Three Essays on Gender and Inequality

Césarine Boinet

A thesis presented in fulfillment of the requirements
for the degree of Doctor of Philosophy

Department of Economics
University of Strathclyde
September 19, 2025

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

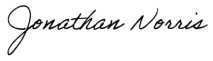
Date: September 19, 2025

Signed: 

Statement of Contribution of Co-Authors

Chapters 1 to 3 consist either of a papers submitted for publication, or working papers during the period of candidature. The undersigned hereby certify that:

1. They meet the criteria for co-authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the working paper in their field or expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the working paper;
3. There are no other authors of the publication according to these criteria and;
4. There are no potential conflicts of interest that require to be disclosed to (a) granting bodies, (b) the editor or publisher of any journal or other publication, and (c) the head of the responsible academic unit.

Chapter 1			
Contributor	Contribution	Contributor's Signature	Publication *Title*, Date or Status
C. Boinet	Conceptualization Empirical analysis Co-writing		*The Motherhood Penalty: Gender Norms, Comparative Advantage, and Occupational Sorting* Previously released as IZA Discussion Paper No. 17334 under a different title: *The Motherhood Penalty: Gender Norms, Occupational Sorting, and Labor Supply*
J. Norris	Conceptualization Supervision Analytical support Co-writing		
A. Romiti	Conceptualization Supervision Analytical support Co-writing		

Chapter 2			
Contributor	Contribution	Contributor's Signature	Publication *Title*, Date or Status
C. Boinet	Conceptualization Data cleaning Data analysis Co-writing		*Beliefs on Children's Human Capital Formation and Mothers at Work* Released as IZA Discussion Paper No. 17574
J. Norris	Conceptualization Survey implementation Data analysis Supervision Co-writing		
A. Romiti	Conceptualization Data analysis Supervision Co-writing		
Z. Shi	Conceptualization Survey coding and implementation Data collection Data cleaning		
P. Telemo	Conceptualization Data analysis Theoretical framework Co-writing		

Chapter 3			
Contributor	Contribution	Contributor's Signature	Publication *Title*, Date or Status
C. Boinet	Conceptualization Data cleaning Empirical analysis Co-writing		*Universal No More? Poverty and Child Care Consequences of Means-Testing*
L. Panico	Conceptualization Supervision		
A. Romiti	Conceptualization Supervision Analytical support Co-writing		

Acknowledgements

Completing this thesis would not have been possible without the support, encouragement, and occasional restoration of my self-confidence. I am deeply grateful to the people who provided me with these essential pillars, making my PhD journey significantly smoother.

First and foremost, I would like to thank my supervisors, Dr. Agnese Romiti and Dr. Jonathan Norris. Their unwavering support, encouragement, and thoughtful, constructive feedback consistently pushed my work toward higher standards. Beyond their academic support, their kindness and understanding provided the emotional strength I needed to persevere through the challenges of this PhD journey.

I would also like to thank the Institute for Inspiring Children's Futures for their financial support and for providing me with the incredible opportunity to engage with interdisciplinary research and wonderful colleagues. The writing retreats, seminars, and insightful conversations were invaluable — sparking new ideas and broadening my perspective. A special thank you to Jen for her support and encouragement throughout this journey.

I am equally grateful to the Department of Economics at the University of Strathclyde for providing the resources and a supportive environment that enabled me to pursue my research in the best conditions. In addition, the time and feedback offered by faculty members, as well as all the support provided by Pr. Julia Darby as Director of Postgraduate Research.

I would also like to extend a special thanks to the Centre for Research on Social Inequalities (CRIS), which hosted me from June to December 2024. In particular, I am deeply grateful to Professor Lidia Panico for her trust in me, her appreciation of my research, and for offering me the incredible opportunity to continue working alongside her and learning from her expertise. I sincerely hope this marks the beginning of a long and fruitful collaboration.

To my fellow PhD colleagues — Annie, Arnold, Baland, Ben, Bethany, Geoff, Grant, Jon, Joseph, Kat, Lateef, Rory, Sabin, Sam, Slawek, Yuri, and Zhan — thank you for your camaraderie, support, and for making this journey both pleasant and memorable.

I would also like to deeply thank my friends from Glasgow: Meg, Cairan, Becca, Toni, Ainsley, Cosmin, Rachel... A PhD journey can feel like a long and arduous hike, but with the right people by your side, every step becomes more manageable and the journey far more enjoyable.

Enfin, à mes amis, collègues et à ma famille : un immense merci. À ma famille, d'abord, pour le soutien indéfectible que j'ai reçu, même à distance. À vous, la "Klébar style" — quel chemin parcouru depuis les bancs dauphinois ! À toi, Julie, pour ton

appui sans faille depuis plus de dix ans : pour m'avoir encouragée à vivre cette aventure à l'étranger, pour m'avoir soutenue dans les moments de doute, et pour m'avoir poussée à toujours viser plus haut. Enfin, cette thèse n'aurait sans doute pas vu le jour sans le soutien précieux de mes collègues de France Stratégie, au cours des derniers mois de rédaction, que je remercie sincèrement.

Pour Sibelle...

Abstract

This thesis explores how gender norms and public policies shape maternal labor supply, occupational sorting, and investments in children. Across three empirical chapters, it examines the individual- and household-level mechanisms through which beliefs and institutional constraints affect economic behaviors, using evidence from the United Kingdom (UK) and France. Collectively, these three chapters contribute to the literature on labor economics, gender, and child development by uncovering the persistent influence of norms, beliefs, and economic policy on women's labor market trajectories and family choices.

The first chapter, *The Motherhood Penalty: Gender Norms, Comparative Advantage, and Occupational Sorting* investigates how pre-birth gender norms influence women's labor market outcomes around motherhood in the UK. Using longitudinal data from the British Household Panel Survey, we quantify motherhood penalties in earnings and hours worked, showing that they are significantly larger for women who hold more traditional gender norms prior to the first childbirth. Our analysis further reveals that occupational sorting plays a central role in mediating these penalties. Traditional mothers are more likely to sort into occupations that allow for greater reductions in working hours after childbirth, whereas egalitarian mothers' labor market outcomes are more strongly influenced by comparative advantage — proxied by hourly wage differentials — within the household. Occupational sorting accounts for 80% of the short-run earnings penalty gap between the two groups, and fully explains the difference in hours worked. These findings highlight the significance of anticipatory occupational choices, shaped by gender norms, in driving post-motherhood labor market trajectories.

The second chapter, *Beliefs on Children's Human Capital Formation and Mothers at Work* investigates how beliefs about gender roles in child-rearing influence perceptions of maternal employment and, in turn, may constrain women's labor supply. Using a survey experiment that equalizes earnings potential across genders, we find that women are systematically perceived to hold an absolute advantage in child-rearing, irrespective of comparative advantage within the household. These expectations and beliefs persist even when mothers and fathers are described as having equal time available, with mothers expected to invest more time in their children's skill development. In addition, we show on the one hand that children's outcomes tend to be underestimated when mothers work full-time, but, on the other hand, that providing accurate information about actual performance leads to more accurate beliefs and attenuates concerns about negative impacts. Overall, our findings illustrate how persistent misperception about children's human capital formation under maternal

employment can reinforce gendered labor market behaviors, and point to a potential of informational interventions in shifting perceptions and, in turn, attitudes.

Finally, the third chapter, *Universal No More? Poverty and Child Care Consequences of Means-Testing*, evaluates the unintended consequences of replacing a universal child benefit with a means-tested scheme in 2015 in France. Exploiting quasi-experimental variation from income thresholds based on pre-determined income from two years prior, we assess the causal impact of benefit reductions on material conditions, labor supply, and childcare use. We find that households facing reduced benefits experienced an increase in material deprivation and subjective financial hardship. Despite these financial pressures, we find no evidence of compensatory changes in parental labor supply, including among the self-employed. However, households responded by changing their use of non-parental childcare, moving away from more expensive arrangements (*e.g.*, childminders) toward lower-cost and publicly subsidized daycare. Our findings thus indicate that even modest cuts to benefit generosity can lead to significant changes in both material conditions and the type of childcare used, even in the absence of labor supply responses.

Taken together, these three papers highlight how both informal institutions, such as gender norms and beliefs, and formal institutions, notably public policies, shape parental decisions related to labor supply, caregiving, and child-rearing. By analyzing the interplay between normative expectations, belief formation and policy design, this thesis aims to advance our understanding of the structural and behavioral mechanisms driving gender inequalities in the labor market and within households.

Contents

List of Figures

List of Tables

Introduction

This thesis explores how gender norms and public policies shape maternal employment, labor supply, occupational sorting, and investments in children. Across three empirical chapters, it examines the individual- and household-level mechanisms through which beliefs and institutional constraints affect economic behaviors, using evidence from the United Kingdom and France. Altogether, these three essays contribute to the literature on labor economics, gender, and child development by uncovering the persistent influence of norms, beliefs, and economic policy on women's labor market trajectories and family choices.

Chapter 1 is entitled *The Motherhood Penalty: Gender Norms, Comparative Advantage, and Occupational Sorting*. In this chapter, we examine the contribution of gender norms in shaping women's labor market trajectories and occupational choices around motherhood in the United Kingdom. Using longitudinal data from the British Household Panel Survey (BHPS), we are able to construct a composite score reflecting prospective parents' gender norms, reflecting their opinion about a woman's role at home capturing the interplay between work, parenthood, and family life. This enables us to conduct our analyses in three steps. First, we quantify the impact of gender norms on earnings and labor supply post-childbirth. Second, we investigate the role of pre-birth comparative advantage within couples on these outcomes. Third, we examine the interaction between occupational characteristics — accounting for their degree of family-friendliness — and pre-birth gender norms, and how these contribute to the overall motherhood penalty.

Our main results show that women traditional mothers experience a 18-percentage-point (pp) higher motherhood penalty in earnings and a 20-pp higher motherhood penalty in hours worked compared to egalitarian mothers. In addition, we find that egalitarian parents seem to value the comparative advantage mechanism to determine post-childbirth labor supply decisions, while this is not true for parents with more traditional norms. Finally, we demonstrate that accounting for occupational sorting significantly reduces the average earnings penalty for both traditional and egalitarian mothers, driven entirely by the number of hours worked for traditional women. In particular, occupational sorting explains 80% of the short-run earnings penalty gap between traditional and egalitarian mothers and eliminates the difference in hours worked penalties entirely. These results suggest that traditional women self-select into occupations more conducive to balancing family responsibilities, that in turn have a substantial impact on their earnings trajectory in response to motherhood.

Given the robustness of our main results to a range of tests, no observed shift in

gender norms as a reaction to parenthood, and the absence of mobility across occupations and industries, this chapter provides significant evidence for understanding the perseverance and heterogeneity of the motherhood penalty across countries. Indeed, norms and beliefs do not only serve as a marker of women's post-childbirth behavioral responses, they also guide their pre-birth labor market attachment. Thus, family policies aimed at reducing post-birth gaps between traditional and egalitarian parents will remain ineffective if they do not consider the influence of norms and beliefs on pre-birth occupational sorting.

Chapter 2 is entitled *Beliefs on Children's Human Capital Formation and Mothers at Work*. In this chapter, guided by a conceptual framework, we develop and conduct a survey experiment, using the online platform Prolific, to understand whether women are perceived as holding an absolute advantage in child-rearing, ultimately affecting children's human capital formation. If such beliefs exist, this could explain why some mothers may face pressure to sort out of the labor market after giving birth, and thus the remaining gender gaps. More precisely in this chapter, we follow X key steps, guided by strong and robust evidence.

First, we estimate how beliefs on children's futures outcomes change within-individuals when mothers work longer hours than fathers with an equalized earnings potential. To do so, we present randomized hypothetical vignettes to $N = 1,056$ recruited participants in which (i) we pin down the earnings potential across a mother and a father, and (ii) vary which parent works longer hours in the labor market. In each set, we iterate the wage of the parent working longer hours in exactly the same way and elicit participants' expectations on the hypothetical child's likelihood of graduating from university and their earnings rank at age 30. Our first main result (Result 1) shows that on average, participants expect worse outcomes for children when mothers work longer hours than fathers. Participants significantly reduce the expected likelihood of graduation by just under 1% and earnings rank by 0.67 percentile. While these magnitudes are not large, they demonstrate that even with the uncertainty of earnings differentials removed, beliefs on absolute advantage persist.

Second, we aggregate these expectations at the individual-level to extract an individual-specific measure of these beliefs. Notably, this enables us to assess how beliefs vary by participants' background, including information on their own-mother's employment history, and their own experience of the motherhood penalty (for women only) collected from life-history questionnaires. Our second main result (Result 2) demonstrates that the role model effect from participants seeing their own-mother work full-time when they were young fully wipes out beliefs on absolute advantage. In addition, among women, beliefs about absolute advantage are strongest for those who experienced the highest post-birth employment penalties, strengthening the case that norms relate to the degree of motherhood penalties (*cf.* Chapter 1).

Third, we study mechanisms that can give rise to variation in beliefs about absolute advantage and focus on three key dimensions that may shape this distribution of beliefs: (i) differences in preferences, (ii) differences in the productivity of time investments, and (iii) differences in resource allocation. First, parents may expect gender differences in preferences, believing that even with equal time available, men are less likely than women to allocate time toward their child(ren)'s human

capital formation. Second, parents may perceive mothers as more productive than fathers for an equal amount of time spent on child(ren)'s human capital formation. Third, parents may expect mothers to allocate more resources to skill investments. Altogether, to study whether these mechanisms matter, we introduce a new series of vignettes and randomize features across participants. We find that differences in preferences (i) are an important driver of variation in beliefs, and this is particularly true for those with strong beliefs on absolute advantage (Result 3).

Fourth, we complement our survey with an information treatment about children's performance on tests when mothers work full-time and assess how the provision of information affects beliefs, gender norms, and policy views. Using the Millennium Cohort Study (MCS), we draw statistics on children's performance when mothers work full-time *versus* part-time, and show that they perform similarly in terms of GCSE pass rates (respectively 73% vs. 75%) — a fact participants tended to underestimate. Half the participants were informed of the pass rates, leading to more accurate beliefs, a 5.2% reduction in expected behavioral problems, and a 22% shift in written responses away from the idea that full-time maternal work harms children. Gender norms also became slightly more liberal. Policy support changes were smaller but pointed toward increased backing for initiatives such as free childcare hours, especially among subgroups with stronger beliefs on absolute advantage. While overall effects on policy preferences were weak, they suggest that different forms of information delivery, such as narratives, could have greater and more lasting impacts.

This chapter provides thus strong and robust evidence on the existence of beliefs on absolute advantage, giving weight to their perceived role in women's labor market decisions post-childbirth. In particular, parents hold different beliefs about mothers' and fathers' preferences about their free time, and the time investment toward their child(ren)'s human capital accumulation. Information about how well children actually do when mothers work full-time further shifts people toward accurate perceptions and reduces the degree of harm they expect for the child. Thus, at least some degree of beliefs on absolute advantage centered around time preferences, can be malleable to removing uncertainty with information.

Chapter 3 is entitled *Universal No More? Poverty and Child Care Consequences of Means-Testing*. This chapter analyzes the behavioral and welfare effects of a major 2015 reform to France's family allowance system, which shifted from a universal to a means-tested scheme. The reform introduced two income thresholds that reduced benefits for middle- and higher-income families — halving transfers for households above the first threshold, and reducing them to a quarter for households above the second threshold. Because eligibility was based on pre-determined income from two years prior, the reform generated quasi-experimental variation in benefit reductions. We exploit this design using nationally representative panel data from the Statistiques sur les Conditions de Vie et les Ressources (SRCV), part of the EU-SILC, for the years 2010 to 2019 to identify the causal effects of this policy on financial well-being, parental labor supply, and the use of non-parental childcare.

The analysis unfolds in three steps. First, we estimate the average impact of the reform on household-level financial conditions, including both objective (*e.g.*, income, poverty and deprivation indicators) and subjective (*e.g.*, self-reported financial hard-

ship) measures. Second, we assess potential behavioral responses in parental labor supply, including employment status, hours worked, and labor force participation. Third, we investigate changes in the use of non-parental childcare, and whether the reform reallocated time spent by children in various childcare arrangements.

Our results show that reductions in family allowances significantly increased household deprivation and financial strain, especially for families experiencing the largest cuts. Households in the quarter-benefits group saw a 3.3 percentage point increase in their likelihood of being poor, and a 40% rise in self-reported financial hardship, relative to their pre-reform baseline mean. Material deprivation also increased, as affected households found it harder to cover essential needs such as food, clothing, and housing. In contrast, we find no significant changes in parental labor supply across gender, nor across any subgroup more likely to respond to income shocks, such as self-employed parents, and dual-parent households with an unemployed or inactive partner. Finally, we document behavioral responses in childcare use: families with young children reallocated time away from costlier forms of non-parental care (*e.g.*, childminders) and toward more affordable alternatives (*e.g.*, public daycare centers). These shifts were especially pronounced among the quarter-benefits group, suggesting that reduced transfers and thus tighter budgets constraints increased their sensitivity to "out-of-pocket" childcare costs.

These findings, robust to a range of tests, contribute to the literature on the unintended consequences of targeted cash transfer reductions. While the reform aimed to redirect resources toward lower-income families including single-parent households, our results suggest that even modest income losses among middle- and higher-income families can generate measurable effects on household choices, and financial well-being. In particular, the increase in material deprivation — as captured by the unability to afford essential items — underscores the importance of weighting these trade-offs against the generated fiscal savings in the context of such reforms.

Taken together, these three chapters of this thesis highlight how both informal institutions, such as gender norms and beliefs, and formal institutions, notably public policies, shape parental decisions related to labor supply, caregiving, and child-rearing. By analyzing the interplay between normative expectations, belief formation and policy design, this thesis aims to advance our understanding of the structural and behavioral mechanisms driving gender inequalities in the labor market and within households.

Chapter 1

The Motherhood Penalty: Gender Norms, Comparative Advantage, and Occupational Sorting

1.1 Introduction

Juggling motherhood and career poses a significant challenge for women. Although women have caught up and surpassed men in human capital accumulation, gender gaps in the labor market still persist (Goldin, 2014, Olivetti, Pan, and Petrongolo, 2024), and are mostly concentrated among parents (Bertrand, Goldin, and Katz, 2010, Blau and Kahn, 2017, Kleven, Landais, and Søgaard, 2019, Cortés and Pan, 2023). This well-documented phenomenon, commonly known as the *child penalty* or the *motherhood penalty*, contributes to the persistence of gender inequalities throughout the life course. Yet, the understanding of the mechanisms contributing to the varying magnitudes of the motherhood penalty — both across and within countries — remains inconclusive, whether attributed to factors such as the policy environment (*e.g.*, Kleven, Landais, and Søgaard, 2021; Lassen, 2021; Rabaté and Rellstab, 2022; Andresen and Nix, 2022; Kleven et al., 2024), or family income maximization (Costa Dias et al., 2021).[1] Potential mechanisms driving this observed pattern include women's comparative advantage in child-rearing (Becker, 1985). Alternatively, gender norms may also explain these outcomes, as they shape behavioral expectations for women and men (Seguino, 2007; Bertrand, 2011; Bursztyn, González, and Yanagizawa-Drott, 2020; Porter and Serra, 2020; Cortés et al., 2024; Boelmann, Raute, and Schönberg, 2025). An emerging finding is that the motherhood penalty persists despite the implementation of family policies (Andresen and Nix, 2023; Bailey et al., 2025; Kleven et al., 2024),[2] which would be consistent with the idea that gender norms can affect labor market choices. Traditional gender norms often suggest perceptions of women as better suited for domestic duties and child rearing regardless of their relative performance in the labor market. This influence may manifest itself in two key ways in the labor market: first, by prompting women to

[1] Kleven, Landais, and Leite-Mariante (2024) provide a comprehensive study of the child employment penalty across countries.

[2] To our knowledge, Glogowsky et al. (2025) is the only study that finds that generous family leave policies can have substantial effects on gender inequality in earnings in the German context.

reduce their involvement in the labor market post-motherhood, and second, by guiding women's occupational choices to align with these norms and beliefs. These two implications guide our research questions.

In this paper, we study how pre-birth gender norms shape women's labor market trajectories and occupational choices around motherhood using the British Household Panel Survey (BHPS) — a representative panel of the UK's population spanning over 18 years (1991–2009). We measure gender norms through a composite score derived from six questions about a woman's role at home capturing the interplay between work, parenthood, and family life. We then use the event study methodology proposed by Kleven et al. (2019) to examine four main findings about motherhood labor market penalties: (i) how pre-birth norms drive the magnitude of the motherhood penalty in earnings, hours worked and wages; (ii) the role of comparative advantages in productivity differentials within couples, and across pre-birth gender norms in driving motherhood penalties; (iii) the role of occupational characteristics in explaining motherhood penalties, and their link with pre-birth gender norms; and (iv) the dynamics of gender norms and occupational mobility around the first childbirth.

We start by examining labor market responses after the first childbirth of men and women.[3] Our first main result shows that pre-birth gender norms are important for the magnitude of the motherhood penalty in earnings and hours worked. Indeed, traditional women experience a significant 18-percentage-point (pp) higher reduction in earnings, and a 20-pp larger decrease in the number of hours worked — compared to their more egalitarian counterparts.[4] In the recent literature, Boelmann, Raute, and Schönberg (2025) find that exposure to peer women with more egalitarian norms increased post-birth labor supply among West German women, and Moriconi and Rodríguez-Planas (2021) show that gender norms significantly influence women's employment across various European Union countries. Furthermore, Mensinger and Zimpelmann (2024) model norms as scaling women's elasticity of labor supply, and predicting their post-birth employment penalties. Thus, our first result contributes to this recent evidence base supporting a conclusion that gender norms have enduring effects on women's career trajectories, notably around motherhood.

We then move further and investigate alternative drivers of penalties through relative labor market comparative advantages within couples, and their heterogeneity across gender norms. Neoclassical models (*e.g.*, Becker, 1985) suggest that families maximize household earnings, and therefore prioritize the market work of the parent with the greater comparative advantage.[5] We find that the size of the earnings

[3]While Kleven et al. (2019) also employ the BHPS for their analysis on the United Kingdom, our analytical sample slightly differs because we condition on parents working at least once pre-birth (*i.e.*, reporting at least one year of positive labor earnings) in order to characterize their pre-birth occupational characteristics.

[4]We also show that these results are robust to how we classify prospective parents' norms, not confounded by socioeconomic status, unlikely to be driven by differences in fertility patterns by gender norms, and hold across a range of additional robustness checks.

[5]In this literature, men typically have a lower labor supply elasticity compared to women due to their comparative advantage in market work, leading to gender differences in labor market attachment and gendered reactions to the first childbirth (*e.g.*, Mincer and Polachek, 1974; Akerlof and Kranton, 2000).

penalty increases for women whose partner has a comparative advantage in the labor market before the first childbirth, especially when the difference in comparative advantage within couples is large. This pattern, however, is distinctly different across pre-birth gender norms, and women with more egalitarian norms are the primary drivers of responses to comparative advantage.[6]

The literature on motivated beliefs indicates that individuals derive utility directly from holding certain beliefs, thus incentivizing them to avoid information or actions that might challenge these beliefs (Bénabou and Tirole, 2016). A theoretical application of this concept to family dynamics is discussed in Akerlof and Rayo (2020). They suggest that different families may be invested in opposite narratives, with incentives to reinforce their respective narratives. When applied to the motherhood penalty, this theory implies that families adhering to more traditional gender roles and norms have an incentive to disregard evidence contradicting these roles — such as the woman's comparative advantage. Thus, acting on this comparative advantage would harm their utility even if it improved their budget. Conversely, more egalitarian families may be invested in narratives where the woman's comparative advantage does not conflict with the narrative, making it more likely for the woman to maintain the same labor market attachment post-birth. Our results indicate that the response to comparative advantage is particularly strong for women holding more egalitarian norms. To the best of our knowledge, this finding is novel in the literature and supports the interpretation of gender norms through the theoretical framework proposed by Akerlof and Rayo (2020).

Finally, our paper links pre-birth occupational features, such as family-friendliness, with pre-birth gender norms and their relative contribution to the motherhood penalty. To our knowledge, our paper is the first to explore how these factors intersect. First, in line with findings from the literature (*e.g.*, Costa Dias, Joyce, and Parodi, 2020), we show that occupational characteristics explain about 19% of the overall gender gap in earnings that emerges after parenthood. Second, we then look at the impact of occupational sorting across gender norms, and find that accounting for occupational sorting reduces the *short-run* motherhood earnings penalty by $15pp$ for traditional mothers and $8pp$ for egalitarian mothers, with the reduction entirely driven by hours worked especially for traditional mothers.[7] Occupational sorting accounts for 80% of the earnings penalty gap between the two groups and fully eliminates the difference in hours worked penalties. These results suggest that traditional women self-select into occupations more conducive to balancing family responsibilities, that in turn have a substantial impact on their earnings trajectory in response to motherhood. We then find no evidence of mobility across occupations or industries, nor of differential changes in gender norms by gender from before to after the first childbirth. This suggests that pre-birth gender norms drive a degree of

[6]Recent evidence from the literature finds little support for a comparative advantage mechanism in explaining the motherhood earnings penalty. Siminski and Yetsenga (2022) find that comparative advantage plays little to no role in the sexual division of labor within couple households. Andresen and Nix (2022) show that controlling for measures of predetermined relative labor market productivity differences between spouses as a proxy for comparative advantage does not eliminate motherhood penalties among heterosexual couples.

[7]We focus on short-run results, as they are least likely to be affected by potential violations of the identifying assumptions (see Kleven, Landais, and Søgaard (2019) for a discussion).

occupational sorting prior to childbirth.

The evidence here has significant implications for understanding the perseverance and heterogeneity of the motherhood penalty across countries. Norms and beliefs do not only serve as a marker of women's post-childbirth behavioral responses, they also guide their pre-birth labor market attachment. Thus, family policies aimed at reducing post-birth gaps between traditional and egalitarian parents will remain ineffective if they do not consider the influence of norms and beliefs on pre-birth occupational sorting.

1.2 Data

1.2.1 The British Household Panel Survey

Our main dataset is the British Household Panel Survey (BHPS) — a nationally representative survey covering the years 1991 to 2009. Over the course of 18 years, it includes comprehensive information on a random sample of individuals, who are interviewed annually. The first wave of the panel consists of around 5,500 households and 10,300 individuals, sampled from 250 areas of Great Britain.[8] The BHPS provides comprehensive data on a range of aspects, including information on children, and detailed individual- and household-level data on earnings, labor supply, occupation, and other variables relevant for our analysis, such as attitudes towards gender norms. We constructed our final dataset by exploiting the longitudinal dimension of the original data, and identifying parents and their first childbirth.

1.2.2 The Quarterly Labour Force Survey

In addition, we make use of the UK Quarterly Labour Force Survey (LFS) to supplement our analysis, and to capture pre-birth occupational characteristics of prospective parents. The survey includes approximately 36,000 respondent households each quarter since 1993, aiming for a representative sample of the UK population.[9] The quarterly survey adopts a panel design, retaining households for five consecutive quarters, while introducing a rotational replacement of one-fifth of the sample every quarter. Consequently, this means there is an 80% sample overlap across consecutive waves. Since we are interested in using LFS to rank occupations based on the extent to which they accommodate family commitments, we construct various measures, described in Subsection 1.2.4.3 below, and rank occupations, for each available quarter between 1993:Q1 and 2008:Q4. We then match these measures and ranks to individual records in the BHPS, by quarter-year and reported occupation.

[8] To continue tracking individuals after 2009 and up to the present day, the initial BHPS was succeeded by a follow-up survey known as Understanding Society. This longitudinal study includes about 40,000 individuals, among which approximately 8,000 are from the original BHPS households. In our study, we focus exclusively on the BHPS segment of the survey, as it provides richer and more frequent data on interviewees' attitudes towards gender norms.

[9] Since we lack information for the years 1991 and 1992, we cannot characterize occupational characteristics of BHPS parents for those two particular years. However, this should not present an issue regarding missing data, as all pre-birth characteristics — including those on occupations — are computed as an average across pre-treatment (*i.e.*, pre-birth) years. In addition, we omit observations for the first quarter of 2001 due to missing information on occupations in the LFS.

1.2.3 Final Sample

The implementation of our event study analyses, as described in Section 1.3 below, follows the methodology of Kleven et al. (2019). We define our analytical sample in the BHPS in three steps. First, we retain parents observed at least five times within our event window, spanning from five years before the first childbirth to ten years after.[10] They must as well be observed at least once before and after childbirth. Second, considering factors related to fertility and the labor market, we apply age restrictions and focus on individuals experiencing their first childbirth between the ages of 20 and 45. Finally, we keep individuals who were working, at least one year, prior to the first childbirth and trim the top and bottom first percentiles of the annual earnings distribution to mitigate the impact of outliers.[11] Following these sample restrictions, we compute our main indicators for gender norms and occupational characteristics, respectively outlined in Subsection 1.2.4.2 and Subsection 1.2.4.3. We further provide descriptive statistics in the Appendix for our analytical sample of 755 parents (equivalent to $8,350$ person-year observations), spanning from 1991 to 2009, categorized by their gender (Table A.1) and pre-birth gender norms (Table A.2).

1.2.4 Variables of Interest

1.2.4.1 Labor Market Outcomes

Our main aim is to examine the impact of the first childbirth on the motherhood penalty in earnings, defined as the annual individual labor income reported by the primary BHPS respondent. Our analysis focuses on two different margins potentially explaining the motherhood penalty in earnings — the intensive margin of labor supply, and wages.[12]

We measure the intensive margin of labor supply by the self-reported number of hours worked per week in each job, excluding any overtime hours.[13] This is calculated as the sum of weekly hours worked in the main job and, if applicable, any additional job(s) held by the respondent. Hourly wages are computed using monthly labor income and the number of hours worked (both self-reported), excluding over-

[10] Kleven et al. (2019) adhere to a minimum threshold of eight observations per parent within the event window. Nonetheless, to ensure an adequate sample size for subgroup analyses, we choose to relax this criterion to a minimum of five observations. We confirm in the Appendix, Subsection A.3.1, Figure A.24, that our main result split by gender norms (Figure A.3) is robust to restricting the analysis to parents observed at least eight times.

[11] In essence, to characterize parents' pre-birth occupational characteristics, we drop those who reported zero labor earnings for the five years prior to the first childbirth, and retain parents who reported at least one pre-birth observation of positive labor earnings. We further show, in the Appendix, Subsection A.3.1, that our results are robust to different sample selection criteria for pre-birth employment (Table A.10), as well as for various trimming versions of the annual earnings distribution (Table A.11).

[12] We also include findings related to the extensive margin of labor supply (see Figure A.6 reporting results for labor force participation), based on the self-reported current labor force status, and consider those who are either self-employed, employed, unemployed, or on maternity leave. However, as outlined in Subsection 2.3.1, it is important to note the limitation in terms of external validity of these results, since our sample is predominantly skewed towards working parents.

[13] This choice is determined by a high rate of missing values for the number of overtime hours worked among working parents ($\approx$11%).

time hours.[14]

Table 1.1 below provides descriptive statistics illustrating overall gender differences in our labor market outcomes. It shows that, on average, men earn significantly more and work longer hours per week than women. Additionally, men have higher hourly wages. Figure A.1 in the Appendix displays the trend of these labor market outcomes by gender over time, adjusting for year effects. These findings point to a consistent negative impact of being a woman on all outcomes, with the most pronounced effects observed for labor earnings and hours worked.

Table 1.1: Labor Market Outcomes by Gender

	Men	Women	Diff. (Men – Women)	Diff.(%)	S.E.	N
Earnings	21514.099	13602.162	7911.937***	58.167	246.723	8350
Weekly hours worked	45.584	32.561	13.023***	39.998	0.369	8342
Hourly wages	9.569	8.245	1.325***	16.071	0.138	8342
LFP	0.989	0.947	0.042***	4.446	0.004	8347

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents the mean of our four labor market outcomes, by gender. Earnings correspond to self-reported annual labor earnings. Weekly hours worked refer to the self-reported number of hours worked per week in each job, excluding any overtime hours. Hourly wages are computed using labor earnings and the number of self-reported hours worked, also excluding overtime hours. Both labor earnings and hourly wages are expressed in British pounds (£). LFP stands for labor force participation, and is a binary variable set to one if the respondent is either self-employed, employed, unemployed, or on maternity leave. Note that the sample is restricted to individuals with at least one pre-treatment observation with strictly positive labor earnings.

1.2.4.2 Gender Norms

The BHPS provides a wide range of questions on gender norms, answered solely by the primary survey respondent. Therefore, our focus is on individuals rather than on possible dynamics within couples. We consider six questions as in Flèche, Lepinteur, and Powdthavee (2020), detailed in Table 1.2 below, where respondents indicate their level of agreement with each statement, every odd survey wave.

Table 1.2: Gender Norms Variables

	Answer categories
A pre-school child is likely to suffer if his or her mother works All in all, family life suffers when the woman has a full time job A woman and her family would all be happier if she works Both the husband and wife should contribute to the household income Having a full-time job is the best way for a woman to be an independent person A husband's jobs is to earn money; a wife's job is to look after the home and family	1. Strongly agree 2. Agree 3. Neither agree, nor disagree 4. Disagree 5. Strongly disagree

Notes: Gender norms variables are asked to the BHPS main survey respondent, every odd survey wave, starting in wave 1, until wave 17.

[14]We compute hourly wages by dividing the monthly labor income by 4.3 $\times$ the number of weekly hours worked — excluding overtime hours, as we also lack information on whether these are paid at the same rate as regular hours.

We reverse code the measures where needed and compute the within-respondent and *pre-birth* average across these measures to form a gender norms score (see, *e.g.*, Farré and Vella, 2013; Flèche, Lepinteur, and Powdthavee, 2020). This score ranges from one (indicating more traditional attitudes) to five (more egalitarian attitudes), and represents the average response to the above gender norms questions across all pre-treatment years (*i.e.* before the first childbirth). We also calculate this individual gender norms score for each every other survey year over both the pre- and post-birth years to assess whether norms change after the birth of the first child.

Then, we classify prospective parents as holding more or less traditional attitudes compared to their counterparts in two steps. First, we regress our pre-birth gender norms score on age fixed effects using our analytical sample, and derive the residuals from this regression. Second, we compute the median value of these residuals, and split the sample into two groups — above and below this median. Thus, individuals above the median are labeled as egalitarian, while those below are labeled as traditional. Individuals defined as traditional by our categorization tend to think that women should not work as much as men for various reasons — such as for the child or the overall family well-being. This gender norms indicator is the one we will mostly use for heterogeneity throughout the paper.

Table 1.3: Gender Norms Before and After Childbirth

	Post-birth	Pre-birth	Diff. (Post – Pre)	S.E.	N
Panel A: overall	3.274	3.409	-0.135***	0.018	4450
Panel B: by gender					
Women	3.374	3.494	-0.120***	0.026	2108
Men	3.190	3.323	-0.133***	0.025	2342
P-value (women − men)	0.000	0.000			
Panel C: egalitarian norms					
Average score	3.540	3.797	-0.257***	0.023	2249
Child suffers if mother works	3.792	3.639	0.153***	0.042	2244
Family suffers if mother works	3.677	3.969	-0.292***	0.041	2246
Woman happier if she works†	2.983	3.211	-0.228***	0.033	2245
Husband and wife should contribute to income†	3.465	3.903	-0.438***	0.038	2246
Full-time job best for women's independence†	3.129	3.636	-0.506***	0.042	2246
A wife's job is to stay at home	4.187	4.415	-0.228***	0.033	2248
Panel D: traditional norms					
Average score	3.013	3.007	0.006	0.020	2171
Child suffers if mother works	3.024	2.640	0.385***	0.041	2169
Family suffers if mother works	2.916	2.926	-0.010	0.042	2167
Woman happier if she works†	2.658	2.688	-0.030	0.030	2165
Husband and wife should contribute to income†	3.046	3.121	-0.075**	0.037	2166
Full-time job best for women's independence†	2.744	2.944	-0.199***	0.039	2166
A wife's job is to stay at home	3.680	3.732	-0.051	0.039	2169

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents the means of the average gender norms score — ranging from 1 (more traditional views) to 5 (more egalitarian views) — equal to the within-individual average answer to the six questions presented in Table 1.2, and the average score by individual question. Means are displayed post-birth and pre-birth, and a negative (positive) difference between the post- and pre-birth scores indicates more traditional (egalitarian) attitudes. † refers to variables reverse coded.

Table 1.3 provides an overview of gender norms both before and after the first childbirth. Panel A displays these attitudes within our analytical sample, while Panel

B breaks down the data by gender, and Panel C by pre-birth gender norms. This descriptive analysis allows us to track the evolution of gender norms over time, and to discern potential drivers behind any shifts observed post childbirth. The findings point to a general trend towards more traditional gender norms following the first childbirth, as shown by a significant decrease in the gender norms score. This shift appears similar in magnitude across genders, suggesting that the effect is not gender specific. Interestingly, while egalitarian parents tend to adopt more traditional attitudes post-childbirth, this does not apply to those with traditional gender norms pre-birth.[15] Thus, at least descriptively, the birth of the first child is associated with increased traditional norms, particularly among individuals who held more egalitarian beliefs prior to the first childbirth.

1.2.4.3 Occupational Characteristics

To proxy the level of flexibility and/or family-friendliness of each occupation prior to the first childbirth, we also construct two indices (based on scores) using LFS data and the 2-digit-level occupation.[16] The first index is based on the average number of hours worked within each occupation, including overtime hours, as longer-hours occupations usually constitute a good proxy for less family-friendly occupations (*e.g.*, Goldin, 2014; Bertrand, Kamenica, and Pan, 2015). The second index, drawn from Costa Dias, Joyce, and Parodi (2020), calculates the proportion of part-time workers within each occupation to measure the degree of temporal flexibility within different occupations. Using these two indices, we establish quarterly and yearly rankings of occupations, with the lowest value indicating the most flexible and family-friendly occupations. We then match these rankings to individual records in the BHPS by quarter-year, and reported occupation. This provides us with quarterly-updated information on job flexibility for our sample of BHPS participants.

Table 1.4 presents descriptive evidence of pre-birth occupational characteristic scores across gender, and within each gender group, by pre-birth gender norms, providing some preliminary insights. First, women exhibit significantly ($p < 0.01$) lower scores than men for both rankings, indicating that they work in family-friendlier occupations compared to men. Second, comparing within gender across pre-birth norms, women who hold traditional gender norms prior to the first childbirth demonstrate a similar inclination towards family-friendlier occupations. They have significantly lower scores, for both ranks, relative to their egalitarian counterparts ($p < 0.01$). However, men exhibit statistically significant differences in their pre-birth sorting into longer-hours occupations based on their gender norms, although their differences in magnitudes within rank and across gender norms are relatively small. Given the significant difference in scores between genders, it appears that the tem-

[15] Grinza et al. (2022), using Understanding Society data, showed a notable shift in women's attitudes towards more traditional views upon entering parenthood, with no significant impact observed for men. Nevertheless, their results are not directly comparable to ours (*e.g.*, those presented in Figure 1.4), as they employ a different identification approach, and a different sample selection.

[16] We use the Standard Occupational Classifications (SOC) 1990 and 2000, at the 2-digit level, encompassing the minor occupation group of the main survey respondent in the BHPS. Indeed, in the public version of the BHPS, the SOC90 was released at the 2-digit level, while the SOC00 was released at the 3-digit level. For comparison purposes between SOC90 and SOC00, we can only work at the 2-digit level since we do not have the 3-digit information before 2000.

poral flexibility of occupations — reflected by shorter working hours and a higher prevalence of part-time employment — emerges as an important driver for women, and particularly for those with traditional pre-birth gender norms.

Table 1.4: Occupational Characteristics Before Childbirth by Gender

	Men	Women	Diff. (Men – Women)	S.E.	N
Rank 1: working hours	44.644	27.893	16.751***	0.436	7980
Egalitarian	43.846	29.579	14.267***	0.593	4037
Traditional	45.288	25.272	20.016***	0.658	3882
P-value (egal. – trad.)	0.018	0.000			
Rank 2: part-time workers	29.837	16.315	13.521***	0.424	6583
Egalitarian	29.243	17.698	11.545***	0.575	3344
Traditional	30.245	14.352	15.892***	0.643	3188
P-value (egal. – trad.)	0.131	0.000			

Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. This table presents the means of our two occupational characteristics scores, based on our two ranking measures of family-friendliness, by gender. Specifically, a lower value here represents working in family-friendlier occupations, *i.e.*, respectively for each rank, in occupations (1) with lower working hours, and (2) higher shares of part-time workers. The above ranks range from 10 to 99, with 10 (99) being the minimum (maximum) value.

Based on these rankings we also create indicators denoting whether prospective parents are employed in occupations deemed more or less family-friendly. We construct these as we did for gender norms in a two-step process. First, we regress separately each pre-birth family-friendliness score on age fixed effects and derive the residuals from this regression. Next, we compute the median value of these residuals, and split the sample into two groups — above and below this median — indicating the relative family-friendliness of their occupations compared to others. Individuals with values below the median are considered as working in family-friendlier occupations compared to those above the median.

1.3 Methodology

1.3.1 Empirical Strategy

Our empirical strategy relies on the event study design, as first proposed by Kleven et al. (2019), and our discussion here follows theirs. This event study design involves a staggered setup, where the treatment (*i.e.*, the first childbirth) occurs at different times across individuals. We estimate equation (1.1) below for each gender g and outcome Y_{it}:

$$Y_{it}^{g} = \sum_{t\neq -1} \beta_t^g \mathbb{1}[event_{it} = t] + \sum_{y=1991}^{2009} \gamma_y^g \mathbb{1}[year_{it} = y] + \sum_{k=20}^{45} \alpha_k^g \mathbb{1}[age_{it} = k] + v_{it} \tag{1.1}$$

The first term on the right hand side corresponds to event time dummies, ranging from five years before the first childbirth denoted by $t = 0$ (the event) to ten years

after.[17] These dummies are computed relative to the year prior to the first childbirth ($t = -1$), which means that the estimated event time coefficients (*i.e.*, $\widehat{\beta_t^g}$) will measure the impact of having a child relative to the year just before the first childbirth. The second term refers to year fixed effects to control for calendar time and business cycle effects. Finally, the third term refers to age fixed effects to control for life-cycle trends. The OLS estimates of equation (1.1) above enable us to derive two main measures — the within-gender impact of having a child, and the motherhood penalty, described in the next subsections.

1.3.2 Within-Gender Penalty

We compute, for both men and women, the percentage change in the outcome due to the first childbirth for all t. To quantify this, we first estimate the predicted outcome in absence of a child $\tilde{Y}_{it}^g = \sum_s \gamma_y^g \mathbb{1}[year_{it} = y] + \sum_k \alpha_k^g \mathbb{1}[age_{it} = k] + u_{it}$. Second, we use the average of this predicted outcome to normalize the within-gender effect of having a child, as such:

$$P_t^g = \frac{\widehat{\beta_t^g}}{E[\widetilde{Y}_{it}^g]}, \forall t \in [-5; 10] \tag{1.2}$$

When presenting our graphical results in Section 1.4, we plot the estimated values of P_t for both men and women, across each event time.

1.3.3 Motherhood Penalty

We compare mothers and fathers as in Kleven et al. (2019), and express the motherhood penalty as being the percentage change[18] in the outcome of men m compared to women w, as follows:

$$P_t = \frac{\widehat{\beta_t^m} - \widehat{\beta_t^w}}{E[\widetilde{Y}_{it}^w | t]} \tag{1.3}$$

We further provide three different specifications of this motherhood penalty. First, the average gap, *i.e.*, the mean of the motherhood penalty across the 10 years after becoming parent. Second, the short-run penalty, which corresponds to the mean of P_t in the first three years after the first childbirth, *i.e.*, for $t \in [0; 3]$. Third, the long-run penalty, which provides an estimate of the average of all P_t, for $t \in [7; 10]$. Yet, as highlighted by Kleven et al. (2019), the results can become less informative as we move further away from the event time $t = 0$, because the smoothness assumption may not longer hold in the long run due to for example having another child in the event window $t \in [1; 10]$. We should therefore exercise caution when interpreting the longer-term results. Nonetheless, results remain robust when we change the

[17] The "event-time" variable is computed using the interview year minus the year of birth of the first child due to limited information in the public version of BHPS.

[18] As mentioned in Section 3.3, we merge BHPS and LFS based on pre-birth reported occupations in the BHPS. This means that, for the post-birth periods, it is possible that respondents do not work and report 0 as labor earnings. We therefore cannot make use of the log specification or any log-like re-scaling (Chen and Roth, 2023).

definition of the control group to men and women with no children, which should minimize concerns about unobserved factors.[19]

1.3.4 Subgroup Differences

To understand the drivers of the motherhood penalty as defined above, we estimate equation (1.1) by different pre-birth indicators such as gender norms and occupational characteristics. To examine whether effects are statistically different between subgroups, we estimate a simple static difference-in-differences specification given by the following equation:

$$Y_{iat} = \alpha + \beta W_i + \theta D_{it} + \gamma_a + \lambda_t + v_{iat} \tag{1.4}$$

Where α is the intercept, W_i is a dummy variable for gender and is equal to 1 if the individual is a woman, and D_{it} is our treatment indicator equal to 1 if the individual is a woman and has a child at time t, and 0 otherwise. Additionally, γ_a and λ_t are age, and year fixed effects, respectively. We finally run a t-test on the difference across subgroups, using a Wald test.

1.3.5 Internal Validity

This event study design relies on several assumptions, which we discuss below. First, the estimation of the motherhood penalty considers men as a counterfactual for women, under the assumptions that (i) men's labor market outcomes remain unaffected by childbirth, and (ii) the timing of the first childbirth is exogenous to relative expectations in labor market outcomes within the couple (Angelov, Johansson, and Lindahl, 2016). These seem substantiated by a clear discontinuity in women's labor market outcomes coinciding with the first childbirth, while no such discontinuity is observed for men (*e.g.*, Figure A.3).

Second, the analysis rests on the validity of the Stable Unit Treatment Value Assumption (SUTVA), which requires that outcomes of each treated individual are independent from untreated individuals' outcomes, thus assuming no spillover effects to the control group. Graphical inspection of event study point estimates (*e.g.*, Figure A.3) suggests that this assumption is likely met, as earnings trajectories for men remain consistent with pre-treatment trends. To this extent, all pre-birth estimates are close to zero and non-significant, which additionally points to the absence of pre-trends. We further demonstrate in Section 1.5 that our results are robust to a placebo test on treatment timing and to heterogeneity in treatment effects across groups and time.

Finally, concerns may arise regarding the endogeneity of our gender norms and occupational flexibility indicators. On the one hand, women's perceptions of their roles in the workplace (*i.e.*, their gender norms) may affect their labor market decisions and at the same time the latter can in turn influence their gender norms

[19]We also estimate the earnings impact by number of children and display our results in Subsection A.2.8 in the Appendix. The results suggest that effects are substantial for the first and only child, and additional children tend to increase the short-run penalty, while the long-run penalty becomes not significant.

(Moriconi and Rodríguez-Planas, 2021). On the other hand, parenthood can also affect gender norms. However, we measure gender norms pre-birth, and show that parenthood does not affect them in a significant way as these exhibit relatively stable patterns over time (Figure 1.4, Subsection 1.4.4) and, more importantly, there is no differentiated effect across gender. Similarly, we compute our occupational index pre-birth, and further show in Subsection 1.4.2 that parenthood has a negligible influence on occupational mobility.[20]

1.4 Results

1.4.1 Industry, Occupations, and Gender Norms

There is supportive evidence in the literature that gender norms held by parents prior to the first childbirth drive women's earnings and labor supply. We add to this literature by showing, in the Appendix Subsection A.2.1, that men do not experience any impact on their labor earnings trajectories from becoming parents, regardless of their pre-birth gender norms, while women do. We further add to this literature by trying to understand, below, whether occupations and/or industry influence the dynamics of the motherhood penalty — and whether these dynamics differ between parents with egalitarian *versus* traditional gender norms. To do so, we estimate our main specification (equation (1.1)) on our main labor market outcomes, with and without industry and occupation fixed effects, separately by pre-birth gender norms.[21]

However, it is important to highlight a limitation of this method. Occupation and industry codes are only recorded for individuals who are employed at the time of interview. Periods in which individuals are out of work are therefore excluded from the analysis. This implies that the inclusion of occupation or industry fixed effects restricts the sample to working parents, and the results should be interpreted as conditional on employment. While such an approach is standard in descriptive decompositions, any selection into or out of employment around childbirth may itself be systematically related to gender norms, potentially leading to bias in the reported effects.

We provide graphical representations of our findings for annual labor earnings (Figure 1.1), weekly hours worked (Figure, A.7) and hourly wages (Figure, A.15 in the Appendix), and focus here on the short-run penalties because we lack long-term observations when we include industry fixed effects. The inclusion of occupation fixed effects in our event study regressions leads to a substantial reduction in the short-run earnings gaps, more pronounced for traditional mothers (Figure 1.1, Sub-

[20] An additional concern is that our approach, which follows recent literature, is prone to potential heterogenous effects by mothers' age at first birth. Melentyeva and Riedel (2023) shows that accounting for this potential heterogeneity with a stacked DiD strategy and a restriction on the control group yields larger estimates of earnings losses after childbirth. Thus, our estimates can be considered a lower-bound of the effects.

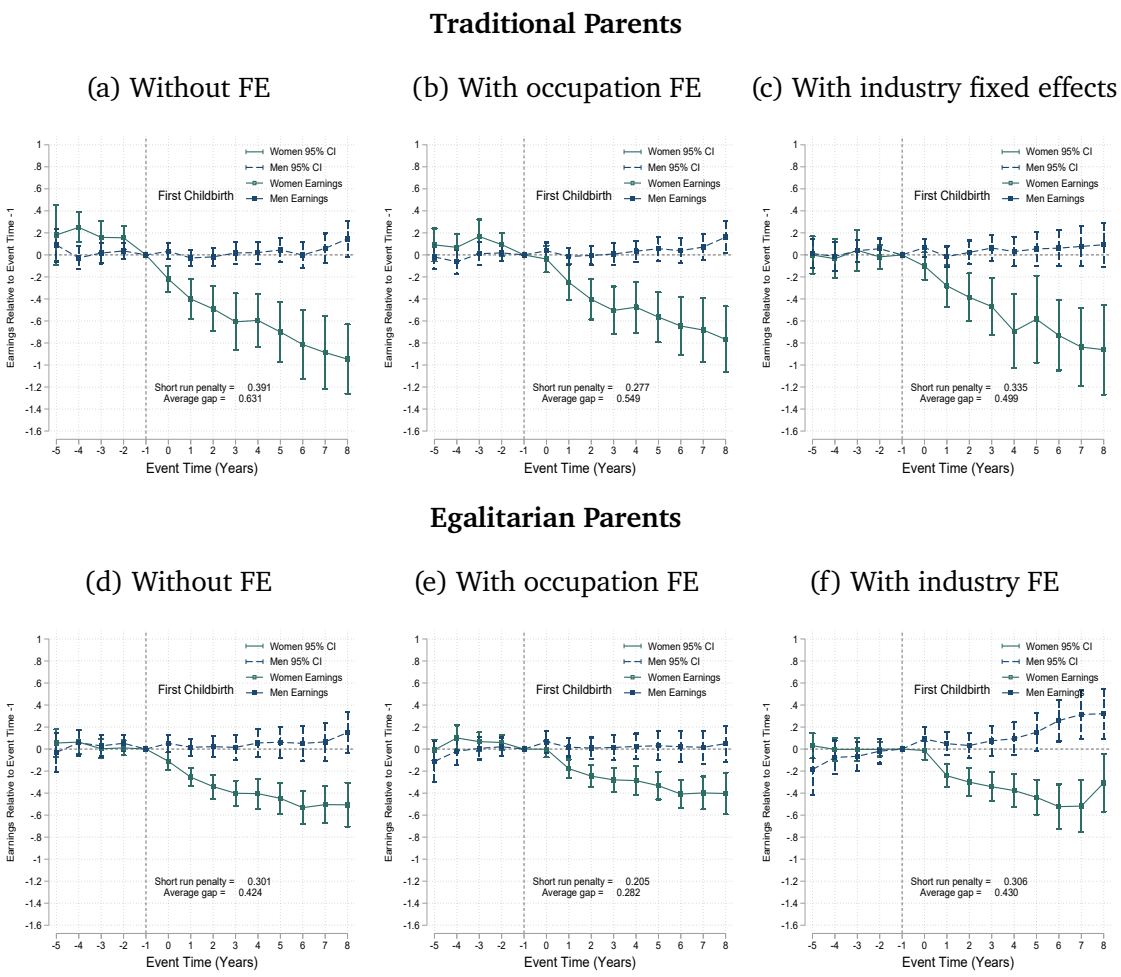
[21] To account for industry, we use the 4-digit industry classification provided in BHPS and aligned with the UK Standard Industrial Classification (SIC), encompassing both the 1980 and 1992 versions, across all panel waves. To account for occupations, we use the Standard Occupational Classifications (SOC) 1990 and 2000, at the 2-digit level, encompassing the minor occupation group of the main survey respondent in BHPS, as mentioned in Subsection 1.2.4.3.

figures (b) and (e)). In particular, controlling for occupation fixed effects, the difference in short-run earnings gaps between traditional and egalitarian women decreases by 8 percentage points (pp) — from $10pp$ to $2pp$. This suggests that occupational sorting accounts for a substantial share (80%) of the difference in the short-run motherhood penalty in earnings between traditional and egalitarian women.

On the other hand, the inclusion of industry fixed effects also leads to a decrease in the short-run gap in earnings between traditional and egalitarian women from $10pp$ to $4pp$. Thus, occupational sorting seems to play a much larger role than industry sorting in explaining short-run earnings losses after childbirth, as 80% of the initial earnings gap between traditional and egalitarian mothers is due to occupation choices, while it is only 40% for industry effects. However, it should be stressed that, because occupation and industry are only observed for individuals in employment, the reported estimates capture dynamics among employed parents. Thus, selection into employment, which we cannot observe in this decomposition, could influence the magnitude of our effects.

In addition, our results indicate that accounting for occupational sorting completely removes the short-run difference in hours worked penalties between traditional and egalitarian women (see Figure A.7 in the Appendix). Consistent with the descriptive insights reported in Table 1.4, our results seem to indicate a tendency for traditional women (as opposed to egalitarian women) to self-select into occupations that better accommodate family responsibilities and facilitate a reduction in working hours upon motherhood (*e.g.*, Figure A.4). We believe that these results, which combine gender norms and occupational sorting in a dynamic way, are novel to the literature on the motherhood penalty.

Figure 1.1: Impact of Parenthood on Earnings (Un)Conditional on Industry and Occupations



Notes: OLS results for equation (1.1) on annual labor earnings, by pre-birth gender norms (traditional *versus* egalitarian). Results in Figures (a) and (b) are without fixed effects, while Figures (b) and (e) present results controlling for 2-digit occupation fixed effects, and Figure (c) and (f) controlling for 4-digit for industry fixed effects. FE stands for fixed effects.

1.4.2 Labor Market Mobility Around Childbirth

Here, we aim to examine whether childbirth affects labor market mobility, considering mobility across occupations, industries, and sectors (private *versus* public). This analysis allows us to examine the dynamics of occupational sorting, in particular whether the latter precedes the first childbirth — and is therefore linked with pre-existing gender norms — or whether there is labor mobility around the first childbirth. If sorting precedes the first childbirth, rather than being a consequence of it, then this would corroborate our result on gender norms, suggesting that traditional women tend to choose occupations *prior to the birth* that are more accommodating to family responsibilities compared to their egalitarian counterparts and men.

1.4.2.1 Occupations

First, we examine how the first childbirth affects job mobility. We measure occupations at the 2-digit level, and further create a binary variable taking the value one if the individual changed occupation between $t-1$ and t; and zero otherwise. As previously mentioned, our results are conditional on employment, since occupation and industry codes are not observed during non-working periods. The event study results presented in the Appendix, Figure A.17, suggest no effect of the first childbirth on occupational mobility.

We also descriptively examine in Table 1.5 the share of individuals working in family-friendlier occupations, based on our two constructed ranks, separately pre- and post-treatment considering different time-points. These are defined as follows: (i) post-birth ($t \geq 0$), (ii) pre-birth ($t < 0$), (iii) between five years and two years prior to parenthood ($t \in [-5; -2]$), and (iv) between two years prior to parenthood and ten years after ($t \in [-2; -10]$). This suggests that sorting into specific occupations, enabling shorter working hours for instance, occurs mainly before the first childbirth, as there seems to be no job mobility induced by the first childbirth.

Table 1.5: Share of Parents Working in Family-Friendly Occupations

	Mean	Min	Max	N
Rank 1: working hours				
Post-birth: $t \geq 0$	0.500	0.000	1.000	5351
Pre-birth: $t < 0$	0.501	0.000	1.000	2939
$t \in [-5; -2]$	0.506	0.000	1.000	2302
$t \in [-2; 10]$	0.498	0.000	1.000	5988
Rank 2: part-time workers				
Post-birth: $t \geq 0$	0.484	0.000	1.000	3493
Pre-birth: $t < 0$	0.521	0.000	1.000	2660
$t \in [-5; -2]$	0.532	0.000	1.000	2177
$t \in [-2; 10]$	0.483	0.000	1.000	3976

Notes: This table presents the share of individuals working in family-friendly occupations (as defined in Subsection 3.3.3) *i.e.*, in occupations with shorter working hours (rank 1) and higher shares of part-time workers (rank 2), at different time-points. These time-points represent the (i) post-birth ($t \geq 0$) and (ii) pre-birth ($t < 0$) periods, as well as (iii) between five years and two years prior to parenthood ($t \in [-5; -2]$), and (iv) between two years prior to parenthood and ten years after ($t \in [-2; -10]$).

1.4.2.2 Industry

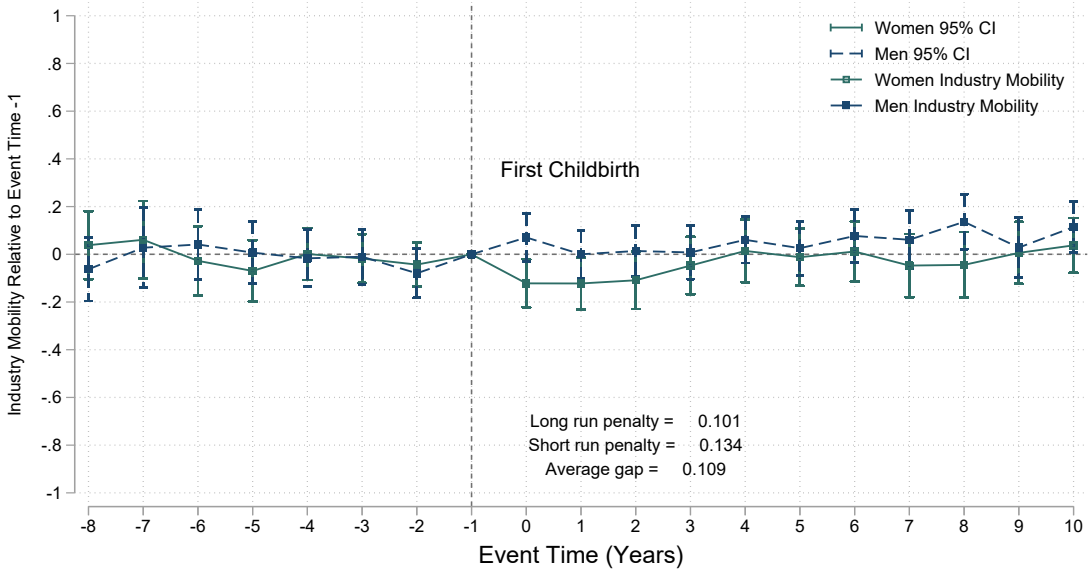
As for occupations, we examine here whether there is any mobility between industries. To do so, we create a binary variable set to one if a respondent changed industry (reflected by a change in the 4-digit SIC code) between $t-1$ and t, and zero otherwise. We use this variable as an outcome and run our main event study specification (equation (1.1)). Figure 1.2 reports the results of this exercise, and points to no effect as well as no difference by gender of parenthood on mobility between industries. These findings suggest that individuals tend to sort into specific industries prior to their first childbirth — rather than changing the industry sector they work in — as a consequence of becoming parents.

1.4.2.3 Public *versus* Private Sectors

Finally, we construct our main outcome to capture whether the respondent works in the public sector at time t, regardless of where they were working at time $t-1$.[22] We run this regression to investigate whether childbirth affects moving into the public sector, under the assumption that working in the public sector is associated with amenities such as shorter working hours, higher flexibility, and job security, as well as employment protection often preferred by workers with care responsibilities (Chassamboulli and Gomes, 2023; Niederle and Vesterlund, 2007; Anghel, de la Rica, and Dolado, 2011). Our results are presented in Figure A.18 in the Appendix,

[22]This information is reported by BHPS respondents. In each wave the respondent is asked the following: "Do you work for a private firm or business or other limited company, or do you work for some other type of organisation?". We use this variable as opposed to the alternative variable indicating whether parents work in the private sector at time $t-1$ and in the public sector at time t, as the latter does not have enough variation. This occurs for fewer than 50 observations.

Figure 1.2: Impact of Parenthood on Mobility Between Industries



Notes: OLS results for equation (1.1) on a dummy variable equal to 1 if the participant changed industry (reflected by a change in the 4-digit SIC code provided in BHPS) between $t-1$ and t; 0 otherwise. This figure shows the estimated impact of having a first child on any mobility between industries. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

and point to no effect of the first childbirth on sorting into the public sector, as well as no difference by gender. This further suggests that sorting into a specific sector, *e.g.* characterized by family-friendlier features, happens prior to becoming a parent.

1.4.3 Pre-Birth Comparative Advantage

Here, we explore alternative mechanisms potentially driving the motherhood penalty, and particularly differences in comparative advantage between women and men, *i.e.*, differences in relative labor market productivity pre-birth. As we discussed in the introduction, if couples have different labor supply elasticity (*e.g.*, Becker, 1985), they may aim to maximize total household income choosing to prioritize the market work of the parent with the greater comparative advantage. The parent with the lowest comparative advantage would assume a relatively larger share of domestic responsibilities with effects on their earnings trajectory and the motherhood penalty.[23] Gender norms can in turn affect the labor market response to parenthood based on differences in comparative advantage pre-birth within the couple.

We assess the role of comparative advantage, proxied by labor market productivity, through pre-birth hourly wage differentials. Specifically, for each couple, we compute the *average* difference in hourly wages prior to their first childbirth, and further construct a binary indicator. The latter equals one if the father had a higher

[23]For descriptive evidence on this issue, see Costa Dias et al. (2021).

average hourly wage than the mother before the first childbirth, thereby identifying which partner held the comparative advantage. To quantify the impact of parenthood on earnings by pre-birth comparative advantage, we run our main specification (equation (1.1)) split by the comparative advantage indicator, and present our results in Figure 1.3.

We first present the results using the full sample (Panels (a) and (b)). Given that median within-couple wage differences are minimal (under one pound), couples with similar earnings may not perceive a strong comparative advantage for either partner. In such cases, decisions may be driven by expectations of future career potential or individual preferences regarding career progression. Conversely, a larger wage differential likely makes the existing comparative advantage more salient, and more influential in decision-making. Thus, we further present results dropping observations between the 45^{th} and the 55^{th} percentiles of the pre-birth wage difference distribution (Panels (c) and (d)), and dropping observations between the third and the seventh deciles of the wage difference distribution (Panels (e), and (f)). The results presented below suggest that the role of comparative advantage in the motherhood penalty starts to emerge for the 30-70 trimming sample, where comparative advantage can be more salient within the couple, though it is not consistently significant across trimming levels.[24] We now turn to explore whether comparative advantage mechanisms may operate differently across pre-birth gender norms.

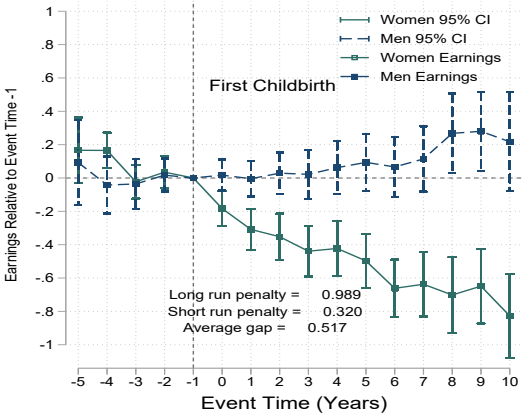
[24]Table A.6 in the Appendix shows the results of a pooled difference-in-differences model broken down by pre-birth comparative advantage. These results show that the motherhood penalty in earnings is higher (and statistically significantly different) for women with no comparative advantage, especially for the 30-70 trimming sample.

Figure 1.3: Impact of Parenthood on Earnings by Pre-Birth Comparative Advantage

No Trimming

(a) C.A.: man = 1

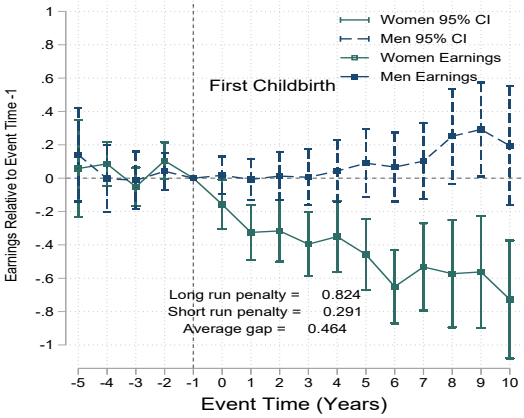
(b) C.A.: man = 0



Trimming 45-55 Percentiles

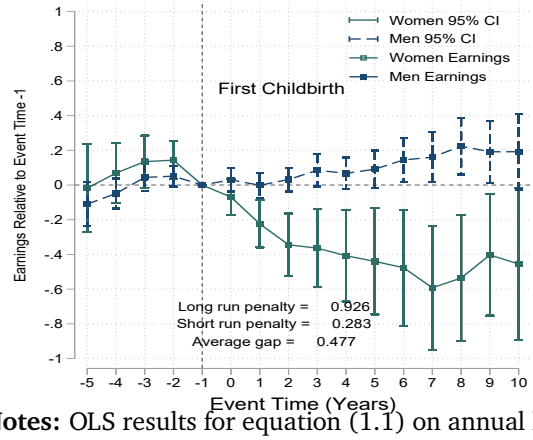
(c) C.A.: man = 1

(d) C.A.: man = 0

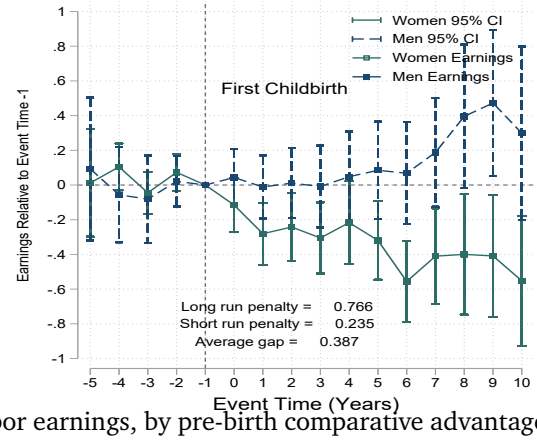


Trimming 30-70 Percentiles

(e) C.A.: man = 1



(f) C.A.: man = 0



Notes: OLS results for equation (1.1) on annual labor earnings, by pre-birth comparative advantage. On the left hand side (Figures (a), (c) and (e)), results are displayed for when men had a comparative advantage (C.A.: man = 1). On the right hand side (Figures (b), (d) and (f)), results are displayed for when women had a comparative advantage (C.A.: man = 0). Results are displayed for the full distribution of the pre-birth average wage difference (Figures (a) and (b)), dropping observations between the 45^{th} and 55^{th} percentiles of this distribution (Figures (c) and (d)), and dropping observations between the third and seven deciles of this distribution (Figures (e) and (f)). Refer to Figure A.3 for the definition of the gaps.

Next, we examine the different role of comparative advantage across traditional and egalitarian women. To do so, we estimate a triple-difference specification — outlined in equation (1.5) below — to analyze how the average effects of motherhood on labor earnings change by comparative advantage and across gender norms:

$$Y_{iat} = \alpha + \beta W_i + \theta Post_t + \delta CA_i + \phi(W_i \times Post_t) + \psi(W_i \times CA_i) + \chi(Post_t \times CA_i) + \eta(W_i \times Post_t \times CA_i) + \gamma_a + \lambda_t + u_{iat} \tag{1.5}$$

In equation (1.5), α denotes the intercept, W_i is a dummy variable for gender and is equal to 1 if the individual is a woman, $Post_t$ is a dummy denoting the post-childbirth period, and CA_i is a dummy set to 1 if a man had a comparative advantage (*i.e.*, a greater average hourly wage than their spouse or partner) prior to the first childbirth. Additionally, γ_a and λ_t are respectively age, and year fixed effects. We report the estimates of this regression in Table 1.6 below using the same sample restrictions as in Figure 1.3.

Overall and across the different samples, these results suggest that the comparative advantage mechanism is heterogeneous across gender norms. Results become clearer as we trim the sample but remain qualitatively consistent for the untrimmed sample as well (columns (1)–(2)), suggesting that the comparative advantage mechanism operates for egalitarian women only. The strongest results are shown in columns (5)–(6) for the 30-70 trimming, where there is a clear difference in pre-birth comparative advantage between men and women. Here, the earnings penalty for egalitarian mothers with no comparative advantage is significantly larger (about 56-pp higher) than for egalitarian mothers with comparative advantage.[25] In contrast, the motherhood earnings penalty is similar for traditional women regardless of their comparative advantage. This evidence substantiates the role of the comparative advantage mechanism in shaping earnings trajectories for egalitarian women around motherhood, while it appears to play no role for traditional women.

A theoretical explanation, from Akerlof and Rayo (2020), and in line with these results, is that households may be invested in different narratives, and the utility retrieved from mothers' post-birth labor supply differs across gender norms — even under the same change in their budget constraint. Our results therefore underscore the significance of individual beliefs and preferences as credible sources of disparities in the magnitude of the motherhood penalty in earnings, notably as a driver of post-birth labor supply. However, this evidence is much stronger if gender norms do not change as a consequence of parenthood, which we turn to test in the next subsection.

[25]The estimated motherhood penalty coefficients for women with and without a comparative advantage — $\hat{\phi}$ and $\hat{\eta}$ respectively — are normalized (divided) by women's pre-birth average of labor earnings.

Table 1.6: Impact of Parenthood on Earnings by Pre-Birth Comparative Advantage and Gender Norms

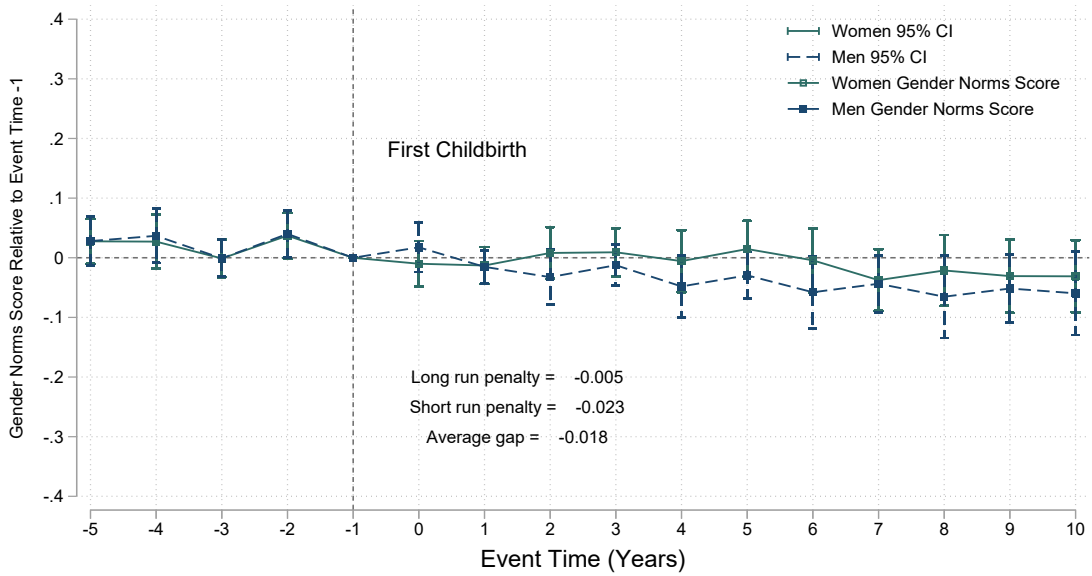
	No Trimming		Trimming 45-55 Percentiles		Trimming 30-70 Percentiles	
	(1) Egalitarian	(2) Traditional	(3) Egalitarian	(4) Traditional	(5) Egalitarian	(6) Traditional
Post	779.144	536.576	821.989	-651.197	-286.140	-292.770
	(1407.559)	(1377.524)	(1597.458)	(1501.612)	(2166.917)	(2000.516)
Woman	269.810	-578.290	1838.793	1386.107	1634.960	2378.324
	(1091.798)	(1727.153)	(1286.732)	(3110.570)	(1679.402)	(3576.378)
Woman × Post **(a)**	-5672.439***	-9918.804***	-4977.637**	-8400.923**	-2082.884	-9230.018**
	(1634.092)	(2039.481)	(2009.218)	(3404.234)	(2514.737)	(3991.487)
C.A.: man	3141.909***	1497.397	4176.617***	1541.844	3459.363**	1811.065
	(1113.477)	(1217.349)	(1310.376)	(1343.366)	(1715.001)	(1779.461)
Post × C.A.: man	1146.050	1303.435	2459.047	3451.944**	4112.726*	2905.540
	(1718.321)	(1517.967)	(1935.641)	(1653.130)	(2427.993)	(2057.312)
Woman × C.A.: man	-3753.624**	-4592.155**	-5985.445***	-7495.896**	-6865.149***	-8840.579**
	(1463.475)	(1908.151)	(1696.103)	(3160.515)	(2089.856)	(3638.224)
Woman × Post × C.A.: man **(b)**	-2596.019	406.913	-4464.942*	-1928.940	-8291.040***	-380.818
	(2148.303)	(2293.892)	(2545.386)	(3535.670)	(3117.326)	(4128.194)
Normalized Average Effects (%)						
(a)	-41.61	-72.76	-34.74	-58.64	-14.07	-62.34
(a)+(b)	-60.65	-69.77	-65.91	-72.10	-70.07	-64.91
Observations	4150	4131	3203	3115	2405	2516
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. OLS results for equation (1.5) on annual labor earnings for egalitarian and traditional parents, as defined in Subsection 1.2.4.2. We display results for the full distribution of the pre-birth wage difference in columns (1) and (2). In addition, we drop observations between the 45^{th} and 55^{th} percentiles of this distribution in columns (3) and (4), and we drop observations between the third and seventh decile of this distribution in columns (5) and (6). The variable 'C.A.: man' is a dummy set to 1 if a man had a greater comparative advantage pre-birth than her female partner, 0 if a woman had a greater comparative advantage pre-birth, and 'post' is a dummy set to 1 for the post-childbirth period. Having a greater comparative advantage pre-birth translates in having a greater hourly wage pre-birth. Further, we report the normalized average effects for the coefficients of interested (a) $\hat{\phi}$ and (b) $\hat{\eta}$, where we divide (a) and (a) + (b) by the pre-birth average of women's labor earnings, separately by trimming scenario. These display the relative drop in labor earnings due to childbirth relative to the average of all women pre-birth.

1.4.4 Gender Norms Around Childbirth

We examine here whether gender norms are affected by the event of having a child, whether there is any gendered effect, and whether effects vary by pre-birth gender norms. If, for example, women (but not men) become more traditional after becoming parents, part of the differential effects by gender norms could be driven by compositional changes. Descriptive evidence presented earlier in Table 1.3 suggests that individuals tend to become overall slightly more traditional after the first childbirth, particularly those who hold more egalitarian gender norms prior to the first childbirth. To quantify the impact of the first childbirth on gender norms, we run our main event study specification using the gender norms score as an outcome (Figure 1.4 below), and further split our regressions by the original gender norms indicator for when individuals have pre-birth egalitarian *versus* traditional gender norms (Figure A.8 in the Appendix).

Figure 1.4: Impact of Parenthood on Gender Norms



Notes: OLS results for equation (1.1) on gender norms ("gender norms score") defined in Subsection 1.2.4.2 presented alongside the motherhood penalties (short, long and average gaps). The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

In summary, our findings indicate that becoming parents does not alter the relative difference in gender norms between men and women, regardless of the gender norms they hold pre-birth. Therefore, potential changes in gender norms cannot threaten the validity of our results, and we can carefully conclude that pre-birth gender norms constitute an important driver of labor supply responses to motherhood, contributing to the motherhood penalty in earnings.[26]

[26] In addition, we also test whether the first childbirth affects a broader measure of norms, *i.e.*, social norms. Results are presented in the Appendix, Figure A.26, and confirm that the first childbirth does

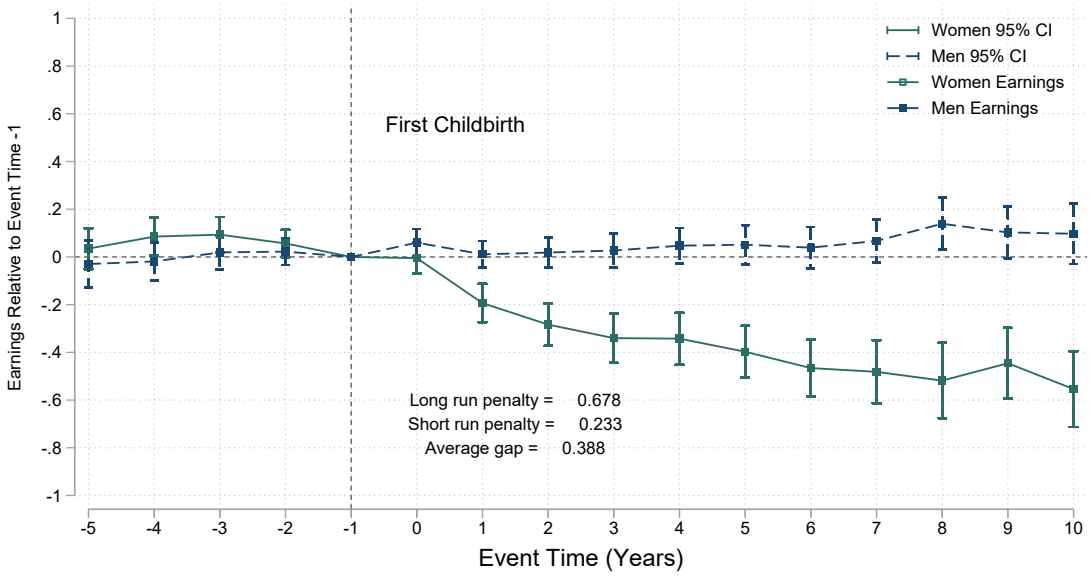
1.4.5 Occupational Characteristics

We now turn to the relationship between gender inequalities and occupational characteristics and examine the link between pre-birth occupational characteristics, occupational sorting and gender norms, as well as their contribution to the motherhood penalty.

1.4.5.1 Differences Between Occupations

We examine first the potential overall influence of differences between occupations on the extent of the motherhood penalty. We estimate our main specification on earnings controlling for 2-digit occupation fixed effects and report our findings in Figure 1.5 below. Compared to the model without occupation fixed effects, the average gap in labor earnings diminishes by 9 percentage points (pp), which indicates that about 19% of the overall motherhood penalty can be attributed to differences between occupations. We also report the results of estimating the average effect in the Appendix Table A.7, where we control for occupation fixed effects. The inclusion of occupation fixed effects leads the normalized average effect to reduce by $4pp$, which confirms that between-occupation differences only explain a small part of the overall motherhood penalty.

Figure 1.5: Impact of Parenthood on Earnings Conditional on Occupations



Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in annual labor earnings conditional on 2-digit occupation fixed effects. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

not imply more conservative and/nor more traditional social norms.

1.4.5.2 Family-Friendly Occupations

Next, we examine the influence of flexible occupations on the magnitude of the motherhood penalty using the two measures we developed in Subsection 1.2.4.3 to capture pre-birth occupational characteristics in terms of family-friendliness. We report our results in Subsection A.2.6.2 in the Appendix. Overall, our findings suggest that family-friendly occupations do not have a significant role in shaping the overall magnitude of the motherhood penalty in earnings. These conclusions align with those of Costa Dias, Joyce, and Parodi (2020), implying that differences in job characteristics contribute only a small proportion to the overall motherhood penalty in earnings.

1.5 Robustness Checks

In this section, we examine the robustness of our results to potential failures of the main assumptions. First, we run a placebo exercise that can help shed light on potential pre-trends or anticipation effects. Second, we employ the method developed by Sun and Abraham (2021) to account for potential heterogeneity in treatment effects. Finally, we examine the robustness of our main results to an alternative control group, using individuals who remain childless as control group.

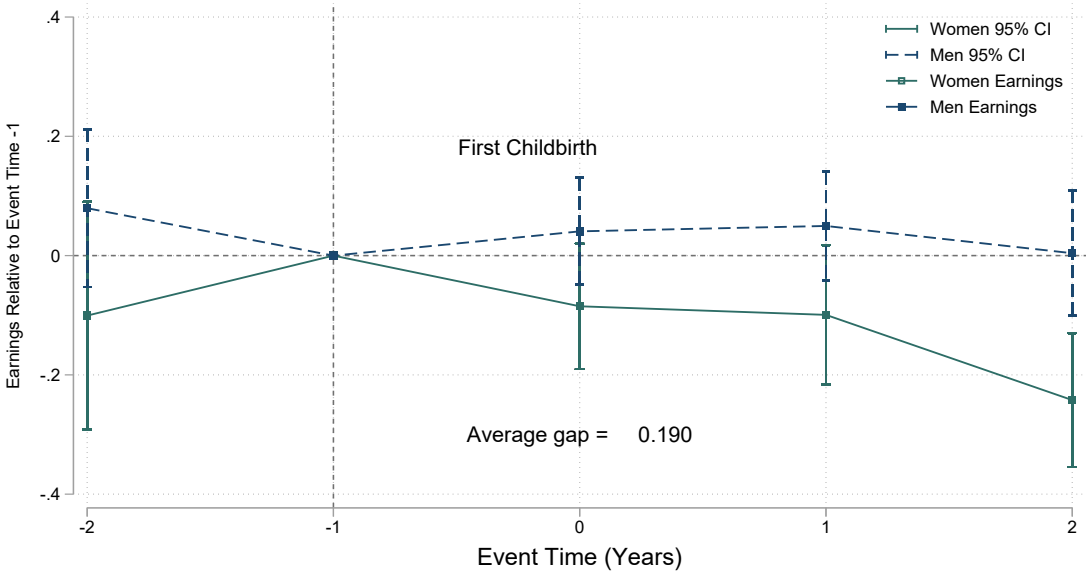
1.5.1 Pre-Trends

The main assumption behind an event study design is the parallel trends assumption. A common way to test for this is to investigate whether the treatment appears to have an effect on the outcome before it actually occurs (Freyaldenhoven, Hansen, and Shapiro, 2019). Therefore, we run a placebo event study regression, in which we set the event date to three years before the actual event, and exclude post-treatment observations, thus focusing on not-yet-parents over the period $[-5;0]$.

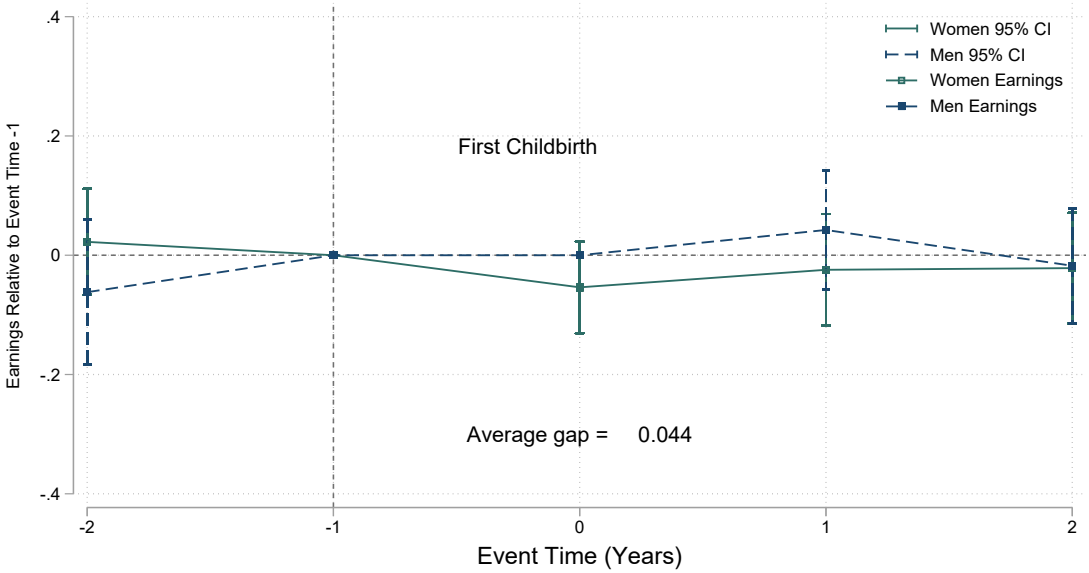
We can thus display the placebo test results (below in Figure 1.6) for the event window $[-2;2]$ with $t = 0$ representing the event 'first childbirth'. We undertake this analysis for parents with both traditional and egalitarian gender norms before childbirth, to validate our main findings. The results of this placebo exercise suggest the absence of pre-trends, as well as no anticipatory behavior on the part of prospective parents, regardless of their gender norms pre-birth. None of the post-treatment point estimates are statistically different between men and women, across gender norms, ruling out potential failures of the parallel trends assumption.

Figure 1.6: Placebo Test: Motherhood Penalty in Earnings by Pre-Birth Gender Norms

(a) Traditional pre-birth



(b) Egalitarian pre-birth



Notes: OLS results of equation (1.1) on annual labor earnings for (a) traditional and (b) egalitarian parents, as defined in Subsection 1.2.4.2. This figure presents a placebo test for the motherhood penalty in earnings to test for the timing of fertility, where the first child year of birth has been set to 3 years prior the actual year of birth. For each $t \in [-2;2]$ and gender g, we present the percentage effects of parenthood on earnings. The average gap represents, for $t \geq 0$, the mean of P_t.

1.5.2 Heterogeneous Treatment Effects

As mentioned in Section 1.3, our event study regressions provide the average effects, assuming an homogeneous treatment effect across all treated individuals and using not-yet-treated-parents as a control. A caveat of such setup is that the event can induce heterogeneous responses not accounted for by the standard event study set up we implemented, that can contaminate leads and lags by effects from other periods (Sun and Abraham, 2021).[27]

To verify that our results are not contaminated by heterogeneous treatment effects, we employ the interaction-weighted method developed by Sun and Abraham (2021). Figure A.20 in the Appendix report the results for labor earnings. We carry out this exercise separately for egalitarian and traditional parents, and pool different cohorts together to increase our control group's sample size.[28] Naturally, we expect different magnitudes of the motherhood penalty because the control group definition has changed. Yet, we should still observe i) a penalty in earnings, *i.e.*, a significant gender gap opening post-birth, and ii) parallel trends pre-birth. The results of this exercise align with our main results, indicating that traditional women suffer from a larger motherhood penalty in earnings compared to egalitarian women.

1.5.3 Difference-in-Differences Event Study

Here, we follow Kleven, Landais, and Søgaard (2019) and run an additional robustness check by employing individuals who remain childless as an alternative control group. This also allows us to assess potential gender-specific effects of parenthood. Those without children as of 2009 — the last observed year in our survey — are assigned "placebo" births based on the age distribution of first childbirth among parents. We provide additional methodological details in Subsection A.2.9 in the Appendix, along with the results stratified by the pre-birth indicator of gender norms attitudes. For the control group, this indicator is calculated across all (placebo) pre-birth years.

The difference-in-differences event study results for individuals with traditional attitudes are presented in Figure A.22, while those for egalitarian parents are shown in Figure A.23. Overall, they confirm the main findings as per our primary specification (equation (1.1)). First, women, both with and without children, exhibit similar pre-trends but experience a clear divergence following the first childbirth notably in the short run. Second, men, whether parents or not, remain unaffected regardless of the timing of the event and regardless of their pre-birth gender norms, supporting the choice of using them as a appropriate control group for women. Finally, pre-existing gender norms contribute to the widening gap in earnings in the short-term observed among women, with traditional mothers experiencing a more pronounced and long-lasting decrease in earnings than egalitarian women.

[27] See Goodman-Bacon (2021); Callaway and Sant'Anna (2021); Roth et al. (2023), among others, for recent developments in this literature.

[28] The last treated cohort (*i.e.*, last "first childbirth" year) in our data is 2008, and contains only 30 observations for egalitarian parents, and 54 observations for traditional parents. We thus decided to pool together those with a first childbirth year equal to 2004, 2005, 2006, 2007 and 2008 to construct a larger control group.

1.6 Conclusion

Despite the implementation of various family policies — such as parental leave, or early childcare policies — the motherhood penalty remains persistent and large, and its heterogeneity across countries remains puzzling, not least regarding the factors that influence it (Kleven et al., 2024). In light of previous findings from this literature, this paper investigates how pre-birth gender norms and comparative advantage mechanisms in productivity shape women's occupational choices and labor market trajectories around motherhood. Using data from the BHPS, we employ an event study design, as first proposed by Kleven, Landais, and Søgaard (2019), to quantify the impact of motherhood on various labor market outcomes in the United Kingdom. We explore how these penalties vary across different factors commonly discussed in the literature.

First, we examine how pre-birth norms drive the magnitude of the motherhood penalty in earnings and hours worked. Traditional mothers experience an 18-percentage-point larger penalty in earnings and a 20-percentage-point larger penalty in hours worked compared to egalitarian mothers. Gender norms can therefore guide the labor supply responses of mothers and serve as a marker for their labor market attachment, pointing to larger gaps for traditional women due to their preference for shorter working hours (Mensinger and Zimpelmann, 2024).

Second, we explore alternative drivers of penalties through relative comparative advantages in productivity within couples. Our findings show that, particularly when the difference in comparative advantage within couples is large, the earnings penalty increases for women whose partner had a comparative advantage in the labor market prior to the first childbirth. This pattern is distinctly different across pre-birth gender norms, and women with more egalitarian norms drive responses to comparative advantage. Therefore, we argue that more egalitarian parents may prioritize the market work of the parent with the greater comparative advantage to rationally maximize household income, while this mechanism does not seem to apply to traditional parents. To the best of our knowledge, this finding is novel in the literature, and speaks to the literature on motivated beliefs. Our results therefore support the interpretation of gender norms as narratives families may want to pursue (Akerlof and Rayo, 2020), guiding behaviors in line with these beliefs to derive a greater utility (Bénabou and Tirole, 2016).

Finally, we examine the role of occupational features, such as family-friendliness, in explaining motherhood penalties, and illustrate the significance of pre-birth gender norms in shaping occupational preferences, and sorting patterns *pre-birth*. We find that occupational characteristics explain about 19% of the overall gender gap in earnings that emerges after parenthood, with distinctive patterns across gender norms. On the one hand, accounting for occupational sorting reduces the *average* motherhood earnings penalty by $13pp$ for traditional mothers and $12pp$ for egalitarian mothers, with the reduction entirely driven by hours worked for traditional mothers. On the other hand, in the short-run, occupational sorting accounts for 80% of the earnings penalty gap between the traditional and egalitarian women, and fully eliminates the difference in hours worked penalties. These results suggest that traditional women sort into occupations more conducive to balancing family respon-

sibilities, that in turn have a substantial impact on their earnings trajectory upon motherhood. We then find no evidence of mobility across occupations or industries, nor changes in gender norms across genders from before to after the first childbirth, suggesting that pre-birth gender norms drive a degree of occupational sorting *prior* to the first childbirth — further explaining part of the differences in post-childbirth experiences between traditional and egalitarian women.

These results hold significant implications for policies that aim at reducing post-birth gaps in the labor market, particularly between traditional and egalitarian women. These policies can remain ineffective if they do not consider the influence of norms and beliefs on pre-birth sorting into specific occupations, and how they contribute to the overall earnings gaps.

Our main findings remain robust across various robustness checks, including a placebo test, heterogeneous treatment effects, and an alternative definition of our control group. Yet, our survey data have limitations when it comes to exploring in more details occupational sorting, and job characteristics. We provide a descriptive analysis using the UK Labour Force Survey, suggesting that more traditional women tend to sort pre-birth into "family-friendly" jobs — characterized by higher proportions of part-time workers and shorter working hours. More detailed data on occupational characteristics will allow further research to examine more in depth the contribution of occupational sorting, and job features in the motherhood penalty.

Chapter 2

Beliefs on Children's Human Capital Formation and Mothers at Work

2.1 Introduction

Across North America and Europe, mothers experience a 25-30% drop in employment after childbirth (Kleven, Landais, and Leite-Mariante, 2024), contributing to the well-documented gender gap in earnings (Bertrand, 2011, 2020; Goldin, 2006). Moreover, this motherhood penalty has proven resistant to policy interventions (Kleven et al., 2024) and is particularly pronounced for new mothers (Bertrand, Goldin, and Katz, 2010; Blau and Kahn, 2017; Cortés and Pan, 2023; Kleven, Landais, and Søgaard, 2019). Economic theory offers two main explanations related to child-rearing. The first is that women have a lower earnings potential than men in the labor market, leading to a comparative advantage in child-rearing (Becker, 1985). However, recent empirical studies find limited support for this mechanism (Andresen and Nix, 2022; Kleven, Landais, and Søgaard, 2021; Siminski and Yetsenga, 2022) and the persistence of gender pay gaps within firms, and among highly skilled women points to the need for alternative explanations (Blau and Kahn, 2017; Card, Cardoso, and Kline, 2016). The second explanation links traditional gender norms to the beliefs that women have an absolute advantage in child-rearing relative to men (Cortés et al., 2024). Consider, for instance, a policy that reduces gender discrimination in hiring and promotions. Even if successful in equalizing earnings potential across genders, these beliefs may still pressure women to contribute less to the labor market.

While beliefs on absolute advantage in child-rearing are important to understand, they are not straightforward to measure. In general, belief elicitation can be obscured by many combinations of preferences (Manski, 2004), and in particular, beliefs on absolute advantage can be obscured by differences in earnings potential. This difficulty leaves the empirical relevance of beliefs on absolute advantage an open question. Moreover, the existence of such beliefs would call for policies that reduce informational frictions by conveying how children are affected when mothers remain in the workforce. Yet, it is unclear whether information can meaningfully shift beliefs about children's outcomes when mothers work. In this paper, we address beliefs about how well children do when mothers maintain careers and responses to infor-

mation on children's actual performance when mothers work.

We begin with a model demonstrating that both comparative and absolute advantage lead women to spend more time at home. Importantly, beliefs that women hold an absolute advantage in child-rearing imply that policies equalizing earnings potential will not be sufficient to close motherhood penalties. Yet, observing labor supply itself, does not reveal whether women are perceived to have an absolute advantage. We use our model to define a distribution of beliefs on absolute advantage by comparing expectations of a child's future human capital accumulation when a mother works long hours in the labor market *versus* a father, assuming their earnings potential is equalized. Motivated by this target belief distribution, we develop a component of our survey to estimate these beliefs.[1]

Our first of four contributions introduces a new survey design to elicit beliefs on absolute advantage. We run our survey with parents in England recruited through Prolific. Participants are presented with vignettes where we pin down the earnings potential across a mother and father and vary which parent works longer hours in the labor market. We equalize earnings potential in the comparisons to capture expectations on a mother relative to a fathers role in child-rearing. To do so, each participant answers multiple scenarios in which the mother works longer hours at different wage rates and again in which the father works longer hours. We iterate the wage of the parent working longer hours in exactly the same way across gender and elicit participants' expectations on two domains: the hypothetical child's likelihood of graduating from university, as a measure of human capital accumulation, and their earnings rank at age 30, as a measure of long-term economic success.[2]

This design allows us to estimate how beliefs on children's future outcomes change within-individuals when mothers work long hours relative to fathers. Our first contribution, Result 1 in Section 2.3.3, is that, on average, people expect worse outcomes for children when mothers work long hours relative to fathers. Participants significantly reduce the expected likelihood of graduation by just under 1 percentage point (pp) and earnings rank by 0.67 percentile. These results demonstrate that even with the uncertainty of earnings differentials removed, beliefs on absolute advantage persist. We also provide additional evidence ruling out alternative interpretations, strengthening the validity of our method and interpretation.

Second, beliefs related to motherhood may form over the life-cycle around family narratives (Akerlof and Rayo, 2020) and subsequently affect labor supply decisions. To study this, we assess how beliefs vary by participants' background. In particular, we focus on the employment history of their own-mother and their own labor supply and birth penalty. We find that having a mother who worked full-time while they were young fully wipes out beliefs on absolute advantage consistent with a theory that family narratives shape children's perceptions into adulthood (Akerlof and

[1] We abstract from the role played by formal childcare as our primary focus is to provide an additional explanation to the observed motherhood penalty. Therefore, we focus on comparing men and women under the assumption that both face the same outside option in terms of formal childcare. See Boneva et al. (2022) for a study that examines beliefs about the impact of maternal labor supply on children and family outcomes, where the alternative to maternal time for the child is formal childcare.

[2] Attanasio, Cattan, and Meghir (2022), Boneva and Rauh (2018), and Kiessling (2021) also asked about expected earnings at age 30 for a hypothetical child but in a design related to returns to parenting practice. Additionally, we ask about the probability of graduating college as a measure more directly linked to parental time investment and long-term human capital accumulation.

Rayo, 2020). Moreover, this reinforces a literature on the inter-generational transmission of gender norms (Alesina, Giuliano, and Nunn, 2013; Fernández and Fogli, 2009). Next, we find that weaker expectations of mothers' absolute advantage in child-rearing predicts greater labor supply for women and a lower employment motherhood penalty.[3] Altogether, these results form our second contribution, Result 2 in Section 2.3.4: beliefs on absolute advantage have roots in childhood exposure to maternal labor supply and in turn impact labor supply and child-rearing decisions today.

Third, we study mechanisms that can give rise to variation in beliefs about absolute advantage and help explain the mental models people hold. Specifically, we highlight three key dimensions that may shape this distribution of beliefs. First, people may expect *differences in preferences*, believing that even with equal time available, men are less likely than women to allocate time toward investments into children's skills. Second, people may hold expectations on *differences in productivity of time investments*, where they perceive mothers as more productive than fathers for an equal amount of time spent on investments. Third, people may hold expectations on *differences in resource allocation*, where they expect mothers to allocate more resources to skill investments. If mothers are expected to have more resource control as they earn a higher share of the household budget, then this expectation would work to offset the beliefs on absolute advantage we just outlined.

To study whether any of these mental mappings matter, we introduce a new series of vignettes and randomize features across participants. The specific details are described in Section 2.3.5 and the evidence is summarized by Result 3. It is differences in expectations on time investments, but not other dimensions, that we find are important. We also show that this is particularly true for those with strong beliefs on absolute advantage. A mental model of differences in time preferences across genders is consistent with beliefs on absolute advantage originating from how women and men form preferences on free time. Differences in these preferences could then put significant pressure on women's labor market decisions, as it suggests a perceived tradeoff between maintaining a career at the expense of time investments to children. The gain in resources from a mother working would then need to offset the expected time loss. In Section 2.4.2.1, we study responses to an open ended question that further confirms this mental mapping is what respondents tend to have in mind.

Fourth, we complement our survey with an information treatment, providing evidence on whether beliefs about mothers are responsive to information. Given the intergenerational link we find for beliefs and how they predict women's labor supply, it is unclear whether such beliefs can be shifted with mere information. To test this, we construct an information treatment, drawing a new measure based on children's performance on tests when mothers work full-time.[4] Based on a longitudinal survey from the UK, we collect an incentivized expectation on the average share of passed qualification tests (GCSEs) taken during adolescence for children whose

[3]From life histories, we estimate the motherhood employment penalty for women in our sample. We show in the Appendix, Section B.4.1.3 that mothers in our sample have experienced a very similar penalty as documented in the literature by Kleven, Landais, and Leite-Mariante (2024).

[4]Our design is in the spirit of Haaland and Roth (2020) and Haaland and Roth (2023) who use incentivized beliefs as information experiments on beliefs about immigration and racial discrimination.

mother worked full-time during their primary school years.[5] We find that participants tend to underestimate how well the children of full-time mothers actually do. Next, half of the sample is randomly allocated to get the correct information, forming our treatment. Following this, we draw another incentivized quantitative belief as an outcome, but this time, we focus on the share of children with an abnormal level of behavioral problems. Participants also provide written responses explaining what guided their quantitative answer. We categorize these based on whether their written responses reflect worse outcomes when mothers work full-time. We also draw a scale of self-reported gender norms about the role of mothers.

Participants react to the information updating beliefs, leading to Result 4 in Section 2.4.2.1. We find that they reduce expectations on children's problem behaviors by 5.2% and consistently move from underestimating how well children do when mothers work full-time toward more accurate beliefs, *i.e.*, closer to the truth. Also, in their written responses, they shift views away from worse expectations for the child by 22%, and on self-reported norms, they move toward more liberal gender norms, by about 3% of the mean. Further, we show that belief updating in response to the information is robust to a range of concerns, including the risk of experimenter demand effects.[6] Finally, we find that when participants expect children to do worse with a full-time working mother, their written responses demonstrate this is driven by expectations of reduced time investments into the child, which is entirely consistent with our evidence forming Result 3. Information thus decreases this expectation in the written responses. Altogether, we find strong evidence that beliefs about children when mothers maintain careers can respond to information which corrects uncertainty improving misperceptions.

Next, we look at policy support. Given their close link to labor supply decisions, shifting these beliefs can raise awareness of the importance — or absence — of public family policies. We then examine how our intervention increases support for policies that help mothers return to work and balance career and family life. With our focus on a small fact as the information treatment, it is hard to expect large changes in policy views, because variation is likely more limited. Indeed, in Section 2.4.2.2, this is what we generally find. Results are only suggestive but go in the direction of increasing support for policies that may facilitate mothers returning to work. Further, in an obfuscated one-week follow-up, we look at additional policy support measures. While aggregate results are null, we find significant increases in support for more free childcare hours across multiple subgroups who held stronger views on absolute advantage measured in our main survey. Overall, the treatment weakly improves support for policies helping women with children work longer. We see this as a motivator for further work to understand how the intensity and type of information both move beliefs toward accuracy and updates political will.[7] Nevertheless, our

[5]The incentive is a cash bonus if the answer is within a small bandwidth around the actual pass rate.

[6]At the end of the main survey, we ask respondents to respond in a text box with what they thought the survey was about. We coded their answers based on whether they appeared to understand that the survey was about perceptions of mothers working and dropped these respondents in a robustness check.

[7]For instance, stories with accurate information over statistics may serve to be more salient and long-lasting in how they influence policy positions (Graeber, Roth, and Zimmermann, 2024).

evidence demonstrates that beliefs here are malleable.

Summary.

Altogether, our paper shows that beliefs on absolute advantage exist and have intergenerational links, giving weight to their perceived role in labor market decisions post child-birth. We then show that indeed these beliefs are linked with labor supply and child-rearing decisions. Next, we find that people expect mothers to put more free time into children's skill-building, suggesting their mental model of absolute advantage is around differences in the preferences that mothers and fathers hold. Information about how well children actually do when mothers work full-time then shifts people toward accurate perceptions and reduces the degree of harm they expect for the child. Thus, a degree of beliefs on absolute advantage centered around time preferences can be shifted by removing uncertainty through information.

Related literature.

Our work contributes to the literature on gender gaps and motherhood penalties (Blau and Kahn, 2017; Cortés and Pan, 2023; Kleven, Landais, and Søgaard, 2019). Under-representation of women in the labor market has economic consequences through being costly in terms of economic efficiency (Hsieh et al., 2019). We help understand one driver of mothers' sorting out of the labor market through societal beliefs on absolute advantage, and contribute by effectively estimating these beliefs. Additionally, we demonstrate the mechanisms people have in mind that inform these beliefs and represent targets for correcting uncertainty and misinformation.

Thus, we relate to the literature on gender norms and the role that they can play in constraining women's behavior and preferences for work (Blau and Kahn, 2017; Cortés and Pan, 2023). Recent evidence from Norway (Andresen and Nix, 2022) shows that motherhood penalties differ substantially between women in heterosexual couples and same-sex couples, suggesting that gender norms may play an important role. Perceptions of gender norms, however, can be incorrect. Progressiveness in one's country or local area is generally under-estimated (Bursztyn et al., 2023), and information about this misperception can lead to more positive views and an increase in women's labor supply (Bursztyn, González, and Yanagizawa-Drott, 2020; Cortés et al., 2024). Women also are expected to generally take more socially informed decisions than men even when actual attitudes are not different (Exley et al., 2024). Moreover, Settele (2022) shows that perceptions of the gender pay gap can be inaccurate but responsive to information about the actual size of the gap. So far, beliefs about women appear substantially heterogeneous but to some degree malleable. We turn attention to understanding the form and strength of beliefs on absolute advantage in child-rearing, what characterizes them, and whether correcting expectations on simple facts about how well children do when mothers work full time shifts beliefs toward more accuracy and less harmful expectations.

Our study is also related to a literature examining gender differences in decisions around work and job search. Wage growth in part-time relative to full-time work is often over-estimated and can bias decisions between full-time and part-time work (Backhaus, Schäper, and Schrenker, 2023; Blesch et al., 2023). This can be

important for gender gaps in labor markets, as women are typically observed to work fewer hours than men and are more likely to work part-time (Cortés and Pan, 2019; Goldin, 2014). Women also tend to sort into less demanding jobs in terms of working time (Maestas et al., 2023; Wiswall and Zafar, 2017), with job amenities important factors that women, more so than men, consider in the decision making about their job (Hotz, Johansson, and Karimi, 2018; Wasserman, 2022). Moreover, women tend to have weaker bargaining power and less optimism about future earnings (Card, Cardoso, and Kline, 2016; Cortés et al., 2023). Our study speaks to this literature, because beliefs on mothers' absolute advantage for children's development can explain why women sort into more flexible jobs requiring shorter hours and why they tend to hold weaker bargaining power. This is particularly salient given the expectations on time preferences that we estimate, where we find that women are expected to spend more time with children when free than men. This may act to pressure mothers out of work without substitutes to alleviate these expectations.

Additionally, we add to some recent work on the impact of paternity leave expansion (Farré et al., 2024) and on how fathers use time during paternity leave (González, Guirola, and Hospido, 2024). These studies show that in Spain children had more developmental delays after paternity leave increased, while fathers spent less time on developmental activities and more time on leisure. Our survey experiment expands this literature, showing that indeed people expect mothers to spend more time on investments than fathers. Thus, beliefs on women's absolute advantage may partly be based on observations of fathers' time-use, putting pressure on mothers to compensate. We then show that when expectations on children's outcomes for full-time working mothers are underestimated, information on actual performance can correct misconceptions toward accuracy.

We further contribute to a growing literature on parental time investments and parental beliefs about returns to parental time for children's skill development (Attanasio, Cattan, and Meghir, 2022; Boneva et al., 2022; Boneva and Rauh, 2018; Kiessling, 2021). Parental time with children is increasing in many countries (Aguiar and Hurst, 2007; Borra and Sevilla, 2019), due partially to increasing returns to education and competition in the education market (Ramey and Ramey, 2009). One recent study examines beliefs about the effects of mothers' decision to work on children's skill development (Boneva et al., 2022). They find that beliefs on children's skills and family outcomes increase when mothers move from no work to part-time work — effects partially driven by increases in income — but decrease when moving into full-time work. Our paper explores a related though different mechanism, by focusing on beliefs about absolute advantage thereby intentionally removing a mechanism operating via income effects. We then show that beliefs shift in response to information and that this works through a shift in the mental model people have in mind.

The remainder of this paper establishes our conceptual framework in Section 2.2 and then moves through each of our four main results. In Section 2.3.1, we describe our sample and, through the rest of Section 2.3, we describe our survey design and estimation of beliefs on absolute advantage, as well as the mechanisms outlining Results 1 to 3. We then finish the paper in Section 2.4, assessing responses to our information treatment establishing Result 4.

2.2 Conceptual Framework

Our first objective is to effectively measure beliefs about women's absolute advantage in child-rearing — referred to throughout the paper as 'women's absolute advantage'. To provide structure for our analysis, we develop a simple Beckerian model of household labor division. The model provides us with an economic framework to interpret these beliefs and guides the development of a target beliefs distribution for estimation.

The model draws on Siminski and Yetsenga (2022) and assumes a household that allocates parental time between the home and the workplace. The home-time of the mother (m), the father (f), together with earnings (e), feed into their child's human capital production function as follows:

$$\text{Child human capital:} \quad HC(m,f,e) = m^{\rho_m} f^{\rho_f} e^{\rho_e}.$$

Here, ρ_m, ρ_f, and ρ_e represent the household's beliefs regarding the elasticity of maternal time, paternal time, and earnings, respectively, in producing human capital. We say that the mother has an *absolute advantage in child rearing* if $\rho_m > \rho_f$. Note that this definition incorporates both cases where the mother is more effective at child-rearing per hour spent and cases where mothers allocate more of their home-time to child-rearing activities. Indeed, in section 2.3.5 we use our survey design to disentangle these two sources of absolute advantage.

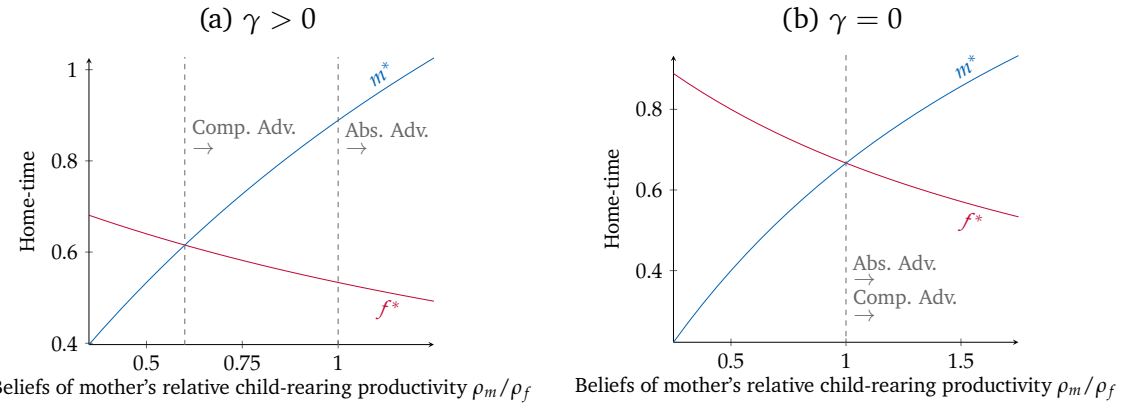
Each parent has an endowment of one unit of time, which can be allocated either to the home or the workplace. The father earns a wage rate of W, while the mother earns $(1-\gamma)W$. The parameter $\gamma \in [0,1]$ reflects the earnings gap between the mother and father, for instance due to the existence or lack of family-friendly workplace policies. The household's budget constraint is given by:

$$\text{Budget constraint:} \quad e = (1-m)(1-\gamma)W + (1-f)W.$$

The key insight from the model is that household members will specialize in market work or home-production according to *comparative advantage*. This can stem either from differences in market productivity (captured by γ) or from differences in presumed child-rearing productivity (captured by ρ_m and ρ_f). In Figure 2.1, we highlight this in two empirically relevant cases, with the formal derivation presented in Appendix B.1.1. Panel (a) of Figure 2.1 demonstrates the case where $\gamma > 0$, hence the mother faces a wage penalty, or equivalently, fathers have an absolute advantage in market work. In this scenario the mother specializes in home-production even in the absence of gendered beliefs (*i.e.*, when $\rho_m = \rho_f$). In contrast, panel (b) shows the case where $\gamma = 0$, so that wage potentials are equalized. Here, the mother will only undertake the majority of home-production responsibility if she is believed to have an *absolute advantage in child-rearing* (*i.e.*, if $\rho_m > \rho_f$). Insofar as people hold beliefs that mothers' have an absolute advantage in child-rearing, policies aimed at closing the earnings potential gap of mothers and fathers (captured by a reduction in γ) will not be sufficient to equalize labor market outcomes. Thus, beliefs on mothers' absolute advantage are highly relevant for policy outcomes.

Our survey experiment allows us to isolate beliefs of absolute advantage from

Figure 2.1: Optimal Child-Rearing Allocation under Different Beliefs



Note: *Comp. Adv.* and *Abs. Adv.* refers to the mother's comparative and absolute advantage in child-rearing.

beliefs of comparative advantage, by holding earnings fixed in scenarios where we vary whether the mother works longer hours ($MWL = 1$) or the father works longer ($MWL = 0$). Using the language of the model, the beliefs that we target empirically can be written as

$$\theta_{i,e} := \widetilde{HC}_i(\underbrace{h_s, h_\ell}_{MWL=1}, e) - \widetilde{HC}_i(\underbrace{h_\ell, h_s}_{MWL=0}, e). \tag{2.1}$$

$\widetilde{HC}_i$ represents person i's beliefs of the human capital of a child growing up with family income e and parental home-time inputs h_ℓ, h_s representing long and short hours respectively such that $h_\ell > h_s$. It is easy to show that $\theta_{i,e} < 0$ *if and only if* $\rho_m > \rho_f$. Hence, empirically testing the sign of $\theta_{i,e}$ is equivalent to testing whether the mother is believed to have an absolute advantage in child-rearing in the model, giving our empirical results a close model analogy. In the next section, we describe how our survey experiment is structured to capture these beliefs and to investigate the mechanisms driving them.

2.3 Hypothetical Beliefs Elicitation: Design and Results

In this section, we address five key points. First, we define our sample selection, recruitment, and demographics. Second, we present a hypothetical design through vignettes to elicit beliefs on a child's future outcomes when a mother works longer hours relative to a father. Third, we describe our estimation strategy and results to study within-person average estimates relevant to equation (2.1). Fourth, we describe how we empirically extract individual perceptions, by approximating an individual-level measure of equation (2.1), and how these vary across individuals' characteristics. Fifth, we investigate channels that can give rise to these beliefs.

2.3.1 Sample

We conducted our experiment on the online platform Prolific, recruiting 1,056 participants.[8] We had two main inclusion criteria, requiring participants to be (i) parents of at least one child aged 18 or below, and (ii) currently residing in England.[9] Throughout this paper, we follow our pre-registered analysis plans with some minor deviations on extended results. We point these out where relevant and describe them further in Appendix Section B.5.

We contrast our participants' demographics with current parents living in the United Kingdom using the latest wave (2022) of Understanding Society (US 2022). For comparison purposes, we restrict the US 2022 sample to parents of at least one child aged 18 or below and who live in England. We use only the latest wave of Understanding Society to be as close as possible to contemporaries of our respondents. In the Appendix, Table B.1, we show that our sample is similar to Understanding Society on some dimensions but over-sampled on higher education and monthly net earnings. Later, we will also show that our results are not fully driven by those with high education or income. Additionally, we will re-weight some of our key analyses in robustness checks, showing that our evidence and conclusions are unaffected. These weights are constructed with a standard "raking" procedure described in the Appendix, Section B.2.1. Finally, based on life histories, we show that mothers in our sample have experienced an average $29pp$ drop in employment probability post-child birth (see Figure B.10 and Section B.4.1.3). This is entirely consistent with the 25-30% employment penalities observed across the US and Europe by Kleven, Landais, and Leite-Mariante (2024), indicating that our sample looks very similar in terms of employment experiences and parenthood relative to the wider population.

2.3.2 Hypothetical Design

Framing.

We use six hypothetical scenarios in vignettes to elicit participants' beliefs on children's human capital accumulation in response to women *versus* men working longer hours in the labor market. The following is the text participants see to set the stage for the scenarios, and we further provide screenshots of the online survey in the Appendix, Subsection B.4.2.

> We are interested in your beliefs about children's future outcomes, comparing families with different financial resources and time demands.
>
> **Setup:** Please imagine an average family in your community. Suppose this family consists of a father and a mother who are both employed, and they have a boy (girl, *randomized*) who is aged 10 (4, *randomized*). Suppose household expenditure decisions are made jointly by the father

[8] The survey design is browser-based and built using the oTree framework (Chen, Schonger, and Wickens, 2016).

[9] We focus on England only, as later in the surevy we will use a metric (GCSE pass rates, see Subsection 2.4.1) mainly known in England. In other countries of the United Kingdom, the names, content, as well as the exam requirements of GCSEs are typically different.

> and the mother, and this hypothetical family spends 10% (20%, *randomized*) of their total income on the child's educational and extracurricular activities such as clubs, tutoring, music, sports, etc.
>
> We will show you different scenarios, and ask your opinion about the likelihood that the child will be successful in education and the labor market. There are no clear right or wrong answers, and we know these questions are difficult. Please try to consider each scenario carefully and tell us what you believe the likely outcomes will be.

Randomization in the setup.

We randomize several features in the setup. These are whether the participant reads that the family has a boy or a girl, the age of the child (4 years old *versus* 10), and the share of income (10% *versus* 20%) spent on the child's educational and extracurricular activities (denoted by SSE_i below). These randomized features enable us to assess whether participants paid attention to the vignettes, and to later assess whether beliefs differ across these features. Table B.3 in the Appendix confirms that these features are balanced across participants.

Scenarios and outcomes.

Next, for each participant, we iterate through a set of scenarios (six in total) — presenting three scenarios per page — and varying two components: (i) whether the father or mother works longer hours, and (ii) the hourly wage of the parent who works longer hours. An example scenario is as follows:

> The **father** works **35** hours per week at a wage of **£12** per hour.
>
> The **mother** works **42** hours per week at a wage of **£17** per hour.

We then ask each participant their beliefs on the probability that the hypothetical child will eventually graduate from university, using a 0–100 scale with a slider. Additionally, we ask them for the child's earnings rank at age 30 relative to other 30-year-olds in terms of percentile rank using a 1–99 scale on a slider.[10] We iterate on the scenarios, and at each, re-collect these expectations/beliefs for those two dimensions. Example images of what the participants see here are presented in the Appendix, Section B.4, Figures B.14, and B.15.

Randomization in the scenarios.

Table 2.1 below contains the design for iterating through scenarios. Participants work through two pages, one for a mother and one for a father working longer hours, with each containing three scenarios. Importantly, the wages they see when a mother works longer hours will be exactly the same as in scenarios with the father. To avoid order effects, we randomize whether each participant starts with the man or woman

[10]To familiarize participants with the scaling used throughout the survey, we provide them with an "introduction to scale" (see Figure B.12 in the Appendix), common to all participants, before displaying the hypothetical scenarios.

working longer hours. We also randomly draw the ordering of wages shown within each page so that participants do not move sequentially through lower to higher wage changes. In all cases, we hold constant the wage of the parent working fewer hours.

We further randomize whether the wage profile of the hypothetical parent working longer hours has a lower bound of either £12 or £17 and an upper bound of either £22 or £27. This allows across participants for the overall wage profile to range from £12 to £27. We contrast the distribution of weekly household labor income across wage profiles in our design with the distribution drawn from the 2022 wave of the Family Resources Survey. Overall, Figure B.8, in the Appendix, shows that we have good coverage over this distribution in England — although our hypothetical distribution does not cover the top 25% of the earnings distribution.

Finally, the weekly number of hours worked is randomized across participants. Half of the sample sees both parents working full-time with one of them working longer hours (42 *versus* 35 hours per week), and the other half sees a full-time working parent and a part-time working parent (36 *versus* 20 hours per week). The former is referred to as the "FT–FT" design, while the latter is referred to as the "FT–PT" design. We will use this later for heterogeneity.

Table 2.1: Design of Hypothetical Scenarios

	Man Works More		Woman Works More	
	w_m	w_f	w_m	w_f
$k=1$	£17 (£12)	£17 (£12)	£17 (£12)	£17 (£12)
$k=2$	£22 (£17)	£17 (£12)	£17 (£12)	£22 (£17)
$k=3$	£27 (£22)	£17 (£12)	£17 (£12)	£27 (£22)

Notes: This table presents the design of our hypothetical scenarios, where w_m is the man's hourly wage, and w_f is the woman's. Participants here are randomized into either the higher or lower wage profile (in parentheses).

Attention and confidence.

First, we regress each of our collected expectations (graduation likelihood and earnings rank) on the randomized features in the vignette setup and a pre-registered set of controls. Results are reported in Table 2.2.[11] We see strong responses on a number of design features consistent with our participants paying attention to the design details. Particularly, seeing a large share of the family budget allocated to educational activities for the child or seeing a higher wage profile strongly increases positive expectations. Second, we follow Haaland, Roth, and Wohlfart (2023) to test participants' attention to the survey and confidence in their answers. Before completing the hypothetical scenarios, we provide participants with a paragraph of text, wherein we ask them to report that their favorite color is "turquoise". Below this

[11]Where participants' characteristics are controlled for in this study, we use the following pre-registered set: participant's gender, a quadratic in age, an indicator for whether they have at least a university degree, employment status (full-time *versus* part-time or less), and ethnicity (white *versus* non-white).

paragraph, we ask participants "what is your favourite colour?". In our survey, 95% (1,003) of our participants passed this attention check, suggesting strong attention to our survey. Next, after the hypothetical vignettes, we ask participants to what extent they are sure about their answers. 75% of participants (795) reported being at least somewhat sure of their answers.[12] Later, we perform robustness checks (see Subsection 2.3.3) using these screeners, to test the reliability of our estimates.

Table 2.2: Design Effects Across Participants

	(1) ℙ(graduate)	(2) Earnings Rank
Child is a girl	1.019	1.858*
	(1.06)	(1.00)
Child is aged 4	-0.107	0.451
	(1.07)	(1.02)
SSE_i: 20%	2.807***	1.370
	(1.06)	(1.01)
FT–FT profile	2.107**	2.077**
	(1.06)	(1.01)
High wage profile	7.658***	5.280***
	(1.07)	(1.00)
Mother shown first	0.050	0.569
	(1.06)	(1.00)
Participants	1056	1056
Observations	6336	6336
Individual Controls	Yes	Yes

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on individuals. The probability to graduate (ℙ(graduate)) is scaled between 0 and 100. Earnings Rank is the percentile ranking expected for the child at age 30 among other 30 year-old. SSE_i is an indicator for seeing the share of budget spent on educational expenditures at 20% instead of the 10% in the vignette setup. The FT–FT design presents both parents as full-time with one working longer hours (42 *vs*. 35). The "mother shown first" variable is equal to 1 when scenarios with $MWL = 1$ (mother works longer hours) were shown first or 0 when scenarios with $MWL = 0$ (father works longer hours) were shown first. Individual controls include the pre-registered set of participants' characteristics.

2.3.3 Results: Hypothetical Beliefs Elicitation

We now test whether beliefs about children's future outcomes vary based on whether in a family the mother or the father works longer hours.

[12]We provide screenshots of the attention check and confidence questions that participants actually see in the Appendix, Section B.4 (see Figures B.13 and B.20).

Empirical strategy on gendered beliefs.

Empirically, we provide estimates for the within-person average difference in beliefs, holding constant the earnings potential of the mother and the father. We approximate an average related to the individual measure we defined in equation (2.1) of the conceptual framework. Based on our design, this leads to the following estimation target:

$$\delta = \frac{1}{N}\frac{1}{K}\sum_{i=1}^{N}\sum_{k=1}^{K}\theta_{i,k}(MWL),$$

where we look at a within person average difference in a child's future human capital accumulation over K different levels of earnings potential. In the survey, our collected expectations on a child's future outcomes ($y^o_{i,j,k} = [y_{i,j,k}{}^{\text{graduation}}, y_{i,j,k}{}^{\text{rank}}]$) vary within individuals and the wage levels (k). These wage levels change in exactly the same way for scenarios with a mother working longer hours ($MWL_{j=1}$) versus a father ($MWL_{j=0}$). We aggregate the within-person difference in these measured beliefs over MWL_j. If there are no gendered beliefs, then the average change in beliefs will be the same ($\hat{\delta} = 0$), regardless of who works the longer hours. While an estimate of $\hat{\delta} < 0$ will be consistent with beliefs that women hold an absolute advantage. In this case, average expectations are that it is more harmful for children's human capital accumulation if women work longer hours compared to men. Following our pre-registration the main specification is:

$$y^o_{i,j,k} = \alpha_0 + \delta MWL_j + \tau_k + \mu_i + \epsilon_{i,j,k}. \tag{2.2}$$

Participant fixed effects are captured by the vector μ_i and vignette household income fixed effects by τ_k.[13] In some specifications, we replace μ_i with the pre-registered set of participant's characteristics, which are the following: gender, a quadratic in age, an indicator for whether they have at least a university degree, employment status (full-time *versus* part-time or less), and ethnicity (white *versus* non-white).[14]

Average estimates of gendered beliefs.

Results for each outcome (graduation likelihood, earnings rank) based on equation (2.2) are presented in Table 2.3.

Result 1 *Beliefs on children's future outcomes are on average worse when a mother works longer hours compared to when a father works longer hours for the same wage.*

Our estimates of $\hat{\delta}$ return significant and negative effects for scenarios with the mother working longer hours. In these scenarios, participants reduced their expected probability that the child will graduate university by nearly 1 pp and earnings rank at age 30 by about 0.67^{th} of a percentile. While these magnitudes are not large, they suggest that even when earnings potential is equalized between mothers and fathers,

[13]Note: by construction the MWL indicator can only vary within individuals, thus including or omitting individual fixed effects will not change the estimates, as we demonstrate in results.

[14]Note: 6 respondents listed "other" or "prefer not to say" for gender. We set these to 0 and control for an indicator flagging them.

Table 2.3: Beliefs About Mothers Working Longer

	(1) $\mathbb{P}$(graduate)	(2) $\mathbb{P}$(graduate)	(3) Earnings Rank	(4) Earnings Rank
$MWL_{j=1}$	-0.933***	-0.933***	-0.668**	-0.668**
	(0.299)	(0.299)	(0.268)	(0.268)
Mean Dep. Var	56%	56%	49^{th}	49^{th}
Participants	1056	1056	1056	1056
Observations	6336	6336	6336	6336
Individual Controls	Yes	No	Yes	No
Individual Fixed Effects	No	Yes	No	Yes
Scenario Income Fixed Effects	Yes	Yes	Yes	Yes

Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. Standard errors are clustered on individuals. Expectations on the child's probability to graduate ($\mathbb{P}$(graduate)) range between 0 and 100 with a mean of 56 representing a 56% expected likelihood. The expected percentile earnings rank when the child is 30 years old lies between 1 and 99. Individual controls include the pre-registered set of participants' characteristics.

hesitancy over the mother working longer hours may remain. This is consistent with beliefs that mothers can hold an absolute advantage in child rearing, and it forms our main result of Section 2.3.

In the remainder of this section, we further unpack and add context to our interpretation of these results. We additionally consider some heterogeneity, but to maintain focus, we defer comment until later.

Interpretation and discussion.

While our preferred interpretation of Result 1 is that people hold beliefs of women's absolute advantage in child-rearing, a potential alternative explanation is that participants inferred unobserved characteristics of the hypothetical families from the vignettes. Most notably, they might associate a father working part-time with lower family-level human capital than when the mother works part-time. If so, responses could reflect beliefs about the intergenerational transmission of human capital rather than the direct effect of mothers working longer hours. We do not believe, however, that this alternative interpretation is plausible. If it was, we would not expect to find a link between beliefs and actual labor supply of respondents or any intergenerational link between beliefs and the labor supply of respondents' mothers. Yet, we will show in Section 2.3.4 that in fact beliefs are highly correlated both with respondents' own labor supply and the labor supply of their mothers growing up. Additionally, we offer further discussion and empirical evidence that runs against this alternative interpretation in the Appendix, Section B.2.2.1.

Heterogeneity analysis.

The heterogeneity analysis at the heart of this paper is an observation of strong persistence in perceptions of absolute advantage: women whose mothers worked part-time during their upbringing are more likely to hold such beliefs, and individuals

with beliefs of women's absolute advantage are more likely to have lower labor supply and suffer larger motherhood penalties. We will discuss this finding in detail in Section 2.3.4.

Before turning to this, we briefly highlight a few interesting findings from our pre-registered heterogeneity analysis. One striking result is that the negative effects of mothers working longer hours on expected child outcomes are concentrated in the lowest income quartiles and become statistically insignificant in the top quartile. This is visualized in Figure B.1, which groups scenarios by the hypothetical family's income level. Note that these income levels overlap reasonably well with the actual UK family income distribution, as shown in Figure B.8. Our interpretation of this result is that people view time and money as substitutable inputs into child human capital. While maternal time is considered an important input at lower income levels, higher earning families are able to substitute for this with pecuniary resources. In Appendix B.1.2 we make this point formally by showing that an extension to our baseline model where time and monetary inputs are gross substitutes and exhibit decreasing returns to scale successfully replicates the observed pattern: the effect of mothers working longer decreases as household earnings increase.

Furthermore, across participant characteristics (Appendix, Tables B.4 and B.5), we find that the negative effects in MWL scenarios are somewhat stronger among men, those born outside the UK, university degree holders, part-time or less employed participants, and those who voted conservative, other, or none, at the last UK General Election. While the university result appears surprising, less surprisingly, beliefs on absolute advantage appear to be strongest within more conservative groups. Finally, in Table B.6, we assess heterogeneity in the MWL effect by design features. For brevity, we provide a discussion in Appendix Section B.2.2.3.

Robustness checks.

Finally, we test the robustness of our key finding (Result 1) by implementing different sample restrictions and checks, and report these in the Appendix, Table B.7. First, we exclude those who reported being unsure or very unsure about their answers to the vignette scenarios. Second, we exclude participants who did not pass the attention check. Third, as an additional check against inattention, we exclude participants with the 5% lowest and highest response times. Fourth and last, we re-weight our sample to match the national population distribution (see Subsection B.2.1). Our main result is robust to all of these checks, with the coefficients on MWL_j in Table B.7 about the same magnitude as the ones we find in Table 2.3.

2.3.4 Labor Supply and the Intergenerational Transmission of Beliefs

Having established that people *on average* hold beliefs of women's absolute advantage in child-rearing, we now want to explore where these beliefs come from and how they influence individual behavior. In this section, we therefore study whether beliefs of absolute advantage vary with early-life exposure to maternal employment, interpreted as a role model effect, before turning to whether these beliefs predict individuals' own labor supply choices.

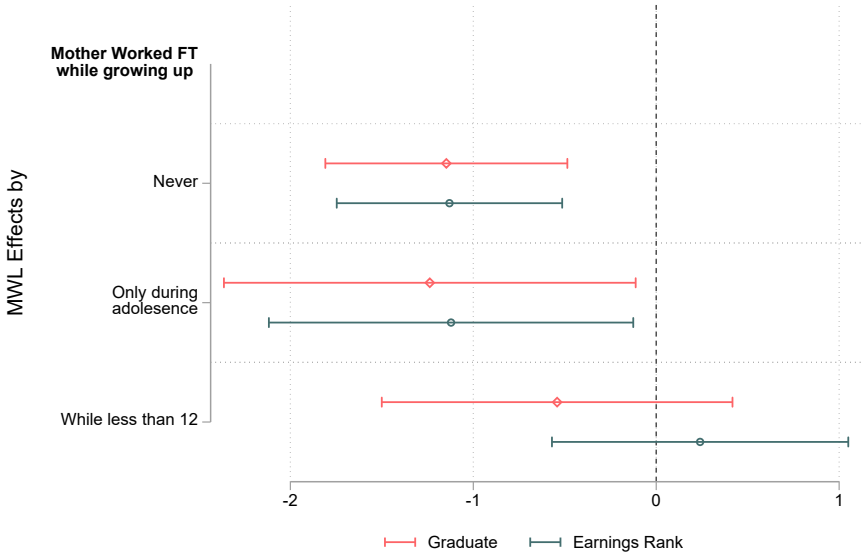


Figure 2.2: Role Model Effects on Beliefs of Absolute Advantage

Notes: Bars represent 95% confidence intervals, where standard errors are clustered on individuals. Each coefficient is obtained from separate OLS regressions estimating equation (2.2) on the subsample defined by the displayed label and for our two expected outcomes: (i) the probability for the child to graduate ($\mathbb{P}$(graduate)), and (ii) the earnings rank of the child at age 30 (Earnings Rank). "FT" stands for full-time.

Effect of own mother's labor supply on beliefs.

Beliefs related to identity likely form over the life cycle, and theoretically should include the family narrative (Akerlof and Rayo, 2020; Bénabou and Tirole, 2002). In this light, beliefs on absolute advantage may well persist from childhood experiences. We find strong support for the hypothesis that beliefs of absolute advantage are correlated with childhood experiences. In our survey, respondents were asked about their mothers' labor supply when they were growing up, distinguishing between whether the mother was working full-time or part-time during childhood (<12 years old) and adolescence (12–18 years old). Figure 2.2 presents the results on heterogeneity in $\hat{\delta}$ based on these experiences.[15] Respondents whose mothers never worked full-time during their upbringing hold significantly stronger beliefs of absolute advantage, with lower expectations for both the earnings rank and graduation probability of children with mothers working longer compared to fathers. In contrast, for individuals whose own mothers worked full-time before they turned 12, we find no evidence that they hold beliefs of mothers' absolute advantage. These results point to a strong role model effect, adding to a literature on the intergenerational transmission of gender norms through maternal behavior (Alesina, Giuliano, and Nunn, 2013; Fernández and Fogli, 2009). Given this intergenerational link, we next look at whether beliefs inform own labor supply decisions when grown.

[15]Appendix Table B.5 presents the full regression results.

Individual perceptions

We aim to study how individual beliefs of women's absolute advantage in child-rearing shape actual labor supply behavior. To do so, we construct an individual-level measure of perceptions of women's absolute advantage that will be used to predict labor supply behavior, aggregating elicited beliefs to individual average perceptions. In our design, this equates to the following formulation:

$$\theta_i = \frac{1}{K}\sum_{k=1}^{K}\theta_{i,k}(MWL),$$

where we have collapsed the difference in beliefs around MWL_j to the individual-level.

For each expectation outcome o, we label the individual perceptions of mothers working longer hours compared to fathers as ${\theta_i}^{\text{graduate}}$ for the probability of the child to graduate from university and ${\theta_i}^{\text{rank}}$ for the earnings rank at age 30. For each of these θ_i^o measures, the scale is increasing in more positive views about children's future outcomes when women work longer hours relative to fathers, with 0 implying no expected difference. We report the distribution of these measures in the Appendix, Figure B.2, and in the Appendix, Figure B.3, we show that these two measures are consistent with one another.

Effect of beliefs on labor supply.

We now use the estimated individual perceptions and examine how they relate to participants' actual labor supply behavior. Our outcomes consist of self-reported individual information on two current labor supply measures from our end-of-survey questionnaire — weekly working hours and an indicator for working full time. We also use our continuous measure of the employment motherhood penalty as described in Subsection B.4.1.3, in the Appendix. Results are presented in Table 2.4, where we first regress each outcome on the individual belief measures (considering θ_i^{grad}, and θ_i^{rank} in separate regressions) including our standard control set (column 1 and 2). Additionally, we follow the method designed by Gillen, Snowberg, and Yariv (2019) to isolate a joint beliefs measure from the two domains and estimate the correlation between this measure and our outcomes (column 3). This method relies on an IV approach to minimize measurement errors arising from estimating regressions on a relatively small sample, as well as leveraging the two distinct domains we consider (college graduation and earning rank) to isolate the common underlying variation.[16]

[16]Kiessling (2021) provides a similar application of this method by relating participants' elicited perceived returns to investments in neighborhood and parenting styles to actual measures of neighborhood assessment and parenting behavior. The IV strategy follows Gillen, Snowberg, and Yariv (2019), and consists in first duplicating the number of observations, then using the estimated θ_i^{grad}, and θ_i^{rank} once as a regressor and once as an instrument. For each of our outcomes, we estimate the following equations, which yields an estimate of δ_1.

$$\begin{pmatrix} y_i \\ y_i \end{pmatrix} = \begin{pmatrix} \delta_0^{grad} \\ \delta_0^{rank} \end{pmatrix} + \delta_1 \begin{pmatrix} \theta_i^{grad} \\ \theta_i^{rank} \end{pmatrix} + \begin{pmatrix} \delta_2^{grad} X_i \\ \delta_2^{rank} X_i \end{pmatrix} + \nu_i \tag{2.3}$$

Table 2.4: Relationship Between Beliefs and Participants' Labor Supply

	Hours Worked			Full Time			Motherhood Penalty		
Panel A: Women	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
θ^{graduate}	0.844			0.023			0.026*		
	(0.541)			(0.020)			(0.014)		
θ^{rank}		1.229**			0.040**			0.034**	
		(0.591)			(0.020)			(0.014)	
$\theta(IV)$			1.546***			0.047**			0.044***
			(0.588)			(0.021)			(0.016)
Observations	525	525	1050	525	525	1050	514	514	1028
Mean Dep. Var.	26.12	26.12	26.12	0.39	0.39	0.39	-0.19	-0.19	-0.19
Panel B: Men									
θ^{graduate}	-0.090			-0.014					
	(0.563)			(0.017)					
θ^{rank}		-0.046			0.007				
		(0.441)			(0.018)				
$\theta(IV)$			-0.123			-0.006			
			(0.644)			(0.023)			
Observations	525	525	1050	525	525	1050			
Mean Dep. Var	37.74	37.74	37.74	0.86	0.86	0.86			

Notes: p <0.10; ** p <0.05, *** p <0.01. Standard errors, in parentheses, are bootstrapped with 1,000 replications. θ values are standardized to z-scores. We run OLS regressions of three labor supply measures as indicated in the headings on individual perceived returns over (1) the probability for the child to graduate (θ^{graduate}), (2) the expected earnings rank of the child at age 30 (θ^{rank}), and (3) a joint measure from separate IV estimates yielding a measurement error-corrected latent estimate across the two other measures. The implementation of this IV strategy requires duplicating the number of observations. Hours Worked corresponds to the participant's weekly number of hours worked. Full time corresponds to an indicator set equal to one if the participant is currently employed full time. Motherhood Penalty is our continuous measure of motherhood penalty in employment. Panel A reports results for women, and Panel B for men. Individuals controls include a quadratic in age, an indicator for whether they have at least a university degree, and ethnicity (white *versus* non-white). Six participants listed "other" or "prefer not to say" for gender are excluded.

Despite substantial labor market convergence, men still exhibit stronger labor market attachment than women (Olivetti, Pan, and Petrongolo, 2024), both in terms of labor force participation and working time. Gender gaps in wage rates largely favor men, and women — especially married women — exhibit higher labor supply elasticity than men (Chetty, Friedman, and Saez, 2013; Eissa and Hoynes, 2004; Keane, 2011). All of these factors may lead men's labor supply to be less sensitive to changes in beliefs. Additionally, men are already largely sorted into full-time jobs, leaving less margin for adjustment. Women, however, may respond more strongly to changes in beliefs in terms of observed labor supply behavior both because of their different position in the labor market and because these beliefs are more salient and directly apply to women's labor supply choices. Our evidence in Table 2.4 is consistent with this interpretation. More positive perceptions of women working yields

$$\text{instrumenting} \begin{pmatrix} \theta_i^{grad} \\ \theta_i^{rank} \end{pmatrix} \text{ with } Z = \begin{pmatrix} \theta_i^{rank} & 0_N \\ 0_N & \theta_i^{grad} \end{pmatrix}.$$

heterogeneous predictions across women and men: they predict higher labor supply for women only, with point estimates much smaller and generally close to zero for men. Although qualitatively consistent across the two domains, these correlations for women are stronger when we use beliefs on earnings rank, and generally stronger when we account for potential measurement errors in the IV exercise. In terms of magnitude of effects, if we focus on the joint estimate in column (3), a one-standard-deviation increase in positive perceptions of women working corresponds to a 1.546-hour increase in weekly worked hours — equivalent to about a 6 percent increase relative to the average.[17] Similarly, the same increase in perceptions is associated with about 5pp increase in the likelihood of working full time — equivalent to 13 percent relative to the average — and a 4pp reduction in the employment motherhood penalty- equivalent to 21 percent relative to the average.

For men, even without changes in labor supply, beliefs could affect how they allocate their time through substituting leisure for time invested into children's skill development. To test whether this is the case, we estimate the same specification as in Table 2.4, examining how perceptions of women working relate to parents' self-reported time spent on skill-enhancing and outdoor activities. We report these results in Table B.8 in the Appendix. Results point to a link between more positive views on mothers' working and the time fathers spend in skill-enhancing activities and complement the results on mothers' labor supply: a one-standard-deviation increase in positive perceptions of women working corresponds to almost half an hour increase in time spent in activities helping children develop their skills — equivalent to about an 11 percent increase relative to the average. While we cannot rule out all forms of potential biases, overall our results speak to the link between our elicited measure of beliefs and actual labor market choices.

The measures of individual-level beliefs suggest that perceptions of mothers' absolute advantage in child rearing are predictive of women's actual labor supply behavior — but not men's. More positive perceptions of women working are associated with stronger labor market attachment across all three of our labor supply measures: working hours, likelihood of full-time employment, and the motherhood employment penalty.

Summary.

We find that respondents own mother's labor supply predict beliefs of absolute advantage, and that beliefs of absolute advantage in turn are predictive of women's labor supply. Taken together, this suggests that beliefs of women's absolute advantage in child-rearing can be an important driver of the intergenerational persistence in labor supply from mother to daughter. We summarize these findings in Result 2 below. Next, we turn to elicit views on the channels that may explain these beliefs on absolute advantage.

Result 2 *Beliefs of women's absolute advantage in child-rearing are driven by individuals whose own-mothers worked part-time during their upbringing. Furthermore,*

[17]Hours worked include those reporting zero hours (8% of the sample). Results remain qualitatively consistent and similarly significant when we restrict the sample to observations with positive working hours.

when these beliefs are weaker (stronger) they predict higher (lower) labor supply among women.

2.3.5 Mechanisms for Variation in Beliefs

Now, we consider what channels may give rise to variation in the beliefs distribution we outlined in equation 2.1 and document in Result 1. We propose three main possibilities. First, people may hold beliefs about differences in *preferences*. This would imply that people believe mothers and fathers hold different valuations for time spent outside of work, whereby they expect women value spending more of their free time investing in a child's skills than do men. Second, people may hold beliefs about differences in the *productivity* of time investments. In this case, they may presume mothers have an absolute advantage because they believe mothers are more productive in producing a child's skills than a father for the same amount of time spent. Third, people may believe that for a given budget, the parent who works longer, or earns more, makes the *resource allocation* decisions for monetary investments to a child. A presumption that mothers will allocate more of the budget to these monetary investments would push in the opposite direction of the other two channels.[18] Here, we investigate each of these potential mechanisms to understand the mental models informing beliefs about absolute advantage.

To quantify the importance of these mechanisms, we introduced new vignettes where key features were randomized across participants to target each of the proposed channels. For each channel, we briefly describe the experimental setup below, while the full survey details are in the Appendix, Subsection B.4.2 — in particular, see pages 5 to 8.

We will also assess heterogeneity for each of these based on beliefs of absolute advantage. To do this parsimoniously, we define a collapsed belief measure, θ_i, as the simple average of the individual-level belief differences in expected graduation likelihood (θ_i^{grad}) and expected earnings rank (θ_i^{rank}). That is,

$$\theta_i = \frac{1}{2}(\theta_i^{\text{grad}} + \theta_i^{\text{rank}}).$$

This measure captures an individual's overall belief about whether children fare worse when mothers, rather than fathers, work longer hours, and is used throughout this section and the next to split participants into those holding more negative views ($\theta_i < 0$) versus those who do not ($\theta_i \geq 0$).[19]

Finally, we close the section by providing more evidence in contrast to the alternative interpretation of Result 1, which we previously discussed. Here, we test the possibility that people react to the scenarios based on different expectations about the likelihood of mothers and fathers holding a university degree based on who works longer hours rather than expectations about different working arrangements within the couple.

[18] Evidence from the literature generally rejects the income pooling hypothesis, indicating that when mothers are in charge of resources (as opposed to fathers), expenditures on children tend to increase (Bobonis, 2009; Hoddinott and Haddad, 1995; Lundberg, Pollak, and Wales, 1997).

[19] This approach simplifies the presentation, as the two components are conceptually aligned and highly correlated (Appendix Figure B.3).

Expectations on time preferences: design.

Beliefs on differences in preferences imply people will expect a mother to spend more time on activities with their child than a father given the same free time. To investigate this, we present respondents with a child aged 11 who will soon take the Key Stage 2 national test.[20] We randomize across participants whether both (*versus* neither) hypothetical parents have a university education and, importantly, whether the father (mother) has a busy week ahead with only the mother (father) free to help. We then ask how much time they expect will be spent helping the child study for the test and how much time they expect will be spent on extracurricular activities.[21] All participants are informed of the average time (30 minutes) spent per week on teaching activities by parents in the 2014-2015 UK Time Use Survey to give them a common contextual reference.[22]

Table 2.5: Expectations on Time

	(1)	(2)	(3)	(4)	(5)	(6)
	All Participants		By θ			
			< 0		≥ 0	
Panel A: Time spent on test help						
Mother (*father*) free to help	12.945*	6.299	18.492*	-1.266	5.301	10.901
	(6.816)	(9.467)	(9.747)	(13.856)	(9.872)	(13.549)
Both parents (*neither*) have a university education	37.728***	30.843***	32.876***	11.545	42.964***	48.532***
	(6.890)	(9.957)	(9.823)	(14.143)	(9.870)	(14.101)
Both parents have a uni education × Mother free to help		13.603		42.239**		-11.006
		(13.615)		(19.256)		(19.420)
Mean Dep. Var	149.242	149.242				
Panel B: Time spent on extracurricular	(1)	(2)	(3)	(4)	(5)	(6)
Mother (*father*) free to help	2.962	-8.204	4.998	-12.234	-0.448	-4.990
	(6.760)	(9.303)	(9.700)	(13.052)	(9.667)	(13.492)
Both parents (*neither*) have a university education	38.096***	26.528***	33.010***	14.406	41.961***	37.445***
	(6.857)	(9.755)	(10.012)	(14.000)	(9.719)	(13.869)
Both parents have a uni education × Mother free to help		22.855*		36.840*		8.927
		(13.601)		(19.389)		(19.326)
Mean Dep. Var	161.061	161.061	157.157	157.157	164.805	164.805
Participants	1056	1056	517	517	539	539
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: p <0.10; ** p <0.05, *** p <0.01. OLS results for the expectations on time spent on test help (panel A) in minutes per week, and time spent on extracurricular activities (panel B) in minutes per week, with the hypothetical child. Italicized words in parenthesis correspond to the reference category. All specifications include controls for the pre-registered set of participants' characteristics. Robust standard errors in parentheses.

[20] A key stage refers to a level within the education systems of England, whereby a certain level of educational knowledge is expected from students. Key Stage 2 concerns pupils aged 7 to 11 who take SATs, and is particularly well known by English parents. See the UK Government website for more information.

[21] Both are answered by moving a slider in 10-minutes increments that can range from 0 to 10 hours.

[22] Currently, this is the last edition of this survey, and the sample is restricted to parents who have at least one child in the 10 to 14 age range. We further inform participants of this.

Expectations on time preferences: results.

We regress the expectations for time spent studying and on extracurricular activities on an indicator for seeing the scenario with the mother free instead of the father, and an indicator for seeing the scenario with university educated parents, while controlling for participants' characteristics. These are reported in column (1) of Table 2.5 and in column (2) we add an interaction between these two randomized features. Respondents expect mothers to dedicate more time than fathers, particularly for study help (about 13 minutes more) and in scenarios where the presented parents are university educated. While in column (2), results are not significant, we find that the marginal effect of a mother having free time in the university educated scenarios is significant and about 20 minutes longer in study time than fathers $(6.3 + 13.6)$.[23] Finally, looking beyond our pre-registered plans in columns (3) to (6), these results appear stronger for those with more negative views about women working longer hours as captured by our θ_i measure. We view this last step as exploratory, as it is beyond our pre-defined plans, but suggestive that those holding strong views on absolute advantage do have in mind different time preferences between mothers and fathers. Altogether, these results are consistent with beliefs on time preferences where mothers are expected to spend more time on educational activities than fathers.[24]

Expectations on productivity: design.

Now, we aim to explore beliefs on differences in the productivity of a given time investment over mothers relative to fathers. After answering the expected time questions, participants move to the next survey page where we continue the setup of the previous question. Now, however, we fix the time the parent who is free spends helping the child prepare for the test. For instance, if a participant was randomized to see that the "mother" was free on the previous question, this continues here and we pin down the time spent. We also randomize this between 30 minutes (shorter time) or 1 hour 30 minutes (longer time). Participants are asked how well they think the child will do compared to other students in terms of percentile rank on the Key Stage 2 test. To answer, they drag a slider ranging from the 1^{st} to the 99^{th} percentile.

Expectations on productivity: results.

We regress participants' expected percentile rank for the child at the Key Stage 2 national test on the three randomized features and participants' characteristics. The features include an indicator for the mother being free instead of the father, an indicator for the scenario where both parents have a university education, and an indi-

[23] We have not reported these calculations but can make them available on request.

[24] We abstract from discussing whether individual perceptions of differential time use across mothers and fathers are accurate. However, as a reference, our descriptive analysis from the 2014-2015 UK Time Use Survey, shows that, when comparing full-time working mothers and fathers—controlling for household income, education, number of children, and whether the diary was completed on a weekday or weekend—mothers with at least one child aged 5 to 10 spend, on average, about 30 more minutes per week on reading and teaching to the child(ren). This difference, however, is not statistically significant.

Table 2.6: Expectations on Performance

	(1)	(2)	(3)	(4)	(5)	(6)
	All Participants		By θ			
			< 0		≥ 0	
Panel C: Expected rank at test						
Mother (*father*) free to help	0.002	-0.017	-0.009	-0.032	0.012	-0.003
	(0.012)	(0.020)	(0.017)	(0.029)	(0.016)	(0.029)
1h30 (*30 minutes*) of help	0.101***	0.084***	0.090***	0.066***	0.113***	0.101***
	(0.012)	(0.017)	(0.017)	(0.024)	(0.016)	(0.023)
Both parents (*neither*) have a university education	0.072***	0.071***	0.053***	0.055**	0.089***	0.087***
	(0.012)	(0.017)	(0.017)	(0.024)	(0.016)	(0.023)
Mother free to help $\times$ 1h30 of help		0.035		0.047		0.024
		(0.024)		(0.034)		(0.033)
Mother free to help $\times$ Both parents have a uni education		0.003		-0.003		0.004
		(0.024)		(0.035)		(0.033)
Mean Dep. Var	42nd	42nd	42nd	42nd	42nd	42nd
Participants	1056	1056	517	517	539	539
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: p <0.10; ** p <0.05, *** p <0.01. OLS results for the expectations on the child's test performance (rank among his peers). Italicized words in parenthesis correspond to the reference category. All specifications include controls for the pre-registered set of participants' characteristics. Robust standard errors in parentheses.

cator for seeing the time spent on studying help as 1.5 hours (longer time) instead of 0.5 hour. These are reported in column (1) of Table 2.6, and in column (2), we add an interaction between the mother being free and two additional randomized features.[25] Remember that the amount of the time investment is pinned down in the scenario here, thus the comparison is between a mother being free versus a father for a given time investment. We do not see strong evidence for disparate beliefs on the productivity of time investments when mothers are free relative fathers. The point estimates in columns (2) and (4) do suggest that when the time spent is 1.5 hours participants expect small gains when it is the mother helping, but these estimates are insignificant. Overall, the evidence does not support expectations around differences across genders in the productivity of time spent.

Expectations on resource allocation.

Another possibility is that people expect more resources to be allocated to a child's educational activities when the mother earns a larger share of the household budget. If so, we think this would work in the opposite direction of beliefs about absolute advantage potentially offsetting them when a mother works longer hours. Yet, we find no evidence for this in column (1) of Table 2.7. We randomize participants to see a mother (father) earning a larger share of the family budget and ask them for the expected share of the family budget spent on the child's educational and extracurricular activities.[26] Regressing this expectation on an indicator for those who

[25]For transparency, we did not mention these interactions in the pre-registration plan directly, although we had noted we would analyze results by design features this was directly about the previous section. Thus, we put these forward with caution along with the splits by the θ_i measures.

[26]Participants read a scenario again with a child aged 11 that reports the father (mother) earning a net monthly income of £1,500 and the mother (father) earning £2,500, randomizing which parent

see the mother earns more, and controlling for respondents' characteristics, returns a tight null. Additionally, we show in the Appendix, Table B.10, that the results are also null when we further split by negative and positive values of θ_i.[27] Thus, differences in resource allocation do not appear to drive beliefs.

Table 2.7: Expectations on Resource Allocation and Parental Education

	(1) Resource Allocation	(2) $\mathbb{P}$(University Graduate)		
		Mother	Father	Difference
Mother (*father*) earns more	0.007			
	(0.009)			
Works full-time (*part-time*)		0.114***	0.082***	0.032
		(0.013)	(0.013)	(0.023)
Participants	1056	1056	1056	1056
Individual Controls	Yes	Yes	Yes	Yes

Notes: p <0.10; ** p <0.05, *** p <0.01. OLS results for the expectations on (1) resource allocation, and (2) parental education. Italicized words in parenthesis correspond to the reference category. All specifications include controls for the pre-registered set of participants' characteristics. Robust standard errors in parentheses.

Expectations on parental education.

Finally, people may expect part-time working fathers to be much less skilled than part-time working mothers. If so, this could explain expectations on children's future outcomes when a mother relative to a father works full-time. To explore this, we present a mother (father) working 36 hours per week for £27 per hour, while the father (mother) works 20 hours per week for £17 per hour. We then ask respondents the likelihood for each parent to hold a university degree and regress these answers on an indicator for having seen that parent working full-time in the scenario, including respondents' characteristics. The results show that there is a part-time to full-time expected education gradient. Respondents expect that a university degree is more likely for either the mother or father when they work full-time relative to part-time (columns 2 and 3 of Table 2.7). However, the difference in this expectation across mothers and fathers is negligible and not significant. Results further split by negative and positive values of θ_i^0 (Appendix, Table B.10) also remain negligible.[28] Thus, we see no evidence that differences in skill expectations drive beliefs when comparing mothers working longer hours relative to fathers.

earns more. Participants are then asked what share of income they expect to be spent on the child's educational and extracurricular activities.

[27]This analysis was not pre-registered. See Section B.5 in the Appendix for more details.

[28]Transparency: we pre-registered this design but only realized after the survey collection the best way to use the information to address the question at hand.

Summary.

Our evidence points to beliefs on time preferences. It is suggestive that people expect mothers to invest more time in developing a child's skills than fathers with similar free time. This would be consistent with a version of absolute advantage, where different beliefs originate from how parents form preferences in allocating their free time. In this case, mothers are believed to hold an absolute advantage because people expect them to allocate more of their time at home to productive inputs for children's human capital.

Result 3 *Respondents expect mothers to spend more time on investments to a child's skills relative to a father with the same free time, especially when a parent has higher education.*

Together our evidence in Section 2.3 can be summarized through three results. First, we find strong evidence of beliefs on absolute advantage (Result 1), suggesting that even with earnings potentials equalized gender gaps may remain. Second, role model effects during childhood are important for shaping later beliefs, and beliefs then impact women's labor supply and motherhood penalties (Result 2). Third, a mental model of differences in preferences across mothers and fathers for time investments into children appears an important element of comparisons informing participants' beliefs (Result 3).

2.4 Information Experiment: Design and Results

So far, our evidence is consistent with beliefs that mothers have an absolute advantage in child-rearing. However, our evidence that beliefs about mothers are intergenerationally rooted and link with labor supply choices, could mean these beliefs are sticky, unresponsive to information. Thus, it is not clear whether individuals update them in response to new information that removes uncertainty on how well children do when mothers work. This has important implications for policy design if providing information can ultimately influence labor supply decisions by changing people's beliefs.

The overarching goal of this section is to test whether individuals revise their beliefs in response to new information about children's outcomes when mothers work longer hours. We focus on how people react to a statistical fact, both in terms of their beliefs about children's prospects when mothers maintain full-time careers, and their perceptions of the role of mothers. We further examine how information affects support for policies that promote mothers' labor market opportunities. Our sample of participants remains the same as defined in Section 2.3.1, and throughout this section, we assess whether responses to information vary by the individual-level belief measure (θ_i) to test whether information responses are symmetric or asymmetric over prior beliefs. Our design uses incentivized beliefs in a similar structure as Haaland and Roth (2023) who study beliefs about racial discrimination and Haaland and Roth (2020) who study beliefs about immigration and the labor market.

2.4.1 Treatment Design and Outcomes

To form an information treatment, we use a nationally representative survey in the UK that follows children and their parents.[29] We calculate the share of children passing five or more of their secondary school GCSEs with a grade C/4 or higher. This pass rate is a common metric in school league tables in England, which will likely be familiar to our sample of parents.[30] We split this metric by families where the mother worked, on average at least full-time hours per week when the child was aged 5 and 7 *versus* families where the mother worked part-time or less. This gives us a comparison based on the children's primary school years. We draw our calculations from dual-parent homes in England and compare families whose parents have similar income and education levels. We provide more details on the data and our calculations in the Appendix, Subsection B.4.1.2.

Prior beliefs on GCSE pass rates.

We first inform participants of this GCSE pass rate for families where the mothers worked part-time or not at all. We then collect an incentivized belief about this pass rate for families where the mother works at least full-time hours. Below is the text participants read.

> We, as researchers at the University of Strathclyde, have calculated the share of children passing five or more GCSEs with a grade of C/4 or higher.
>
> Among families where the mother worked part-time or not at all, around **73**% of children passed five or more GCSEs with a C/4 or higher. This information is also shown visually in the graph below.
>
> We then computed this statistic for families with similar income and education levels but where the mother worked full-time (35 hours or more). In these families, what percentage of children do you believe eventually passed five or more GCSEs with a C/4 or higher?
>
> You will gain £1.50 if your answer is within 2 percentage points of the true number.

We also present this information and question visually in a graph (see the Appendix, Figure B.21). Participants then respond by dragging a slider between 0 and 100% in increments of one percentage point. Let us note that participants are informed that this comparison is made on mothers who work full-time or not when the children are of primary school age (5 and 7). We capture priors on this GCSE pass rate before giving half the sample the true information.

While the 'GCSE prior' is conceptually distinct from our earlier vignette-based belief measure (θ_i), where θ_i holds *all else equal* by design, this incentivized belief

[29] For this, we use the the Millennium Cohort Study (MCS), a nationally representative sample of children born in the year 2000.

[30] See the UK Government website for further information about GCSEs results in 2023. Also see the GCSEs' subject content, by field.

still captures perceived costs or benefits of maternal full-time work. Its main advantage is that it is incentivized, which has been shown to yield more truthful survey answers in particular for politically sensitive questions (Bullock et al., 2015; Prior, Sood, Khanna, et al., 2015). As such, this belief provides a valuable complement to our earlier measure. Notably, we once again find strong support for beliefs of women's absolute advantage. As can be seen in Panel (a) in Figure B.4, displaying the distribution of GCSE priors, we find that over 70% of participants expect children of full-time working mothers to perform worse compared to their part-time counterparts.

Information treatment.

In reality, our data reveals that among families with similar education and incomes, children with full-time working mothers do just as well, if not slightly better, than those with part-time or less working mothers. We emphasize to the reader (not participant) that this does not imply mothers' time with children is not important. In fact, the substitution model we previously discussed would explain the descriptive pattern, suggesting that when mothers work full-time sufficient inputs are substituted to keep children's skills even. We make no implication to participants and only use the statistical fact as a treatment. However, the underestimation of children's performances when mothers work longer hours we observe points to a potential role for information provision to correct these misperceptions. We thus randomize participants to either receive the actual pass rate when mothers work longer hours — the fact — or to the control group with no information. Those assigned to the treatment are shown the statistic in text and graphically. The text is below and the visual aid is reported in the Appendix, Figure B.22.

> For mother's working full-time (35 hours or more), adjusted to have similar education and income levels as mothers working fewer than 35 hours, we found that around **75**% of their children eventually passed five or more GCSEs with a C/4 or higher.
>
> This means these children did about **2 percentage points better** compared to those of mothers working less than 35 hours per week.

Outcomes.

We follow the information treatment by collecting a set of outcomes. Details about the wording and descriptive statistics are provided in the Appendix, Subsection B.4.2. Our primary interest is on whether people respond to information updating how they think about the effects of a mother working full-time. We look at these views in three ways.

First, we collect another incentivized belief, focusing this time on externalizing behavioral problems when mothers work part-time or less compared to full-time or more across families with similar education and income.[31] Relative to our question on GCSE pass rates, we use a different response scale and change its direction to

[31]The share of children at high risk of behavioral problems when the mother works full-time is 16.573%.

mitigate concerns over numerical anchoring (per suggestions in Haaland, Roth, and Wohlfart, 2023). In the Appendix, Figure B.4, we show both the collected GCSE expectations and the behavioral expectations both have a high degree of variation for identifying effects. Below is the text shown to participants to elicit this behavioral belief. Participants are asked to report their expectation in a text box.

> The data that we used to calculate the share of children passing five or more GCSEs also provides information on the children's externalizing behavioral problems at age 7 (*e.g.*, conduct problems and hyperactivity/inattention).
>
> Among families where the mother worked part-time or not at all, **out of 100 children** aged 7, we found that around **17** had an abnormal level of behavioral problems.
>
> We then computed this statistic for families with similar income and education levels but where the mother worked full-time (35 hours or more). In this group, **out of 100 children**, how many do you believe had an abnormal level of behavioral problems?
>
> You will gain £1.50 if your answer is within 2 points of the true number.

Second, we ask them to write in complete sentences what guided their answer to this question (see Appendix, Figure B.25). We use this to gain insight on what participants are really thinking about when responding to the quantitative question on abnormal problem behaviors. We coded their responses as suggesting it is harmful for the child when the mother works full-time, not harmful, or an unclear answer (13% classed unclear).[32] As discussed in the Appendix, Section B.5, our classification scheme departs from our pre-registered classification plan as we realized afterwards the best way to code these responses to test our objectives.

Third, we ask five questions related to gender norms on the role of mothers in the family. Participants are asked to indicate their level of agreement on a 1–5 scale with each statement. The questions participants see are presented in the Appendix, Table B.26, and are drawn from the British Household Panel Survey, as these are commonly used in the literature (*e.g.*, Flèche, Lepinteur, and Powdthavee 2020). We sum answers to these into a scale where higher values reflect more liberal views. In the Appendix, Section B.3.1 we demonstrate that both the GCSE and our θ_i measure predict expectations on behavior, as well as self-reported norms.

Finally, we collect government policy views about subsidized childcare and paternity leave policies, whereby participants indicate their level of agreement. Details are in the Appendix, Figure B.24. We then code these into binary outcomes for high support split by the median.

[32] We read and manually coded each response. We experimented with textual analysis on a training set, but participants use a wide range of language in their responses leading standard machine learning tools to classify poorly (accuracy of 60%). Participants, however, provided often rich answers, and we could clearly classify nearly all responses outside of a small percentage (13%).

Obfuscated follow-up.

We invited participants back one week later and continued our look at self-reported policy views. Participants received a generic invitation from Prolific to take a five-minute survey, which did not reveal the connection to the main survey.[33] Among the 1056 participants of the first survey, 86% (893 respondents) took part in the obfuscated follow-up. Further, we asked four questions but only two of these relate to our research questions and are about policies to lower the cost for mothers to work, *e.g.*, on childcare policies. The additional questions serve to obscure a link between this survey and the original. These questions, as well as answer modalities, are presented in the Appendix, Subsection B.3.2.1.

2.4.2 Information Treatment Effects

We now assess whether information on children's performance when mothers work full-time is relevant to beliefs about the impact of mothers working full-time. We study information treatment effects on the following: (i) an incentivized quantitative scale about children's abnormal level of behavioral problems; (ii) responses to an open-ended question about what guided their answer to the quantitative scale; and (iii) self-reported gender norms. We finally turn to an extension on policy support, which includes a one-week follow-up.

2.4.2.1 Belief Updating

Information effects: approach.

We look at three outcomes. On each of these, we estimate a treatment effect for the exposure to information about the GCSE pass rates when mothers work full-time hours given by the following:

$$y_i = \beta_0 + \gamma D_i + \sum_{j=1}^{J} \beta_j X_{ij} + \epsilon_i. \tag{2.4}$$

The outcome y_i is first the behavior beliefs for each individual i rescaled to lie between 0 and 1; second, the open ended harmful/not-harmful categorization; and third, the gender norms score with higher values representing more liberal norms. Exposure to the information treatment is captured by $D_i = 1$ and is otherwise equal to 0. X_{ij} is a vector of individual and pre-determined demographic variables.

We also test whether this information treatment effect is heterogeneous. To do this, we first disaggregate the information effect across under- and over-estimators of the GCSE initial beliefs, estimating equation (2.4) separately by these dimensions. Second, we repeat this exercise but use the hypothetical scenario-based individual beliefs (θ_i). Here, we split the sample by an indicator for whether a person has strictly negative perceptions relative to null or positive perceptions about mothers working longer hours. Later, we also look at heterogeneity around participants' characteristics and past experiences.

[33] This survey was opened on 25 July 2024 a week after the main survey. We kept it open until 27 July 2024.

Information effects: results.

In Table 2.8, we find that information leads to more positive perceptions on the impact of mothers working full-time.

In Panel A, the outcome is beliefs about behavioral problems when mothers work full-time. The average response to information is a reduction (or improvement) in beliefs by about 5.2 pp. In column group (2), we show that this information treatment effect is driven by participants who under-estimated GCSE pass rates. In column group (3) we look at results split by participants' positive *versus* negative θ_i. The information treatment is generally homogeneous across these groups.[34] Taken together, those who were wrong and below the true GCSE pass rate on average updated their beliefs regardless of their perceptions of absolute advantage captured in the θ_i measure. Moreover, in the Appendix, we report a homogeneous pattern of treatment effects across participants' characteristics (Table B.13), and past experiences (Table B.14). This is also confirmed by a mostly homogenous pattern in a causal forest (see the Appendix, Table B.15 and Athey and Imbens, 2016; Athey and Wager, 2019), further suggesting that information can move beliefs.[35]

In Panel B, we leverage the open-ended responses on what guided their answers to the quantitative question on behavioral problems. Our purpose here is to study whether participants are really thinking about and shifting views on what happens to children when mothers work full-time. We drop participants whose answers were unclear (141 unclear participants) and code responses as either expecting worse outcomes for children when mothers work full-time (harmful) or expectations children will do just as well if not better (not-harmful category).[36] Information leads to a 22% decrease in the likelihood that participants' written answer expresses views that children will suffer (column 1). What participants write about strongly suggests their responses on the quantitative problem behavior scale truly capturing their views on what happens to children when mothers work full-time. Additionally, when we split by prior belief measures, we see the same pattern of information effects in column groups (2) to (4) as we saw on the quantitative scale.

We further categorize the written responses based on whether they specifically expect lower time investments when a mother works full-time. Our prior evidence indicates that the mental model underlying worse expectations for children when mothers work relates to expectations on time investments into the child. Here, we find that nearly all of the respondents (96%) whose written answers fall into the "harmful" category relate to expectations on worse time investments for the child. Meanwhile, respondents who did not expect worse outcomes tended to highlight that when a mother works full-time the child might have more opportunities to be positively exposed to a wider range of experience or for there be a no net impact.

[34]Recall that the θ_i measures pinned down earnings potential between women and men and capture a particular portion of beliefs. Expectations on the GCSE pass rate may share variance with these measures but include other dimensions or simply uncertainty. When asking about the GCSE pass rate, we compare scenarios where mothers work longer versus shorter hours, controlling for similar education and income levels at household level. However, differences in earning potential among mothers with different working hours are not explicitly controlled for.

[35]Causal forests represent a machine learning approach to estimating heterogeneity. See Athey and Wager (2019) for an introduction and an application.

[36]Later we will drop these from our analysis on beliefs as a robustness check (see Table B.11).

Table 2.8: Belief Updating and Information Effects

	(1) All Participants	(2) By GCSE Beliefs		(3) By θ	
		Under-	Over-	< 0	≥ 0
Panel A: Incentivized beliefs					
Treatment	-0.052***	-0.066***	-0.015	-0.064***	-0.039***
	(0.007)	(0.009)	(0.010)	(0.011)	(0.010)
Difference: *p*-value		0.000		0.078	
Mean Dep. Var	0.213	0.227	0.169	0.218	0.207
Panel B: Open Q: harmful/not harmful					
Treatment	-0.219***	-0.275***	-0.087	-0.223***	-0.211***
	(0.031)	(0.035)	(0.064)	(0.044)	(0.045)
Difference: *p*-value		0.009		0.856	
Mean Dep. Var	0.611	0.667	0.455	0.649	0.572
Panel C: Gender norms					
Treatment	0.546***	0.873***	-0.219	0.404	0.675**
	(0.206)	(0.236)	(0.398)	(0.286)	(0.299)
Difference: *p*-value		0.017		0.510	
Mean Dep. Var	17.462	16.999	18.818	17.277	17.640
Participants	1056	787	269	517	539
Individual Controls	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. This table presents OLS results of equation (2.4) each outcome listed in the panels. All specifications include controls for the pre-registered set of participants' characteristics. Results are presented (1) for the full sample of participants, (2) by under- and over-estimators of the GCSE initial beliefs, and (3) by negative and positive values of θ.

We find the written responses consistent with expectations that mothers time with children matters but there are also substitutable experiences that can be beneficial. Altogether, we find that respondents think about the information and shift their views in response toward less negative perceptions for children and away from a model of reduced time investments when mothers work full-time.

To further corroborate that information leads to belief updating on views about mothers, we finally turn to a self-reported scale of gender norms about the role of mothers. In Table 2.8 and Panel C, we again find a similar pattern for the estimated information effects. Treated respondents express significantly more liberal views (column 1) and this is driven by GCSE under-estimators (column group 2). We again see evidence that the information response here is homogeneous across the θ_i measure, suggesting a degree of malleability in these beliefs. The magnitude of the effect is small relative to the mean, but this is reassuring that our information effect is yielding some thoughtful reflection and a small shift in views.

Information effects: robustness.

We test whether our results on belief updating are robust to participants' uncertainty, their attention, risk of experimenter demand effects, and lack of clarity in their written answer to the open-ended question. In particular, experimenter demand effects

would be a problem if respondents receiving the information try to give answers they think we want. While De Quidt, Haushofer, and Roth (2018) suggest that demand effects are minimal in practice, we try to rule these out by asking participants at the end of the survey to tell us what they think the survey was about in a text box open-ended entry. We classify those who appear to understand our focus on perceptions about mothers at work and children as at risk of a demand effect and drop them from the analysis.[37] Across all of these checks, reported in columns (1) to (5) in the Appendix Table B.11, we find our results remain robust. We next re-weight our estimates based on the population weights discussed in Section 2.3.1, showing in column (6) that our results remain unchanged. Finally, we apply a post-double selection Lasso (Belloni, Chernozhukov, and Hansen, 2014) to the selection of pre-determined controls based on all possible variables we could use, showing in column (7) that again our conclusions hold.

Information effects and the degree of learning.

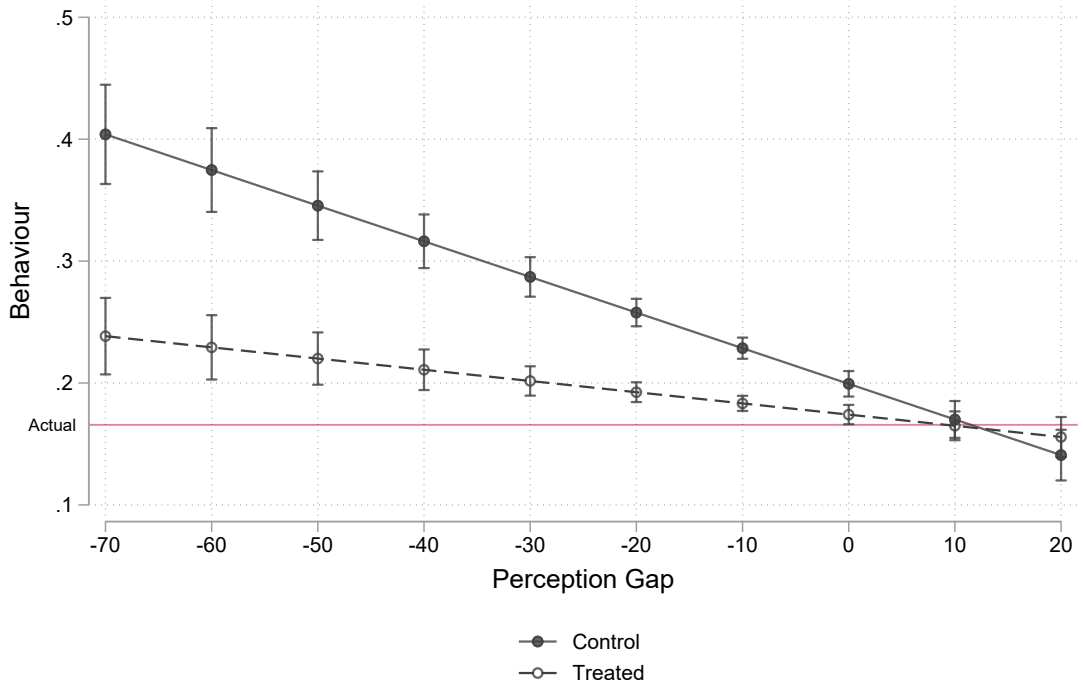
GCSE under-estimators drive responses to information, but do they become more accurate? We now interact the information treatment with the perception gap (PG_i). This is the difference between the prior GCSE belief and the actual GCSE pass rate when mothers work full-time, *i.e.*, 75%. Participants with a positive perception gap are over-estimators on GCSE pass rates, while those with a negative gap are under-estimators. We use this, as in Haaland and Roth (2023), to assess the degree of learning in response to the information treatment. We look at the quantitative belief on behavioral problems and estimate the following specification:

$$y_i = \beta_0 + \beta_1 D_i + \beta_2 D_i \times PG_i + \beta_3 PG_i + \sum_{j=1}^{J} \beta_j X_{ij} + \varepsilon_i. \tag{2.5}$$

The regression results are reported in the Appendix, Table B.12, and in Figure 2.3, we report the visual representation of the results based on the full sample of participants (column 1 of Table B.12). Belief updating is stronger when a respondent's perception gap is more negative. Treatment moves respondents toward the correct answer, not away from it, with treated respondents becoming more accurate than the control group. We further split the results by beliefs on absolute advantage in Table B.12, and strikingly, we find that those who held stronger views on absolute advantage, as measured by negative θ_i measure, drive this accuracy update in response to information. Our evidence here is again reassuring that participants respond thoughtfully to the information.

[37] We use a Random Forest Classifier to predict the risk of a demand effect, preprocessing text data by converting to lowercase, removing special characters, stop words, and HTML tags, and applying lemmatization. Text embeddings are generated using the pre-trained SentenceTransformer model (all-MiniLM-L6-v2) to capture semantic nuances. To address class imbalance, we used a pipeline with SMOTE and optimized hyper-parameters through GridSearchCV with StratifiedKFold cross-validation. We first trained the model on 30 classifications and then used it to predict labels for the remaining data. The model achieved an overall label prediction accuracy of 93.33%, highlighting the robustness of the classifier and preprocessing steps.

Figure 2.3: Belief Updating and the Perception Gap



Notes: This figure displays means of the behavior belief across the distribution of the perception gap and over treatment status. Estimates are based on equation (2.5), for the full sample of participants.

Summary.

We find clear evidence of belief updating in response to information. Our information treatment demonstrates that when mothers work full-time during the primary school years, children do at least just as well on later important GCSE exams compared to when mothers work part-time. While participants tend to underestimate how well children actually do when mothers work full time, receiving the correct information leads to more positive perceptions of mothers working full-time regardless of their prior views on absolute advantage. Thus, well tailored and delivered information can be a useful tool to support accurate beliefs among parents on the impact of mothers working.

Result 4 *Information leads to belief updating on the effect of mothers working full-time reducing expectations on children's abnormal behavioral problems, leading to more positive views and a shift toward more liberal gender norms on the role of mothers.*

2.4.2.2 Policy Support

Our findings point to the existence of beliefs about women's absolute advantage in child-rearing, and show that these beliefs are predictive of women's labor supply decisions. We also find that showing people how well children perform when mothers work full-time reduces negative perceptions of the consequences of women working longer and leads to more liberal gender norms.

Given how closely these beliefs are tied to labor supply decisions, shifting them can increase awareness of the importance — or absence — of public family policies. We then examine how exposure to our intervention increases support for policies that help mothers return to work after childbirth and balance career and family life.

In the main survey, we asked participants how strongly they agreed with support for policy to increase subsidized childcare and paternity leave policies. One-week later, in an obfuscated follow-up, we collected more views on policy support.[38] The results of the follow-up survey are reported in the Appendix, Section B.3.2.1. Based on our pre-registration, these collected support measures on a 1 to 5 scale are split by the median to binary high/low support measures.

Policy support: main survey.

Our focus on a small fact as the information treatment, means that it is hard to expect large changes in policy views, as variation is likely more limited on agree or disagree type questions. Consistent with this, in the main survey, we do not see significant information effects on policy support, although, we do see some heterogeneous effects. The results are reported in Table 2.9. On support for childcare policies we see flat nulls, although nearly 80% of participants agreed with this question. On support for paternity leave policies there is more variation to leverage (about 64% agree), and here we see suggestive evidence of a positive information effect. This positive effect on support is stronger and significant for GCSE under-estimators.[39] Turning to heterogeneity by participants' characteristics (Appendix Table B.16) and past experiences (Appendix Table B.17) we are too under-powered to make strong claims, but again, we see suggestive evidence of positive effects from information on support for paternity leave. This suggestion is particularly strong, though not always significant, among females, those born outside the UK, university degree holders, part-time workers, and those with lower income. Interestingly, the results of the follow-up study are broadly consistent in terms of groups with the results on policy support in the main survey. We find evidence that the information significantly increases support for additional free childcare hours for women, those born outside the UK, and university degree holders, and suggestive evidence for part time workers. The consistency of results across groups between the main and the follow-up surveys suggests that policy may play an important role in shaping views through information provision.

[38]These are about a proposal in the UK government to expand free childcare to 30 free hours per week for parents earning less than £60,000 per year and a proposal to create new nurseries in high-need areas by converting space in existing primary schools. We also asked participants how many hours of free childcare they would support to provide a more continuous scale (also pre-registered).

[39]These results are more efficient and significant for both GCSE under-estimators and those with $\theta_i < 0$, if we use a Lasso to select the optimal control set from all possible controls we have. Results not shown but available upon request.

Table 2.9: Information Effects on Policy Support

	All controls					LASSO selected controls				
	(1)	(2)		(3)		(4)	(5)		(6)	
	All	By GCSE Beliefs		By θ		All	GCSE		By θ	
		Under-	Over-	< 0	≥ 0		Under-	Over-	< 0	≥ 0
Panel A: Subsidized childcare policies										
Treatment	-0.005	-0.012	0.014	0.006	-0.020	0.045	-0.010	0.018	0.009	-0.016
	(0.024)	(0.029)	(0.046)	(0.035)	(0.034)	(0.029)	(0.029)	(0.046)	(0.035)	(0.034)
Difference: *p*-value		0.633		0.592						
Mean Dep. Var	0.804	0.795	0.829	0.807	0.801	0.804	0.795	0.829	0.807	0.801
Panel B: Paternity leave policies										
Treatment	0.040	0.066*	-0.022	0.064	0.016	0.045	0.075**	-0.020	0.077*	0.023
	(0.029)	(0.034)	(0.056)	(0.042)	(0.041)	(0.029)	(0.034)	(0.057)	(0.042)	(0.041)
Difference: *p*-value		0.170		0.402						
Mean Dep. Var	0.638	0.624	0.680	0.642	0.635	0.638	0.624	0.680	0.642	0.635
Participants	1056	787	269	517	539	1056	787	269	517	539
Individual Controls	Yes	Yes	Yes	Yes	Yes	LASSO	LASSO	LASSO	LASSO	LASSO

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. This table presents OLS results of equation (2.4) for the outcome in each panel. These are a binary variable set to one if the participant indicated a strong level of agreement (above the median) with subsidized childcare policies (panel A), and another with paternity leave policies (panel B). All specifications include controls for the pre-registered set of participants' characteristics.

2.5 Conclusion

In this paper, we study beliefs about children's skills when mothers work. Beliefs that mothers have an absolute advantage in child-rearing relative to men imply that gender gaps in labor markets will remain even as earnings potential is equalized across gender. We describe a target belief distribution as one that pins down earnings potential between a mother and father and captures differences in expectations for a child's future human capital between this mother or father working longer hours in the labor market. With a survey designed around vignettes of a family with a mother or father working longer hours, we elicit these beliefs on absolute advantage and show they are present and vary substantially over respondents.

Beliefs about women's absolute advantage are particularly predicted by those whose own-mother did not work full-time while they were growing up and among women who themselves had a strong employment motherhood penalty. Our evidence demonstrates that beliefs on absolute advantage are shaped by role model effects during formative childhood years, are highly correlated with post-birth labor market trajectories and are predictive of higher labor supply for women - consistent with gender norms and beliefs having an important effect on gender gaps.

Individuals' beliefs about the impact of mothers working likely stem from a mental model of mothers relative to fathers time use and productivity with children. To investigate this, we introduced new vignettes but randomized participants across features that let us test whether respondents have in mind expectations on differences in preferences between mothers and fathers for time investments into children's skills, differences in the productivity of investments, or differences in resource allocation or parental skill when mothers relative to fathers work longer hours.

The evidence points toward differences in preferences where participants expect

that with equivalent time mothers will spend more time on skill investments with children than will fathers. Our analysis of a qualitative, open-ended question further demonstrates that when mothers work full-time participants largely tend to expect lower time investments into children. An important implication, is that mothers who are deciding whether to maintain a career may face pressure from expectations that their children will suffer from lower time investments. This is likely the most salient for families without the resources to pay for costly high quality childcare as a substitute, consistent with our evidence that beliefs on absolute advantage dissipate when we show respondents vignette scenarios with overall higher household resources.

Finally, we investigate whether factual information about children's outcomes when mothers work full-time will lead people to reduce expectations of harm for children. We first find that people tend to underestimate how well children perform when mothers work full time. Our information treatment effects show that those participants who initially underestimated outcomes do respond to this information. Treated respondents move toward more accurate expectations and in their qualitative responses demonstrate a lower expectation of harm for children. Thus, there is a role for policy to target misinformation about how well children do when mothers work.

Brought together this paper forms a broad profile of evidence on beliefs about children when mothers work. It gives empirical weight to the idea that beliefs influence gender gaps in labor markets through a new approach to elicit these beliefs. Thus, policy solely focused on equalizing earnings potential may remain ineffective for closing gender gaps. Moreover, we are able to show how these beliefs vary across the population and the mental models that can inform them. These mental models then create targets for policy effort to reduce uncertainty and misinformation. Our evidence on responses to the information treatment is then encouraging and suggests room for belief updating.

Chapter 3

Universal No More? Poverty and Child Care Consequences of Means-Testing

3.1 Introduction

How do families adjust when family benefits are cut? Financial support to families with children is a central pillar of social protection systems aimed at reducing child poverty and creating a supporting environment for early childhood development. France has long prioritized such investments, allocating nearly 3% of its GDP in 2014 to family benefits — among the highest in the OECD (OECD, 2024). Yet, one in four children lives in financially constrained households in France, a 10-percentage-point increase over the past decade (UNICEF, 2023). Public concern over the sustainability of the French social welfare system (*Sécurité Sociale*) and over the targeting of such spending has led to successive reforms from 2013 onwards, including the 2015 reform of family allowances (*allocations familiales*).

On July 1, 2015, France implemented a major reform of its family allowance system, transitioning from a universal to a means-tested model. While households with two or more children previously received benefits irrespective of income, the reform introduced two income thresholds: families between the first and second thresholds saw their benefits halved (half-benefits group) while those above the second received only a quarter (quarter-benefits group). Because eligibility was determined by household income from two years prior, the reform created quasi-experimental variation ideal for exploiting the causal effects of benefits cuts on households living with children. More precisely, we use a French nationally representative panel data from the EU-SILC — *Statistiques sur les Ressources et Conditions de Vie* (SRCV) — for the years 2010 to 2019 inclusive and focus on households with two children or more — thus eligible for family allowances.[1] This household panel is particularly appealing because, in addition to providing rich information on all individuals within interviewed households, administrative data on taxes, income, and benefits are merged to households. This enables us to focus on households, individuals (parents), and on their children to quantify the impact of the 2015 reform on three main dimensions, further highlighting the different trade-offs parents may face in reaction to such benefits

[1]Households with only one child are not entitled to family allowances in France. This holds true before and after the reform.

cuts.

First, we quantify the impact of the reform on household financial resources and deprivation. In particular, we consider on the one hand rich household financial measures provided in the data — including objective measures of savings, poverty, as well as self-reported hardship. Although the reform did not affect households' savings, we find an increase in the likelihood of being classified as poor by 3.3 percentage points (pp) relative to their baseline mean for the most affected families, who received a quarter of the initial amount of benefits. The largest effects are found for the probability of reporting financial hardship, with a 24% and 40% increase for the half- and quarter-benefits groups, respectively. These highlight the importance of possible financial difficulties these middle income parents may have faced and induced by the decrease in household financial resources. In addition to these, we consider measures of deprivation, such as a binary variable for being deprived or the number of essential items a household is unable to afford, and further investigate which items were most affected by the reform. We find that cutting benefits for families, either by half or by three-quarter, had some significant and sizeable effects with respect to their baseline means. The half-benefits and quarter-benefits groups see their likelihood of being deprived increased by $3pp$ and $4.2pp$ respectively. Similarly, the number of items they are unable to afford increased by 0.21 and 0.22, respectively, with a significant increase in housing deprivation for the quarter-benefits group, and more pronounced increases in food and clothing deprivations for the half-benefits group.

Second, we investigate the effects of the reform on parental labor supply, including their likelihood of being employed and/or employed full-time, their weekly hours worked, as well as their labor force participation. We find no effect of the reform on any margins of parental labor supply. Further, we find no difference between self-employed prior to the reform — whose labor supply is typically more elastic — and employed parents. The reform thus did not seem to have affected parental labor supply, regardless of their pre-reform employment status.

Third and last, we examine whether the reform affected the distribution of non-parental childcare among children aged 0 to 5. We find evidence of a reallocation away from more expensive formal care (notably childminders) and toward more affordable non-parental care (daycare centers), particularly for children aged 0 to 2 — suggesting a price-sensitive substitution response. Among children aged 3 to 5, reductions in daycare use appear largely offset by greater uptake of free preschool, while effects for older children are negligible. At the extensive margin, households whose benefits were reduced were significantly less likely to use childminders, and slightly more likely to use daycare. These results suggest that the reform led to a reduced access to formal care, potentially reducing children's exposure to arrangements known to support early cognitive development (Berger, Panico, and Solaz, 2021).

Overall, our main results suggest important implications of benefits cuts affecting middle income families: an increase in household deprivation both objectively measured and self-reported, and a reallocation of family resources in terms of childcare arrangements away from more costly and toward more subsidized and thus cheaper options, with no effect on parental labor supply.

The relationship between financial resources and child development has been extensively studied in the literature, and a large body of work investigates how income-support policies affect children's outcomes. For instance, Dahl and Lochner (2012) find that expansions in the Earned Income Tax Credit (EITC) led to improved child achievement, and Chetty, Friedman, and Rockoff (2014) similarly highlight that income constraints in early life associate with poorer non-cognitive outcomes. In parallel, financial resources affect child development by enabling access to enriched learning environments and reducing household stress (Kalil, 2015). More recently, Page (2024) synthesized evidence from the quasi-experimental literature and concluded that income increases generally improve child outcomes, particularly in early childhood and among low-income families. All of these, in turn, have been linked with long-term effects on health, well-being, and socioeconomic stability in adulthood (Cunha and Heckman, 2007; Heckman, Pinto, and Savelyev, 2013; Hoynes, Schanzenbach, and Almond, 2016; Almond, Currie, and Duque, 2018).

While these studies underscore the role of income transfers in shaping child outcomes, they often conceptualize poverty narrowly, relying on income-based definitions. However, a growing body of work — building on the capabilities approach developed by Sen (1982) — argue for a broader understanding based on the concept of *deprivation*. Deprivation captures not only insufficient financial resources but also the inability to access socially perceived necessities due to limited opportunities or broader structural constraints. This multidimensional perspective has been reflected in empirical work across various disciplines that incorporate indicators of unmet material and social needs — for example, the inability to afford adequate clothing, maintain a comfortably heated home, or participate in social activities such as hosting friends or relatives — thus moving beyond income alone as a measure of economic hardship (*e.g.*, Townsend, 1979; Leturcq and Panico, 2019). Building on this approach, our paper adds to this literature by estimating the impact of a benefit cut on a broader set of household- and individual-level outcomes, including concrete dimensions of deprivation mentioned earlier (*e.g.*, inability to afford adequate clothing, housing, or participate in social activities), as well as children's participation in non-parental childcare.

Our paper adds to the literature in two additional ways. First, while most existing evidence documents the effects of welfare expansion — such as increases in child tax credits or universal cash transfers (*e.g.*, Milligan and Stabile, 2011; Dahl and Lochner, 2012; Hoynes, Schanzenbach, and Almond, 2016; Averett and Wang, 2022) — there is comparatively less evidence on the effects of benefit cuts, particularly for middle income households. Thus, we add to the few block of evidence focusing on child benefits reductions (*e.g.*, Currie and Gruber, 1996). To our knowledge, this paper is the first to attempt to quantify the causal effects of benefit reductions on outcomes such as material deprivation. Without assuming symmetry between the effects of benefit increases and reductions, this remains an unexplored area.

Second, as we document both the extensive- and intensive-margin responses to "out-of-pocket" childcare costs changes induced by the reform, we also contribute to the literature examining the determinants of non-parental childcare use and the role of affordability in shaping its demand. While most of the literature highlights substitution effects between informal and formal care when subsidized childcare

is introduced (Blau and Tekin, 2007; Baker, Gruber, and Milligan, 2008; Havnes and Mogstad, 2011), we document evidence of substitution between non-parental modes of care and in particular evidence of substitution for cheaper options, *i.e.*, between childminders and daycare centers. Our results therefore highlight parental responsiveness in reallocating care arrangements under reduced household resources. This is even more important in the French context, characterized by a mixed public-private provision. Indeed, in France, daycare centers (*crèches*) are more often publicly managed and subsidized by municipalities and national programs, whereas professional childminders (*assistantes maternelles*) operate predominantly as independent providers within the private sector, with some public financial support to families through tax credits (Barrère-Maurisson and Lemière, 2006; Onape, 2024).

As part of our analysis, we further examine whether parental labor supply responded to the reduction in family benefits, treating this as a potential mechanism through which household resources and childcare choices might interact. While existing literature documents labor supply responses to benefit changes among low-income or single-parent households (*e.g.*, Moffitt and Zahn, 2019; Klose, 2020; Agostinelli, Borghesan, and Sorrenti, 2021), evidence is more limited for middle- and higher-income families. Our results show no significant change in labor market outcomes following the reform, including employment status, labor force participation, weekly hours worked, or the probability of working full time. Additional heterogeneity analyses by gender, pre-reform marital and employment statuses reveal no differential effects, suggesting that labor supply did not constitute a channel of adjustment to the benefit reduction for middle- and higher-income groups.

Finally, our paper opens further discussion on the potential unintended consequences of the reform for children's early childhood development. Our findings suggest that some higher-income families affected by the benefit reduction reallocated non-parental childcare time away from more expensive providers (childminders) toward more affordable arrangements (daycare centers). While existing evidence suggests that such within-formal-care substitution has ambiguous or limited effects on cognitive outcomes of children from higher-income backgrounds (*e.g.*, Havnes and Mogstad, 2011; Berger, Panico, and Solaz, 2021), the resulting change in peer composition within daycare settings may generate spillover effects that affect not only treated children but also their peers. On the one hand, evidence from France (Goux and Maurin, 2007) and other settings (Duflo, Dupas, and Kremer, 2011; Cascio, Schanzenbach, and Whitmore, 2013; Bursztyn and Jensen, 2015; Cornelissen et al., 2018) suggests that peer effects in early education can significantly enhance outcomes for lower-income children, particularly when they are exposed to peers from higher and wealthier socioeconomic backgrounds. However, the literature also highlights that these effects are likely to be heterogenous and depend on children's relative position (or rank) within the peer group (Lavy, Silva, and Weinhardt, 2012; Carrell, Sacerdote, and West, 2013; Elsner and Isphording, 2017; Denning, Murphy, and Weinhardt, 2023; Kiessling and Norris, 2023). Indeed, exposure to higher-achieving peers can serve as role models by fostering aspiration and motivation, but may also lead to discouragement or disengagement when the gap in achievement or resources between peers is perceived as too large to overcome. Similarly, children from higher-income families may experience reduced engagement or slower progress if increased

classroom heterogeneity affects instructional pace or peer interactions (Fongoni et al., 2022). While our analysis does not directly examine these child-level outcomes, our findings point to peer spillovers as a theoretically and policy-relevant channel, which should be examined more directly in future research.

3.2 Background

France has a long-standing tradition of prioritizing social welfare policies, particularly through its generous allocation of resources to family benefits. In 2014, France allocated 3% of its GDP to family benefits, surpassing the OECD average (2%) and closely rivaling Denmark, which leads with 3.6% of its GDP dedicated to similar policies (OECD, 2024).[2] This substantial fiscal commitment reflects France's proactive stance in addressing societal challenges, particularly demographic concerns such as declining birth rates and an ageing population. In 2013, the French Government initiated discussions concerning the deficit of the social welfare system (*Sécurité Sociale*).

Guided by a desire to enhance fiscal equity and to better target financial support towards lower-income households, and particularly single-parent households, France introduced a significant reform to its family allowances (*allocations familiales*) system. Prior to the reform, these allowances were universal, and provided equally to all households with at least two children, regardless of income. However, from July 1st, 2015, this universality was altered with the introduction of a means-tested approach. The reform introduced two income thresholds based on a household's total taxable income from *two years prior*, with variations depending on the number of children in the household. By leveraging the quasi-experimental variation induced by the reform, we can estimate its impact on families and children. Specifically, the reform introduced two income thresholds defining our treatment status, such that (1) households with income at $t-2$ above the first threshold but below the second threshold receive half of the usual benefit, and (2) households with income at $t-2$ above the second threshold receive only one-quarter of the standard allowance. Consequently, this implies that households below the first threshold continued to receive the full amount of family allowances.

[2]OECD (2024), Family benefits public spending indicator, accessed December 5, 2024.

Table 3.1: Income Thresholds and Family Allowances Amount in Euros (€)

	Annual Household Income	
	(1) Half Benefits	(2) Quarter Benefits
Number of children		
$N=2$	67,140	89,491
$N=3$	72,735	95,086
For any additional child	5,595	5,595
	Family Allowances Amount (€)	
	(1) Half Benefits	(2) Quarter Benefits
Number of children		
$N=2$	64.67	32.34
$N=3$	147.53	73.76
For any additional child	82.86	41.44
Bonus for children aged 14+	32.34	16.17

Notes: Information retrieved from the French Senate website. This table presents the income thresholds introduced by the 2015 family allowances reform, as well as the amount of family allowances per month, by treatment status. The annual household income from two years prior to the reform (2013) and the number of children at the time of the reform determine the eligibility to family allowances post-reform, to a full rate or to a reduced rate, *i.e.*, to (1) half benefits, or (2) quarter benefits.

Table 3.1 above summarizes the income thresholds introduced by the reform and the corresponding family allowance amounts per month. As previously mentioned, the new amount of family allowances received by the family varies based on the total number of children within the household, and on the treatment status, *i.e.* based on their household taxable income's (at $t2-$) position with respect to the thresholds introduced by the reform. For instance, after the reform, a two-child household with an annual income below 67,140€ continues to receive the full family allowance (129.34€ in total for two children), whereas a household earning between 67,140€ and 89,491€ is now entitled to only half of this amount for two children (64.67€). Income thresholds increase with the number of children. Indeed, for households with three children, those earning between 72,735€ and 95,086€ receive half benefits (147.53€), while those above 95,086€ receive a quarter rate (73.76€). Note that for each additional child beyond the third, the allowance is set at 82.86€ per additional child for households receiving half benefits and 41.44€ per child for those in the quarter-benefit group. This means that for instance a four-child household in the half-benefits group would be eligible to 230.39€ ($147.53 + 82.86$) per month for all four children.

To better understand the coverage of the reform in terms of household income, we plot the cumulative distribution of annual household income in Figure C.1 in the Appendix, by family-type, against the income thresholds introduced by the reform. Overall, this descriptive evidence highlights that even though the thresholds appear high (*cf.* Table 3.1), the reform seems to affect 35% of households with two children,

and about the same share of households with three children.

3.3 Data

3.3.1 Statistiques sur les Ressources et Conditions de Vie

We use the Statistiques sur les Ressources et Conditions de Vie (SRCV) — a panel survey conducted annually by the Institut National de la Statistique et des Etudes Economiques (INSEE) since 2004 — to assess the impact of the 2015 family allowances reform on families and children. The SRCV serves as the French component of the European Union's Statistics on Income and Living Conditions (EU-SILC), surveying approximately 12,000 households each year.[3] Within these households, every individual over 16 also completes an individual questionnaire, covering around 26,000 individuals in total.

SRCV data are particularly suitable for our analysis on two key aspects. First, detailed information on income and social benefits is precisely measured as it is directly extracted from administrative data sources.[4] This distinct feature allows us to accurately determine whether a household is affected by the reform by precisely identifying whether its income falls below or exceeds the established thresholds, while also providing detailed information on benefit receipt.

Second, SRCV data provide rich information at both the household and individual levels across two annually surveyed domains. The primary domain covers individuals' sociodemographic characteristics in interviewed households, along with detailed data on income, financial status, living conditions, employment, housing, education, and health. The secondary domain, consisting of *ad hoc* modules, rotates on a three- to six-year basis and generally focuses on themes such as housing conditions, children's living conditions, indebtedness, and well-being.

3.3.2 Final Samples

We use data from the SRCV survey (2010–2019) to construct two distinct analytical samples aimed at capturing the effects of the 2015 family allowances reform. The full period includes five pre-reform years (2010–2014) and five post-reform years (2015–2019).

Sample of adults.

The first sample includes individuals aged 18 to 65 (over 99% of the adult population in our dataset, with a few outliers aged over 65), surveyed by the SRCV team, and living in households with at least two children. We impose this condition to ensure that all families are eligible in principle for family allowances, since households with only one child are not entitled to these benefits. Using information on household

[3] In addition, 3,000 households are added each year as part of a refreshment sample.

[4] Household taxable income is extracted from tax records, while information on social benefits is based on data from the relevant national agencies, including the National Family Allowance Fund (Caisse Nationale des Allocations Familiales), the Agricultural Social Mutual Fund (Mutualité Sociale Agricole), and the National Pension Insurance Fund (Caisse Nationale d'Assurance Vieillesse).

structure, we merge relevant household-level variables to individuals, such as taxable income, used to construct our treatment status. Specifically, we construct a time-varying treatment status for each household (thus individual) and year, assigning them to one of three groups — full benefits, half benefits, or quarter benefits — based on their taxable income and number of dependent children, in accordance with the eligibility thresholds introduced by the 2015 reform (see Table 3.1).[5] The final adult estimation sample includes 20,207 individuals across 7,763 households. Table C.1 in the Appendix presents key demographic characteristics, disaggregated by treatment status. This sample will be used to analyze the effects of the reform on household-level outcomes, such as financial resources and deprivation, as well as parental labor supply.

Sample of children.

The second sample includes children aged 0 to 12, corresponding to the age range covered by the SRCV childcare module. We construct a panel of children observed annually between 2010 and 2019, ensuring each child lives in a household with at least two children, to satisfy eligibility conditions for family allowances. As with the adult sample, we assign treatment dynamically, updating each year based on the household's taxable income and number of dependent children. The final estimation sample includes 11,235 children, aged seven on average, and approximately 49% of them are girls. No additional selection is based on personal characteristics beyond eligibility and age-range constraints. This sample is used to estimate the effects of the reform on children's time allocation across different non-parental childcare arrangements, as reported in SRCV by their parents.

3.3.3 Variables of Interest

3.3.3.1 Household Financial Resources and Deprivation

Before turning to the analysis of deprivation measures, we first examine objective measures of poverty and financial well-being. Panel A of Table 3.2 reports descriptive statistics on key economic indicators, including household income and the amount of family allowances — extracted from administrative datasources, as mentioned in Subsection 3.3.1 — as well as measures of savings, poverty and self-reported financial hardship. We further describe how we construct and employ these variables in the Appendix, Subsection C.2.2.

As expected, households in the full-benefits group exhibit systematically lower levels of income and savings compared to those in the half-benefits or quarter-benefits groups, reflecting their lower position in the income distribution. This is consistent with the reform's design, which targeted higher-income households for benefit reductions. In particular, total household savings were substantially higher in the

[5]Each wave of SRCV includes extracted tax information from the previous year. We use this structure to assign treatment status dynamically, updating it each year based on household income from two years prior (*e.g.*, income in the 2014 wave reflects 2013 income). This allows us to define annual treatment status in accordance with both the eligibility thresholds and changes in household income or size over time.

quarter-benefits group (€72,649) relative to the half-benefits group (€32,835) and the full-benefits group (€10,535) prior to the reform. However, savings levels exhibit a downward trend across all groups, with the largest absolute decline observed in the quarter-benefits group (−€2,176). In addition, measures of subjective financial strain suggest that households in the half-benefits and quarter-benefits groups were significantly more likely to report financial difficulties post-reform, with an increase of 2.9 and 2.2 percentage points, respectively. While these trends do not establish causal relationships, they underscore the importance of examining how benefit reductions may have contributed to changes in financial hardship.

Table 3.2: Descriptive Statistics for Household Financial Resources and Deprivation

	Full Benefits			Half Benefits			Quarter Benefits		
	Before	After	Difference	Before	After	Difference	Before	After	Difference
Panel A: Financial and economic measures									
Gross annual household income	45910	46844	933***	78788	78891	103	144550	137901	-6649***
Net annual household income	38464	38794	329***	62794	61802	-992***	110537	100409	-10128***
Family allowances amount	4443	4445	2	3202	2926	-276***	2807	1893	-914***
Total household savings	10535	11455	919***	29835	30870	1034	72649	70473	-2176
Household savings above median	0.343	0.341	-0.002	0.643	0.624	-0.018**	0.769	0.746	-0.023***
Increase in total savings over last 12 months	0.162	0.163	0.001	0.283	0.282	-0.001	0.367	0.365	-0.002
In poverty	0.235	0.270	0.035***	0.000	0.001	0.000	0.000	0.002	0.002**
Poverty in living standards	0.208	0.206	-0.002	0.046	0.046	0.000	0.014	0.021	0.007***
Consider themselves in financial need	0.393	0.401	0.008	0.127	0.156	0.029***	0.105	0.127	0.022***
Panel B: Deprivation measures									
Unable to									
Eat meat (fish) every other day	0.091	0.109	0.018***	0.018	0.031	0.013***	0.009	0.017	0.008***
Buy new clothes	0.172	0.184	0.011***	0.034	0.043	0.009**	0.012	0.021	0.009***
Own two pairs of good shoes	0.133	0.131	-0.002	0.027	0.029	0.002	0.010	0.013	0.003
Replace worn-out furniture	0.382	0.364	-0.019***	0.116	0.115	-0.001	0.043	0.069	0.026***
Keep home at comfortable temperature	0.081	0.080	-0.001	0.023	0.021	-0.002	0.012	0.016	0.004*
Host relatives	0.129	0.130	0.002	0.031	0.039	0.009**	0.016	0.020	0.004
Do gifts	0.122	0.120	-0.002	0.023	0.027	0.004	0.011	0.008	-0.003
Afford holidays	0.395	0.366	-0.030***	0.121	0.117	-0.004	0.047	0.047	-0.000
Cover unexpected expense	0.490	0.521	0.031***	0.156	0.173	0.016**	0.067	0.074	0.007
Number of deprived items	1.990	1.997	0.007	0.544	0.592	0.047*	0.221	0.281	0.059***
House deprivation	0.394	0.378	-0.016***	0.122	0.121	-0.000	0.048	0.074	0.026***
Clothing deprivation	0.211	0.223	0.012***	0.046	0.055	0.009**	0.015	0.026	0.011***
Food deprivation	0.091	0.109	0.018***	0.018	0.031	0.013***	0.009	0.017	0.008***
Leisure deprivation	0.573	0.589	0.016***	0.208	0.216	0.008	0.087	0.097	0.010*
Observations	21045	17157	38202	5447	5804	11251	5289	5883	11172

Notes: This table reports the means of household-related variables by treatment status, before and after the reform, including measures of financial and economic conditions (Panel A), and measures of deprivation (Panel B).

Turning to household deprivation, Panel B of Table 3.2 reports summary statistics for individual deprivation items, capturing whether households are (un)able to afford key consumption goods and services. To this extent, the head of each interviewed household reports the items their household can afford, such as eating meat or fish every other day, purchasing new clothing, or affording holidays. For all nine items considered, responses are recoded to indicate the proportion of items that households are *unable to afford*, and we report their distribution, by item. Across all items, full-benefit households report significantly higher levels of material deprivation relative to those in the half-benefits or quarter-benefits groups.

To provide a more structured analysis, we finally construct five alternative deprivation measures, also reported in Panel B, divided in two sets. The first is a deprivation score, which sums across the nine individual deprivation items to provide an aggregate measure of material hardship. Our descriptive statistics indicate that,

on average, households in the full-benefits group report being unable to afford 1.99 out of the nine items, compared to 0.54 in the half-benefits group and 0.22 in the quarter-benefits group.

The second set of measures captures deprivation by category: (i) *house deprivation* is defined as the inability to either keep the home at a comfortable temperature or replace furniture (correlation $= 0.31, p < 0.01$) affecting 39.4% of full-benefit households prior to the reform, compared to 11.6% and 4.3% for the half-benefits and quarter-benefits groups, respectively; (ii) *clothing deprivation* is based on the inability to buy new clothes or own two pairs of good shoes (correlation $= 0.55, p < 0.01$), and is reported by 21.1% of full-benefit households, relative to 4.6% in the half-benefits group and 1.5% in the quarter-benefits group; (iii) *food deprivation* captures whether households are unable to afford meat or fish every other day and is reported by 9.1% of full-benefit households, but only 1.8% of half-benefits households and 0.9% of quarter-benefits households; and (iv) *leisure deprivation*, which is the broadest category, reflecting the ability to host relatives, do gifts, afford holidays, or cover unexpected expenses (between-item correlations > 0.35, $p < 0.01$), which affected 57.3% of full-benefit households, 20.8% of half-benefit households, and 8.7% of quarter-benefit households before the reform.

Overall, across all our deprivation measures, our descriptive evidence highlights the strong association between income levels and material deprivation, reinforcing the need to investigate how benefit reductions may have affected consumption patterns in light of the 2015 family allowances reform.

3.3.3.2 Parental Labor Supply

To investigate the impact of the reform on parental labor supply, we consider several outcomes related to employment and income. First, we define a binary variable set to one if the individual reports being either in full-time or part-time employment or self-employment. Second, we add measures related to the type of employment and report the share of self-employed individuals — given their higher flexibility in adjusting labor supply than employees — as well as the share of individuals with more than one job. Third, we look at the labor force participation rate, defined as the share of individuals who are either employed or unemployed. Fourth, we look at the intensive margin of labor supply, *i.e.*, the usual number of hours worked per week reported, from which we further derive a binary variable set to one for individuals who report working usually 35 hours or more per week, as this constitutes the legal duration of work for a full-time job.[6] Lastly, we consider the individual annual labor income in euros (€) before tax.

Descriptive statistics for the full sample (Panel A) and disaggregated by gender (Panels B and C) as presented in Table 3.3 below. Overall, our descriptive statistics indicate that, across all groups, labor supply outcomes increase as benefit levels decline. In addition, men consistently exhibit higher (full-time) employment rates, labor force participation, hours worked, and annual labor income compared to women, regardless of the benefit group they belong to. On the other hand, while women's annual labor income remain below that of men, their labor supply is higher for those

[6]See the official French administration website.

exposed to larger benefit cuts. Indeed, for instance, the share of women working full time is significantly larger in the quarter-benefits group — the most affected by benefits cuts — than in the full-benefits group — those not affected by the reform. These patterns suggest that, at least at the descriptive-level, reductions in benefits may disproportionately influence secondary earners, potentially shifting household labor supply dynamics. However, we need to rely on exogenous variation in benefit to tease out the causal relationship.

Table 3.3: Descriptive Statistics for Parental Labor Supply

	Full Benefits			Half Benefits			Quarter Benefits		
	Before	After	Difference	Before	After	Difference	Before	After	Difference
Panel A: All individuals									
In employment	0.650	0.651	0.002	0.763	0.777	0.015*	0.724	0.738	0.014*
Self-employment	0.062	0.075	0.014***	0.066	0.061	-0.005	0.124	0.111	-0.014**
More than one job	0.345	0.336	-0.009*	0.224	0.201	-0.023***	0.272	0.242	-0.030***
Labor force participation	0.649	0.640	-0.009*	0.765	0.769	0.004	0.724	0.730	0.006
Weekly hours worked	36.574	36.355	-0.219	39.032	38.774	-0.258	42.146	41.998	-0.148
Full-time employment	0.757	0.764	0.006	0.805	0.830	0.025***	0.836	0.855	0.019**
Annual labor income	14049	14475	426***	24280	25623	1343***	35145	37460	2315***
Observations	17157	21045	38202	5804	5447	11251	5883	5289	11172
Panel B: Men									
In employment	0.720	0.709	-0.011*	0.800	0.804	0.004	0.744	0.768	0.024**
Self-employment	0.083	0.096	0.013***	0.080	0.071	-0.009	0.148	0.138	-0.010
More than one job	0.267	0.269	0.002	0.181	0.173	-0.008	0.250	0.214	-0.036***
Labor force participation	0.721	0.694	-0.027***	0.810	0.797	-0.014	0.745	0.760	0.015
Weekly hours worked	40.214	39.676	-0.538***	42.657	41.776	-0.881***	46.676	45.668	-1.007***
Full-time employment	0.922	0.917	-0.005	0.942	0.946	0.005	0.947	0.950	0.003
Annual labor income	17811	17760	-51	29381	29972	591	43811	47184	3373***
Observations	8119	10036	18155	2961	2766	5727	3024	2770	5794
Panel C: Women									
In employment	0.585	0.600	0.014**	0.724	0.749	0.025**	0.702	0.707	0.005
Self-employment	0.041	0.056	0.015***	0.050	0.049	-0.001	0.098	0.082	-0.017*
More than one job	0.415	0.395	-0.021***	0.269	0.231	-0.038***	0.297	0.272	-0.025**
Labor force participation	0.583	0.591	0.008	0.718	0.741	0.023*	0.700	0.698	-0.002
Weekly hours worked	32.556	32.905	0.349*	34.881	35.454	0.574**	36.872	37.837	0.965***
Full-time employment	0.575	0.604	0.029***	0.648	0.701	0.053***	0.706	0.747	0.041***
Annual labor income	10620	11525	904***	19018	21094	2077***	25614	27173	1559*
Observations	9038	11009	20047	2843	2681	5524	2859	2519	5378

Notes: Means of individuals' labor supply outcomes after and before the reform, as well as the difference, by treatment status introduced in Section 3.2. Results are presented for all individuals (Panel A) and further split by gender (Panels B and C).

3.3.3.3 Childcare Attendance

The 2015 reform of family allowances reduced benefit levels for certain households, and thus likely influenced how families allocated resources for childcare. Since family allowances are intended to help cover childcare costs, the reform is expected to have affected the time spent by children in different childcare arrangements, particularly for younger children.

Before describing our sample of children, it is useful to contextualize childcare in the French setting, where the structure of childcare provision and its affordability vary substantially by the age of the child.

According to a recent report from Onape (2024), parental care remains the main form of childcare for children under age 3 in France — 56% of children are primarily cared for by parents during the week (between 8am and 7pm) while 44% use formal or informal care arrangements. Among these, 21% are primarily cared for by childminders (*assistantes maternelles*) and in-home providers, 18% in daycare (*crèches*), 3% by grandparents or relatives, 2% attend pre-school.[7] However, many families combine multiple arrangements: nearly 50% of children under three use a mix of parental and formal care, where childminders and daycare constitute the majority of available formal care places in France, with 31.5 and 23.3 spots per 100 children respectively (Onape, 2024).

Parents' demand for daycare and childminders likely depends on availability but also on the out-of-pocket expenses associated with these formal care types.[8] These vary based on household income, the structure (daycare *vs.* childminders), and the number of hours of care. As an example drawn from Onape (2024), a dual-earner couple with a child in full-time care (162 hours per month) would pay €306 per month for a childminder, but €138 for daycare. For higher-income households (over five times the minimum wage), a childminder becomes slightly more cost-effective than a crèche, though the difference is minor. Thus, overall, for middle-income households relying on full-time care, childminders tend to be more expensive than crèches. In addition, the relative affordability of each option also depends on the number of hours of care. For shorter durations (*e.g.*, 72 hours/month), childminders are often more advantageous for middle- and high-income families. We provide additional information on how daycare and childminders operate in France, in the Appendix Subsection C.2.1.

Table 3.4: Summary of Childcare Modes in the French Context

Childcare Mode	Age Range	Applicable Subsidies
Daycare	0 to 3	PAJE (Prestation d'Accueil du Jeune Enfant); CMG (Complément de libre choix du Mode de Garde); Municipal subsidies; Family income-based reduction.
Professional childminder	0 to 6	PAJE; CMG; tax deductions.
Cared by a relative	0 to 12	
Pre-school	3 to 6	(Mostly) publicly funded; Families may pay for canteen and extracurricular activities.
School	6 to 18	(Mostly) publicly funded; Families may pay for canteen and extracurricular activities.
After-school	3 to 17	Income-based reduction.

Notes: Age ranges are inclusive. PAJE and CMG refer to subsidies received by parents, see Appendix Subsection C.2.1 for further details.

[7] In some priority education areas (*zones d'éducation prioritaire*), children can attend pre-school from the age of 2. These areas are designated by the French Ministry of Education to provide additional support to schools serving disadvantaged communities, and thus often offer earlier access to publicly funded pre-school.

[8] Parents can indeed benefit from subsidies, and tax credits to cover for these expenses.

When children reach the age of three, they become eligible for publicly-funded pre-school (*école maternelle*), and from six onwards, school enrollment is mandatory. Notably, a 2019 law lowered the compulsory school age from six to three. Given our data spans the period 2010–2019 and given pre-school enrollment became mandatory for the school start of September 2019, the latter may be captured in our last wave if the household interviewee was conducted between September and December 2019. We further perform checks later dropping 2019, and demonstrate that our main result remain robust without including this year (see Table C.10).

Finally, SRCV does not distinguish between public and private schooling, although public education is predominant in France. As of 2022, 87% of primary and 79% of secondary school students were enrolled in public schools according to data from the French Ministry of Education.[9]

We present in Table 3.5 below the average number of hours per week that children spend in each childcare arrangement, broken down by age group. Panel A includes children aged 0 to 2 — those not yet eligible for (free) pre-school — who are typically cared for by a parent, a crèche, or a childminder. Panel B includes children aged 3 to 5, who may attend pre-school while still relying partially on other care. Panel C includes children aged 6 to 12, who are enrolled in primary school but may also require after-school arrangements (e.g., relatives or childminders).

Overall, our descriptive evidence indicates a gradient in the use of professional childminders and daycare centers based on the treatment status. Looking at younger children (Panel A), it is interesting to note that children in families with higher income — thus entitled to either half or quarter benefits after the reform — tend to spend more time in daycare centers and/or with professional childminders than children in the full-benefits group. This reinforces evidence on the socioeconomic gap in early childhood education and care (ECEC) enrollment, with lower-income families underrepresented.[10] Further, for older children (*e.g.*, Panel B) this gradient remains in the same direction. Such gradient in the use of formal childcare arrangements motivates the need to understand whether the reform affected this distribution of time spent in different childcare arrangements by child's age.

[9]Link to the dataset.

[10]In France, in 2014, ECEC attendance among families in the lowest-income tertile was 65 percentage points lower than that of families in the highest-income tertile. Despite the availability of subsidies and means-tested fees, 19% of families living below the poverty line were utilizing ECEC services by 2018 (Carbuccia, 2024).

Table 3.5: Descriptive Statistics for Children's Time Spent in Different Childcare Arrangements

	Full Benefits			Half Benefits			Quarter Benefits		
	Before	After	Difference	Before	After	Difference	Before	After	Difference
Panel A: 0–2 years-old									
Hours per week									
In daycare	8.440	9.046	0.606	17.509	19.971	2.462*	19.487	19.821	0.334
With a professional childminder	2.160	1.741	-0.419*	5.429	2.808	-2.621***	6.719	6.328	-0.391
Cared by a relative	1.847	1.901	0.054	2.396	2.079	-0.317	2.077	3.084	1.007
Observations	2215	1693	3908	380	343	723	224	229	453
Panel B: 3–5 years-old									
Hours per week									
In daycare	1.966	2.017	0.051	4.406	3.861	-0.544	4.724	4.820	0.096
With a professional childminder	0.937	0.561	-0.376***	1.707	1.576	-0.131	3.779	2.265	-1.514***
In pre-school	23.536	24.133	0.597**	24.340	25.187	0.848	24.660	25.192	0.532
Cared by a relative	1.521	1.353	-0.168	1.570	1.649	0.079	1.709	2.286	0.576
Observations	3074	2473	5547	536	577	1113	377	366	743
Panel C: 6–12 years-old									
Hours per week									
In school	26.876	27.449	0.573***	27.926	28.414	0.488*	28.305	29.296	0.992***
With a professional childminder	0.166	0.129	-0.037	0.329	0.239	-0.090	1.334	1.152	-0.182
Cared by a relative	0.981	0.998	0.017	1.330	1.489	0.158	1.315	1.159	-0.156
Observations	3074	2473	5547	536	577	1113	377	366	743

Notes: This table reports the average weekly hours spent by children in each childcare arrangement, disaggregated by age groups: 0 to 2 years (Panel A), 3 to 5 years (Panel B), and 6 years and older (Panel C). See Table 3.4for additional information on childcare arrangements in the French context. Estimates are presented for the periods before and after the reform, stratified by treatment status, as defined in Section 3.2.

3.4 Empirical Strategy

Below, we present the different strategies employed to quantify the impact of the reform on households, parental behavior, and the time children spend in different childcare facilities.

3.4.1 Household-Level Analysis

To estimate the effect of the 2015 reform, we adopt a difference-in-differences (DiD) framework. The reform reduced the generosity of family allowances for higher-income households, creating three distinct groups of households:

1. Households that continued receiving full benefits after 2015 (control);
2. Households whose benefits were halved (treated group 1), and;
3. Households whose benefits were reduced to a quarter (treated group 2).

Treatment status is defined at the household level based on their household income at $t-2$ with respect to the thresholds introduced by the reform. We denote by $D_h^{1/2}$ an indicator equal to one if a household h experienced a reduction half benefits, and by $D_h^{1/4}$ an indicator equal to one if a household h experienced a reduction to quarter benefits. The omitted category corresponds to households that continued receiving full benefits. To make this distinction explicit, we write the baseline specification as:

$$Y_{ht} = \alpha + \beta Post_t + \gamma_1 D_h^{1/2} + \gamma_2 D_h^{1/4} + \theta_1 (Post_t \times D_h^{1/2}) + \theta_2 (Post_t \times D_h^{1/4}) + \lambda_t + \delta_h + \epsilon_{ht}, \tag{3.1}$$

where Y_{ht} is the outcome of interest for household h in year t, α is the intercept, $Post$ is a binary indicator for the post-treatment period, $D_h^{1/2}$ and $D_h^{1/4}$ indicate whether the household was subject to the half- or quarter-benefit reduction — the omitted group consists of households that continued receiving full benefits. We further include year (λ_t) and household (δ_h) fixed effects. Thus, the coefficients γ_1 and γ_2 capture baseline differences between treatment and control groups, while θ_1 and θ_2 estimate the reform's effect on the treated groups relative to the control group.

In addition, to examine the dynamic effects of the reform, we estimate an event study specification. The event time (*event*) is defined relative to the reform year 2015 denoted $t = 0$, negative values indicate pre-reform years, and positive values post-reform years. Note that we exclude exclude $t = -1$ as the reference period. As before, treatment status distinguishes between three groups: (1) households continuing to receive full benefits (control), (2) households whose benefits were reduced to one-half ($D_h^{1/2} = 1$), and (3) households whose benefits were reduced to one-quarter ($D_h^{1/4} = 1$). The event study specification is:

$$Y_{ht} = \alpha + \sum_{t \neq -1} \beta_t \mathbf{1}[event = t] + \gamma_1 D_h^{1/2} + \gamma_2 D_h^{1/4} + \sum_{t \neq -1} \theta_{1t} \left(\mathbb{1}[event = t] \times D_h^{1/2} \right) + \sum_{t \neq -1} \theta_{2t} \left(\mathbb{1}[event = t] \times D_h^{1/4} \right) + \lambda_t + \delta_h + v_{ht}. \tag{3.2}$$

Here, β_t captures year-specific deviations in outcomes for the full-benefit group relative to the baseline year, while γ_1 and γ_2 capture average baseline differences between the treated and control households. The dynamic effects of the reform are given by θ_{1t} and θ_{2t}, which trace the relative evolution of outcomes for the half- and quarter-benefit groups respectively, compared with the full-benefit group.

3.4.2 Individual-Level Analysis

For the analyses at the individual-level, *i.e.*, for parents and children, we use the specifications outlined below. We estimate a difference-in-differences model at the individual-level modifying Equation (3.2) as follows:

$$Y_{it} = \alpha + \beta Post_t + \gamma_1 D_i^{1/2} + \gamma_2 D_i^{1/4} + \theta_1 (Post_t \times D_i^{1/2}) + \theta_2 (Post_t \times D_i^{1/4}) + \delta_i + \zeta X_{it} + \lambda_t + \varepsilon_{it}. \tag{3.3}$$

The differences with Equation (3.2) are twofold. First, we replace household fixed effects with individual fixed effects (δ_i). Second, we we add individual-level controls, specifically age and age squared, contained in X_{it}. Treatment status is still determined at the household level, but each individual inherits the treatment assignment of their household. Accordingly, $D_i^{1/2}$ indicates whether individual i belongs to a household subject to the half-benefit reduction, and $D_i^{1/4}$ indicates whether they belong to a household subject to the quarter-benefit reduction. The omitted category consists of individuals in households that continued to receive full benefits. The coefficients γ_1 and γ_2 capture baseline differences between the treated and control groups, while θ_1 and θ_2 measure the relative impact of the reform on the half- and quarter-benefit groups.

As for households, we also estimate an event-study specification, given by Equation (3.4) below. The terms are defined as in Equation (3.2), except that we again replace household fixed effects with individual fixed effects and control for age and age squared in X_{it}.

$$Y_{it} = \alpha + \sum_{t \neq -1} \beta_t \, \mathbb{1}[event = t] + \gamma_1 D_i^{1/2} + \gamma_2 D_i^{1/4} + \sum_{t \neq -1} \theta_{1t} \left(\mathbb{1}[event = t] \times D_i^{1/2} \right) + \sum_{t \neq -1} \theta_{2t} \left(\mathbb{1}[event = t] \times D_i^{1/4} \right) + \delta_i + \zeta X_{it} + \lambda_t + \nu_{it}. \tag{3.4}$$

Thus, the parameters γ_1 and γ_2 capture baseline differences between treated and control individuals, while θ_{1t} and θ_{2t} trace the dynamic evolution of outcomes for the

half- and quarter-benefit groups relative to the full-benefit group. In addition, pre-treatment coefficients provide a direct assessment of the plausibility of the parallel-trends assumption, which we further discuss in the following subsection.

3.4.3 Internal Validity

A key concern when using DiD and event study designs is whether the main assumptions necessary for identification hold in our setting. We discuss below the main threats to validity and provide robustness checks to mitigate these concerns.

Parallel trends assumption.

The validity of our DiD approach relies on the assumption that, absent the reform, treated and control groups would have followed parallel trends in our outcome variables. We assess this assumption in several complementary ways. First, we provide graphical evidence (*e.g.*, Figures 3.1, and C.3) from our event study that pre-reform estimates are statistically indistinguishable from zero. The absence of significant pre-trend deviations suggests that treated and control groups followed parallel trajectories before the reform. Second, we implement a placebo analysis by assigning a false reform date (2013) and show in Figure 3.3, that there is no significant effect before the actual policy change, further indicating that parallel trends hold. Third, we conduct a placebo test in the treatment group by assigning a "fake" treatment status to one-child households — who were *de facto* not eligible for family allowances. We then run our main specification using income thresholds that only apply to families with two or more children. As shown in Table 3.10, we find no consistent effects (compared with the DiD results), suggesting that our results are not driven by different time trends across income levels, unrelated to the policy. Fourth, we implement a triple-difference (DDD) strategy that compares trends in eligible (two-child or more) and ineligible (one-child) families across income groups. Results from Tables 3.11 and 3.12 further rule out differential time trends across income groups unrelated to the reform, and provide supporting evidence for the robustness of our main results. Finally, across all specifications, we include household or individual fixed effects, as well as year fixed effects, to account for potential unobserved heterogeneity and common time shocks.

Policy anticipation.

Another potential threat is that households could have anticipated the 2015 reform and adjusted their behavior before implementation, consequently violating the exogeneity of treatment timing.[11] First, given that eligibility to the new amounts of family allowances is based on income from two years prior, households could not manipulate their taxable income contemporaneously to remain eligible. Second, as discussed above, if parents anticipated the policy, we should see any significant effect before 2015, which we do not observe. Overall, we find no evidence of systematic violations of the exogeneity of treatment timing assumption, which is reassuring.

[11]The reform was first announced on the 22nd of December 2014, when published as an official Law to fund the Social Security budget for 2015. See official text here.

3.5 Results

3.5.1 Effects on Households

3.5.1.1 Financial Resources

We begin by examining how the 2015 family allowances reform affected household financial conditions. Specifically, we look at changes in three objective measures of these conditions, and one subjective measure. These are respectively (1) household savings, (2) poverty, (3) poverty in living conditions, and (4) a self-reported measure for being in financial need. Further description of how we construct these variables is provided in the Appendix, Subsection C.2.2, as well as the pairwise correlation matrix between these measures (Table C.11). As shown in Table C.11, we find statistically significant and moderate negative correlations between savings and poverty ($r = -0.29$, $p < 0.01$), as well as between savings and poverty in living conditions ($r = -0.32$, $p < 0.01$). Moreover, the subjective measure of financial hardship is positively correlated with both objective measures of poverty ($r = 0.34$, $p < 0.01$) and poverty in living conditions ($r = 0.20$, $p < 0.01$), suggesting a meaningful association and a consistent pattern between subjective and objective dimensions of household financial resources.

Our difference-in-differences (DiD) results for these outcomes are presented in Table 3.6 below and we focus on the interaction terms, which show how each outcome changed for families whose benefits were reduced to (a) half or (b) a quarter of the original amount compared to those who continued to receive the full benefit.

Our DiD results indicate that cutting benefits for families, either by half or by three-quarter, had some significant and sizeable effects on both objective and subjective measures of household financial conditions.

As for the objective measures, although the reform did not affect households' savings likelihood, it increased the likelihood of being classified as poor by 3.3 percentage points (pp) relative to their baseline mean for families who receive a quarter of the initial amount of benefits. This objective effect is further confirmed when looking at results for the self-reported measure of financial hardship, highlighting the importance of possible financial difficulties parents may have faced and induced by the decrease in household financial resources. Indeed, benefits cuts raised the probability of reporting financial hardship by $3pp$ for the half-benefits group and by $4.2pp$ for the quarter-benefits group. Relative to their respective pre-reform means of 12.7% and 10.5%, these represent roughly a 24% and 40% increase in self-reported hardship.

In addition, we present the event study estimates of our model in Figure C.2 in the Appendix. Overall, our results are qualitatively consistent with the DiD results and confirm that the reform had a lasting impact on household financial resources and well-being, with affected families experiencing declines in net income as well as increases in poverty, poverty related to living conditions, and self-reported financial hardship. However, while point estimates suggest that families in the quarter-benefits group (red) experienced larger effects, the confidence intervals overlap with those of the half-benefits group (blue), indicating that differences between the two groups should be interpreted with caution. Thus, our estimates point to a gradual

Table 3.6: Average Treatment Effects on Household Financial Resources and Deprivation

	(1) Savings Above Median		(2) Poverty		(3) Poverty in Living Conditions		(4) Self-Reported Hardship	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Half Benefits	0.021		-0.080***		-0.015		-0.022**	
	(0.016)		(0.009)		(0.010)		(0.011)	
Post	0.102***	0.105***	-0.032*	-0.029	-0.063***	-0.064***	-0.054***	-0.054***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.017)	(0.017)	(0.016)	(0.016)
Post × Half Benefits	-0.015		0.014		0.019		0.030**	
	(0.019)		(0.011)		(0.012)		(0.012)	
Quarter Benefits		0.018		-0.181***		-0.033**		-0.045***
		(0.028)		(0.021)		(0.015)		(0.014)
Post × Quarter Benefits		-0.031		0.033***		0.019		0.042***
		(0.022)		(0.012)		(0.012)		(0.010)
Mean Dep. Var	0.643	0.769	0.000	0.000	0.046	0.014	0.127	0.105
Observations	11251	11172	11251	11172	11251	11172	11251	11172
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.1) on (1) an indicator for household savings above median, (2) a poverty indicator, (3) a poverty in living conditions indicator, and (4) a self-reported measure of financial hardship, as defined in Subsection 3.3.3. Standard errors, clustered at the household-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. The two treatment statuses are defined as: (a) "Half Benefits" if a household with at least two children receives half of the initial amount of benefits after the reform, and (b) "Quarter Benefits" if the household receives a quarter of the initial amount of benefits, after the reform. These are both compared to a control group not affected by the reform, *i.e.*, households with at least two children whose income falls below the first threshold introduced by the 2015 reform.

divergence post-reform, particularly for poverty and self-reported financial hardship, where effects become more pronounced in later years.

3.5.1.2 Deprivation

We then turn to examine the impact of the reform on household deprivation. Table C.2 in the Appendix presents the DiD estimates for two deprivation measures: a binary indicator for deprivation and a deprivation score reflecting the number of essential items a household is unable to afford. While overall deprivation declined over time, the results indicate that households affected by benefit reductions experienced a relative worsening of their material well-being. Indeed, households in the half-benefits group became $3pp$ ($p < 0.05$) more likely to be deprived post-reform, compared to their baseline mean of 4.2%. For the quarter-benefits group, their likelihood of being deprived increased by $4.2pp$ ($p < 0.01$) with respect to their baseline mean of 0.3%. Similarly, the deprivation score increased for both groups, and the effect is even more pronounced for the quarter-benefits group (0.22 item, $p < 0.01$) given their baseline mean is equal to 0.16. These findings suggest that the reform had a negative impact on household financial stability, leading to greater material hardship and deprivation among those affected.

These results are confirmed by graphical evidence displayed in Figure 3.1 from our event study results where both treatment groups experienced a similar upward trend in our two deprivation measures post-reform, with overlapping confidence intervals (CIs). This suggests that differences between the half-benefits and quarter-

benefits groups are not statistically different, even though estimates appear more precisely estimated for the quarter-benefits group. The effects are particularly pronounced from two years after the reform onward, reinforcing the idea that financial strain may accumulate over time.

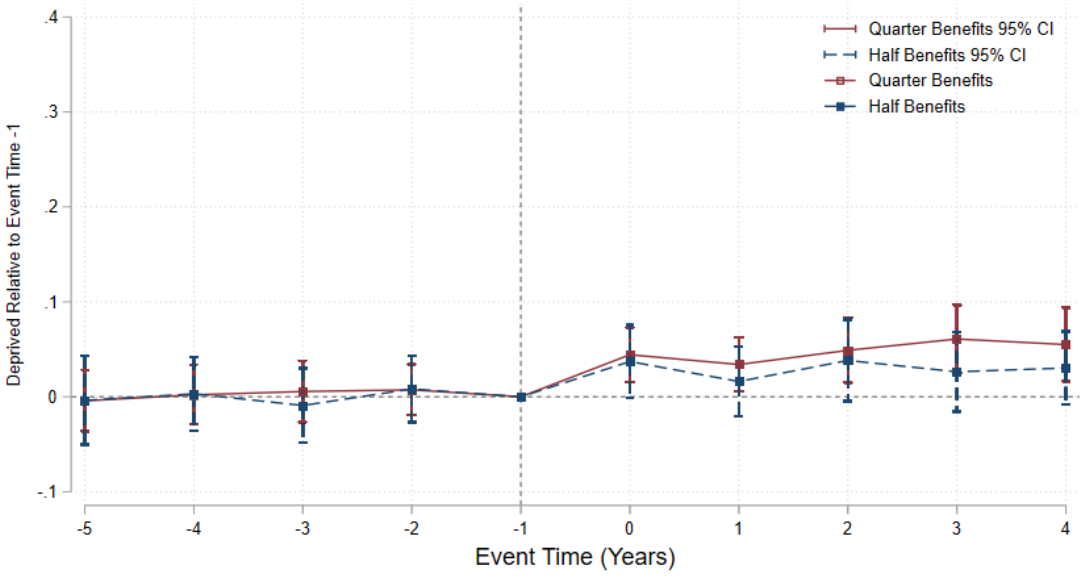
Building on the increase in the number of items a household is unable to afford, we finally investigate which items were most significantly affected by the reform. Our DiD results (Appendix Table C.4) show that household deprivation increased significantly across all categories, with stronger effects for families facing the lowest cuts.

For the half-benefits group, deprivation rose by 0.041 item for clothing ($p < 0.01$), 0.021 item for food ($p < 0.05$), and 0.031 item for housing ($p < 0.1$). Relative to baseline means of 0.039, 0.019, and 0.108 respectively, these translate into a 105% increase in food deprivation, 105% in clothing deprivation, and 29% in food deprivation. For the quarter-benefits group, we find a larger effect on housing deprivation (0.073, $p < 0.01$), but somewhat smaller increases in clothing (0.037, $p < 0.01$) and food (0.013, not significant) deprivation compared to the half-benefits group. These results suggest that the largest benefit cuts mainly affected housing-related deprivation, while increases in food and clothing deprivation were more pronounced among the half-benefits group. Event study results (Figure 3.2) support this pattern, showing a gradual rise in item-specific deprivation over time.

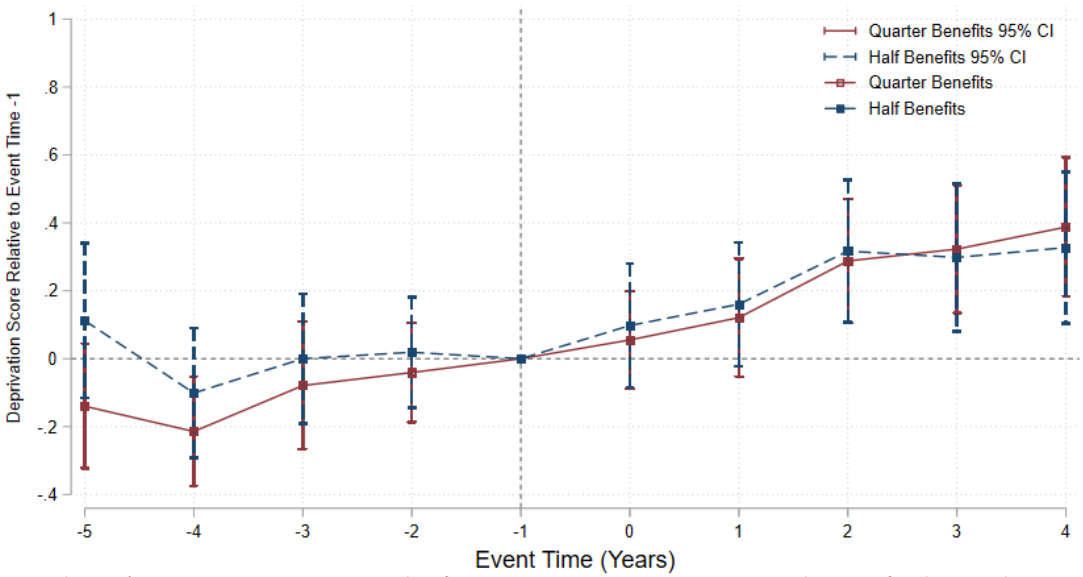
Finally, we explore whether the effects of the reform on deprivation vary depending on household structure, as different family types may have different vulnerabilities to benefit cuts. In particular, we focus on three subgroups: single-parent households, who typically face higher risks of poverty due to relying on a single income (Panel A), dual-parent households with an unemployed partner (Panel B), and dual-parent households with an inactive partner (Panel C),who may be especially sensitive to income shocks given the (temporary) absence of earnings from one adult in the household. For each group, we define household status based on pre-reform characteristics (at least one year prior) and estimate subgroup-specific effects on the likelihood of being deprived and on the number of deprivation items. As shown in the Appendix Table C.5, we find no clear evidence of heterogeneous effects: point estimates are broadly similar across groups and close to the full-sample estimates (Table C.2). Although effects appear somewhat larger for dual-parent households with an unemployed partner (Panel B), the differences are not statistically significant and should be interpreted cautiously due to limited statistical power in our subgroup analyses.

Figure 3.1: Impact of the Reform on Overall Deprivation

(a) Likelihood of Being Deprived



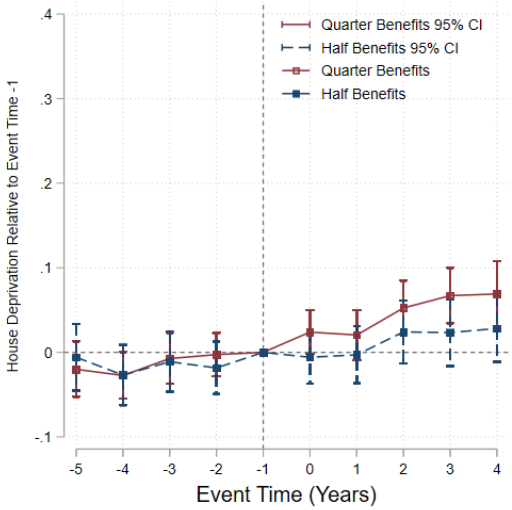
(b) Deprivation Score



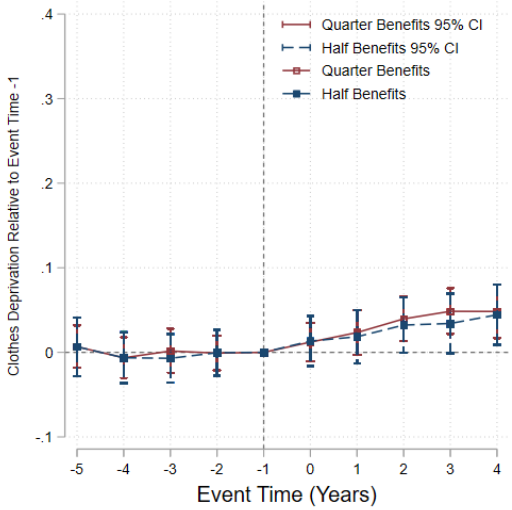
Notes: These figures present OLS results for Equation (3.2) on (a) an indicator for being deprived (provided by SRCV), (b) the number of items the household is unable to afford, as described in Subsection 3.3.3. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status, and for the corresponding treated group.

Figure 3.2: Impact of the Reform on Items of Deprivation

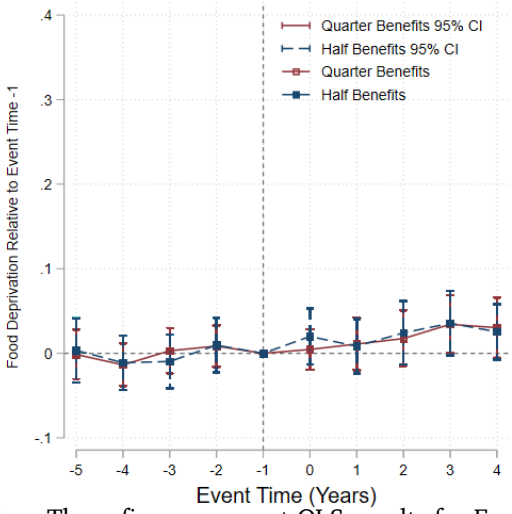
(a) House Deprivation



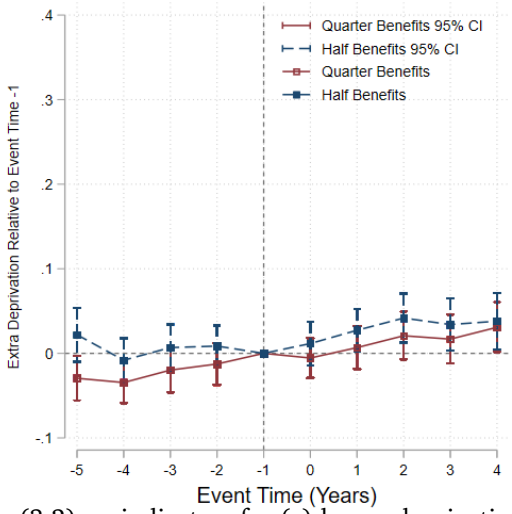
(b) Clothing Deprivation



(c) Food Deprivation



(d) Leisure Deprivation



Notes: These figures present OLS results for Equation (3.2) on indicators for (a) house deprivation, (b) clothing deprivation, (c) food deprivation, and (d) leisure deprivation, as described in Subsection 3.3.3. There are two treatment statuses defined as (1) "Half Benefits" if the household receives half of the initial amount of benefits, and (2) "Quarter Benefits", if the household receives a quarter of the initial amount of benefits, after the reform. These are both compared to a control group not affected by the reform, *i.e.*, households with at least two children whose income falls below the first threshold introduced by the 2015 reform. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group.

3.5.2 Effects on Parental Labor Supply

Next, we turn to investigate whether the reform had any effect on labor supply. We present in Table 3.7 below, OLS results for our DiD specification (Equation (3.3)) on our main labor supply outcomes. Results are presented for all individuals (Panel A), and further split by gender (Panels B and C). Overall, our results suggest that the reform had no effect on parental labor supply. It is however important to note that the point estimates for hours worked are always positive, which is consistent at least qualitatively, with a potential increase in working time.

Further, we estimate Equation (3.4) for the full sample. Although our event study results are imprecisely estimated, we also find no significant effect of the reform on the likelihood of being employed, in the labor force or in full-time work, nor on the number of hours worked (see Figure C.3 in the Appendix). This suggests that benefits cuts did not affect parental labor supply.

Further, we run heterogeneity analysis focussing on self-employed individuals, as they should have more flexibility than employees in adjusting their working time, and estimate Equation (3.4) this time breaking down our sample by individuals' employment status pre-reform — whether they were self-employed or not. Although self-employment could be the result of the policy, making the stratification less relevant, the proportion of individuals switching from wage employment to self-employment changed very little after the reform, and across both treatment groups. In the half-benefits group, the share of self-employed fell from approximately 6.58% before the reform to 6.08% after. Likewise, in the quarter-benefits group the share of self-employed moved from about 12.43% to 11.08%, respectively before and after the reform. These minor shifts thus suggest that employment-related transitions were relatively rare and similarly small across both groups, making this stratification relevant. Results of this exercise are presented in Table 3.8 below.

Consistent with our main findings, we do not observe statistically significant effects of the reform on employment, labor force participation, hours worked, or full-time status when disaggregating by pre-reform employment status. Among self-employed individuals (Panel B), point estimates are small and imprecisely estimated across all outcomes, with no indication of systematic changes in labor supply following the reform. For individuals who were already employed prior to the reform (Panel C), we again find no statistically significant effect, although the point estimates for hours worked are positive, particularly for the quarter-benefits group. These findings mirror those from Table 3.7, and further support the conclusion that the reform did not induce significant changes in parental labor supply on average, including among groups theoretically more flexible in adjusting their working schedule.

Lastly, as for deprivation, we investigate heterogeneity regarding marital status (single-parent household vs. not) and working status (unemployed vs. inactive partner). Results of this exercise are presented in the Appendix Table C.6. Overall, there seems to be no significant heterogeneous effect of the reform on parental labor supply. Although point estimates are not significant, we observe small but positive adjustments in weekly hours worked (column 3) and on the likelihood of working full-time (column 4).

Table 3.7: Average Treatment Effects on Parental Labor Supply, by Gender

	(1) Employment		(2) Labor Force Participation		(3) Hours Worked		(4) Full-Time	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: All individuals								
Half Benefits	0.027***		0.034***		1.481***		0.006	
	(0.008)		(0.008)		(0.359)		(0.010)	
Post	-0.074	-0.071	-0.060	-0.061	-3.067	-2.759	-0.105	-0.097
	(0.090)	(0.096)	(0.093)	(0.099)	(3.575)	(3.966)	(0.099)	(0.104)
Post × Half Benefits	-0.010		-0.012		0.261		0.014	
	(0.009)		(0.010)		(0.421)		(0.012)	
Quarter Benefits		0.022*		0.028**		1.009		0.003
		(0.013)		(0.013)		(0.625)		(0.017)
Post × Quarter Benefits		0.011		0.007		0.709		0.002
		(0.011)		(0.012)		(0.552)		(0.014)
Mean Dep. Var	0.763	0.724	0.763	0.722	30.125	30.751	0.805	0.836
Observations	11234	11165	11251	11172	11251	11172	8731	8190
Panel B: Men								
Half Benefits	0.028***		0.045***		1.890***		0.010	
	(0.010)		(0.011)		(0.555)		(0.009)	
Post	-0.096	-0.047	-0.091	0.001	-4.080	-0.698	-0.085	-0.020
	(0.121)	(0.132)	(0.125)	(0.144)	(5.197)	(6.296)	(0.090)	(0.094)
Post × Half Benefits	0.001		-0.009		0.209		0.009	
	(0.012)		(0.013)		(0.632)		(0.011)	
Quarter Benefits		0.032*		0.026		1.559		0.004
		(0.017)		(0.019)		(0.987)		(0.019)
Post × Quarter Benefits		0.016		0.029*		1.171		0.008
		(0.015)		(0.016)		(0.846)		(0.013)
Mean Dep. Var	0.800	0.744	0.809	0.744	34.607	34.982	0.942	0.947
Observations	5718	5789	5727	5794	5727	5794	4621	4377
Panel C: Women								
Half Benefits	0.026**		0.024**		1.067**		0.002	
	(0.011)		(0.012)		(0.455)		(0.018)	
Post	-0.046	-0.081	-0.028	-0.097	-1.788	-3.754	-0.120	-0.171
	(0.111)	(0.116)	(0.117)	(0.112)	(4.063)	(3.971)	(0.167)	(0.168)
Post × Half Benefits	-0.019		-0.012		0.392		0.019	
	(0.014)		(0.014)		(0.558)		(0.023)	
Quarter Benefits		0.011		0.029		0.432		0.004
		(0.018)		(0.019)		(0.761)		(0.029)
Post × Quarter Benefits		0.008		-0.014		0.289		-0.004
		(0.017)		(0.018)		(0.705)		(0.025)
Mean Dep. Var	0.724	0.702	0.715	0.697	25.500	26.099	0.648	0.706
Observations	5516	5376	5524	5378	5524	5378	4110	3813
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3) on labor supply outcomes introduced in Subsection 3.3.3. We focus here on (1) a binary variable for being employed, (2) a binary indicator for labor force participation, (3) the usual number of weekly hours worked, and (4) an indicator for full-time employment. Standard errors, clustered at the individual-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group.

Table 3.8: Average Treatment Effects on Parental Labor Supply, by Pre-Reform Employment Status

	(1) Employment		(2) Labor Force Participation		(3) Hours Worked		(4) Full-Time	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: All individuals								
Half Benefits	0.027***		0.034***		1.481***		0.006	
	(0.008)		(0.008)		(0.359)		(0.010)	
Post	-0.074	-0.071	-0.060	-0.061	-3.067	-2.759	-0.105	-0.097
	(0.090)	(0.096)	(0.093)	(0.099)	(3.575)	(3.966)	(0.099)	(0.104)
Post × Half Benefits	-0.010		-0.012		0.261		0.014	
	(0.009)		(0.010)		(0.421)		(0.012)	
Quarter Benefits		0.022*		0.028**		1.009		0.003
		(0.013)		(0.013)		(0.625)		(0.017)
Post × Quarter Benefits		0.011		0.007		0.709		0.002
		(0.011)		(0.012)		(0.552)		(0.014)
Mean Dep. Var	0.763	0.724	0.763	0.722	30.125	30.751	0.805	0.836
Observations	11234	11165	11251	11172	11251	11172	8731	8190
Panel B: Self-employed at baseline								
Half Benefits	0.035		0.047		-0.082		0.032	
	(0.033)		(0.037)		(2.138)		(0.049)	
Post	-0.173	-0.306	-0.204	-0.258	-22.481	-32.692	0.354	0.173
	(0.468)	(0.412)	(0.470)	(0.414)	(24.439)	(23.959)	(0.540)	(0.516)
Post × Half Benefits	-0.005		0.007		2.106		0.040	
	(0.046)		(0.055)		(2.727)		(0.060)	
Quarter Benefits		-0.010		-0.009		-0.972		0.029
		(0.030)		(0.032)		(1.684)		(0.059)
Post × Quarter Benefits		0.028		0.029		1.694		-0.032
		(0.045)		(0.048)		(2.215)		(0.059)
Mean Dep. Var	0.712	0.761	0.698	0.763	32.027	35.109	0.830	0.861
Observations	546	882	547	882	547	882	401	653
Panel C: Employed at baseline								
Half Benefits	0.027***		0.034***		1.607***		0.008	
	(0.008)		(0.008)		(0.365)		(0.010)	
Post	-0.083	-0.076	-0.071	-0.072	-3.273	-3.038	-0.128	-0.125
	(0.094)	(0.102)	(0.096)	(0.105)	(3.724)	(4.114)	(0.102)	(0.107)
Post × Half Benefits	-0.012		-0.014		0.023		0.008	
	(0.009)		(0.010)		(0.432)		(0.013)	
Quarter Benefits		0.031**		0.039**		1.827***		0.003
		(0.014)		(0.015)		(0.693)		(0.018)
Post × Quarter Benefits		0.004		-0.003		0.271		0.002
		(0.012)		(0.013)		(0.592)		(0.015)
Mean Dep. Var	0.768	0.716	0.768	0.714	29.975	30.123	0.802	0.836
Observations	10423	9916	10439	9923	10439	9923	8125	7260
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3) on labor supply outcomes introduced in Subsection 3.3.3. For the number of weekly hours worked, we recode missing values as zeros for individuals who are not working. Standard errors, clustered at the individual-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group.

3.5.3 Effects on Childcare Attendance

We now turn to examining the impact of the reform on the distribution of time children spend across different non-parental childcare arrangements. Table 3.9 presents estimates of the intensive-margin effects of the 2015 family allowances reform on weekly hours spent in each type of care among children aged 0 to 5 — a group for whom childcare decisions are especially sensitive to changes in household resources and service affordability.[12] Estimates for older children (aged 6 to 12) are provided in Appendix Table C.7, but are not the focus of this section. The table is structured by age group, allowing us to assess both average effects across early childhood (Panel A) and heterogeneity in behavioral responses across narrower age bands (Panels B to D).

Panel A provides a benchmark by reporting average effects for the full sample of children aged 0 to 5. We observe small but suggestive signs of a reallocation of formal care hours in response to the reform. Specifically, children in families receiving a quarter of the previous benefit amount spend 5.2 fewer hours per week in daycare after the reform (significant at the 10% level), while time in childminder care also declines by 2.2 hours, although this estimate is not statistically significant. These estimates, while modest, point to a reduction in formal care use under reduced benefit generosity.

Panel B focuses on children aged 0 to 2 — an age range in which parental childcare decisions are particularly constrained by financial affordability and access. This group is especially relevant in the context of the 2015 reform, which modified the generosity of family allowances in a way that altered the effective out-of-pocket prices of different childcare modes. Our DiD estimates reveal a marked pattern of substitution between formal care types, particularly between childminders and daycare centers. In the half-benefits group, the reform is associated with an increase of 4.3 hours per week in daycare use, coupled with a symmetric decrease of 4.3 hours in childminder care. A similar, albeit less precisely estimated, pattern is observed in quarter-benefit groups, where the increase in daycare hours (6.7) is offset by a decrease in childminder hours (6.4). This near one-for-one substitution is particularly telling considering the French institutional context: daycare centers (*crèches*) are more heavily subsidized and therefore more affordable to families than licensed childminders (*assistantes maternelles*), even after the reform (Onape, 2024). Reduced benefit levels likely raised households' effective (out-of-pocket) price of both forms of care, but disproportionately so for childminders. Thus, the observed reallocation is consistent with a price-sensitive intensive-margin response: when faced with reduced household resources, families re-optimize by reallocating time toward lower-cost non-parental care options.

For children aged 3 to 5 (Panel C), substitution patterns are less clear. While the post-reform period is associated with a significant drop in daycare hours (-17.8 hours in quarter-benefit regions), this is largely offset by an increase in preschool hours ($+38.6$ hours, significant at the 1% level). This shift likely reflects institutional transitions: many children in this age range become eligible for free pre-school (*école maternelle*), which is publicly provided — although at the time of the reform not

[12] We also present graphical evidence of our results by child's age, estimating Equation (3.4). Results for children aged 0 to 5, 0 to 2, and 3 to 5 can be found respectively in Figures C.4, C.5, C.6.

mandatory until the age of 6 — and thus financially more attractive regardless of changes in cash transfers. The reduction in daycare hours here may therefore not reflect financial pressure *per se*, but rather a natural movement into the education system. Interestingly, there is no significant change in childminder or relative care use, consistent with the idea that preschool absorbs most of the formal care demand for this age group.

Finally, among the oldest age group (Panel D), results are largely imprecise and statistically insignificant across all care types. This likely reflects the near-universal enrollment in free preschool by this age, which flattens any potential substitution effects. The muted responses here reinforce the idea that financial incentives embedded in family allowances matter more when alternative public options are not yet fully accessible (as in the 0–2 group) or are just becoming relevant (3–5 group).

Next, Table C.8 in the Appendix reports the extensive-margin treatment effects of the 2015 reform, measuring the probability that a child attends a given type of non-parental childcare. For children aged 0 to 2 (Panel B), we observe that the reform was associated with a broad increase in the use of non-parental care across all types for families in the full-benefits group, as captured by the positive and significant Post coefficients. However, this expansion is significantly muted among families receiving reduced benefits. In particular, children in half-benefits households were 19 percentage points less likely to attend a childminder post-reform relative to those in the full-benefits group ($p < 0.05$). Although not statistically significant, the sign and magnitude of the interaction term for quarter-benefits households are consistent with this pattern.

Consistent with our intensive-margin findings, we also observe suggestive evidence of increased daycare attendance among families with reduced benefits: the estimated effects are $+8.6$ and $+13.4$ percentage points for the half- and quarter-benefits groups, respectively, relative to pre-reform levels. While these estimates are not statistically significant, they are directionally aligned with the reallocation of hours toward daycare documented earlier, and further underscore the role of relative price sensitivity in shaping families' childcare decisions — consistent with prior work in the literature (*e.g.*, Blau and Tekin, 2007).

One concern over regional heterogeneity in childcare provision can arise here, particularly in the French context. Indeed, according to the Onape (2024) report, while the national average was about 60 formal care places per 100 children under age three in 2022, coverage rates vary widely across regions. Some areas in the West and South-West of France offer significantly higher access, while other areas — including parts of the Paris region and the Mediterranean zone — remain under-served. In addition, the availability of childminders is greater in rural and less urbanized areas than in large metropolitan zones. These structural differences across regions could thus lead to biased estimates of the average effects. Thus, to account for persistent regional differences in childcare supply and parental options, we include fixed effects for broad geographical zones, based on the eight-category territorial units provided in the SRCV and used in official French statistics.[13] Our results,

[13]The regional classification includes: (1) Ile-de-France, (2) Paris Bassin (surrounding central and northern regions), (3) North, (4) East, (5) West, (7) South-West, (8) Center-East, and (9) Mediterranean. The first category (Île-de-France) is used as reference.

presented in the Appendix Table C.9, remain robust to the inclusion of these fixed effects. For instance, among children aged 0–2, the substitution away from more expensive childminders toward daycare remains evident and of similar magnitude. Similarly, the effects observed for the 3–5 age group, including a shift from daycare to preschool are preserved. These similarities reinforce the interpretation that our results are not driven by compositional changes in the geography of childcare supply, but rather actually reflect parents' behavioral responses to reduced benefit generosity.

Finally, although this is not directly the purpose of this paper, we further open discussion on the potential unintended consequences of this reform for early childhood development, as outlined in the introduction. Our results suggest that some higher-income families affected by the reform reallocated non-parental childcare time away from professional childminders toward more affordable daycare centers. This price-sensitive substitution effect may have altered the peer composition within daycare settings and, in turn, generated spillover effects on children's early learning environments. Indeed, empirical evidence shows that exposure to peers from more advantaged backgrounds in early education can improve outcomes for children from lower-income families (*e.g.*, Goux and Maurin, 2007; Duflo, Dupas, and Kremer, 2011), but the effects of peer composition are likely to be heterogeneous and to depend on children's relative position within the peer group (*e.g.*, Elsner and Isphording, 2017; Denning, Murphy, and Weinhardt, 2023). In particular, large gaps in achievement or resources may limit the potential benefits of exposure by inducing discouragement or disengagement among relatively lower-ranked children. While we cannot directly estimate these effects, our findings highlight a theoretically relevant channel that may contribute to changes in the distribution of children's skills, and that calls for further empirical investigation.

Table 3.9: Average Treatment Effects on Time Spent in Different Childcare Facilities

	(1) Daycare		(2) Childminder		(3) Relative		(4) Pre-school	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: 0–5 years-old								
Half Benefits	0.360		1.383		-0.379		-0.304	
	(1.241)		(0.841)		(0.435)		(0.865)	
Post	-15.568***	-14.589***	-3.140***	-2.894***	-0.013	-0.140	59.393***	59.097***
	(1.451)	(1.519)	(0.775)	(0.781)	(0.638)	(0.663)	(1.056)	(1.124)
Post × Half Benefits	-1.568		-2.457**		-0.285		0.739	
	(1.510)		(1.019)		(0.570)		(1.020)	
Quarter Benefits		1.707		1.859		0.429		-2.319
		(2.205)		(1.261)		(0.747)		(1.632)
Post × Quarter Benefits		-5.185*		-2.162		-1.084		3.177*
		(2.710)		(1.902)		(0.954)		(1.671)
Mean Dep. Var	9.857	10.409	2.643	4.355	1.860	2.217	15.135	15.508
Observations	1833	1196	1835	1195	1828	1188	1836	1196
Panel B: 0–2 years-old								
Half Benefits	-2.053		1.639		-0.156			
	(2.638)		(1.679)		(1.037)			
Post	32.899***	30.588***	6.301***	7.639***	5.918***	4.561**		
	(3.936)	(4.077)	(2.414)	(2.592)	(2.097)	(1.912)		
Post × Half Benefits	4.305		-4.256*		-1.184			
	(3.427)		(2.530)		(1.788)			
Quarter Benefits		-5.405		3.352		0.289		
		(5.308)		(2.840)		(0.544)		
Post × Quarter Benefits		6.672		-6.434		-0.859		
		(6.473)		(4.564)		(0.909)		
Mean Dep. Var	18.679	19.656	4.185	6.521	2.245	2.586		
Observations	722	453	723	453	718	449		
Panel C: 3–5 years-old								
Half Benefits	-0.045		0.112		-0.504		-0.348	
	(1.309)		(0.976)		(0.535)		(1.636)	
Post	-17.287***	-17.807***	-4.739***	-3.293***	-1.311	-1.322	38.370***	38.558***
	(1.947)	(2.124)	(1.294)	(1.260)	(1.174)	(1.252)	(2.745)	(2.937)
Post × Half Benefits	-0.239		-0.867		0.524		1.517	
	(1.575)		(1.113)		(0.557)		(1.847)	
Quarter Benefits		2.276		2.327*		-0.480		-3.589
		(2.510)		(1.264)		(1.267)		(3.285)
Post × Quarter Benefits		-3.909		-1.230		0.351		2.465
		(3.773)		(1.726)		(1.052)		(3.585)
Mean Dep. Var	4.123	4.771	1.639	3.032	1.611	1.993	24.779	24.855
Observations	1111	743	1112	742	1110	739	1113	743
Panel D: 4–5 years-old								
Half Benefits			-0.069		-0.441		-0.553	
			(0.600)		(0.851)		(2.025)	
Post			0.175	1.198	0.942	0.872	1.629	2.248
			(1.573)	(1.594)	(1.681)	(1.882)	(3.817)	(4.018)
Post × Half Benefits			0.087		0.207		1.019	
			(0.871)		(0.820)		(2.181)	
Quarter Benefits				0.542		-1.283		-2.284
				(1.259)		(2.798)		(5.093)
Post × Quarter Benefits				-1.107		0.072		4.391
				(1.132)		(1.566)		(3.311)
Mean Dep. Var			0.914	2.731	1.446	1.770	28.028	28.133
Observations			787	516	785	513	788	517
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

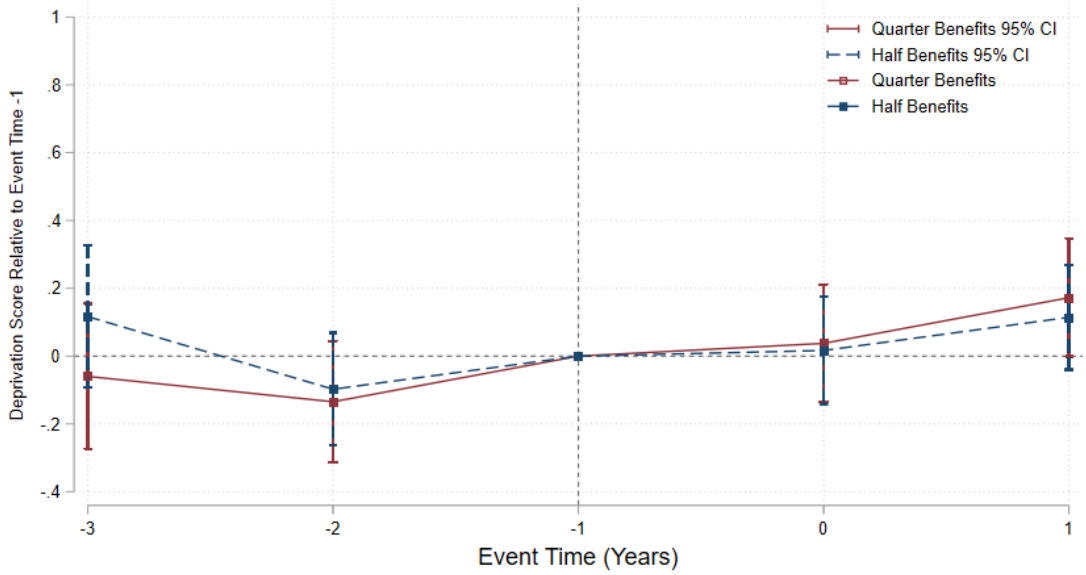
Notes: OLS results for Equation (3.3) on the number of hours per week spent (1) in daycare center, (2) with a professional childminder, (3) cared by a relative, and (4) in pre-school. Standard errors, clustered at the child-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status.

3.6 Robustness

3.6.1 Placebo Tests

We first run a placebo test by changing the policy date, and imply it to be set two years prior to the actual date (2013 instead of 2015), as outlined in Section 3.4. By virtually pretending the reform happened earlier than it actually did, we check whether our main results on the deprivation score captures any pre-existing differences in trends before the policy occurred. Our main results, presented in Figure 3.3 below, indicate no significant effect in this score before the actual policy change, further reinforcing the assumption that differences between groups can be attributed to the reform itself, rather than pre-existing trends.

Figure 3.3: Placebo Test in Policy Timing



Notes: This figure presents a placebo test in time, where the policy date is virtually set to be 2013 (instead of 2015). We then estimate Equation (3.2) on the deprivation score (number of items the household is unable to afford), and present our estimated coefficients here.

We further run a placebo test in treatment groups to control for changes across our income groups that may be unrelated to the policy. In particular, we compare families with one child — not eligible to family allowances by definition — and assign them a "fake" treatment status based on their household position with respect to thresholds introduced by the reform and given in Table 3.1. To implement this exercise, we consider income thresholds for a two-child household — €67,140 and €89,491 respectively for the half- and quarter-benefits groups. We then run our main DiD specification to test for the additional robustness of our main results on the number of weekly hours spent in non-parental childcare, and present our results in Table 3.10 below. As opposed to our main results on the intensive-margin effects (Table 3.9), we find opposite signs and magnitudes here, thus no evidence of sim-

ilar effects, strengthening the credibility of our main results.[14] Indeed, in Panel B and for both groups, even if not significant, at least qualitatively results point to an increase in the number of hours spent with a childminders and a relative decrease in the number of hours spent in daycare. Reassuringly, this suggests that our results are not driven by unrelated income-based variation or coincidental trends near the reform thresholds.

Table 3.10: Placebo Test in Treatment Group

	(1) Daycare		(2) Childminder		(3) Relative		(4) Pre-school	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: 0–5 years-old								
Half Benefits	3.856		-1.730		0.716		1.532	
	(2.683)		(1.598)		(1.136)		(2.747)	
Post	-11.866***	-11.460***	-2.305	-1.888	-1.792	-1.513	87.025***	87.055***
	(3.586)	(3.725)	(2.028)	(2.229)	(2.082)	(2.136)	(3.854)	(4.171)
Post × Half Benefits	-5.514		0.346		0.385		0.090	
	(3.461)		(2.070)		(1.564)		(3.668)	
Quarter Benefits		-2.391		-3.079		-1.021		1.070
		(6.371)		(3.180)		(2.047)		(6.223)
Post × Quarter Benefits		-1.111		2.797		1.659		-2.039
		(6.313)		(2.907)		(1.711)		(5.747)
Mean Dep. Var	18.752	17.324	3.168	5.186	2.958	3.202	14.346	15.276
Observations	521	333	525	333	518	332	306	185
Panel B: 0–2 years-old								
Half Benefits	1.644		-2.826		0.178			
	(3.895)		(2.838)		(1.737)			
Post	48.964***	48.684***	10.177**	10.579**	4.870	5.208		
	(6.296)	(6.607)	(4.912)	(4.996)	(3.381)	(3.568)		
Post × Half Benefits	-2.940		2.333		0.682			
	(5.450)		(4.171)		(3.027)			
Quarter Benefits		0.057		-2.710		-1.289		
		(8.683)		(5.717)		(3.826)		
Post × Quarter Benefits		-0.013		1.648		0.888		
		(7.922)		(5.166)		(3.235)		
Mean Dep. Var	25.412	22.940	4.433	6.801	2.632	2.907		
Observations	330	216	330	216	326	215		
Panel C: 3–5 years-old								
Half Benefits	-3.712		-0.538		-0.488		1.782	
	(6.307)		(3.042)		(2.362)		(4.039)	
Post	-35.302***	-36.895***	-6.240*	-5.284	-7.263*	-7.557*	58.672***	59.756***
	(7.056)	(7.639)	(3.396)	(3.567)	(4.190)	(4.396)	(6.798)	(7.277)
Post × Half Benefits	-0.534		0.484		2.462		0.628	
	(6.253)		(2.415)		(3.145)		(5.402)	
Quarter Benefits		-7.586		-3.338		-1.627		3.291
		(11.106)		(3.078)		(3.678)		(10.776)
Post × Quarter Benefits		7.527		2.268		0.822		-7.949
		(8.043)		(2.431)		(3.614)		(9.422)
Mean Dep. Var	7.246	6.957	1.026	2.205	3.510	3.744	22.513	23.880
Observations	191	117	195	117	192	117	195	117
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3) using only families with one child on the number of hours per week spent (1) in daycare center, (2) with a professional childminder, (3) cared by a relative, and (4) in pre-school. Standard errors, clustered at the child-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status.

[14] Placebo results for children aged 4 and 5 are not reported in Table 3.10 due to insufficient sample sizes within subgroups.

3.6.2 Triple-Difference

To provide additional evidence regarding the robustness of our findings, we implement a triple-difference (DDD) design that exploits the fact that households with only one child were never eligible for family allowances — either before or after the 2015 reform. These one-child families therefore serve as a natural comparison group for two-child (and larger) families who are eligible and thus potentially affected by the reform. In doing so, we rule out potential effects that are driven by different time trends across income levels and that are unrelated to the policy we consider.

We construct the DDD by comparing changes in outcomes over time between households above and below the income thresholds introduced by the reform, separately for families with only one child and those with two or more children. While one-child families were never eligible for family allowances, we assign them to the same income categories used for two-child families under the reform — specifically, below the first threshold (full-benefit group), between the first and second thresholds (half-benefit group), and above the second threshold (quarter-benefit group). This allows us to recover income-specific time trends among families unaffected by the policy, which we then use to difference out potential confounding trends in the treated populations.

Formally, we estimate the difference-in-differences among families with at least two children — comparing those above and below the income thresholds pre- and post-reform — and subtract from it the analogous difference-in-differences among one-child families. This triple-difference identifies the differential effect of the reform on treated families, netting out any general income-based divergence in trends that would have occurred in the absence of the policy, and can be formalized as follows:

$$Y_{ht} = \alpha + \beta Post_t + \gamma D_h + \theta OC_h + \phi(D_h \times Post_t) + \psi(D_h \times OC_h) + \chi(Post_t \times OC_h) + \eta(D_h \times Post_t \times OC_h) + \delta_h + \lambda_t + u_{ht} \tag{3.5}$$

Where α denotes the intercept, $Post_t$ is a dummy denoting the post-reform period, D_h is the treatment indicator at the household-level, and OC_h is a dummy variable set to one if a household is a one-child household and is equal to zero if a household is a two-children household. Additionally, δ_h and λ_t are respectively household (individual, based on the specification), and year fixed effects. Since the reform introduced two treatment intensities, we estimate separate DDD effects for families whose benefits were halved and those whose benefits were reduced to a quarter.

First, we report estimates of this regression for households in Table 3.11 below. Our results confirm our evidence presented in Section 3.5.1, whereby half- and quarter-benefits households experienced a relative increase in material deprivation and financial hardship following the reform. Importantly, these effects do not appear among one-child households, who were never eligible for the allowance: the triple interaction terms are small and statistically insignificant across both outcomes. This strengthens the validity of our identification strategy by confirming that the estimated effects are not driven by income-related trends around the reform thresholds.

Second, we report estimates for the triple-difference specification on children's

time in non-parental care in Table 3.12. Although imprecisely estimated, we can be pretty confident in the validity of our main results presented in Table 3.9. Indeed, among eligible children in the half-benefits group (Panel A), the reform led to a significant reduction in time spent with childminders (-5.2 hours, $p < 0.1$), while no significant change is observed for daycare use. In contrast, one-child households — who were not eligible for the reform — did not exhibit similar patterns, and the triple interaction terms are statistically insignificant — although not different from zero — across both care types. Results for the quarter-benefits group (Panel B) follow a similar pattern, with no significant triple-difference effects. Given that these estimates run counter the direction of our main effects, these estimates reinforce the interpretation that our main findings reflect the causal effect of the reform on non-parental care usage, rather being driven by unrelated income variation across groups or general trends in childcare.

Table 3.11: Triple-Difference for Households

	(1) Deprivation Score	(2) Self-Reported Hardship
Panel A: Half-Benefits		
Post	-0.382***	0.055***
	(0.064)	(0.021)
Treatment	-0.141***	-0.180***
	(0.045)	(0.017)
Post × Treatment	0.139**	0.033
	(0.057)	(0.021)
One-Child Household	-0.035	0.037**
	(0.053)	(0.016)
Post × One-Child Household	-0.040	-0.041**
	(0.057)	(0.018)
Treatment × One-Child Household	0.081	-0.058**
	(0.070)	(0.024)
Post × One-Child Household × Treament	-0.049	0.041
	(0.086)	(0.030)
Panel B: Quarter-Benefits		
Post	-0.366***	0.039*
	(0.066)	(0.021)
Treatment	-0.332***	-0.363***
	(0.069)	(0.026)
Post × Treatment	0.151***	0.044*
	(0.055)	(0.023)
One-Child Household	-0.049	0.036**
	(0.057)	(0.018)
Post × One-Child Household	-0.043	-0.040**
	(0.059)	(0.018)
Treatment × One-Child Household	0.151*	-0.014
	(0.079)	(0.030)
Post × One-Child Household × Treament	-0.022	0.023
	(0.088)	(0.034)
Observations	116189	117395
Household Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes

Notes: OLS results for Equation (3.5) and for households on (1) the deprivation score (number of items a household is unable to afford), and (2) a self-reported measure for financial hardship.

Table 3.12: Triple-Difference for Children

	(1) Daycare	(2) Childminder
Panel A: Half-Benefits		
Post	44.120***	8.561***
	(4.320)	(3.126)
Treatment	-3.297	2.583
	(3.376)	(2.117)
Post × Treatment	3.675	-5.228*
	(4.491)	(3.100)
One-Child Household	5.647**	1.314
	(2.194)	(1.659)
Post × One-Child Household	0.065	0.343
	(2.433)	(1.835)
Treatment × One-Child Household	5.489	-5.465*
	(4.795)	(3.195)
Post × One-Child Household × Treament	-5.425	7.286
	(6.716)	(4.580)
Panel B: Quarter-Benefits		
Post	41.697***	10.766***
	(4.540)	(3.323)
Treatment	-4.363	0.272
	(5.910)	(3.169)
Post × Treatment	-0.611	-1.607
	(8.883)	(4.065)
One-Child Household	4.805**	3.098*
	(2.241)	(1.687)
Post × One-Child Household	1.437	-1.292
	(2.536)	(1.879)
Treatment × One-Child Household	4.608	-1.613
	(8.247)	(5.315)
Post × One-Child Household × Treament	4.088	1.598
	(11.203)	(5.901)
Observations	5668	5670
Individual Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes

Notes: OLS results for Equation (3.5) on the number of hours per week spent by children aged less than two in (1) daycare, and (2) with a professional childminder.

3.7 Conclusion

This paper investigates the consequences of a major policy shift in France's family allowance system, which in 2015 replaced a universal-benefit system with a means-tested scheme for households with two or more children. In particular, we leverage rich household data from the SRCV and exploit the reform's income thresholds (based on household income from two years prior) as a source of exogenous variation in family benefits, in order to understand how families adjust to reductions in child-related cash transfers. Our analysis focuses on three key domains: household material conditions, parental labor supply, and the organization of childcare for young children.

We find that benefit cuts significantly increased material deprivation and self-reported financial hardship among affected families, with no corresponding increase in parental labor supply. Instead, families seemed to have responded by adjusting the type of non-parental childcare used for their children, moving away from costlier options such as childminders and toward more affordable, and often publicly subsidized, alternatives (such as daycare centers). These results suggest that even if benefit cuts reduced available household resources, they did not imply compensatory increases in labor supply, but instead altered the composition of early childhood care.

Our main results speak to a broader literature emphasizing the role of financial resources in shaping children's environments and long-run outcomes. While prior work has focused primarily on the benefits of income support expansions, we add to the disparate evidence on the asymmetric effects of benefit reductions, particularly on multidimensional measures of deprivation. These are crucial in order to understand how such benefits can shape households' living conditions, proxying for children's living environment. The fact that households did not offset the loss in household resources through increased work effort, but instead reduced consumption and restructured childcare, suggests limited substitutability between time and money in our context.

These findings carry important implications for policy design. While the 2015 reform preserved benefits for lower-income families, our results show that cuts targeted at middle- and upper-middle-income households nonetheless led to increased material hardship and reallocation away from more expensive forms of formal childcare. In particular, families reduced use of childminders — who provide more individualized care — and shifted toward daycare centers, which, while more affordable, offer different types of group-based interactions. Although existing evidence suggests that this substitution may have limited effects on children from higher-income families (Havnes and Mogstad, 2011; Berger, Panico, and Solaz, 2021), these behavioral responses may nonetheless affect early learning environments through changes in peer composition, and the resulting exposure to peers from higher and wealthier socioeconomic backgrounds. Nonetheless, a growing body of the literature highlights that these peer composition effects are likely to be heterogeneous, based on children's relative position within the ability distribution (*e.g.*, Elsner and Isphording, 2017; Denning, Murphy, and Weinhardt, 2023). Future research should examine whether such channel influences children's early cognitive and socio-emotional development in the French context, and assess their magnitude and persistence over

time. In addition, these trade-offs should be carefully weighed against the generated fiscal savings when evaluating the reform — as also emphasized by a Cour des Comptes report.[15]

[15]Cour des Comptes, 2017, *Les réformes récentes des prestations familiales et de la fiscalité des familles : des transferts massifs, des questions en suspens*. Available here.

Conclusion

This thesis contributes to the understanding of how gender norms, beliefs, and public policies shape labor market trajectories, household choices, and children's living conditions. Through three distinct chapters, it provides empirical evidence on the mechanisms that underlie persistent inequalities in earnings, labor supply, and households' material conditions, with a focus on the role of expectations and institutional design.

Chapter 1 investigates the role of gender norms in shaping the motherhood penalty in the United Kingdom. Using longitudinal data from the British Household Panel Survey and an event-study approach, the analysis first demonstrates that women with more traditional gender norms before childbirth face significantly larger declines in earnings and working hours after becoming mothers than those with more egalitarian views. Second, the comparative advantage mechanism only seems to apply to egalitarian mothers, while no such pattern emerges for mothers with more traditional views. In other words, when egalitarian women earn more than their male partner before the first childbirth, they do not experience an significant drop in their earnings and hours worked, while traditional women do. Finally, based on the evidence that traditional women work in occupations with fewer hours worked, higher shares of part-time workers and that there is no job mobility after the first childbirth, the analysis further accounts for occupational sorting. Results show that accounting for occupational sorting significantly reduces the average penalty in earnings and hours worked for both egalitarian and traditional women, and explains 80% of the short-run earnings penalty gap between them. Thus, traditional women seem to sort pre-birth into occupations that facilitate a larger reduction in hours worked post-motherhood, which in turn have a substantial impact on their earnings trajectory. Overall, these results highlight the importance of pre-existing norms and occupational trajectories in explaining persistent gender inequalities in the labor market.

Chapter 2 examines whether beliefs about the consequences of maternal labor supply for children's outcomes can be explained by perceptions of a "mom advantage" in child-rearing or, in other words, an absolute advantage. Using a novel online survey experiment conducted with parents in England, this chapter shows that, even when earnings potential is equalized, participants still believe that children perform worse when mothers (rather than fathers) work long hours. These beliefs are shaped by participants' intergenerational experiences — particularly their own mother's employment status during childhood — and predict lower labor supply and larger employment penalties for women. The underlying mechanism appears to rest

on expectations that mothers will spend more time on skill-enhancing activities with children than fathers. However, when shown factual information about how well children perform when mothers work full-time, respondents seem to update their beliefs and express more supportive attitudes toward family policies and more liberal gender norms, suggesting that perceptions are malleable. These results are important

Chapter 3 finally analyzes how changes in family policy affect parental behavior, using the 2015 reform of the French family allowance system. The reform reduced benefits for higher-income families with two or more children through the introduction of two income thresholds based on pre-determined income from two years prior. Specifically, households above the thresholds received either half or a quarter of the original benefit, while those below continued to receive full benefits. Exploiting this quasi-experimental variation induced by the reform, the analysis focuses on outcomes likely to respond to changes in household financial resources, including material hardship, labor supply, and childcare choices. Results of difference-in-differences and event-study designs show that benefit reductions significantly increased material deprivation and self-reported financial hardship, particularly among the most affected households. This loss in household resources did not get compensated by increases in labor supply, including among those self-employed or single-parent households. Instead, reduced benefits shifted non-parental childcare arrangements for young children, away from private and costlier care (*e.g.*, childminders) and toward more subsidized and group-based care, such as daycare centers. These results underscore how moderate income losses can affect household living conditions and childcare decisions.

Bibliography

Agostinelli, Francesco, Emilio Borghesan, and Giuseppe Sorrenti (2021). "Welfare, Workfare and Labor Supply: A Unified Evaluation". In: pre-published.

Aguiar, Mark and Erik Hurst (2007). "Measuring Trends in Leisure: The Allocation of Time over Five Decades". In: *The Quarterly Journal of Economics* 122.3, pp. 969–1006.

Akerlof, George A. and Rachel E. Kranton (2000). "Economics and Identity". In: *The Quarterly Journal of Economics* 115.3, pp. 715–753.

Akerlof, Robert and Luis Rayo (2020). "Narratives and the Economics of the Family". In: CEPR Discussion Paper 3674941. Pre-published.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn (2013). "On the Origins of Gender Roles: Women and the Plough". In: *The Quarterly Journal of Economics* 128.2, pp. 469–530.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). "Childhood Circumstances and Adult Outcomes: Act II". In: *Journal of Economic Literature* 56.4, pp. 1360–1446.

Andresen, Maike E. and Emily Nix (2023). "You Can't Force Me Into Caregiving: Paternity Leave and the Child Penalty". Mimeo.

Andresen, Martin Eckhoff and Emily Nix (2022). "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples". In: *Journal of Labor Economics* 40.4, pp. 971–1004.

Angelov, Nikolay, Per Johansson, and Erica Lindahl (2016). "Parenthood and the Gender Gap in Pay". In: *Journal of Labor Economics* 34.3, pp. 545–579.

Anghel, Brindusa, Sara de la Rica, and Juan Jose Dolado (2011). "The Effect of Public Sector Employment on Women's Labour Market Outcomes". In: SSRN Scholarly Paper 1889991. Pre-published.

Athey, Susan and Guido Imbens (2016). "Recursive Partitioning for Heterogeneous Causal Effects". In: *Proceedings of the National Academy of Sciences* 113.27, pp. 7353–7360.

Athey, Susan and Stefan Wager (2019). "Estimating Treatment Effects with Causal Forests: An Application". In: *Observational Studies* 5.2, pp. 37–51.

Attanasio, Orazio, Sarah Cattan, and Costas Meghir (2022). “Early Childhood Development, Human Capital, and Poverty”. In: *Annual Review of Economics* 14.1, pp. 853–892.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman (2019). “Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes”. In: *American Economic Journal: Applied Economics* 11.3, pp. 338–381.

Averett, Susan L and Yang Wang (2022). “Labor Supply Responses to Cash Transfers: Evidence from the Alaska Permanent Fund”. In: *Labour Economics* 75, p. 102150.

Backhaus, Teresa, Clara Schäper, and Annekatrin Schrenker (2023). “Causal Misperceptions of the Part-Time Pay Gap”. In: *Labour Economics* 83, p. 102396.

Bailey, Martha, Tanya Byker, Elena Patel, and Shanthi Ramnath (2025). “The Long-Run Effects of California’s Paid Family Leave Act on Women’s Careers and Childbearing: New Evidence from a Regression Discontinuity Design and US Tax Data”. In: *American Economic Journal: Economic Policy* 17.1, pp. 401–431.

Baker, Michael, Jonathan Gruber, and Kevin Milligan (2008). “Universal Child Care, Maternal Labor Supply, and Family Well-Being”. In: *Journal of Political Economy* 116.4, pp. 709–745.

Barrère-Maurisson, Marie-Agnès and Séverine Lemière (2006). “Entre statut professionnel et politique familiale : l’emploi des assistantes maternelles en France”. In: *Enfances Familles Générations. Revue interdisciplinaire sur la famille contemporaine* (4).

Becker, Gary S. (1985). “Human Capital, Effort, and the Sexual Division of Labor”. In: *Journal of Labor Economics* 3.1, S33–S58.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014). “Inference on Treatment Effects after Selection among High-Dimensional Controls”. In: *The Review of Economic Studies* 81.2, pp. 608–650.

Bénabou, Roland and Jean Tirole (2002). “Self-Confidence and Personal Motivation”. In: *The Quarterly Journal of Economics* 117.3, pp. 871–915.

— (2016). “Mindful Economics: The Production, Consumption, and Value of Beliefs”. In: *Journal of Economic Perspectives* 30.3, pp. 141–164.

Berger, Lawrence M., Lidia Panico, and Anne Solaz (2021). “The Impact of Center-Based Childcare Attendance on Early Child Development: Evidence From the French Elfe Cohort”. In: *Demography* 58.2, pp. 419–450.

Bertrand, Marianne (2011). “Chapter 17 – New Perspectives on Gender”. In: *Handbook of Labor Economics*. Vol. 4. Elsevier, pp. 1543–1590.

— (2020). “Gender in the Twenty-First Century”. In: *AEA Papers and Proceedings*. Vol. 110. American Economic Association, pp. 1–24.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz (2010). "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors". In: *American Economic Journal: Applied Economics* 2.3, pp. 228–255.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015). "Gender Identity and Relative Income within Households". In: *The Quarterly Journal of Economics* 130.2, pp. 571–614.

Bertrand, Marianne and Jessica Pan (2013). "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior". In: *American Economic Journal: Applied Economics* 5.1, pp. 32–64.

Blau, David and Erdal Tekin (2007). "The Determinants and Consequences of Child Caree Subsidies for Single Mothers in the USA". In: *Journal of Population Economics* 20.4, pp. 719–741.

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations". In: *Journal of Economic Literature* 55.3, pp. 789–865.

Blesch, Maximilian, Philipp Eisenhauer, Peter Haan, Boryana Ilieva, Annekatrin Schrenker, and Georg Weizsäcker (2023). *Biased Wage Expectations and Female Labor Supply*. Tech. rep.

Bobonis, Gustavo J. (2009). "Is the Allocation of Resources Within the Household Efficient? New Evidence from a Randomized Experiment". In: *Journal of Political Economy* 117.3, pp. 453–503.

Boelmann, Barbara, Anna Raute, and Uta Schönberg (2025). "Wind of Change? Cultural Determinants of Maternal Labor Supply". In: *American Economic Journal: Applied Economics* 17.2, pp. 41–74.

Boneva, Teodora, Marta Golin, Katja Kaufmann, and Christopher Rauh (2022). "Beliefs About Maternal Labor Supply". In: 4309199. Pre-published.

Boneva, Teodora and Christopher Rauh (2018). "Parental Beliefs about Returns to Educational Investments—The Later the Better?" In: *Journal of the European Economic Association* 16.6, pp. 1669–1711.

Borra, Cristina and Almudena Sevilla (2019). "Competition for University Places and Parental Time Investments: Evidence from the United Kingdom". In: *Economic Inquiry* 57.3, pp. 1460–1479.

Bullock, John G, Alan S Gerber, Seth J Hill, and Gregory A Huber (2015). "Partisan Bias in Factual Beliefs about Politics". In: *Quarterly Journal of Political Science* 10, pp. 519–578.

Bursztyn, Leonardo, Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott (2023). *How Are Gender Norms Perceived?* Tech. rep. National Bureau of Economic Research.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". In: *American Economic Review* 110.10, pp. 2997–3029.

Bursztyn, Leonardo and Robert Jensen (2015). "How Does Peer Pressure Affect Educational Investments?" In: *The Quarterly Journal of Economics* 130.3, pp. 1329–1367.

Callaway, Brantly and Pedro H. C. Sant'Anna (2021). "Difference-in-Differences with Multiple Time Periods". In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 200–230.

Carbuccia, Laudine (2024). "Demand-Side Determinants of the Socioeconomic Gap in ECEC Enrollment in France". report. Laboratoire Interdisciplinaire d'Evaluation des Politiques Publiques ; Caisse Nationale d'Allocations Familiales, 91 pages.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women". In: *The Quarterly Journal of Economics* 131.2, pp. 633–686.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West (2013). "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation". In: *Econometrica* 81.3, pp. 855–882.

Cascio, Elizabeth U, Schanzenbach, and Diane Whitmore (2013). "The Impacts of Expanding Access to High-Quality Preschool Education". In: Working Paper Series.

Chassamboulli, Andri and Pedro Gomes (2023). "Public-Sector Employment, Wages and Education Decisions". In: *Labour Economics* 82, p. 102345.

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

Chen, Jiafeng and Jonathan Roth (2023). "Logs with Zeros? Some Problems and Solutions". In: *The Quarterly Journal of Economics*, qjad054.

Chetty, Raj, John N Friedman, and Emmanuel Saez (2013). "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings". In: *American Economic Review* 103.7, pp. 2683–2721.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff (2014). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood". In: *American Economic Review* 104.9, pp. 2633–2679.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg (Dec. 2018). "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance". In: *Journal of Political Economy* 126.6, pp. 2356–2409.

Cortés, Patricia, Gizem Koşar, Jessica Pan, and Basit Zafar (2024). "Should Mothers Work? How Perceptions of the Social Norm Affect Individual Attitudes Toward Work in the US". In: *Review of Economics and Statistics*, pp. 1–28.

Cortés, Patricia and Jessica Pan (2019). "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap". In: *Journal of Labor Economics* 37.2, pp. 351–398.

— (2023). "Children and the Remaining Gender Gaps in the Labor Market". In: *Journal of Economic Literature* 61.4, pp. 1359–1409.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar (2023). "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab". In: *The Quarterly Journal of Economics* 138.4, pp. 2069–2126.

Costa Dias, Monica, Robert Joyce, and Francesca Parodi (2020). "The Gender Pay Gap in the UK: Children and Experience in Work". In: *Oxford Review of Economic Policy* 36.4, pp. 855–881.

Costa Dias, Monica, Camille Landais, Alison Andrew, and Oriana Bandiera (2021). *The Careers and Time Use of Mothers and Fathers*. Institute for Fiscal Studies.

Cunha, Flavio and James Heckman (2007). "The Technology of Skill Formation". In: *American Economic Review* 97.2, pp. 31–47.

Currie, Janet and Jonathan Gruber (1996). "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women". In: *Journal of Political Economy* 104.6, pp. 1263–1296.

Dahl, Gordon B. and Lance Lochner (2012). "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit". In: *American Economic Review* 102.5, pp. 1927–1956.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth (2018). "Measuring and Bounding Experimenter Demand". In: *American Economic Review* 108.11, pp. 3266–3302.

Denning, Jeffrey T., Richard Murphy, and Felix Weinhardt (2023). "Class Rank and Long-Run Outcomes". In: *The Review of Economics and Statistics* 105.6, pp. 1426–1441.

Duflo, Esther, Pascaline Dupas, and Michael Kremer (2011). "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya". In: *American Economic Review* 101.5, pp. 1739–1774.

Eissa, Nada and Hilary Williamson Hoynes (2004). "Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit". In: *Journal of Public Economics* 88.9-10, pp. 1931–1958.

Elsner, Benjamin and Ingo E. Isphording (2017). "A Big Fish in a Small Pond: Ability Rank and Human Capital Investment". In: *Journal of Labor Economics* 35.3, pp. 787–828.

Exley, Christine L, Oliver P Hauser, Molly Moore, and John-Henry Pezzuto (2024). "Believed Gender Differences in Social Preferences". In: *The Quarterly Journal of Economics*, qjae030.

Farré, Lídia, Libertad González, Claudia Hupkau, and Jenifer Ruiz-Valenzuela (2024). "Paternity Leave and Child Development. Discussion Paper No. 2024." In: *Centre for Economic Performance*.

Farré, Lídia and Francis Vella (2013). "The Intergenerational Transmission of Gender Role Attitudes and Its Implications for Female Labour Force Participation". In: *Economica* 80.318, pp. 219–247.

Fernández, Raquel and Alessandra Fogli (2009). "Culture: An Empirical Investigation of Beliefs, Work, and Fertility". In: *American Economic Journal: Macroeconomics* 1.1, pp. 146–177.

Flèche, Sarah, Anthony Lepinteur, and Nattavudh Powdthavee (2020). "Gender Norms, Fairness and Relative Working Hours within Households". In: *Labour Economics* 65, p. 101866.

Fongoni, Marco, Jonathan Norris, Agnese Romiti, and Zhan Shi (2022). *Reference Dependent Aspirations and Peer Effects in Education*. IZA - Institute of Labor Economics.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro (2019). "Pre-Event Trends in the Panel Event-Study Design". In: *American Economic Review* 109.9, pp. 3307–3338.

Gillen, Ben, Erik Snowberg, and Leeat Yariv (2019). "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study". In: *Journal of Political Economy* 127.4, pp. 1826–1863.

Glogowsky, Ulrich, Emanuel Hansen, Dominik Sachs, and Holger Lüthen (2025). "The Evolution of Child-Related Gender Inequality in Germany and the Role of Family Policies, 1960–2018". In: *European Economic Review* 175, p. 105018.

Goldin, Claudia (2006). "The Quiet Revolution That Transformed Women's Employment, Education, and Family". In: Working Paper Series 11953. Pre-published.

— (2014). "A Grand Gender Convergence: Its Last Chapter". In: *American Economic Review* 104.4, pp. 1091–1119.

González, Libertad, Luis Guirola, and Laura Hospido (2024). *Fathers' Time-Use while on Paternity Leave: Childcare or Leisure?* Tech. rep. Barcelona School of Economics.

Goodman-Bacon, Andrew (2021). "Difference-in-Differences with Variation in Treatment Timing". In: *Journal of Econometrics* 225.2, pp. 254–277.

Goux, Dominique and Eric Maurin (2007). "Close Neighbours Matter: Neighbourhood Effects on Early Performance at School". In: *The Economic Journal* 117.523, pp. 1193–1215.

Graeber, Thomas, Christopher Roth, and Florian Zimmermann (2024). "Stories, Statistics, and Memory". In: *The Quarterly Journal of Economics* 139.4, pp. 2181–2225.

Grinza, Elena, Francesco Devicienti, Mariacristina Rossi, and Davide Vannoni (2022). "How Entry into Parenthood Shapes Gender Role Attitudes: New Evidence from The UK". In: *Feminist Economics* 28.4, pp. 194–220.

Haaland, Ingar and Christopher Roth (2020). "Labor Market Concerns and Support for Immigration". In: *Journal of Public Economics* 191, p. 104256.

— (2023). "Beliefs about Racial Discrimination and Support for Pro-Black Policies". In: *Review of Economics and Statistics* 105.1, pp. 40–53.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (2023). "Designing Information Provision Experiments". In: *Journal of Economic Literature* 61.1, pp. 3–40.

Havnes, Tarjei and Magne Mogstad (2011). "Money for Nothing? Universal Child Care and Maternal Employment". In: *Journal of Public Economics*. Special Issue: International Seminar for Public Economics on Normative Tax Theory 95.11, pp. 1455–1465.

Heckman, James, Rodrigo Pinto, and Peter Savelyev (2013). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes". In: *American Economic Review* 103.6, pp. 2052–2086.

Hoddinott, John and Lawrence Haddad (1995). "Does Female Income Share Influence Household Expenditures? Evidence from Côte d'Ivoire". In: *Oxford Bulletin of Economics and Statistics* 57.1, pp. 77–96.

Hotz, V Joseph, Per Johansson, and Arizo Karimi (2018). "Parenthood, Family Friendly Workplaces, and the Gender Gaps in Early Work Careers". In.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond (2016). "Long-Run Impacts of Childhood Access to the Safety Net". In: *American Economic Review* 106.4, pp. 903–934.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2019). "The Allocation of Talent and U.S. Economic Growth". In: *Econometrica* 87.5, pp. 1439–1474.

Kalil, Ariel (2015). "Inequality Begins at Home: The Role of Parenting in the Diverging Destinies of Rich and Poor Children". In: *Families in an Era of Increasing Inequality: Diverging Destinies*. Ed. by Paul R. Amato, Alan Booth, Susan M. McHale, and Jennifer Van Hook. Cham: Springer International Publishing, pp. 63–82.

Keane, Michael P (2011). "Labor Supply and Taxes: A Survey". In: *Journal of Economic Literature* 49.4, pp. 961–1075.

Kiessling, Lukas (2021). "How do Parents Perceive the Returns to Parenting Styles and Neighborhoods?" In: *European Economic Review* 139, p. 103906.

Kiessling, Lukas and Jonathan Norris (2023). “The Long-Run Effects of Peers on Mental Health”. In: *The Economic Journal* 133.649, pp. 281–322.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante (2024). “The Child Penalty Atlas”. In: *Review of Economic Studies*, rdae104.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller (2019). “Child Penalties across Countries: Evidence and Explanations”. In: *AEA Papers and Proceedings* 109, pp. 122–126.

— (2024). “Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation”. In: *American Economic Journal: Economic Policy* 16.2, pp. 110–149.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (2019). “Children and Gender Inequality: Evidence from Denmark”. In: *American Economic Journal: Applied Economics* 11.4, pp. 181–209.

— (2021). “Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families”. In: *American Economic Review: Insights* 3.2, pp. 183–198.

Klose, Sophie-Charlotte (2020). “Identifying Latent Structures in Maternal Employment: Evidence on the German Parental Benefit Reform”. In: *arXiv preprint arXiv:2011.03541*.

Lassen, Anne Sophie (2021). “Gender Norms and Specialization in Household Production: Evidence from a Danish Parental Leave Reform”. In: Working Paper / Department of Economics. Copenhagen Business School 04-2021. Pre-published.

Lavy, Victor, Olmo Silva, and Felix Weinhardt (2012). “The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools”. In: *Journal of Labor Economics* 30.2, pp. 367–414.

Lei, Ziteng and Shelly Lundberg (2020). “Vulnerable Boys: Short-Term and Long-Term Gender Differences in the Impacts of Adolescent Disadvantage”. In: *Journal of Economic Behavior & Organization* 178, pp. 424–448.

Leturcq, Marion and Lidia Panico (2019). “The Long-Term Effects of Parental Separation on Childhood Multidimensional Deprivation: A Lifecourse Approach”. In: *Social Indicators Research* 144.2, pp. 921–954.

Lundberg, Shelly J, Robert A Pollak, and Terence J Wales (1997). “Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit”. In: *Journal of Human Resources*, pp. 463–480.

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till von Wachter, and Jeffrey B. Wenger (2023). “The Value of Working Conditions in the United States and the Implications for the Structure of Wages”. In: *American Economic Review* 113.7, pp. 2007–47.

Manski, Charles F. (2004). “Measuring Expectations”. In: *Econometrica* 72.5, pp. 1329–1376.

Melentyeva, Valentina and Lukas Riedel (2023). “Child Penalty Estimation and Mothers’ Age at First Birth”. In: ECONtribute Discussion Paper. Pre-published.

Mensinger, Tim and Christian Zimpelmann (2024). “How Gender Role Attitudes Shape Maternal Labor Supply”. In: Discussion Paper Series - CRC TR 224. Pre-published.

Milligan, Kevin and Mark Stabile (2011). “Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions”. In: *American Economic Journal: Economic Policy* 3.3, pp. 175–205.

Mincer, Jacob and Solomon Polachek (1974). “Family Investments in Human Capital: Earnings of Women”. In: *Journal of Political Economy* 82 (2, Part 2), S76–S108.

Moffitt, Robert A. and Matthew V. Zahn (June 2019). “The Marginal Labor Supply Disincentives of Welfare: Evidence from Administrative Barriers to Participation”. In: Working Paper Series 26028. Pre-published.

Moriconi, Simone and Núria Rodríguez-Planas (2021). “Gender Norms and the Motherhood Employment Gap”. In: CESifo Working Paper No. 9471 3982034. Pre-published.

Niederle, Muriel and Lise Vesterlund (2007). “Do Women Shy Away From Competition? Do Men Compete Too Much?” In: *The Quarterly Journal of Economics* 122.3, pp. 1067–1101.

Olivetti, Claudia, Jessica Pan, and Barbara Petrongolo (2024). “The Evolution of Gender in the Labor Market”. In: *Handbook of Labor Economics*. Ed. by Christian Dustmann and Thomas Lemieux. Vol. 5. Elsevier, pp. 619–677.

Onape (2024). *Les Derniers Chiffres de Référence de l’accueil Des Jeunes Enfants*. Tech. rep.

Page, Marianne E. (Sept. 2024). “New Advances on an Old Question: Does Money Matter for Children’s Outcomes?” In: *Journal of Economic Literature* 62.3, pp. 891–947.

Porter, Catherine and Danila Serra (2020). “Gender differences in the choice of major: The importance of female role models”. In: *American Economic Journal: Applied Economics* 12.3, pp. 226–254.

Prior, Markus, Gaurav Sood, Kabir Khanna, et al. (2015). “You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions”. In: *Quarterly Journal of Political Science* 10.4, pp. 489–518.

Rabaté, Simon and Sara Rellstab (2022). “What Determines the Child Penalty in the Netherlands? The Role of Policy and Norms”. In: *De Economist* 170.2, pp. 195–229.

Ramey, Garey and Valerie A Ramey (2009). “The Rug Rat Race”. In: *National Bureau of Economic Research*. Working Paper No. 15284.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe (2023). "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature". In: *Journal of Econometrics* 235.2, pp. 2218–2244.

Sanbonmatsu, Kira (2002). "Gender Stereotypes and Vote Choice". In: *American Journal of Political Science* 46.1, p. 20.

Schober, Pia and Jacqueline Scott (2012). "Maternal Employment and Gender Role Attitudes: Dissonance among British Men and Women in the Transition to Parenthood". In: *Work, Employment and Society* 26.3, pp. 514–530.

Seguino, Stéphanie (2007). "Plus Ça Change? Evidence on Global Trends in Gender Norms and Stereotypes". In: *Feminist Economics* 13.2, pp. 1–28.

Sen, Amartya (1982). *Poverty and Famines: An Essay on Entitlement and Deprivation*. OUP Oxford. 276 pp.

Settele, Sonja (2022). "How Do Beliefs about the Gender Wage Gap Affect the Demand for Public Policy?" In: *American Economic Journal: Economic Policy* 14.2, pp. 475–508.

Siminski, Peter and Rhiannon Yetsenga (2022). "Specialization, Comparative Advantage, and the Sexual Division of Labor". In: *Journal of Labor Economics* 40.4, pp. 851–887.

Sun, Liyang and Sarah Abraham (2021). "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects". In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 175–199.

Townsend, Peter (1979). *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living*. University of California Press. 1220 pp.

UNICEF, IRC (2023). "Innocenti Report Card 18: Child Poverty in the Midst of Wealth". In: *UNICEF Global Office of Research and Foresight*.

Wasserman, Melanie (2022). "Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents". In: *The Review of Economic Studies* 90.3, pp. 1535–1568.

Wiswall, Matthew and Basit Zafar (2017). "Preference for the Workplace, Investment in Human Capital, and Gender". In: *The Quarterly Journal of Economics* 133.1, pp. 457–507.

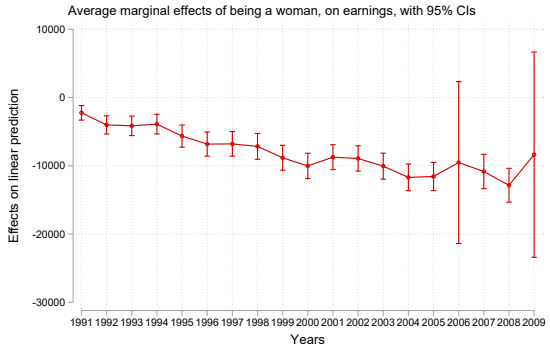
Appendix A

Additional Material for Chapter 1

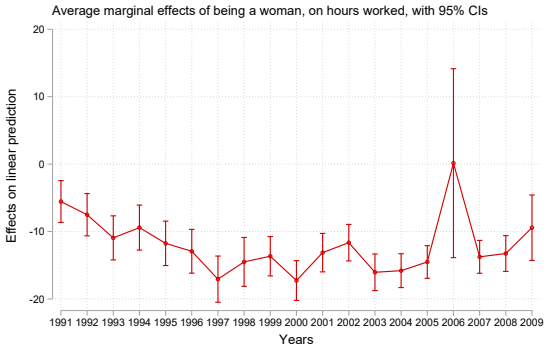
A.1 Additional Descriptive Statistics

Figure A.1: Evolution of Labor Market Outcomes by Gender

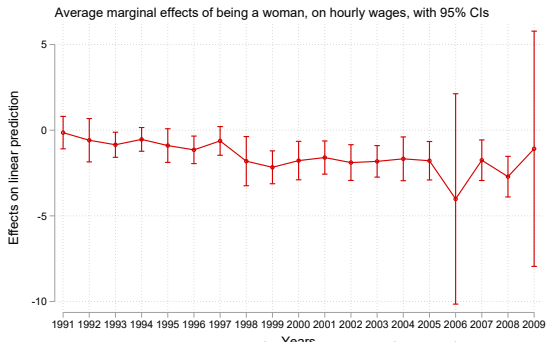
(a) Labor earnings



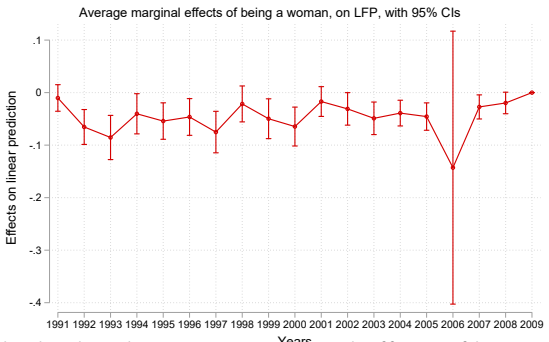
(b) Weekly hours worked



(c) Hourly wages



(d) Labor force participation



Notes: Figures A.1a, A.1b, A.1c, and A.1d respectively display the average marginal effects of being a woman (as opposed to a man) on annual labor earnings, the weekly number of hours worked (excluding overtime), hourly wages, and labor force participation (LFP) for our analytical sample. Specifically, we regress our labor market outcomes on an indicator for being a woman, controlling for year fixed effects. The spikes arising for certain years (*e.g.*, 2006) are explained by the low ($N < 10$) sample sizes for these specific years driven by our sample restrictions.

Table A.1: Sample's Summary Statistics by Gender

	Men	Women	Diff. (Men – Women)	S.E.	N
Age	34.839	33.444	1.395***	0.145	8350
Age at parenthood	32.864	31.899	0.965***	0.095	8350
First child age	3.527	3.344	0.183**	0.093	8350
Second child age	2.162	2.275	-0.113	0.095	5086
Household size	3.054	2.945	0.109***	0.024	8350
Number of children	1.798	1.715	0.083***	0.016	8350
Education					
Primary	0.049	0.028	0.021***	0.004	8350
Low secondary	0.054	0.024	0.031***	0.004	8350
Low secondary/vocational	0.234	0.282	-0.049***	0.010	8350
High secondary/vocational	0.138	0.115	0.023***	0.007	8350
Higher vocational	0.234	0.255	-0.021**	0.009	8350
First degree	0.232	0.253	-0.021**	0.009	8350
Higher degree	0.058	0.043	0.015***	0.005	8350
Marital status					
Never married	0.229	0.259	-0.030***	0.009	8347
Married	0.697	0.668	0.029***	0.010	8347
In a civil partnership	0.000	0.000	-0.000	0.000	8347
Separated	0.019	0.024	-0.005*	0.003	8347
Divorced	0.053	0.047	0.006	0.005	8347
Widowed	0.003	0.001	0.001	0.001	8347

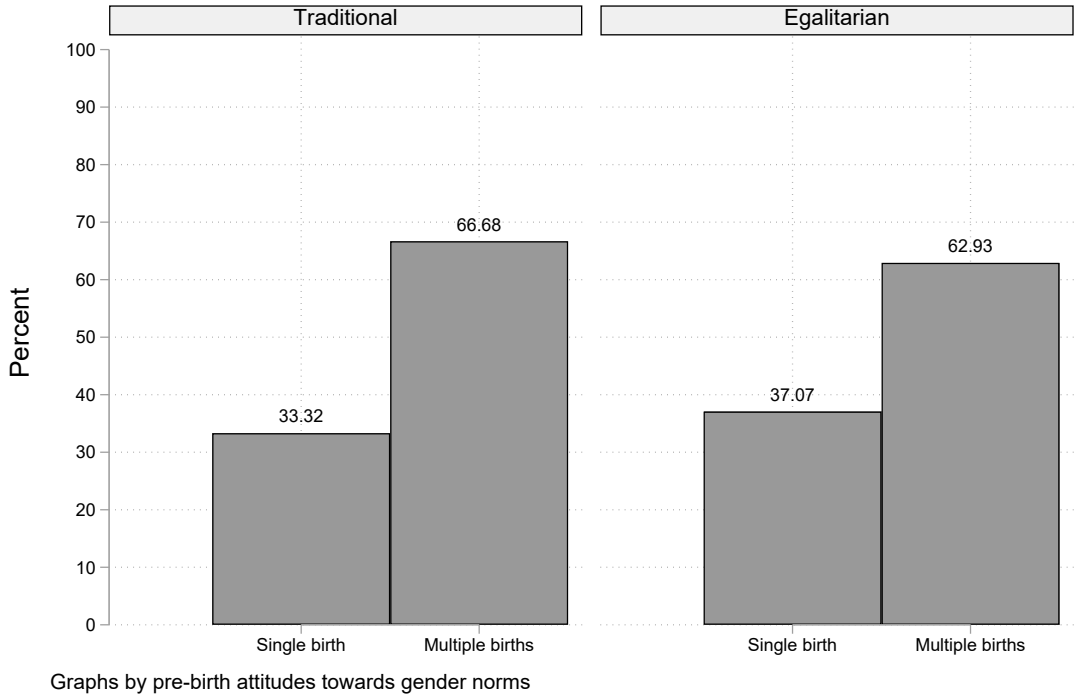
Notes: Summary statistics for our analytical sample, by respondent's gender. * p <0.10; ** p <0.05, *** p <0.01. Levels of educational attainment follow ISCED coding, *i.e.*, Primary — level 1; Low secondary — levels 2A/B; Low secondary/vocational — level 3C; High secondary/vocational — level 3A; Higher vocational — level 5B; First degree — level 5A; Higher degree — level 6.

Table A.2: Sample's Summary Statistics by Pre-Birth Gender Norms

	Traditional	Egalitarian	(Trad. – Egal.)	S.E.	N
Age	34.456	33.961	0.495***	0.147	8281
Age at parenthood	32.660	32.226	0.433***	0.096	8281
First child age	3.461	3.424	0.038	0.093	8281
Second child age	2.322	2.117	0.205**	0.096	5022
Household size	2.994	3.009	-0.015	0.024	8281
Number of children	1.775	1.732	0.043***	0.016	8281
Education					
Primary	0.048	0.031	0.017***	0.004	8281
Low secondary	0.047	0.034	0.013***	0.004	8281
Low secondary/vocational	0.291	0.223	0.068***	0.010	8281
High secondary/vocational	0.128	0.127	0.001	0.007	8281
Higher vocational	0.249	0.241	0.008	0.009	8281
First degree	0.190	0.293	-0.104***	0.009	8281
Higher degree	0.047	0.051	-0.004	0.005	8281
Marital status					
Never married	0.237	0.252	-0.015	0.009	8278
Married	0.688	0.677	0.011	0.010	8278
In a civil partnership	0.000	0.000	-0.000	0.000	8278
Separated	0.025	0.018	0.007**	0.003	8278
Divorced	0.046	0.053	-0.007	0.005	8278
Widowed	0.004	0.000	0.004***	0.001	8278

Notes: Summary statistics for our analytical sample, by respondent's pre-birth gender norms. * p <0.10; ** p <0.05, *** p <0.01. Levels of educational attainment follow ISCED coding, see Table A.1 for a detailed description of categories.

Figure A.2: Single vs. Multiple Births by Pre-Birth Gender Norms



Notes: This figure presents the distribution of single *versus* multiple births by respondent's pre-birth gender norms.

A.2 Additional Results

A.2.1 Pre-Birth Gender Norms

In this subsection, we estimate our event study regressions on our main labor market outcomes, stratified by the indicator reflecting individuals' pre-birth gender norms (defined in Subsection 1.2.4.2). Overall, our findings highlight the substantial influence of individual beliefs and gender norms on the magnitude of the motherhood penalty in earnings (Figure A.3), with distinct patterns emerging following childbirth between traditional and egalitarian mothers. Traditional women experience a higher reduction in labor force participation in the short run compared to their egalitarian counterparts (Figure A.6), as well as a significantly larger decrease in hours worked following the first childbirth (Figure A.4).

Yet, a potential identification issue may arise if the effects observed for traditional women are influenced by different fertility patterns compared to those of egalitarian women. In the Appendix, Figure A.2, we report the breakdown of single *versus* multiple births (*i.e.*, with two children or more) by respondent's pre-first-childbirth gender norms. The distribution of multiple births is similar between egalitarian and traditional parents, indicating that disparities in fertility patterns, based on gender norms, are unlikely to explain the large effects we find.

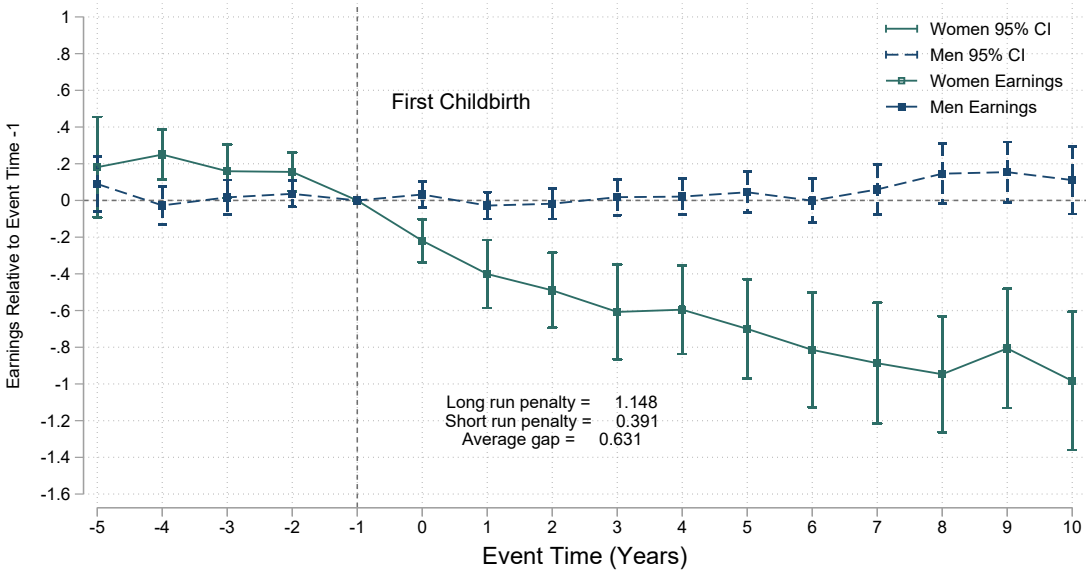
In addition, to support the robustness of these pre-birth gender norms results, we conducted different checks, which we report in the Appendix, Section A.3. First, our results can be interpreted as conservative as we remove outliers by trimming the bottom and top first percentiles of the earnings distribution. We show in Table A.11 that our results are qualitatively similar (i) without trimming the earnings distribution, and (ii) by trimming only the top first percentile to retain all observations with zero labor earnings. Second, we change the definition of traditional and egalitarian by dividing our sample based on the *average* value of the pre-birth gender norms score, rather than using the *median* value, and run our main specification (equation (1.1)) on labor earnings split by this new definition of gender norms. Graphical evidence reported in Figure A.25 in the Appendix indicates that our main result remains robust to this new definition.

Further, we try to disentangle the effect of socioeconomic status (SES) from that of gender norms. While the latter refers to societal expectations and perceptions regarding gender roles, SES encompasses factors such as income, education, or occupation. Given their correlation (0.13, $p < 0.01$), we examine, in Subsection A.3.2.3 in the Appendix, whether the observed results persist when controlling for socioeconomic factors, and show that our findings remain robust. This suggests that our main results are not sensitive to changing the criterion we use to classify individuals based on gender norms, and are not simply driven by differences in SES.

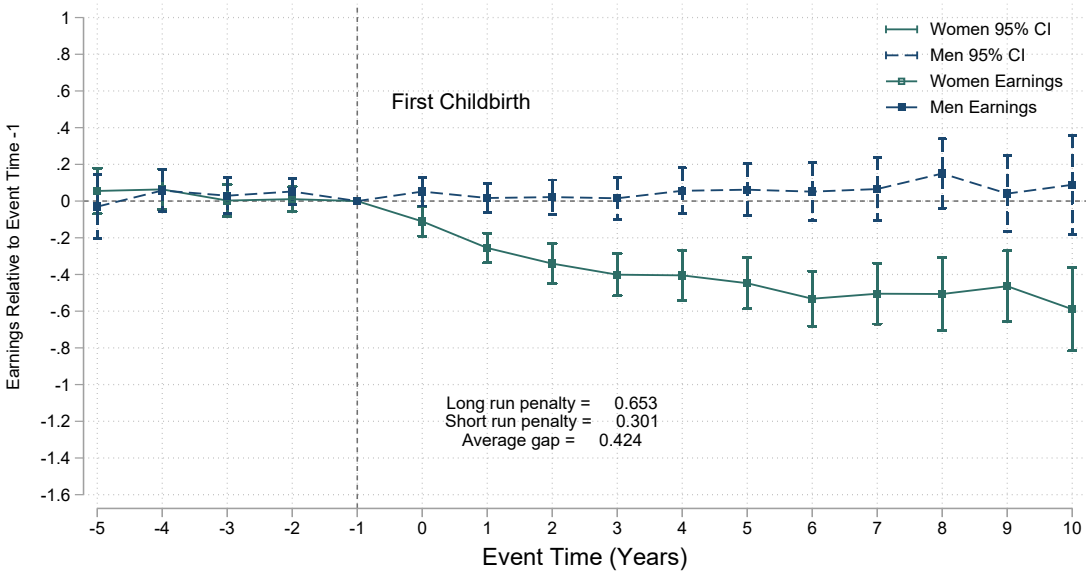
Finally, this reduction in hours worked may come together with additional family commitments. We explore this dimension descriptively in the next subsection, and we see consistent evidence that traditional women report a higher rate of family commitments potentially interfering with career choices than their more egalitarian peers.

Figure A.3: Impact of Parenthood on Earnings by Pre-Birth Gender Norms

(a) Traditional



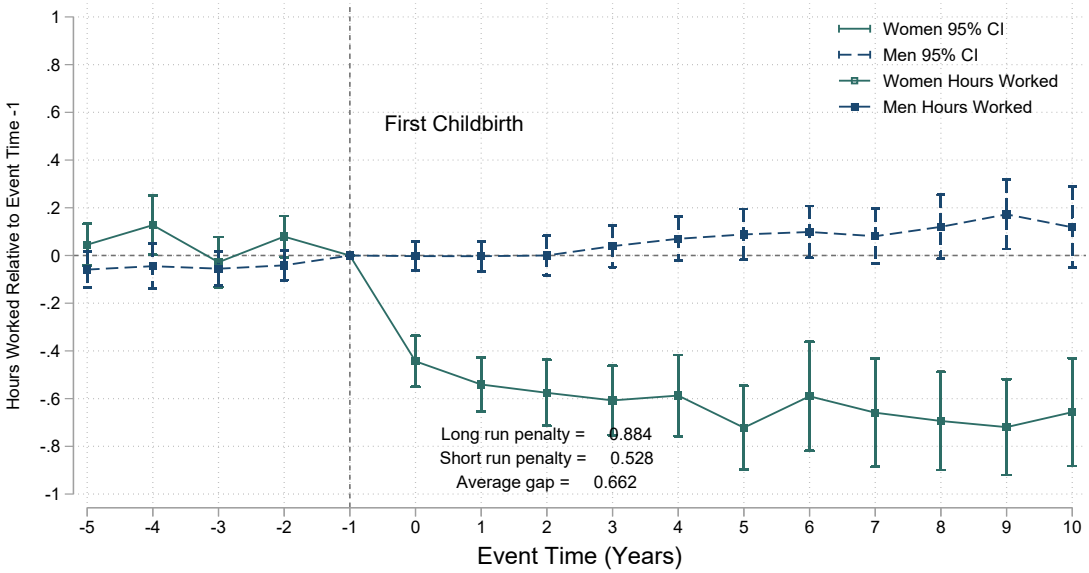
(b) Egalitarian



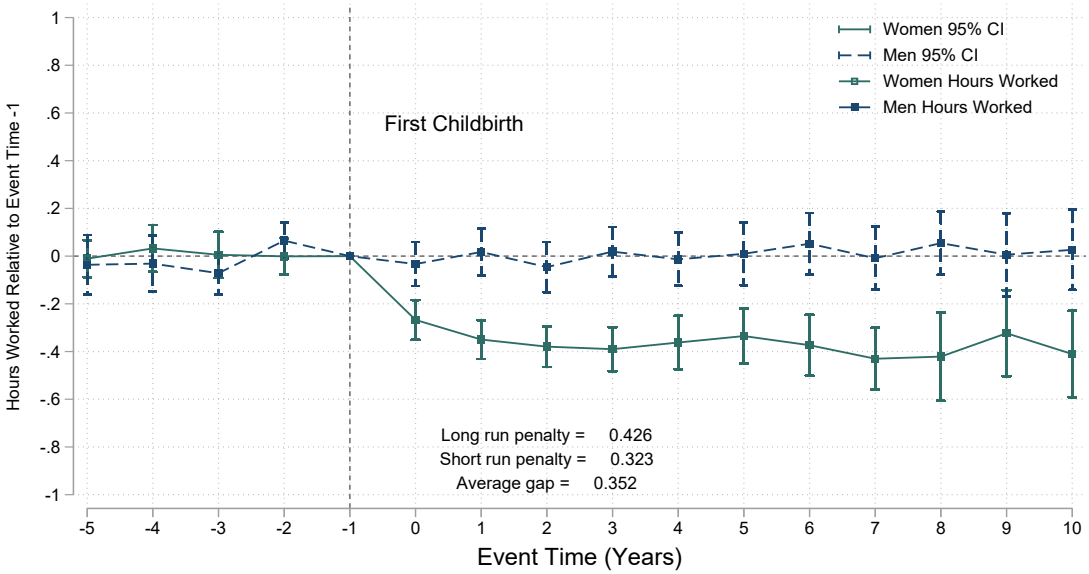
Notes: OLS results for equation (1.1) on annual labor earnings presented alongside the motherhood penalties (short, long and average gaps) for parents with pre-birth (a) traditional and (b) egalitarian norms, as defined in Subsection 1.2.4.2. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. Estimated average effects for egalitarian and traditional women obtained from equation (1.4) are statistically different ($p < 0.0452$)

Figure A.4: Impact of Parenthood on Hours Worked by Pre-Birth Gender Norms

(a) Traditional pre-birth



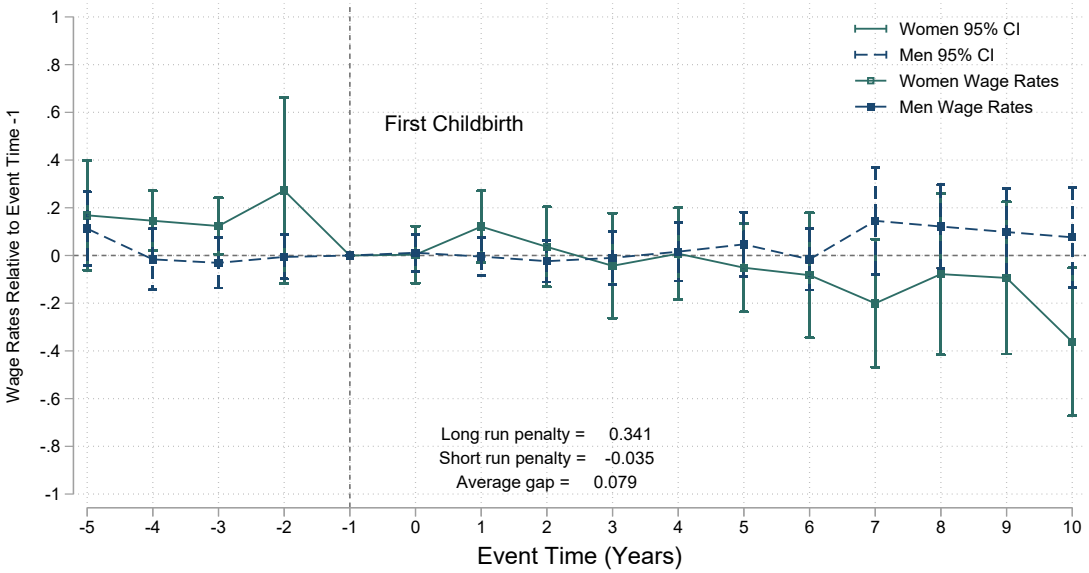
(b) Egalitarian pre-birth



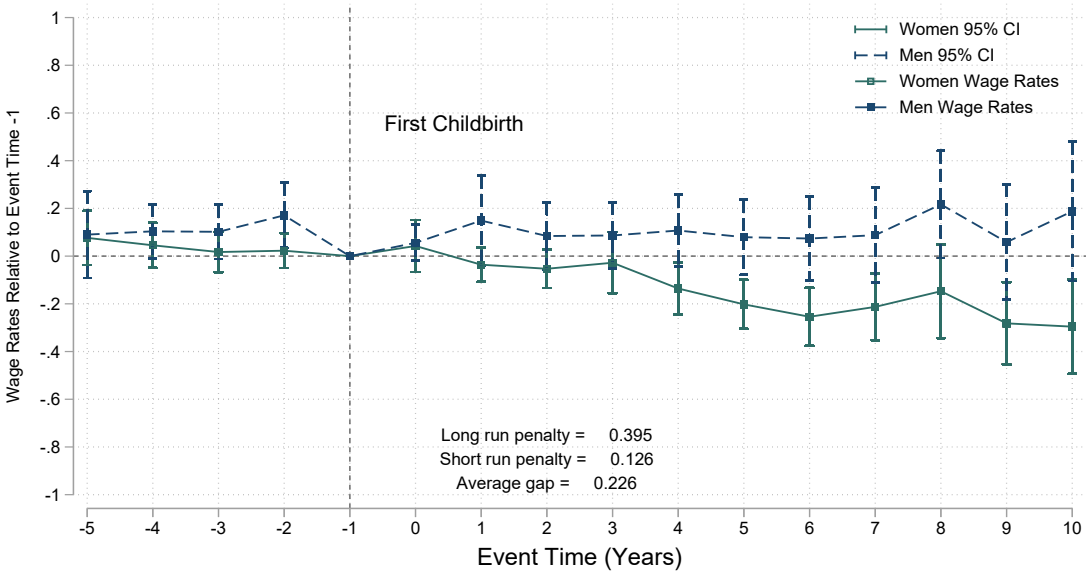
Notes: OLS results for equation (1.1) on our analytical sample defined in Subsection 2.3.1 presented alongside the motherhood penalties (short, long and average gaps) in weekly hours worked for parents with pre-birth (a) traditional and (b) egalitarian norms, as defined in Subsection 1.2.4.2. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. Estimated average effects for egalitarian and traditional women obtained from equation (1.4) are statistically different ($p < 0.0104$).

Figure A.5: Impact of Parenthood on Hourly Wages by Pre-Birth Gender Norms

(a) Traditional pre-birth



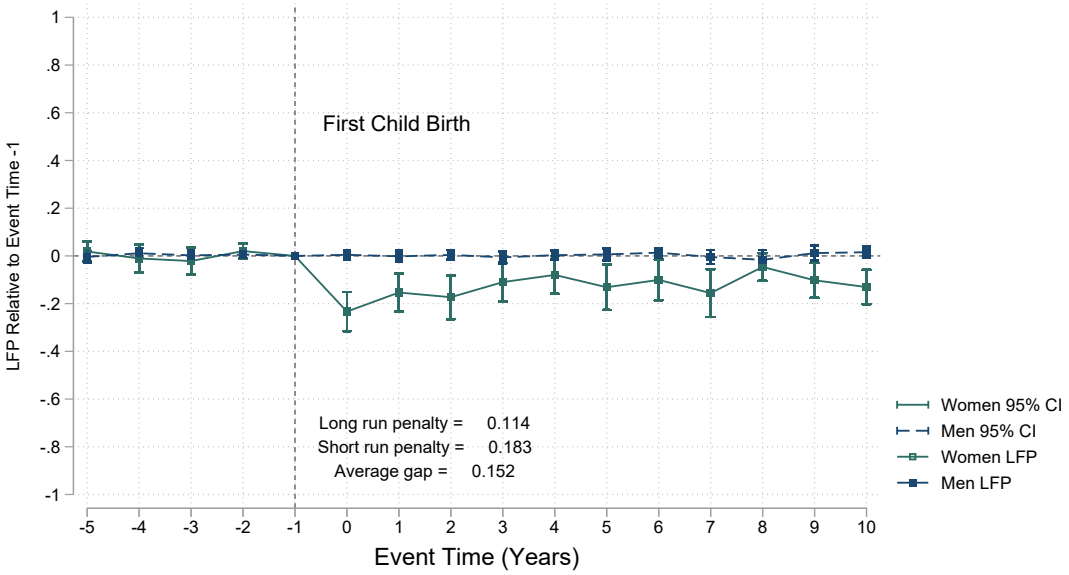
(b) Egalitarian pre-birth



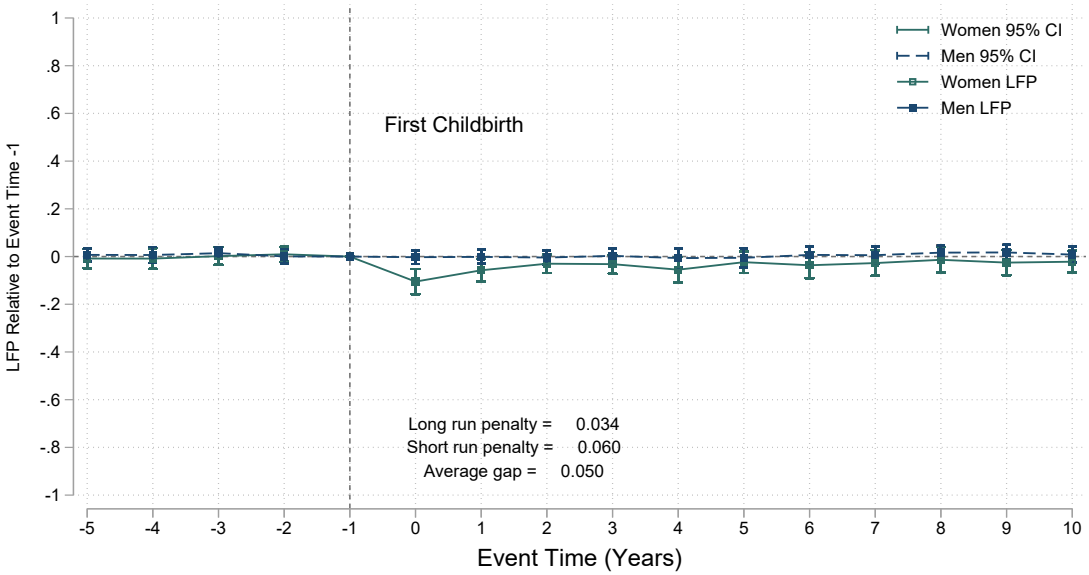
Notes: OLS results for equation (1.1) on our analytical sample defined in Subsection 2.3.1 presented alongside the motherhood penalties (short, long and average gaps) in hourly wages for parents with pre-birth (a) traditional and (b) egalitarian norms, as defined in Subsection 1.2.4.2. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. Estimated average effects for egalitarian and traditional women obtained from equation (1.4) are not statistically different ($p < 0.2569$).

Figure A.6: Impact of Parenthood on Labor Force Participation by Pre-Birth Gender Norms

(a) Traditional pre-birth



(b) Egalitarian pre-birth



Notes: OLS results for equation (1.1) on our analytical sample defined in Subsection 2.3.1 presented alongside the motherhood penalties (short, long and average gaps) in labor force participation (LFP) for parents with pre-birth (a) traditional and (b) egalitarian norms, as defined in Subsection 1.2.4.2. LFP is a dummy set to 1 if the individual is either self-employed, employed, unemployed, or on maternity leave. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. LFP stands for labor force participation. The relatively low gaps are partially due to the condition imposed on parents, for working at at least one year prior to the birth. Estimated average effects for egalitarian and traditional women obtained from equation (1.4) are statistically different ($p < 0.0009$).

A.2.2 The Interplay Between Family Commitments and Employment

We focus here on how family commitments can affect jobs on various aspects, with a gender dimension. We expect women to over-report (compared to men) such situation, and especially if they hold traditional norms. Table A.3 below presents descriptive evidence by gender for such variables provided in BHPS. Unsurprisingly, women tend to report a significantly higher number of family commitments affecting their employment, compared to men. Specifically, nearly 8% of women indicate that family commitments have resulted in reduced working hours. Consequently, various forms of family commitments appear to have a noticeable effect on women's employment, whereas they seem to pose a considerably lesser challenge for men. As a result, these job-related decisions interacted with family commitments can significantly influence women's career trajectories and lead to larger gender inequalities in response to childbirth.

We further anticipate that these decisions vary depending on their gender norms. We provide descriptive evidence by pre-birth gender norms in Table A.4 below, for the full sample (Panel A) and for women only (Panel B). First, we observe a noteworthy distinction between traditional and egalitarian parents in the way family commitments interact with the number of hours worked. This may be attributed to the fact that, on average, egalitarian parents work longer hours than traditional ones, thus having a larger margin to reduce their working hours. Second, shifting our attention to women specifically, we see that family commitments required traditional women to leave jobs significantly more than their egalitarian counterparts, but also prevented job search. Thus, family commitments appear to hinder traditional women from taking up new employment opportunities, implying a lower level of job mobility for them.

Lastly, Table A.5 below explores the relationship between family commitments and women's employment, categorized by the type of occupation held pre-birth defined in Subsection 1.2.4.3. We display below results for the "working hours" index. Our findings show that women in occupations with shorter working hours report significantly more family commitments that interact with their working hours. This relationship may be due to the greater flexibility of shorter-hours occupations, which allows women to adjust their hours as needed, particularly when they have family obligations.

Table A.3: Interaction Between Family Commitments and Job by Gender

	Men	Women	Diff. (Men – Women)	S.E.	N
Any family commitment affected job	0.0504	0.1903	-0.1400***	0.0112	3076
Family commitment prevented job search	0.0174	0.0768	-0.0594***	0.0074	3073
Family commitment prevented taking job	0.0072	0.0399	-0.0327***	0.0053	3070
Family commitment prevented job change	0.0210	0.0456	-0.0245***	0.0064	3070
Family commitment required job change	0.0090	0.0349	-0.0259***	0.0052	3068
Family commitment required leaving job	0.0042	0.0605	-0.0563***	0.0061	3070
Family commitment led to less working hours	0.0156	0.0826	-0.0669***	0.0075	3070

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents the means of variables related to the interaction between family commitments and job, by gender.

Table A.4: Interaction Between Family Commitments and Job by Pre-Birth Gender Norms

	Traditional	Egalitarian	Diff. (Trad. – Egal.)	S.E.	N
Panel A: full sample					
Any family commitment affected job	0.1100	0.1166	-0.0066	0.0115	3058
Family commitment prevented job search	0.0500	0.0364	0.0136*	0.0074	3055
Family commitment prevented taking job	0.0207	0.0240	-0.0034	0.0054	3052
Family commitment prevented job change	0.0301	0.0323	-0.0022	0.0063	3052
Family commitment required job change	0.0188	0.0234	-0.0046	0.0052	3051
Family commitment required leaving job	0.0338	0.0261	0.0077	0.0062	3052
Family commitment led to less working hours	0.0351	0.0570	-0.0219***	0.0075	3052
Panel B: women only					
Any family commitment affected job	0.2088	0.1713	0.0375*	0.0210	1401
Family commitment prevented job search	0.1013	0.0546	0.0467***	0.0142	1399
Family commitment prevented taking job	0.0376	0.0419	-0.0043	0.0106	1398
Family commitment prevented job change	0.0426	0.0445	-0.0019	0.0110	1398
Family commitment required job change	0.0344	0.0356	-0.0012	0.0099	1398
Family commitment required leaving job	0.0835	0.0432	0.0403***	0.0128	1398
Family commitment led to less working hours	0.0687	0.0915	-0.0227	0.0148	1398

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents the means of variables related to the interaction between family commitments and job, for our analytical sample (Panel A) and for women only (Panel B).

Table A.5: Interaction Between Family Commitments and Job by Occupation

	Longer-hours	Shorter-hours	Diff. (Longer – Shorter)	S.E.	N
Any family commitment affected job	0.1770	0.2119	-0.0349	0.0220	1401
Family commitment prevented job search	0.0732	0.0806	-0.0074	0.0149	1399
Family commitment prevented taking job	0.0394	0.0393	0.0001	0.0109	1398
Family commitment prevented job change	0.0339	0.0682	-0.0343***	0.0117	1398
Family commitment required job change	0.0306	0.0435	-0.0128	0.0104	1397
Family commitment required leaving job	0.0656	0.0475	0.0181	0.0133	1398
Family commitment led to less working hours	0.0701	0.1072	-0.0371**	0.0155	1398

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents the means of variables related to the interaction between family commitments and job, for our analytical sample by type of occupation defined in Subsection 1.2.4.3. Specifically, longer-hours occupations refer to occupations above the median of the number of working hours across occupations, while shorter-hours occupations are the ones below the median.

A.2.3 Pre-Birth Comparative Advantage

Table A.6: Impact of Parenthood on Earnings by Pre-Birth Comparative Advantage

	No Trimming		Trimming 45-55 Percentiles		Trimming 30-70 Percentiles	
	(1) Man has C.A.	(2) Woman has C.A.	(3) Man has C.A.	(4) Woman has C.A.	(5) Man has C.A.	(6) Woman has C.A.
Post	2010.218***	526.192	3361.518***	-361.573	3396.829***	39.076
	(741.326)	(1144.011)	(866.206)	(1368.415)	(924.391)	(1863.618)
Woman	-4188.921***	-70.111	-5021.481***	1499.120	-5764.594***	1987.606
	(681.236)	(1023.816)	(762.161)	(1377.497)	(856.253)	(1725.659)
Woman × Post	-8540.921***	-6945.492***	-9483.866***	-5552.221***	-9740.908***	-4532.034**
	(940.650)	(1333.321)	(1057.296)	(1792.665)	(1204.894)	(2227.387)
Observations	5034	3316	4074	2287	3333	1617
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. OLS results for equation (1.4) on annual labor earnings, further split by the pre-birth comparative advantage indicator. We display results for the full distribution of the pre-birth wage difference in columns (1) and (2). In addition, we drop observations between the 45^{th} and 55^{th} percentiles of this distribution in columns (3) and (4), and we drop observations between the third and seventh decile of this distribution in columns (5) and (6). C.A. stands for comparative advantage.

A.2.4 Industry, Occupations, and Gender Norms

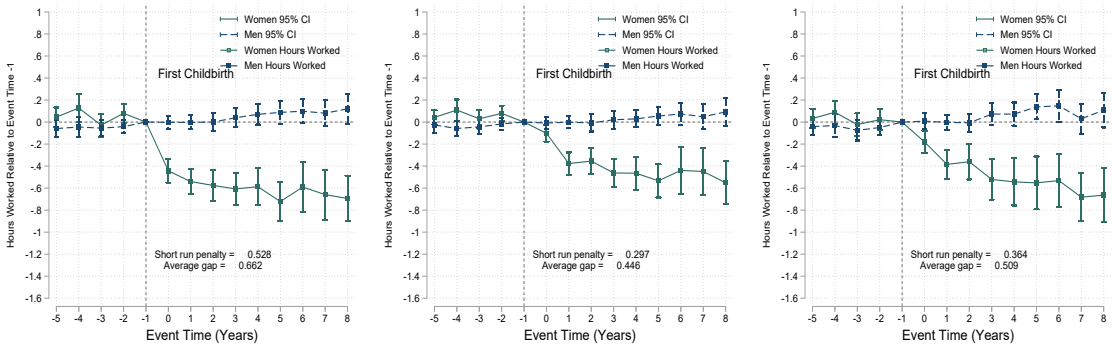
Figure A.7: Impact of Parenthood on Hours Worked (Un)Conditional on Industry and occupations

Traditional Parents

(a) Without fixed effects

(b) With occupation fixed effects

(c) With industry fixed effects

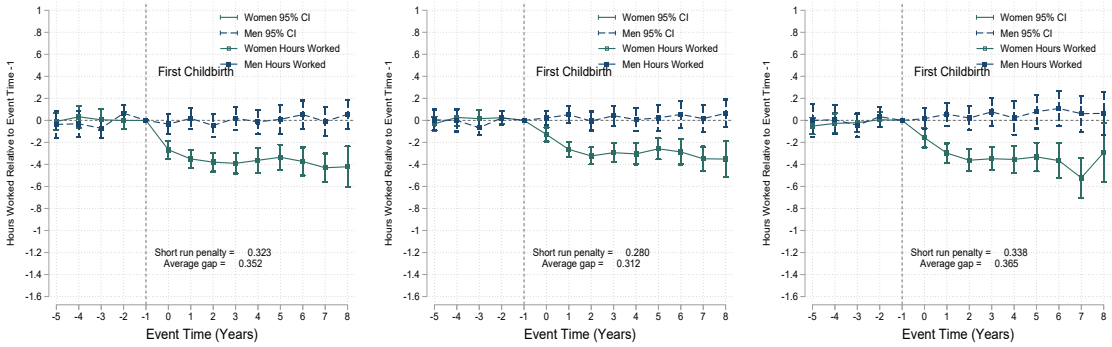


Egalitarian Parents

(d) Without fixed effects

(e) With occupation fixed effects

(f) With industry fixed effects

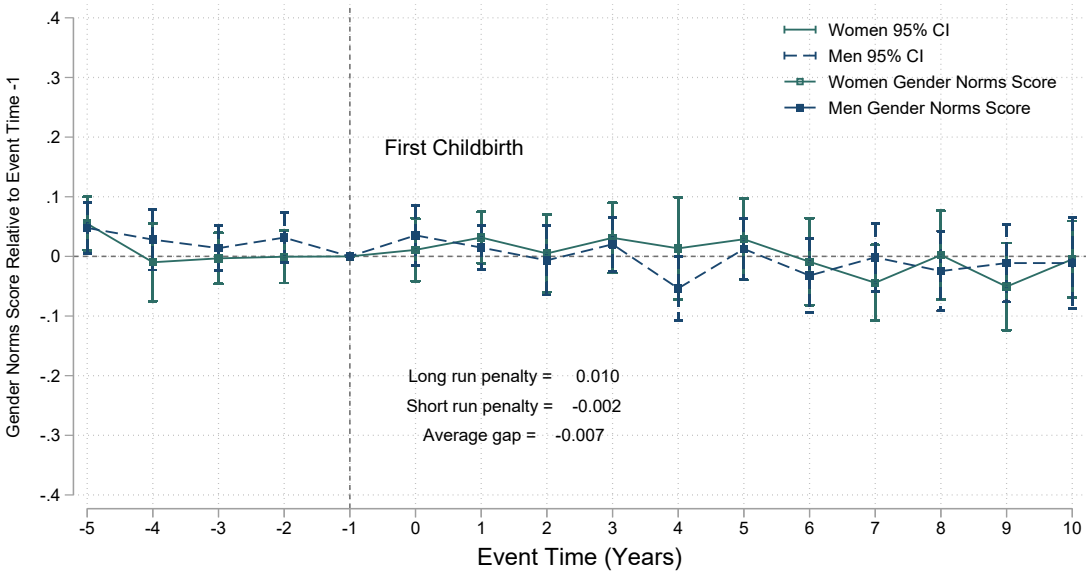


Notes: OLS results for equation (1.1) on the number of weekly hours worked, by pre-birth gender norms (traditional *versus* egalitarian). Results in Figures (a) and (b) are without fixed effects, while Figures (b) and (e) present results controlling for 2-digit occupation fixed effects, and Figure (c) and (f) controlling for 4-digit for industry fixed effects.

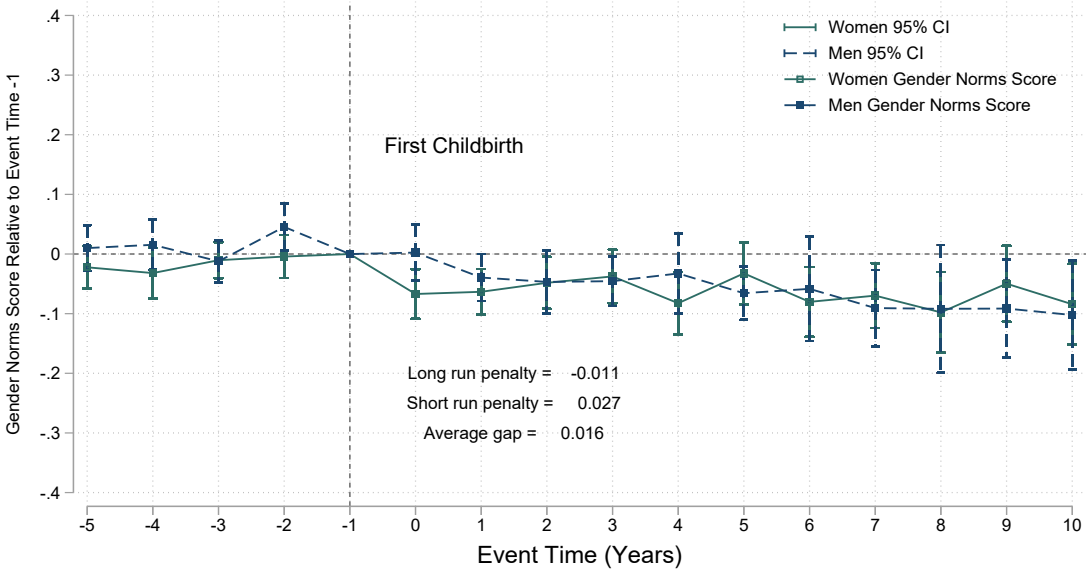
A.2.5 Gender Norms Around Childbirth

Figure A.8: Impact of Parenthood on Gender Norms by Pre-Birth Gender Norms

(a) Traditional pre-birth



(b) Egalitarian pre-birth



Notes: OLS results for equation (1.1) on the gender norms score, for (a) individuals with traditional and (b) egalitarian gender norms pre-birth — as defined in Subsection 1.2.4.2 — to investigate the impact of the first childbirth on gender norms. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

A.2.6 Occupational Characteristics

A.2.6.1 Differences Between Occupations

Below are the results of equation (1.4) on earnings unconditional (column 1) and conditional (column 2) on occupation fixed effects.

Table A.7: Post-Birth Gendered Trajectory of Labor Earnings

	Annual Labor Earnings	
	(1)	(2)
Woman	-2558.248***	-2322.985***
	(584.245)	(542.065)
Post	1361.975**	2147.064***
	(624.944)	(553.801)
Woman × Post	-7989.865***	-7436.672***
	(775.178)	(685.697)
Observations	8350	7980
Year Fixed Effects	Yes	Yes
Occupational Fixed Effects	No	Yes
Age Fixed Effects	Yes	Yes
Normalized Average Effect (%)	-58.61	-54.55

Notes: * p <0.10; ** p <0.05, *** p <0.01. OLS results for equation (1.4) on our analytical sample unconditional (column 1) and conditional on occupation fixed effects. The 'post' dummy takes the value 1 if $t \geq 0$, and corresponds to the post-childbirth period. The normalized average effect corresponds to the estimated average effect ($\hat{\theta}$) divided by the baseline mean *i.e.*, women's pre-birth average labor earnings.

A.2.6.2 Family-Friendly Occupations

We present below our main event study regression on labor earnings, hours worked, and hourly wages, stratified by our three measures reflecting occupations' family-friendliness pre-birth to further investigate if there is any heterogeneity in the motherhood penalty. Overall, our results, based on our two measures reflecting pre-birth occupational characteristics and their family-friendliness, point to no significant effect of such characteristics in shaping the overall magnitude of the motherhood penalty. Our conclusions are in line with Costa Dias, Joyce, and Parodi (2020) and suggest that differences in job characteristics only partially explain the different magnitudes of the overall motherhood penalty.

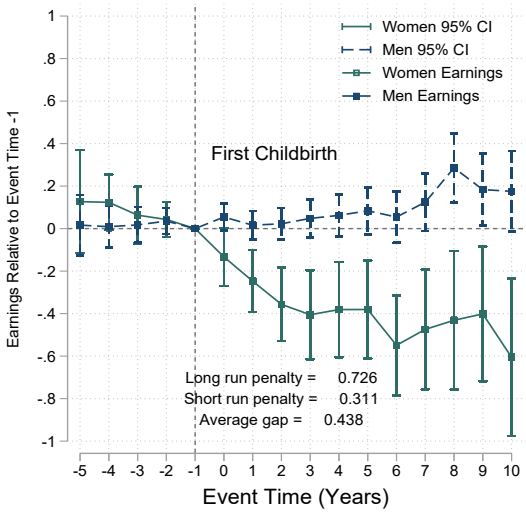
Working hours. We present first our results on labor earnings (Figure A.9), hours worked (Figure A.10), and hourly wages (Figure A.11), stratified by whether respondents worked in (a) shorter- or (b) longer-hours occupations before the first childbirth. Although the penalty in hours worked for shorter-hours occupations is significantly lower than that experienced by women in longer-hours occupations, it seems to have no effect on the penalty in earnings. Indeed, the average effects for

women, based on their pre-birth working hours profile are not statistically different for labor earnings ($p < 0.1410$), and wage rates ($p < 0.2268$), and significant for weekly hours worked ($p < 0.0004$). This suggests that family-friendlier occupations do not contribute to shaping the motherhood penalty in earnings. Yet, these results are only suggestive, as the SUTVA condition, assuming no spillover effects on the control group, appears no longer valid in the long run, leading to an overestimation of the motherhood penalty for these groups.

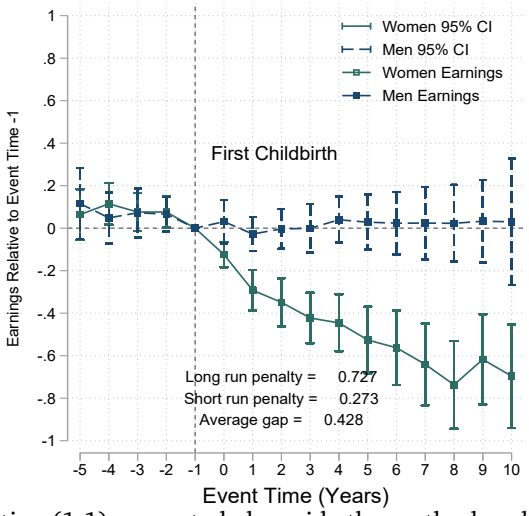
Share of part-time workers. Finally, we investigate whether working, pre-birth, in occupations with different exposures to part-time workers, reflecting more flexible occupations (see Costa Dias, Joyce, and Parodi, 2020), helps understand the magnitude of the motherhood penalty. Results for labor earnings, hours worked, and hourly wages are presented in Figures A.12, A.13, and A.14 below. The average effects, for each subgroup, are not statistically different for earnings ($p < 0.9417$), weekly hours worked ($p < 0.4605$), and hourly wages ($p < 0.1748$), confirming that family-friendly occupations do not contribute to shaping the magnitudes of the motherhood penalty in earnings.

Figure A.9: Impact of Parenthood on Earnings by "Working Hours" Index

(a) Shorter-hours occupations



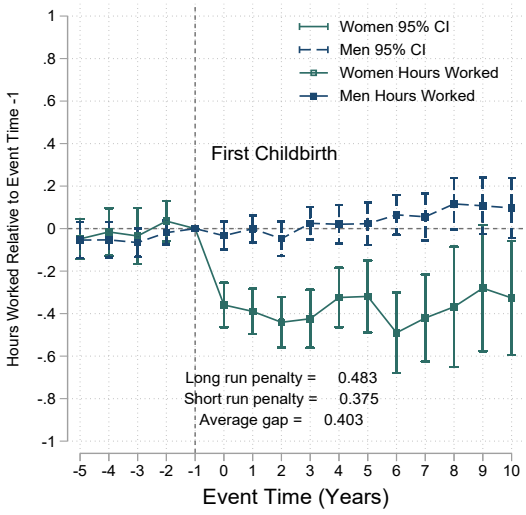
(b) Longer-hours occupations



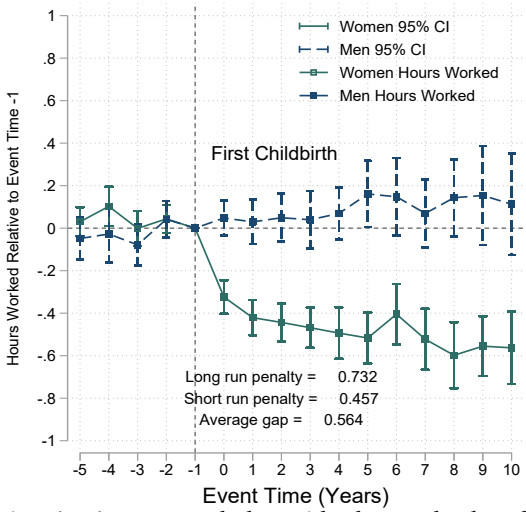
Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in annual labor earnings for parents working pre-birth in (a) shorter-hours and (b) longer-hours occupations, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

Figure A.10: Impact of Parenthood on Hours Worked by "Working Hours" Index

(a) Shorter-hours occupations



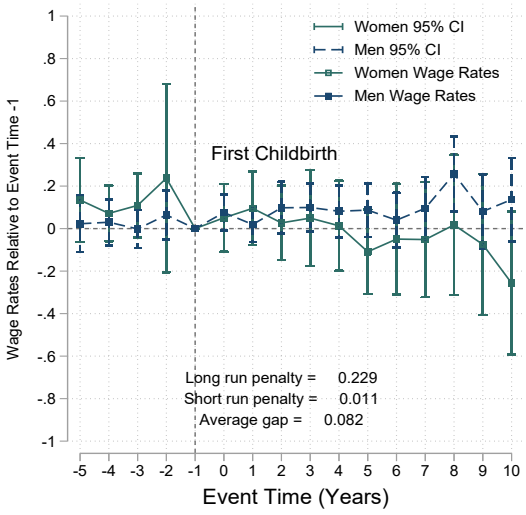
(b) Longer-hours occupations



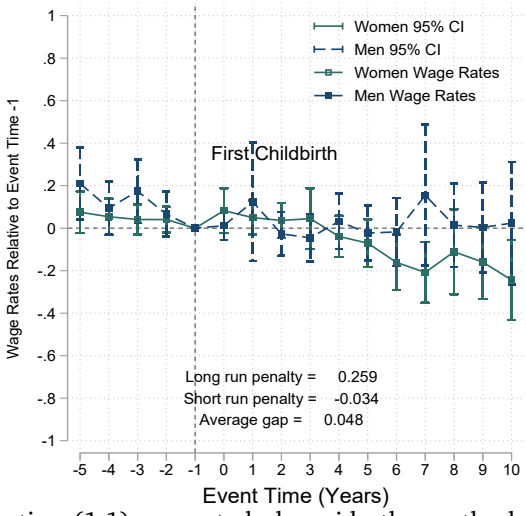
Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in weekly hours worked for parents working pre-birth in (a) shorter-hours and (b) longer-hours occupations, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

Figure A.11: Impact of Parenthood on Hourly Wages by "Working Hours" Index

(a) Shorter-hours occupations



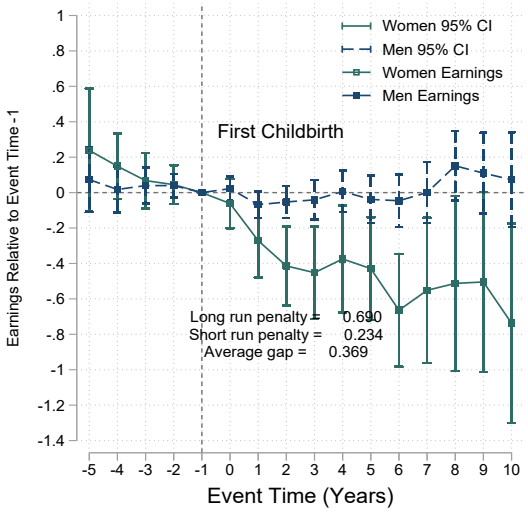
(b) Longer-hours occupations



Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in hourly wages for parents working pre-birth in (a) shorter-hours and (b) longer-hours occupations, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

Figure A.12: Impact of Parenthood on Earnings by "Share of Part-Time Workers" Index

(a) Higher shares of PT workers occupations

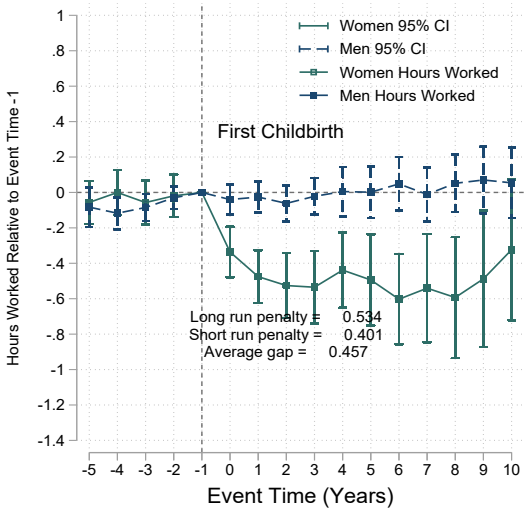


(b) Lower shares of PT workers occupations

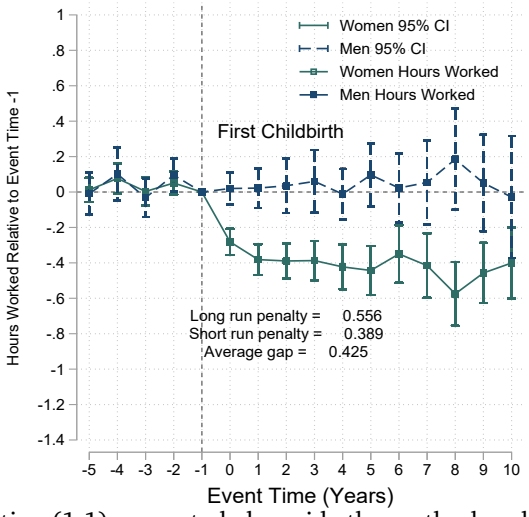
Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in annual labor earnings for parents working pre-birth in occupations with (a) higher shares of part-time workers and (b) lower shares, before childbirth, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. PT stands for part-time.

Figure A.13: Impact of Parenthood on Hours Worked by "Share of Part-Time Workers" Index

(a) Higher shares of PT workers occupations



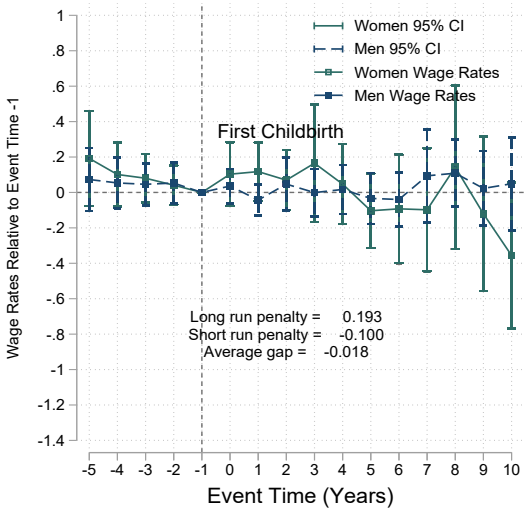
(b) Lower shares of PT workers occupations



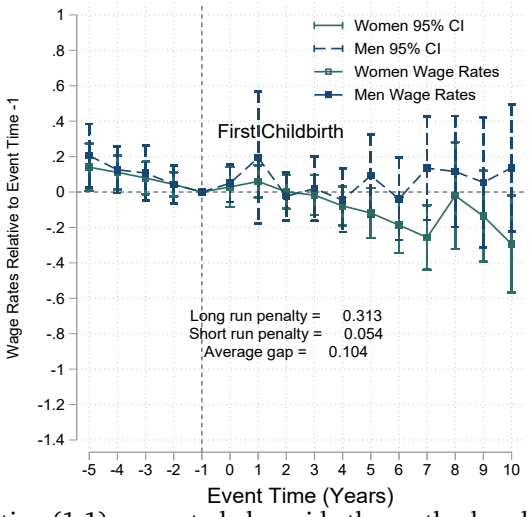
Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in weekly hours worked for parents working pre-birth in occupations with (a) higher shares of part-time workers and (b) lower shares, before childbirth, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. PT stands for part-time.

Figure A.14: Impact of Parenthood on Hourly Wages by "Share of Part-Time Workers" Index

(a) Higher shares of PT workers occupations



(b) Lower shares of PT workers occupations



Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in hourly wages for parents working pre-birth in occupations with (a) higher shares of part-time workers and (b) lower shares, before childbirth, as defined in Subsection 1.2.4.3. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10. PT stands for part-time.

A.2.6.3 Occupational Characteristics and Gender Norms

Table A.8: Motherhood Penalties (Un)Conditional on Occupations by Pre-Birth Gender Norms

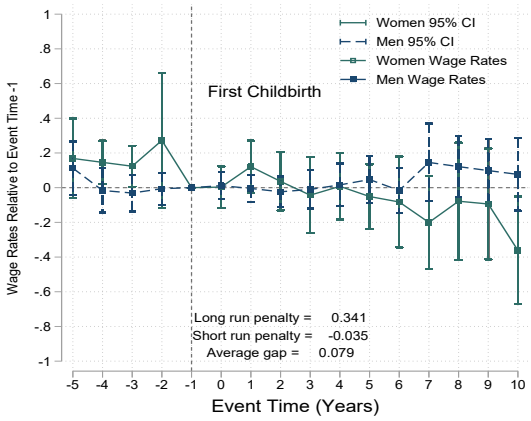
	(1)	(2)	(3)	(4)
	Traditional		Egalitarian	
Average gaps (%)				
Labor earnings	0.61***	0.48***	0.43***	0.31*
	(.12)	(.13)	(.09)	(.09)
Hours worked	0.58***	0.39***	0.38***	0.34***
	(.09)	(.09)	(.08)	(.08)
Hourly wages	0.09	-0.02	0.22***	0.15**
	(.1)	(.1)	(.08)	(.08)
Short-run gaps (%)				
Labor earnings	0.39***	0.24**	0.29***	0.21***
	(.09)	(.09)	(.06)	(.07)
Hours worked	0.48***	0.26***	0.34***	0.29***
	(.07)	(.06)	(.07)	(.06)
Hourly wages	-0.03	-0.09	0.12	0.08
	(.08)	(.08)	(.07)	(.07)
Long-run gaps (%)				
Labor earnings	1.10***	0.88***	0.68***	0.47***
	(.24)	(.22)	(.18)	(.15)
Hours worked	0.78***	0.59***	0.45***	0.44***
	(.16)	(.15)	(.13)	(.12)
Hourly wages	0.34**	0.14	0.40***	0.25*
	(.17)	(.16)	(.15)	(.13)
Occupational FE	✗	✓	✗	✓

Notes: This table displays the motherhood penalties in annual labor earnings, weekly hours worked, and hourly wages by pre-birth gender norms. Specifically, the motherhood penalties presented here are the average, short- and long-run gaps, respectively encompassing the 10-year period following parenthood (average gaps), the 3 years after (short-run gaps), and between 7 and 10 years after parenthood (long-run gaps). Specifications in columns (1) and (3) are unconditional on occupation fixed effects, while specifications in columns (2) and (4) control for 2-digit occupation fixed effects. Standard errors, bootstrapped with 1000 replications, are presented in parentheses. Results are statistically significant from zero in a two-sided test at * 10%, ** 5%, and *** 1%; NS otherwise.

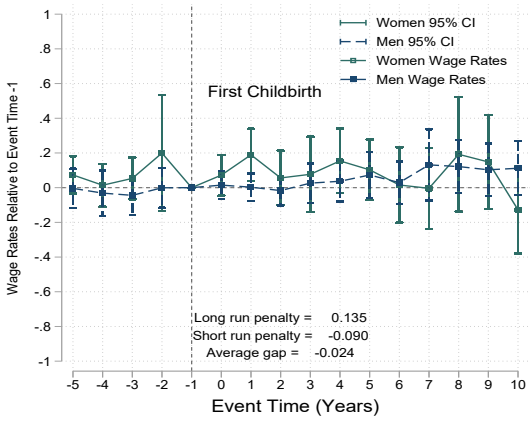
Figure A.15: Impact of Parenthood on Hourly Wages by Pre-Birth Gender Norms (Un)Conditional on Occupations

Traditional

(a) Without occupation fixed effects

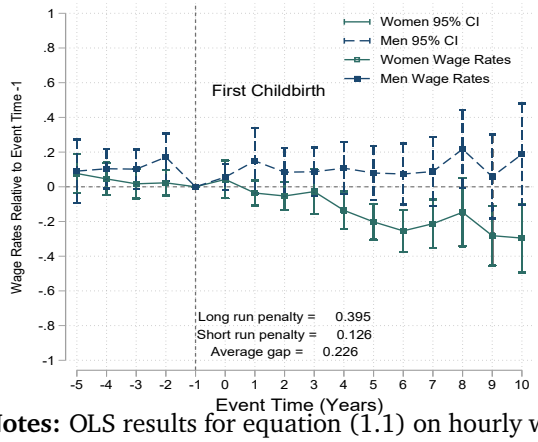


(b) With occupation fixed effects

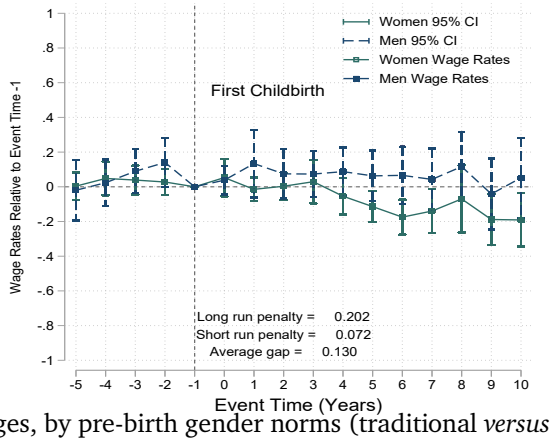


Egalitarian

(c) Without occupation fixed effects

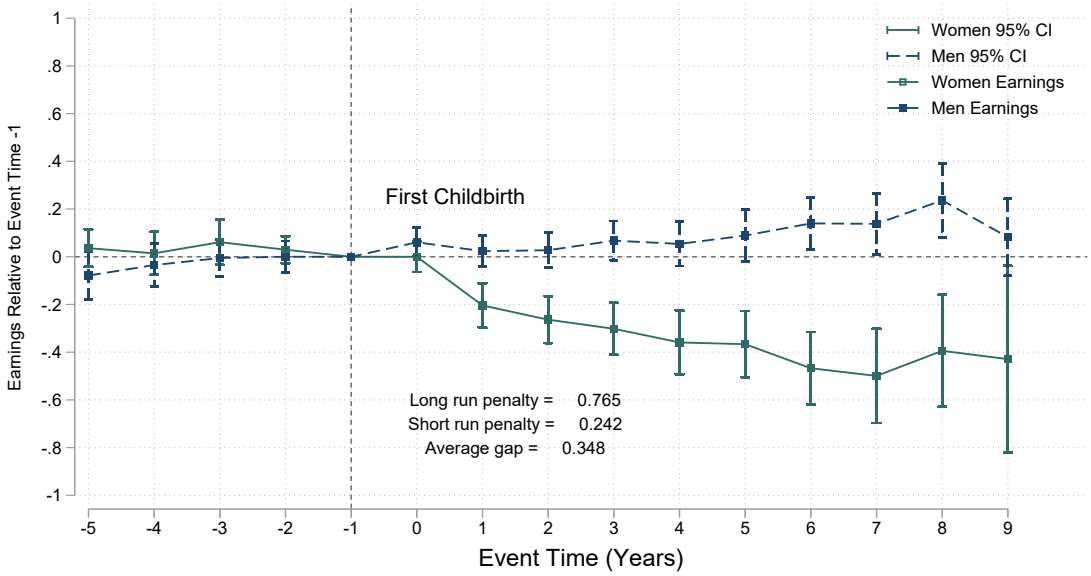


(d) With occupation fixed effects



Notes: OLS results for equation (1.1) on hourly wages, by pre-birth gender norms (traditional *versus* egalitarian). Figures (a) and (c) on the left hand side display results unconditional on occupation fixed effects, while Figures (b) and (d) on the right hand side display results controlling for 2-digit occupation fixed effects.

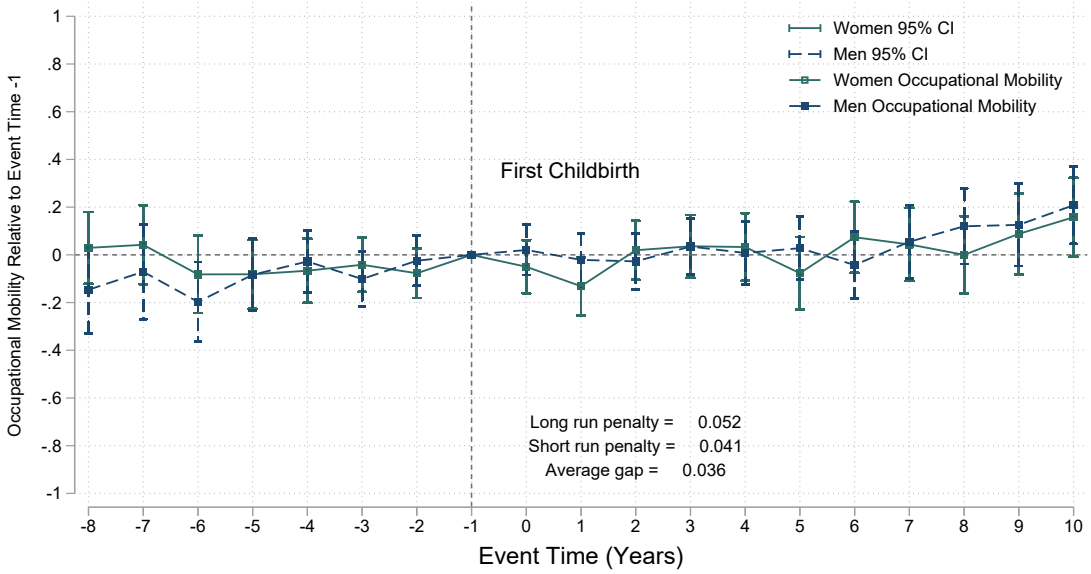
Figure A.16: Impact of Parenthood on Earnings (Un)Conditional on Occupation and Industry Fixed Effects



Notes: OLS results for equation (1.1) presented alongside the motherhood penalties (short, long and average gaps) in annual labor earnings, conditional on 2-digit occupation fixed effects — Standard Occupational Classification (SOC) codes — and 4-digit industry fixed effects — Standard Industrial Classification (SIC) codes. We cannot display men's estimates for $t = 10$ due to lack of observation. Thus, the average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 9, encompassing the 9-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 9.

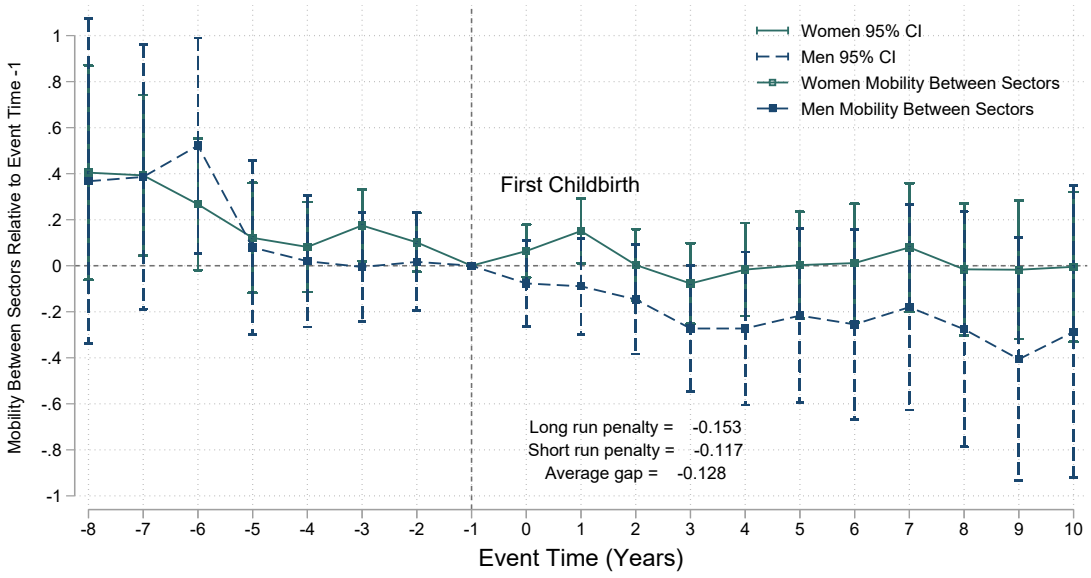
A.2.7 Mobility Around Childbirth

Figure A.17: Impact of Parenthood on Occupational Mobility



Notes: OLS results for equation (1.1) on a dummy variable equal to 1 if the participant changed occupation between $t-1$ and t, 0 otherwise. This figure shows the estimated impact of having a first child on any job mobility. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

Figure A.18: Impact of Parenthood on Mobility Between Sectors



Notes: OLS results for equation (1.1) on a dummy variable equal to 1 if the respondent works in the public sector in time t, regardless of where they were working in $t-1$. This figure shows the estimated impact of having a first child on mobility towards the public sector. The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

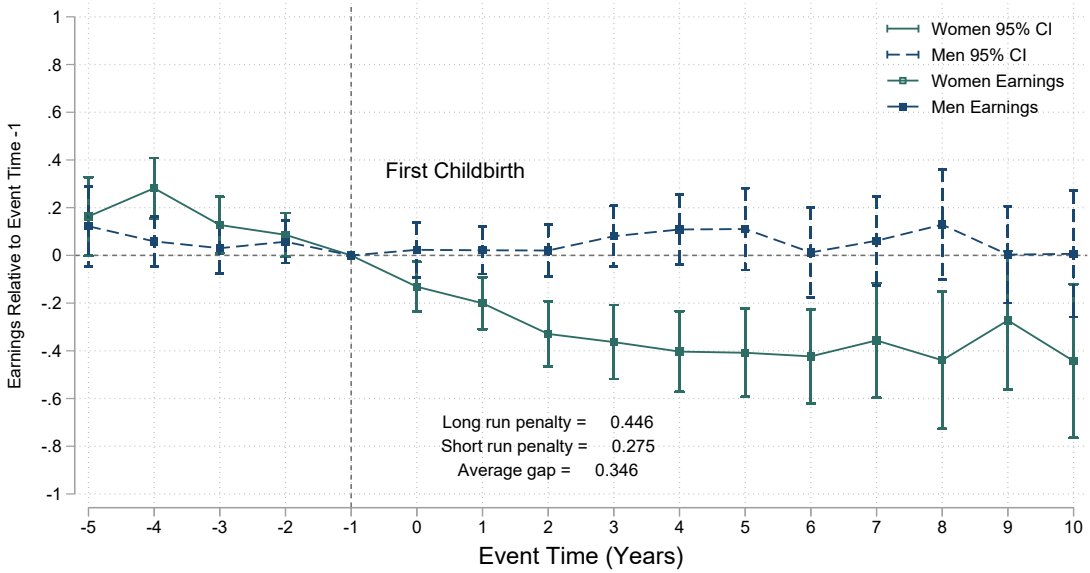
A.2.8 Robustness

A.2.8.1 Earnings Impacts by Number of Children

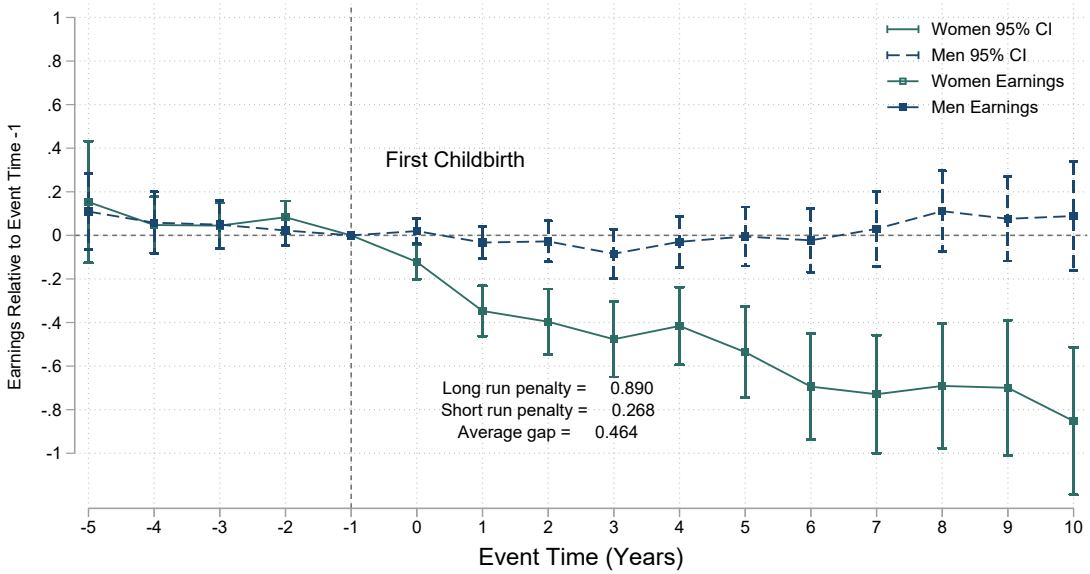
In the idea of replicating Kleven, Landais, and Søgaard (2019) and particularly the checks provided in the online appendix, we study the earnings impacts by number of children as of the last year of the panel. We do not have enough observations to study separately the impact of the third or the fourth child, so we study separately the earnings impacts for one-child parents, two-child parents and parents that have three or more kids. We do not report results for the last group because point estimates are too imprecisely estimated due to sample size.

Figure A.19: Impact of Parenthood on Earnings by Number of Children

(a) One-child parents



(b) Two-children parents

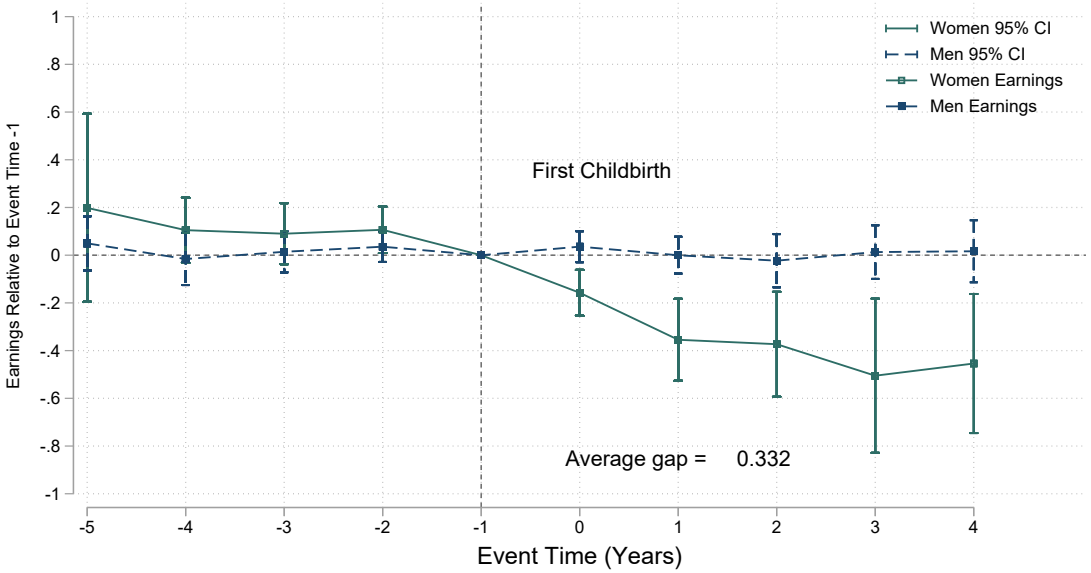


Notes: The figures show the impacts of children on earnings, for parents of (a) only one child as of 2009, and (b) two children as of 2009.

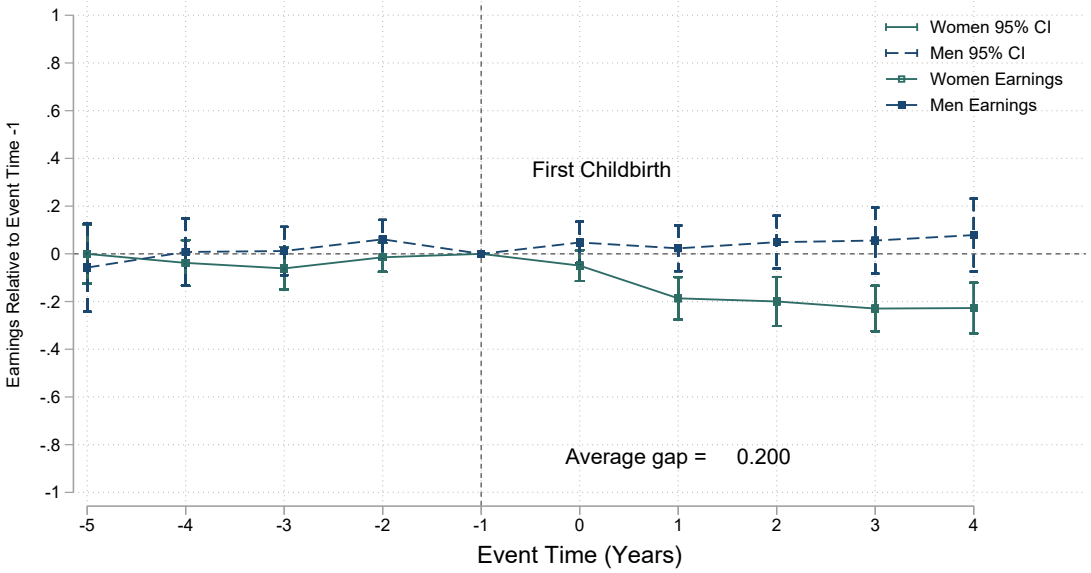
A.2.8.2 Heterogeneous Treatment Effects: Sun and Abraham (2021)'s Interaction Weighted Estimates

Figure A.20: Impact of Parenthood on Earnings by Pre-Birth Gender Norms

(a) Traditional pre-birth



(b) Egalitarian pre-birth

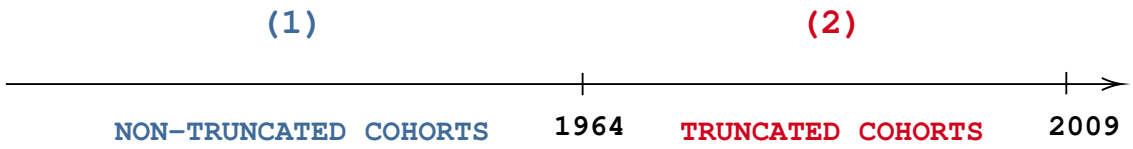


Notes: Both figures show the estimated impact of having a first child on annual labor earnings within each gender group using Sun and Abraham (2021)'s estimator to account for heterogeneous treatment effects, for (a) traditional and (b) egalitarian parents, as defined in Subsection 1.2.4.2. We use the last treated cohorts (year of the first childbirth being either 2004, 2005, 2006, 2007 or 2008 due to sample size issues) as a control group. Therefore, we plot $P_t^g = \hat{\beta}^g / \tilde{Y}^g$, where the counterfactual outcome is the last treated cohort's.

A.2.9 DiD Event Study

We follow Kleven, Landais, and Søgaard (2019) who use men and women who never have children as controls, and further assign placebo births using the actual distribution of age at first child among those who have children — the treated. Our last observed year is 2009, which further implies the following truncation:

Figure A.21: Control Cohorts Set-Up



1. Non-truncated cohorts born in 1964 or before (45 or older) unlikely to have children after the end of the survey in 2009 — in blue (1) above (Figure A.21).

2. Truncated cohorts born after 1964 (aged less than 45) that do not have children but might have some after the end of the survey — in red (2) above.

Therefore, for those born after 1964 and younger than 45 (the truncated cohorts), we select those who are most likely never to have children based on a linear probability model as in Kleven, Landais, and Søgaard (2019), and estimated separately for men and women.[1] The selection criteria within the truncated cohorts also relies on Kleven, Landais, and Søgaard (2019) and consists in selecting a subset n_c of respondents with the highest estimated probabilities of never having children, written as $n_c = N_c \times \mathbb{P}_{<1964}$, with N_c the total number of individuals in the cohort, and $\mathbb{P}_{<1964}$ the average share of childless respondents before 1964. Our control group is then constituted of 1) those born before 1964 without children, and 2) those selected as described above, and born after 1964.

We now need to assign placebo births to both cohorts within the control group. Again, considering the truncation, we distribute the age at first child following Kleven, Landais, and Søgaard (2019),[2] and can now implement event studies that compare our treatment group — those who have their first child and meet our analytical sample restrictions — to a control group — a panel of those who do not have children as of 2009, have been assigned placebo births and also meet our analytical sample restrictions. Descriptive statistics for this new analytical sample are presented in Table A.9 below.

We estimate the impact of children as such, and run our main event study specification without controlling for year dummies, as in Kleven, Landais, and Søgaard

[1]The linear probability model corresponds to estimating $P[\kappa_{iT} = 0] = X'\beta$ where κ_{iT} is a dummy for zero lifetime fertility, and X includes the following dummy variables: mother's and father's educational qualifications in five categories, within-cohort quantiles of labor earnings, a dummy for holding a college degree, and 13 dummies for government office regions.

[2]For the non-truncated cohorts, we assign a log-normal distribution of age at first child ($A_{c,e} \sim LN(\hat{\mu}_{c,e}, \hat{\sigma}^2_{c,e})$) where $\hat{\mu}_{c,e}$ is the observed mean within each cohort-college-educated *vs.* college-educated cohorts cells, and $\hat{\sigma}^2_{c,e}$ their variance. For the *truncated cohorts*, we assign a log-normal distribution of age at first child ($A_{c,e} \sim LN(\tilde{\mu}_{c,e}, \hat{\sigma}^2_{c,e})$) where $\tilde{\mu}_{c,e}$ is the predicted average age at first child obtained by estimating a linear trend on the older cohorts.

Table A.9: Alternative Sample's Summary Statistics by Treatment Status

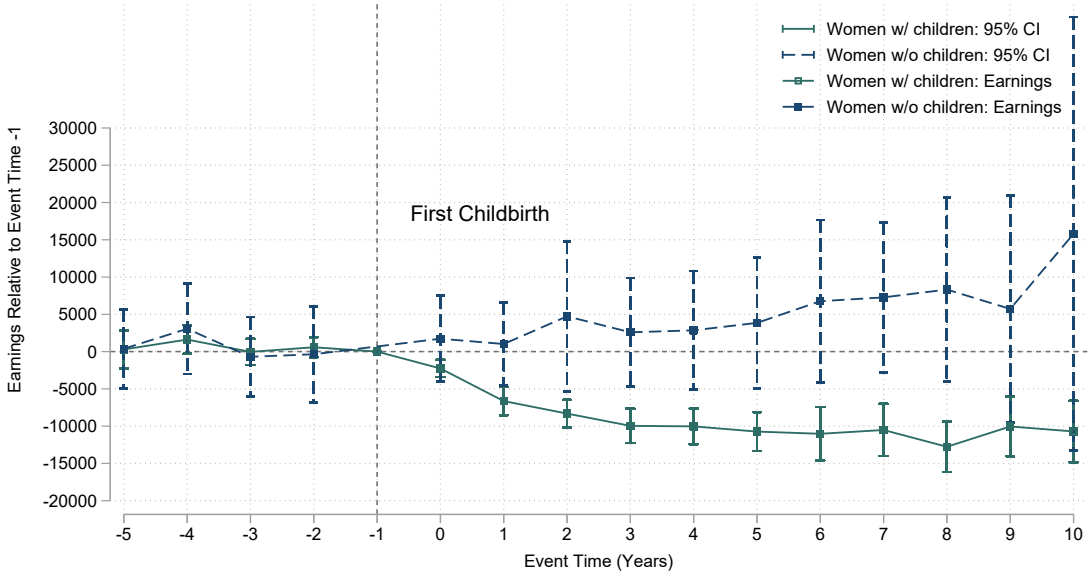
	Treated	Control	Diff. (Treated – Control)	S.E.	N
Age	29.191	34.087	-4.896***	0.097	20458
Age at parenthood	25.507	32.372	-6.865***	0.047	22256
Household size	3.105	3.004	0.102***	0.018	20458
Education					
Primary	0.075	0.045	0.030***	0.003	20224
Low secondary	0.064	0.035	0.029***	0.003	20224
Low secondary/vocational	0.313	0.254	0.059***	0.006	20224
High secondary/vocational	0.188	0.133	0.056***	0.005	20224
Higher vocational	0.187	0.233	-0.046***	0.006	20224
First degree	0.146	0.243	-0.097***	0.006	20224
Higher degree	0.022	0.056	-0.035***	0.003	20224
Marital status					
Never married	0.596	0.261	0.335***	0.007	20450
Married	0.333	0.666	-0.333***	0.007	20450
In a civil partnership	0.001	0.000	0.001*	0.000	20450
Separated	0.021	0.020	0.001	0.002	20450
Divorced	0.046	0.051	-0.005*	0.003	20450
Widowed	0.003	0.001	0.002***	0.001	20450
Labor market outcomes					
Labor earnings	11773.503	17196.701	-5423.198***	196.697	20209
Weekly hours worked	29.749	33.229	-3.479***	0.275	20432
Hourly wage	8.411	10.945	-2.534***	0.136	16818
LFP	0.793	0.897	-0.104***	0.005	20453

Notes: Summary statistics for the new analytical sample, by treatment status. * p <0.10; ** p <0.05, *** p <0.01. Levels of educational attainment follow ISCED coding, see Table A.1 for a detailed description of categories. Treated respondents correspond to our analytical sample described in Subsection 2.3.1, while the control cohort is a panel of those who do not have children as of 2009 (last survey year) and have been assigned placebo births. These also meet our analytical sample restrictions. LFP stands for labor force participation.

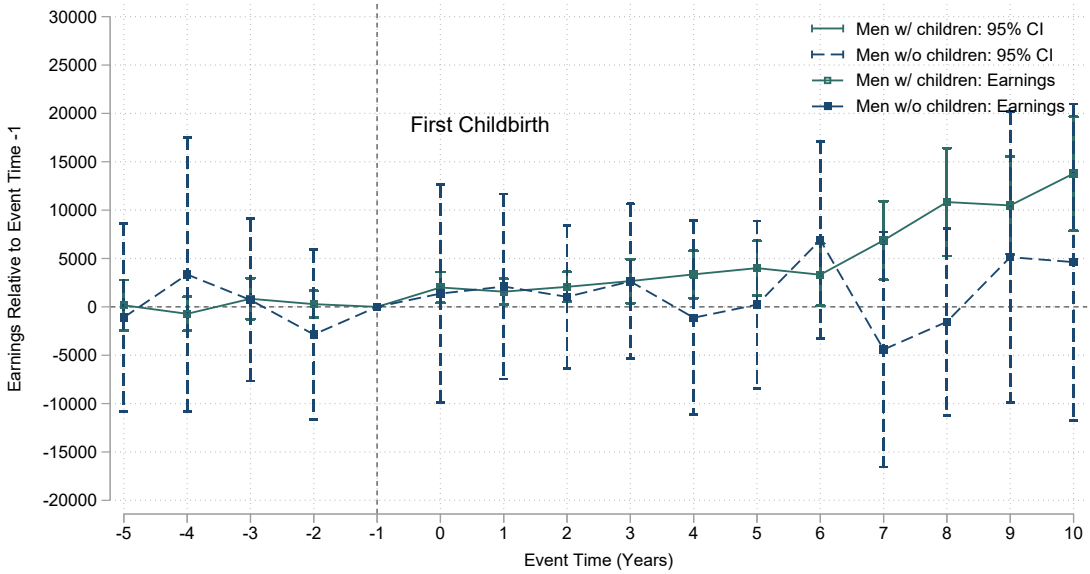
(2019), on labor earnings, by pre-birth gender norms. Overall, this implementation check confirms the key findings concerning the standing of the motherhood penalty by gender norms. Mothers experience a sharp decrease in their short-run labor earnings as a reaction to the first child birth, while men, on the other hand, are unaffected by fatherhood. The reduction is significantly more important for mothers with traditional views pre-birth.

Figure A.22: Impact of Parenthood on Earnings for Traditional Parents in DiD-ES Design

(a) Pre-birth traditional women who have children *vs*. women who do not



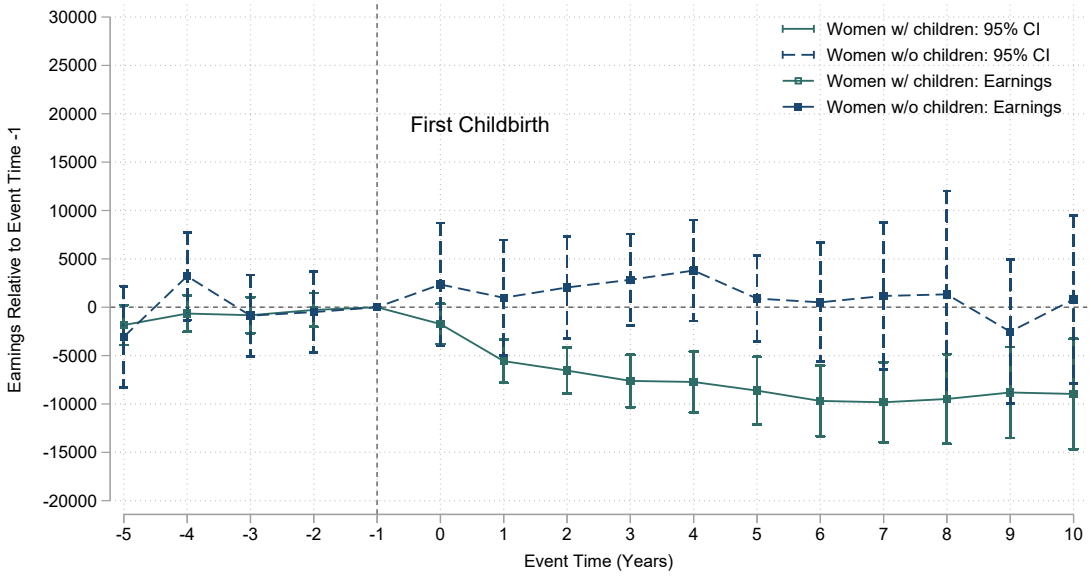
(b) Pre-birth traditional men who have children *vs*. men who do not



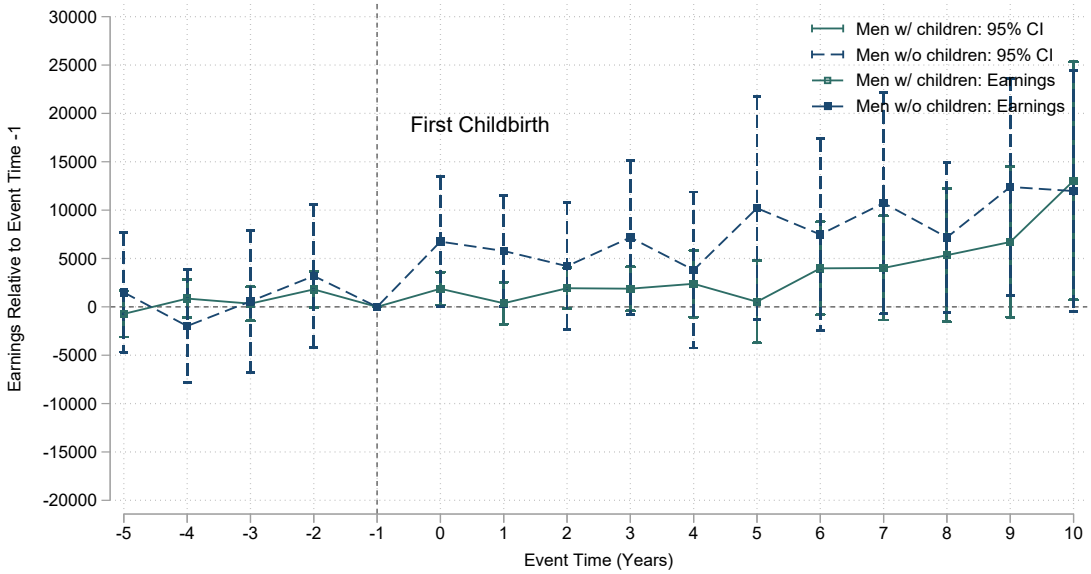
Notes: The figures show the evolution of labor earnings relative to the year before the first child for (a) women and (b) men with traditional gender norms pre-birth, as defined in Subsection 1.2.4.2. We compare those with children (treated) to those without children as of 2009 — last available year of BHPS — but that have been assigned placebo births (control). In the legend of both graphs, w/ stands for with, and w/o for without.

Figure A.23: Impact of Parenthood on Earnings for Egalitarian parents in a DiD-ES Design

(a) Pre-birth egalitarian women who have children *vs*. women who do not



(b) Pre-birth egalitarian men who have children *vs*. men who do not



Notes: The figures show the evolution of labor earnings relative to the year before the first child for (a) women and (b) men with egalitarian gender norms pre-birth, as defined in Subsection 1.2.4.2. We compare those with children (treated) to those without children as of 2009 — last available year of BHPS — but that have been assigned placebo births (control). In the legend of both graphs, w/ stands for with, and w/o for without.

A.3 Technical Details

A.3.1 Sample Checks

A.3.1.1 Condition on Pre-Birth Employment

We present in Table A.10 below the average gaps in earnings under two scenarios of pre-birth employment, (1) including parents who reported at least one year of positive labor earnings prior to the first childbirth, and (2) dropping observations with zero annual labor earnings prior to the first childbirth. Note that the version (1) is the one used throughout the paper. The results of this check appear qualitatively similar across the two sample restrictions, confirming the robustness of our main findings.

Table A.10: Average Gaps in Earnings Under Various Pre-Birth Employment Conditions

	(1)	(2)
Average gaps (%)		
Full sample	47.7	47.6
Traditional	60.8	59.5
Egalitarian	43.4	44.4
Difference (in pp)	17.4	15.1
Individuals	755	761
Observations	8350	8368

Notes: Average gaps in annual labor earnings retrieved from estimating equation (1.1) under two scenarios of pre-birth employment. In column (1), the sample includes parents who reported at least one year of positive labor earnings prior to the first childbirth. This is the version used throughout the paper. In column (2), we drop observations with zero annual labor earnings prior to the first childbirth.

A.3.1.2 Trimming of the Annual Labor Earnings Distribution

Table A.11 presents the average gaps in earnings for various trims of the annual earnings distribution, showing qualitatively consistent gaps across different versions.

Table A.11: Average Gaps in Earnings Under Various Trimming Conditions

	(1) Trimming 1st and 99th Percentiles	(2) Trimming 99th Percentile	(3) No Trimming
Average gaps (%)			
Full sample	47.7	73.6	80.9
Traditional	60.8	115.8	124.8
Egalitarian	43.4	53.6	58
Difference (in pp)	17.4	62.2	62.6
Individuals	755	755	755
Observations	8350	9230	10223

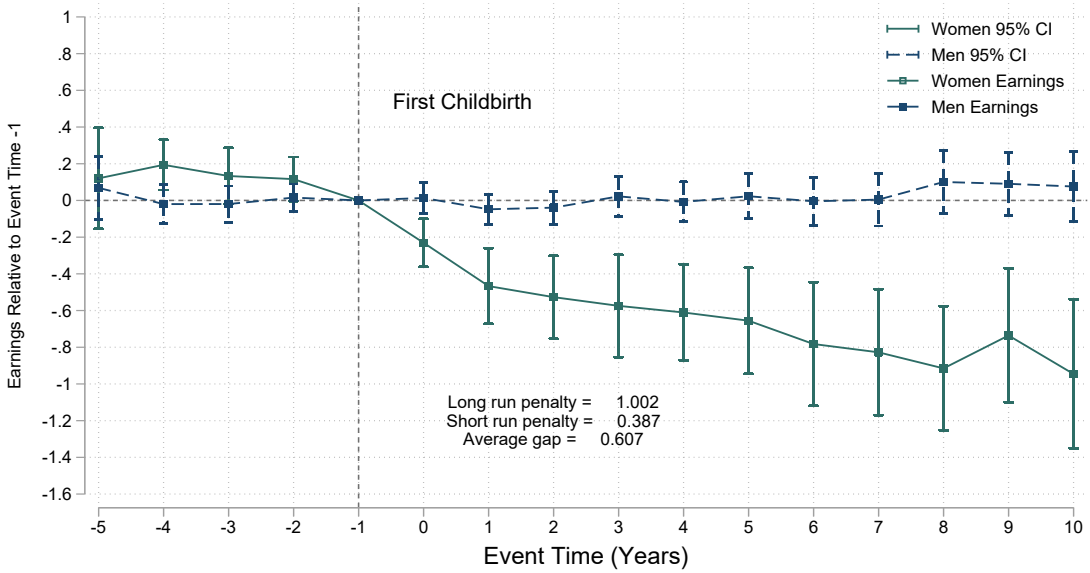
Notes: Average gaps in annual labor earnings retrieved from estimating equation (1.1) trimming (1) the bottom and top percentiles (1%) of the annual earnings distribution, (2) only the top 1%, and (3) not trimming this distribution. Results in column (1) correspond to the approach used throughout the paper.

A.3.1.3 Observations Within the Event Window

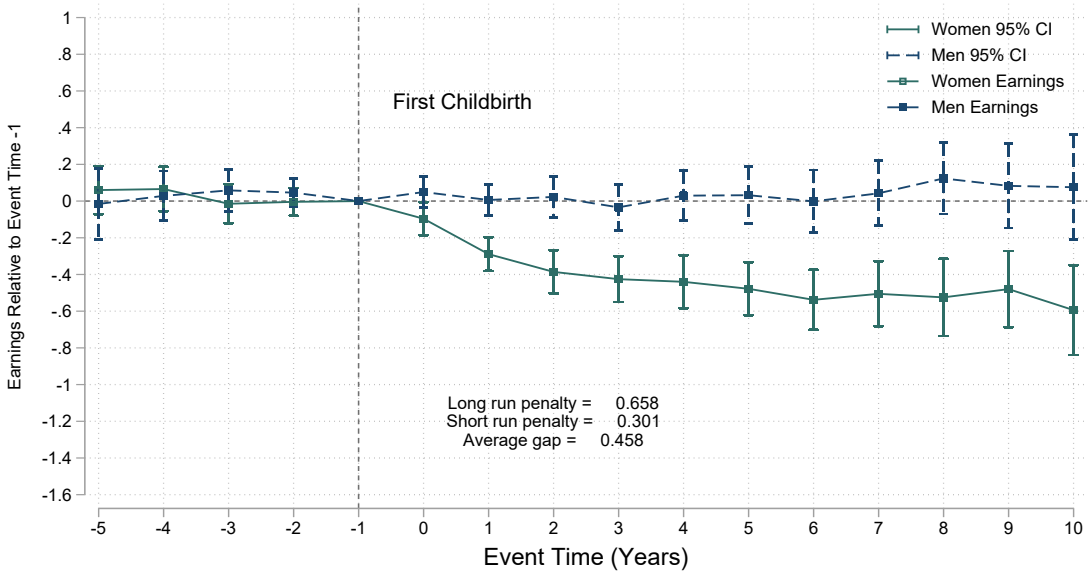
We run our main specification (equation (1.1)) on earnings by pre-birth gender norms, now retaining at least eight (instead of five in the paper) observations within the event window as done in Kleven et al. (2019). Graphical results displayed in Figure A.24 below confirm that our findings remain consistent with this sample selection criterion, substantiating the robustness of our results.

Figure A.24: Impact of Parenthood on Earnings by Pre-Birth Gender Norms

(a) Traditional pre-birth



(b) Egalitarian pre-birth



Notes: This figure illustrates the impact of motherhood on earnings, for (a) traditional and (b) egalitarian parents as defined in Subsection 1.2.4.2, with parents observed at least *eight* times within the event window, as opposed to our criterion of five observations in the paper. With at least eight observations within the event window, we have a sample of 608 parents and 7,600 observations. When conditioning on five observations to increase sample size (see Figure A.3), traditional women exhibited an average earnings gap of approximately 61%, while egalitarian women showed a gap of 43%. These figures closely align with our main result, confirming the robustness of our findings to different sample selection criteria.

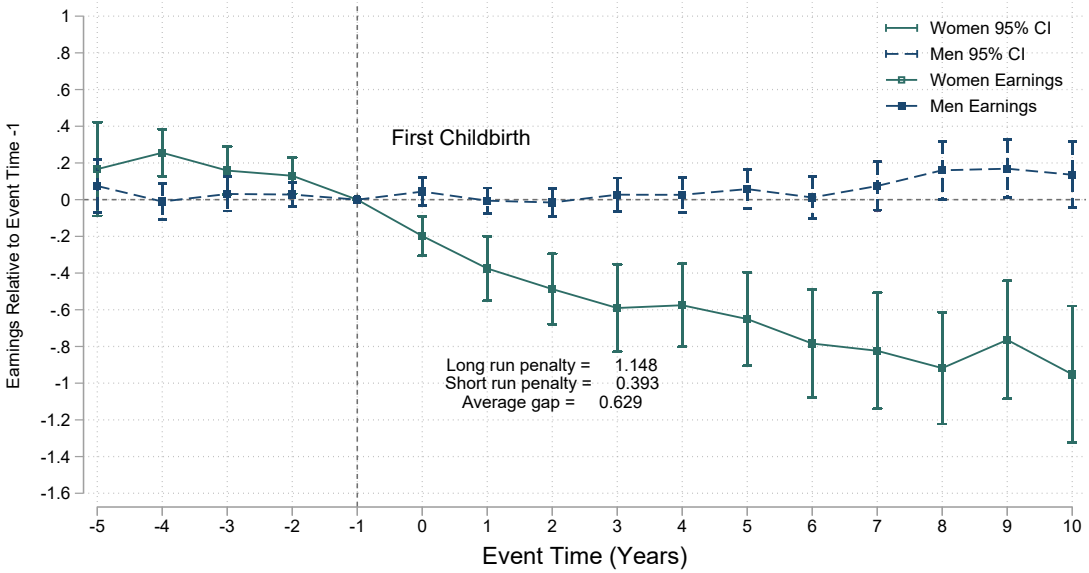
A.3.2 Reliability of the Gender Norms Indicator

A.3.2.1 Indicator Construction

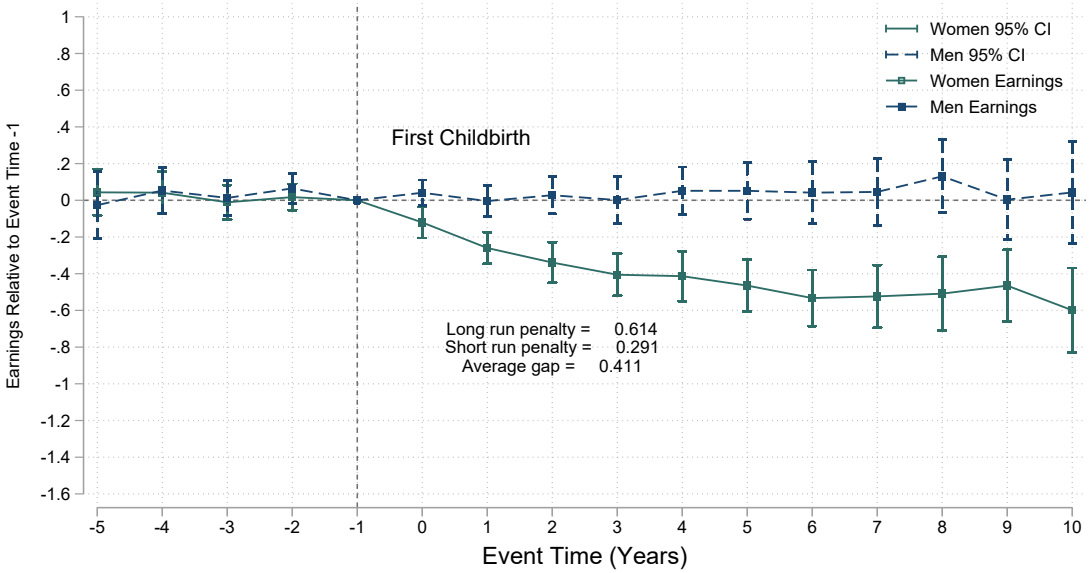
To ensure that the results based on gender norms are not influenced by the way we constructed our indicator, we modify its definition, and instead of dividing the sample into two groups based on the *median* value of the pre-birth gender norms score, we divide the sample in two groups based on the *average* value of the pre-birth gender norms score. Individuals below the average value are considered more 'traditional' while individuals above the mean are considered more 'egalitarian'. We run the same event study regressions as presented in Figure A.3 but with this new indicator. Figure A.25 below presents the percentage effects of parenthood on earnings for the two new panels. The main specification in the paper suggests that traditional women suffer from a 18-pp-larger motherhood penalty in earnings than more egalitarian women (Figure A.3). With this new specification, we find a similar pattern experienced by traditional and egalitarian women, with qualitatively similar magnitudes as per our main specification. The results of these checks are reassuring for the validity of our previous conclusions, particularly on the role of pre-birth gender norms in contributing to the motherhood penalty in earnings.

Figure A.25: Impact of Parenthood on Earnings by Pre-Birth Gender Norms (New Definition)

(a) Traditional pre-birth



(b) Egalitarian pre-birth



Notes: OLS results for equation (1.1) on annual labor earnings by the new definition of pre-birth gender norms for (a) traditional and (b) egalitarian parents using a new definition. Here, we redefine traditional *versus* egalitarian by dividing our sample based on the *average* value of the pre-birth gender norms score, rather than using the *median* value (as defined in Subsection 1.2.4.2). We further illustrate the motherhood penalty (P_t, equation (1.3)) for the short-term (within three years of becoming parents), and the long-term (in the last three observed years, $t \in [7;10]$), as well as the average gap (within ten years of becoming parents).

A.3.2.2 Measuring Gender Norms: The Role of Social Norms

The concept of norms has been advocated in the economics literature as being an important driver of individuals' behaviors (*e.g.*, Akerlof and Kranton, 2000). However, norms are difficult to capture and hence to describe as they are largely internalized within individuals, who are mainly unaware of such norms. This calls for carefulness in categorizing people as being either 'traditional' or 'egalitarian'. This is why we often use more relative expressions, such as "more traditional" or "more egalitarian" and divide our sample by the median value, enabling individuals to be relatively more or less egalitarian or traditional compared to their peers. Additionally, the notion of "gender norms" we use and how the variables reflect this concept within BHPS is based on pseudo-arbitrary choices, despite some authors using the same variables (Flèche, Lepinteur, and Powdthavee, 2020; Schober and Scott, 2012 and Grinza et al., 2022).[3] To make sure we are actually capturing "gender norms", we tested the effect of childbirth on a broader measure of norms — "social norms."[4] We assess this as, first, we would expect the effect of childbirth to be smaller or null on social norms, and, second, as we assume that there is no incentive *a priori* to become more conservative as regards social norms upon parenthood.[5] Table A.12 below presents the 6 variables asked to reflect individuals' social norms. These are asked to the primary survey respondent in BHPS, collected at each odd wave, commencing from wave 1 and up to wave 17. The answer categories are also presented below.

Table A.12: Social Norms Variables

	Answer categories
Ordinary people get their fair share of the nation's wealth There is one law for the rich, and one law for the poor Private enterprise is the best way to solve Britain's economic problems Major public services and industries ought to be in state ownership It is the government's responsibility to provide a job for everyone who wants one Strong trade unions are needed to protect the working conditions and wages of employees	1. Strongly agree 2. Agree 3. Neither agree, nor disagree 4. Disagree 5. Strongly disagree

Notes: Social norms variables are asked to the main survey respondent, every odd wave, starting wave 1, until wave 17.

The scale is reversed where needed (Questions 3, 4 and 5 in the above Table), to enable the highest value to reflect more progressive social norms. A lower value thus corresponds to more conservative norms.

As for gender norms variables, we construct a score reflecting whether individuals are more or less progressive as regards social norms. We take the average answer to these six questions to compute this score — where one is the minimum value (conservative norms), and six is the maximum value (more progressive norms). As in Subsection 1.4.4, we run our main specification described in equation (1.1) on this social norms score in order to understand whether such norms are affected by

[3]We are cautious in interpreting our results however. The data provider of BHPS classified the 6 variables we are using as reflecting "gender roles" which confirms our intuitions of capturing the right concept.

[4]Also referred as "social opinions" in BHPS.

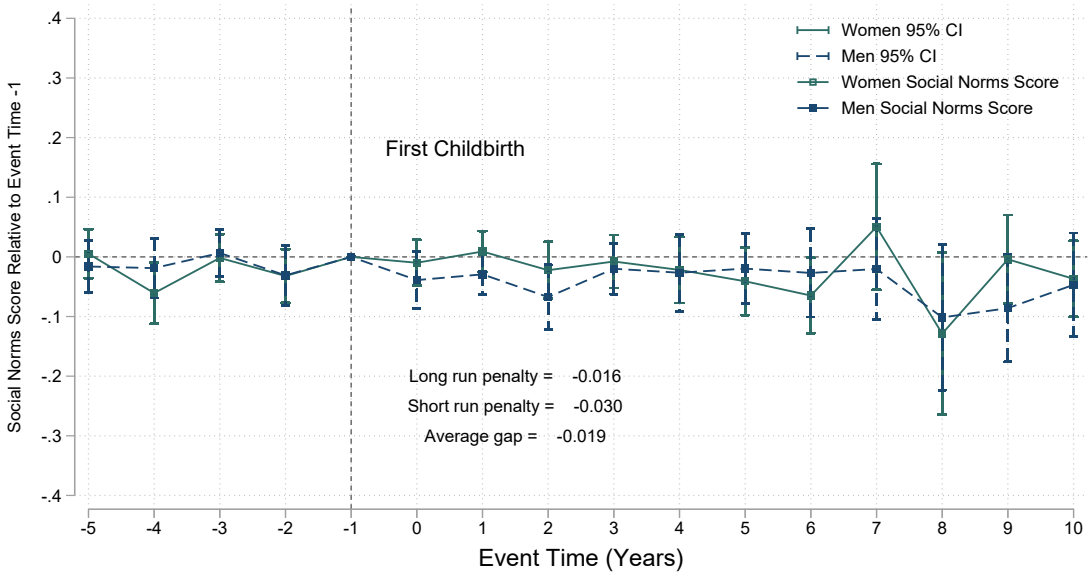
[5]The unconditional correlation between social norms pre-birth and gender norms pre-birth (both continuous scores) is equal to 0.0926 ($p < 0.01$).

the event of having a child, and whether there is any gendered effect. Similarly, this exercise enables to understand if there is any compositional change that could affect the validity of our main results split by gender norms.

Intuitively, we expect the effect of the first childbirth to be smaller or null on social norms, as one would think there is no obvious incentive to become more conservative on such social aspects after parenthood. Yet, social norms can also constitute a broader measure, themselves including — and reflecting — gender norms. Indeed, conservative social opinions can be associated with a broader traditional background, and therefore more traditional gender norms (*e.g.*, Sanbonmatsu, 2002)

OLS results for equation (1.1) on social norms are displayed in Figure A.26 below and confirm the validity of our main results split by gender norms. Indeed, we do not see any clear emerging pattern after parenthood for mothers and fathers, as well as no gender difference.

Figure A.26: Impact of Parenthood on Social Norms



Notes: OLS results for equation (1.1) on the social norms score, *i.e.*, the estimated impact of having a first child on the social norms score within each gender group (P_t^g, as described in equation (1.2)). The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

A.3.2.3 Disentangling Socioeconomic Status and Gender Norms

The aim here is to investigate whether our findings are influenced more by socioeconomic status (SES) than by gender norms, as these positively correlate (0.1158, $p < 0.01$). The majority of available data in BHPS concerning respondents' social class relies on information related to current or past occupations (*e.g.*, Goldthorpe Social Class, RG Social Class, Cambridge Scale provided by the survey), which may not be ideal for assessing labor market outcomes. Consequently, we opt to use educational attainment, specifically a binary variable denoting whether participants hold at least a college degree, to proxy broadly for socioeconomic status.

First, we run the pooled difference-in-differences (as in equation (1.4)) for four different panels presented in Table A.13: lower education (Panel A), higher education (Panel B), traditional (Panel C), and egalitarian (Panel D), and we can finally compare the magnitudes and signs of the estimated average effects ($\hat{\theta}$). Second, we run the same specification but interacting the treatment dummy with an indicator for holding a college degree or less, and stratifying the regressions by pre-birth gender norms. Results are presented in Table A.14. Finally, Figure A.27 presents the results of our event study regressions split by college education.

Overall, our results suggest two key things. First, SES as captured by the highest educational level at the individual level, does not significantly contribute in shaping the trajectory of labor earnings as both groups experience a significant drop in their labor earnings of almost comparable magnitudes. Second, a gap opens up between traditional women and egalitarian women with a significant difference. It therefore suggests that our results are not driven entirely by the socioeconomic status, as norms hold stronger than education in shaping the magnitude of the motherhood penalty.

Table A.13: Post-Birth Gendered Earnings Trajectory by Education and Pre-Birth Gender Norms (1/2)

	Panel A	**Panel B**	**Panel C**	**Panel D**
	Low education	Higher education	Traditional	Egalitarian
Female	-2852.872***	-3165.579***	-3280.310***	-2103.550***
	(539.229)	(1113.260)	(949.018)	(738.070)
Post	1521.725**	2323.349*	1437.921*	1277.890
	(624.076)	(1268.511)	(782.739)	(985.646)
Female × Post	-6984.086***	-9428.358***	-9981.738***	-6950.029***
	(758.897)	(1538.996)	(1117.266)	(1069.696)
Observations	5698	2652	4131	4150
Normalized Average Effect (%)	-.614	-.532	-.767	-.492
Year Fixed Effects	Yes	Yes	Yes	Yes
Age Fixed Effects	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. OLS results for equation (1.4) on annual labor earnings. High education corresponds to holding at least a college qualification. The 'post' dummy takes the value 1 if $t \geq 0$, and corresponds to the post-childbirth period. The normalized average effects divide the estimated average effect ($\hat{\theta}$) for each panel, by the baseline mean for each panel, respectively: for Panel A, by the pre-birth labor earnings average of lower educated women; for Panel B, by the pre-birth earnings average of higher educated women; and for Panel C (D), by the pre-birth earnings average of traditional (egalitarian) women.

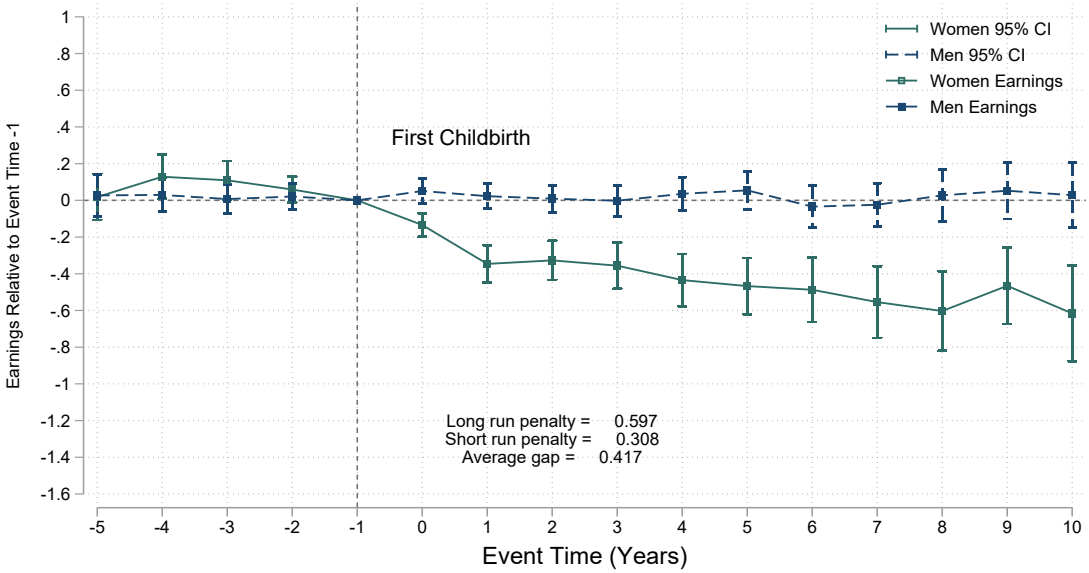
Table A.14: Post-Birth Gendered Earnings Trajectory by Education and Pre-Birth Gender Norms (2/2)

	Panel A	**Panel B**
	Egalitarian	Traditional
Post	5374.782***	4521.043***
	(1584.104)	(1636.486)
Woman	-2207.265*	-4112.518**
	(1241.140)	(2063.843)
Post × Woman	-10000.534***	-10197.460***
	(1845.637)	(2592.788)
Low education	-4827.698***	-6193.421***
	(1150.020)	(1204.404)
Post × Low education	-5975.784***	-3133.573*
	(1720.026)	(1778.996)
Woman × Low education	-106.311	1038.815
	(1466.661)	(2141.316)
Woman × Post × Low education	5140.831**	650.305
	(2145.319)	(2760.199)
Observations	4150	4131
Normalized Average Effect (%)	.364	.05
Year Fixed Effects	Yes	Yes
Age Fixed Effects	Yes	Yes

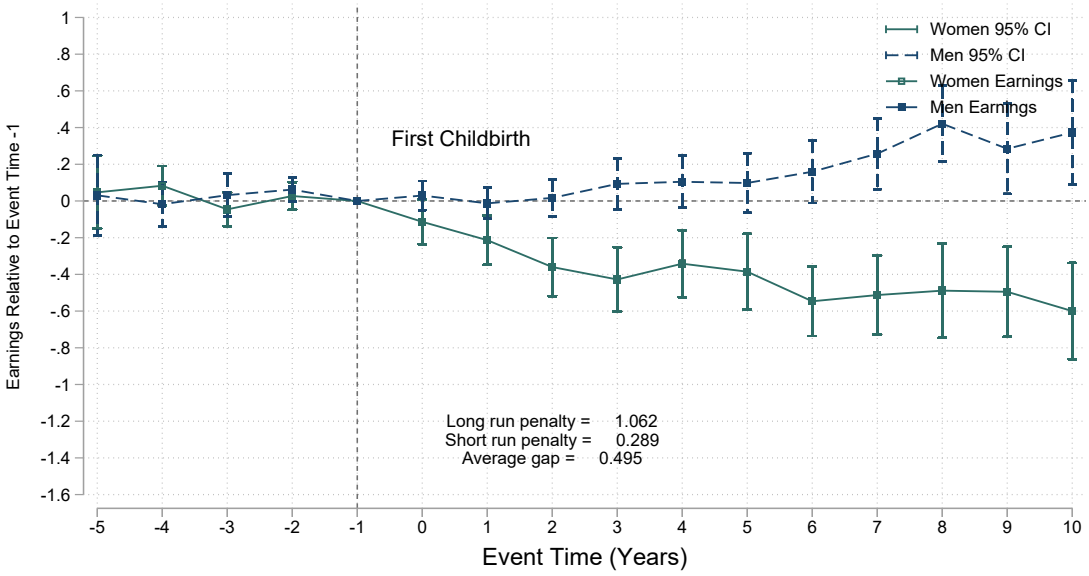
Notes: * p <0.10; ** p <0.05, *** p <0.01. OLS results for equation (1.4) on annual labor earnings for egalitarian and traditional parents, as defined in Subsection 1.2.4.2. The 'post' dummy takes the value 1 if $t \geq 0$, and corresponds to the post-childbirth period. The normalized average effects correspond to the estimated average effect ($\hat{\theta}$) divided by the baseline mean for each panel, respectively: for Panel A, by the pre-birth labor earnings average of egalitarian women; for Panel B, by the pre-birth earnings average of traditional women.

Figure A.27: Impact of Parenthood on Earnings by Education

(a) Does not have a college degree



(b) Does have at least a college degree



Notes: OLS results for equation (1.1) on annual labor earnings for parents who do not have a college degree (a) and those who have at least a college degree (b). Both figures show the estimated impact of having a first child on earnings within each gender group (P_t^g, as described in equation (1.2)). The average gap indicates the mean of P_t (as described in equation (1.3)) for t spanning from 0 to 10, encompassing the 10-year period following parenthood. The short-run gap represents the mean of P_t during the initial 3 years after the first childbirth, covering t from 0 to 3, while the long-run gap corresponds to the average of all P_t for t from 7 to 10.

Appendix B

Additional Material for Chapter 2

B.1 Theoretical Appendix

B.1.1 Model Solution

This appendix shows how our two results illustrated in figure 2.1 holds generally in our model framework. Specifically, we solve the model and show that (i) when $\gamma > 0$ and $\rho_m = \rho_f$, mothers supply less labor due to a comparative advantage in child-rearing, and (ii) when if $\gamma = 0$ and $\rho_m > \rho_f$, mothers supply less labor due to an absolute advantage in child-rearing.

Substituting in the budget constraint, we write the household's optimization problem as

$$HC(m^*, f^*, e^*) = \max_{m,f} \; m^{\rho_m} f^{\rho_f} \left[(1-\gamma)W(1-m) + W(1-f)\right]^{\rho_e}.$$

Taking logarithms and solving the first-order conditions gives:

$$\begin{aligned}
\frac{\partial \log HC}{\partial m} = 0 \iff \frac{\rho_m}{m} &= \frac{\rho_e(1-\gamma)}{(1-\gamma)(1-m)+(1-f)}, \\
\frac{\partial \log HC}{\partial f} = 0 \iff \frac{\rho_f}{f} &= \frac{\rho_e}{(1-\gamma)(1-m)+(1-f)}.
\end{aligned}$$

Combining these gives:

$$\frac{f^*}{m^*} = \frac{\rho_f}{\rho_m}(1-\gamma).$$

The two results follow directly from this optimality condition.

B.1.2 Model Extension

This section presents an extended version of our conceptual framework, adjusting the child human capital production function to allow for different substitutability between time and financial inputs and for differential returns to scale. Unlike our baseline model, this augmented version can capture the empirical observation that the

effect of mothers working longer hours on child development diminishes as household earnings increase.

We start by showing the limitations of the baseline model in replicating this observation. Specifically, we demonstrate that models where time and money inputs are *q-complements* – such as our baseline model – cannot replicate the observed negative effect. A general child human capital production function featuring time inputs from each parent and earnings as inputs can be written as $HC = f(T(m,f),e)$ where $T(m,f)$ represents the aggregation of parental time inputs in domestic work, and e is earnings. We assume that f is increasing in both inputs. Time and earnings are considered *q-complements* if the marginal productivity of one input rises with the level of the other, i.e. $\frac{\partial^2 f}{\partial T \partial e} > 0$.

Let h_ℓ and h_s represent long hours and short hours at-home, respectively. As before, we define our empirical target θ as

$$\theta := f(T(\underbrace{h_s, h_\ell}_{MWL=1}),e) - f(T(\underbrace{h_\ell, h_s}_{MWL=0}),e).$$

Since f is increasing in T, we have $\theta < 0$ if and only if $T(MWL=1) < T(MWL=0)$. Our empirical finding is that for cases with $\theta < 0$ we observe $\frac{\partial \theta}{\partial e} > 0$. However, in the model

$$\frac{\partial \theta}{\partial e} = \frac{\partial f(T(MWL=1),e)}{\partial e} - \frac{\partial f(T(MWL=0),e)}{\partial e}$$

which is *negative* if $\frac{\partial^2 f}{\partial T \partial e} > 0$.

To reconcile the model with the data, we propose an augmented model incorporating a CES (Constant Elasticity of Substitution) aggregator for time and earnings inputs, with a parameter to control returns to scale. The revised human capital function is given by:

$$HC = \left((m^{\rho_m} f^{\rho_f})^\sigma + e^\sigma\right)^{\alpha/\sigma},$$

where $\sigma \in (-\infty, 1)$ governs the substitutability between T and e while $\alpha \in (0,\infty)$ governs the returns to scale with respect to T and e. Using this definition, we derive:

$$\frac{\partial \theta}{\partial e} = \alpha \left((h_s^{\rho_m} h_\ell^{\rho_f})^\sigma + e^\sigma\right)^{\frac{\alpha}{\sigma}-1} e^{\sigma-1} - \alpha \left((h_\ell^{\rho_m} h_s^{\rho_f})^\sigma + e^\sigma\right)^{\frac{\alpha}{\sigma}-1} e^{\sigma-1}.$$

Thus, $\frac{\partial \theta}{\partial e} > 0$ if and only if

$$\left((h_s^{\rho_m} h_\ell^{\rho_f})^\sigma + e^\sigma\right)^{\frac{\alpha}{\sigma}-1} > \left((h_\ell^{\rho_m} h_s^{\rho_f})^\sigma + e^\sigma\right)^{\frac{\alpha}{\sigma}-1}.$$

After some manipulation it is clear that three parameter ranges are relevant for evaluating this inequality.

$$\begin{cases} \text{If } \sigma < 0 \text{ or } 0 < \sigma < \alpha & \text{we require } h_s^{\rho_m - \rho_f} > h_\ell^{\rho_m - \rho_f}. \\ \text{If } \sigma > \alpha & \text{we require } h_s^{\rho_m - \rho_f} < h_\ell^{\rho_m - \rho_f}. \end{cases}$$

Since $h_\ell > h_s$ and $\rho_f > \rho_m$ (indicating $\theta < 0$ — women hold an absolute advantage

in domestic work), only the second inequality holds. We conclude that θ is increasing in earnings if and only if $\sigma < \alpha$, i.e. if the degree of substitutability between time and earnings is greater than the degree of returns to scale. Since σ is bounded above by 1 this requires that T and e are both gross substitutes ($\sigma > 0$) and that there is decreasing returns to scale ($\alpha < 1$).

B.2 Hypothetical Beliefs Elicitation: Additional Results

B.2.1 Sample

Table B.1: Sample Representativeness

	National Population		Sample		Sample	
	Mean	SE	Unweighted Mean	SE	Weighted Mean	SE
Gender*						
Man	0.46	0.01	0.50	0.02	0.48	0.02
Woman	0.54	0.01	0.50	0.02	0.52	0.02
Age	43.69	0.22	38.46	0.22	40.98	0.35
Born in the UK	0.90	0.12	0.82	0.01	0.91	0.01
Ethnicity						
Asian	0.10	0.01	0.08	0.01	0.06	0.01
Black	0.04	0.00	0.10	0.01	0.04	0.01
Mixed	0.02	0.00	0.02	0.00	0.02	0.01
White	0.83	0.01	0.80	0.01	0.86	0.02
Other	0.01	0.00	0.01	0.00	0.01	0.00
Education						
No qualification	0.03	0.00	0.00	0.00	0.00	0.00
Other	0.05	0.01	0.01	0.00	0.05	0.01
GCSE or equivalent	0.19	0.01	0.10	0.01	0.22	0.02
A-levels or equivalent	0.21	0.01	0.23	0.01	0.22	0.02
Degree or higher	0.52	0.01	0.66	0.01	0.51	0.02
Monthly net income (£)						
0-500	0.12	0.01	0.08	0.01	0.13	0.02
500-1000	0.09	0.01	0.06	0.01	0.10	0.01
1000-1500	0.16	0.01	0.12	0.01	0.15	0.02
1500-2000	0.17	0.01	0.19	0.01	0.18	0.02
2000-2500	0.16	0.01	0.20	0.01	0.15	0.02
2500-3000	0.11	0.01	0.12	0.01	0.11	0.02
3000+	0.18	0.01	0.09	0.01	0.07	0.01
Single parent	0.17	0.01	0.12	0.01	0.20	0.02
Number of children aged 0-16*	1.52	0.02	1.78	0.03	1.83	0.04
Participants	6237		1056		1056	

Notes: Means and standard errors (SE) of the key demographic information for a nationally representative sample (column 1) as well as for our survey participants (column 2). The national population figures are drawn from the relevant population of respondents to the 2022 Understanding Society wave, and weighted using the corresponding cross-sectional weight. The * indicates variables targeted through our sampling approach. See paragraph below for the description of our sample's weighting approach.

Weighting approach.

For some dimensions, our sample is different from the corresponding national population distribution of parents in England. In particular, those who hold a degree or higher are over-represented in our sample, and lower income categories are under-represented. Also, we find that our sample's average age, the shares of individuals born in the UK, and the share of single parents are somewhat lower than the national distribution. For all categories, we construct respective initial weights corresponding to the national population proportion divided by that in the sample. For instance, the weight for holding a degree or higher is equal to $w_{degree} = \frac{0.5248}{0.6553}$, while the weight

for no qualification is $w_{none} = \frac{0.0275}{0.009}$, etc., and we use 4-digit proportions to avoid zeros. As age is a continuous variable, we normalize its weight to lay on a scale from 0 to 1. Finally, we combine these initial weights for participants by multiplying all initial weights w_d.

Table B.2: Participants' Descriptive Statistics

	Treated	Control	Diff.	Overall
Gender				
Man	0.48	0.51	-0.03	0.50
Woman	0.52	0.47	0.05	0.50
Age	38.14	38.79	-0.65	38.46
Born in the UK	0.83	0.82	0.01	0.82
University graduate	0.66	0.65	0.00	0.66
Ethnicity				
Asian	0.06	0.09	-0.03	0.08
Black	0.11	0.09	0.02	0.10
Mixed	0.03	0.02	0.01	0.02
White	0.79	0.80	-0.00	0.80
Other	0.01	0.01	0.01	0.01
Vote at last UK General Election				
Conservative	0.10	0.10	0.00	0.10
Labour	0.46	0.48	-0.02	0.47
Liberal Democrat	0.11	0.10	0.02	0.10
Green Party	0.07	0.07	-0.00	0.07
Reform UK	0.09	0.08	0.01	0.09
Other	0.02	0.03	-0.02	0.02
None	0.16	0.14	0.01	0.15
Full-time employment	0.60	0.64	-0.04	0.62
Weekly hours worked	31.05	32.68	-1.63	31.86
Monthly net income (£)				
Low	0.46	0.45	0.02	0.45
Medium	0.33	0.31	0.02	0.32
High	0.21	0.25	-0.04	0.23
Single parent	0.12	0.13	-0.02	0.12
Number of children aged 0-16	1.78	1.78	-0.00	1.78
Partner's monthly net income (£)				
No partner	0.06	0.09	-0.03	0.07
Low	0.43	0.44	-0.00	0.43
Medium	0.32	0.31	0.01	0.31
High	0.25	0.26	-0.01	0.25
Participants	525	531		1056

Notes: This table displays the means of the key demographic information for our survey participants, by treatment status. Differences are statistically significant at the following levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note that six respondents listed "other" or "prefer not to say" for gender. For the participant's and the partner's (if any) monthly net income, we group the categories into tertiles. For the participant's income, "Low" encompasses income between 0 and £2,000,"Medium" considers income between £2,000 and £3,000, while "High" corresponds to monthly net income greater than £3,000. For the partner's, we construct tertiles of the original variable and for those who have a partner. The "Low" category corresponds to incomes between 0 and £1,500, the "Medium" category corresponds to incomes between £1,500 and £3,000, while the "High" category is for incomes above £3,000.

B.2.2 Hypothetical Beliefs Elicitation

Table B.3: Randomization in Set-Up

	Mean	SD	N
Child is a girl (vs. boy)	0.49	0.50	517
Child is aged 4 (vs. 10)	0.51	0.50	542
SSE_i: 20% (vs. 10%)	0.51	0.50	540

Notes: Total number of participants = 1056. This table presents descriptive statistics for the randomization in set-up. For instance, 49% of our sample (*i.e.*, 517 participants) got displayed, in the hypothetical scenarios, a boy child. SSE_i refers to the share of income spent on the child's educational and extracurricular activities.

B.2.2.1 Alternative Interpretations of Main Result

This subsection provides further support for our preferred interpretation of our main result (Result 1).

Inferred unobservable characteristics

One concern is that respondents may infer different unobservable characteristics in families where the father, rather than the mother, works longer hours. If so, responses may not reflect a *ceteris paribus* comparison, and our estimates could be biased by perceived differences in parental human capital rather than beliefs about maternal labor supply.

We view this interpretation as unlikely for two reasons. First, our heterogeneity analysis (Section B.2.2.2, Table B.5), shows that variation in $\hat{\delta}$ is strongly correlated with participants' own labor supply histories. Women who reduced their labor supply after childbirth have significantly lower values of $\hat{\delta}$, as do those whose mothers worked shorter hours when they were growing up. These patterns suggest that participants are drawing on internalized beliefs and personal experiences rather than inferring differences across vignette families.

Second, we directly tested whether individuals inferred differences in parental human capital across scenarios. We probed participants on the likelihood of the father and mother having a University degree in settings where we varied which parent worked full-time relative to part-time. The results, shown in Table 2.6, indicate that participants generally assign a higher probability of having a university degree to both mothers and fathers when they work full-time. However, this differential is not significantly different between mothers and fathers and point estimates suggest that part-time mothers are penalized slightly more than part-time fathers. If anything, this pattern would bias our estimated effect toward zero, not away from it.[1]

[1]Importantly, in this exercise we kept the family income level fixed in the scenarios where the mother or father worked full-time at a level of £1,312. 'As can be seen 'in Figure B.1, this is within the range where we find that $\hat{\delta} < 0$.

Order effects

We find that the effects are more substantial when participants saw the father working longer hours on the first page (and thus mothers on the second page). Because we randomize the order, meaning that the page order is orthogonal to the MWL indicator, then an order effect whereby people simply change beliefs based on the page number will not bias our $\hat{\delta}$-estimate in equation 2.2. For instance, if participants always downgrade their expectations on the second-page, then when mothers are shown first as working longer hours, expectations with fathers on the second-page will be pushed up by any beliefs on absolute advantage but pushed down by this order effect hiding the true degree of beliefs on absolute advantage. While for those randomized to first see fathers working longer hours, then on the second-page with mothers, both beliefs on absolute advantage and the order effect can widen the differences in expectations. Importantly, because we randomize the order, the net effect over all participants removes any order effect leaving the beliefs we are after.[2] Alternatively, this heterogeneity may arise because the gender difference in the two hypotheticals becomes more salient when participants encounter first the normative family arrangement with the father working longer hours priming attention to their core beliefs. In any case, we find no reason to be concerned.

B.2.2.2 Heterogeneity by participants' characteristics and past experiences

[2]Indeed, when we include a scenario order as a control variable in our regression, estimates remain unchanged.

Table B.4: Heterogeneity in Beliefs by Participants' Characteristics

	Median Age		Gender		Ethnicity		Born in the UK		University Degree		Working Hours		Participant's Income		Number of Children		Vote at last UK General Election		
	≥ 38	< 38	Female	Male	White	Non-white	Yes	No	Yes	No	FT	PT or none	$\geq$ Median	$<$ Median	< 2	≥ 2	Conservative	Liberal	Other or None
$\mathbb{P}$(graduate): $MWL_{j=1}$	-1.11***	-0.73	-0.86*	-1.02**	-0.59*	-2.26***	-0.63**	-2.37***	-1.15***	-0.52	-0.87**	-1.04**	-0.51	-1.48***	-0.97**	-0.89**	-0.94	-0.73*	-1.69***
	(0.39)	(0.46)	(0.45)	(0.40)	(0.32)	(0.79)	(0.31)	(0.88)	(0.37)	(0.51)	(0.37)	(0.50)	(0.39)	(0.46)	(0.40)	(0.44)	(0.74)	(0.38)	(0.63)
Earnings Rank: $MWL_{j=1}$	-0.64*	-0.70*	-0.36	-0.97***	-0.48*	-1.40*	-0.48*	-1.56**	-0.69**	-0.63	-0.44	-1.04**	-0.19	-1.29***	-0.74**	-0.57	-1.14*	-0.24	-1.76***
	(0.36)	(0.40)	(0.40)	(0.36)	(0.28)	(0.73)	(0.28)	(0.78)	(0.33)	(0.45)	(0.33)	(0.46)	(0.35)	(0.41)	(0.37)	(0.39)	(0.67)	(0.32)	(0.66)
Participants	563	493	525	525	840	216	871	185	692	364	657	399	596	460	613	443	194	679	183
Observations	3378	2958	3150	3150	5040	1296	5226	1110	4152	2184	3942	2394	3576	2760	3678	2658	1164	4074	1098
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scenario Income Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * $p < 0.10$; ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered on individuals. Each coefficient is obtained from separate OLS regressions estimating equation (2.2) on the subsample defined by the displayed label and for our two expected outcomes: (i) the probability for the child to graduate ($\mathbb{P}$(graduate)), and (ii) the earnings rank of the child at age 30 (Earnings Rank). "FT" stands for full-time, while "PT" stands for part-time.

Table B.5: Heterogeneity in Beliefs by Participants' Past Experiences

	Own-Mothers' FT Employment				By the Median Motherhood Penalty	
	Yes: when you were < 12	No: when you were < 12	Yes: when you were ≥ 12	No: when you were ≥ 12	$\geq$ | Median Penalty |	$<$ | Median Penalty |
$\mathbb{P}$(graduate): $MWL_{j=1}$	-0.54	-1.13***	-0.93**	-1.02**	-1.42**	-0.05
	(0.58)	(0.35)	(0.46)	(0.40)	(0.60)	(0.68)
Rank: $MWL_{j=1}$	0.24	-1.07***	-0.38	-0.96***	-1.09*	0.40
	(0.49)	(0.32)	(0.40)	(0.36)	(0.57)	(0.58)
Participants	339	708	494	547	261	253
Observations	2034	4248	2964	3282	1566	1518
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Scenario Income Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Standard errors are clustered on individuals. Each coefficient is obtained from separate OLS regressions estimating equation (2.2) on the subsample defined by the displayed label and for our two expected outcomes: (i) the probability for the child to graduate ($\mathbb{P}$(graduate)), and (ii) the earnings rank of the child at age 30 (Earnings Rank). "FT" stands for full-time, while "PT" stands for part-time. For participants' own-mothers' employment during the participant's childhood (own-mother's employment: < 12), and adolescence (own-mother's employment: ≥ 12), we drop observations that were listed as not applicable ($N = 16$). We also collected this information for their father's, but 93% had a full-time working father when they were less than 12, and 90% when they were adolescents. Thus, we do not report results due to small cell sizes for the part-time or less category. Finally, we calculated each woman's motherhood employment penalty relative to men. The median of this distribution is approxitamely a -18 percent drop in the likelihood of FT employment after childbirth. We split the sample by women with a high penalty (a $\geq$ absolute value of the median penalty) or a low penalty ($<$ absolute value of the median penalty).

B.2.2.3 Heterogeneity by design features.

We report heterogeneous effects for the MWL effects stratified by the randomized hypothetical design features in Table B.6. First, we observe stronger negative effects in scenarios with a boy (as opposed to scenarios with girls). The results, based on the gender of the vignette child, point toward different expectations on the needs of children and the role of mothers. While beyond the scope of this paper, this finding aligns with evidence that boys, especially in disadvantaged families, are more responsive to parental inputs (Autor et al., 2019; Bertrand and Pan, 2013; Lei and Lundberg, 2020). If boys are viewed as less resilient than girls and maternal time is perceived as 'higher-quality' time, then we should indeed expect stronger beliefs that their development and educational outcomes may suffer more from reduced parental input quality when mothers work longer.

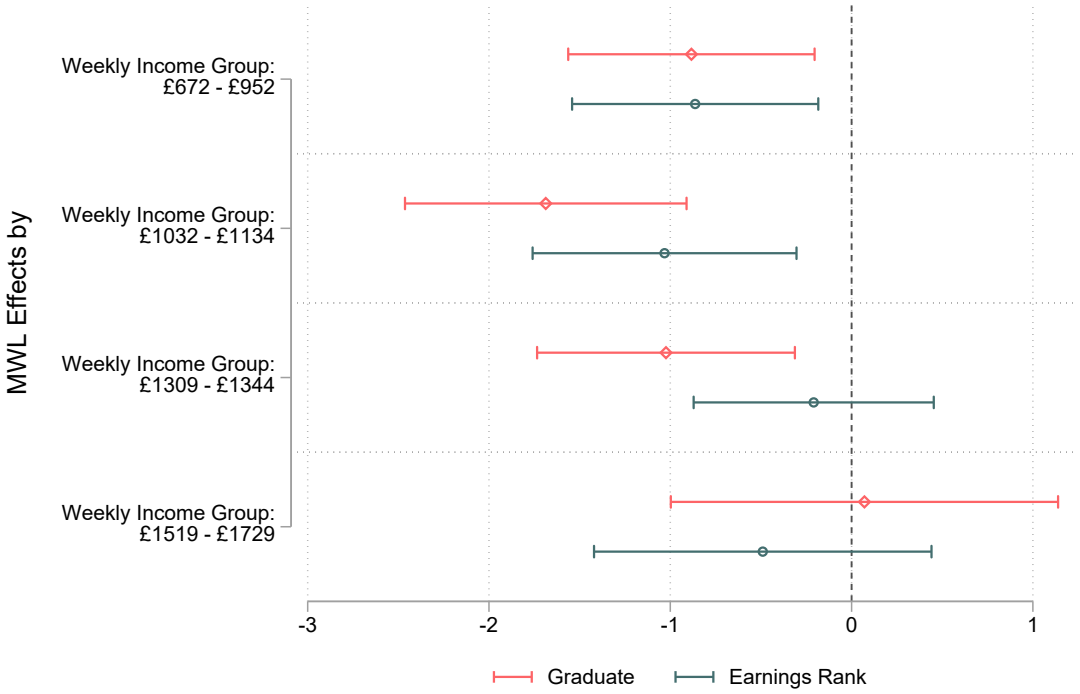
Second, the effects are stronger in scenarios with, on average, lower hypothetical household incomes and scenarios with a lower allocation of resources (10% instead of 20%) to the child's educational activities. We already showed, in Section 2.3.3, that in fact beliefs only fade out at quite high family incomes. Thus, these beliefs are not merely related to expectations on low income families. These results do suggest that beliefs on absolute advantage may place pressure on mothers who have less disposable resources to invest in the child (lower share of the budget allocated to the child) when deciding whether to continue careers.

Table B.6: Heterogeneity in Beliefs by Design Features

	Child's Gender		Child's Age		SSE		Working Hours Profile		Wage Profile		First Shown	
	Girl	Boy	4	10	SSE: 20%	SSE: 10%	FT–FT	FT–PT	Higher	Lower	Mother	Father
$\mathbb{P}$(graduate): $MWL_{j=1}$	-0.57	-1.28***	-1.11***	-0.75*	-0.53	-1.36***	-0.57	-1.30***	-0.45	-1.50***	-0.29	-1.59***
	(0.43)	(0.41)	(0.41)	(0.44)	(0.41)	(0.44)	(0.42)	(0.42)	(0.40)	(0.45)	(0.43)	(0.41)
Earnings Rank: $MWL_{j=1}$	-0.41	-0.91**	-0.66*	-0.68*	-0.59	-0.75*	-0.56	-0.78**	-0.34	-1.05**	-0.34	-1.00***
	(0.37)	(0.38)	(0.37)	(0.39)	(0.36)	(0.39)	(0.38)	(0.38)	(0.34)	(0.43)	(0.39)	(0.36)
Participants	517	539	542	514	540	516	533	523	567	489	532	524
Observations	3102	3234	3252	3084	3240	3096	3198	3138	3402	2934	3192	3144
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scenario Income Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Standard errors are clustered on individuals. We estimate equation (2.2) for each feature, *i.e.*, when the hypothetical child is a girl, a boy, is 4, etc. and report here the $\hat{\delta}$ associated with $MWL_{j=1}$ for each of those regressions. The FT–FT design presents both parents as full-time with one working longer hours (42 vs. 35), while the FT–PT design presents a full-time parent *versus* a part-time parent (36 vs. 20). The "first shown" column corresponds to the set of scenarios shown first to the participant — either $MWL = 1$ (mother works longer hours) or $MWL = 0$ (father works longer hours) in the beliefs elicitation survey.

Figure B.1: Heterogeneity by Hypothetical Family Income



Notes: Error bars represent 90% confidence intervals. We group scenarios by four equally sized bins of the weekly income for the hypothetical family.

Table B.7: Robustness — Beliefs about Mothers Working Longer

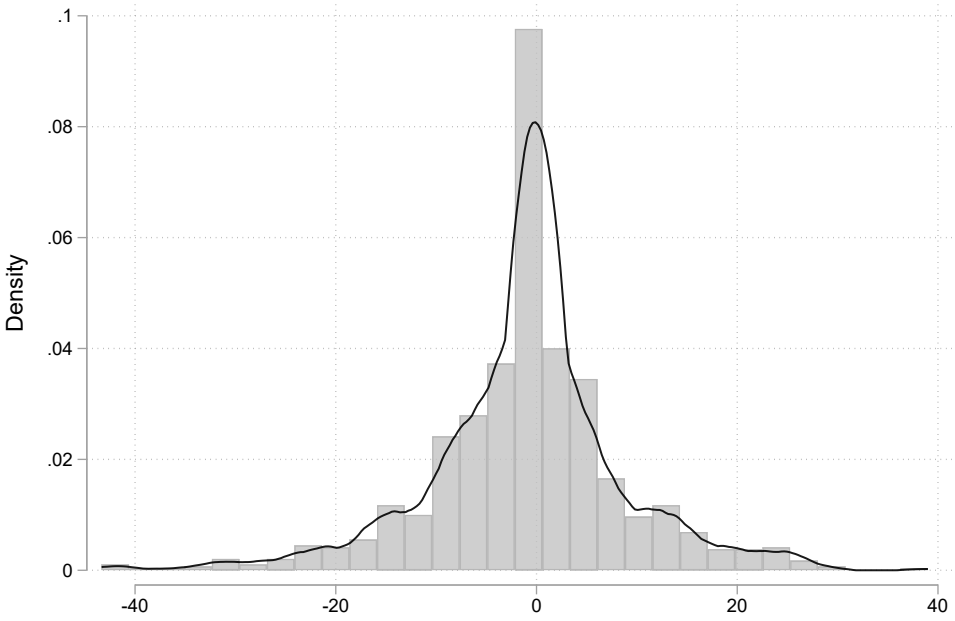
	(1) At Least Somewhat Certain		(2) Screener		(3) Response Time		(4) Re-Weighting	
	ℙ(graduate)	Earnings Rank	ℙ(graduate)	Earnings Rank	ℙ(graduate)	Earnings Rank	ℙ(graduate)	Earnings Rank
$MWL_{j=1}$	-1.230***	-0.870***	-0.717**	-0.556**	-0.925***	-0.736***	-0.788**	-0.840**
	(0.360)	(0.322)	(0.300)	(0.270)	(0.313)	(0.283)	(0.377)	(0.356)
Mean Dep. Var	56^{th}	47^{th}	56^{th}	52^{nd}	56^{th}	49^{th}	55^{th}	49^{th}
Participants	795	261	1003	53	950	950	1056	1056
Observations	4770	1566	6018	318	5700	5700	6336	6336
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scenario Income Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. OLS results for equation (2.2) — including individual and scenario income fixed effects — on the expected probability for the child to graduate (ℙ(graduate)), and the percentile ranking expected for the child at age 30 among other 30 year olds (Earnings Rank). Standard errors are clustered on individuals. We present results for four checks in four distinct blocks of columns. In (1), we restrict the sample to participants at least somewhat certain about their answer to the hypothetical scenarios, while in (2), we keep participants who passed the "turquoise" screener discussed in Subsection 2.3.2. In (3), we drop participants with the lowest and highest 5% response times, and in (4) we re-weight our sample to match the national population distribution (see Subsection B.2.1).

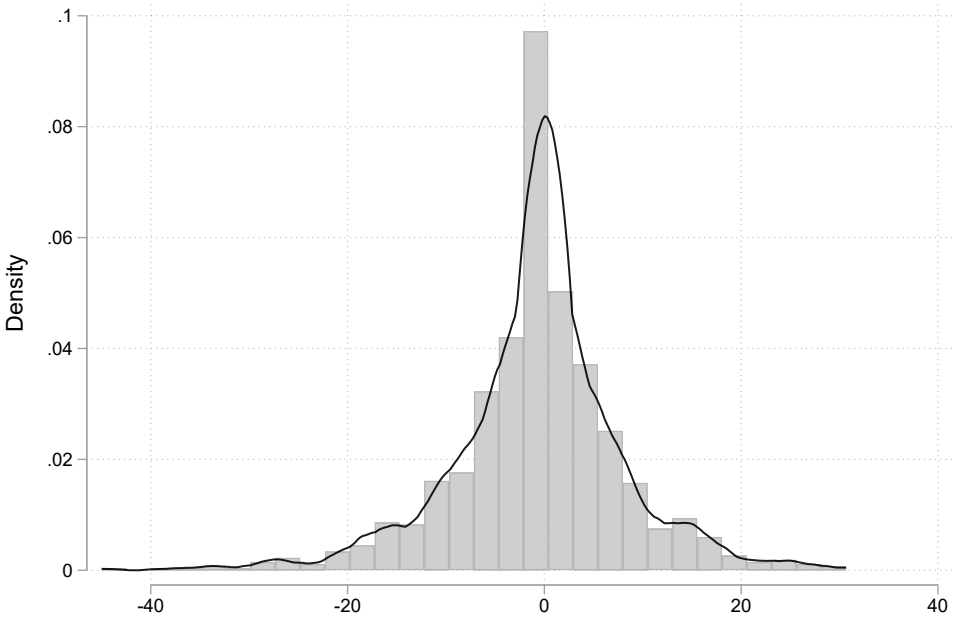
B.2.3 Individual Perceptions

Figure B.2: Distribution of Perceived Returns

(a) Perceived Returns: θ^{graduate}

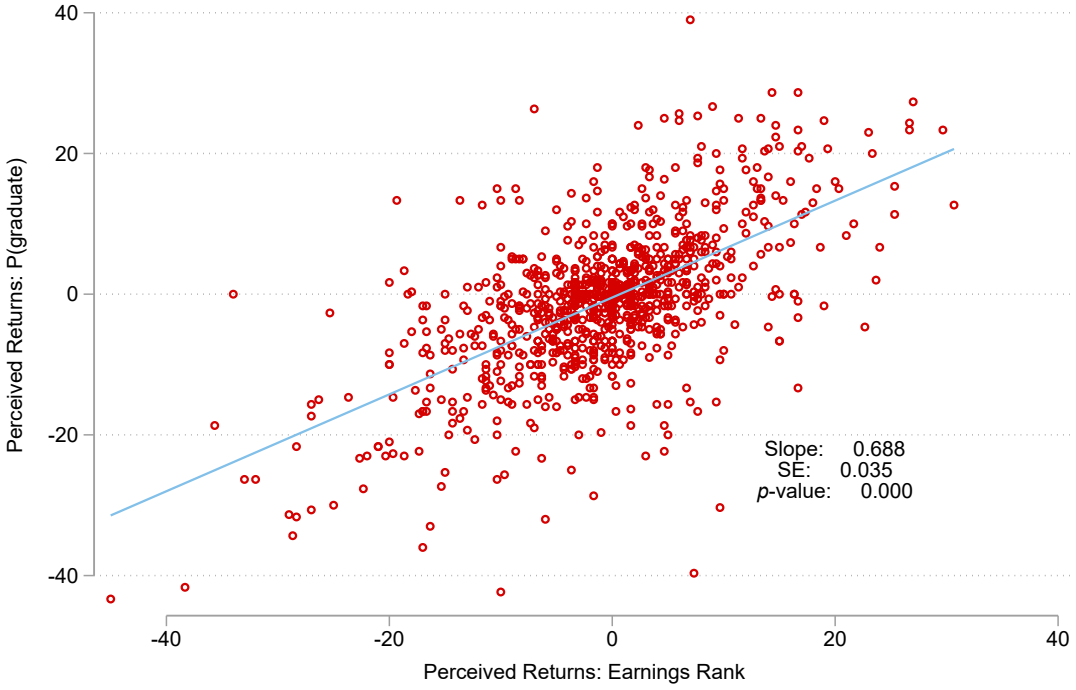


(b) Perceived Returns: θ^{rank}



Notes: $N = 249$. Distribution of our two perceived returns with kernel density plot over (a) the probability for the child to graduate from university (θ^{graduate}), and (b) the expected earnings rank of the child at age 30 (θ^{rank}).

Figure B.3: Relationship of Individual-Level Beliefs Measures



Notes: This figure presents a scatter plot with a line of best fit for our individual-level perceptions over the probability of graduating from university (θ^{graduate}) on the y-axis, and the earnings rank of the child at age 30 (θ^{rank}) one the x-axis.

Table B.8: Relationship Between Beliefs and Participants' Behavior: Time With Children

	Skills Time			Outdoor Time		
Panel A: Women	(1)	(2)	(3)	(1)	(2)	(3)
θ^{graduate}	-0.140			0.099		
	(0.137)			(0.119)		
θ^{rank}		-0.152			-0.016	
		(0.129)			(0.114)	
$\theta(IV)$			-0.218			0.063
			(0.140)			(0.118)
Observations	525	525	1050	525	525	1050
Mean Dep. Var.	4.97	4.97	4.97	4.94	4.94	4.94
Panel B: Men						
θ^{graduate}	0.300**			0.084		
	(0.135)			(0.131)		
θ^{rank}		0.198			0.031	
		(0.139)			(0.123)	
$\theta(IV)$			0.450**			0.104
			(0.181)			(0.163)
Observations	525	525	1050	525	525	1050
Mean Dep. Var	4.20	4.20	4.20	4.32	4.32	4.32

Notes: p <0.10; ** p <0.05, *** p <0.01. Standard errors, in parentheses, are bootstrapped with 1,000 replications. We run OLS regressions on three main outcomes reflecting participant's behavior with their child(ren) and on the labor market for the control group only. Specifically, these regressions look at the associations between the participant's behavior and (1) their perceived returns over the probability for the child to graduate (θ^{graduate}), (2) the expected earnings rank of the child at age 30 (θ^{rank}), and (3) both dimensions. Skills Time (resp., Outdoor Time) corresponds to the number of hours spent per day by the participant helping their child(ren) develop their skills (resp., doing outdoor activities with them). ln(Hours Worked) corresponds to the log of the participant's weekly number of hours worked. Individuals controls include participant's gender (for Panel A), a quadratic in age, an indicator for whether they have at least a university degree, employment status (full-time *versus* part-time or less), and ethnicity (white *versus* non-white). Six participants listed "other" or "prefer not to say" for gender. We code these as 0 and control for an indicator flagging them in Panel A, and exclude them in Panels B and C. The number of participants does not vary between panels as those six participants belong to the treatment group.

B.2.4 Mechanisms for Variation in Beliefs

Table B.9: Outcomes Descriptive Statistics

	Mean	SD	Min	Max	N
Minutes spent helping the child					
Prepare for the test	149.24	116.05	0.00	600.00	1056
Doing extracurricular activities	161.06	112.01	0.00	600.00	1056
Expected rank at the test	0.42	0.20	0.01	0.99	1056
Mother: $\mathbb{P}$(University Graduate)	0.51	0.22	0.00	1.00	1056
Father: $\mathbb{P}$(University Graduate)	0.46	0.22	0.00	1.00	1056
Share of income spent on extracurricular activities	0.18	0.15	0.00	0.97	1056

Notes: This table displays descriptive statistics for the outcomes collected to investigate the mechanisms in variation in beliefs, introduced in Subsection 2.3.5.

Table B.10: Expectations on Resource Allocation and Parental Education by Elicited Beliefs

	(1)	(2)
	By θ	
	< 0	≥ 0
Panel A: Resource Allocation		
Mother Earns More	0.006	0.008
	(0.013)	(0.012)
Panel B: $\mathbb{P}$(University Graduate)	(1)	(2)
Difference (Mother $-$ Father)		
Works Full-Time	0.066	-0.001
	(0.034)	(0.032)
Participants	517	539
Individual Controls	Yes	Yes

Notes: p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. All specifications include controls for the pre-registered set of participants' characteristics.

B.3 Information Treatment: Additional Results

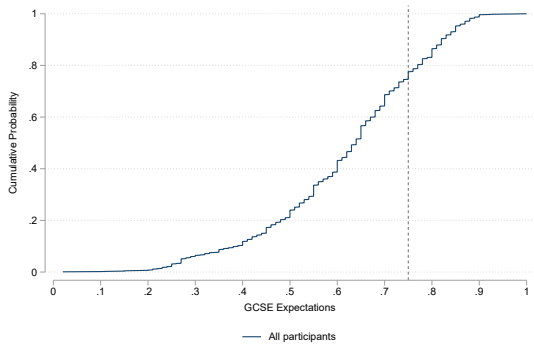
B.3.1 Belief Updating

Heterogeneity in beliefs about children's skills.

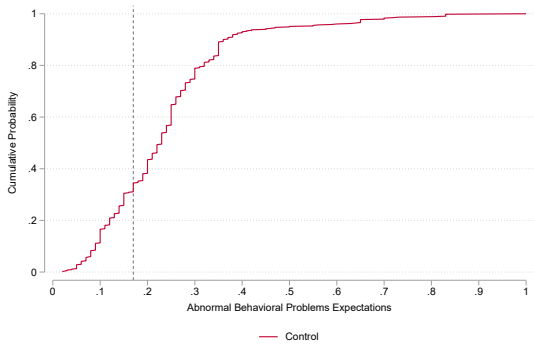
We look at how our measures of beliefs vary with the posterior beliefs (expectations on behaviors) and gender norms. We demonstrate, in the Appendix, Figure B.4, that the cumulative distributions of GCSE pass rates and the posterior expectations have a high degree of variation for estimating our information treatment effect. Next, using the control group only and conditional on participant characteristics, we show that all of our beliefs measures ($GCSE_i$, $\theta_i^{\text{graduate}}$, θ_i^{rank}) associate with expectations on the share of children with an abnormal level of behavioral problems (Appendix, Figure B.5) and self-reported gender norms (Appendix, Figure B.6). Overall, these beliefs measures strongly relate to behavioral expectations and self-reported liberal norms, suggesting they can serve as priors to help understand either homogeneous or heterogeneous responses to information. Finally, we find that expectations on GCSE pass rates only partially, but positively, associate with $\theta_i^{\text{graduate}}$ and θ_i^{rank} (Appendix, Figure B.7). Thus, as we discussed above, these GCSE expectations are weakly correlated with the θ_i^o measures, which remove differences in earnings potential between a mother and a father. The expectations on GCSE pass rates may capture a wider range of misinformation, uncertainty, or other dimensions, since earnings potential are not fixed and the question is asked in a more open way. In our next assessment of the information treatment effects, this gives us a useful way to distinguish belief updating around different dimensions in the priors.

Figure B.4: CDF of Incentivized Beliefs

(a) GCSE Pass Rates



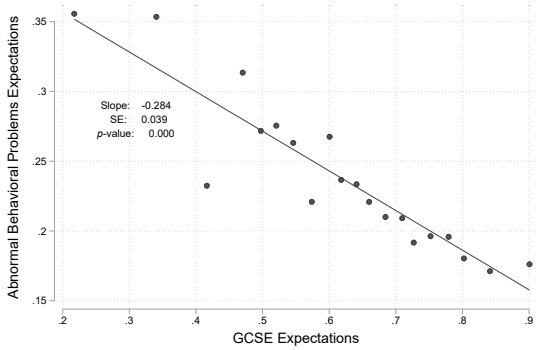
(b) Behavioral Problems



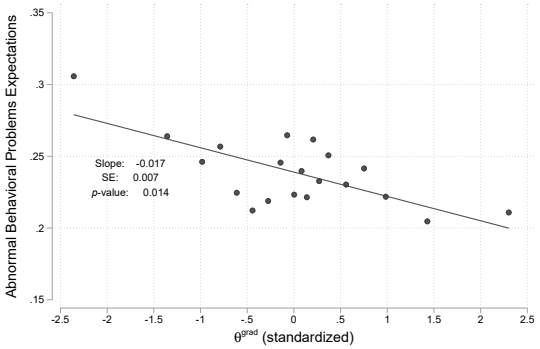
Notes: Figure (a) shows the cumulative distribution function for beliefs about the passing rate of 5 or more GCSEs (with at least C/4) in families where mothers work 35 hours or more per week. Figure (b) also reports the cumulative distribution function, but for beliefs about the share of children having an abnormal level of behavioral problems when the mother works full-time, using only control group respondents. In both panels, the short-dashed lines respectively indicate the true levels of (a) for GCSE pass rates (75%), and (b) behavioral problems (17%).

Figure B.5: Associations of Beliefs Measures and the Posterior

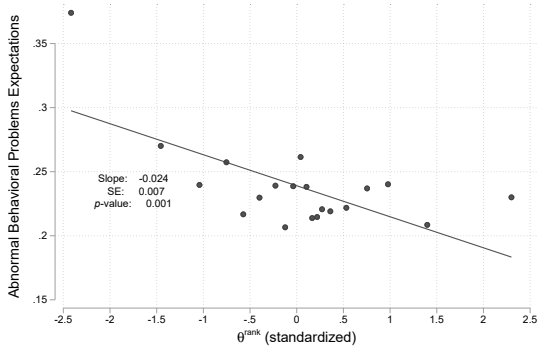
(a) Expectations on GCSE pass rates



(b) Expectations on $\mathbb{P}$(graduate)



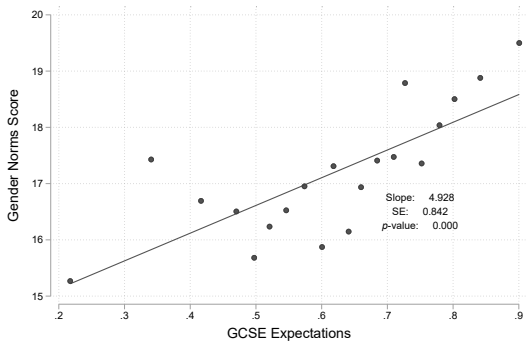
(c) Expectations on Earnings Rank



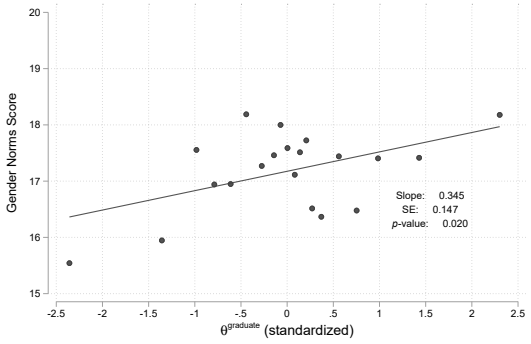
Notes: These figures present binscatter plots to highlight the relationships between the posterior belief, expectations on abnormal behavioral problems when mothers work full-time, and (a) the expectation for GCSE pass rates when mothers work full-time ($GCSE_i$), (b) expectations on the child's probability to graduate from university on the expected probability for the child to graduate ($\theta_i^{\text{graduate}}$), and (c) the expected earnings rank of the child at age 30 (θ_i^{rank}). In (b) and (c) we standardize the θ_i^o measures around 0 with an standard deviation of 1. We always control for participant characteristics, and we only use the information treatment control group for this analysis ($N = 531$).

Figure B.6: Associations of Beliefs Measures and Gender Norms

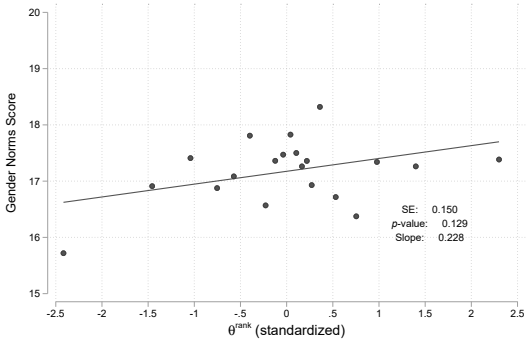
(a) Expectations on GCSE pass rates



(b) Expectations on $\mathbb{P}$(graduate)



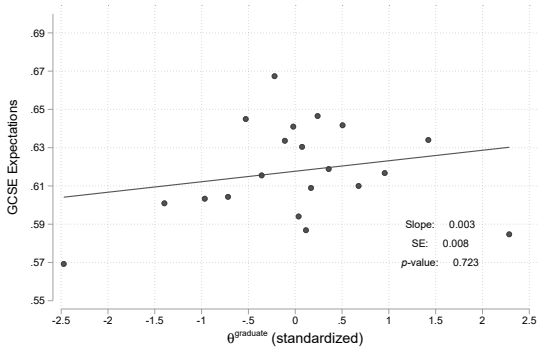
(c) Expectations on Earnings Rank



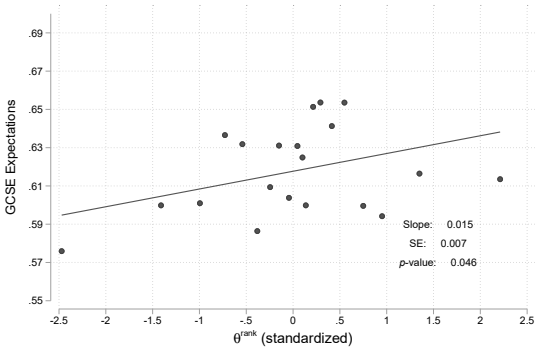
Notes: These figures present binscatter plots to highlight the relationships between gender norms and (a) the expectation for GCSE pass rates when mothers work full-time ($GCSE_i$), (b) expectations on the child's probability to graduate from university on the expected probability for the child to graduate ($\theta_i^{\text{graduate}}$), and (c) the expected earnings rank of the child at age 30 (θ_i^{rank}). These are for the control group only ($N = 531$). In (b) and (c), we standardize the θ_i^o measures around 0, and with a standard deviation of 1. All specifications include controls for the pre-registered set of participants' characteristics.

Figure B.7: Associations of GCSE Expectations and Elicited Beliefs

(a) Expectations on ℙ(graduate)



(b) Expectations on Earnings Rank



Notes: These figures present binscatter plots to highlight the relationships between the the expectation for GCSE pass rates when mothers work full-time ($GCSE_i$) and (a) expectations on the hypothetical child's probability to graduate from university ($\theta_i^{\text{graduate}}$), and (b) expectations on the earnings rank of the child at age 30 (θ_i^{rank}), for the control group only ($N = 531$). We standardize the θ_i^o measures around 0, and with a standard deviation of 1. All specifications include controls for the pre-registered set of participants' characteristics.

Table B.11: Robustness: Belief Updating and Information Effects

	(1) At Least Somewhat Certain	(2) Screener	(3) Response Time	(4) Risk of Demand Effect	(5) Unclear Mechanism	(6) Re-Weighting	(7) PDS Lasso
Panel A: Incentivized beliefs							
Treatment	-0.053***	-0.052***	-0.059***	-0.047***	-0.049***	-0.050***	-0.056***
	(0.009)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)
Mean Dep. Var	0.217	0.209	0.209	0.215	0.211	0.208	0.213
Panel B: Open Q: harmful/not harmful							
Treatment	-0.211***	-0.220***	-0.235***	-0.189***	-0.219***	-0.221***	-0.219***
	(0.036)	(0.032)	(0.033)	(0.036)	(0.031)	(0.040)	(0.033)
Mean Dep. Var	0.630	0.607	0.614	0.610	0.611	0.620	0.611
Panel C: Gender norms							
Treatment	0.609**	0.486**	0.514**	0.399*	0.569**	0.596**	0.565***
	(0.240)	(0.213)	(0.219)	(0.228)	(0.223)	(0.244)	(0.208)
Mean Dep. Var	17.382	17.513	17.518	17.295	17.395	17.150	0.611
Participants	795	1003	950	796	915	1056	1056
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. This table presents OLS results of equation (2.4) for each outcome listed in the panels. In columns (1) and (2) respectively, we keep participants at least somewhat certain of their answers on the hypothetical vignettes, and who passed the "turquoise" screener. In the next columns, we drop participants (3) with the 5% lowest and highest response times, (4) more at risk of a demand effect, and (5) those whose answer was unclear to the open-ended question on what guided their answer about children with abnormal level of behavioral problems. Risk of demand effect was coded from an open question at end of the survey where respondents were asked what they thought the survey was about. We classify answers that allude to perceptions about mothers working and children as "at risk of a demand effect" and drop those (column 4) from the analysis. Finally, in (6), we re-weight our sample to match the national population distribution (see Subsection B.2.1), and in (7), we perform a PDS Lasso approach including all of participants' demographics for the control variables selection (Belloni, Chernozhukov, and Hansen, 2014).

Table B.12: Belief Updating and the Perception Gap by Prior Beliefs

	(1) All Participants	(2) By θ	
		< 0	≥ 0
Perception Gap × Treatment	0.002***	0.003***	0.001*
	(0.001)	(0.001)	(0.001)
Perception Gap	-0.003***	-0.003***	-0.003***
	(0.000)	(0.001)	(0.001)
Treatment	-0.024***	-0.025**	-0.022**
	(0.008)	(0.012)	(0.011)
Mean Dep. Var†	0.213	0.218	0.207
Participants	1056	517	539
Individual Controls	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. †The mean of the dependent variable is calculated for the control group only. This table shows OLS results of equation (2.5), where the outcome is the incentivized beliefs on the share of children out of 100 scoring above the threshold for abnormal behavioral problems when mothers work full-time. We rescaled this to lie between 0 and 1 for interpretation purposes. Results are presented (1) for the full sample of participants, (2) by negative and positive values of θ. Individual controls follow our pre-registered set defined previously; however, this heterogeneity analysis here by prior beliefs was not pre-registered.

Table B.13: Heterogeneity in Information Treatment Effects on Incentivized Beliefs by Participant's Characteristics

	Median Age		Gender		Ethnicity		Born in the UK		University Degree		Working Hours		Household Income		Number of Children		Voted at last UK General Election		
	≥ 38	< 38	Female	Male	White	Non-white	Yes	No	Yes	No	FT	PT	$\geq$ Median	$<$ Median	< 2	≥ 2	Conservative	Liberal	Other or None
Treatment	-0.06***	-0.04***	-0.06***	-0.04***	-0.04***	-0.09***	-0.04***	-0.10***	-0.05***	-0.05***	-0.04***	-0.07***	-0.07***	-0.05***	-0.04***	-0.06***	-0.06***	-0.05***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Participants	563	493	525	525	840	216	871	185	692	364	657	399	261	253	613	443	194	679	183
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. We estimate separately equation (2.4) on the participant's posterior beliefs (about the share of children having an abnormal level of behavioral problems in families where the mother works full-time), for each participant's characteristic, *i.e.*, for when the participant is a man, a woman, has a degree, etc. We display here $\hat{\gamma}$ associated with D_i for each of these regressions. Estimates for other genders than male and female not displayed due to sample size issues. The outcome is rescaled to lie between 0 and 1. For the columns "Vote at last UK General Elections", we asked participants: "Which party did you choose as your primary vote in the last UK General Election?" and provided them with a list of candidate parties. We condensed information into three categories as follows: Conservative Party and Reform UK; labor, Liberal Democrats, and Green Party; and other or none.

Table B.14: Heterogeneity in Information Treatment Effects on Incentivized Beliefs by Participant's Past Experiences

	Own-Mothers' FT Employment				By Median Birth Penalty	
	Yes: when you were <12	No: when you were <12	Yes: when you were ≥ 12	No: when you were ≥ 12	$\geq \lvert$ Median Penalty $\rvert$	$< \lvert$ Median Penalty $\rvert$
Treatment	-0.06***	-0.05***	-0.05***	-0.06***	-0.07***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Participants	339	708	494	547	261	253
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. Robust standard errors in parentheses. We estimate separately equation (2.4) on the participant's posterior beliefs about the share of children having an abnormal level of behavioral problems in families where the mother works full-time, for each participant's past experience, *i.e.*, if they had a full-time working mother when aged <12, if they did not, etc., and we display here $\hat{\gamma}$ associated with D_i for each of these regressions.

Table B.15: Causal forest: heterogeneity in the CATES by individual characteristics

	(1) High Predicted CATE	(2) Low Predicted CATE	(3) Diff.	(4) p-value
FT working mother when age <12	0.326	0.322	0.004	0.897
FT working mother when age ≥ 12	0.474	0.475	-0.001	0.971
High motherhood penalty	0.532	0.482	0.050	0.256
Age	38.472	38.456	0.015	0.973
Age^2	1532.521	1528.407	4.114	0.908
Woman	0.509	0.485	0.025	0.424
White	0.773	0.818	-0.045	0.067
Born in the UK	0.814	0.835	-0.021	0.374
University graduate	0.644	0.667	-0.023	0.438
FT employment	0.604	0.640	-0.036	0.228
ln(Household income)	8.235	8.277	-0.042	0.092
Number of children	1.716	1.847	-0.131	0.015
Vote: conservative	0.189	0.178	0.011	0.634
Vote: other or none	0.188	0.159	0.028	0.223
Median CATEs		-0.053		
Observations	528	528	1,056	

Notes: We report summary statistics as the mean for each participants' characteristic split by those above or below the median of the conditional average treatment effect (CATEs) estimated via a Causal Forest. We also report the difference between the means in column (3) and p-values in column (4).

B.3.2 Policy Support

Table B.16: Heterogeneity in Information Treatment Effects on Policy Support by Participant's Characteristics

	Median Age		Gender		Ethnicity		Born in the UK		University Degree		Working Hours		Participant's Income		Number of Children		Vote at last UK General Election		
	≥ 38	< 38	Female	Male	White	Non-white	Yes	No	Yes	No	FT	PT or none	$\geq$ Median	$<$ Median	< 2	≥ 2	Conservative	Liberal	Other or None
Panel A: Subsidized childcare policies																			
Treatment	-0.00	-0.00	0.01	-0.02	-0.02	0.04	0.00	-0.02	0.01	-0.03	-0.02	0.02	-0.01	0.04	-0.02	0.01	-0.00	0.01	-0.06
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)	(0.03)	(0.06)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.05)	(0.03)	(0.04)	(0.06)	(0.03)	(0.07)
Panel B: Paternity leave policies																			
Treatment	0.03	0.05	0.08*	0.01	0.03	0.09	0.03	0.09	0.05	0.02	0.02	0.09*	0.06	0.10*	0.03	0.04	0.08	0.04	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.06)	(0.03)	(0.07)	(0.03)	(0.05)	(0.04)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.07)	(0.04)	(0.07)
Participants	563	493	525	525	840	216	871	185	692	364	657	399	261	253	613	443	194	679	183
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. We estimate separately equation 2.4 on two collected outcomes for policy support in distinct panel, and for each participant's characteristic, *i.e.*, for when the participant is a man, a woman, has a degree, etc. We display here $\hat{\gamma}$ associated with D_i for each of these regressions. For panels A and B respectively, we construct binary variables set to one, indicating a strong level of agreement to subsidized childcare policies and paternity leave policies. We define strong by being above the median response as defined in the pre-registration. Individuals controls follow our previously defined set. Estimates for other genders than male and female not displayed due to sample size issues. For the columns "Vote at last UK General Elections", we asked participants: "Which party did you choose as your primary vote in the last UK General Election?" and provided them with a list of candidate parties. We condensed information into three categories as follows: Conservative Party and Reform UK; labor, Liberal Democrats, and Green Party; and other or none.

Table B.17: Heterogeneity in Information Treatment Effects on Policy Support by Participant's Past Experiences

	Own-Mothers' FT Employment				By Median Birth Penalty	
	Yes: when you were < 12	No: when you were < 12	Yes: when you were ≥ 12	No: when you were ≥ 12	$\geq \mid$ Median Penalty $\mid$	$< \mid$ Median Penalty $\mid$
Panel A: Subsidized childcare policies						
Treatment	0.05	-0.02	0.03	-0.02	-0.01	0.04
	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.05)
Panel B: Paternity leave policies						
Treatment	0.09*	0.03	0.08*	0.02	0.06	0.10*
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)
Participants	339	708	494	547	261	253
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. We estimate separately equation (2.4) on two outcomes for policy support, and for each participant's past experience, *i.e.*, if they had a full-time working mother when aged < 12, if they did not, etc., and we display here $\hat{\gamma}$ associated with D_i for each of these regressions. Specifically, panel A displays results for a strong level of agreement with subsidized childcare policies, while panel B displays results for a strong level of agreement with paternity leave policies. We define strong by being above the median response as defined in the pre-registration. Individuals controls follow our previously defined set.

B.3.2.1 Follow-up Survey

Policy support: follow-up survey.

Here, we discuss our analyses based on the one-week follow-up. The outcomes are a continuous measure on the number of free childcare hours supported and two binary support measures on childcare policy and the conversion of existing primary schools into school-based nurseries. Results are in Tables B.19, B.20, and B.21. In Table B.19, we find no significant effects. We do find, however, that for the number of free childcare hours the information effect is suggestive of a positive effect. Next, in Table B.20, we find this information effect on free childcare hours is larger and significant for the same groups we saw suggestive evidence of policy support in the main survey. Information significantly increases the supported number of free childcare hours among women (1.7 hours higher), those born outside the UK (3.53 hours higher), and university degree holders (1.65 hours higher).[3] Strikingly, some of these groups were among those in Section 2.3 for whom we saw strong beliefs of absolute advantage.[4] Again, we interpret this as pointing toward the ability of information to alleviate uncertainty for participants on the production of children's skills when mothers work.

We further look at heterogeneity on the binary support measures. Again, we see mainly null results. For support on converting existing space in primary schools to new nurseries, we even see some suggestive negative information effects, though none are significant. This measure, though, combines changing primary school space with nursery space potentially obfuscating views on support for mothers to work. This seems likely the case given our positive results on support for more free childcare hours. We also look at heterogeneity by participants' past experiences in Table B.21 with generally insignificant results but again a pattern of positive information effects on the number of free childcare hours. Overall, we must emphasize that our results on policy support are not strong. Nevertheless, they generally suggest information nudges policy views toward more support for women to be able to work.

[3] We also see insignificant but similar sized effects among part-time workers and those with lower income as we did in the main survey.

[4] Per our pre-registration, we also associated our measures of absolute advantage with these policy support measures at the follow-up using the information control group (Table B.18). The sample size with the control group is too small to reliably pick up links to these policy support measures where variation can already be limited. We do see some positive link (θ_i^{rank} specifically, with the number of free childcare hours but do not interpret these strongly).

Table B.18: Follow-Up – Associations Between Elicited Beliefs and Policy Support

	(1) Free Childcare Hours			(2) Childcare Policies			(3) Nursery Policies		
$\theta^{graduate}$	-0.017		-0.082	0.000		-0.002	0.000		-0.001
	(0.046)		(0.059)	(0.002)		(0.003)	(0.002)		(0.003)
θ^{rank}		0.074	0.129**		0.004	0.005		0.001	0.002
		(0.052)	(0.064)		(0.003)	(0.003)		(0.003)	(0.003)
Participants	445	445	445	445	445	445	445	445	445
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. This table presents OLS results for the control group (main survey) and for three outcomes collected during the follow-up survey: (1) a continuous measure for the number of hours of free childcare participants think parents should receive, (2) a binary indicator for a strong level of agreement (above the median) to expanding free childcare for parents earning less than £60,000, and (3) a binary indicator for a strong level of agreement to increasing childcare supply through the conversion of existing primary schools into 'school-based' nurseries. Robust standard errors in parentheses. Individuals controls follow our previously defined set.

Table B.19: Follow Up — Information Treatment Effects on Policy Support

	(1) All Participants	(2) By GCSE Beliefs		(3) By θ	
		Under-	Over-	< 0	≥ 0
Panel A: Free childcare hours					
Treatment	0.539	0.880	-0.491	1.038	0.022
	(0.681)	(0.788)	(1.377)	(0.987)	(0.931)
Difference: p-value		0.381		0.449	
Mean Dep. Var	30.015	29.921	30.292	29.855	30.166
Panel B: Childcare policies					
Treatment	-0.021	-0.040	0.035	-0.009	-0.038
	(0.033)	(0.038)	(0.066)	(0.047)	(0.046)
Difference: p-value		0.320		0.646	
Mean Dep. Var	0.548	0.538	0.575	0.521	0.573
Panel C: Nursery policies					
Treatment	-0.028	-0.059*	0.045	0.013	-0.061
	(0.031)	(0.035)	(0.061)	(0.043)	(0.044)
Difference: p-value		0.131		0.224	
Mean Dep. Var	0.699	0.711	0.664	0.714	0.684
Participants	893	667	226	434	459
Individual Controls	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. This table presents OLS results of equation (2.4) for three outcomes presented in separate panels. Panel A presents results for a continuous measure of the number of hours of free childcare participants think parents should receive, while panels B and C present results for separate binary variables set to one if the participant indicated a strong level of agreement to expanding free childcare for parents earning less than £60,000 (panel B), and a strong level of agreement to increasing childcare supply through the conversion of existing primary schools into 'school-based' nurseries (panel C). All specifications include controls for the pre-registered set of participants' characteristics.

Table B.20: Follow-Up — Heterogeneity in Information Treatment Effects on Policy Support by Participant's Characteristics

	Median Age		Gender		Ethnicity		Born in the UK		University Degree		Working Hours		Participant's Income		Number of Children		Vote at last UK General Election		
	≥ 38	< 38	Female	Male	White	Non-white	Yes	No	Yes	No	FT	PT or none	$\geq$ Median	$<$ Median	< 2	≥ 2	Conservative	Liberal	Other or None
Panel A: Free childcare hours																			
Treatment	-0.48	1.56*	1.72**	-0.47	0.14	2.31	0.01	3.53**	1.65**	-1.31	0.03	1.55	1.73	1.83	0.05	1.12	-1.68	0.83	1.58
	(0.97)	(0.95)	(0.87)	(1.03)	(0.74)	(1.62)	(0.74)	(1.65)	(0.83)	(1.17)	(0.86)	(1.12)	(1.29)	(1.25)	(0.87)	(1.09)	(1.61)	(0.81)	(1.87)
Panel B: Childcare policies																			
Treatment	-0.04	0.00	0.03	-0.06	-0.03	0.02	-0.04	0.08	0.01	-0.07	-0.05	0.03	0.04	0.01	0.00	-0.05	0.07	-0.03	-0.13
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.07)	(0.04)	(0.08)	(0.04)	(0.05)	(0.04)	(0.05)	(0.07)	(0.07)	(0.04)	(0.05)	(0.07)	(0.04)	(0.08)
Panel C: Nursery policies																			
Treatment	-0.06	0.01	-0.06	0.01	-0.04	0.04	-0.03	-0.01	0.01	-0.09*	0.00	-0.07	-0.09	-0.06	-0.06	0.01	-0.09	-0.02	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.07)	(0.03)	(0.08)	(0.04)	(0.05)	(0.04)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.07)	(0.04)	(0.08)
Participants	480	413	436	452	703	190	733	160	589	304	561	332	213	213	516	377	173	570	150
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p <0.10; ** p <0.05, *** p <0.01. Robust standard errors in parentheses. We estimate separately equation 2.4 on three collected outcomes from the follow-up for policy support in distinct panel, and for each participant's characteristic, *i.e.*, for when the participant is a man, a woman, has a degree, etc. We display here $\hat{\gamma}$ associated with D_i for each of these regressions. Panel A presents results for a continuous measure of the number of hours of free childcare participants think parents should receive, while panels B and C present results for separate binary variables set to one if the participant indicated a strong level of agreement to expanding free childcare for parents earnings less than £60,000 (panel B), and to increasing childcare supply through the conversion of existing primary schools into 'school-based' nurseries (panel C). Individuals controls follow our previously defined set. Estimates for other genders than male and female not displayed due to sample size issues. For the columns "Vote at last UK General Elections", we asked participants: "Which party did you choose as your primary vote in the last UK General Election?" and provided them with a list of candidate parties. We condensed information into three categories as follows: Conservative Party and Reform UK; labor, Liberal Democrats, and Green Party; and other or none.

Table B.21: Follow-Up — Heterogeneity in Information Treatment Effects on Policy Support by Participant's Past Experiences

	Own-Mothers' FT Employment				By Median Birth Penalty	
	Yes: when you were <12	No: when you were <12	Yes: when you were ≥ 12	No: when you were ≥ 12	$\geq$ | Median Penalty |	$<$ | Median Penalty |
Panel A: Free childcare hours						
Treatment	1.35	0.23	1.17	0.23	1.73	1.83
	(1.35)	(0.80)	(1.03)	(0.93)	(1.29)	(1.25)
Panel B: Childcare policies						
Treatment	-0.07	0.00	-0.01	-0.02	0.04	0.01
	(0.06)	(0.04)	(0.05)	(0.05)	(0.07)	(0.07)
Panel C: Nursery policies						
Treatment	-0.04	-0.02	-0.09**	0.04	-0.09	-0.06
	(0.06)	(0.04)	(0.05)	(0.04)	(0.06)	(0.06)
Participants	279	607	410	470	213	213
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

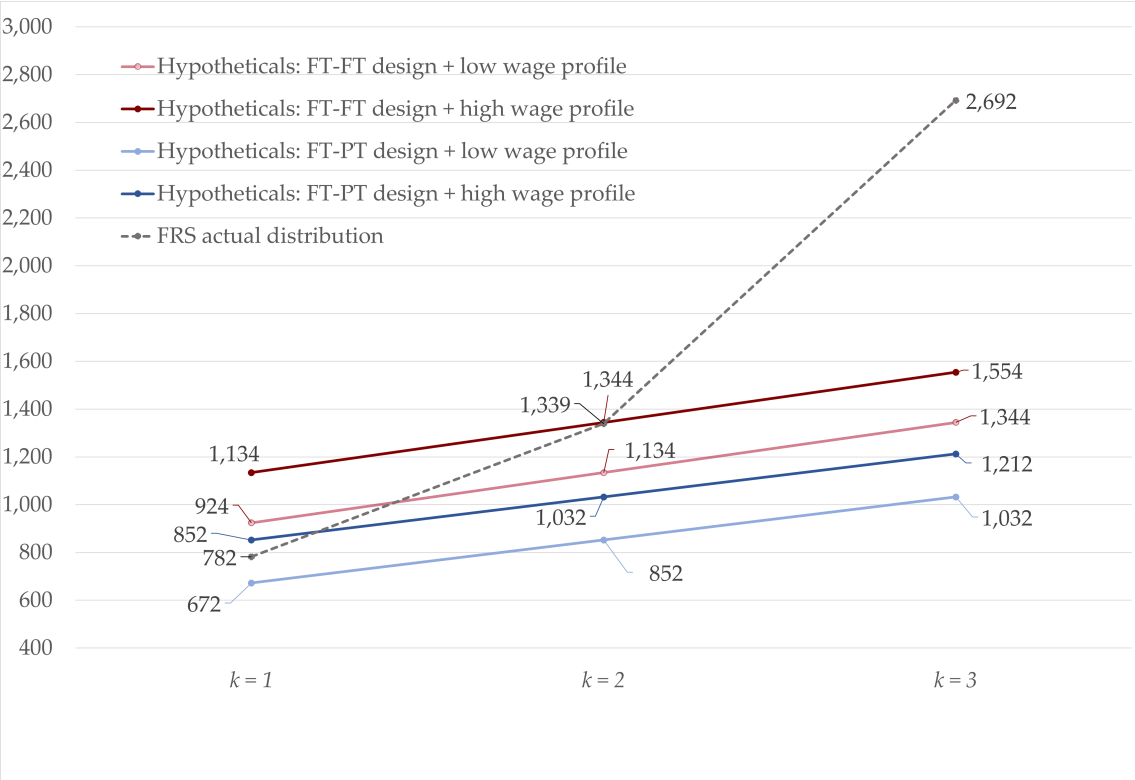
Notes: * $p<0.10$; ** $p<0.05$, *** $p<0.01$. Robust standard errors in parentheses. We estimate separately equation (2.4) on three outcomes for policy support from the follow-up survey, and for each participant's past experience, *i.e.*, if they had a full-time working mother when aged <12, if they did not, etc., and we display here $\hat{\gamma}$ associated with D_i for each of these regressions. Specifically, for panels A and B, we construct binary variables set to one, indicating a strong level of agreement to subsidized childcare policies and paternity leave policies. We define strong by being above the median response as defined in the pre-registration. For panel C, we use the continuous gender norms score where higher values imply more liberal views toward the role of mothers. Individuals controls follow our previously defined set.

B.4 Survey Technical Details and Questionnaires

B.4.1 Survey Technical Details

B.4.1.1 Beliefs Elicitation Survey

Figure B.8: Weekly Household Labor Income Distribution (£)



Notes: This table presents the distribution of weekly household labor income across the hypothetical wage profiles introduced in Table 2.1, contrasted with tertiles of the weekly household income distribution (dashed-line), drawn from the Family and Resources Survey (FRS) 2022-2023. To be consistent with our survey sample selection, we restrict the FRS data to dual-parent families in England in which both parents work.

B.4.1.2 Information Treatment Construction

We present below details on how we build the information treatment *i.e.*, the share of children passing five or more GCSEs with a grade C/4 or higher, as well as the share of children having an abnormal level of behavioral problems. For both metrics, we use the Millennium Cohort Study (MCS) also described below.

The Millennium Cohort Study.

The Millennium Cohort Study (MCS) follows the lives of around 19,000 young people ($N = 18,818$) born across England, Scotland, Wales and Northern Ireland in 2000-02. The MCS offers a wide range of measures tracking the cohort members' physical, socio-emotional, cognitive, and behavioral development over time, along

with detailed insights into their daily lives, behaviors, and experiences. Additionally, it provides comprehensive data on economic conditions, parenting practices, relationships, and family life, as reported by both resident parents. [5] For our analysis, we use the parents' reported data on various factors, such as their education levels, employment statuses, and weekly working hours, as well as the joint net household income. We then combine this data with relevant information about their child(ren), specifically focusing on non-cognitive outcomes and later GCSE pass rates.

Data management.

We construct two different datasets, respectively, for the two metrics we want to construct: (i) the share of children having an abnormal level of behavioral problems, and (ii) the share of children passing five or more GCSEs with a grade C/4 or higher. This leads us to work with different waves of MCS. For the first metrics (i), we make use of sweeps 1 (9 months of the child), 3 (age 5), and 4 (age 7). Specifically, we use sweep 1 to get information on parents' highest educational achievement. The final education variable we consider is a 3-category variable (High, Medium and Low), as presented in Table B.22 below. We extract the mother's working hours from sweeps 3 and 4 (ages 5 and 7) corresponding to the child's primary school age, and use sweep 4 to get infomation on parental income.[6] This wave also corresponds to the time of the SDQ measurement (see next paragraph for more details). Thus, we merge information reported from sweeps 1, 3 and 4, keeping England only, as well as dual-parents families.[7] This leaves us with a sample of $N = 6,787$ children. For the second metrics *i.e.*, the information treatment (ii), we make use of the same waves as above but also include sweeps 7 (age 17), when the GCSEs outcomes are measured. In turn, we merge parents' reported information from sweeps 1, 3, 4 and 7. Keeping only England, as well as dual-parents families, we end up with a sample of $N = 5,457$ children.

Metrics.

The two metrics have been constructed at the individual (child) level. First, to obtain the share of children having an abnormal level of behavioral problems, we mainly use wave 4 (age 7) in which parents respond to the Strengths and Difficulties Questionnaire (SDQ). Specifically, we focus on the externalizing (behavior) score — ranging from 0 to 20 — corresponding to the sum of the conduct and hyperactivity scales.[8] We calculate this score for each child aged 7, and further create a dummy variable equal to 1 if this score is equal or greater than 11. Indeed, children with scores between 11 and 20 are considered as having an "abnormal" level of behavioral problems.[9] Second, to obtain the share of 5 or more GCSEs passed with a grade of C/4

[5] See the MCS website for a more detailed description of the survey.

[6] The income variable provided by MCS is a 19-category variable, ranging from less than £1,600 a year, to £100,000 or more.

[7] We do not consider the second wave (age 3 of the child) because it does not correspond to the primary school age.

[8] See the Early Intervention Foundation website.

[9] This threshold for the abnormal level of behavioral problems is provided by the official SDQ website.

Table B.22: Education Coding Scheme

3-category coding	9-category coding	Questionnaire items included
High education	1. Higher degree	Higher degree (A)
	2. Bachelor's degree	First degree (A) Professional qualifications at degree level (V)
	3. HE below degree	Diplomas in higher education (A) Nursing or other medical qualifications (V)
Medium education	4. A-level	A/AS/S level (A) NVQ/SVQ/GSVQ Level 3 (V)
	5. Trade apprenticeship	Trade apprenticeship (V)
	6. GCSE A-C	O-level/GCSE grades A-C (A) NVQ/SVQ/GSVQ Level 2 (V)
Low education	7. GCSE D-G	GCSE grades D-G (A) NVQ/SVQ/GSVQ Level 1 (V)
	9. None	None of these (A & V)

Notes: We excluded category 8, corresponding to "other qualification including oversea" for consistency purposes. (A) stands for academic, (V) for vocational.

or higher, we use the wave 7 (age 17) in which pupils are asked about their educational attainment.[10] Since the exam conditions and requirements vary in the United Kingdom, we restrict our analysis to England only where students are expected to take nine GCSEs subjects, among which 3 of them are compulsory — Maths, English and Science.[11] Thus, we calculate the within-person number of GCSEs passed and create a dummy variable set to 1 if they have achieved 5 or more GCSEs with grades ranging from A* to C(4).

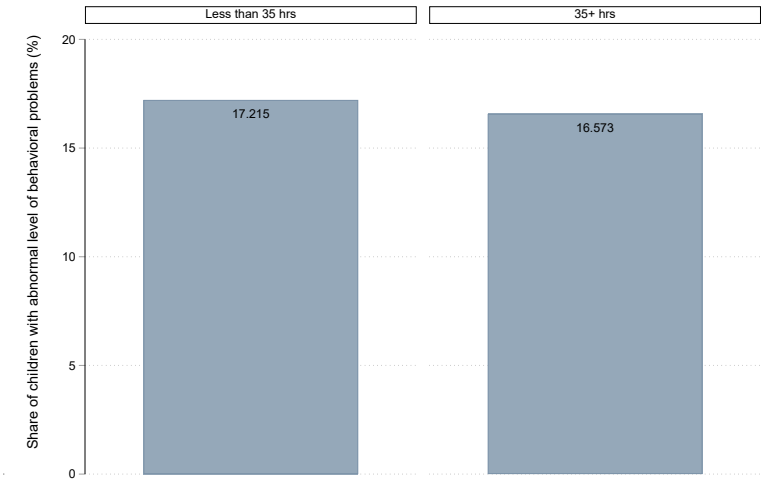
Estimation.

To derive both final metrics, we proceed in three steps, separately for each metric. First, we estimate a probit model where the dependent variable is a dummy set to 1 if the mother worked part-time or less during primary school years of the child, and the independent variables are categorical variables for the mother's and father's education, as well as income categories. Second, we generate predicted values from the probit model and convert them to probabilities using a normal cumulative distribution function. These give the probability that the mother works part-time based on the observed variables (education and income). Third, we create weights to adjust for the likelihood of being part-time based on those observed characteristics. Thus, we provide the average for both metrics in Figure B.9 for when the mother worked part-time or less during the child's primary school age, and the re-weighted average for when the mother worked full-time.

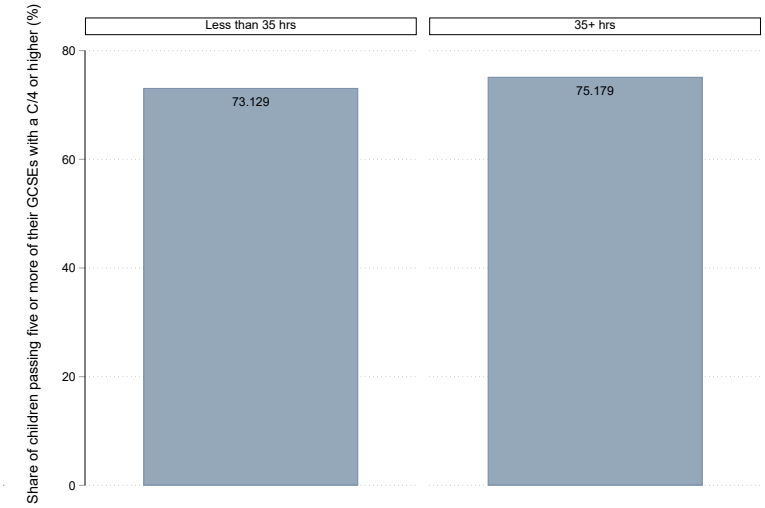
[10] At the age of 17, we expect pupils to have taken their GCSEs.
[11] See this website for a full description of GCSE requirements.

Figure B.9: Distribution of Constructed Metrics

(a) Behavioral Problems



(b) GCSE Pass Rates



B.4.1.3 Motherhood Penalty

In order to produce the heterogeneity results by participant's past experiences, we present below how we construct an indicator for women above the median motherhood penalty (*e.g.*, Kleven, Landais, and Søgaard, 2019).

Information collection.

At the end of the survey, after collecting the demographic information, we ask information about the personal employment history of the participant from age 16 up to their current age. This enables us to construct a yearly panel of participants, with updated information (from years 1972 to 2024) on whether they were (i) in education, (ii) employed full-time, (iii) employed part-time (iv) unemployed, (v) retired, (vi) stay-at-home parent, or (vii) in any other type of activity. Specifically, participants are shown a table, presented in Figure B.31 in the Appendix.

Sample.

We keep participants who experienced the first childbirth during the standard reproductive period, between the ages of 18 and 45. We end up with a sample of N$= 1,025$ participants, and $23,951$ observations.

Event study analyses.

We first conduct the event-study methodology proposed by Kleven et al. (2019), and present below, in Figure B.10, a graphical representation of these regressions, for men and women, which aligns with previous results from the literature (*e.g.*, Kleven, Landais, and Søgaard, 2021; Kleven et al., 2024). These results clearly show that men's likelihood of being employed does not seem to be affected by the first childbirth, while women's employment likelihood in the 10 years after the first childbirth reduces by a significant margin of $\approx 29\%$ with respect to men.

Indicator construction.

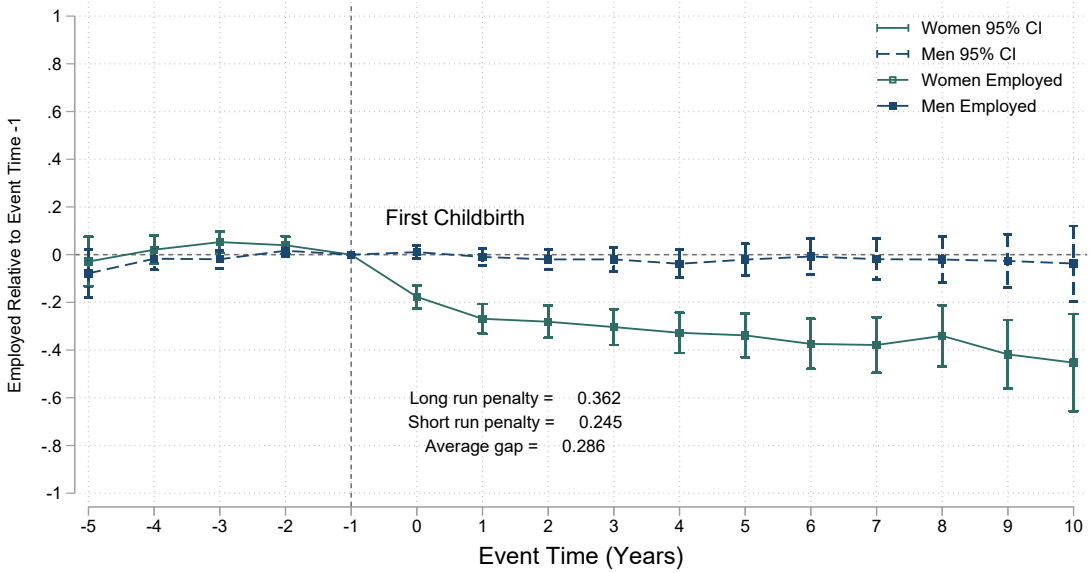
Our aim is to construct a binary indicator flagging women who experienced a motherhood employment penalty above the median of the sample.[12] To do this, we estimate each woman's motherhood employment penalty — the change in the likelihood of being employment after childbirth relative to men controlling for age and post birth fixed effects. Specifically, we interact a post-birth indicator with individual fixed effects for women (while treating men's fixed effects as the reference group). This specification allows us to compute the individual specific post birth employment penalty for each woman. This specification yields a woman-specific estimate of the change in employment likelihood following the birth of her first child.

To do this, we estimate each woman's individual motherhood employment penalty—the change in her probability of being employed after childbirth relative to men—using a static difference-in-differences model. Specifically, we interact a

[12]Given that men are unaffected by parenthood (see Figure B.10), we focus on women only.

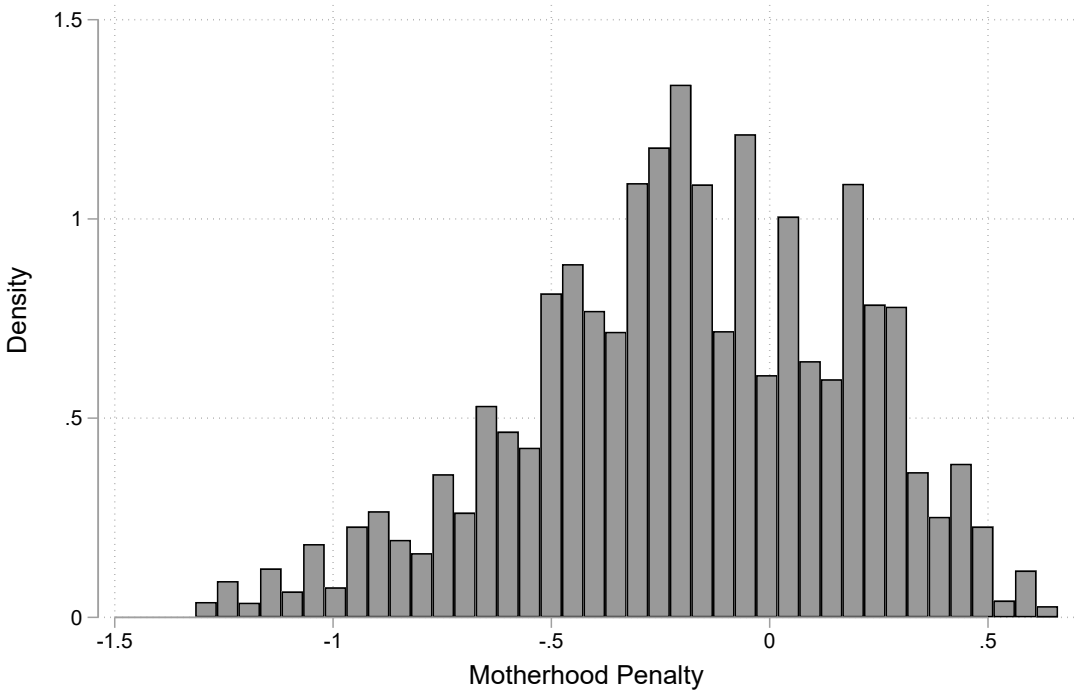
post-birth indicator with individual fixed effects for women controlling for age and post-birth fixed effects (while treating men's fixed effects as the reference group). This specification allows us to estimate the individual specific post-birth employment penalty for each woman. We further create a binary indicator for women whose penalty is above the sample median ($\approx -19\%$). We present below in Figure B.11 a histogram of their calculated penalties.

Figure B.10: Impact of Parenthood on Employment



Notes This figure presents the impact of having a child on a binary variable set to one for being employed at time t, by gender. It further indicates the percentage by which women are falling behind men (1) in the long-run ('long run penalty'), *i.e.*, between seven and ten years post-parenthood, (2) in the short-run ('short run penalty'), *i.e.*, in three years after the first childbirth, and (3) on average ('average gap'), *i.e.*, in the ten years following parenthood.

Figure B.11: Distribution of Motherhood Penalties



Notes: This figure presents an histogram of women's motherhood penalties, *i.e.*, the impact of motherhood on the likelihood of being employed with respect to men.

B.4.2 Survey Screenshots and Questionnaires

B.4.2.1 Main Survey

Figure B.12: Page 1/18

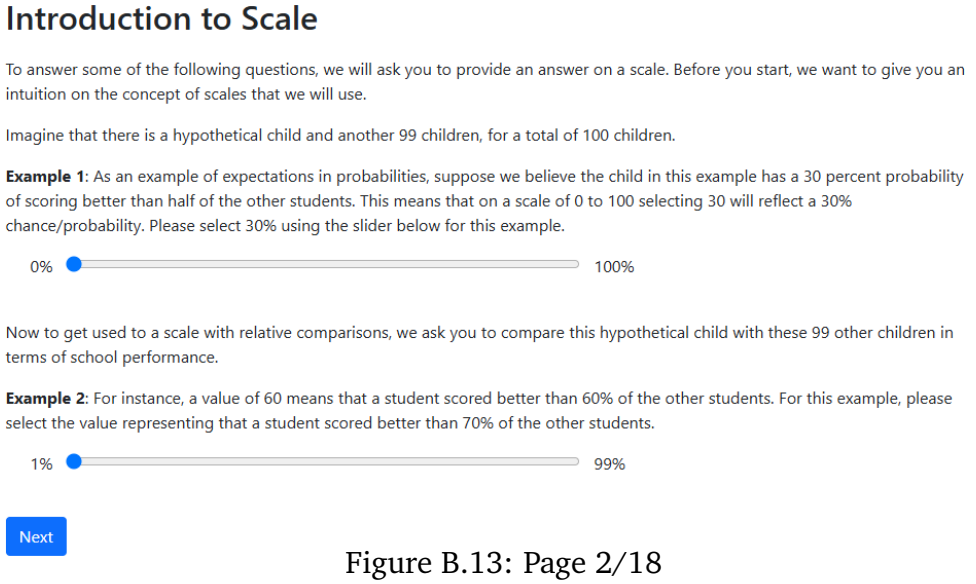
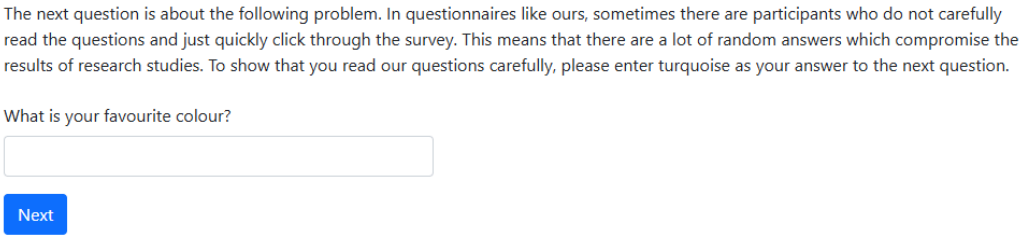


Figure B.13: Page 2/18



Notes: This figure shows the “turquoise” screener.

Figure B.14: Page 3/18

We are interested in your beliefs about children's future outcomes, comparing families with different financial resources and time demands.

▼ **Setup:**

Please imagine an average family in your community. Suppose this family consists of a father and a mother who are both employed, and they have a **boy** who is aged **4**. Suppose household expenditure decisions are made jointly by the father and the mother, and this hypothetical family spends **10%** of their total income on the child's educational and extracurricular activities such as clubs, tutoring, music, sports, etc.

We will show you different scenarios, and ask your opinion about the likelihood that the child will be successful in education and the labour market. There are no clear right or wrong answers, and we know these questions are difficult. Please try to consider each scenario carefully and tell us what you believe the likely outcomes will be.

Scenario 1: The **father** works **42** hours per week at a wage of **£17** per hour.
The **mother** works **35** hours per week at a wage of **£12** per hour.

- What do you think is the probability that this child will eventually graduate from university?

  0% 100%

- When the child is 30 years old, how do you think this child's earnings will compare to other 30 years old?

  1% 99%

Scenario 2: The **father** works **42** hours per week at a wage of **£12** per hour.
The **mother** works **35** hours per week at a wage of **£12** per hour.

- What do you think is the probability that this child will eventually graduate from university?

  0% 100%

- When the child is 30 years old, how do you think this child's earnings will compare to other 30 years old?

  1% 99%

Scenario 3: The **father** works **42** hours per week at a wage of **£22** per hour.
The **mother** works **35** hours per week at a wage of **£12** per hour.

- What do you think is the probability that this child will eventually graduate from university?

  0% 100%

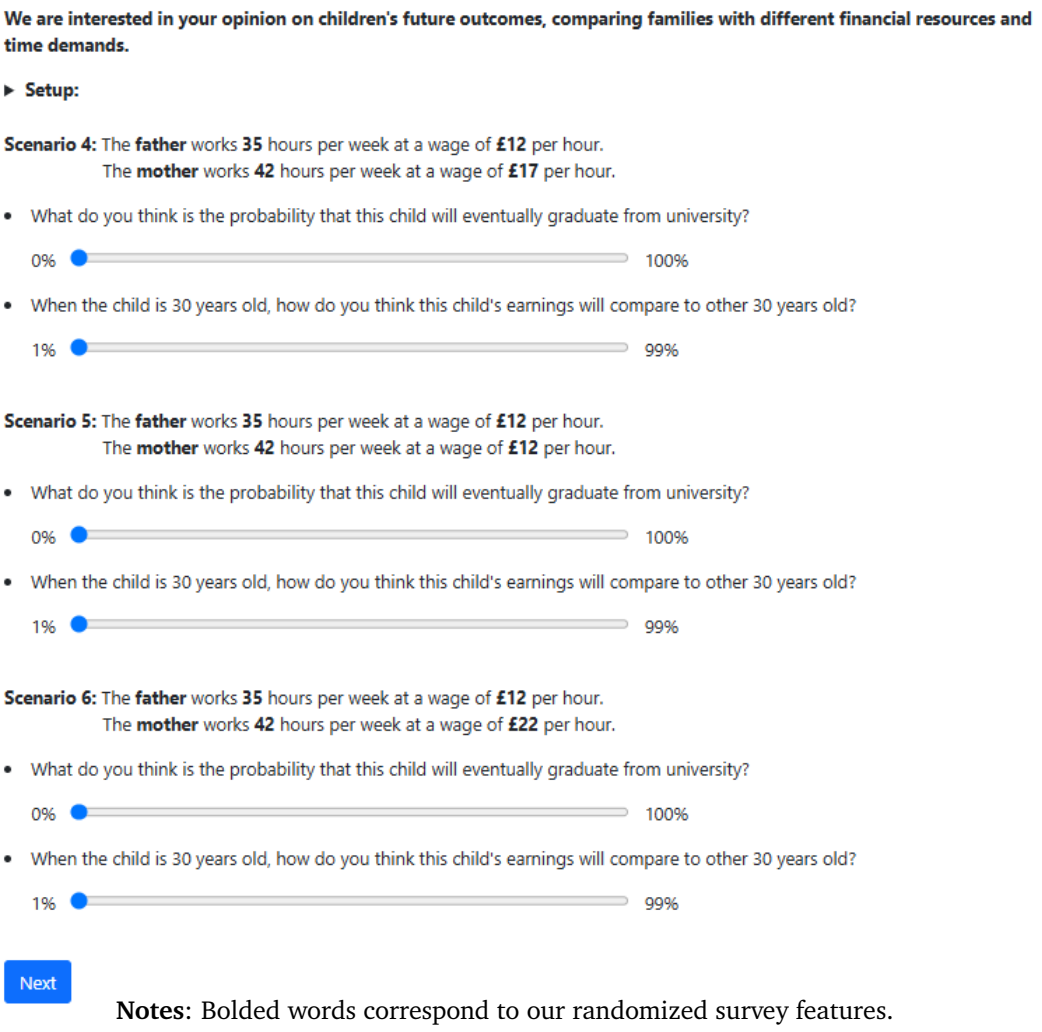
- When the child is 30 years old, how do you think this child's earnings will compare to other 30 years old?

  1% 99%

Next

Notes: Bolded words correspond to our randomized survey features.

Figure B.15: Page 4/18



Notes: Bolded words correspond to our randomized survey features.

Figure B.16: Page 5/18

For context, in the 2013 edition of the British Time Use Survey, parents of at least one child aged 10-14 on average spent 30 minutes per week teaching their children.

Imagine a family whose child aged **11** has the **Key Stage 2** national test upcoming. **Both parents** have a University education. The **father** is very busy this week and only the **mother** has time to help over the week ahead.

How many hours do you expect will be spent helping the child study for the test over the week ahead?

0 time

Next, how many hours do you expect will be spent with the child in extracurricular activities such as sports, art, reading for fun, etc.?

0 time

Next

Figure B.17: Page 6/18

▼ **Scenario:**
Imagine a family whose child aged 11 has the Key Stage 2 national test upcoming. Both parents have a University education. The **father** is very busy this week and only the **mother** has time to help over the week ahead.

Now, suppose that the **mother** will dedicate **30 minutes** in the upcoming week to help the child prepare for the test.

How well do you think the child will do compared to other students?

1% 99%

Next

Figure B.18: Page 7/18

Imagine a family where:

- The **father** works **36** hours per week, earning an hourly wage of **£27**.
- The **mother** works **20** hours per week, earning an hourly wage of **£17**.

How likely do you think it is that each parent has a University education?

Mother 0 100%

Father 0 100%

Next

Figure B.19: Page 8/18

Imagine a family with one child aged **11**, where the **mother's** monthly net income is **£1,500** and the **father's** monthly net income is **£2,500**.

What percentage of income do you believe the family will spend on the child's educational and extracurricular activities such as clubs, tutoring, music, sports, etc?

0 100%

Next

Figure B.20: Page 9/18

How sure are you about your answers to the previous questions under the hypothetical setting?

- ○ Very sure
- ○ Sure
- ○ Somewhat sure
- ○ Unsure
- ○ Very unsure

Next

Figure B.21: Page 10/18

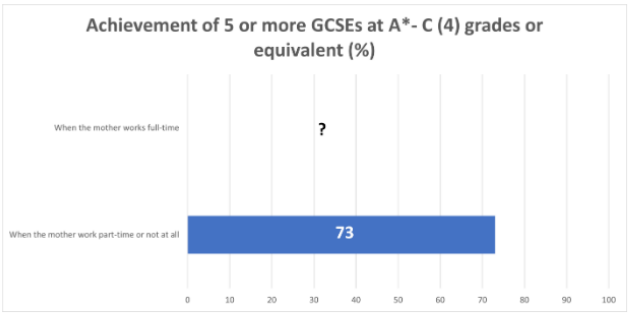
We, as researchers at the University of Strathclyde, have calculated the share of children passing five or more GCSEs with a grade of C/4 or higher.

Among families where the mother worked part-time or not at all, around **73%** of children passed five or more GCSEs with a C/4 or higher. This information is also shown visually in the graph below.

We then computed this statistic for families with similar income and education levels but where the mother worked full-time (35 hours or more). In these families, what percentage of children do you believe eventually passed five or more GCSEs with a C/4 or higher?

You will gain £1.50 if your answer is within 2 percentage points of the true number.

0% 100%



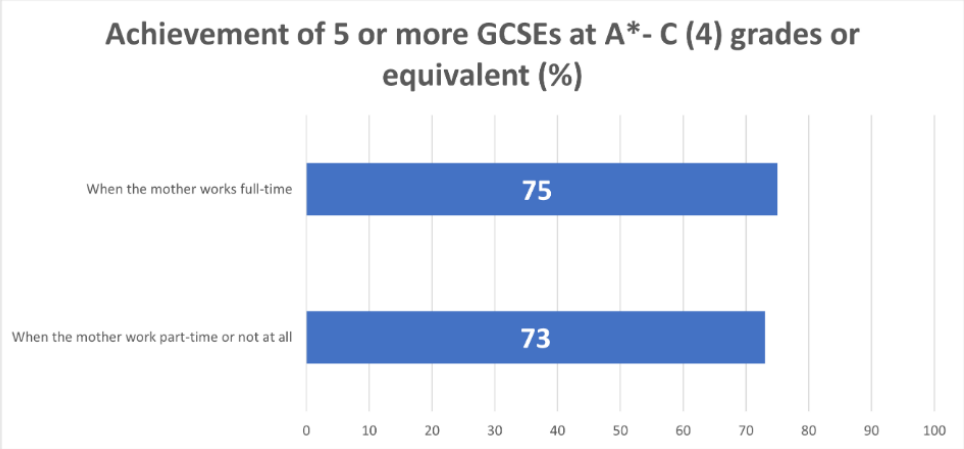
Note: Authors' calculations using the Millennium Cohort Study, a nationally representative sample of families and children born in the year 2000. Mother's working hours correspond to the average number of weekly hours worked, when the child was aged between 5 and 7. Families with mothers working full-time (35 hours or more) are adjusted to have similar income and education as families with mothers working less.

Next

Figure B.22: Page 11/18

For mother worked full-time (35 hours or more), adjusted to have similar education and income levels as mothers working fewer than 35 hours, we found that around **75%** of their children eventually passed five or more GCSE's with a C/4 or higher.

This means these children did about **2 percentage points better** compared to those of mothers working less than 35 hours per week.



Next

Figure B.23: Page 12/18

Research shows that after having children, women experience a drop in labour earnings. This is often explained by the fact that, due to childcare responsibilities, they sort into jobs that offer lower wages but are more flexible and do not require long hours. We are interested in your opinion on two questions about government policy.

Do you think that the government should offer...

- ... policies such as subsidized childcare and universal childcare to help women with children work longer hours?

○ Strongly disagree ○ Disagree ○ Neither agree or disagree ○ Agree ○ Strongly agree

- ... policies such as paternity leave to help women with children return sooner to work after the childbirth?

○ Strongly disagree ○ Disagree ○ Neither agree or disagree ○ Agree ○ Strongly agree

Next

Figure B.24: Page 13/18

The data that we used to calculate the share of children passing five or more GCSEs also provides information on the children's externalising behavioural problems at age 7 (*e.g.,* conduct problems and hyperactivity/inattention).

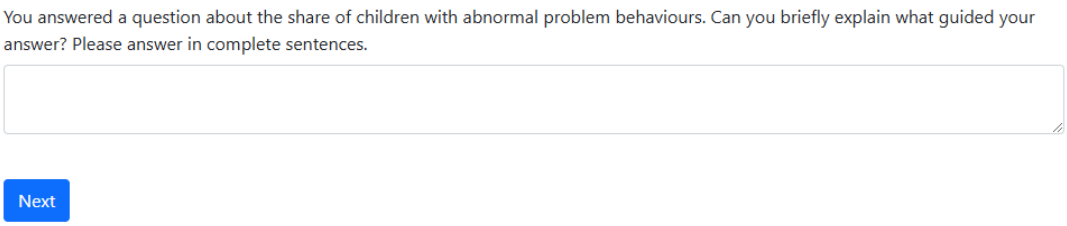
Among families where the mother worked part-time or not at all, **out of 100 children** aged 7, we found that **around 17** had an abnormal level of behavioural problems.

We then computed this statistic for families with similar income and education levels but where the mother worked full-time (35 hours or more). In this group, **out of 100 children**, how many do you believe had an abnormal level of behavioural problems?

You will gain £1.50 if your answer is within 2 points of the true number.

Next

Figure B.25: Page 14/18



Notes: With this open question, we classified participants' answers with two schemes. The first is whether they expect mothers working full-time to cause harm, no harm, or they provide an unclear answer. The second classifies answers into five categories: (1) better resources, (2) lower time investments, (3) no relationship, (4) other, and (5) unclear. We use the first classification to see if information leads respondents to express expectations of less harm from mothers working full-time, and we use the second classification to look at whether information makes respondents less likely to mention the lower time investment model.

Figure B.26: Page 15/18

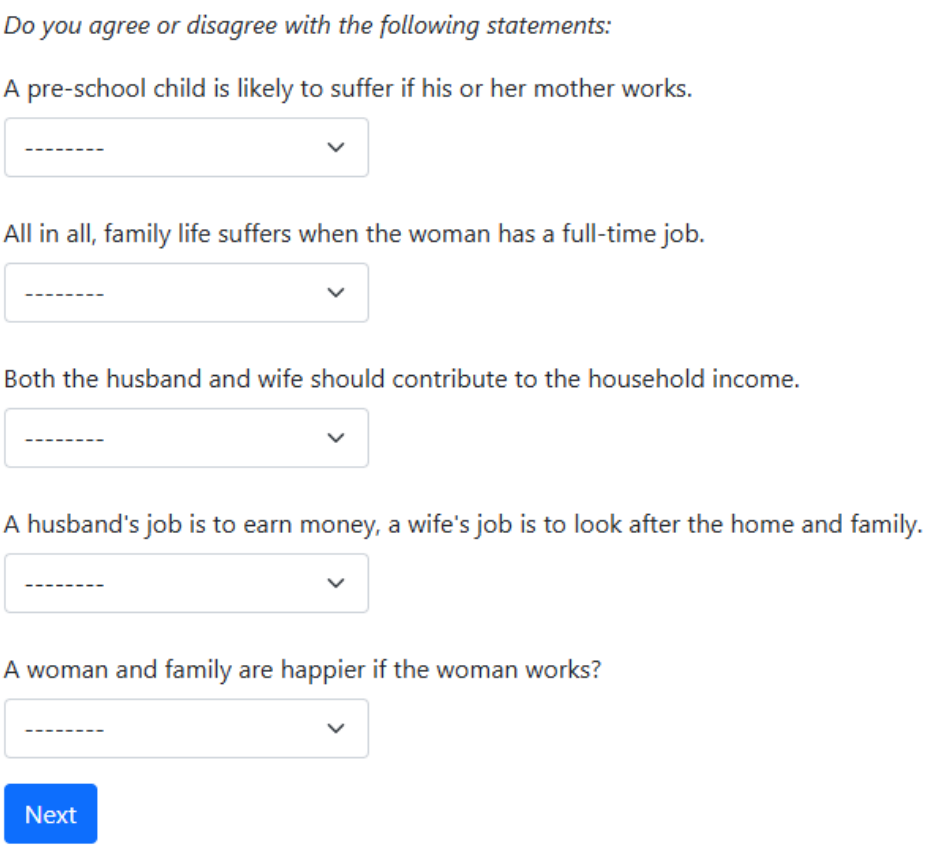


Figure B.27: Page 16/18 — Part 1

Next, we are going to ask some questions about yourself.

What is your age?

What is your gender?

- ○ Male
- ○ Female
- ○ Other
- ○ Prefer not to say

To which of these ethnic groups do you consider you belong?

- ○ Asian or Asian British
- ○ Black, Black British, Caribbean or African
- ○ Mixed or multiple ethnic groups
- ○ White
- ○ Other ethnic group

Are you born in the UK?

- ○ Yes
- ○ No
- ○ Prefer not to say

Which is the highest qualification you have?

- ○ No Qualification
- ○ Other Qualification
- ○ GCSE or equivalent
- ○ A-levels or equivalent
- ○ Degree or Higher

Figure B.28: Page 16/18 — Part 2

Which party did you choose as your primary vote in the last UK General Election?

- ○ Conservative
- ○ Labour
- ○ Liberal Democrat
- ○ Green Party
- ○ Reform UK
- ○ Other
- ○ None

Which of this best describes your current employment situation?

- ○ Self-employed
- ○ Employed
- ○ Unemployed
- ○ Retired
- ○ On maternity leave
- ○ Family care or home
- ○ Full-time student
- ○ Long-term sick or disabled
- ○ On apprenticeship
- ○ Other

What was your (main) job in the week ending last Sunday? Please write your job title below.

Do you work part-time or full-time?

- ○ Part-time
- ○ Full-time
- ○ Not Applicable

How many hours do you typically work per week?

Figure B.29: Page 16/18 — Part 3

What is your personal typical monthly net income?

What is your current marital status?

- Married
- Living as couple
- Widowed
- Divorced
- Separated
- Single, never married
- Civil partnership
- Other

What is your partner's gender?

- Male
- Female
- Other
- Prefer not to say
- I do not have a partner

Does your partner work part-time or full-time?

- Part-time
- Full-time
- Not Applicable

What is your partner's typical monthly net income?

How many children (aged 0-16) do you have in your family?

Figure B.30: Page 16/18 — Part 4

What is the year of birth of your first (eldest) child (adopted or biological)?

How many hours per week do you spend helping your child(ren) develop their skills (e.g. helping with homework/checking workbooks, reading books/telling stories, playing board or card games, etc)? If you have more than one child, please report the overall number of hours.

How many hours per week do you spend doing outdoors activities with your child(ren) (e.g. going to the playground, taking a walk, bringing your child to any sporting activity, going to museums/galleries, etc)? If you have more than one child, please report the overall number of hours.

Thinking back to when you were...	**was your mother working?**	**was your father working?**
less than 12 years old	-------	-------
between 12 and 18 years old	-------	-------

Figure B.31: Page 16/18 — Part 5

We would now like to ask you for some information about your personal history starting at age 16.

Please fill out the table below as follows:

Please state what has happened in your life since you were 16. It is important that you give some answer for every year of your life up to your current age. If you are over 60, please provide your responses up to age 60. If more than one answer applies in a particular year, please select all answers that apply.

At the age of...	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**
I was...																										
in education (*e.g.*, GCSEs, A-levels, college, university)	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
employed full-time	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
employed part-time	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
unemployed	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
retired	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
stay-at-home parent	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
other	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Next

Figure B.32: Page 17/18

If you had to guess, what was the purpose of this study?

Next

Figure B.33: Page 18/18

Thanks for completing all our questions!

Children's GCSE pass rate results:
Your answer about the share of students passing five or more GCSEs with a C/4 or higher does not fall within the range.

Children's behavioural problems results:
Your answer about the number of children out of 100 having abnormal behavioural problems does not fall within the range.

To be paid, please redirect back to Prolific to confirm your participation: https://app.prolific.com/submissions/complete?cc=C8L708C0. Your total payment is: £2.50.
This consists of your base pay, which is £2.50, plus any bonuses if you won them. Bonus payments will be processed after the base payment.
If you want to keep track of your payment, please keep your completion code.

Contacts & Final Report:
If you have questions or concerns about the study, you can contact the researchers at jonathan.norris@strath.ac.uk and agnese.romiti@strath.ac.uk. Please be aware that you will break the confidentiality protocol. For more information and findings of the project, please visit https://sites.google.com/view/svyresults.

B.4.2.2 Obfuscated Follow-Up

Figure B.34: Page 1/6

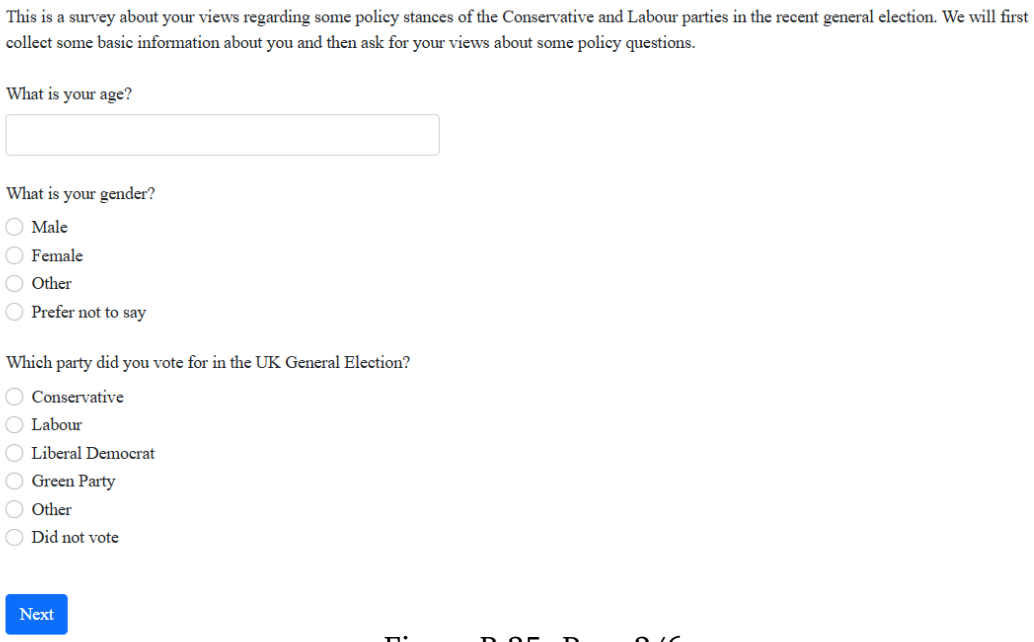


Figure B.35: Page 2/6

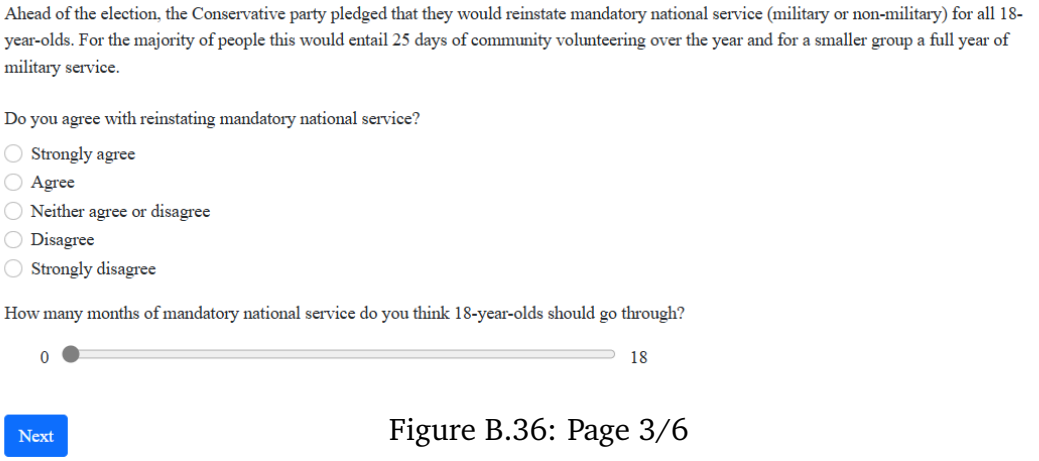


Figure B.36: Page 3/6

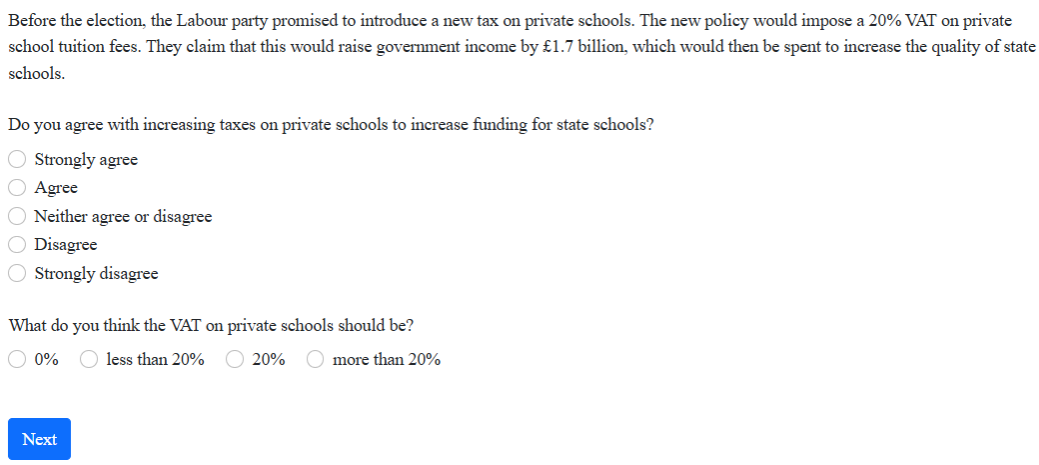


Figure B.37: Page 4/6

In 2023, the government announced a £4 billion expansion of free childcare which would provide parents earning less than £60,000 with 30 free hours of childcare per week. Before the election, the Labour party promised to back this expansion if they won the election.

Do you agree with the policy to increase the hours of free childcare for 2- to 4-year-olds from 15 to 30 hours?

- ○ Strongly agree
- ○ Agree
- ○ Neither agree or disagree
- ○ Disagree
- ○ Strongly disagree

How many hours of free childcare do you think parents of 2- to 4-year-old children should receive?

0 time

Next

Figure B.38: Page 5/6

In a plan to increase the availability of childcare, the Labour Party pledged, prior to the recent election, to convert existing primary schools into 'school-based' nurseries, at a cost of £40,000 per classroom. The party claims that this initiative will create 3,334 new nurseries in high-need areas that currently lack sufficient childcare places. This policy will be funded by the VAT levied on private schools.

Do you agree with this policy?

- ○ Strongly agree
- ○ Agree
- ○ Neither agree or disagree
- ○ Disagree
- ○ Strongly disagree

Next

Figure B.39: Page 6/6

Thanks for completing all our follow-up questions!

To be paid, please redirect back to Prolific to confirm your participation: https://app.prolific.com/submissions/complete?cc=C4JGHDC0.Your total payment is: £0.70.
If you want to keep track of your payment, please keep your completion code.

Contacts & Final Report:
If you have questions or concerns about the study, you can contact the researchers at jonathan.norris@strath.ac.uk and agnese.romiti@strath.ac.uk. Please be aware that you will break the confidentiality protocol. For more information and findings of the project, please visit https://sites.google.com/view/svyresults.

B.5 Departures from Pre-Registered Analysis Plan

We list below figures and tables that were not pre-registered in the analysis plan, and explain our motivation behind conducing such analyses.

Section 2.3.5.

Mechanisms for beliefs, extended analyses.

- We did not pre-register the splits by values of θs (columns 3 to 10) of Tables 2.5 or 2.6, nevertheless we deemed it important to understand how expectations about time investments and productivity of inputs varied by our measures of beliefs on absolute advantage.
- In Table 2.6, we did not pre-register the interactions between "Mother is free to help" and the additional two randomized features: (i) 1h30 *versus* 30 minutes of help, and (ii) both parents *versus* neither have a university education. These dimensions as randomized features were important for understanding variation in the effect of our key feature "mother is free to help".
- Table 2.7, expectations on parental education. We pre-registered the design, but only realized afterwards that looking at the full-time to part-time expected education gradient was the best way to analyze our question on parental education.
- Appendix, Table B.10, extended analysis on expectations about resource allocation. We did not pre-register the splits by values of θs, but we performed these to be consistent with Tables 2.5 and 2.6.

Section 2.4.2.1.

Additional results on belief updating.

- Table 2.8, open-ended question. We departed from our pre-registered classification plan for the open-ended question described in Section 2.4, as we realized later that responses could be best coded to test our objective by "harmful to children when mothers work full-time", "not harmful", or "unclear answer". We also added a second classification to extract more detail.
- Appendix, Table B.7, information treatment effects robustness. We did not pre-register all of our robustness checks on the information treatment, nevertheless each of these provide a useful check to demonstrate our main results here are not sensitive.
- Appendix, Table B.12, perception gap. We did not pre-register the splits by values of θ for this perception gap analysis, but we performed these to be consistent through the paper.

Appendix C

Additional Material for Chapter 3

C.1 Additional Descriptive Statistics and Results

C.1.1 Sample

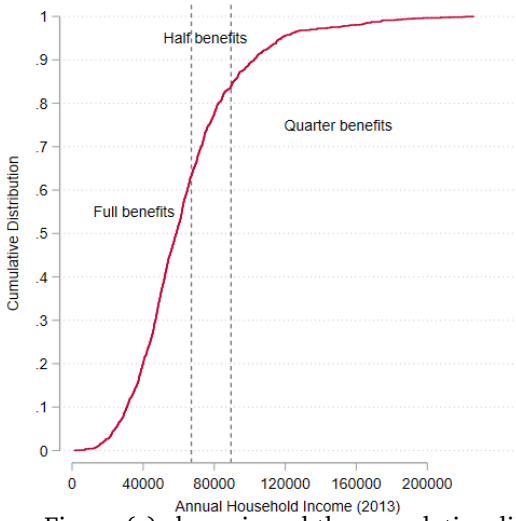
Table C.1: Individual-Level Descriptive Statistics

	Full Benefits	Half Benefits	Quarter Benefits
Woman	0.525	0.491	0.481
	(0.499)	(0.500)	(0.500)
Age	37.590	38.826	39.521
	(11.314)	(11.769)	(12.662)
Single	0.425	0.390	0.372
	(0.494)	(0.488)	(0.483)
Born in France	0.854	0.930	0.923
	(0.353)	(0.254)	(0.267)
University graduate	0.225	0.463	0.567
	(0.418)	(0.499)	(0.496)
Number of children in the household	2.541	2.350	2.393
	(0.864)	(0.669)	(0.662)
Individuals	12655	3654	3896
Observations	38202	11251	11172

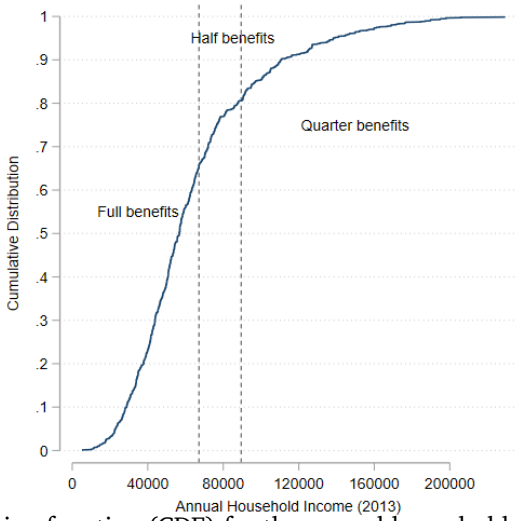
Notes: Means and standard deviations in parentheses of individual characteristics, by treatment status. The sample here corresponds to adults with at least two children in their household, and is further divided into three categories: (1) the control group, "Full Benefits" *i.e.*, households with at least two children whose income fall behind the first threshold introduced by the 2015 reform, and presented in Table 3.1, (2) a first treatment group "Half Benefits" corresponding to households whose income is between the first and second threshold and who receive half of the initial amount of allowances, and (3) a last treatment group "Quarter Benefits" corresponding to households whose income is above the second threshold, and who now receive a quarter of the initial amount of family allowances.

Figure C.1: CDF of Annual Household Income

(a) Households with Two Children



(b) Households with Three Children



Notes: Figure (a) shows in red the cumulative distribution function (CDF) for the annual household income in 2013 of families with two children; the two dashed lines correspond to the income thresholds introduced in 2015 for families with two children, respectively 67,140€ and 89,491€. Figure (b) shows the cumulative distribution function for the annual household income in 2013 of families with three children, and the two dashed lines correspond to the income thresholds introduced in 2015 for these families, respectively 72,735€ and 95,086€.

C.1.2 Effects on Households

Table C.2: Average Treatment Effects on Overall Deprivation

	(1) Deprived		(2) Deprivation Score	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Half Benefits	-0.022**		-0.185***	
	(0.011)		(0.049)	
Post	-0.054***	-0.054***	-0.557***	-0.535***
	(0.016)	(0.016)	(0.078)	(0.078)
Post × Half Benefits	0.030**		0.208***	
	(0.012)		(0.059)	
Quarter Benefits		-0.045***		-0.337***
		(0.014)		(0.082)
Post × Quarter Benefits		0.042***		0.224***
		(0.010)		(0.059)
Mean Dep. Var	0.042	0.003	0.544	0.159
Households	1181	1160	1173	1156
Observations	11251	11172	11159	11117
Household Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.1), ran separately by treatment status, on (1) a binary variable for being deprived, and (2) the number of items the household is unable to afford, as defined in Subsection 3.3.3. Standard errors, clustered at the household-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treatment group. See notes of Table 3.6 for the definition of the treatment status.

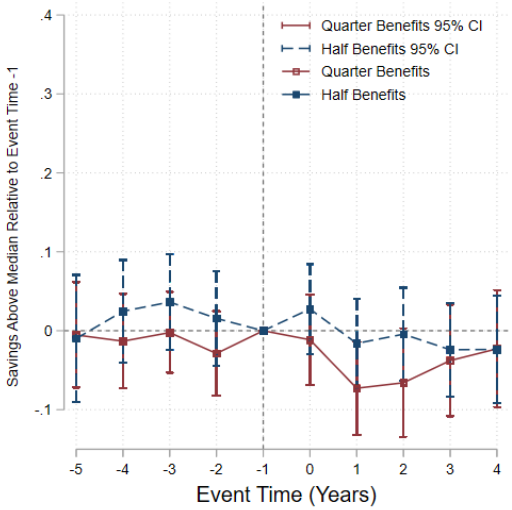
Table C.3: Average (Categorical) Treatment Effects on Overall Deprivation

	(1) Deprived	(2) Deprivation Score
Half Benefits	-0.021**	-0.205***
	(0.009)	(0.044)
Quarter Benefits	-0.037***	-0.261***
	(0.010)	(0.056)
Post	-0.048***	-0.494***
	(0.014)	(0.069)
Post × Half Benefits	0.027**	0.198***
	(0.011)	(0.054)
Post × Quarter Benefits	0.042***	0.206***
	(0.010)	(0.054)
Mean Dep. Var	0.027	0.385
Households	7763	7681
Observations	60630	60027
Household Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes

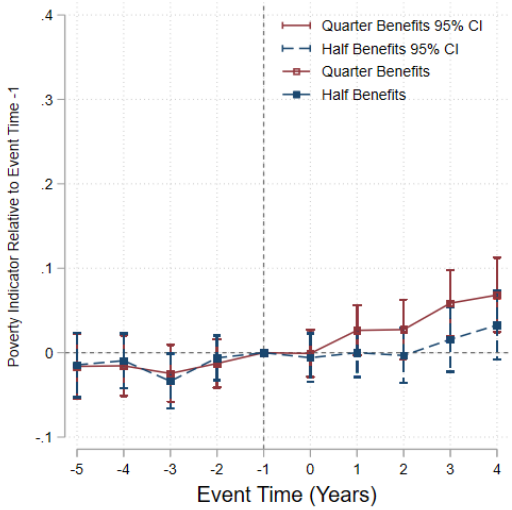
Notes: OLS results for Equation (3.1) using a categorical variable for the treatment status, on (1) a binary variable for being deprived, and (2) the number of items the household is unable to afford, as defined in Subsection 3.3.3. Standard errors, clustered at the household-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) for both treated groups. See notes of Table 3.6 for the definition of the treatment status.

Figure C.2: Impact of the Reform on Poverty Indicators

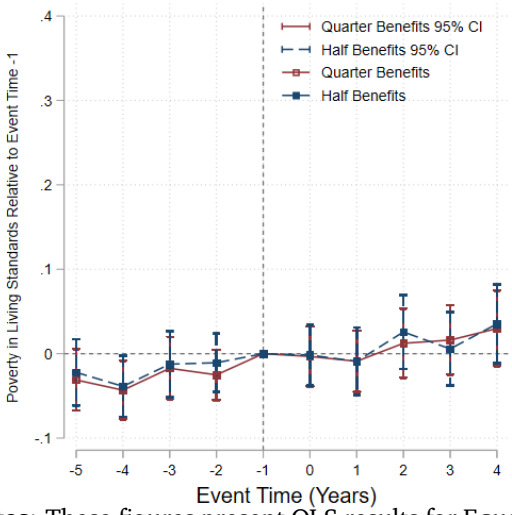
(a) Net Annual Household Income



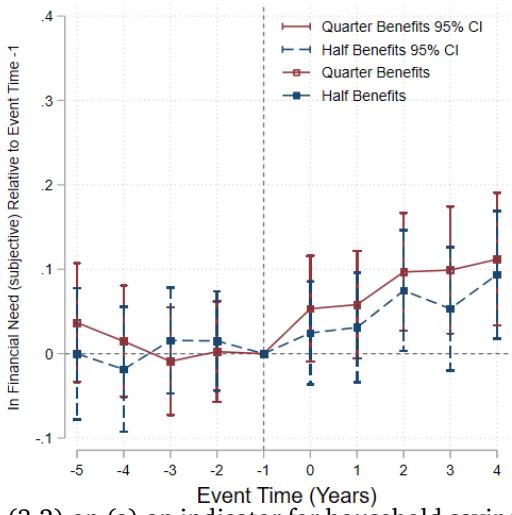
(b) Poverty



(c) Poverty in Living Conditions



(d) Self-Reported Hardship



Notes: These figures present OLS results for Equation (3.2) on (a) an indicator for household savings above median (b) a poverty indicator, (c) a poverty in living conditions indicator, and (d) a self-reported measure of financial hardship, as defined in Subsection 3.3.3. See notes of Table 3.6 for the definition of the treatment status.

Table C.4: Average Treatment Effects on Items of Deprivation

	(1) House Deprivation		(2) Clothes Deprivation		(3) Food Deprivation		(4) Leisure Deprivation	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Half Benefits	-0.019		-0.029**		-0.020**		-0.040***	
	(0.013)		(0.011)		(0.008)		(0.016)	
Post	-0.168***	-0.167***	-0.071***	-0.068***	-0.025**	-0.021*	-0.106***	-0.102***
	(0.021)	(0.021)	(0.018)	(0.018)	(0.012)	(0.013)	(0.020)	(0.019)
Post × Half Benefits	0.031*		0.041***		0.021**		0.025	
	(0.017)		(0.014)		(0.010)		(0.018)	
Quarter Benefits		-0.105***		-0.059***		-0.022*		-0.076***
		(0.024)		(0.014)		(0.013)		(0.026)
Post × Quarter Benefits		0.073***		0.037***		0.013		0.029
		(0.016)		(0.012)		(0.010)		(0.018)
Mean Dep. Var	0.108	0.041	0.039	0.018	0.019	0.004	0.195	0.070
Households	1181	1159	1180	1160	1181	1160	1181	1160
Observations	11249	11170	11237	11163	11244	11170	11249	11172
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.1), ran separately by treatment status, on (1) the number of items deprived, (2) a house deprivation indicator, (3) a clothing deprivation indicator, (4) a food deprivation indicator, and (5) a leisure deprivation indicator, as defined in Subsection 3.3.3. Standard errors, clustered at the household-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treatment group. See notes of Table 3.6 for the definition of the treatment status.

Table C.5: Average Treatment Effects on Overall Deprivation, by Household Baseline Characteristics

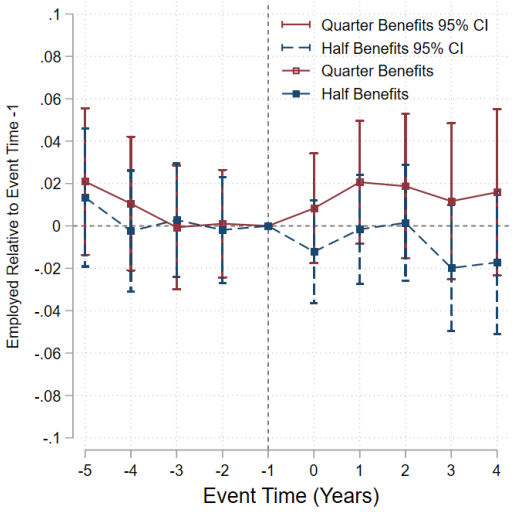
	(1) Deprived		(2) Deprivation Score	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: Single-parent household				
Half Benefits	0.070		0.068	
	(0.061)		(0.223)	
Post	-0.133**	-0.141**	-0.808***	-0.819***
	(0.061)	(0.058)	(0.258)	(0.255)
Post × Half Benefits	-0.078		-0.289	
	(0.083)		(0.339)	
Quarter Benefits		-0.138		-0.501
		(0.097)		(0.426)
Post × Quarter Benefits		0.123		0.126
		(0.103)		(0.395)
Mean Dep. Var	0.177	0.040	1.498	0.673
Observations	530	386	530	386
Panel B: Unemployed partner				
Half Benefits	-0.042***		-0.240***	
	(0.016)		(0.084)	
Post	-0.053*	-0.046	-0.367**	-0.363**
	(0.029)	(0.030)	(0.143)	(0.145)
Post × Half Benefits	0.037**		0.162*	
	(0.018)		(0.098)	
Quarter Benefits		-0.049**		-0.454***
		(0.024)		(0.132)
Post × Quarter Benefits		0.060***		0.280***
		(0.017)		(0.083)
Mean Dep. Var	0.037	0.008	0.543	0.217
Observations	3200	3031	3200	3031
Panel C: Inactive partner				
Half Benefits	-0.012		-0.164***	
	(0.014)		(0.061)	
Post	-0.058***	-0.060***	-0.643***	-0.596***
	(0.019)	(0.019)	(0.096)	(0.097)
Post × Half Benefits	0.026		0.226***	
	(0.016)		(0.073)	
Quarter Benefits		-0.041**		-0.281***
		(0.017)		(0.105)
Post × Quarter Benefits		0.035***		0.194**
		(0.013)		(0.079)
Mean Dep. Var	0.044	0.013	0.545	0.224
Observations	8051	8141	8051	8141
Household Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.1), ran separately by treatment status, on (1) a binary variable for being deprived, and (2) the number of items the household is unable to afford, as defined in Subsection 3.3.3. Panel A corresponds to single-parent households before the reform (at baseline), while Panels B and C correspond to dual-parent households, but with an unemployed partner (Panel B) or with an inactive partner (Panel C) at baseline. Standard errors, clustered at the household-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline and for the corresponding treatment group.

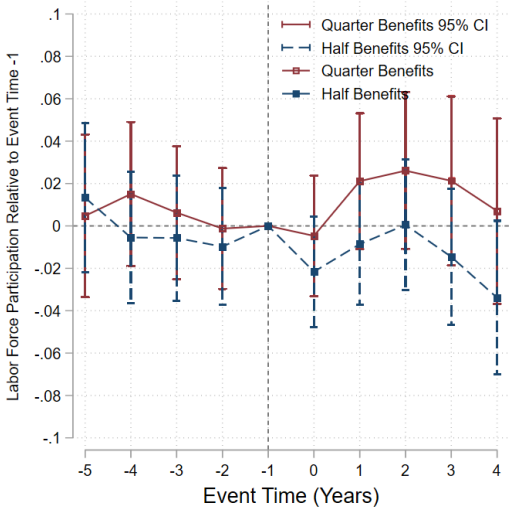
C.1.3 Effects on Parental Labor Supply

Figure C.3: Impact of the Reform on Parental Labor Supply

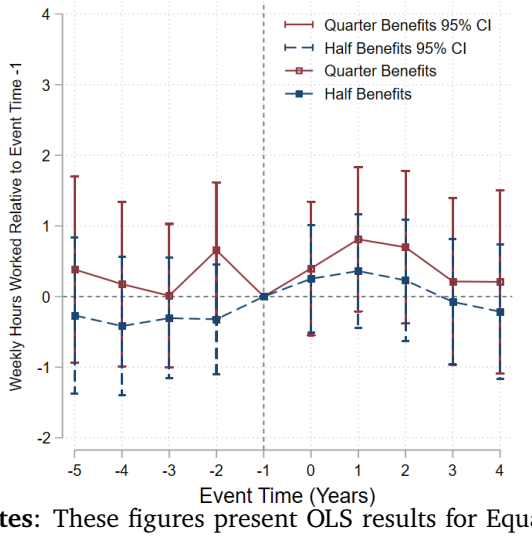
(a) Employed



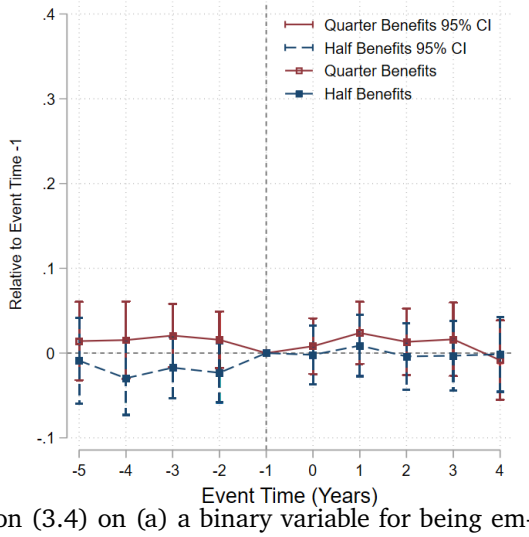
(b) Labor Force Participation



(c) Weekly Hours Worked



(d) Full-Time Work



Notes: These figures present OLS results for Equation (3.4) on (a) a binary variable for being employed, (b) a binary indicator for labor force participation, (c) the usual number of weekly hours worked, and (d) an indicator for full-time employment, as defined in Subsection 3.3.3. See notes of Table 3.6 for the definition of the treatment status.

Table C.6: Average Treatment Effects on Parental Labor Supply, by Household Baseline Characteristics

	(1) Employment		(2) Labor Force Participation		(3) Hours Worked		(4) Full-Time	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: Single-parent household								
Half Benefits	0.060		0.042		0.314		0.011	
	(0.040)		(0.041)		(1.186)		(0.058)	
Post	0.038	0.041	0.141	0.186	5.537	7.935	-0.210	-0.192
	(0.290)	(0.285)	(0.291)	(0.296)	(9.177)	(9.008)	(0.409)	(0.379)
Post × Half Benefits	-0.062		-0.021		1.257		0.009	
	(0.050)		(0.056)		(1.480)		(0.069)	
Quarter Benefits		0.095*		0.062		0.581		-0.028
		(0.056)		(0.064)		(1.985)		(0.045)
Post × Quarter Benefits		-0.058		-0.046		3.063		0.118
		(0.078)		(0.089)		(3.485)		(0.102)
Mean Dep. Var	0.569	0.556	0.551	0.550	37.881	40.148	0.806	0.761
Observations	528	386	530	386	341	264	341	264
Panel B: Unemployed partner								
Half Benefits	0.029**		0.035***		0.553		0.008	
	(0.013)		(0.013)		(0.405)		(0.016)	
Post	-0.067	-0.029	0.011	-0.017	-0.026	0.460	-0.010	0.049
	(0.142)	(0.141)	(0.150)	(0.155)	(4.961)	(5.279)	(0.208)	(0.225)
Post × Half Benefits	0.001		-0.004		0.392		0.009	
	(0.017)		(0.018)		(0.460)		(0.022)	
Quarter Benefits		0.009		0.029		0.274		0.007
		(0.021)		(0.022)		(0.713)		(0.029)
Post × Quarter Benefits		0.022		-0.003		0.871		0.019
		(0.018)		(0.020)		(0.646)		(0.024)
Mean Dep. Var	0.771	0.751	0.777	0.751	39.135	42.457	0.811	0.845
Observations	3196	3029	3200	3031	2535	2301	2535	2301
Panel C: Inactive partner								
Half Benefits	0.024**		0.032***		0.337		0.004	
	(0.009)		(0.010)		(0.297)		(0.012)	
Post	-0.099	-0.108	-0.100	-0.098	0.291	1.332	-0.141	-0.135
	(0.111)	(0.120)	(0.112)	(0.126)	(2.233)	(2.458)	(0.114)	(0.116)
Post × Half Benefits	-0.014		-0.013		0.399		0.014	
	(0.011)		(0.012)		(0.357)		(0.015)	
Quarter Benefits		0.025		0.025		-0.117		0.000
		(0.016)		(0.017)		(0.542)		(0.020)
Post × Quarter Benefits		0.007		0.012		0.008		-0.010
		(0.014)		(0.015)		(0.505)		(0.017)
Mean Dep. Var	0.758	0.709	0.755	0.706	38.973	41.970	0.802	0.830
Observations	8038	8136	8051	8141	6196	5889	6196	5889
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3), ran separately by treatment status on (1) employment, (2) labor force participation, (3) weekly hours worked, and (4) an indicator for full-time employment. Panel A corresponds to single-parent households before the reform (at baseline), while Panels B and C correspond to dual-parent households, but with an unemployed partner (Panel B) or with an inactive partner (Panel C) at baseline. Standard errors, clustered at the individual-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group.

C.1.4 Effects on Childcare Choices

Table C.7: Average Treatment Effects on Time Spent in Different Childcare Facilities

	(1) School		(2) Relative		(3) Childminder	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel C: 6–12 years-old						
Half Benefits	0.007		-0.046		-0.098	
	(0.528)		(0.191)		(0.095)	
Post	15.277***	15.372***	-1.816***	-1.771***	-0.594***	-0.606***
	(0.640)	(0.653)	(0.282)	(0.281)	(0.129)	(0.153)
Post × Half Benefits	-0.109		-0.133		0.008	
	(0.580)		(0.255)		(0.134)	
Quarter Benefits		-0.574		-0.047		0.208
		(0.766)		(0.392)		(0.167)
Post × Quarter Benefits		0.122		-0.160		-0.201
		(0.722)		(0.307)		(0.233)
Mean Dep. Var	28.188	28.822	1.416	1.233	0.281	1.238
Observations	3242	2974	3246	2975	3250	2984
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3), ran separately by treatment status, on the number of hours per week spent (1) in school, (2) cared by a relative, and (3) with a professional childminder. Standard errors, clustered at the child-level, are presented in parenthesis. Results are presented for all children aged 6 to 12. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status.

Table C.8: Average Treatment Effects on Likelihood to Attend Different Childcare Facilities

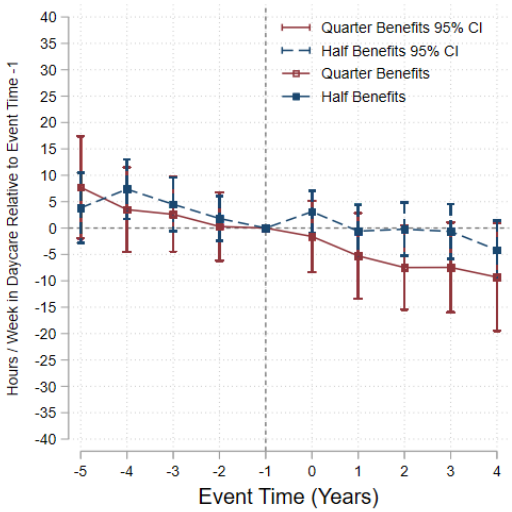
	(1) Daycare		(2) Childminder		(3) Relative		(4) Pre-school	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: 0–5 years-old								
Half Benefits	0.001		0.034		-0.010		0.007	
	(0.041)		(0.031)		(0.031)		(0.030)	
Post	-0.422***	-0.376***	-0.019	-0.015	0.132***	0.119***	2.065***	2.061***
	(0.048)	(0.050)	(0.028)	(0.028)	(0.037)	(0.038)	(0.035)	(0.037)
Post × Half Benefits	-0.018		-0.090**		0.020		-0.004	
	(0.051)		(0.037)		(0.036)		(0.035)	
Quarter Benefits		0.008		0.081		0.029		-0.022
		(0.064)		(0.054)		(0.051)		(0.059)
Post × Quarter Benefits		-0.062		-0.049		-0.049		0.054
		(0.078)		(0.072)		(0.055)		(0.054)
Mean Dep. Var	0.362	0.363	0.114	0.218	0.184	0.177	0.534	0.540
Observations	1833	1196	1835	1195	1828	1188	1836	1196
Panel B: 0–2 years-old								
Half Benefits	-0.055		0.087		0.021			
	(0.081)		(0.057)		(0.058)			
Post	1.174***	1.124***	0.270***	0.332***	0.368***	0.329***		
	(0.125)	(0.129)	(0.082)	(0.088)	(0.089)	(0.094)		
Post × Half Benefits	0.086		-0.192**		0.058			
	(0.104)		(0.080)		(0.077)			
Quarter Benefits		-0.079		0.194*		-0.018		
		(0.136)		(0.117)		(0.089)		
Post × Quarter Benefits		0.134		-0.225		-0.033		
		(0.176)		(0.140)		(0.094)		
Mean Dep. Var	0.571	0.561	0.140	0.243	0.192	0.160		
Observations	722	453	723	453	718	449		
Panel C: 3–5 years-old								
Half Benefits	0.039		0.004		-0.023		0.021	
	(0.056)		(0.044)		(0.051)		(0.053)	
Post	-0.596***	-0.575***	-0.080	-0.058	0.004	0.009	0.908***	0.912***
	(0.087)	(0.091)	(0.062)	(0.064)	(0.076)	(0.079)	(0.093)	(0.099)
Post × Half Benefits	0.027		-0.059		0.029		0.022	
	(0.069)		(0.052)		(0.056)		(0.061)	
Quarter Benefits		0.048		0.110		0.013		-0.061
		(0.084)		(0.079)		(0.087)		(0.118)
Post × Quarter Benefits		-0.018		-0.045		-0.019		0.038
		(0.110)		(0.094)		(0.098)		(0.116)
Mean Dep. Var	0.226	0.242	0.097	0.202	0.179	0.187	0.873	0.863
Observations	1111	743	1112	742	1110	739	1113	743
Panel D: 4–5 years-old								
Half Benefits			0.003		-0.025		0.019	
			(0.062)		(0.081)		(0.056)	
Post			-0.019	0.005	0.143	0.102	-0.276**	-0.283**
			(0.114)	(0.117)	(0.136)	(0.146)	(0.110)	(0.116)
Post × Half Benefits			-0.047		0.009		0.008	
			(0.081)		(0.094)		(0.069)	
Quarter Benefits				0.045		0.002		-0.058
				(0.141)		(0.117)		(0.188)
Post × Quarter Benefits				-0.109		-0.127		0.065
				(0.159)		(0.131)		(0.084)
Mean Dep. Var			0.078	0.205	0.176	0.193	0.958	0.954
Observations			787	516	785	513	788	517
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3), ran separately by treatment status, on indicators set to one if the the number of hours per week spent (1) in daycare center, (2) with a professional childminder, (3) cared by a relative, or (4) in pre-school is greater than zero; 0 otherwise. Standard errors, clustered at the child-level, are presented in parenthesis. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status.

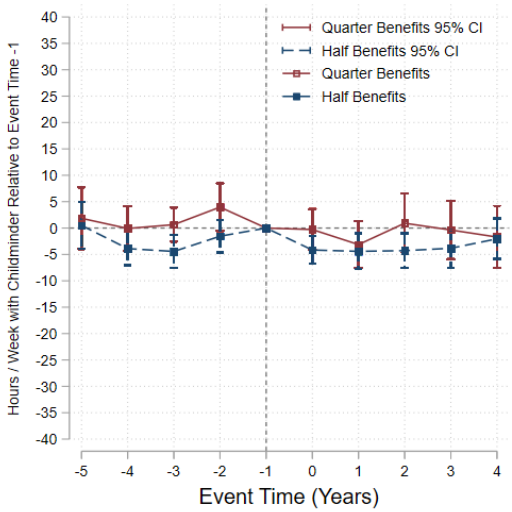
Figure C.4: Impact of the Reform on Time Spent in Childcare Facilities (1/4)

Children Aged 0 to 5

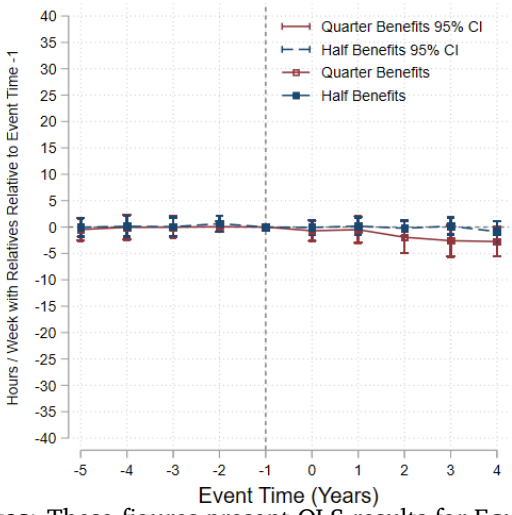
(a) Daycare



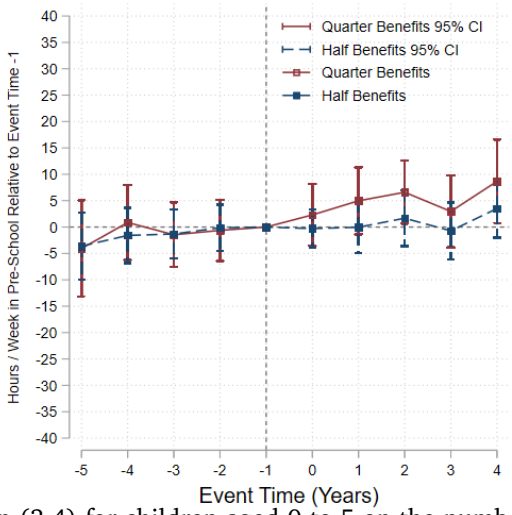
(b) Childminder



(c) Relative



(d) Pre-School

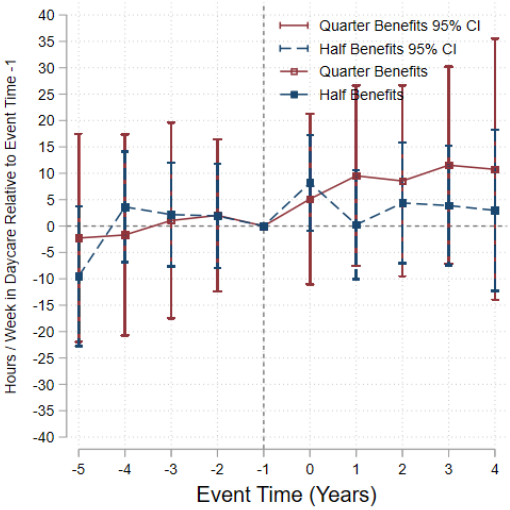


Notes: These figures present OLS results for Equation (3.4) for children aged 0 to 5 on the number of hours per week spent (a) daycare, (b) with a professional childminder, (c) cared by a relative, and (d) in in pre-school, as described in Subsection 3.3.3.

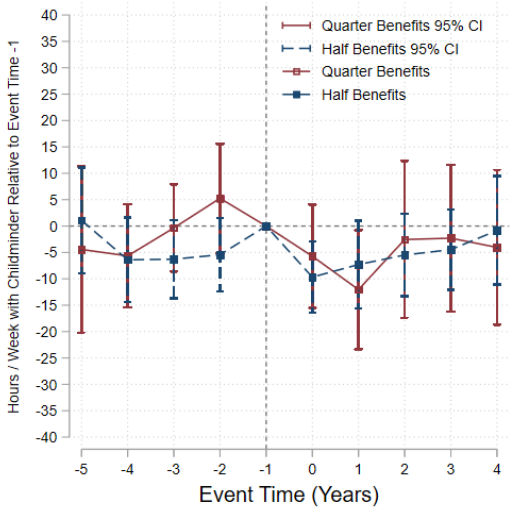
Figure C.5: Impact of the Reform on Time Spent in Childcare Facilities (2/4)

Children Aged 0 to 2

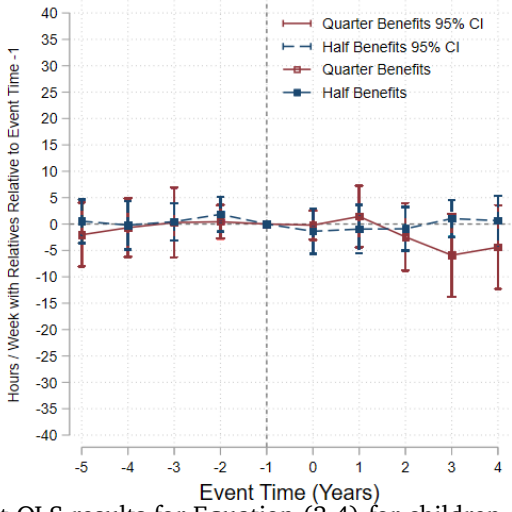
(a) Daycare



(b) Childminder



(c) Relative

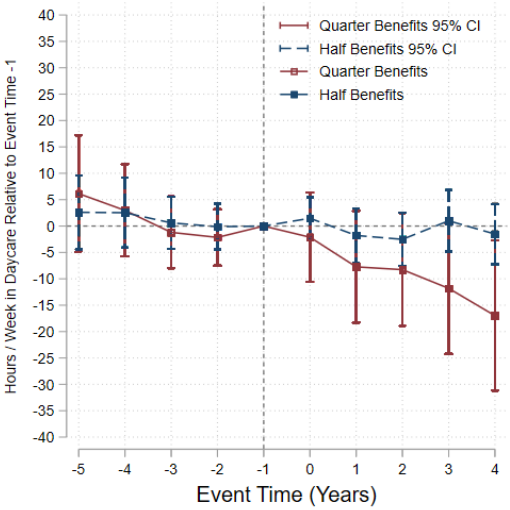


Notes: These figures present OLS results for Equation (3.4) for children aged 0 to 2 on the number of hours per week spent (a) daycare, (b) with a professional childminder, (c) cared by a relative, and (d) in in pre-school.

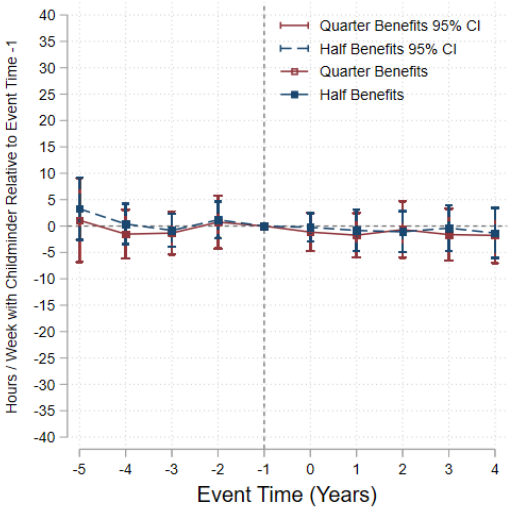
Figure C.6: Impact of the Reform on Time Spent in Childcare Facilities (3/4)

Children Aged 3 to 5

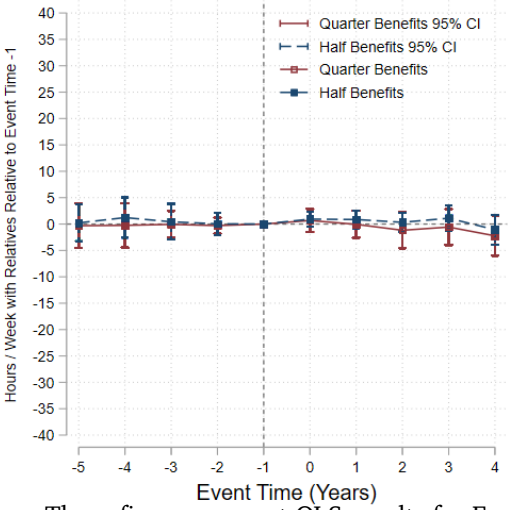
(a) Daycare



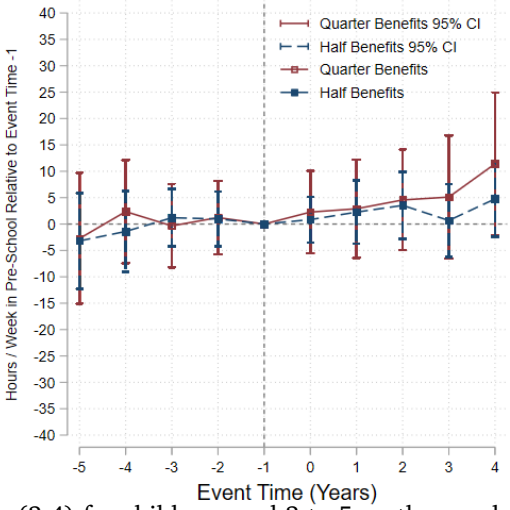
(b) Childminder



(c) Relative



(d) Pre-School

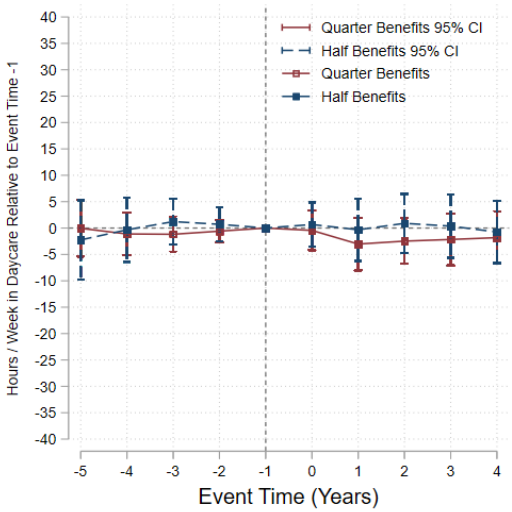


Notes: These figures present OLS results for Equation (3.4) for children aged 3 to 5 on the number of hours per week spent (a) daycare, (b) with a professional childminder, (c) cared by a relative, and (d) in in pre-school.

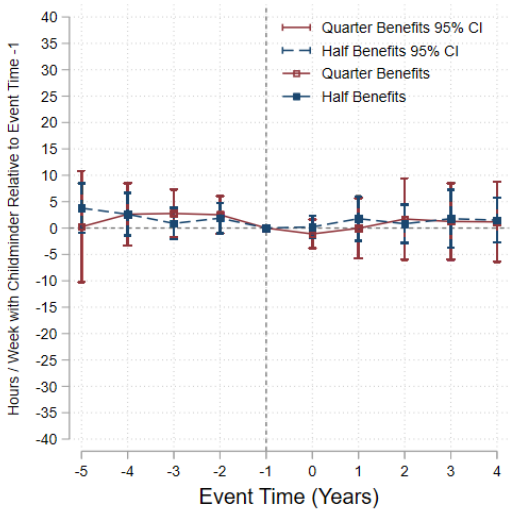
Figure C.7: Impact of the Reform on Time Spent in Childcare Facilities (4/4)

Children Aged 4 to 5

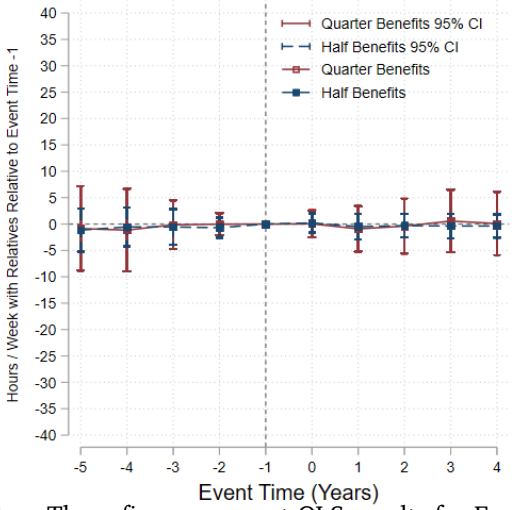
(a) Daycare



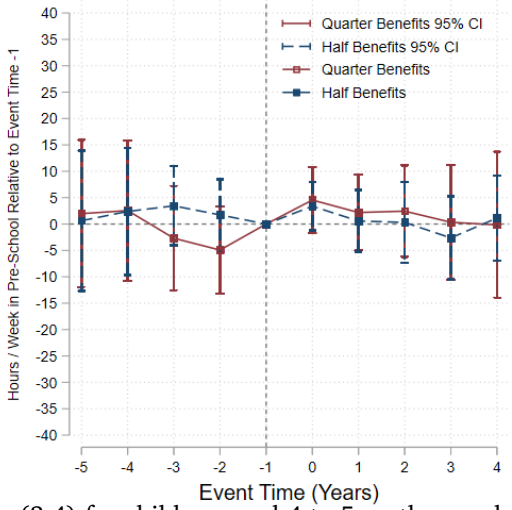
(b) Childminder



(c) Relative



(d) Pre-School



Notes: These figures present OLS results for Equation (3.4) for children aged 4 to 5 on the number of hours per week spent (a) daycare, (b) with a professional childminder, (c) cared by a relative, and (d) in in pre-school.

Table C.9: Average Treatment Effects on Likelihood to Attend Different Childcare Facilities

	(1) Daycare		(2) Childminder		(3) Relative		(4) Pre-school	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: 0–5 years-old								
Half Benefits	0.932		2.018**		-0.270		-0.359	
	(1.590)		(1.008)		(0.514)		(1.064)	
Post	-16.730***	-15.407***	-3.411***	-3.222***	0.053	-0.017	59.011***	58.563***
	(1.654)	(1.730)	(0.943)	(0.977)	(0.690)	(0.707)	(1.222)	(1.318)
Post × Half Benefits	-2.054		-2.797**		-0.258		0.980	
	(1.795)		(1.146)		(0.669)		(1.194)	
Quarter Benefits		3.451		1.895		0.896		-3.364*
		(2.771)		(1.648)		(0.905)		(1.842)
Post × Quarter Benefits		-5.256*		-2.012		-1.590		2.818
		(3.137)		(2.322)		(1.187)		(1.793)
Mean Dep. Var	9.857	10.409	2.643	4.355	1.860	2.217	15.135	15.508
Observations	1833	1196	1835	1195	1828	1188	1836	1196
Panel B: 0–2 years-old								
Half Benefits	-1.675		2.686		0.374			
	(3.527)		(1.921)		(1.023)			
Post	32.084***	29.513***	5.639*	7.751**	6.381**	5.279**		
	(4.813)	(5.003)	(3.083)	(3.456)	(2.573)	(2.484)		
Post × Half Benefits	4.557		-4.551		-1.817			
	(4.136)		(2.786)		(2.260)			
Quarter Benefits		-3.273		5.188		0.289		
		(6.400)		(3.831)		(0.986)		
Post × Quarter Benefits		5.678		-9.068		-1.225		
		(7.980)		(5.656)		(1.627)		
Mean Dep. Var	18.679	19.656	4.185	6.521	2.245	2.586		
Observations	722	453	723	453	718	449		
Panel C: 3–5 years-old								
Half Benefits	0.055		0.702		-0.484		-0.399	
	(1.852)		(1.176)		(0.671)		(2.060)	
Post	-17.899***	-17.650***	-4.688***	-3.194**	-1.066	-0.941	37.969***	37.566***
	(2.242)	(2.369)	(1.563)	(1.588)	(1.413)	(1.481)	(3.122)	(3.330)
Post × Half Benefits	-0.006		-1.600		0.381		2.182	
	(2.075)		(1.389)		(0.656)		(2.220)	
Quarter Benefits		2.021		3.066*		0.415		-5.281
		(2.548)		(1.680)		(0.853)		(3.441)
Post × Quarter Benefits		-1.360		-1.697		-0.406		1.551
		(2.637)		(2.140)		(1.118)		(3.544)
Mean Dep. Var	4.123	4.771	1.639	3.032	1.611	1.993	24.779	24.855
Observations	1111	743	1112	742	1110	739	1113	743
Panel D: 4–5 years-old								
Half Benefits			-0.097		-0.563		0.818	
			(0.830)		(1.212)		(2.636)	
Post			0.598	1.533	0.999	1.162	0.942	1.679
			(1.925)	(2.038)	(2.112)	(2.306)	(4.389)	(4.585)
Post × Half Benefits			0.159		0.262		-0.067	
			(1.124)		(1.117)		(2.700)	
Quarter Benefits				0.714		0.217		-6.068
				(1.624)		(1.458)		(4.101)
Post × Quarter Benefits				-1.276		-0.475		6.441*
				(1.272)		(1.383)		(3.691)
Mean Dep. Var			0.914	2.731	1.446	1.770	28.028	28.133
Observations			787	516	785	513	788	517
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TU Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3), ran separately by treatment status, on indicators set to one if the the number of hours per week spent (1) in daycare center, (2) with a professional childminder, (3) cared by a relative, or (4) in pre-school is greater than zero; 0 otherwise. Standard errors, clustered at the child-level, are presented in parenthesis. This specification includes Territorial Unit (TU) fixed effects, in addition to the usual ones. The mean of the dependant variable is calculated at baseline (before the reform) and for the corresponding treated group. See notes of Table 3.6 for the definition of the treatment status.

Table C.10: Average Treatment Effects on Time Spent in Different Childcare Facilities Excluding 2019

	(1) Daycare		(2) Childminder		(3) Relative		(4) Pre-school	
	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits	(a) Half Benefits	(b) Quarter Benefits
Panel A: 0–5 years-old								
Half Benefits	0.448		1.106		-0.392		-0.459	
	(1.247)		(0.843)		(0.443)		(1.306)	
Post	-12.539***	-12.060***	-3.446***	-2.691***	-0.017	-0.272	65.031***	64.640***
	(1.311)	(1.353)	(0.738)	(0.738)	(0.571)	(0.586)	(1.609)	(1.697)
Post × Half Benefits	-1.341		-3.115***		-0.342		0.627	
	(1.563)		(1.028)		(0.610)		(1.491)	
Quarter Benefits		1.415		1.554		0.538		-3.219
		(2.152)		(1.208)		(0.734)		(2.462)
Post × Quarter Benefits		-5.512**		-2.473		-0.996		3.976*
		(2.739)		(1.950)		(0.978)		(2.382)
Mean Dep. Var	10.015	10.281	2.619	4.410	1.868	2.236	20.041	19.878
Observations	1640	1070	1642	1069	1635	1062	1232	835
Panel B: : 0–2 years-old								
Half Benefits	-1.648		1.247		-0.239			
	(2.638)		(1.660)		(1.070)			
Post	31.104***	28.735***	4.322**	6.387***	5.202***	3.634**		
	(3.456)	(3.547)	(2.165)	(2.440)	(1.973)	(1.845)		
Post × Half Benefits	5.722		-5.654**		-1.472			
	(3.486)		(2.554)		(2.035)			
Quarter Benefits		-4.924		3.263		0.281		
		(5.209)		(2.814)		(0.549)		
Post × Quarter Benefits		7.489		-6.587		-0.873		
		(6.376)		(4.748)		(0.951)		
Mean Dep. Var	18.892	19.534	4.117	6.531	2.250	2.650		
Observations	648	401	649	401	644	397		
Panel C: : 3–5 years-old								
Half Benefits	-0.232		0.023		-0.478		-0.485	
	(1.317)		(0.976)		(0.540)		(1.653)	
Post	-15.046***	-15.554***	-4.276***	-3.196***	-1.536	-1.708	33.739***	33.894***
	(1.829)	(1.997)	(1.169)	(1.214)	(1.069)	(1.153)	(2.586)	(2.777)
Post × Half Benefits	-0.681		-1.078		0.586		1.189	
	(1.642)		(1.103)		(0.568)		(1.891)	
Quarter Benefits		1.900		2.384*		-0.506		-3.457
		(2.463)		(1.275)		(1.300)		(3.341)
Post × Quarter Benefits		-4.333		-1.167		0.321		2.613
		(3.837)		(1.774)		(1.073)		(3.666)
Mean Dep. Var	4.216	4.735	1.640	3.136	1.620	1.989	24.631	24.813
Observations	992	669	993	668	991	665	994	667
Panel D: : 4–5 years-old								
Half Benefits			-0.223		-0.420		-0.502	
			(0.569)		(0.858)		(2.039)	
Post			0.502	1.136	0.547	0.534	0.779	1.953
			(1.395)	(1.526)	(1.641)	(1.840)	(3.598)	(3.806)
Post × Half Benefits			-0.281		0.256		1.135	
			(0.747)		(0.829)		(2.218)	
Quarter Benefits				0.522		-1.329		-2.114
				(1.285)		(2.840)		(5.165)
Post × Quarter Benefits				-1.131		0.014		4.588
				(1.156)		(1.541)		(3.287)
Mean Dep. Var			0.915	2.946	1.425	1.728	27.838	28.030
Observations			704	467	702	464	705	467
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS results for Equation (3.3), ran separately by treatment status, on the number of hours per week spent (1) in daycare center, (2) with a professional childminder, (3) cared by a relative, and (4) in pre-school. In this exercise, we drop the year 2019 — corresponding to the enforcement of compulsory school from the age of three. Standard errors, clustered at the child-level, are presented in parenthesis. The mean of the dependant variable here is calculated at baseline (before the reform) for the corresponding age-group, and includes all treated and control groups. See notes of Table 3.6 for the definition of the treatment status.

C.2 Technical Details

C.2.1 Additional Information on the French Setting

Over the past two decades, formal childcare usage in France — characterized by daycare centers (*crèches*) and professional childminders (*assistantes maternelles*) — has expanded substantially, particularly among children under the age of three. According to a recent report from Onape (2024), between 2002 and 2021, the share of children primarily cared for by a crèche nearly doubled (from 9% to 18%), while reliance on professional childminders also increased (from 13% to 21%), reflecting a broader shift away from exclusive parental care (56% in 2021).[1] Furthermore, by 2022, approximately 50% of children under three had used at least one form of formal childcare subsidized by the family branch, though this usage varied significantly by household structure and income. For instance, lower-income families are three times less likely to use these formal childcare arrangements (23%) compared to higher-income households (74%), and households in which the two parents are inactive are five times less likely (17%) than dual-earner households (88%) highlighting persistent inequalities in access to formal care.

Daycare centers and childminders are two systems differing substantially in their modes of operation, financing structures, and cost implications for families, although both are deeply included within a publicly-supported childcare system. On the one hand, crèches — including municipal crèches, parental co-operatives, and micro-crèches — are typically managed either by local governments or private organizations and receive substantial public subsidies. The most common funding mechanism for these is the Prestation de Service Unique (PSU), a public subsidy. To receive this support, crèches must charge families fees based on a national scale, which takes into account household income and the number of dependent children. This means-tested approach ensures that fees remain affordable, especially for low- and middle-income households. On average, a dual-earner family earning around two times the minimum wage pays approximately €138 per month for full-time care (162 hours per month) in a PSU-funded crèche. Additionally, micro-crèches—smaller facilities with fewer than 10 children—are most often financed indirectly via the Complément de Libre Choix du Mode de Garde (CMG), which reimburses families directly. Municipalities also contribute to the operational and capital costs of crèches, notably through programs like the "crèche investment plans" and bonuses promoting social and geographic diversity in care provision.

Childminders, by contrast, are independent childcare providers certified by local child protection authorities, usually welcoming children into their homes. As of 2023, over 227,000 childminders were active in France, collectively offering 684,600 childcare places representing more than half of the country's formal childcare capacity (Onape, 2024). Families who employ a certified childminder receive financial assistance primarily through the CMG, which partially reimburses wages and covers 100% of social security contributions. This aid is income-dependent and capped. After receiving the CMG, families are also eligible for a tax credit of 50% on remain-

[1] The last five percent for 2021 are distributed amongst grandparents or other relatives than parents (3%) and pre-school (2%).

ing expenses (up to €1,750 per child annually). For a typical dual-earner family on modest income, the average out-of-pocket cost for full-time care by a childminder is about €306 per month, notably higher than for a crèche. Still, childminders offer greater flexibility and personalized care (Barrère-Maurisson and Lemière, 2006), making them a preferred option for many families. While the funding structure for crèches centers on direct public subsidies to institutions, the system for childminders relies more heavily on subsidies and tax incentives paid directly to families. These distinctions reflect broader policy choices: crèches are more regulated and standardized, whereas childminders provide a more individualized and decentralized form of care. Together, these two systems constitute the backbone of France's formal early childhood care infrastructure, with total public spending on childcare for children under three amounting to €16.7 billion in 2023 (Onape, 2024).

C.2.2 Variables Description

C.2.2.1 Household Financial Resources and Deprivation

Poverty measures.

The first measure considered for being *in poverty* is the the binary variable provided by SRCV set to one if an individual lives in a household whose standard of living is below the poverty line. This is determined in relation to the distribution of living standards for the population as a whole, and equal to 60% of the median standard of living. Note that standards of living are calculated as the ratio between disposable income (corrected by an inflation factor for within-household non-response) and the number of consumption units within the household.

We also use the variable provided by SRCV for the poverty in living conditions. Material *living conditions poverty* is defined by the presence of multiple deprivations in areas essential for basic well-being or difficulties in daily life. According to Insee (the French National Institute of Statistics and Economic Studies), a household is considered poor in living conditions if it experiences at least 8 out of 27 defined deprivations or difficulties. These include:

1. Financial constraints: high loan repayments relative to income, insufficient income to balance the household budget, reliance on savings, lack of savings, frequent overdrafts, or self-assessment of financial hardship.
2. Payment delays: inability to pay electricity/gas bills, rent, or taxes over the past year.
3. Consumption restrictions: difficulty maintaining adequate home temperature, affording a week of vacation, replacing furniture, buying warm clothing, purchasing meat, hosting guests, giving gifts, or owning two pairs of shoes.
4. Housing difficulties: overcrowding, lack of essential facilities (bathroom, toilets, heating, or hot water), small or poorly heated homes, humidity, and noise.

Finally, we define a binary variable related to a *subjective measure for being in financial need*. This variable is set to one if the reported (thus subjective measure of) income needed to make ends meet is greater than the household's total income.

Table C.11: Pairwise Correlation Coefficients of Household Financial Resources Measures

	(1) Savings Above Median	(2) Poverty	(3) Poverty in Living Conditions	(4) Self-Reported Hardhsip
(1)	1.0000			
(2)	-0.2874*	1.0000		
(3)	-0.3215*	0.3279*	1.0000	
(4)	-0.2474*	0.3424*	0.2040*	1.0000

Notes: * $p < 0.01$

C.2.2.1.1 Deprivation measures.

For our deprivation measures, we use each item asked in the SRCV, *i.e.*, whether a household is unable (reverse-coded) to:

1. Eat meat (fish) every other day.
2. Buy new clothes.
3. Own two pairs of good shoes.
4. Replace worn-out furniture.
5. Keep home at a comfortable temperature.
6. Host relatives.
7. Do gifts.
8. Afford holidays.
9. Cover unexpected expenses.

Further, we construct a deprivation score, which is the sum of each items the household is unable to afford, as well as four other measures given the pairwise correlation (r) between each item: (i) *house deprivation* is defined as the inability to either keep the home at a comfortable temperature and/or replace furniture ($r = 0.31$, $p < 0.01$); (ii) *clothing deprivation* is based on the inability to buy new clothes and/or own two pairs of good shoes ($r = 0.55, p < 0.01$); (iii) *food deprivation* captures whether households are unable to afford meat or fish every other day; and (iv) *leisure deprivation*, which is the broadest category, reflecting the unability to host relatives, do gifts, afford holidays, and/or cover unexpected expenses (each $r > 0.35, p < 0.01$). The pairwise correlation coefficients of these four deprivation measures are presented in Table C.12 below.

Table C.12: Pairwise Correlation Coefficients of Household Deprivation Measures

	(1) House Deprivation	(2) Clothing Deprivation	(3) Food Deprivation	(4) Leisure Deprivation
(1)	1.0000			
(2)	0.5684*	1.0000		
(3)	0.3448*	0.4366*	1.0000	
(4)	0.5918*	0.3424*	0.2670*	1.0000

Notes: * $p < 0.01$