Traditional Views, Egalitarian Views, and the Child Penalty: Insights from Immigrant Populations in France*

Dominique Meurs, Paris Nanterre University and Ined,
Pierre Pora, Insee, Crest and Ined†

Link to latest draft

Abstract

This study examines whether the child penalty is driven by traditional gender attitudes using French survey data on immigrants. While women with traditional views participate less in the labor force, their child penalty is not larger. Consequently, the child penalty explains a greater share of gender gaps among immigrants with egalitarian views. Comparative analyses across upbringing environments and countries of origin confirm the absence of a causal link between gender attitudes and the child penalty. The findings suggest that as norms become more egalitarian, the child penalty accounts for a growing share of gender disparities in labor outcomes.

Keywords: Gender, Child penalty, Immigrants, Attitudes, Values

JEL Classification: J15, J16, J24, D63

*We thank Christel Colin, Sylvie Le Minez, Marion Leturcq, Éric Maurin, Hélène Périvier, Léa Pessin, Simon Rabaté as well as attendees at Ined, Insee, Sciences Po (LIEPP) and PSE (Chaire Travail) seminars and AASLE (Bangkok, 2024) and AFÉPOP (Marseille, 2025) for useful suggestions. All remaining errors and opinions are ours.

†Corresponding author, pierre.pora@ensae.fr

1 Introduction

In developed countries, the child penalty – the significant decline in women's labor outcomes following the birth of children – remains the leading driver of gender inequality in the labor market (Angelov, Johansson, and Lindahl, 2016; Kleven, Landais, and Søgaard, 2019; Cortés and Pan, 2023). Numerous studies have examined potential explanations for this phenomenon. These investigations have shown that the child penalty is not linked to biological differences in reproductive contributions or productivity disparities in the labor market (Kleven, Landais, and Søgaard, 2021). Evidence from policy reforms suggests that family policies, such as affordable childcare or parental leave, have limited long-term effects on the child penalty (Rabaté and Rellstab, 2022; Kleven et al., 2024). As a result, the role of institutions is largely ruled out, leaving norms and preferences as the most plausible explanation. This hypothesis is further supported by various spatial comparisons, both across countries and regions (Kleven et al., 2019; Rabaté and Rellstab, 2022; Casarico and Lattanzio, 2023; Kleven, Landais, and Leite-Mariante, 2024; Kleven, 2022), as well as more recent studies on migration, encompassing both internal and international migrants (Boelmann, Raute, and Schönberg, 2020; Rabaté and Rellstab, 2022; Kleven, 2022).

While there is now substantial and growing evidence that the child penalty is shaped by norms and preferences, the literature largely remains silent on the exact nature of the underlying beliefs and preferences. This question is significant both theoretically and practically. From a theoretical point of view, it can help explain why the child penalty persists, despite significant convergence in men's and women's outcomes, and the decreasing prevalence of negative attitudes towards working women. Practically, this issue has important policy implications, as the design of effective interventions is likely to require a deep understanding of the beliefs that such policies are intended to address.

This paper explores the hypothesis that the child penalty is influenced by traditional gender-related attitudes. We define these attitudes as negative views on female employment and women's autonomy. Such attitudes are known to be strongly correlated with gender gaps in labor force participation, both across countries (Fortin, 2005) and over time (Fortin, 2015), and to decrease maternal labor supply both at the individual level and through peer effects (Cavapozzi, Francesconi, and Nicoletti, 2021). Changes in these attitudes also help explain why historical shocks that increased female labor force participation during the twentieth century had persistent effects over time (Gay, 2023). Given the prominence of the child penalty as a key driver of gender gaps, this investigation is therefore highly relevant.

To investigate this, we draw on individual-level survey data about opinions and attitudes, which have rarely been utilized in the child penalty literature. Most studies in this field rely on high-quality administrative data, which allows a detailed analysis of labor market histories. However, administrative data are typically not well suited for exploring beliefs and opinions, as they are not collected for this purpose. In contrast, the survey data we use enable us to link individual beliefs and attitudes about female employment and women's autonomy to labor supply decisions that contribute to the child penalty. This allows us to quantify the role of traditional gender-related beliefs and attitudes in the child penalty.

Specifically, we rely on recent French survey data on immigrants (both first- and second-generation) living in France between 2019 and 2020 (TeO2 survey). The key advantage of these data is that they provide individual-level information on gender-related beliefs collected at the time of the survey, alongside retrospective labor market histories and comprehensive fertility decisions, both of which are crucial for estimating the child penalty.

This allows us to make three distinct comparisons that form the core of the paper. First, at the individual level, we compare immigrants with more traditional gender views to those with more egalitarian perspectives. Second, building on evidence that immigrants raised in more traditional family environments tend to hold more traditional gender views as adults, we compare immigrants based on their family environment before the age of 18. Lastly, drawing on the correlation between countries of origin and beliefs about traditional gender roles, we compare second-generation immigrants – who all grew up in France – based on whether their parents were born in countries from which first-generation immigrants are more likely to hold traditional gender views.

Empirically, we find no variation in the child penalty along these dimensions, which appears to contradict the idea that traditional gender beliefs are a major driver of the child penalty. We conclude that this hypothesis is not supported by the data. However, this finding does not imply that norms and preferences are not factors contributing to the child penalty, as norms and preferences have many dimensions beyond the one we focus on. However, it does suggest that the child penalty is largely unrelated to negative views regarding female employment or women's autonomy.

We provide further evidence that this finding does not undermine the notion that such beliefs do influence gender differences in labor supply. Ultimately, it remains true that (i) traditional gender-related attitudes contribute to gender gaps in the labor market, and (ii) the larger share of gender differences in labor outcomes arises from the child penalty, but these two facts appear to be independent of one another. The validity of both statements suggests, in particular, that the significance of the child penalty as a key driver of gender gaps in labor outcomes is greater when attitudes towards female employment and women's autonomy are more positive.

This investigation faces three main empirical challenges. The first concerns the measurement of beliefs and attitudes. We rely on several opinion items from the TeO2 survey, which specifically address gender inequality and women's bodily autonomy, as well as religion and political orientation, to construct a one-dimensional index of gender-related beliefs and attitudes. The second challenge involves the measurement of the child penalty. We build upon the now-standard event-study approach developed by Angelov, Johansson, and Lindahl (2016) and Kleven, Landais, and Søgaard (2019). The third challenge pertains to identifying the effect of beliefs and attitudes, given that we use opinions reported at the time of the survey to explain labor supply decisions made several years prior. To address this issue, we exploit variation that predates these decisions, both in terms of timing and causality, such as differences in immigrants' family environments during childhood and comparisons of second-generation immigrants based on their parents' countries of origin.

This paper contributes to the extensive literature on gender inequality, which has been notably highlighted by the recent Nobel Prize awarded to Claudia Goldin. Empirical research

on this topic has grown significantly in recent years, with two main trends. The first is a shift away from analyses that focus on labor demand towards those that emphasize the labor supply side. Empirically, this shift corresponds to studies that quantify gender inequalities in terms of annual labor earnings — that is, the total salaried earnings over a year, without conditioning on salaried employment — rather than hourly wages, allowing for the inclusion of interruptions in working life.

This shift has been partly driven by a second trend, which emphasizes the role of family life, and particularly parenthood, in generating gender gaps in the labor market (Juhn and McCue, 2017). The spread of contraceptives, by allowing young women to control their childbearing schedules, has led to a change in women's attitudes toward the labor market (Goldin and Katz, 2002; Goldin, 2006), contributing to the narrowing of gender gaps in OECD countries since the 1970s. However, family responsibilities continue to fall disproportionately on women. As organizing daily life often requires balancing family and work, the two spheres compete (Goldin, 2021), generating not only mechanical gender differences in labor market participation and working hours, but also gender pay gaps — either due to higher anticipated separation rates for women (Lazear and Rosen, 1990) or because highly skilled occupations disproportionately reward time availability and long hours (Goldin, 2014).

A particularly salient strand of the growing literature focuses on the child penalty, which refers to the significant decline in mothers' labor outcomes following the arrival of children (Angelov, Johansson, and Lindahl, 2016; Kleven, Landais, and Søgaard, 2019). This effect has now been consistently quantified across a variety of countries (see e.g. Kleven et al., 2019; Meurs and Pora, 2019; Sieppi and Pehkonen, 2019; Quinto, Hospido, and Sanz, 2021; Rabaté and Rellstab, 2022; Casarico and Lattanzio, 2023; Lebedinski, Perugini, and Vladisavljević, 2023), and more recently, globally (Kleven, Landais, and Leite-Mariante, 2024). The evidence confirms that, in highly developed countries, gender inequality in the labor market is primarily driven by the child penalty.

While specific evidence on the contribution of family policies to the child penalty is lacking in the French case studied in this paper, evidence showing that increasing the availability of affordable daycare slots does not improve maternal labor outcomes (Pora, 2020), or that earmarking parental leave does not shift the burden to fathers (Périvier and Verdugo, 2024), is consistent with financial incentives, such as those provided by family policies, not playing a major role in the child penalty.

These negative findings have prompted interest in explanations linked to beliefs and identity, in the spirit of Akerlof and Kranton (2000), thereby connecting the child penalty literature to a body of work that has examined the role of beliefs and attitudes regarding gender in shaping labor supply decisions. While some studies have identified gender norms through anomalies in the data (see e.g. Bertrand, Kamenica, and Pan, 2015), a significant portion of this literature has drawn on self-reported survey data to elicit beliefs and attitudes regarding gender. This approach has enabled researchers to demonstrate that beliefs prioritizing traditional gender roles are associated with lower female labor force participation, both across countries (Fortin, 2005) and over time (Fortin, 2015).

In the specific context of the child penalty, several authors have attempted to estimate the

contribution of beliefs and attitudes by using a variety of spatial comparisons: across countries, based on beliefs measured in the ISSP data (Kleven et al., 2019); across US states, using State-level average beliefs and attitudes regarding gender derived from the GSS data (Kleven, 2022); across Italian regions, based on sensitivity to traditional gender norms captured in the European Value Study (Casarico and Lattanzio, 2023); and across Dutch municipalities, relying on average religious attendance (Rabaté and Rellstab, 2022). Other studies have focused on same-sex couples (Andresen and Nix, 2022). The findings of these studies can vary, with some reporting strong correlations, while others observe limited differences. Additionally, in the former case, even in the most egalitarian areas in terms of gender-related attitudes, the child penalty remains substantial. Individual-level evidence based on elicited beliefs and attitudes is still rare, with the notable exception of Lebedinski, Perugini, and Vladisavljević (2023), who use comparisons based on levels of self-reported religiosity in Russia. In contrast, our paper primarily draws insights from individual-level comparisons enabled by the rich survey data we use, which include directly elicited beliefs about gender, without relying on proxies like religion.

However, even properly measured individual-level beliefs may not provide convincing evidence, as these beliefs may be endogenous outcomes of adulthood experiences and labor supply decisions (Borrell-Porta, Costa-Font, and Philipp, 2018; Kuziemko et al., 2018; Borrell-Porta, Contreras, and Costa-Font, 2023). This is why a significant portion of the literature on the impact of culture on labor supply has focused on immigrants, especially second-generation immigrants. This approach allows researchers to apply what is known as the "epidemiological approach" to culture (Fernández, 2007; Fernández and Fogli, 2009; Fernández, 2011). The intuition behind this approach is that second-generation immigrants offer a valuable case for distinguishing between culture, on the one hand, and institutions and policies, on the other, since they were born and raised within similar institutions and markets, while their beliefs and attitudes are often strongly correlated with those of their parents' countries of origin (Alesina, Giuliano, and Nunn, 2013). Thus, if second-generation immigrants' labor supply decisions are uncorrelated with their parents' origins, it would provide a strong argument against the role of culture and beliefs.

The epidemiological approach has recently gained traction in the child penalty literature. Kleven (2022) finds no correlation between the child penalties of second- or later-generation immigrants in the US and those observed in their parents' countries of origin. Our comparison in Section 5 differs in two important ways. First, we focus exclusively on second-generation immigrants rather than on all later generations sharing the same ancestry. In this context, finding no difference would imply a much faster pace of cultural convergence than what Kleven (2022)'s results suggest. Second, the question we address is conceptually distinct. By linking immigrants' child penalties to those in their countries of ancestry, Kleven (2022) tests whether these penalties are rooted in culturally transmitted norms and attitudes. We take this cultural foundation as given and instead ask whether the child penalty reflects specific hostility toward working women. To that end, we compare child penalties across second-generation immigrants according to the gender-related beliefs held by first-generation immigrants from the same country of origin.

Similarly, Rabaté and Rellstab (2022) examines immigrants in the Netherlands, though it

does not distinguish between generations, complicating interpretation. Other studies complement this perspective: Ichino et al. (2024) combine the epidemiological approach with exogenous tax reforms in Sweden, showing that couples from conservative cultures reinforce traditional childcare roles when taxes are adjusted. Meanwhile, Boelmann, Raute, and Schönberg (2020) reveal that East German mothers retain egalitarian employment attitudes despite migrating to West Germany, while West German mothers partially align with East German norms, suggesting a greater persistence of egalitarian attitudes.

In line with this trend, our paper focuses on immigrants (both first- and second-generation), primarily due to data limitations. We also conduct an empirical investigation similar to the epidemiological approach, though it constitutes only one piece of the evidence we present. Consistent with the other comparisons we examine, we find no correlation between the child penalty and gender-related attitudes inferred from the countries of birth of second-generation immigrants' parents.

The paper proceeds as follows. Section 2 introduces the database used for our first two comparisons and outlines the construction of the variables of interest, including maternity and paternity dates, labor market histories, and the opinion and attitude variable derived from questions in various TeO2 modules. Section 3 describes our empirical framework. The results from our first two comparisons, namely those based on current attitudes and past family environments, are presented in Section 4. Section 5 details our third empirical test, an adaptation of the epidemiological approach focusing on the countries of birth of second-generation immigrants' parents. The paper concludes with a summary of the findings. Appendices include robustness checks and alternative specifications.

2 Data

This paper relies on the *Trajectoires et Origines 2* (TeO2) survey, conducted in France between 2019 and 2020 by Insee (the French statistical office) and Ined (the French institute for demographic studies). This survey focuses on the lived experiences of immigrants and provides rich data on values and attitudes, including opinions on gender inequality and women's bodily autonomy. It also includes a detailed retrospective labor market calendar and comprehensive information about respondents' children.

This section explains how we combine these data elements and presents summary statistics for our final sample.

2.1 *Trajectoires et Origines 2* survey : general overview

The *Trajectoires et Origines 2* (TeO2) survey is a large-scale statistical study conducted by Insee and Ined between July 2019 and November 2020. It includes responses from over 27,000 individuals aged 18 to 59 living in ordinary dwellings in mainland France. This second edition of the TeO survey, initially conducted in 2008–2009, aims to provide national statistics on population diversity and to explore the influence of migratory origins on individual trajectories. Given its focus, the sample intentionally over-represents first- and second-generation immigrants. Further methodological details are provided in Thao Khamsing et al. (2022).

The survey covers a wide array of topics, from migration histories to housing, experiences of discrimination, and cultural practices. For the purposes of this paper, three specific features of the survey are particularly relevant:

- It includes questions on attitudes and opinions, which we use to infer gender-related attitudes;
- It collects detailed information on all children born to immigrant parents, regardless of their current living arrangements;
- It provides a retrospective calendar of labor market outcomes, enabling the reconstruction of entire career trajectories rather than snapshots of current employment.

These three features form the foundation of our analysis, allowing us to explore the connection between child penalties and gender-related attitudes.

2.2 Eliciting gender-related attitudes

A central aspect of this study is the identification of gender-related attitudes from the responses provided in the TeO2 survey. The following sections outline the methodology used to extract and interpret these attitudes.

2.2.1 Relevant survey items

The TeO2 survey provides a comprehensive perspective on the lived experiences of immigrants in France. In addition to standard demographic and occupational data, it includes questions on values, attitudes, religion, and social life. This study relies on six specific survey items to assess the degree of traditional gender views among respondents.

The first three items, extracted from the "Attitudes and Opinions" module, address key aspects of gender-related beliefs:

I_GENRE When there are not many jobs, men are more entitled to work than women. 1. Totally agree; 2. Agree; 3. Disagree; 4. Totally disagree.

I_AVORT A woman can have an abortion for non-medical reasons. 1. Totally agree; 2. Agree; 3. Disagree; 4. Totally disagree.

I_HOMO Same-sex couples should have the same rights as different-sex couples. 1. Totally agree; 2. Agree; 3. Disagree; 4. Totally disagree.

The first item, highlighted by Fortin (2005), is strongly correlated with international gender disparities in labor force participation. The second item captures attitudes toward women's autonomy and their societal role as mothers. The third item is relevant because it touches on adoption and parenthood rights for same-sex couples, linking gender and family structures. Additionally, lower child penalties observed among same-sex couples suggest the influence of gender identity on mothers' labor supply (Andresen and Nix, 2022).

We also include three additional items from other modules of the survey. These items capture complementary dimensions of gender-related attitudes:

A_RHOM and A_RFEM [For respondents who met friends within the two weeks prior to the survey] Among these friends, how many are of the same gender as the respondent? 1. Almost all; 2. More than half; 3. Half; 4. Less than half; 5. Almost none or none.

R_IMPVIE [For respondents who currently practice a religion] How important is religion in your life today? 1. Not at all important; 2. Moderately important; 3. Important; 4. Very important.

I_OPIPOL [On political opinion] Would you say that you are...? 1. On the far left; 2. On the left; 3. Centrist; 4. On the right; 5. On the far right; 6. Neither right nor left.

The first item reflects gender segregation within friendship networks, providing a behavioral perspective on gender identity. The second captures religiosity, which has been linked to more traditional gender attitudes (Seguino, 2011). The third highlights political orientation, a strong predictor of gender equality views (Sevincer et al., 2023).

Both the first and second items apply only to respondents who met specific conditions—having interacted with friends recently or practicing a religion. We account for this by including additional levels in the categorical variables: "no recent social interactions" for the first item and "no religion" for the second.

We acknowledge that combining these diverse dimensions into a single measure is inherently complex. Each item captures a different aspect of gender-related attitudes, and integrating them requires careful consideration. To address potential concerns, we perform a robustness check by focusing exclusively on the **I_GENRE** item, an approach similar to Fortin (2005). Specifically, we divide the sample based on whether respondents totally disagreed with the statement or not. This simplified approach produces results that are largely consistent with our main findings.

2.2.2 Using PCA to infer gender-related attitudes

To explore the relationship between child penalties and gender-related attitudes, this paper requires a low-dimensional—preferably one-dimensional—measure of such attitudes. However, deriving this measure from six multinomial variables is not straightforward. We address this challenge by adopting a Principal Component Analysis (PCA) approach. Specifically, we focus on a restricted sample comprising immigrant parents who responded to all relevant survey items,[1] and estimate the first principal component of the dataset. This synthetic continuous variable is then interpreted as capturing the degree of traditionality in gender-related attitudes.

For ease of comparison, we discretize the first principal component by categorizing individuals as either above or below the estimated population median. Respondents above the median are interpreted as holding more traditional views regarding gender inequality and women's bodily autonomy compared to those below the median.[2]

[1] This includes individuals who reported having no friends of the same gender or no religious affiliation, as these categories are accounted for in the categorical variables.

[2] Baseline estimates compare immigrant parents above and below the median. Similar conclusions are drawn when comparing respondents in the top tertile with those in the bottom tertile. These checks are detailed in a separate note available upon request from the authors.

Consistent with the rationale, immigrants above the median are more likely to favor male employment over female employment, oppose non-medical abortion and equal rights for same-sex couples, maintain more gender-segregated friendships, display higher levels of religiosity, and identify more frequently as politically right-leaning or apolitical.

2.2.3 Interpretation of the PCA analysis

Because the TeO2 survey is not specifically focused on gender and family dynamics, one might question the inclusion of the specific items on values and attitudes used in this analysis. However, we demonstrate that such concerns do not undermine our results. Firstly, immigrants' responses to these items are strongly correlated, indicating that they capture a low-dimensional aspect of values and attitudes. Secondly, summary statistics reveal that these inferred gender-related attitudes are significantly associated with gender gaps in labor market outcomes (see 2.5) and unequal distributions of household chores.

Finally, our results remain robust even when excluding certain survey items. Dropping any particular item does not notably change how respondents are categorized or the overall conclusions regarding child penalties.[3]

2.3 Fertility measurement

The TeO2 survey includes questions about respondents' children, specifically gathering the total number of children they have had over their lifetime, regardless of whether: (i) these children still live with them; (ii) these children were born in France; or (iii) these children currently reside in France. Additional details about each child are also available in the survey data. Of particular importance to this paper is the inclusion of the children's birth dates. Our identification of the child penalty is based on the timing of the birth of the immigrant's first child. In this context, the first child is defined broadly, irrespective of whether: (i) the child was born in France; (ii) the child is still alive; or (iii) the child currently lives in France.

2.4 Labor market outcomes

The TeO2 survey gathered data on immigrants' professional careers. This feature is essential to our analysis as we aim to examine the child penalty over the entire career lifecycle, rather than focusing solely on cross-sectional labor market disparities. Specifically, the survey recorded up to 15 periods of career-related data, beginning either after formal education or at age 14. Each period was required to last at least one year, with the following information provided for each period: (i) the year it began; (ii) the year it ended; and (iii) the respondent's status during that time, including whether they were employed as a salaried worker, unemployed, in education, out of the labor force, or in another status.

We rely on this feature to reconstruct immigrants' labor market participation and employment trajectories over the lifecycle on a yearly basis. Because this paper focuses on the French context, and to ensure comparability with our administrative registers that do not include information on labor market outcomes out of France, for first-generation immigrants we restrict

[3] These results are detailed in a separate note available upon request from the authors.

ourselves to their labor market trajectories from the first year they began to live in mainland France.

2.5 Sample construction and summary statistics

Our baseline sample is drawn from the TeO2 dataset. Specifically, we focus on first- and second-generation immigrants who have lived at least once in a cohabiting relationship that lasted at least 6 months, with at least one child, regardless of whether the child still resides with the parent. Individuals who reported having children before the age of 15 are excluded. To ensure sufficient labor market history for identifying child penalties, we limit our sample to those born before 1995. All combined, these restrictions reduce the sample size from 27,200 to 11,900 respondents. Further, our empirical strategy relies on information about immigrant parents' attitudes, which leads to a reduction in the sample size to 7,000 individuals due to non-responses to relevant survey items. Finally, we also incorporate information on the environments in which immigrants were raised. Non-responses to these specific questions leave us with 6,900 respondents.

Table 1 presents summary statistics for our sample. After appropriate weighting, the sample represents a total population of 2.9 million first- and second-generation immigrant parents living in France. This population includes individuals born between the 1960s and 1990s, with an average year of birth in the 1970s.

We divide the data based on whether respondents hold more or less traditional views regarding gender. A key finding from this exercise is that individuals with less traditional views tend to be slightly older and more educated than those with more traditional gender-related attitudes. These individuals are also more likely to be second-generation immigrants.[4] Finally, those with less traditional views generally had their first child at an older age, particularly among women, and have fewer children overall than those with more traditional views.

Figure 1 shows the lifecycle profiles of labor force participation, segmented by gender and attitudes. While differences in labor market participation between attitude groups are minor for fathers, they are substantial for mothers. This suggests that the gender gap in labor force participation is more pronounced among first and second-generation immigrant parents who hold more traditional views about gender compared to those who hold less traditional views. Specifically, at age 45, the labor force participation gap is 16 percentage points for those with more traditional views, compared to 7 percentage points for those with less traditional views. This supports our interpretation that the first component of the PCA analysis captures attitudes related to gender.

The divergence in labor force participation rate across attitudes groups begins at age 20, which coincides with the start of childbearing ages for our population of interest. The remainder of this paper aims at clarifying whether these fertility events can actually explain this divergence.

[4] More traditional attitudes are less common among immigrants from Western Europe compared to those from African or Middle Eastern countries.

3 Empirical analysis

The empirical analysis of this paper aims to quantify the effect of gender-related attitudes on the child penalty – specifically, how these attitudes influence the impact of children on the labor market outcomes of parents, particularly mothers. An ideal, though infeasible, solution would be to conduct an experiment in which both children and gender-related beliefs are randomly assigned to potential parents. Since we cannot rely on this infeasible solution, our analysis must address three key issues.

Measurement issue In principle, assigning beliefs would have the beneficial side effect of ensuring that beliefs are known to the experimenter. However, this is not the case in practice, which is why we must infer gender-related attitudes from individual responses to our survey data. Our approach is detailed in 2.2.

Endogenous fertility decisions Children are not randomly assigned; rather, they result from fertility decisions that partly depend on parents' expectations regarding labor market outcomes. Consequently, the comparison between parents and non-parents may not yield a causal interpretation. We address this issue by employing an approach that improves on the now standard event-study approach to identify the child penalty (Angelov, Johansson, and Lindahl, 2016; Kleven, Landais, and Søgaard, 2019). Subsection 3.1 provides further details on our approach.

Endogenous gender-related attitudes Even in the best-case scenario, where gender-related attitudes could be observed directly, we would still face the issue of their endogeneity. This issue is particularly salient in this context because we rely on current self-reported beliefs to infer gender-related attitudes, while we measure the effect of children based on past fertility decisions and labor outcomes. As a result, gender-related attitudes may be the product of an *ex-post* rationalization of past fertility and labor supply choices. This would imply that comparisons of child penalties across different gender-related attitudes cannot be interpreted causally. Our solution to this issue is twofold: (i) we ensure that the immigrants we compare across varying gender-related attitudes are as similar as possible in terms of observable characteristics relevant to the child penalty; (ii) we compare immigrants raised in different environments, with respect to dimensions strongly correlated with current gender beliefs, which are causally anterior to the fertility and labor supply decisions that lead to the child penalty. Subsection 3.2 provides further details on these solutions.

3.1 Child penalty estimation

Our approach to estimating the child penalty builds on the widely used event-study method in this context (Angelov, Johansson, and Lindahl, 2016; Kleven, Landais, and Søgaard, 2019). This method relies on a difference-in-differences framework that compares parents from the same generation based on their age at first child. If the impact of children on their parents' labor outcomes is negligible before birth, and if age at first child is not correlated with a

more pronounced upward trend in labor market outcomes, then the labor market trajectories of parents with a later-born first child serve as a valid counterfactual for those with an earlier-born first child. The remainder of this subsection formalizes this intuition, details its implementation, and discusses the relevance of this approach in the context of this paper.

3.1.1 Model and identification

Our analysis builds on the event-study approach proposed by Kleven, Landais, and Søgaard (2019), with a slight improvement by incorporating insights from recent difference-in-differences literature (see Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). Specifically, our approach aims to address identification issues related to the use of two-way fixed effects in settings where treatment effects are likely to be heterogeneous.[5] The presentation of our empirical framework largely follows Callaway and Sant'Anna (2021) and Sun and Abraham (2021).

Our approach relies on within-gender, within-generation (year of birth)[6] across-age-at-first-child comparisons. Specifically, let $Y_{i,t}$ denote the relevant labor outcome – e.g., labor earnings or labor market participation – of individual i at age t. Let $G_i \in \{f, m\}$ denote individual i's gender and B_i her generation (year of birth).[7] Finally, let C_i denote her age when her first child was born ($C_i = \infty$ if she is childless).

Let $Y_{i,t}(c)$ denote the potential labor supply of individual i at time t had she given birth to her first child at age c. Consistently, $Y_{i,t}(\infty)$ represents her labor supply at time t had she chosen to remain childless. By construction:

$$Y_{i,t} = Y_{i,t}(\infty) + \sum_c (Y_{i,t}(c) - Y_{i,t}(\infty))\mathbb{1}\{C_i = c\} \tag{1}$$

Our analysis focuses on the causal effect of parenthood on labor supply. In other words, we are interested in (functionals of) the distribution of random variables $Y_{i,t}(c) - Y_{i,t}(\infty)$, with $c < \infty$. Specifically, we define the gender-generation-cohort-specific average treatment effect on the treated:

$$CATT_{g,b,c,t} = \mathbb{E}[Y_{i,t}(c) - Y_{i,t}(\infty) \mid G_i = g, B_i = b, C_i = c] \tag{2}$$

This quantity corresponds to the effect of being $t-c$ years away from the birth of one's first child, for those individuals of gender g, born in b, who gave birth to their first child at age c. These average treatment effects are not conditional on possible subsequent childbirths. As a result, they incorporate both the causal effect of motherhood at the extensive margin, i.e., choosing to be a parent or not, and at the intensive margin, i.e., choosing to give birth to an additional child for those who are already parents. In other words, the causal effect of parenthood combines that of the first child and all subsequent children, with weights depending on the difference between the time-period t and the timing of the first child's birth c: short-run effects ($t = c$) relate

[5] Considering estimates where the child penalty estimation strictly follows the method developed by Kleven, Landais, and Søgaard (2019) leads to different results regarding post-birth dynamics, but it does not alter our conclusion that the child penalty is the same for both traditional and egalitarian groups.

[6] In practice, due to limited sample size, we group immigrant parents by decennial generation, i.e., those born in the 1960s, 1970s, and so on.

[7] We further discuss the consequences of comparing individuals within their own generation at 3.1.3.

almost exclusively to the extensive margin of fertility, whereas longer-run effects ($t > c$) will integrate a larger share of the consequences of the intensive margin. This is particularly true in a context where most parents choose to have more than one child, as indicated by Table 1. We discuss this concern further in 3.2.1.

To identify these quantities from the data, we make two key assumptions: (i) a parallel trends assumption and (ii) a limited anticipation assumption.

Assumption 1 (Parallel trends in baseline outcomes) *For all g, b, (t, t'), and (c, c'), where $c, c' > 1$ and $c, c' < \infty$, we assume:*

$$\mathbb{E}[Y_{i,t}(\infty) - Y_{i,t'}(\infty) \mid G_i = g, , B_i = b, , C_i = c] = \mathbb{E}[Y_{i,t}(\infty) - Y_{i,t'}(\infty) \mid G_i = g, , B_i = b, , C_i = c'] \tag{3}$$

Assumption 2 (Limited anticipation) *For all t, g, b, and c, if $t < c - 1$, then:*

$$\mathbb{E}[Y_{i,t}(c) - Y_{i,t}(\infty) \mid G_i = g, B_i = b, C_i = c] = 0 \tag{4}$$

Assumption 1 implies that, in the absence of children, the average labor outcomes for parents of the same gender and birth cohort evolve in parallel over time. Assumption 2 assumes that the effect of children on their parents' labor outcomes is zero up to two years before the child's birth. This differs from a full no-anticipation assumption because (i) fertility decisions typically occur in the year prior to childbirth, and (ii) maternity leave often starts in the final year before childbirth, especially if the birth occurs early in the civil year, thus affecting the mother's labor supply.

Under these assumptions, if there is sufficient variation in the timing of childbirth within each gender and cohort group, cohort-specific average treatment effects on the treated (ATTs) can be identified from the data.

Proposition 1 (Difference-in-differences estimand) *For all (g, c, t), with $1 < c < \infty$, the cohort-specific treatment effect is given by:*

$$\begin{aligned} CATT_{g,b,c,t} \quad = \quad & \mathbb{E}[Y_{i,t} \mid G_i = g, B_i = b, C_i = c] \\ & - \mathbb{E}[Y_{i,c-2} \mid G_i = g, B_i = b, C_i = c] \\ & - \mathbb{E}[Y_{i,t} \mid G_i = g, B_i = b, \max(1, c-2, t+1) < C_i < \infty] \\ & + \mathbb{E}[Y_{i,c-2} \mid G_i = g, B_i = b, \max(1, c-2, t+1) < C_i < \infty] \end{aligned} \tag{5}$$

Additionally:

$$\begin{aligned} \mathbb{E}[Y_{i,t}(\infty) \mid G_i = g, C_i = c] \quad = \quad & \mathbb{E}[Y_{i,c-2} \mid G_i = g, B_i = b, C_i = c] \\ & + \mathbb{E}[Y_{i,t} \mid G_i = g, B_i = b, \max(1, c-2, t+1) < C_i < \infty] \\ & - \mathbb{E}[Y_{i,c-2} \mid G_i = g, B_i = b, \max(1, c-2, t+1) < C_i < \infty] \end{aligned} \tag{6}$$

Proposition 1 suggests that, within each gender $\times$ generation group, if we can observe individuals at least two years before the birth of their first child, it is possible to impute the

counterfactual labor supply profiles of parents whose first child is already born, allowing for the identification of cohort-specific ATTs.[8]

Specifically, let $\underline{T(g,b)},,\underline{T(g,b)}+1,,...,,\overline{T(g,b)}-1,,\overline{T(g,b)}$ denote the time periods that can be observed for individuals in group (g,b). For all c, $CATT(g,b,c,t)$ is identified from the data if:

(i) $\underline{T(g,b)} \leq c-2 \leq \overline{T(g,b)}$;

(ii) $\underline{T(g,b)} \leq t \leq \overline{T(g,b)}$;

(iii) $\mathbb{P}(\max(1,,c-2,,t+1) < C_i < \infty,|,G_i = g,,B_i = b) > 0$.

This final condition implies that very long-run effects are generally unidentifiable under these assumptions, as no counterfactual can be inferred from the data after the last mother is about to give birth to her first child. However, as long as we focus on impacts of parenthood within the first 10 years following the birth of the first child, this is not an issue.

3.1.2 Aggregation and estimation

Aggregation The quantities of interest represent the causal effect of parenthood for a given gender and a specified time period. These effects are recovered by aggregating gender-generation-cohort-specific ATTs, weighted proportionally to population size. To enhance comparability with existing literature and across labor outcomes, we measure these effects relative to the counterfactual level.[9] Specifically, let Ω denote the subset of gender-generation-cohort-time-period combinations for which all three conditions hold, as well as $C_i > 1$, ensuring that $CATT(g,b,c,t)$ is identified from the data. We define:

$$\tau(g,s) = \frac{\mathbb{E}[Y_{i,C_i+s}(C_i) - Y_{i,C_i+s}(\infty) \mid (g,B_i,C_i,C_i+s) \in \Omega]}{\mathbb{E}[Y_{i,C_i+s}(\infty) \mid (g,B_i,C_i,C_i+s) \in \Omega]} \tag{7}$$

This expression captures the relative average treatment effect of being s years past the birth of one's first child, for a specific subset of individuals that varies with s. By Proposition 1, $\tau(g,s)$ can be expressed as a function of quantities fully identified from the data.

Finally, to quantify how parenthood amplifies gender gaps in labor market outcomes, we introduce the child penalty $\xi(S)$. This measure represents the average difference in the relative effects of parenthood between mothers and fathers over a specified duration:

$$\xi(S) = \frac{1}{S}\sum_{s=0}^{S}\{\tau(f,s) - \tau(m,s)\} \tag{8}$$

[8] In practice, when dealing with bounded outcomes, such as binary outcomes like labor market participation, we cap counterfactual labor supply profiles to ensure they stay within admissible ranges. For example, for labor market participation, we write:

$$\begin{aligned}\mathbb{E}[Y_{i,t}(\infty) \mid G_i = g,\, C_i = c] \;=\; & \min(1,\max(0,\mathbb{E}[Y_{i,c-2} \mid G_i = g,\, B_i = b,\, C_i = c] \\ & +\mathbb{E}[Y_{i,t} \mid G_i = g,\, B_i = b,\, \max(1,\, c-2\,,t+1) < C_i < \infty] \\ & -\mathbb{E}[Y_{i,c-2} \mid G_i = g,\, B_i = b,\, \max(1,\, c-2,\, t+1) < C_i < \infty]))\end{aligned}$$

[9] Considering results obtained when the child penalty is measured in absolute value does not alter our conclusions.

Estimation The above discussion motivates a straightforward plug-in estimator, where population probabilities and expectations are replaced by their empirical counterparts. The same approach applies to the estimation of $\xi(S)$.

Under standard integrability assumptions, these estimators are asymptotically normal (Callaway and Sant'Anna, 2021). For inference, we rely on a reweighted bootstrap approach (Shao and Tu, 1995), clustered at the individual level. This method is particularly advantageous as it mitigates the risk of empty comparison groups in bootstrap replications—a non-negligible concern given the narrow groups and finite sample in our analysis. Clustering at the individual level is justified by the study design, as the treatment, i.e., parenthood, is assigned at the individual level (Abadie et al., 2022).

3.1.3 Validation exercises

Credibility of the parallel trends assumption Our entire approach relies on the assumption that, in the absence of children, the average labor outcomes of parents of the same gender and generation who had their first child at different ages would have evolved at the same rate. This is the key assumption underpinning the event-study methodology developed by Kleven, Landais, and Søgaard (2019) to study the child penalty. However, this assumption is subject to debate, as the age at first childbirth is strongly linked to the age of entry into stable employment (Landaud, 2021), which, in turn, is closely correlated with educational attainment. Consequently, comparing parents who had their first child later with those who had their first child earlier effectively compares individuals with differing levels of education, who are likely to experience differences in labor market opportunities not only in levels but also in trends.

One potential solution to this issue is to restrict the comparison groups to parents with similar levels of educational attainment. However, this approach is impractical in our context due to the limited sample size in the survey data; further restricting the groups results in highly imprecise estimates.We address this limitation by utilizing large administrative datasets where an educational proxy is available. We demonstrate that narrowing comparison groups by educational levels in these datasets does not alter our findings regarding the aggregate child penalty -— the average widening of gender gaps in labor market outcomes caused by parenthood over time.[10]

Comparison with administrative registers We validate our survey-based results by comparing them with analogous estimates derived from administrative data. This comparison reveals that for labor outcomes observed in both datasets, the estimated child penalty for immigrant parents is remarkably consistent across data sources.[11]

Couple penalties vs. child penalties As discussed by Kleven, Landais, and Leite-Mariante (2024), if living as a couple is associated with anticipated parenthood, part of the effect of parenthood on parents' labor supply may manifest before the arrival of children and materialize as soon as men and women enter a stable relationship. This raises two important questions:

[10] These results are detailed in a separate note available upon request from the authors.

[11] The same applies to these results.

(i) if children arrive shortly after the start of a stable marital life, our approach might capture the gender-biased impact of stable relationships rather than the actual effect of parenthood on labor supply; and (ii) the anticipated effect of motherhood on women's labor supply might be larger for those with more traditional views, even if the direct impact of motherhood is the same.

Appendix A shows that minor modifications to our model allow us to jointly identify the couple penalty and the child penalty, provided there is sufficient variation in the time lapse between the beginning of a relationship and the arrival of children. It also provides empirical estimates of both penalties based on this approach. The key finding is that our estimate of the child penalty is not substantially different from the one derived using this alternative approach. Finally, the appendix demonstrates that both the couple penalty and the child penalty are similar across attitude groups.

3.1.4 Choice of labor outcome

The survey data underlying this study tracks labor market trajectories along only a few dimensions. Specifically, earnings, wages, and hours worked are not included in this dataset, which focuses solely on labor market participation and employment. While analyzing salaried employment is appealing for comparability with administrative registers, we choose to focus on labor market participation. This choice is motivated by the fact that labor market participation, by accounting for the possibility of involuntary unemployment, aligns more closely with the concept of labor supply than realized employment does, because the former is not an equilibrium outcome as the latter is.

Figure 2 presents our estimates of the average impact of parenthood on labor market participation among first- and second-generation immigrants in France. The results show that the decline in labor market participation due to parenthood is 15 percentage points greater for mothers than for fathers.

Although we focus on labor market participation, our results are not sensitive to this specific choice. Estimates similar to our baseline findings, but based on employment-to-population ratios rather than labor market participation rates, lead to similar conclusions.

3.2 Does the comparison of immigrants with different gender-related attitudes have a causal interpretation?

To support the causal interpretation of our comparison of immigrants with different gender-related attitudes, we consider two distinct approaches. The first, which involves a reweighting procedure based on the inverse propensity score, aims to make immigrants as similar as possible across comparison groups in terms of the observable characteristics most relevant to the child penalty. While the implementation is straightforward, it raises concerns about which characteristics should be included. We specifically discuss whether variables describing fertility decisions should be part of the covariate set.

The second approach ensures that our comparisons are immune to the concern that current self-reported beliefs may themselves result from past life events directly relevant to the child penalty. This would be the case, for instance, if these beliefs are an *ex-post* rationalization of

past fertility and labor supply decisions that contribute to the child penalty. To address this, we build on an approach similar to an instrumental variable strategy. Specifically, we first compare immigrants across the environments in which they were raised. We show that immigrants raised in larger families, where parents' allocation of household chores was more gender-unequal, and who had a more religious upbringing, hold more traditional views regarding gender today. Because, in terms of timing, this environment precedes the fertility and labor supply decisions that give rise to the child penalty, it should also be a causal antecedent of these decisions. Assuming that this correlation was equally strong when the relevant fertility and labor supply decisions were made, comparing immigrants across these dimensions is therefore informative regarding the impact of gender-related attitudes on the child penalty.

Even though this idea closely resembles an instrumental variable strategy, we do not go as far as to move from this reduced-form comparison to a Wald estimand. The reason is that for this Wald estimand to have a straightforward causal interpretation, a restriction exclusion assumption must hold. In the context of this paper, the event-study approach to the child penalty identifies the impact of parenthood on parents. The issue is that these quasi-instruments we consider may lead immigrants to change their fertility decisions in addition to altering their gender-related attitudes. For instance, let us assume that a more traditional upbringing makes immigrants more prone to having children, as seems plausible. In this case, the set of parents raised in more traditional families includes: (i) parents who would have become parents even if they had been raised in less traditional families, for whom the child penalty for parents raised in less traditional families serves as a good counterfactual, and (ii) parents who would not have become parents had they been raised in less traditional families, for whom the counterfactual child penalty cannot be identified from the data on immigrants raised in less traditional families. For this reason, we stick to a reduced-form estimate, acknowledging that selection into parenthood may still be at play.

In addition to this comparison across upbringing environments, we develop a related framework that compares second-generation immigrant parents across their parents' countries of birth, based on the observation that (i) first-generation immigrants with similar migration backgrounds tend to hold the same views; and (ii) second-generation immigrants' attitudes correlate with those of first-generation immigrants from the same country as their parents. Section 5 details this framework and performs this comparison using both survey and administrative data.

3.2.1 Comparing immigrants with similar observable characteristics

Interpreting the comparison of child penalties across groups defined by their adherence to more or less traditional views regarding gender and bodily autonomy may prove tricky if these groups differ in characteristics that influence both their views and labor supply decisions. Indeed, as Table 1 suggests, immigrants with more traditional gender attitudes tend to be less educated, more likely to be first-generation immigrants, and usually have more children. Addressing this issue requires techniques aimed at making the groups more similar in terms of observable characteristics.

In the context of this paper, it is helpful to distinguish between two sets of observable characteristics. The first set includes characteristics that are causal antecedents to both attitudes

and labor supply decisions. An example here is first-generation status, as country of birth is determined prior to the formation of attitudes. Clearly, we should compare immigrants who are similar with respect to such observable characteristics (Pearl, 2009).

The second set consists of characteristics that are likely to be consequences, rather than causes, of attitudes and that are likely to influence labor supply decisions. A relevant example here is fertility decisions. Whether we should make immigrants similar with respect to these characteristics depends on the question at hand. Essentially, this is a matter of mediation analysis, rather than an issue of whether our causal estimates are biased.

Dealing with differences in fertility The question, therefore, is whether we should compare immigrants with similar fertility decisions. Ultimately, we choose to make attitude groups similar in terms of their fertility decisions. The reason for this choice is that, while we aim to assess the consequences of gender-related attitudes, of which fertility decisions certainly form a part, our outcome of interest is not an observable outcome *per se*, but the consequences of fertility decisions on labor supply. Particularly relevant in this context is the fact that, while our approach is based on the timing of *ego*'s first child's birth, the effect of parenthood incorporates not only the consequences of this first child on *ego*'s labor force participation but also those of all her potential subsequent children. As such, it would be problematic to compare the effects of fertility across groups with differing fertility outcomes.

While conditioning on fertility is our baseline approach, we also consider estimates based on two different approaches: (i) comparisons based on raw data, and (ii) comparisons based on data in which we condition on the first set of observable characteristics, such as generation, migration status, or educational attainment, but not on the second set, which includes fertility decisions.

Supply or demand effects? Given that gender-related attitudes are strongly correlated with migration status and countries of origin, a natural question that may arise is whether our framework captures labor supply or labor demand effects. Indeed, the demand for immigrant labor may depend on their migration background: for instance, if discrimination against immigrants from certain countries is pervasive in the labor market.

Our answer to this question is threefold. Firstly, our outcome of interest is labor force participation, as opposed to actual employment: it includes the possibility of involuntary unemployment that may result from employers discriminating against a particular group of immigrants at the hiring stage (see 3.1.4). Secondly, what we are comparing across attitude groups are not labor force participation rates *per se*, but triple differences in labor force participation rates (across gender, over time, and between parents who have their first child earlier or later in their lifecycle). As a result, for discrimination or, more generally, labor demand to drive our results, it would require employers to discriminate against actual mothers of a particular migration background, as opposed to immigrants from a particular set of countries. Thirdly, we include in our covariate set a synthetic variable built from the TeO2 data that corresponds to (i) self-reported experiences with discrimination and unequal treatment in the labor market, either at the hiring stage or at the workplace, due to either their origins or the color of their skin; and (ii) whether survey respondents believed that they belong to a group subject to racial

discrimination in France.[12] As a result, after this reweighting, both attitude groups should be considered as having the same experience with labor market discrimination.

In practice, our approach to the conditioning that allows us to make immigrants more similar across attitude groups is based on an inverse propensity score reweighting technique.

3.2.2 Leveraging the environment in which immigrants grew up

Even when comparing immigrants who are very similar in terms of their observable characteristics, a legitimate concern could be that current gender-related attitudes stem from an *ex-post* rationalization of their past decisions which contribute to the child penalty. To address this issue, we propose moving from current gender-related attitudes to variables assigned prior to these attitudes, in an approach similar to an instrumental variable strategy.

To do so, we begin by considering variables that describe the family environment in which immigrants were raised. Specifically, we examine the prominence of religion during immigrants' childhood, the gender imbalance in household chores between their parents, and the number of children born to their mothers. The premise is that these factors shape attitudes in adulthood, and that we should attribute potential differences in child penalties across immigrants raised in different environments along these dimensions to their gender-related attitudes in adulthood.

We begin by regressing binary gender-related attitudes on these variables using a linear probability model estimated via ordinary least squares. We then split our sample of immigrant parents into two groups, based on whether individuals fall above or below the gender-specific estimated population median in predicted attitudes from this regression. The difference in the share of immigrants whose attitudes we have characterized as more or less traditional between these two environmental groups exceeds 30 percentage points, regardless of gender.[13] This suggests that upbringing environments have a strong effect on later attitudes regarding gender and bodily autonomy.

The final step of our approach is to compare child penalties across these upbringing environment groups, after reweighting the data to make these groups more similar in terms of the relevant observable characteristics (see 3.2.1). Inference for these results is based on a reweighted bootstrap, similar to the method discussed in Subsection 3.1. Subsection 4.2 presents our results.

4 Results

4.1 Comparison across self-reported gender-related attitudes

4.1.1 Child penalty

We now present our baseline results, comparing child penalties across groups of immigrant parents based on their self-reported gender-related attitudes. Figure 3 displays the estimates, obtained after reweighting the data to ensure similarity across groups in terms of pre-childbirth observable characteristics (e.g., migration status, generation, educational attainment, marital

[12] In practice, experience with discrimination only deals with experiences within 5 years before the survey.

[13] After reweighting the data to make the groups comparable in terms of relevant observable characteristics.

history, experience with discrimination), as well as fertility and post-childbirth decisions (e.g., age at first childbirth, total number of children).

The differences in child penalties across attitude groups are minimal. In fact, the penalty appears slightly larger for parents with less traditional attitudes, though our estimates lack the precision to draw firm conclusions.[14] Ultimately, the key takeaway is that the child penalty in labor market participation remains consistent regardless of whether immigrant parents hold traditional or progressive views on gender and bodily autonomy.

This result withstands several methodological concerns. First, it is not driven by our use of a composite index for gender-related attitudes; similar results are obtained when comparing groups based on responses to a single survey item about gender inequality in the labor market. Second, the result holds irrespective of the inclusion or exclusion of specific survey items in defining gender-related attitudes. Third, it remains robust to variations in the threshold used to split groupsor in the child penalty specification. Fourth, the finding applies equally to the child penalty in employment, suggesting that: (i) the choice of labor market participation as the baseline outcome does not drive the result, and (ii) for immigrant mothers, motherhood generally does not lead to unemployment, as the magnitude of child penalties is similar across these outcomes. Fifth, it also persists when the child penalty is measured in absolute terms (percentage points) instead of relative to the counterfactual labor market participation rate (percentage). All these checks are detailed in a separate note available upon request from the authors.

Finally, the result is also robust to accounting for the effects of couple life, as shown in Appendix A. This is particularly relevant since selection into parenthood is closely tied to selection into stable relationships, which could have otherwise biased our findings.

4.1.2 Labor market participation

We further explore this topic and show that, among women, both realized and counterfactual labor market participation rates differ significantly across attitude groups. In contrast, these differences are negligible among men. Consequently, the absence of a correlation between child penalties and gender-related attitudes does not contradict the observation that gender disparities in labor market participation are influenced by these attitudes.

To better understand these implications, Table 2 breaks down the contribution of different groups of mothers – categorized by how their child-related decisions change based on their gender-related attitudes – to variations in the child penalty. This analysis relies on a causal interpretation of differences between attitude groups and is based on two simplifying assumptions: (i) that having children always reduces female labor force participation, and (ii) that holding more traditional gender views consistently lowers female labor supply.

[14] The standard error of the child penalty estimate for each attitude group is approximately 0.03, allowing us to detect differences greater than 8 percentage points. To put this into perspective, in the child penalty data published by Kleven et al. (2024), this magnitude corresponds roughly to the difference in the employment child penalty between France and countries such as Italy or the UK. Conversely, drawing firm conclusions on the typical differences observed in our estimates – about 3 to 4 percentage points – would require standard errors that are three times smaller. In practical terms, this would necessitate a 10-fold increase in sample size.

Under these assumptions, only two groups of mothers contribute to attitude-related differences in the child penalty. The first group consists of mothers who, under nontraditional views, would always work regardless of having children but, under traditional views, would leave the labor market because of children; they increase the child penalty. The second group includes mothers who, under nontraditional views, leave the labor market because of children but, under traditional views, would not participate in the labor market even without children; they reduce the child penalty. Furthermore, the first group creates a gap in realized labor supply across attitudes (but not in counterfactual labor supply), while the second group creates a gap in counterfactual labor supply (but not in realized labor supply). Ultimately, our findings suggest that these two groups are approximately equal in size.

We conclude this analysis by comparing the contribution of the child penalty to the average gender gaps in labor force participation across attitude groups. To do so, we consider all individual-age observations in our sample of immigrant parents. We impute a counterfactual labor force participation value by adding the corresponding estimated average impact of parenthood to the observed value for individuals with children. For each age and attitude group, we compute both the realized and counterfactual gender gaps in labor force participation, and then average these values over all ages between 20 and 50. The counterfactual value represents the residual gap, not explained by the child penalty, while the difference between the realized and counterfactual values quantifies the contribution of the child penalty to the overall labor force participation gap.

Figure 4 presents our estimates. The realized gender gap in labor force participation for more traditional immigrants is 20.5 percentage points, more than twice as large as for their more egalitarian counterparts (9.5 p.p.). While the child penalty is very similar across groups, more traditional mothers tend to have their first child at a younger age (see Table 1), which slightly increases the contribution of the child penalty to the gender gap for immigrants with more traditional views: 9.6 p.p. compared to 8.1 p.p. for their less traditional counterparts. Put differently, the child penalty accounts for 47% of the gender gap in labor force participation in the more traditional group, compared to 86% in the less traditional one. Lastly, the residual gap is sizable for more traditional immigrants (10.9 p.p.) but much smaller for the less traditional group (1.4 p.p.). Ultimately, the wider gender gap in the more traditional group (20.5 - 9.5 = 11 p.p.) is more attributable to differences unrelated to the child penalty (10.9 - 1.4 = 9.5 p.p.) than to differences in the child penalty's contribution (9.6 - 8.1 = 1.5 p.p.).

This finding aligns closely with the cross-country comparisons presented by Kleven, Landais, and Leite-Mariante (2024). Their research shows that the proportion of the gender gap attributable to child penalties is higher – often exceeding 100% – in countries such as Denmark, Sweden, the Netherlands, and Australia. They support this conclusion with historical data, framing it within a discussion of the child penalty's relevance at various stages of economic development and structural transformation. In contrast, the immigrants in our study all reside in France, where economic development is consistent across the population. Our evidence therefore suggests that the importance of the child penalty in explaining gender gaps in labor outcomes depends on gender-related attitudes. Put differently, the data indicate that the child penalty becomes a dominant driver of gender disparities in the labor market only when gender

attitudes shift toward more positive views on female employment and women's autonomy.

4.2 Comparison across environments in which immigrants grew up

Since current self-reported attitudes may be influenced by past fertility and labor supply decisions, the above comparison could provide a biased assessment of the role of gender-related attitudes in the child penalty. To address this issue, we compare child penalties across groups of immigrant parents, defined by the environment in which they were raised. This approach relies on two key arguments: (i) these environments are not outcomes of fertility and labor supply decisions made later in life, and (ii) growing up in a more traditional family -— characterized by higher religiosity during childhood, unequal distribution of household chores among *ego*'s parents, or a larger number of siblings -— increases the likelihood of holding more traditional gender views by over 30 percentage points (see 3.2.2).

Figure 5 presents our findings. The key takeaway is that the child penalty in labor market participation is remarkably similar across different upbringing environments. In other words, having children is no more detrimental to the labor market participation of mothers raised in traditional families than to those raised in less traditional families. Since the causal pathway from upbringing environments to the child penalty operates through gender-related attitudes held in adulthood, this finding suggests that such attitudes do not account for the child penalty, reinforcing the causal interpretation of the previous comparison.

This result is not driven by our decision to reweight the data to make the comparison groups comparable in terms of pre-childbirth observable characteristics and fertility decisions. It is also robust to using more restrictive measurements of the upbringing environment, or defining our attitudes based solely on views regarding gender inequality in the labor market. All these checks are detailed in a separate note available upon request from the authors.

In addition to these robustness checks, this result remains unchanged when the child penalty is measured in absolute terms rather than relative to the counterfactual labor market participation rate. Furthermore, for women, both realized and counterfactual participation rates exhibit significant differences across upbringing environments, while no such differences are observed for men. Since comparisons across upbringing environments are less likely to be affected by reverse causality bias than those based on current attitudes (as reported in 2019-2020), this reinforces the conclusion that the child penalty is likely not causally related to gender-related attitudes, while still supporting the idea that these attitudes are a key driver of gender gaps in labor market participation.

5 Additional evidence from second-generation immigrants

In this section, we use administrative data on second-generation immigrant parents to compare the child penalty across groups defined by the prevalence of traditional gender-related attitudes among first-generation immigrant parents from the same country. This approach is similar to the epidemiological method used to study the impact of culture on economic outcomes (Fernández, 2011), and closely follows that of Alesina, Giuliano, and Nunn (2013). The rationale for this comparison is similar to that behind the comparison across upbringing environments: namely,

that countries of origin are assigned before gender-related attitudes are formed. Restricting the comparison to second-generation immigrants has the additional advantage of ensuring that all individuals are born and raised in France, thus growing up within similar institutions.

This comparison has the advantage of requiring only knowledge of immigrants' countries of origin and aggregate information on gender-related attitudes by country, without needing additional individual data on second-generation immigrants, whose fertility and labor supply decisions we track. This enables us to conduct these comparisons using administrative registers rather than survey data. Moving from survey data to administrative registers offers two main advantages: (i) the sample size is much larger, improving the precision of the comparison; and (ii) the data on labor market outcomes are more comprehensive, allowing us to examine the impact of gender-related attitudes on dimensions beyond labor market participation, and making our results more comparable to the standard child penalty literature.

5.1 EDP data: general overview

The Permanent Demographic Sample (*Echantillon Démographique Permanent*, EDP) collects administrative data from birth registers, tax returns, firm records, and welfare benefits recipients' files, as well as information from comprehensive censuses and census surveys. These data pertain to a representative sample of the French population, including all individuals born on one of the first four days of April, July, and October, or between January 2nd and 5th. The EDP tracks the professional, financial, residential, and family situations of these "EDP individuals" annually.

In the context of this paper, the EDP data have three key advantages:

- Combining information from birth certificates with census data allows us to identify both first- and second-generation immigrants, including their country of birth and, for second-generation immigrants, their parents' country of birth;
- Information from birth certificates regarding EDP individuals' children allows us to track their fertility decisions;
- Lastly, payroll tax data enable us to track entire labor market histories.

Taken together, these three features make it possible to estimate the child penalty by country of origin for second-generation immigrants, helping us assess the causal interpretation of our results regarding the link between gender-related attitudes and the child penalty.[15][16]

[15] This dataset also allows us to estimate the child penalty for both first- and second-generation immigrants, which helps validate the quality of the survey data by comparing them with a dataset that has already provided published estimates of the child penalty in France.

[16] Although the data are of high quality, using them requires combining information from multiple sources – census data, birth certificates, tax returns, and payroll tax forms. Therefore, decisions must be made regarding how to handle these various pieces of information, which may raise some issues in terms of methodology. A separate note, available upon request from the authors, details how we address these concerns, particularly regarding the identification of immigrants and migration outflows, the measurement of fertility decisions, and the definition of relevant labor market outcomes.

5.2 Method

Our approach here is based on ideas similar to those outlined in Subsection 3.2, except that instead of considering upbringing environments measured at the individual level, it relies on a more aggregated variation. The intuition is that, while it is possible that one's self-reported beliefs may depend on past decisions, this should not be the case when it comes to the beliefs of their peers. To take advantage of this fact, we first document that first-generation immigrants' friends often come from the same country as they do. We then show that the beliefs of first-generation immigrants are strongly correlated with those of other first-generation immigrants born in the same country. Additionally, we provide evidence that second-generation immigrants whose parents were born in countries where first-generation immigrants are more likely to report traditional gender-related attitudes are also more likely to hold traditional views on gender. This, in turn, motivates the comparison of second-generation immigrants in terms of the child penalty, based on how frequently traditional gender beliefs are held among first-generation immigrants from their parents' countries of origin, which closely mirrors Alesina, Giuliano, and Nunn (2013)'s approach to identifying the effect of cultural norms on economic outcomes.

Lateral transmission We begin by considering survey items related to friendship networks. The proportion of immigrant parents who report having met with at least one friend with the same migration background as themselves in the last two weeks is above 50% for first-generation immigrant mothers and over 60% for first-generation immigrant fathers. This suggests that, in most cases, first-generation immigrants tend to stay close to their peers, which creates an opportunity for the lateral transmission of attitudes.

We then turn to the attitudes themselves. To do so, we compare first-generation immigrant parents' attitudes to those of their peers using a leave-one-out approach. Specifically, we regress *ego*'s gender-related attitudes on the average attitudes of first-generation immigrant parents born in the same country as her, excluding *ego* from the computation of these average attitudes. This regression is based on a linear probability model estimated using ordinary least squares. Table 3 presents our results. These results show that first-generation immigrant parents' attitudes regarding gender are indeed positively correlated within groups defined by their country of birth, which we interpret as evidence of lateral transmission.

Vertical transmission We then turn to second-generation immigrants. Specifically, we regress second-generation immigrant parents' attitudes regarding gender on the average attitudes of first-generation immigrant parents born in the same country as their parents[17], using a linear probability model estimated with ordinary least squares. Table 4 presents our estimates. These results show that second-generation immigrant parents' attitudes closely track those of their parents' peers.

We use the estimated coefficients to predict gender-related attitudes for second-generation immigrant parents in the administrative registers, where the attitudes of first-generation parents' peers were imputed based on the survey results.

[17] If one of the parents is born in France, we use the average attitudes of French parents of the same gender who are neither first- nor second-generation immigrants.

Consistent with the rationale, we find that second-generation immigrant parents whose parents were born in countries from which immigrants in France typically hold more traditional views on gender also tend to hold more traditional views on gender. The difference in actual traditional views on gender between those above the median predicted probability of holding traditional views is about 15 percentage points for fathers and up to 20 percentage points for mothers.

The final step of our approach is to compare child penalties across these origin groups, after reweighting the data to make the groups more similar in terms of relevant observable characteristics (see 3.2.1). Inference regarding these results is based on a reweighted bootstrap procedure similar to the one discussed in Subsection 3.1, with the main difference being that we now need to generate replicates for both the survey data (to account for the uncertainty in the estimated coefficients displayed in 4, upon which the grouping is based) and the administrative registers (to account for the uncertainty in the estimation of the child penalty)

5.3 Results

Figure 6 displays the child penalties in labor earnings for each group defined by the attitudes of second-generation immigrant parents' parents' peers. We find that the aggregate child penalty—the difference in the impact of parenthood on mothers and fathers, averaged over time—is not larger for second-generation immigrants who, based on their parents' countries of birth, are predicted to be more likely to hold traditional views. Specifically, this quantity remains around 30 percentage points, regardless of the group considered. This finding supports our interpretation that the child penalty is not easily explained by traditional gender-related attitudes.

Additional results that separately consider transmission through fathers and mothersdo not alter our conclusion. We also demonstrate that this finding is not dependent on our decision to consider the child penalty in relative rather than absolute terms, the reweighting of the data to make origin groups more comparable, or the restriction of the comparison set to parents who entered the labor market at the same age. All these checks are detailed in a separate note available upon request from the authors.

6 Concluding remarks

This study leverages a rich dataset on first- and second-generation immigrants in France to examine whether traditional gender-related attitudes are a primary driver of the child penalty in labor outcomes. Across three empirical tests, we find that the child penalty remains remarkably consistent across groups, regardless of variations in self-reported beliefs, childhood environments, or country of origin. This uniformity challenges the idea that traditional gender attitudes, as captured in our data, directly explain differences in the child penalty.

Our results confirm that the child penalty is substantial and borne almost entirely by mothers, leading to a 15% reduction in their labor force participation compared to a counterfactual scenario without children, while fathers' labor market outcomes remain unaffected. Notably, we observe significant differences in female labor force participation across groups with varying

gender attitudes, but these differences are present both with and without children. This suggests that the child penalty itself operates independently of the attitudinal dimensions captured by our measures, even if cultural norms influence overall participation levels.

While our findings may seem at odds with previous research that links gender attitudes to labor market outcomes, they underline a critical distinction: gender-related attitudes may shape the baseline levels of participation or other dimensions of inequality, but they do not appear to influence the additional labor market disadvantage triggered specifically by motherhood. This raises important questions about the cultural and institutional determinants of the child penalty.

Our work is motivated by earlier literature showing that family policies often fail at mitigating the child penalty (Rabaté and Rellstab, 2022; Kleven et al., 2024). These findings have prompted a shift in focus toward cultural determinants of gender inequality, which this study aims to explore. However, our results do not directly assess the role of public policies, nor do they imply that such policies are ineffective. Instead, they suggest that even as cultural and institutional shifts reduce overall gender gaps, the child penalty may persist as a structural barrier disproportionately affecting mothers.

Future research should aim to refine the measurement of cultural norms and attitudes to better capture the subtle and multidimensional nature of their influence on the child penalty. Beliefs about motherhood, family responsibilities, and social expectations may play a significant role but remain difficult to quantify with current tools. Cross-national comparisons and longitudinal approaches could also help identify how institutional and cultural contexts interact to shape labor market outcomes for mothers.

These findings suggest a paradoxical outcome: as societies move toward more egalitarian attitudes and reduce overt gender discrimination, the child penalty may become the dominant driver of labor force gender gaps, as appears to be the case in many developed countries (Kleven, Landais, and Leite-Mariante, 2024). Understanding and addressing the specific norms, practices, and structural barriers that sustain the child penalty will be critical for achieving gender equality in labor markets.

References

Abadie, A., S. Athey, G.W. Imbens, and J.M. Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138:1–35.

Akerlof, G.A., and R.E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* 115:715–753.

Alesina, A., P. Giuliano, and N. Nunn. 2013. " On the Origins of Gender Roles: Women and the Plough." *The Quarterly Journal of Economics* 128:469–530.

Andresen, M.E., and E. Nix. 2022. "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples." *Journal of Labor Economics* 40:971–1004.

Angelov, N., P. Johansson, and E. Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34:545–579.

Bertrand, M., E. Kamenica, and J. Pan. 2015. "Gender Identity and Relative Income Within Household." *The Quarterly Journal of Economics* 130:571–614.

Boelmann, B., A. Raute, and U. Schönberg. 2020. "Wind of Change? Cultural Determinants of Maternal Labor Supply." RF Berlin - CReAM Discussion Paper Series No. 2020, Rockwool Foundation Berlin (RF Berlin) - Centre for Research and Analysis of Migration (CReAM), Oct.

Borrell-Porta, M., V. Contreras, and J. Costa-Font. 2023. "Is employment during motherhood a 'value changing experience'?" *Advances in Life Course Research* 56:100528.

Borrell-Porta, M., J. Costa-Font, and J. Philipp. 2018. "The 'mighty girl' effect: does parenting daughters alter attitudes towards gender norms?" *Oxford Economic Papers* 71:25–46.

Callaway, B., and P.H. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225:200–230, Themed Issue: Treatment Effect 1.

Casarico, A., and S. Lattanzio. 2023. "Behind the Child Penalty: Understanding What Contributes to the Labour Market Costs of Motherhood." *Journal of Population Economics* 36:1489–1511.

Cavapozzi, D., M. Francesconi, and C. Nicoletti. 2021. "The impact of gender role norms on mothers' labor supply." *Journal of Economic Behavior & Organization* 186:113–134.

Cortés, P., and J. Pan. 2023. "Children and the Remaining Gender Gaps in the Labor Market." *Journal of Economic Literature* 61:1359–1409.

de Chaisemartin, C., and X. D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110:2964–96.

Fernández, R., and A. Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American economic journal: Macroeconomics* 1:146–177.

Fernández, R. 2011. "Chapter 11 - Does Culture Matter?" In J. Benhabib, A. Bisin, and M. O. Jackson, eds. *Handbook of Social Economics*. North-Holland, vol. 1, pp. 481–510.

—. 2007. "Women, Work, and Culture." *Journal of the European Economic Association* 5:305–332.

Fortin, N.M. 2005. "Gender Role Attitudes and the Labour-Market Outcomes of Women across OECD countries." *Oxford Review of Economic Policy* 21:416–438.

—. 2015. "Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery." *Annals of Economics and Statistics*, pp. 379–401.

Gay, V. 2023. "The Intergenerational Transmission of World War I on Female Labour." *The Economic Journal* 133:2303–2333.

Goldin, C. 2021. *Career and Family: Women's Century-Long Journey toward Equity*. Princeton University Press.

—. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104:1091–1119.

—. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review* 96:1–21.

Goldin, C., and L.F. Katz. 2002. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy* 110:730–770.

Goodman-Bacon, A. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225:254–277, Themed Issue: Treatment Effect 1.

Ichino, A., M. Olsson, B. Petrongolo, and P.S. Thoursie. 2024. "Taxes, Childcare and Gender Identity Norms." Unpublished, forthcoming, *Journal of Labor Economics*.

Juhn, C., and K. McCue. 2017. "Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts." *Journal of Economic Perspectives* 31(1):183–204.

Kleven, H. 2022. "The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach." Working Paper No. 30176, National Bureau of Economic Research, June.

Kleven, H., C. Landais, and G. Leite-Mariante. 2024. "The Child Penalty Atlas." Unpublished, forthcoming *The Review of Economics Studies*.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. 2024. "Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation." *American Economic Journal: Economic Policy* 16:110–149.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimuüller. 2019. "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109:122–26.

Kleven, H., C. Landais, and J.E. Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11:181–209.

—. 2021. "Does Biology Drive Child Penalties? Evidence From Biological and Adoptive Families." *American Economic Review: Insights* 3:183–198.

Kuziemko, I., J. Pan, J. Shen, and E. Washington. 2018. "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" Working Paper No. 24740, National Bureau of Economic Research, June.

Landaud, F. 2021. "From Employment to Engagement? Stable Jobs, Temporary Jobs, and Cohabiting Relationships." *Labour Economics* 73:102077.

Lazear, E.P., and S. Rosen. 1990. "Male-Female Wage Differentials in Job Ladders." *Journal of Labor Economics* 8:S106–S123.

Lebedinski, L., C. Perugini, and M. Vladisavljević. 2023. "Child Penalty in Russia: Evidence From an Event Study." *Review of Economics of the Household* 21:173–215.

Meurs, D., and P. Pora. 2019. "Gender Equality on the Labour Market in France: A Slow Convergence Hampered by Motherhood." *Economie et Statistique / Economics and Statistics* 510-511-512:109–130.

Pearl, J. 2009. *Causality*, 2nd ed. Cambridge University Press.

Pora, P. 2020. "Keep Working and Spend Less? Collective Childcare and Parental Earnings in France." Working paper No. 2020-29, EconomiX.

Périvier, H., and G. Verdugo. 2024. "Where Are the Fathers? Effects of Earmarking Parental Leave for Fathers in France." *ILR Review* 77:88–118.

Quinto, A., L. Hospido, and C. Sanz. 2021. "The Child Penalty: Evidence From Spain." *SERIEs: Journal of the Spanish Economic Association* 12:585–606.

Rabaté, S., and S. Rellstab. 2022. "What Determines the Child Penalty in the Netherlands? The Role of Policy and Norms." *De Economist* 170:195–229.

Seguino, S. 2011. "Help or Hindrance? Religion's Impact on Gender Inequality in Attitudes and Outcomes." *World Development* 39:1308–1321.

Sevincer, A.T., C. Galinsky, L. Martensen, and G. Oettingen. 2023. "Political Ideology Outdoes Personal Experience in Predicting Support for Gender Equality." *Political Psychology* 44:829–855.

Shao, J., and D. Tu. 1995. "Bayesian Bootstrap and Random Weighting." In *The Jackknife and Bootstrap*. New York, NY: Springer New York, pp. 416–446.

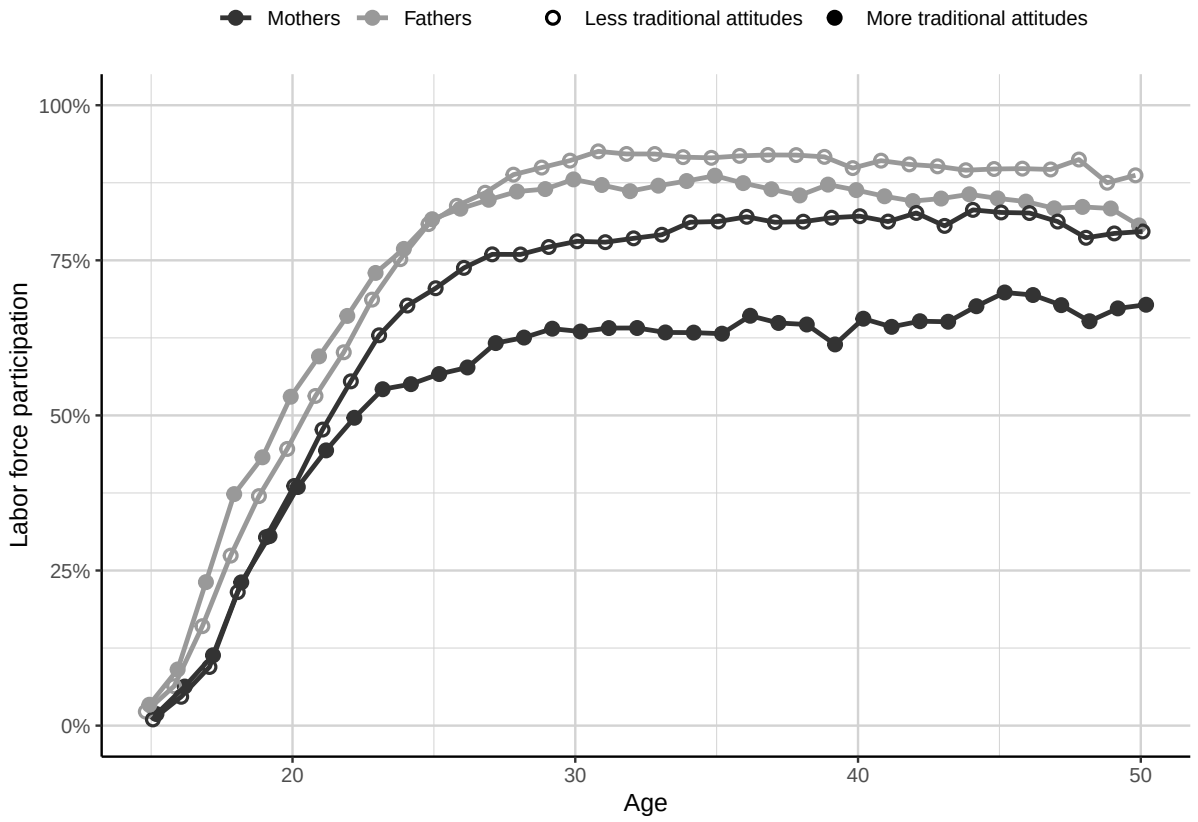
Sieppi, A., and J. Pehkonen. 2019. "Parenthood and gender inequality: Population-based evidence on the child penalty in Finland." *Economics Letters* 182:5–9.

Sun, L., and S. Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225:175–199, Themed Issue: Treatment Effect 1.

Thao Khamsing, W., O. Guin, T. Merly-Alpa, and N. Paliod. 2022. "Enquête Trajectoires et Origines 2. De la conception à la réalisation." Working paper No. 2022/002, Insee.

Figures

Figure 1. Lifecycle profiles of labor force participation: by gender and gender-related attitudes

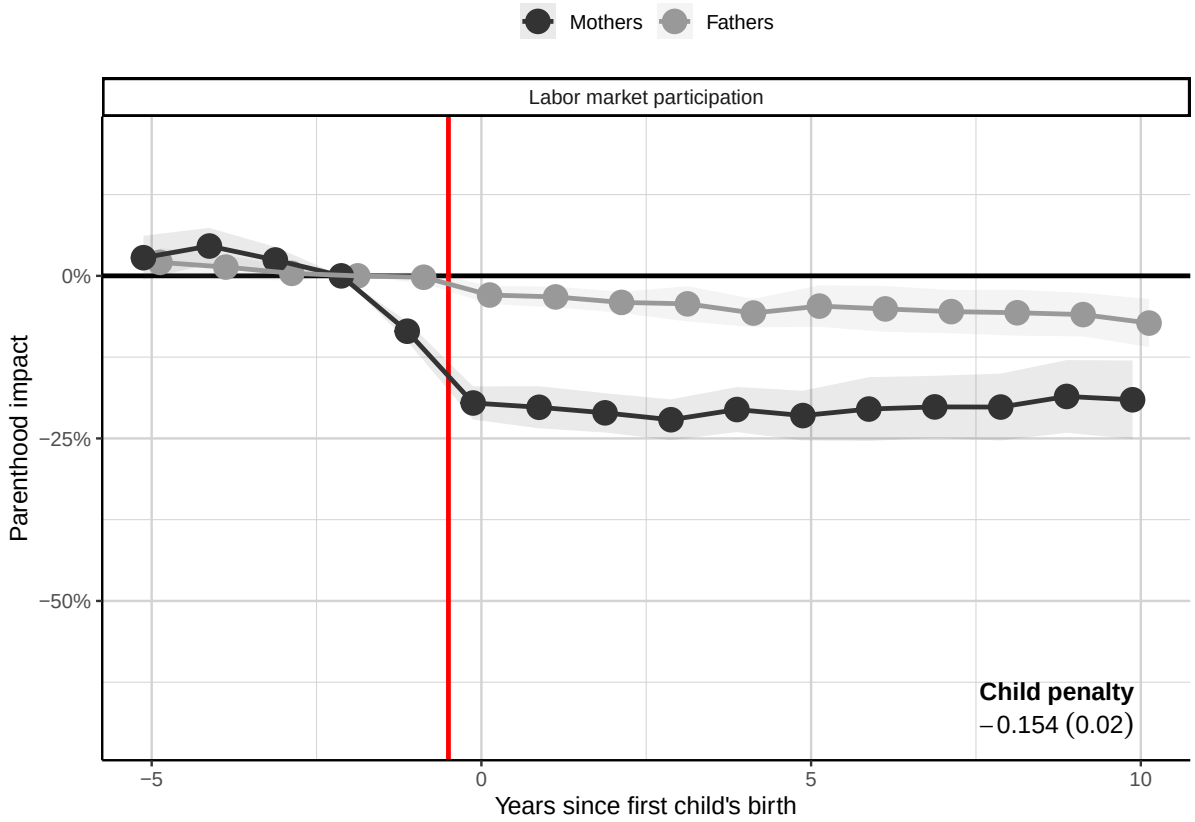


Labor force participation rate for immigrant parents, by age, gender and gender-related attitudes. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure 2. Child penalty in labor market participation for immigrant parents

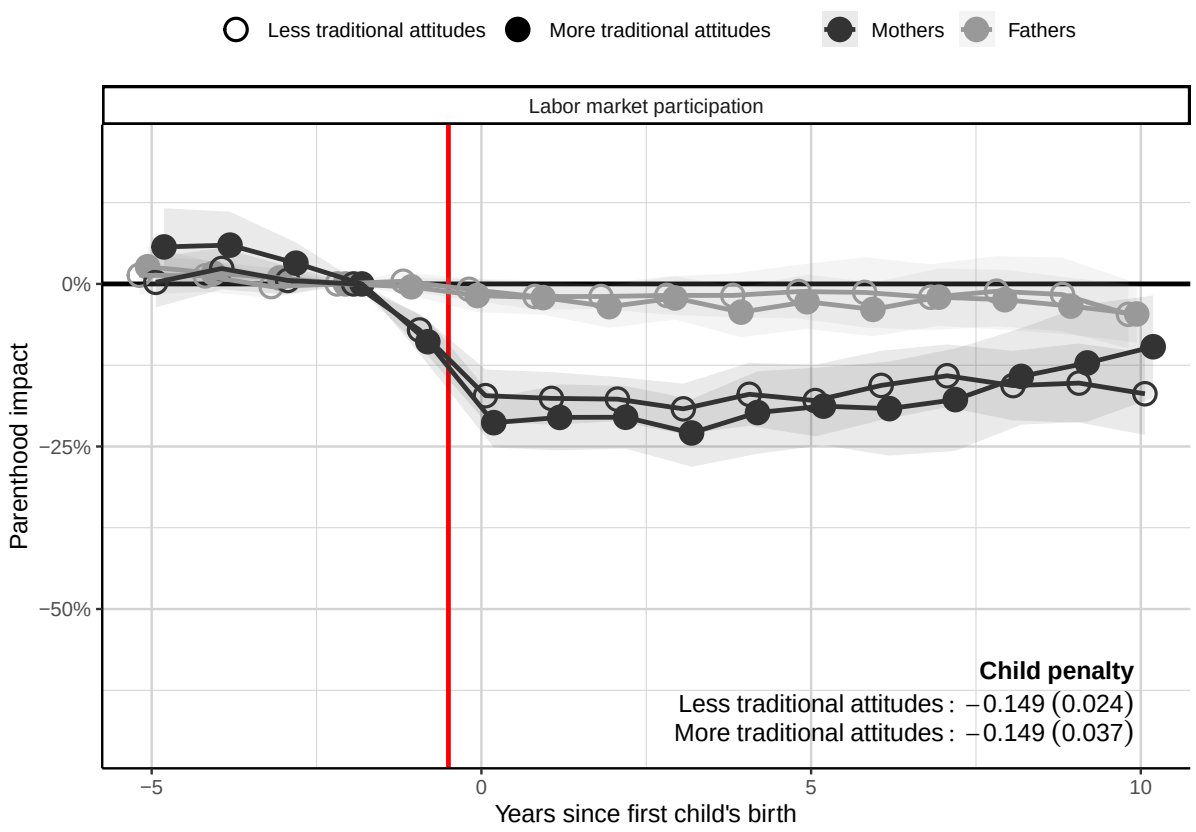


Parenthood impact is identified from comparisons between parents of the same decennial cohort who had their first child at different ages, and displayed relative to the counterfactual labor market participation rate. The child penalty is the difference in parenthood impact between men and women, averaged over the first 10 years after their first child is born. Shaded areas correspond to 95% confidence intervals; they are based on a reweighted bootstrap approach, clustered at the individual level. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure 3. Child penalty in labor market participation for immigrant parents: by self-reported attitudes, balanced on pre-birth characteristics and fertility decisions

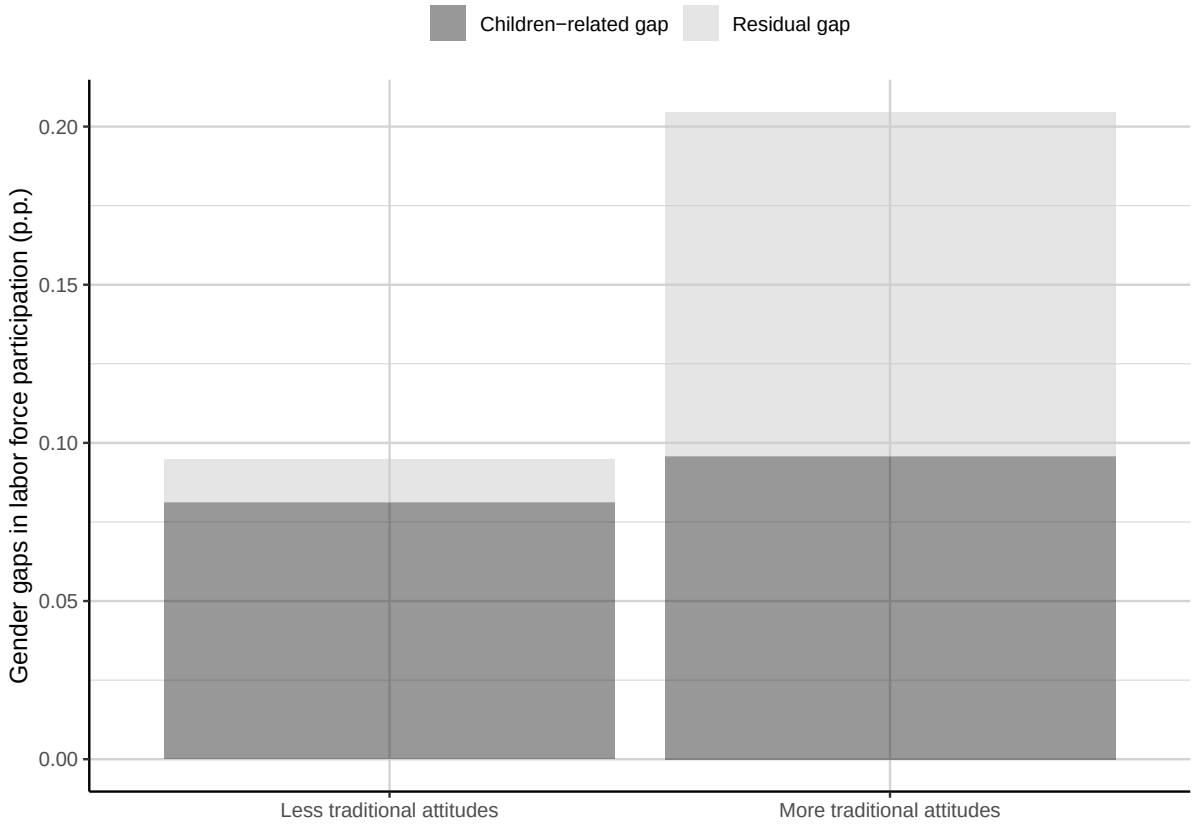


Parenthood impact is identified from comparisons between parents of the same decennial cohort who had their first child at different ages, and displayed relative to the counterfactual labor market participation rate. The child penalty is the difference in parenthood impact between men and women, averaged over the first 10 years after their first child is born. The data are reweighted based on an inverse propensity score approach so as to make attitudes groups similar in terms of pre-birth observables characteristics and fertility decisions. Shaded areas correspond to 95% confidence intervals; they are based on a reweighted bootstrap approach, clustered at the individual level. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure 4. Gender gaps in labor force participation: contribution of the child penalty, by gender-related attitudes

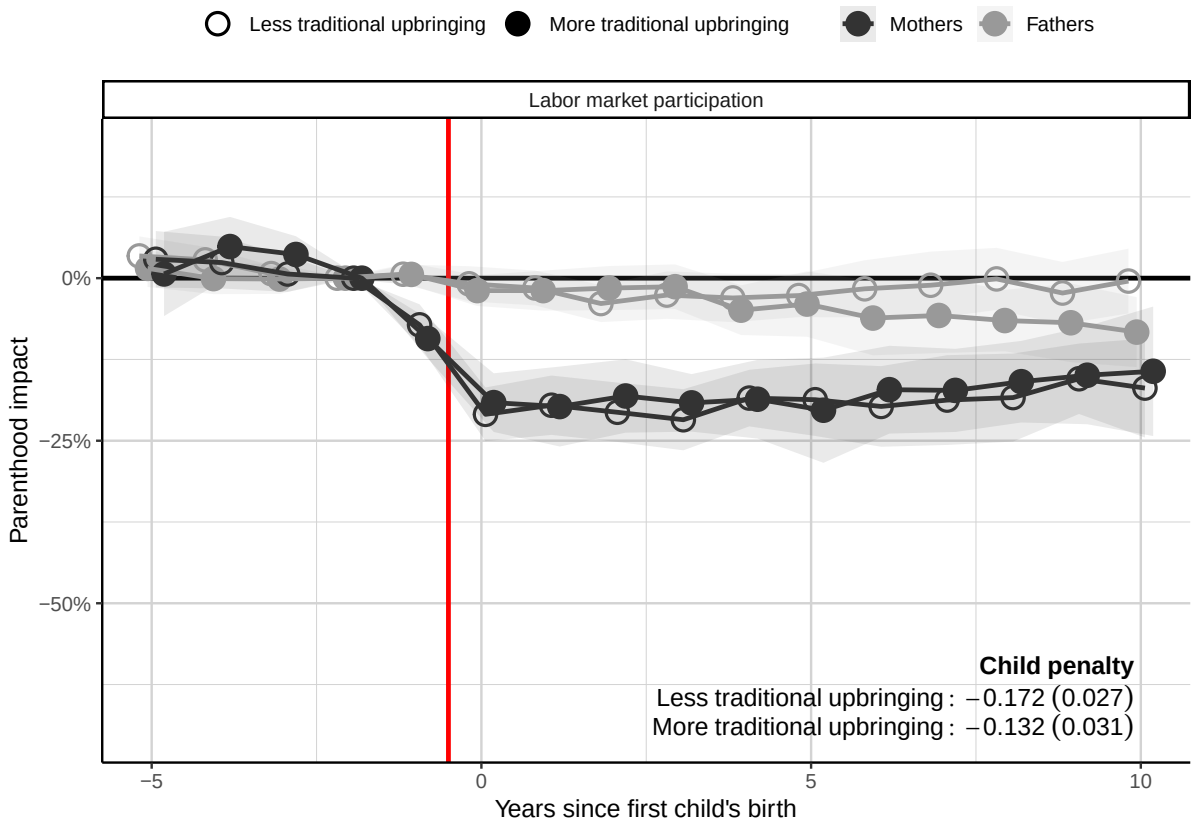


Average labor force participation gaps from age 20 to age 50 for immigrant parents, by gender-related attitudes. This figure displays the contribution of the child penalty to the gender gap, that is the magnitude of the child penalty in labor force participation multiplied by the prevalence of parenthood, and the residual gap, that is the difference between the realized rates and this contribution. Parenthood impact is identified from comparisons between parents of the same decennial cohort who had their first child at different ages. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure 5. Child penalty in labor market participation for immigrant parents: by upbringing environments, balanced on pre-birth characteristics and fertility decisions

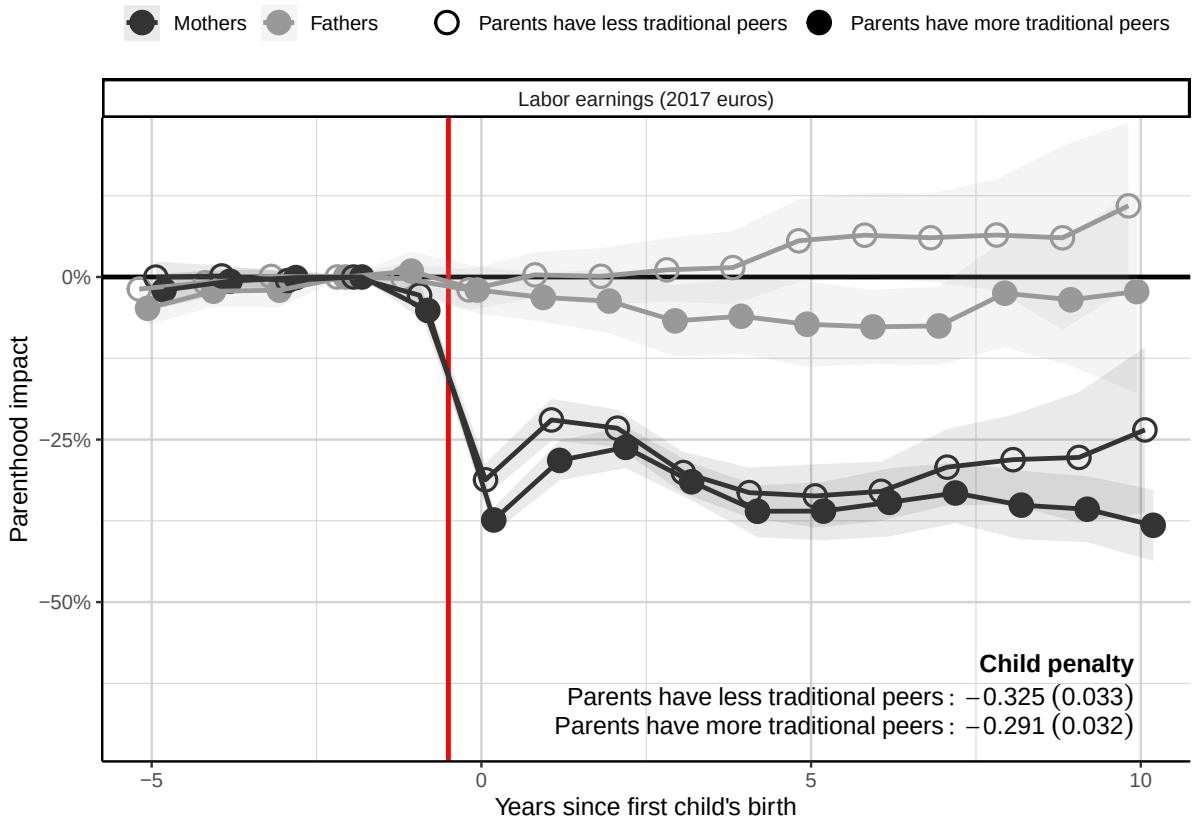


Upbringing environment is measured by *ego*'s father's and mother's having a religion, prominence given to religion in *ego*'s education before the age of 18, gender imbalance between *ego*'s parents' allocation of household chores (daily meals, grocery shopping and clothes washing) before the age of 18 and the number of children born to *ego*'s mother. Parenthood impact is identified from comparisons between parents of the same decennial cohort who had their first child at different ages, and displayed relative to the counterfactual labor market participation rate. The child penalty is the difference in parenthood impact between men and women, averaged over the first 10 years after their first child is born. The data are reweighted based on an inverse propensity score approach so as to make upbringing environments groups similar in terms of pre-birth observables characteristics and fertility decisions. Shaded areas correspond to 95% confidence intervals; they are based on a reweighted bootstrap approach, clustered at the individual level. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure 6. Child penalty in labor earnings for second-generation immigrant parents: by second-generation immigrant parents' parents' peers' attitudes, balanced on pre-birth characteristics and fertility decisions



Parents' peers' attitudes are measured as the share of first-generation female (resp. male) immigrants from *ego*'s mother's (resp. fathers') country of birth with above-median traditional gender-related attitudes. Parenthood impact is identified from comparisons between parents of the same cohort (year of birth) who got their first salaried job at the same time, but had their first child at different ages (at least two years after their first salaried job), and displayed relative to the counterfactual labor market participation rate. The child penalty is the difference in parenthood impact between men and women, averaged over the first 10 years after their first child is born. The data are reweighted based on an inverse propensity score approach so as to make origin groups similar in terms of pre-birth observables characteristics and fertility decisions. Shaded areas correspond to 95% confidence intervals; they are based on a reweighted bootstrap approach, clustered at the individual level.

Population. Second-generation immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), CCMSA, Cnaf, Cnav, DGFiP and Insee, permanent demographic sample (EDP), authors' calculation.

Tables

Table 1. Immigrant parents average outcomes: by gender-related attitudes

	Fathers		Mothers	
	Less traditional	More traditional	Less traditional	More traditional
Sample size	1,591	1,693	1,847	1,725
Population size	690,000	720,000	762,000	691,000
Year of birth	1974	1975	1975	1976
Age at first relationship	23.6	25.0	21.9	22.1
Age at first cohabitation	25.8	27.1	24.0	24.1
Age at first child's birth	31.8	31.1	29.3	27.8
Total number of children	2.0	2.4	2.0	2.4
Age at the end of education	20.8	19.6	21.0	19.8
Self-reported discrimination	0.22	0.29	0.22	0.24
First-generation immigrant	0.42	0.69	0.45	0.66

Traditional gender-related attitudes are defined by scoring above median on the first component of the PCA of the relevant survey items. *Population.* Immigrant parents living in mainland France in 2019-2020 *Source.* Ined and Insee, Trajectoires et Origines 2 survey (2019-2020).

Table 2. Impact of traditional gender-related attitudes on the child penalty, by potential children-related labor supply decisions

Under nontrad. views \ Under trad. views	$(1 \rightarrow 1)$	$(1 \rightarrow 0)$	$(0 \rightarrow 0)$
$(1 \rightarrow 1)$	=	−	=
$(1 \rightarrow 0)$	∅	=	+
$(0 \rightarrow 0)$	∅	∅	=

Under the assumptions that children always decrease mothers' labor supply, the only possible potential children-related labor supply decisions are (i) always to participate in the labor market $(1 \rightarrow 1)$, (ii) to leave the labor force $(1 \rightarrow 0)$ and (iii) never to participate in the labor market $(0 \rightarrow 0)$. Under the additional assumption that holding traditional views always decrease mothers' labor supply, three combinations of potential labor supply decisions are forbidden. The only two combinations that change the child penalty, that is the difference between what women do with children and what they would do without children, depending on the views they hold are $((1 \rightarrow 1) \rightarrow (1 \rightarrow 0))$, that is mothers who would always work regardless of children under nontraditional views, but would leave the labor market due to children under traditional views, who make the child penalty even more negative, and $((1 \rightarrow 0) \rightarrow (0 \rightarrow 0))$, that is mothers who leave the labor market due to children under nontraditional views, and would never participate in the labor market even without children under traditional views, who make the child penalty less negative.

Table 3. Regression of first-generation immigrant parents' attitudes on their peers' attitudes

	Traditional gender-related attitudes	
	Fathers	Mothers
Peers' attitudes	0.57	0.57
	(0.05)	(0.04)
Constant	0.27	0.24
	(0.03)	(0.03)
N	1,909	2,047
R^2	0.07	0.08
F-Statistic	143.03	167.50

First-generation immigrant parents' peers are defined as first-generation immigrant parents of the same gender as them who were born in the same country as them. This variable is computed using a leave-one-out approach. *Population.* First-generation immigrant parents living in mainland France in 2019-2020 *Source.* Ined and Insee, Trajectoires et Origines 2 survey (2019-2020).

Table 4. Regression of second-generation immigrant parents' attitudes on their parents' peers' attitudes

	Traditional gender-related attitudes	
	Fathers	Mothers
Ego's father's peers' attitudes	0.32	0.17
	(0.08)	(0.07)
Ego's mother's peers' attitudes	0.26	0.42
	(0.07)	(0.06)
Constant	0.08	0.09
	(0.04)	(0.04)
N	1,324	1,475
R^2	0.05	0.06
F-Statistic	33.18	47.20

First-generation immigrant parents' peers are defined as first-generation immigrant parents of the same gender as them who were born in the same country as them. *Population.* Second-generation immigrant parents living in mainland France in 2019-2020 *Source.* Ined and Insee, Trajectoires et Origines 2 survey (2019-2020).

A Couple penalties vs. child penalties

A.1 Method

Our approach extends the staggered event-study design to accommodate cases where individuals undergo two treatments. We build on the same foundational elements used in our main empirical analysis.

A.1.1 Model

We keep the same notations as in the core of the paper. Let $Y_{i,t}$ denote the relevant labor outcome – e.g., labor earnings or labor market participation – of individual i at age t. Let $G_i \in \{f, m\}$ denote individual i's gender and B_i her generation (year of birth).[18] Finally, let R_i (respectively C_i) denote her age when she began her first stable relationship (respectively when her first child was born), with $R_i = \infty$ if she never had a stable relationship.

Let $Y_{i,t}(r, c)$ denote the potential labor supply of individual i at time t had she (i) started her first stable relationship at age r; and (ii) given birth to her first child at age c. Consistently, $Y_{i,t}(\infty, \infty)$ represents her labor supply at time t had she chosen (i) never to enter a relationship; and (ii) to remain childless. By construction:

$$Y_{i,t} = Y_{i,t}(\infty, \infty) + \sum_{r,c} (Y_{i,t}(r, c) - Y_{i,t}(\infty, \infty)) \mathbb{1}\{R_i = r, C_i = c\} \tag{9}$$

Because there are two treatments here, the definition of treatment effects, and thus that of the quantities of interest, is not as straightforward as when only parenthood is considered. Specifically, for any individual i and time t, $Y_{i,t}(r, c) - Y_{i,t}(\infty, c)$ represents the partial effect of living as a couple for $t - r$ years, while holding parenthood status constant, whereas $Y_{i,t}(r, c) - Y_{i,t}(r, \infty)$ corresponds to the causal effect of having been a parent for $t - c$ years, while holding relationship status constant. In practice, we define two sets of ATTs:

$$CATT^{R}_{g,b,r,c,t} = \mathbb{E}[Y_{i,t}(r, c) - Y_{i,t}(\infty, c) \mid G_i = g, B_i = b, R_i = r, C_i = c] \tag{10}$$

$$CATT^{C}_{g,b,r,c,t} = \mathbb{E}[Y_{i,t}(r, c) - Y_{i,t}(r, \infty) \mid G_i = g, B_i = b, R_i = r, C_i = c] \tag{11}$$

Here, $CATT^{R}_{g,b,r,c,t}$ denotes the average impact of relationships for a group of individuals defined by the ages at which they began their first relationship and had their first child, while $CATT^{C}_{g,b,r,c,t}$ denotes the average impact of parenthood for the same group.

Our analysis relies on two key assumptions: (i) a parallel trends assumption and (ii) a limited anticipation assumption.

Assumption 3 (Parallel trends in baseline outcomes: relationships) *For all g, b, (t, t'), c, and (r, r'), where $c, c' > 1$ and $c, c' < \infty$, we assume:*

$$\begin{aligned} & \mathbb{E}[Y_{i,t}(\infty, c) - Y_{i,t'}(\infty, c) \mid G_i = g, B_i = b, R_i = r, C_i = c] \\ = \; & \mathbb{E}[Y_{i,t}(\infty, c) - Y_{i,t'}(\infty, c) \mid G_i = g, B_i = b, R_i = r', C_i = c] \end{aligned} \tag{12}$$

[18] We further discuss the consequences of comparing individuals within their own generation in Section 3.1.3.

Assumption 4 (Parallel trends in baseline outcomes: children) *For all g, b, (t, t'), r, and (c, c'), where $c, c' > 1$ and $c, c' < \infty$, we assume:*

$$\begin{aligned} & \mathbb{E}[Y_{i,t}(r,\infty) - Y_{i,t'}(r,\infty) \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] \\ = \; & \mathbb{E}[Y_{i,t}(r,\infty) - Y_{i,t'}(r,\infty) \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c'] \end{aligned} \tag{13}$$

Assumption 5 (Limited anticipation: relationships) *For all t, g, b, r, and c, if $t < r-1$, then:*

$$\mathbb{E}[Y_{i,t}(r,c) - Y_{i,t}(\infty,c) \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] = 0 \tag{14}$$

Assumption 6 (Limited anticipation: children) *For all t, g, b, r, and c, if $t < r-1$, then:*

$$\mathbb{E}[Y_{i,t}(r,c) - Y_{i,t}(r,\infty) \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] = 0 \tag{15}$$

Assumption 3 implies that, in the absence of stable relationships, the average labor outcomes for individuals of the same gender and birth cohort, and who have children at the same age, would evolve in parallel over time. Assumption 4 states that, in the absence of children, the average labor outcomes for individuals of the same gender and birth cohort, and who began their first relationship at the same age, would evolve in parallel over time. Assumption 5 (resp. 6) assumes that the effect of stable relationships (resp. parenthood) on labor outcomes is zero up to two years before the beginning of the first stable relationship (resp. the birth of the first child).

Taken together, these assumptions imply testable restrictions on the data. First, they suggest that before both couple formation and parenthood, average labor outcomes for all (r, c)-cohorts should evolve in parallel. This restriction is similar to the usual "no pre-trend" assumption in difference-in-differences approaches with only one treatment. Secondly, they imply that, before the first child is born, the average impact of relationships should be the same across future parents who began their first relationship at the same age but are expected to have their first child at different ages. They may, however, have different baseline levels.

If the first stable relationship closely follows the first stable job (Landaud, 2021), then Assumption 3 becomes implausible because the sorting of individuals based on the timing of their first relationship is strongly correlated with the trajectory of their potential labor outcomes. In this case, our estimates of the impact of relationships are likely to be biased upwards. One possibility is to consider only the differences in impact between men and women, under the additional assumption that the bias is the same for both genders. However, this approach may fail if the dynamic sorting is, for example, stronger for men than for women, as suggested by Landaud (2021). In that case, our estimates based on gender differences would overestimate the detrimental impact of relationships on women's labor supply.

A.1.2 Identification

Under these assumptions, if there is sufficient variation in the timing of the first stable relationship within each gender × cohort (year of birth) × entry into parenthood group, the group-specific average effects of stable relationships can be inferred from the data:

Proposition 2 (Difference-in-differences estimand: relationships) *For all* (g, r, c, t)*, with* $1 < r < \infty$ *and* $1 < c < \infty$*, the average treatment effect is given by:*

$$\begin{aligned} CATT^{R}_{g,b,r,c,t} \quad = \quad & \mathbb{E}[Y_{i,t} \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] \\ & -\mathbb{E}[Y_{i,r-2} \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] \\ & -\mathbb{E}[Y_{i,t} \mid G_i = g,\, B_i = b,\, \max(1,\, r-2\,, t+1) < R_i < \infty,\, C_i = c] \\ & +\mathbb{E}[Y_{i,r-2} \mid G_i = g,\, B_i = b,\, \max(1,\, r-2,\, t+1) < R_i < \infty,\, C_i = c] \end{aligned} \tag{16}$$

Conversely, if there is sufficient variation in the arrival of children conditional on when the first relationship begins, group-specific average effects of parenthood can be inferred from the data:

Proposition 3 (Difference-in-differences estimand: children) *For all* (g, r, c, t)*, with* $1 < r < \infty$ *and* $1 < c < \infty$*, the average treatment effect is given by:*

$$\begin{aligned} CATT^{C}_{g,b,r,c,t} \quad = \quad & \mathbb{E}[Y_{i,t} \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] \\ & -\mathbb{E}[Y_{i,c-2} \mid G_i = g,\, B_i = b,\, R_i = r,\, C_i = c] \\ & -\mathbb{E}[Y_{i,t} \mid G_i = g,\, B_i = b,\, R_i = r,\, \max(1,\, c-2\,, t+1) < C_i < \infty] \\ & +\mathbb{E}[Y_{i,c-2} \mid G_i = g,\, B_i = b,\, R_i = r,\, \max(1,\, c-2\,, t+1) < C_i < \infty] \end{aligned} \tag{17}$$

A.1.3 Estimation

These identification results suggest a simple plug-in estimator, where expectations are replaced with their empirical counterparts. In practice, we aggregate the impact over time relative to events in the same way as in our baseline approach. Standard errors are computed using the same reweighted bootstrap method as in our baseline estimates, with clusters defined at the individual level.

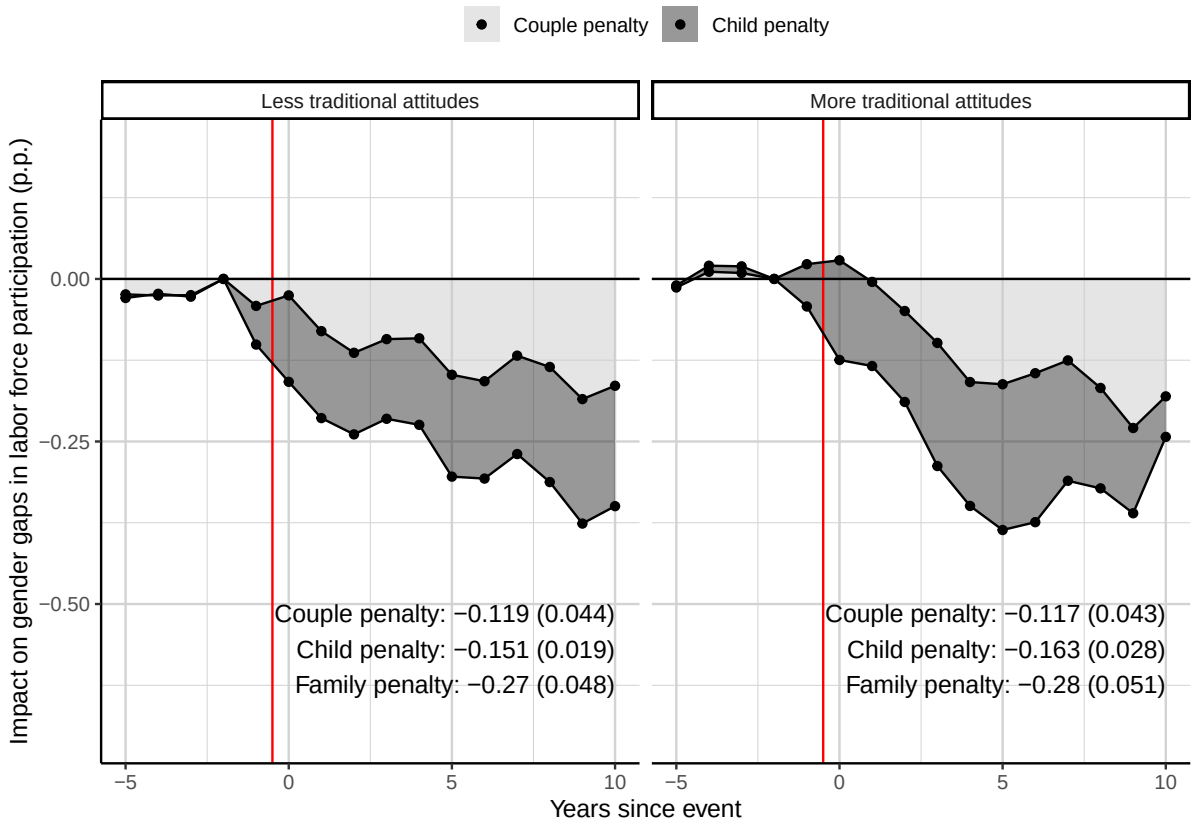
We operationalize the age at which individuals began their first relationship by considering when the relationship started, rather than (i) marriage or (ii) the moment they began living with their first partner. The reason for this choice is that (i) in France, most children are now born outside of marriage; and (ii) cohabitation is more likely to be tied to future plans regarding parenthood than the beginning of a relationship.

A.2 Results

Figure A.1 presents our results on couple and child penalties, jointly estimated by gender-related attitudes. It displays estimates based on comparisons between men and women, with impacts measured in absolute values (percentage points). While the dynamics of the effects differ slightly across attitude groups, the aggregate impacts of both relationships and children on gender gaps in labor force participation do not appear to depend on views regarding gender inequality and women's bodily autonomy. The couple penalty is notably large, especially compared to available estimates of marriage penalties at advanced stages of development, as presented by Kleven et al. (2024). This suggests that the bias resulting from the stronger causal impact of stable jobs on stable relationships for men compared to women may lead to an overestimation of the couple penalty.

Figure A.2 builds on these results to decompose the average gender gap in labor force participation between ages 20 and 50, into (i) one component stemming from the impact of relationships, (ii) one component resulting from the child penalty, and (iii) a residual gap unrelated to family events. For both attitude groups, family events explain more than 100% of the gender gap in labor force participation, implying that the residual gap is negative. In other words, in the absence of relationships and children, our estimates suggest that women would actually participate more in the labor market than men. The contributions of family events to the gender gap in labor force participation are quite similar across attitude groups, although the contribution of the child penalty is slightly smaller for less traditional immigrants, who are more likely to have children later in their lives. Consistent with our baseline estimates, we find that the larger gender gap in labor force participation for immigrant parents with more traditional views regarding gender is mainly due to the residual gap, which is strongly correlated with gender-related attitudes and unrelated to both relationships and parenthood.

Figure A.1. Child penalty in labor market participation for immigrant parents: by self-reported attitudes, balanced on pre-birth characteristics and fertility decisions

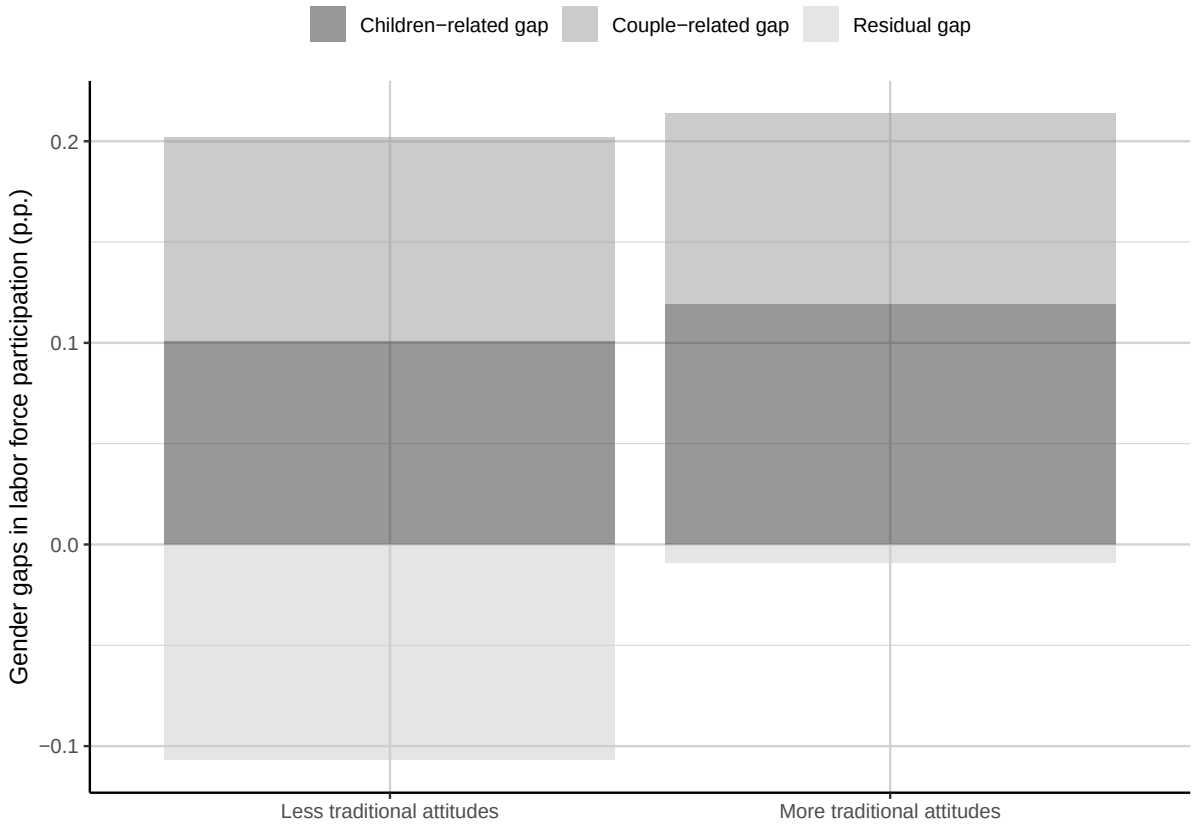


Couple (parenthood) impact is identified from comparisons between parents of the same decennial cohort and who had their first child at the same age (began their first relationship at the same age), but began their first relationship (had their first child) at different ages, and displayed in its absolute value. The couple (child) penalty is the difference in couple impact between men and women, averaged over the first 10 years after their first child is born. The family penalty is the sum of both these penalties. The data are reweighted based on an inverse propensity score approach so as to make attitudes groups similar in terms of pre-birth observables characteristics and fertility decisions. Standard errors are based on a reweighted bootstrap approach, clustered at the individual level. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.

Figure A.2. Gender gaps in labor force participation: contribution of the couple and the child penalties, by gender-related attitudes



Average labor force participation gaps from age 20 to age 50 for immigrant parents, by gender-related attitudes. This figure displays the contribution of the couple penalty and the child penalty to the gender gap. The contribution of the couple (child) penalty is the magnitude of the couple (child) penalty in labor force participation multiplied by the prevalence of stable relationship (parenthood). The residual gap is the difference between the realized gaps the sum of both these contributions. Couple (parenthood) impact is identified from comparisons between parents of the same decennial cohort and who had their first child at the same age (began their first relationship at the same age), but began their first relationship (had their first child) at different ages. First-generation immigrants are only included in the computation after they first arrived in France.

Population. Immigrant parents living in mainland France in 2019-2020.

Source. Ined and Insee, Trajectoires et Origines 2 survey (2019-2020), authors' calculation.