors : Non-professional activities	198
A1	Number of individuals by country of ancestry	205
A2	External validity : Mean comparison tests	206
B1	Estimation results - Both parents are born abroad	208
B2	Estimation results - Only the mother is born abroad	209
B3	Estimation results - Only the father is born abroad	209
C1	Description of gender norms measures	210
C2	Estimation results - Alternative gender norms proxies	211
D1	Robustness check : Controlling for parents' levels of education	214
D2	Correlations between country of ancestry's characteristics	215
D3	Robustness check : Controlling for country of ancestry's characteristics	216
E1	Robustness check : Exclusion of largest groups among countries of ancestry	218
F1	Robustness check : Check for alternative explanations	219
F2	Investigation health implications : Limitations with Instrumental Activities of Daiy Living (IADL)	220

List of Figures

Chapter 1	34
1 Evolution of the time-use gender gap in Germany	45
2 Intra-household time-allocation by family size	56
3 First birth penalties in time-allocation	59
4 Child penalties in well-being	61
5 Heterogeneous effect of family size on maternal time allocation	68
6 Heterogeneous effect of family size on paternal time allocation	69
G1 Profile of compliers, never-takers and always-takers	94
 Chapter 2	 96
1 Effort rates of net childcare costs by country	106
2 Pre-school eligibility rule	107
3 Discontinuity in the probability of pre-school enrollment at the eligibility threshold	111
4 Graphical representation of reduced-form estimations : Maternal labour market outcomes	120
5 Graphical representation of reduced-form estimations : Informal childcare use	122
C1 Distribution of birth quarters	141
C2 Continuity of individual characteristics at the eligibility threshold	142
F1 Composite gender norms indices by country	155

F2	Gender norms index by country	156
F3	Childcare characteristics by country and year	157
G1	Profile of compliers, never-takers and always-takers	167
H1	Results by country	168
 Chapter 3		170
1	Gender Norms Index by country of ancestry	184
2	Association between the Gender Norms Index and the Cognitive Gender Gap	188
C1	Estimation results by year of employment rates measurement	212
C2	Estimation results by year of GGI measurement	213
D1	Estimation results by year of country level control measurement	217

RÉSUMÉ

Au cours du XXe siècle, la place des femmes dans l'économie occidentale a connu de profonds changements. Cependant, d'importants écarts liés au genre persistent aujourd'hui sur le marché du travail et d'autant plus dans la sphère domestique. Cette thèse propose un éclairage sur ces inégalités avec trois analyses empiriques. Le premier chapitre montre que la parentalité constitue un facteur aggravant de l'inégale répartition des tâches au sein du couple en Allemagne. Après la première naissance, les femmes réallouent du temps de travail sur le marché au profit de tâches non-rémunérées alors que l'allocation du temps des pères reste pratiquement inchangée. Au-delà de l'entrée dans la parentalité, le nombre d'enfants intensifie également ces écarts pré-existants. Les femmes déclarent une baisse de satisfaction avec les tâches non-rémunérées de moindre ampleur par rapport aux hommes alors qu'elles portent majoritairement la charge domestique. Ce résultat illustre ainsi un ancrage profond des normes de genre au sein du couple. Le deuxième chapitre évalue l'effet de l'entrée à l'école pré-élémentaire sur l'offre de travail des mères dans une perspective comparative entre pays européens. Les résultats de cette analyse montrent que les programmes éducatifs précoces représentent un levier efficace au travail des mères dans les pays où les normes de genre conservatrices dominent ainsi que dans les pays où la part des dépenses publiques allouée à l'offre de modes de garde intervenant avant l'école pré-élémentaire est plus faible. Le troisième chapitre s'intéresse à la transmission inter-générationnelle des normes de genre et son implication à long terme dans les écarts de santé cognitive entre les femmes et les hommes. Parmi les individus âgés d'au moins 50 ans en Europe, la transmission de normes de genre traditionnelles affecte négativement le fonctionnement cognitif des femmes par rapport aux hommes. Ce résultat s'explique au moins partiellement par des différences de parcours professionnel. Les résultats avancés dans cette thèse appuient la nécessité de l'intervention publique dans la réduction de ces inégalités, tant par une offre de mode de garde accessible et abordable afin d'annihiler la contribution de la parentalité aux inégalités femme-homme que dans l'évolution des normes de genre qui, aujourd'hui encore, empêchent la convergence des femmes et des hommes dans le paysage économique.

MOTS CLÉS

Inégalités liées au genre, Normes, Parentalité, Politiques familiales

ABSTRACT

Throughout the 20th century, women's role in the Western economy underwent profound changes. However, significant gender gaps persist nowadays on the labour market and even more in the domestic sphere. This thesis sheds light on these inequalities with three empirical analyses. The first chapter shows that parenthood emphasizes the unequal division of tasks at the couple level in Germany. After the first birth, women reallocate market working time to unpaid tasks while the allocation of fathers' time remains virtually unchanged. Beyond entry into parenthood, the number of children also intensifies these pre-existing gaps. By exploring the effect of parenthood on parental well-being, this chapter shows that women report a lower drop in satisfaction with unpaid tasks than men after the first birth, even though they bear the major load of domestic work. This result thus illustrates a deep ingraining of gender norms at the couple level. The second chapter evaluates the effect of pre-school enrollment on maternal labour supply in a comparative perspective between European countries. The results of this analysis show that early education programs represent an effective leverage for mothers' labour force in countries where gender norms are more conservative as well as in countries where the share of public expenditure allocated to early childcare provision is lower. The third chapter focuses on the inter-generational transmission of gender norms and its long-term implication in cognitive health gaps between women and men. Among individuals aged at least 50 in Europe, the transmission of traditional gender norms negatively affects the cognitive functioning of women relatively to men. This result is at least partially explained by differences in professional trajectories over the life course. The results put forward in this thesis support the need for public intervention in reducing these inequalities, both through the provision of accessible and affordable childcare in order to eliminate the contribution of parenthood to gender inequalities and through the evolution of gender norms which still prevent the convergence of women and men in the economic landscape.

KEYWORDS

Gender inequalities, Norms, Parenthood, Family policies