Immigration, Parenthood and Child Penalties*

Click here for the latest version

María Alexandra Castellanos†

November 20, 2023

Job Market Paper

Abstract

This paper analyses the impact of an immigration-induced expansion of household services on parenthood costs. In a unique quasi-experimental setting, I exploit the sudden supply shock in domestic and childcare services during the Spanish migration boom. I examine whether the availability of affordable substitutes for household production can reduce gender disparities associated with labour market parenthood penalties. Using a novel individual-level measure of the child penalty and a rich matched employer-employee administrative dataset, I combine a difference-in-differences strategy with a shift-share instrumental-variable design to estimate the causal effect of the shock. I find that the expansion of domestic services, driven by a large inflow of female immigrant workers, reduced the gender gap associated with child penalties for native workers. The responses are driven by two main channels: labour supply and job quality for native mothers. This includes employment in higher-paying firms, as well as better sectoral and occupational attributes. The effect is persistent over time and more pronounced for low-skilled native women, suggesting that affordable substitutes for household production can not only help alleviate gender gaps but also reduce within-gender inequality.

JEL Codes: J16, J31, J61

*Acknowledgements: I am deeply grateful to my advisor Jan Stuhler, for his constructive comments and support and to Daniele Paserman for his valuable advice and insightful discussions. I am also thankful to Luigi Minale, Camila Steffens, Warn N. Lekfuangfu, Jesús Fernández Huertas Moraga, Juan José Dolado, Johannes Schmieder and Felix Stips for their helpful comments as well as the participants of the Micro PhD Workshop at Boston University and the Applied Work-in-Progress seminar at UC3M.

†Universidad Carlos III de Madrid. Email: marcaste@eco.uc3m.es.

1 Introduction

Gender inequality and immigration have increasingly positioned at the core of the political and social agenda worldwide. While considerable progress has been made in reducing gender inequality in recent decades, there are still substantial gaps that persist in the labour market (Goldin, 2006; Olivetti and Petrongolo, 2016). These disparities may vary depending on the specific context, but in developed countries, recent literature has consistently pointed to one major driving factor: the unequal impact of parenthood on men and women (Bertrand et al., 2010; Goldin, 2014; Kleven et al., 2019a,b; Kleven, 2022; Cortés and Pan, 2020). Despite the consensus on the underlying causes, it is not yet clear how to effectively tackle the uneven costs of parenting and its long-run consequences on women's labour market trajectories. This study goes one step further to fill in this gap.

When we turn to the impact of immigration, the debate on the extent to which it affects native labour markets also remains unsettled. With many developed and developing countries experiencing rising immigrant inflows over the past few decades, the perceived burden placed by foreign-born workers and the concern about its economic impact have motivated strong political polarisation (Card et al., 2012). The lack of consensus also extends to the economics literature, where estimating the impact of migration on labour market outcomes remains a challenging exercise (Borjas, 2003; Card and Peri, 2016; Dustmann and Preston, 2012; Dustmann et al., 2016). So far, a considerable number of studies have aimed at understanding the extent of labour market competition among immigrants and natives. Focusing on a different margin, this paper provides causal evidence of the potential role of immigration in reducing gender disparities in the labour market.

To this effect, I exploit one of the major immigration episodes of recent history: the Spanish migration boom of the 2000s. In terms of inflows, between 1999 and 2007, the share of working age foreign-born population jumped from 3% to 14%, making it one of the largest immigration events of the post-war period worldwide (Moraga et al., 2019). Not only the magnitudes were striking, but interestingly this sizable arrival of immigrants translated into the largest push in the supply of domestic services that Spain had ever experienced. In less than 10 years, the number of workers employed in household and care-related services doubled, driven by the arrival of foreign female workers. By 2007, around a third of female immigrants were working in household services. In addition to a larger supply, the new workers providing these services were likely to do so at relatively lower prices (Farré et al., 2011; Carrasco et al., 2008; Bentolila et al., 2008) given that, despite minimum wage legislation, regulation on this sector was scant with many domestic workers being employed in the informal market (ILO, 2013). Combining all these characteristics, the above-mentioned historical context provides a unique large-scale quasi-experimental setting to causally estimate whether an increase in the availability of affordable domestic services can alleviate child penalties.

To identify the causal effect of an expansion of household services on parenthood penalties, I leverage regional and temporal variation from the arrival of female immigrants sorting into these sectors in a difference-in-differences framework with continuous variation in treatment. Considering that immigrant residential choices need not be random but may respond to local labour market conditions, I use a shift-share instrument that exploits networks from immigrants' past-settlements (Altonji and Card, 1991; Card, 2001) to predict female immigrant shares. From a methodological perspective, the sizable and sudden shock that characterised this immigration boom, along with important changes in the country of origin composition, provides a strong case for using this type of instrument (Jaeger et al., 2018).[1] While the instrument aims to deal with potential endogeneity stemming from immigrants' geographical sorting, I also conduct a set of placebo tests which show that the shock was not correlated with outcomes before the migration boom started (pre-period). In addition, the tests confirm that there were no pre-existing trends driving the results. These findings therefore also provide suggestive evidence in favour of the difference-in-differences parallel trends identifying assumption.

Regarding the outcome variable of interest, I construct a novel measure of child penalty at the individual level using an imputation-based counterfactual method. The reasons are two-fold. First, since I am interested in estimating the impact of household services on child penalties for parents, using the latter as an outcome itself becomes a key choice. Second, an individual level analysis allows me to evaluate in detail margins of adjustment, conduct a richer heterogeneity analysis and evaluate responses in terms of the quality of worker-job matches.

To compute individual parenthood penalties, I rely on several advantages of my data. In particular, I use the Continuous Sample of Working Lives (*MCVL*), a rich matched employer-employee administrative dataset from social security records that allows to track workers entire labour history. I merge individuals in the main dataset with household records from the municipality registry to obtain exact dates of birth of all workers in the sample and their cohabitants, including children. Thus, the high-quality longitudinal dimension of the MCVL allows me to follow workers throughout their entire working lives, linking childbirth events to their career paths. Exploiting the information of individuals before parenthood, I estimate their individual-specific trajectory in the absence of kids to recover a time-varying counterfactual measure of earnings.[2] I then define individual

[1] As argued by Jaeger et al. (2018), if the spatial distribution of immigrant inflows is stable over time, the instrument is likely to be correlated with ongoing responses to previous supply shocks. Structural breaks in the components of the instrument can deal with static sources of endogeneity. As noted by Moraga et al. (2019) the lagged impact of previous inflows can expected to be negligible in the Spanish context.

[2] The approach is adapted from the intuitive imputation method proposed by Borusyak et al. (2022)

child penalties by looking at realised earnings after the birth of the first child with respect to their individual estimated counterfactuals. My method therefore takes into account that workers may have heterogeneous income profiles subject also to their own fertility preferences. To the best of my knowledge, this is the first paper applying an imputation based counterfactual approach to measure individual child penalties. In addition, considering that my penalty measure is subject to parametric functional form assumptions on the earnings equation, I assess the robustness of my findings to the outcome definition. Specifically, I employ a matching design using propensity scores. In this approach, I match *treated* units (parents) with comparable *control* units (non-parents) and use these pairs to create an alternative measure of the individual parenthood penalty.

I find that the larger availability of domestic services propelled by the inflow of female immigrants reduced the natives' earnings gender gap associated with parenthood by an average of 29%. The results are statistically significant and become larger from the mid-to-late 2000s, when the arrival of foreign-born workers was fuelled by the two EU enlargements that took place. Moreover, the effects are driven by a reduction in child penalties for women. In this respect, I find that after childbirth, mothers earnings are about 4.1% higher if living in a local labour market with a larger share of female immigrants prone to work in household services (75th percentile of exposure) compared to those in less exposed regions (25th percentile). In contrast, the impact for men is closer to zero and not statistically significant. The differential effects are intuitive considering that men are less likely to experience a *fatherhood* penalty to start with, reflecting the unequal allocation of household and childcare responsibilities that lead to gender disparities in the first place.

Further, the impact is concentrated on less-skilled native mothers. This finding is consistent with what one would expect theoretically: if an expansion of the supply of household services leads to a reduction of caregiving prices, the effect should mostly occur among those families who are at the margin, i.e. those that were relatively more resource constrained relative to those that might have already been able to afford these services at higher prices. These results are robust to different definitions of skills, including alternative classifications in terms of education, pre-birth income and occupation. Moreover, they are robust to the exclusion of non-qualified labourers, ruling out that the effects were driven by direct competition or complementarities in the labour market at the bottom of the skill distribution.

Access to domestic services reduced mothers' child penalties through two main channels: labour supply and job quality. To shed light on these mechanisms, I explore the

which allows to recover individual specific heterogeneity for workers who are observed before and after an event that defines treatment status. In my case this event would be parenthood.

impact of the female immigration shock on the probability of employment, days of work and a series of job attributes including type of contract, daily wages, firm conditions and mobility to higher-paying jobs. I show that a larger provision of household services seems to help mothers return earlier to the labour market. Mothers living in more exposed areas were about 1.8 percentage points more likely to be employed one to three years after childbirth and 1.2 percentage points more likely after up to 5 years. Over the first five years, mothers worked for about 0.6 extra days per month. Moreover, mothers seemed to earn, on average 2% to 3% higher daily earnings, with almost no detectable effect on the probability of working full-time as opposed to part-time. Furthermore, I show that the expansion of these services allowed mothers to work in better-remunerated jobs in terms of occupation and sectoral attributes, and in higher-paying firms. Finally, using the sample derived from the matching approach and looking at women with comparable pre-birth characteristics, I test explicitly whether the expansion of these services benefited mothers and non-mothers alike. I find that only mothers' earnings display a positive and significant response. Altogether, these findings suggest that accessible and affordable substitutes for household production not only allowed for a faster reintegration of mothers into the labour market but also placed them in better career trajectories in terms of the types of jobs they could access.

This paper contributes to several strands of the literature. First, it adds to the growing literature on gender economics and child penalties. Following the approach of Kleven et al. (2019a,b); Kleven (2022), these empirical studies replicated for several countries,[3] use an event-study design around the timing of birth of the first child to assess the evolution of earnings and other labour market outcomes once female and male workers become parents. They therefore consider having a child as the *treatment* variable and the strategy recovers one average effect for all workers. In this study, I provide a measure of individual child penalties that can be used as an *outcome* in individual-level regressions. In effect, while I use this outcome to assess the effects of larger domestic services availability, it could also be employed to assess the impact of other types of labour market shocks or family-related policies. Moreover, my contribution to this line of research is not only methodological but also adds to the conceptual discussion on measuring parenthood costs (Bensnes et al., 2023). Previous literature has commonly conceptualised these penalties in a *before-after* comparison: i.e. how much earnings decline after childbirth with respect to the year before becoming parents. In turn, I assess the decline in earnings with respect to a *contemporary* counterfactual in the absence of kids.

My paper also speaks to the literature on the effects of childcare provision on maternal labour market outcomes (Gelbach, 2002; Berlinski and Galiani, 2007; Baker et al.,

[3] See also Kleven et al. (2023); Bertrand et al. (2010); Angelov et al. (2016); Sieppi and Pehkonen (2019); de Quinto et al. (2021); Andresen and Nix (2022).

2008; Berthelon et al., 2023a,b; Kleven et al., Forthcoming). In Spain, Nollenberger and Rodríguez-Planas (2015) found that the expansion of public childcare for 3-year-old kids in the early 1990s increased maternal employment substantially, revealing the important limitations associated with motherhood and external caregiving availability. In line with the results of my heterogeneity analysis, Berthelon et al. (2023a,b) find that extending school hours has a more positive impact on low-skilled mothers. I contribute to this literature in two key aspects. First, by looking at a more specific measure of gender disparities. Second, by assessing a different external source of childcare which is potentially more flexible than regular childcare centre arrangements. The rigidity of the latter along with persistent cultural norms might be the reason why in some contexts, traditional family policies have not been able to address gender inequality (Kleven et al., Forthcoming). I acknowledge that the greater allocation of care responsibilities to mothers also extends to other aspects of household production. Unlike childcare facilities, these can be substituted with domestic services, along with informal care. In fact, I find that the impact is stronger during the first three years after childbirth. In addition, while most of this literature focuses on maternal employment, the penalty measure I construct could also be used to further evaluate the effectiveness of these policies by looking at the impact on parenthood costs.

Further, this paper also contributes to the literature on the effects of immigration on native women. Cortés and Tessada (2011) find a positive effect on the labour supply of high-skilled women in the US. Cortés and Pan (2013) find that a visa program in Hong Kong for foreign domestic workers increased the labour supply of college-educated mothers of young children. In Italy, Barone and Mocetti (2011) report positive effects in the intensive margin only (i.e. more hours at work) with no effect on the extensive margin. In a cross-country comparison, Forlani et al. (2015) find a positive relationship between low-skilled immigration and labour supply with significant effects at the intensive margin for high-skilled women and at the extensive one for unskilled. Finally, in Spain, Farré et al. (2011) find an increase in high-skilled female labour supply as a result of immigration.

I expand on this literature in several aspects. First, while most of the evidence is centred on women's labour supply, I focus on earnings for both women and men and study directly the relationship between immigration and parenthood penalties in the labour market. Only a few other papers provide related evidence on gender gaps. An exception is a recent paper of Cortés and Pan (2019) who find suggestive evidence that low-skilled immigration allowed young women to enter occupations with higher returns to overwork, shifting women toward higher quantiles of the male wage distribution; and Llull (2021) who develop a structural model to analyse the effects of immigration on the gender gap in the US labour market. None of these papers, however, explore directly child penalty effects. To the best of my knowledge, this paper is the first one to investigate this outcome in the context of immigration of female domestic workers.

Second, almost all empirical studies on this field rely on cross-sectional information from Labour Force Surveys or similar sources to build native labour market outcomes. By contrast, the rich dataset of matched employer-employee records that I exploit here delivers several methodological advantages. First, as mentioned before, the high-quality longitudinal dimension, along with the information of children's exact date of birth, are crucial for the definition of my outcome variable. Moreover, I am able to control exhaustively for worker pre-determined characteristics (i.e. before childbirth events). Certainly, my scope is different. I focus on first-time parents and associated penalties rather than on women more generally, as most of the literature does. However, the restriction is relevant considering the role of parenthood on the persistence of gender inequality. In any case, being able to control for workers' labour market history constitutes an important advantage regardless of the outcome and sample definition.

In addition, the access to granular geographical information for workers and firms at the municipality level also allows me to construct a detailed measure of local labour markets. I define more than 125 local labour markets whereas previous studies for Spain exploit geographical variation at the aggregate province level (50).[4] To do so, I benefit not only from the detailed information from social security registers which I use for natives' outcomes, but also from the full records of the municipal Local Registry (*Padrón Continuo Municipal*) which I use to measure immigrant population by country of origin.[5] Moreover, by combining a shift-share IV with a difference-in-differences design, I am not only able to test explicitly for pre-trends, which none of the above-mentioned studies has tested for, but also to better disentangle the immigration effect from general area-fixed effects.

As mentioned earlier, I show that the responses are mostly driven by relatively lower-skilled mothers, a result that differs from previous studies such as Cortés and Tessada (2011) who find an impact on very high-skilled women. Both sample definition and context-specific characteristics can play a role in these discrepancies. First, I focus on a novel outcome variable and look at a particular point of workers' career –entering parenthood–. Second, in terms of historical context, Cortés and Tessada (2011) evaluate

[4]This includes Gonzalez and Ortega (2011) who look at native labour market effects of immigration combining a spatial and skill-cell approach folllowing (Borjas, 2003), Farré et al. (2011) who study the impact on female employment, Monras et al. (2018) who assess the effect of a legalisation amnesty. An exception is Moraga et al. (2019) who exploit variation at the census tracts level since they look at the impact of immigration on natives residential choices; so this extremely fine geographical level instead of a local labour market serves their purpose better.

[5]Immigrant population can arguably be measured with high accuracy given that, in Spain, registration in the municipality guarantees access to public health and education regardless of legal status and cannot be used by the police for deportation purposes. Thus, this is a key institutional advantage that provides incentives for undocumented workers to be registered.

the US labour market between the 1980-2000, where the new immigration inflows as share of the population were considerably lower than the shock experienced by Spain. By that time, the share of female immigrants in the domestic sector in the US was large (nearly 25% according to these authors). However, it is not comparable to that of Spain where, following the immigration boom, around 60% of workers in that sector were foreign-born. Hence, both factors translate into a larger supply shock of domestic services. From a policy perspective, finding that mothers at the lower part of the skill distribution are more responsive suggests that higher provision of affordable substitutes for household production through immigration may not only help alleviate gender disparities overall but also to reduce within-gender inequality.

The remainder of the paper is organised as follows: Section 2 documents the main features of the Spanish migration boom. Section 3 describes the data sources. Section 4 describes in detail the methodology used to measure child penalties at the individual level, the empirical design and identification strategy. Section 5 reports the main empirical results, Section 6 includes several extensions and robustness checks. Finally, Section 7 concludes.

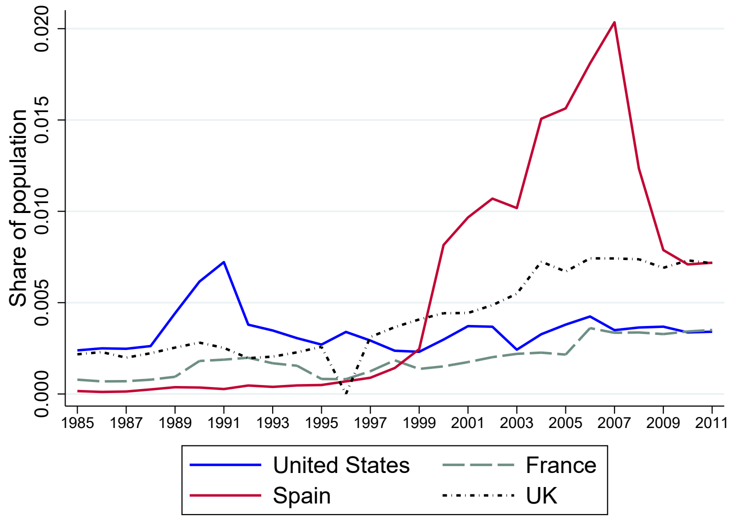
2 Background: Immigration in Spain

The sudden shock of immigrant inflows that Spain experienced between 1999 and 2008 not only contrasts sharply with other traditionally important hosts of overseas citizens but also with its own previous history. Figure 1 compares the inflows of foreign nationals as a proportion of the total population for four selected countries. According to the OECD International Migration Database, between 1985-1995, the share of foreign national inflows represented, on average, around 0.04% of Spain's total population. This trend spiked up by the end of the decade when the Spanish migration boom started. By 2000, the inflows of foreign nationals were 2.3 times as large as the 1999 figure, reaching 0.8% of the total population and as much as 2.0% by 2007.

In terms of stocks, the share of the foreign-born population in Spain rose from 3.1% in 1999 to 11.6% in 2007. The figure was even higher for the population between 18-65 years old, reaching 14.2% in the same year and 16.1% by 2008. With an annual average growth rate of almost 17%, these large labour-supply shifts exceeded a 30% growth rate in 2001 and 2002. From 2008, with the beginning of the global financial crisis, the net inflow of immigrants started to decline with the outflows surpassing the inflows between 2010 and 2014.[6] Yet, the overall stocks have remained relatively constant at around 13%-14% of the population until 2019.

[6]OECD, International Migration Database.

Figure 1: Inflows of foreign population by recipient country



Notes: Source: OECD - International Migration Database.

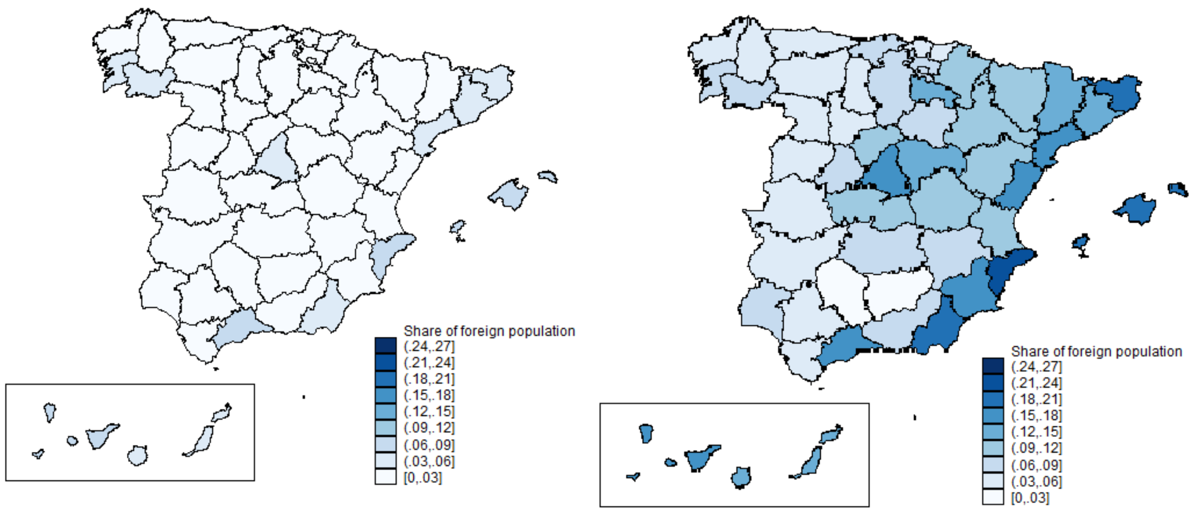
In addition to the sizable shifts in immigrant shares, two other features make the Spanish migration boom an attractive case to study from a methodological perspective. First, immigration displayed important geographical heterogeneity (Figure 2). In absolute numbers, Madrid and Barcelona (the largest cities in Spain) attracted the largest number of immigrants. However, in terms of population shares other provinces stood out such as Alicante and the Balearic Islands in the South-East and East (23.1% and 20.5% by 2007) or Almería and Málaga in the South coast (18.9% and 16.9% respectively). In some areas, the proportion of the foreign population became more than twice as large as the initial figures. In the West and some Southern inland, the shares reached between 3% and 6%, and as much as 9% in other provinces like Ourense in Galicia.

Second, there was also high variation in terms of the composition by country of origin. Three groups from culturally and geographically diverse backgrounds became especially relevant by the mid-2000s: South America, Central and Eastern Europe (driven mainly by Romania) and Morocco. The sudden rise of immigrants from South America was to a great extent explained by the severe economic crisis that affected Ecuador in the years 1998-2000 resulting in large outflows of Ecuadorians to Spain and the US.[7] By 1999 the group accounted for less than 1% of the foreign-born population. Three years later, it represented around 10%. Other important countries of origin were Colombia and Argentina which also experienced financial and political instability at the time. Furthermore, different reasons seemed to motivate immigrants' location choices such as geographical proximity in the case of Morocco, in contrast to (old) family linkages for some South American immigrants (Gonzalez and Ortega, 2011). Likewise, it has been documented that early

[7]See Bertoli et al. (2011) for an extended analysis of the Ecuadorian migration experience.

Figure 2: Share of foreign-born population in Spain by province

(a) 1999 (b) 2007



Notes: Source: Municipal Local Registry.

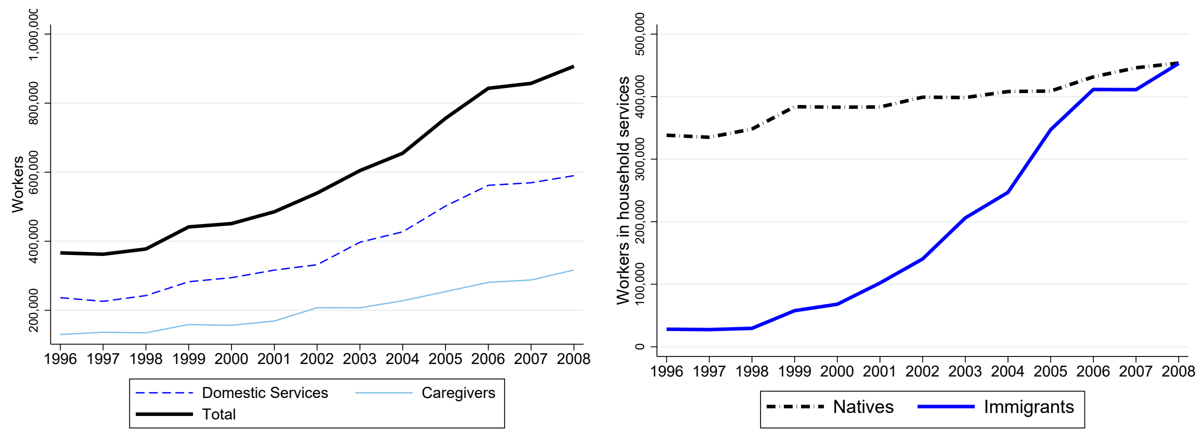
settlements of Romanians in provinces like Madrid, Castellón and Almería followed social, family and religious networks through the Adventist Church (Maza et al., 2013; Bernat and Viruela, 2011). In fact, according to the 2007 Spanish National Immigrant Survey, over 80% of immigrants reported having a local contact (a friend or relative) that could help them upon arrival (Farré et al., 2011). Representing between 12% to 33%, some of these groups quickly surpassed other traditional origin groups, such as Germany and France. As it will be discussed further in Section 4.2, these characteristics provide a considerably favourable setting when it comes to the implementation and identification through the early-settlement instrument.

The immigration boom pushed the supply of workers in the domestic service sector to unprecedented levels. As shown in Figure 3a, between 1999 and 2008, the number of employed workers in the sector doubled, from around 440,000 to nearly 900,000 by the end of the period. This surge was almost entirely driven by the arrival of foreign-born workers, as illustrated by the fact that the number of native workers in this sector remained relatively stable (Figure 3b). By 2008, the number of immigrant workers employed in domestic and caregiving services accounted for 65% and 25% of each sector, respectively (Appendix Figure H.1a). Moreover, according to Labour Force Survey data, this sector became one of the largest employers of female immigrant workers, with nearly one-third of women employed in it (Appendix Figure H.1b).

Compared to other countries, this sector became particularly important for the immigrant workforce and the overall economy in Spain. In the late 2000s, approximately 12.2% of immigrant workers, both men and women, were employed by private households (Figure 4a). This participation rate was considerably higher than in other major immi-

Figure 3: Evolution of workers in the domestic and caregiving sector

(a) All workers (b) Native and immigrant workers

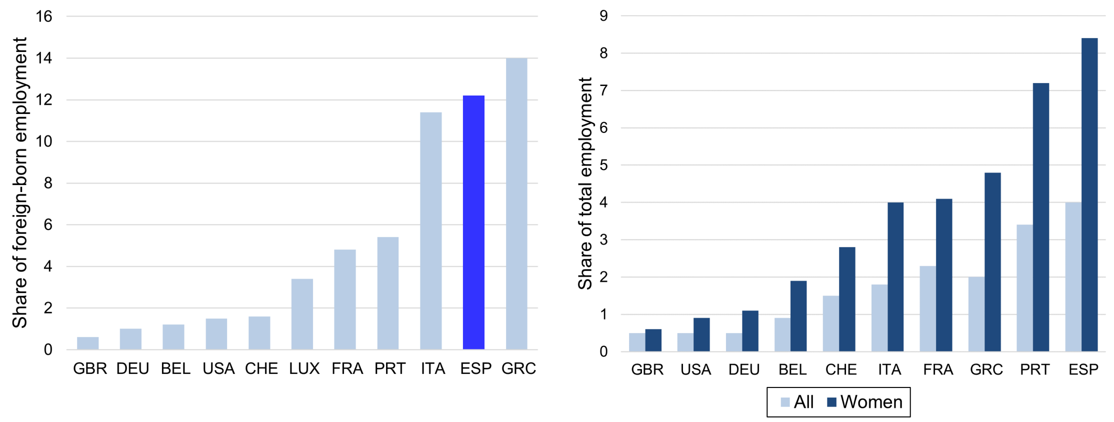


Notes: Panel a) displays the evolution of workers in the domestic and caregiving sector irrespective of origin country. Panel b) looks at thee evolution of workers from both sectors splitting by native and immigrant origin. Source: Labour Force Survey.

grant host countries, like Germany, the UK and the US, where the share was below 2%. Among Western European countries, only Italy and Greece displayed similar figures to Spain. Furthermore, due to the sizable labour supply shock that Spain experienced, the domestic service sector became one of the largest in Europe, accounting for 8.4% of total employment (Figure 4b).

Figure 4: Domestic Workers in Europe and the US

(a) Foreign-born workers employed in private households (b) Domestic workers as share of total employment



Notes: Panel a) Share of foreign-born workers employed in private households by 2007. Source: OECD (2009) based on European LFS and Current Population Survey for the United States. Panel b): Domestic workers (native and immigrants) as share of total employment by 2008-2010. Source: ILO (2013).

Almost half of the foreign-born women working in these sectors were from Ecuador, Romania and Colombia, precisely the countries that led the migration boom (Appendix Table H.1, and Figure H.4), along with Morocco. Although Romania joined the EU in the 2007 enlargement, by the early 2000s 1 out of 10 immigrants in household services came from this country. While inflows from Morocco have been among the largest in Spain, their participation in these services was relatively smaller. In contrast to other countries of origin, the inflow of Moroccan men and women exhibited the largest differences in terms of gender composition. The number of men arriving from this country was almost twice the number of women, working mostly in the construction sector which experienced a boom prior to the Great Recession in Spain. At any rate, the share of Moroccans in household services also increased slightly by the end of the period while the rest of the providers came mostly from other Latin American countries besides Colombia and Ecuador. Key to my empirical design, the share of women from these countries also exhibitied substantial spatial variation (Figure H.2).

Finally, when it comes to quantifying immigration inflows, the Spanish legal framework provides an important advantage relative to other countries. As explained in Dolado and Vázquez (2008), a great majority of immigrants have traditionally entered the country irregularly or via a tourist visa (that was supposed to last for 3 months). Only a fraction of these workers acquired legal status through one of the legalisation amnesties that have taken place.[8] However, despite a person's legal status, registration in the municipal Local Registry (*Padrón Municipal*) grants access to local services such as public education and health.[9] Since immigrants were not asked about their legal status when registered, the Local Registry has turned into one of the most representative and reliable statistical sources for measuring the immigrant population in this country (Moraga et al., 2019; Amuedo-Dorantes and De La Rica, 2011; Gonzalez and Ortega, 2011).

3 Data

Earnings and Employment

To construct natives' labour market outcomes I exploit the microdata from the *Continuous Sample of Working Lives (Muestra Continua de Vidas Laborales, MCVL, in Spanish)*. This matched employer-employee administrative dataset is obtained by combining social security registers, income tax returns, and municipality registration records for a 4% non-stratified random sample of the population that has had any relationship with the social security for a given year. The dataset accounts for nearly 1.1 million individuals each year and allows for the construction of the labour market histories of workers in the sample back to 1967, with earnings data back to 1980. The construction of the estimation dataset follows closely the design of De la Roca and Puga (2017) combining the 2006-2017 waves

[8] 1985, 1991, 1996, 2000, 2001 and 2005

[9] The incentives to register were further enhanced following the approval of the Law 4/2000, *"Law 4/2000, 11th of January, about rights and liberties of foreigners in Spain and their social cohesion"*.

of the MCVL and incorporating specific sample restrictions relevant for this analysis.[10]

The MCVL reports two sources of earnings: social security and tax records. The former are censored for around 12.5% of the sample, from which nearly 10% corresponds to top-coded observations. The latter corresponds to uncensored income information from matched tax records based on worker and firm identifiers; however, this information is only available from 2006. Given that the time dimension of the analysis goes back to the late 1990s, I focus on social security earnings. In those instances where individuals were employed in multiple jobs simultaneously, I consider the job with the highest earnings. For each individual, I define earnings by aggregating all income from labour at the annual level given that immigrant stocks (obtained separately) are available with this frequency. Additionally, based on the distribution of earnings by industry of each year I winsorise them at the 2.5th upper and lower percentiles.

For employment status, I generate a dummy variable that takes a value of one if a worker was employed at least one month of the year. Additionally, since I am interested in the impact on parenthood-related costs, the sample is restricted to Spanish-born workers between 25 and 50 years old, that is, the age range when it is more likely that they had recently become parents. As a result, to measure parenthood costs I restrict the sample to workers who are observed in the dataset before and after having their first child. More details of the construction of this outcome are provided in section 4.1. Lastly, I restrict the sample to workers who had at least one working spell registered in the social security before 1999, at the beginning of the immigration boom.

In terms of workers' characteristics, information is drawn on their age (exact date of birth) sex and education level (below secondary, secondary and college).[11] Experience and tenure can be computed with high precision using the employment history. In addition, I observe 10-digit NACE occupational codes. For sectors, I use 2-digit level industries. Regarding location, the MCVL enables the identification of workplace at the municipality level for firms operating in municipalities with a population of at least 40,000. Using this information, workers are matched to urban areas as defined by Spain's Ministry of Housing in 2008.[12] These areas account for around 69% of workers in the sample. Additionally, to

[10] In terms of representatives, notwithstanding its retrospective design, a comparison of the MCVL with other data sources– such as the Spanish Labour Force Survey from previous studies–, have shown that the past cross-sectional age-distributions of male earnings remain representative up to the late 1980s, while discrepancies for women's become reasonably small in the 1990s. (Bonhomme and Hospido, 2017; Fernández et al., 2010). Thus, this should not be a major issue for this study since it mainly focuses on labour market outcomes from 1999, when the migration boom started.

[11] The MCVL reports education from the Local Registry. According to De la Roca and Puga (2017) a complete national update of the educational attainment was implemented in 1996, with a subsequent update by most municipalities in 2001. Since 2009 the Ministry of Education directly reports individuals' highest educational attainment to the National Statistical Institute which is then used to update the Local Registry records.

[12] See De la Roca and Puga (2017) for a complete description of urban area definitions.

include rural areas, workers who do not belong to any of the identified urban areas are grouped together at the province level. The analysis excludes Ceuta and Melilla, the two Spanish autonomous cities located in Africa. In total, the initial sample assigns workers to 129 regions: 79 urban areas and 50 rural areas.[13]

Family structure
To identify mothers and fathers, as well as the timing of childbirth I use the cohabitants' information available in the MCVL, which originates from the Local Registry. For each year, these records provide details on the date of birth and sex of every person living with each worker. Since the actual relationship with cohabitants is not explicitly provided, I assume that an individual is a parent if a child is born when the worker was between 18 to 45 years old. Additionally, I identify the birth order. As already mentioned, for the empirical analysis I focus on workers' outcomes after the birth of the first child.

Immigration inflows
The immigration inflows are constructed using the information from the Local Registry (*Padrón Continuo Municipal*). The dataset corresponds to the population of all individuals registered in the municipality where they live. A key feature of these records is that despite a person's legal status, registration grants access to local services such as public education and health in the municipality of residence. Immigrants are not asked about their legal status when registered and the authorities cannot use these records for deportation purposes. Moreover, registration served as proof of residence which was required to be eligible for the regularisation amnesties that took place over this time period.[14] These features provide strong incentives for immigrants to register, turning the Local Registry into one of the most reliable and representative sources of information to measure the immigrant population in Spain (Amuedo-Dorantes and De la Rica, 2008; Moraga et al., 2019).

Immigrants are defined as foreign-born workers. Since I am interested in the effect of the larger supply of domestic and care services, I restrict the inflows to working-age female immigrants from 18 to 65 years old. Additionally, I restrict the sample to the top 12 origin countries whose workers were more likely to sort into these services. Based on their specialisation rates according to the Labour Force Survey (LFS), I include Ecuador, Romania, Morocco, Colombia, Argentina, Bolivia, Peru, Dominican Republic, Brazil, Paraguay, Poland and Bulgaria. In total, employees from these countries account for nearly 80% of immigrant workers in the sector. The UK, origin of one of the largest immigrant inflows to Spain is excluded since a large proportion of British people arrived as retirees and were not working in domestic and related services. For the same reasons,

[13]This excludes urban areas that have a too small population to be identified in the MCVL or for which employment cells are extremely small (less than 10 workers), e.g. Blanes & Lloret de Mar.

[14]2001, 2005 and 2006. See Monras et al. (2018).

I exclude France, Germany and other high-income countries that were among the top immigrants' origins in Spain. Aggregating these inflows, I build regional cells based on urban areas and rural province-level areas as described in the previous sections (Appendix Figure H.2). I then match the immigration inflows using the workplace location of workers from the MCVL.

4 Methodology

4.1 Measuring parenthood costs

This section provides the specifics on the construction of my main outcome variable. To determine whether immigration has affected the labour market costs associated with parenthood I propose a new measure based on the concept of child penalty. While the classic event-study regression used in the child penalty literature delivers *one* average treatment effect for all treated workers, this measure provides me with an outcome at the individual level that I can use later in my main regression specification. To define an individual penalty, I conceptualise it in terms of the following question: What would be the earnings trajectory in the absence of kids? I start by defining an imputation-based counterfactual inspired by the recent developments in the literature on event-studies.[15]

For illustrative purposes, let Ω_0 be the set of non-parents or *untreated* in terms of the child-related event. These are workers who have either not become parents yet (but will do so later) or those that will never have a kid while observed (*never treated*). Let Ω_1 be the sample of parents or *treated* (in terms of children-related event) where a person becomes treated from the year of the birth of the first child onwards. Using both samples, I follow the steps outlined below to construct what I define as the *parenthood earnings penalty* (PEP henceforth):

1. First, using the sample of untreated observations only ($i_t \in \Omega_0$), I estimate by OLS the following *mincerian* fixed-effects regression, separately for men and women:

$$Y_{it}^g = A_{it}'\lambda_i^g + X_{it}'\Gamma^g + \alpha_t^g + \varepsilon_{it}, \tag{1}$$

where Y_{it}^g denotes earnings for worker i at year t of each gender g, λ_i^g captures time-invariant individual heterogeneity associated to the vector of individuals A_{it}. Thus, to some extent the design acknowledges potential heterogeneous income profiles driven by

[15]The imputation procedure is similar to Borusyak et al. (2022), however, unlike these authors, I only use the first steps to produce an outcome measure of parenthood labour market costs

individual preferences. X_{it} are common covariates including age fixed effects and interactions of age and education that capture life-cycle patterns, whereas α_t captures business cycle trends via time fixed-effects. I recover the estimates $\hat{\lambda}_i^g, \hat{\Gamma}_i^g, \hat{\alpha}_t^g$ from these regressions.

2. Using the estimated parameters from step 1, I predict the counterfactual earnings for each treated (parents) observation ($i_t \in \Omega_1$) by setting:

$$\hat{Y}_{it}^g(0) = A'_{it}\hat{\lambda^g}_i + X'_{it}\hat{\Gamma^g} + \hat{\alpha_t^g} \tag{2}$$

3. Finally, the *parenthood earnings penalty* is defined as the difference between the observed and counterfactual earnings, normalised by the counterfactual earnings in the absence of kids:

$$PEP_{it}^g = \frac{Yit^g - \hat{Y}_{it}^g(0)}{\hat{Y}_{it}^g(0)} \tag{3}$$

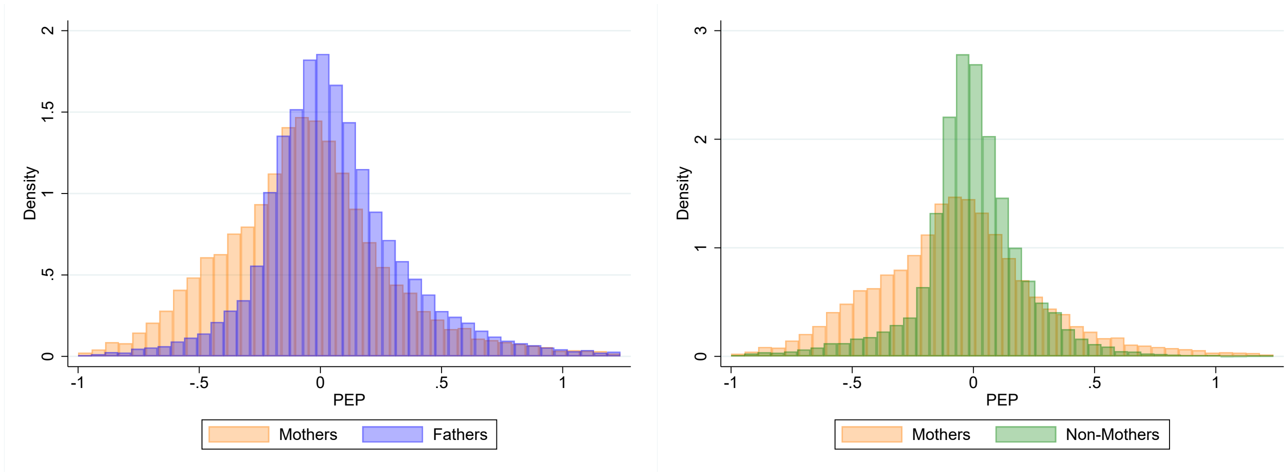
Note that to estimate the individual counterfactual, parents should be observed before and after having a child. The observations during the pre-parenthood period enter the first regression (step 1) allowing to estimate individual fixed effects that will then be used for the counterfactual simulation exercise. Based on this definition, workers experience an unconditional child penalty when $PEP_{it} < 0$ and a premium when $PEP_{it} > 0$. If the immigration shock allowed parents to better reconcile family and work, one would expect this shock to have a *positive* effect on the PEP, which is equivalent to a reduction in parenthood-related costs.

Figure 5a depicts the distribution of the parenthood penalty (PEP) for mothers and fathers, where the sample is restricted to currently employed individuals. As can be observed, a larger concentration of mothers experience an (unconditional) penalty (i.e. PEP < 0) compared to fathers, as indicated by their right-skewed distribution. This asymmetry is consistent with the child-penalty evidence, where women's earnings are largely and disproportionately affected following childbirth compared to their male counterparts. Complementary, Appendix Figure B.1b plots the distribution of the PEP for mothers and fathers considering all workers regardless of working status. The point mass at zero in that Figure reflects the asymmetry in non-employment for mothers compared to fathers. I test formally for the differences in distributions in Appendix B using the Goldman and Kaplan (2018) methodology. I observe that the differences between fathers and mothers are significant over the whole range of the empirical distribution, in favour of men.

In addition, I plot the distribution of an artificial counterfactual penalty for non-treated women: i.e., those who are not yet mothers. Figure 5b compares its distribution

Figure 5: Parenthood Penalty Distribution

(a) Employed parents (b) Artificial penalty for non-mothers



to the one of mothers, making the skewness of mothers even more evident. On the contrary, the distribution for non-mothers is thin-tailed and symmetric, both features being somewhat expected as they were part of the sample used to build the counterfactual earnings. Nonetheless, the fit provides reassuring evidence regarding the accuracy of the counterfactual construction, showing that the PEP for mothers is not a result of overestimated counterfactual earnings for women, which would mechanically lead to larger penalties for them.

Descriptive statistics of the penalty outcome and the rest of variables are shown in the Appendix Table A.1. On average, the gap between fathers' and mothers' was of 15 percentage points for the selected sample.

4.2 Empirical Strategy

To estimate the impact of immigration on parenthood costs I leverage geographical and temporal variation of the female immigrant distribution. Specifically, I regress the *parenthood earnings penalty* PEP for worker i in year t on the share of foreign-born working-age women M^f living in area r at time t, divided by the sum of the total foreign-born (I) and native (N) working-age female local population as of 1999, according to the following baseline specification:

$$PEP_{it}^{k} = \alpha + \beta_1 D \cdot {M^f}_{r(i99)t} + \beta_2 D \cdot {M^f}_{r(i99)t} \cdot g_i + \theta X_{i(pre-birth)} + \rho_{gt} + \delta_{gr} + e_{it}, \quad (4)$$

where

$${M^f}_{r(i99)t} = \frac{I^f_{r,t}}{(I+N)^f_{r,99}},$$

g is a gender binary variable that takes a value of 1 for men, and k denotes the number of years after the birth of their first child. Two subsamples are considered to assess the impact of immigration on parenthood costs. The first one is a subsample of parents who had their first kid in the last 5 years: $k = 5$, thus over the period before entering primary school; and the second one consists of a subsample of parents that had a kid in the last 3 years: $k = 3$ (before entering kindergarten). In terms of immigrant shares, I restrict the supply shock to the immigrant female workers specialised in domestic services. Moreover, normalising by 1999 population allows me to reduce endogeneity concerns stemming from local demographic responses.[16]

As shown in Figure 1 the Spanish migration boom was sudden, with rising inflows over the 2000s depicting a strong break in the trend slope compared to the 1990s decade. D is a binary variable that captures this change: it takes a value of zero before the immigration boom started (1994-1998) and a value of one from 1999 onwards (1999-2007). I add an interaction term of the migration shock with the indicator for gender g. This difference-in-differences with treatment intensity type of design allows to recover average treatment effects pooling all years together *before* and *after* the boom started. $X_{i(pre-birth)}$ is a vector of worker's predetermined characteristics i.e. before an individual became a parent, including industry at the 2-digit level and experience. I also control for age and workers' education. Lastly, I include interactions of gender with area and time fixed-effects. Robust standard errors are clustered at the area level to allow for serial correlation within areas over time. Additionally, I estimate an alternative specification that allows me to test directly for the presence of pre-trends. More details are provided below in Section 5.1.

I therefore exploit geographic, time and gender variation as in a triple difference approach. The estimated coefficient of β_1 in regression (4) captures the average effect of migration on women's motherhood penalty, $\beta_1 + \beta_2$ the effect on fathers, and β_2 the impact on the gender gap between fathers and mothers parenthood costs. Recall that given its definition, the PEP has an inverse relationship with the costs that parents are facing: positive values denote a premium, negative a penalty. Therefore, $\beta_1 > 0$ (positive) would imply that immigration reduced the penalty for women, and $\beta_1 + \beta_2 > 0$ that it reduced the penalty for men. Under that scenario, $\beta_2 < 0$ (negative) would imply that the shock reduced the gap between the two groups, favouring women by a larger extent.

Given the continuous nature of the immigration shock, a causal interpretation of these parameters in a difference-in-differences framework may be subject to what Callaway et al. (2021) refer to as a *strong parallel trends* assumption.[17] In essence, it requires that on

[16]See Moraga et al. (2019) for evidence on the effect of immigration on native population mobility and residential choices.

[17]Assumption 5 of Callaway et al. (2021): $\mathbb{E}\,[Y_t(d) - Y_{t-1}(0)] = \mathbb{E}\,[Y_t(d) - Y_{t-1}(0)\,|\,D = d]$ for an

average, across all amounts of treatment, or which they refer to as *"doses"*, there is no selection into a particular *dose*. To validate this assumption, I apply an instrumental variable (IV) strategy that allows me to isolate exogenous variation of the immigration shock. I discuss my identification strategy further in the next section.[18]

4.3 Identification

The previous specification could be directly estimated by OLS. However, these estimates might be subject to endogeneity problems stemming from two main sources. First, labour supply and thus immigrants' (and natives') location choices may respond to local labour market conditions including region-specific demand shocks. For instance, immigrants could sort into those regions where employment and wage prospects are better leading to upward biased OLS coefficients. Second, as pointed out by Carrasco et al. (2008), lower-skilled immigrants tend to cluster in activities that are less appealing to the native labour force, and thus where native employment and wages are lower. In that case, OLS could overestimate any negative effects derived from a (positive) labour supply shock.

To deal with these identification threats, I make use of a past-settlement instrumental variable (IV) strategy, originally proposed by Altonji and Card (1991) and Card (2001). This shift-share type of instrument leverages the historical geographic distribution of immigrants and country of origin networks. It takes advantage of the tendency of new immigrants to settle in regions where people from the same origin are already concentrated. This IV has been widely used in the literature analysing the impact of immigration in many host countries, including Spain in the work of Gonzalez and Ortega (2011) and Farré et al. (2011). More recently, this IV has been used by Moraga et al. (2019) to analyse the impact of migrant location on native residential preferences, Castellanos (2020) to examine the impact of immigration in native local labour markets outcomes and Özgüzel (2021) to assess the impact of return migration during the Great Recession. I define the instrument by:

$$Z_{rt(99)} = \sum_o \frac{I^f_{o,r,t_0}}{I^f_{o,t_0}} \times \frac{I^f_{o,t}}{(I+N)^f_{r,1999}} \tag{5}$$

where the share of immigrants of country of origin o located in a region r in an initial time period t_0 is multiplied by the national stock of working-age immigrants of origin o at

outcome Y and treatment D.

[18]Notice that the growing TWFE literature is mainly focused on dealing with problems associated to the staggered adoption of a binary treatment, which does not correspond to my context. Recent contributions aiming to deal with problems of negative weighting and treatment effect heterogeneity in those settings include: De Chaisemartin and d'Haultfoeuille (2020); Sun and Abraham (2021); Callaway and Sant'Anna (2021); Borusyak et al. (2022). A detailed review of these and other methods can be found in Roth et al. (2023); De Chaisemartin and d'Haultfoeuille (2023).

the national level, all normalised by 1999 local population. As discussed above, I restrict the sample of countries of origin to those that fuelled the migration boom and whose workers were more likely to sort into the domestic sector.

I set $t_0 = 1998$ which is the earliest year for which disaggregated data with *detailed* country of origin information at the municipality (and therefore urban area level) can be obtained.[19] As is well known, the validity of the instrument relies on satisfying the exogeneity and relevance conditions. For the exogeneity identification assumption to hold, labour market trends during the 2000s should not be correlated with immigrant shares in the 1980s and 1990s. The contrasting location patterns of the main origin groups that became particularly important in the 2000s wave suggest that non-economic reasons motivated their choices, making the identification assumption more plausible. For instance, apart from Madrid and Barcelona which have historically concentrated an important share of immigrants from different origins, there were large settlements of South Americans in the Canary Islands and Galicia (northeast of Spain) in the early 1990s. These regions had historical ties and kinship relationships with Spanish ancestors who had emigrated at the beginning of the 20th century (Gonzalez and Ortega, 2011). By contrast, many Moroccans clustered along the South-East coast, which is geographically closer to their home country.

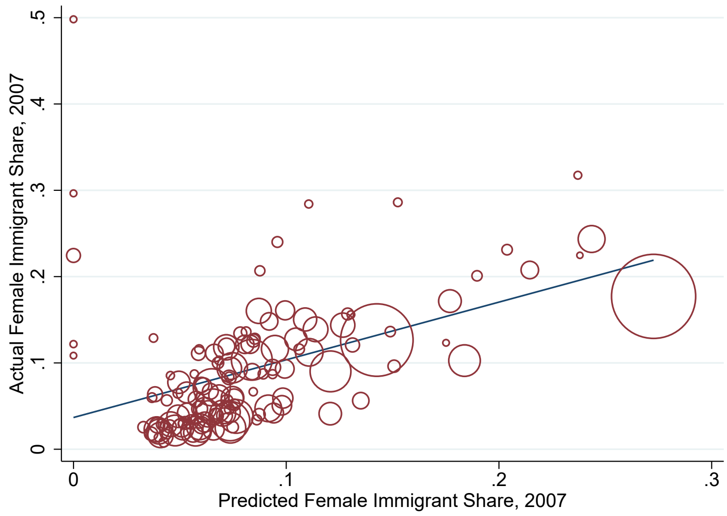
Furthermore, the sudden change in the composition of immigrant origins and the sharp increase during the migration boom strengthen the plausibility of the exogeneity assumption of the past-settlement instrument. By the late 1990s, the largest share of foreign-born citizens came from higher-income countries, particularly the UK and Germany, with the EU-28 accounting for nearly 42.4% while Latin America represented less than 15%. However, by the mid 2000s, most immigrants arrived from Latin America (mainly Ecuador, Colombia and Argentina) which accounted for more than 40% of the immigrant population. Another emigration region that became increasingly important was Central and Eastern Europe (CEE), mainly driven by Romania. By 2007 the country alone accounted for 9.7% of the foreign-born population, just two percentage points below Morocco's share of 11.8%. As highlighted by Moraga et al. (2019) and in line with Jaeger et al. (2018), in view of the sharp change in magnitude and composition of immigrants to Spain, one could expect, if any, only a negligible lagged impact of previous inflows from new dominant origin groups.

The first stage results are provided in Appendix A.2, Tables A.2 and A.3 which show

[19]The Local Registry microdata is only available from 1998 onwards. Unfortunately, the 1991 Census with municipality information only allows to identify country of birth based on 17 aggregated country-of-origin groups, it is therefore not possible to distinguish the key origin countries that provide domestic services individually. Another advantage of using the Local Registry is that it contains 100% of the population registered, unlike the restricted 10% sample from the 1991 census.

that the shift-share IV also satisfies the relevance condition based on F-Statistic values. Complementary to this evidence, Figure 6 shows that there is a strong correlation between actual and predicted female immigrant shares.

Figure 6: First Stage: Actual vs. Predicted share of foreign-born women



Notes: Each circle represents one urban or rural area. The size is proportional to the population of each local labour market. The shock and predicted immigrant shares are obtained from the first stage regression that considers only female inflows from the countries of origin specialised in domestic services provision.

5 Main Results

I start by interpreting the estimates obtained from estimating equation (4) above. Table 1 shows the average effect of immigration on *parenthood earnings penalty* comparing the main IV results with those from an OLS regression on the actual female immigrant share. I focus on parents that had their first kid from 1 to 5 years before time t or equivalently, whose first kid is between 1 and 5 years old. I exclude the year that the kid was born (year zero), when most mothers are likely to be in parental leave.[20] The marginal treatment effects are readily available in the results tables: for each table the row *Mothers* correspond to the estimated coefficient $\hat{\beta}_1$, *Fathers* to $\hat{\beta}_1 + \hat{\beta}_2$ and the *Gap* represents the estimated interaction term $\hat{\beta}_2$ which is the difference between fathers' and mothers' parenthood penalty. Column (1) provides OLS estimates, column (2) the preferred IV specification, column (3) varies in the fixed effects included, (4) removes the time effects since childbirth control and (5) includes all fixed-effects, individual controls and adds additional regional

[20] At the time, mothers in Spain had 16 weeks of job protected leave with 6 weeks compulsory and exclusively reserved for them and up to 10 that could be shared with the father since 1999. Law 3/1989 and Law 39/1999. I later show that this restriction does not alter substantially the results.

controls. To account for differential gender trajectories, column (6) adds full interactions of gender fixed effects with the controls. The table displays clustered robust standard errors at the area level for all coefficients. The results suggest that OLS estimates for both men and women would potentially be downward biased. According to the IV estimates in column 3, a 10 percentage point increase in the share of female immigrants potentially specialised in domestic services in a local labour market, allowed to close the gap in earnings loss associated with parenthood by an average of 5.5 percentage points. Scaling by the interquartile range of immigration exposure– measured as the change in immigrant share between 1999 and 2007, the estimates imply that the gap in a local labour market at the 75th percentile of immigration exposure was reduced by about 4.4 percentage points (0.55x8.8) compared to a region at the 25th percentile of exposure, or nearly 29% of the average gap of the sample.

Table 1: Impact of female immigration on PEP: OLS and IV

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
Mothers	0.249	0.472***	0.553***	0.468***	0.513***	0.514***
	(0.148)	(0.127)	(0.107)	(0.130)	(0.132)	(0.130)
Fathers	-0.161	-0.078	-0.027	-0.080	-0.016	-0.0399
	(0.102)	(0.092)	(0.055)	(0.091)	(0.095)	(0.0944)
Gap: Fathers - Mothers	-0.410**	-0.551***	-0.580***	-0.548***	-0.530***	-0.553***
	(0.149)	(0.134)	(0.103)	(0.135)	(0.134)	(0.134)
Gap: Wild Bootsrap p-value	*0.090*	*0.013*	*0.033*	*0.013*	*0.015*	*0.012*
N	150,816	150,816	150,816	150,816	150,816	150,816
Areas	126	126	126	126	126	126
$\bar{R}^2$	0.108	0.0653	0.0563	0.0653	0.0673	0.0770
K-P		99.46	121.5	101.9	99.73	99.93
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓		✓	✓	✓
Time since childbirth	✓	✓	✓		✓	✓
Regional Controls					✓	✓
Gender x Controls						✓

Notes: The results correspond to pooled regressions following the specification of equation 4. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. All regressions control for age, age squared, education and predetermined characteristics before childbirth including groups of experience and 2-digit industry fixed effects. Column (5) includes predetermined controls (1999) for unemployment, share of agriculture, services, construction and manufacture employment at the province level. Column (6) adds full-interactions between gender and the controls. All regressions are weighted by a linear function of children's age (More details on weighting are available in section 6). Robust standard errors in parentheses, clustered at the area level.
* p<0.05, ** p<0.01, *** p<0.001. In addition, Wild Bootstrap p-values clustered at the area level are reported for the gender gap estimated coefficient $\hat{\beta}_2$. These are calculated with 1000 repetitions.

The previous estimated effect is largely driven by a sharp fall in the parenthood penalties for mothers. Specifically, a 10 percentage point increase in the share of immigrants raised the average earnings for mothers by 4.7%. Recalling that my outcome variable measures their current earnings compared to their *individual* counterfactual in the absence of kids, the positive effect for mothers suggests that the arrival of foreign-born caregivers led to an increase in their PEP, therefore reducing the motherhood penalty. In contrast, the effects for men are much smaller and statistically non-significant. The results are similar if I remove the area fixed-effects (3) or the time since childbirth control (4). They practically remain unchanged if I control for pre-existent regional characteristics (i.e., in the pre-shock period) (5) and if I allow for gender-specific variation in the controls (6). For simplicity, I use the specification in column (2) as baseline in the remainder of the paper.

5.1 Pre-trends and dynamic effects

To further address potential endogeneity concerns, I analyse whether the results respond to the presence of pre-existing trends in the labour market. In particular, I consider a slight modification of regression (4) where I keep the migration shock constant over time at its 2007 level and estimate the following equation for each year $t \in \{1994, 2008\}$.

$$PEP_{it}^k = \alpha^k + \beta_{1t}^k \widehat{M^f}_{r(i99)} + \beta_{2t}^k \widehat{M^f}_{r(i99)} * g_i + g_i + \theta^k X_{i,\text{pre-birth}} + e_{it}, \tag{6}$$

where

$$M^f{}_{r(i99)} = \frac{I^f_{r,2007}}{(I+N)^f_{r,1999}},$$

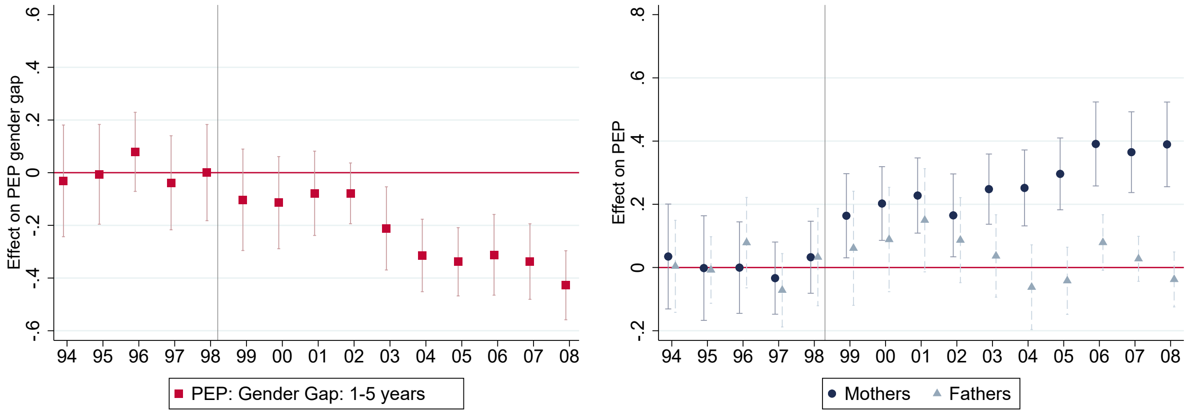
is instrumented with the shift-share IV described before. While the measure for migration inflows $M^f_{r,(i999)}$ remains invariant at its 2007 values for all regressions, the left-hand side varies to capture the change in parenthood costs over time. I expand the time horizon backwards and forward and estimate this equation for the years 1994 to 1998 (pre-shock), and all the years of the boom period until 2008 (post-shock). This specification allows for the identification of pre-existing trends by visual inspection serving as a placebo test for causality claims. Accordingly, one would expect β_{1t} and β_{2t} to be close to zero during the pre-shock period and statistically significant afterwards. To strengthen this pre-trends analysis, I complement its results with an event-study specification in Appendix E and a long-difference regional approach (Appendix C).

Figure 7a illustrates the impact of immigration on the gender earnings gap associated with child penalties. Specifically, it plots the difference in the effects between men and women on the PEP ($\hat{\beta_{2t}}$).[21] As can be seen, the effects are insignificant in the pre-shock period while they become negative and statistically significant after the mid-2000s.

[21] Recall that β_{2t} recovers the differential effect of men relative to women. Given initial gaps on the parenthood earnings penalties in favour of men, a negative coefficient implies a reduction in the differential costs of parenthood or equivalently, a gap closure of magnitude $\hat{\beta_{2t}}$

Figure 7: Effect of female immigration on parenthood penalties

(a) Gender gaps (b) Impact by sex



Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

When analysing the timing of the effects it is important to consider that the immigration boom can be broken down into two stages: (i) the early 2000s when the largest inflows arrived from Ecuador before visa restrictions were implemented in 2003, and (ii) the mid-late 2000s when two important EU enlargements took place fuelling further the arrival of immigrants from Eastern Europe.[22] Therefore, in line with the previous evidence, the stronger effect towards the end of the decade can be rationalised by the fact that there was a greater accumulated supply of domestic workers and that parents had been exposed to the new market conditions for a longer period.

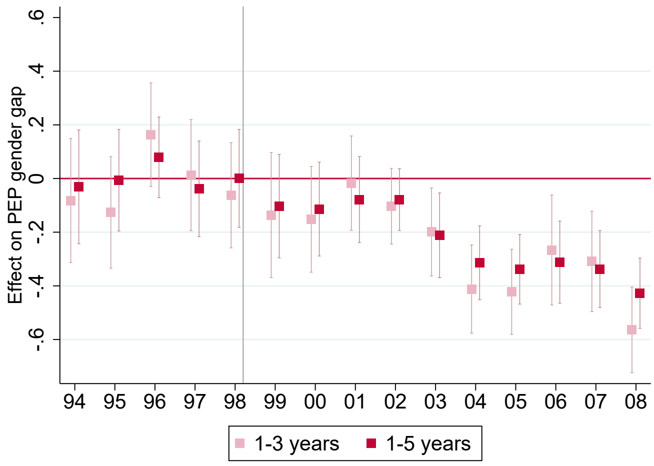
Figure 7b in turn presents the average effect on parenthood earnings penalty (PEP) by sex. Consistent with previous results, the effects are driven by a decrease in the penalties for mothers implied by the positive response on their PEP. In contrast to women, the impact of immigration on the PEP for men remains close to zero over time. It is worth bearing in mind that men generally experience smaller child penalties compared to women (see Appendix G). In addition, there is a null effect for both men and women between 1994-1998, suggesting that the effect is not recovering pre-existing trends in local labour markets for parents. Overall, these findings indicate that the immigration-driven shock in the supply of household services may have contributed to mitigate the gender gap resulting from asymmetries in the costs of parenthood. Moreover, the results hold when I implement a more conventional event-study approach using 1998 as baseline year (see Appendix E).

Additionally, in Appendix F I analyse the short-time penalty effects. To do so, I repeat

[22]The EU enlargements took place in 2004 and 2007. Romania and Bulgaria joined the EU in the last one but the inflows from these countries were already growing at the beginning of the decade, as shown in the Appendix Figure H.4.

the exercise for workers that became parents over the past three years relative to time t. The impact of immigration follows a similar pattern as before, with a more pronounced effect for women taking place towards the end of the 2000s. As a result, there is a greater reduction in the gender gap during the initial years after childbirth, as illustrated in Figure 8.

Figure 8: Effect of immigration on parenthood penalties
Gender Gap in the short-term



Notes: Sample restricted to parents between 25 to 50 years old in a given year. Light markers show the effect for parents who had their first kid in the last 3 years, dark markers the effect for parents that had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

5.2 Who benefits from the shock? Heterogeneity Analysis

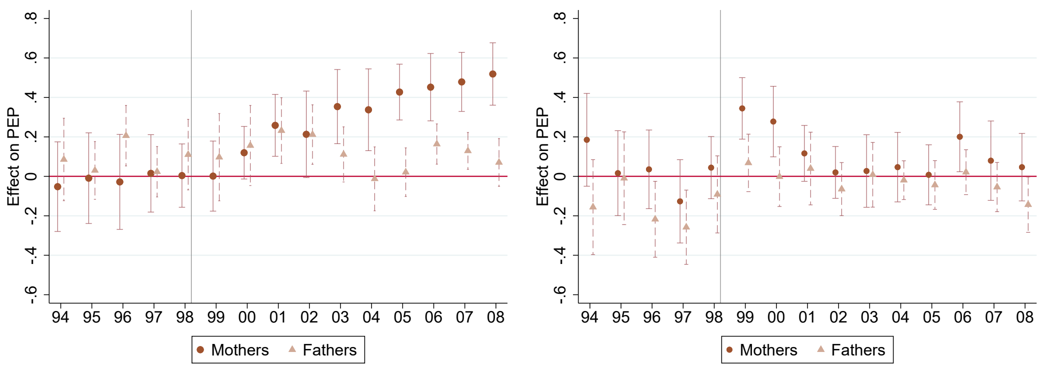
A larger provision of household services may affect low and high-skilled natives differently. Previous research has primarily focused on how the influx of female immigrants impacts high-skilled native women (Cortés and Tessada, 2011; Cortés and Pan, 2019). However, the impact can vary depending on the severity of the shock and the extent to which it influences local domestic service prices. While intuitively high-skilled or high-income households are more likely to employ domestic workers, if these services were already available –albeit in short supply– in the local labour market before the shock, wealthier families might have already been able to afford them, experiencing only a minor response to the immigrant-induced expansion. In such a case, when facing a price reduction, one would expect a stronger response from those workers who are at the margin. Hence, it would be reasonable to find substantial effects for lower-skilled women or more generally for families with tighter budget constraints.[23]

[23] A reason to exclude low-skilled women from the analysis would be to avoid potential effects from direct labour market competition or complementarities: i.e., low-skilled native women being promoted to *better* jobs as immigrants occupy positions at the bottom of the distribution. I provide several robustness

Figure 9 examines the impact of female immigration by educational level of natives: namely, I compare the effect on parents with secondary education or less (Figure 9a) with those that have a college degree or some tertiary education (Figure 9b). The analysis reveals that relatively lower-skilled mothers benefit the most from the shock, compared to their higher-skilled counterparts. While women with higher educational levels experience a reduction in their penalties at the beginning of the immigration boom, the effect is more persistent for those with up to secondary education. As pointed out earlier, this finding suggests that the shock had a more significant and enduring impact on individuals who were relatively more resource-constrained, and therefore for whom facing a reduction in household prices might have had more meaningful consequences. The effect was likely to be intensified as the inflows of foreign-born women kept on rising.

Figure 9: Effect of immigration by education
First kid in the last 5 years

(a) Low (b) High



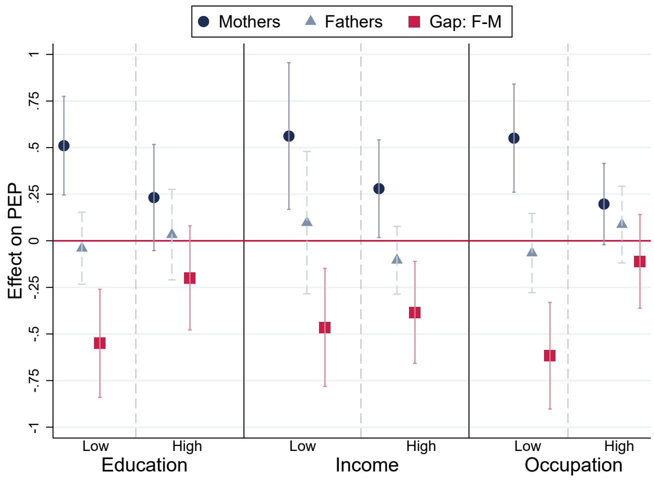
Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. *Low education* refers to workers with up to completed secondary education, *high education* denotes college education. Robust standard errors clustered at the area level. 95% confidence intervals.

I complement the previous evidence by looking at the effects by income and occupational groups. Regarding the former, workers are split into those who are below and above the median income, based on the earnings distribution by gender on the year before childbirth. Considering that some women might opt out of the labour force during pregnancy, I consider the worker's highest annual income of the last three years before the first child is born. In terms of occupation, the MCVL provides information for ten social security contribution categories associated to the skills required to perform a job. In particular, I classify the first five categories as high-skilled: 1. Engineers, college graduates and senior managers 2. Technical engineers and graduate assistants 3. Administrative and technical managers 4. Non-graduate assistants 5. Administrative officers; and the last five categories as low-skilled: 6. Subordinates. 7. Administrative assistants 8. First and second class officers; 9. Third class officers and technicians. 10. Labourers. For each proxy

checks to alleviate these concerns.

of skill definition (education, income, occupation) I estimate equation 4 including skill interaction terms and its correspondent linear and interacted fixed-effects. The results are presented in Table A.6 and summarised in Figure 10. As shown in the Figure, both high and low-skilled women experience a positive response in earnings to the migration shock, with the latter group exhibiting a larger effect. When zooming into more detailed education groups, it becomes evident that the effect is primarily driven by women who have completed secondary education. There is no significant effect for mothers with an education level below this, and a smaller positive effect for those with a college degree. (Appendix Table A.7).

Figure 10: Heterogeneity: Effect of immigration on parenthood penalties by skill group



Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

5.3 Mechanisms

The reduction in motherhood penalties can be driven by two main channels: labour supply and job quality. If access to domestic services allows mothers to return to the labour force sooner after giving birth one could expect the response to be reflected at the extensive margin. Likewise, assuming that these services effectively allow for a better work-family balance, workers might also respond at the intensive margin. Moreover, women might opt for steeper career trajectories. Fewer time constraints could promote career advancement prospects as mothers might become more inclined to seize opportunities for professional growth rather than settling for jobs that are more accommodating to caregiving responsibilities. In other words, with a reduction of work-family conflict, mothers could select more demanding jobs that align better with their competencies and preferences, resulting in improved job matching. This channel would be reinforced if the reduction in career-interruptions diminished the risk of skill depreciation which could lower future earnings potential.

Labour supply

I test the first channel by evaluating the effect of the immigration shock on women's labour supply. To do so, I analyse the probability of being in a part-time contract, being employed and number of days workers per month. For the first two outcomes, I estimate a linear probability model changing the dependent variable of equation 4 to these labour supply outcomes. Table 2 provides the results for all women (Panel A) and for each skill group based on educational attainment (Panel B). Moreover, I look at the effects for those who had a kid in the last 5 and 3 years separately. As shown in columns (1) and (2), the effect on the probability of being in a part-time job after childbirth is small and, albeit negative, not statistically different from zero.

Table 2: Effect of immigration on mothers' labour supply

		Part-time contract		Employment		Days (month)	
		1-5 years	1-3 years	1-5 years	1-3 years	1-5 years	1-3 years
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All		-0.032	-0.051	0.141	0.215	6.336*	7.912*
		(0.086)	(0.103)	(0.106)	(0.114)	(3.047)	(3.291)
Panel B. Education							
Secondary		0.042	-0.003	0.183	0.284*	6.795	10.68*
		(0.121)	(0.152)	(0.130)	(0.142)	(3.652)	(4.373)
Tertiary		-0.005	0.037	-0.013	-0.014	0.730	0.164
		(0.096)	(0.120)	(0.083)	(0.094)	(2.864)	(3.306)
	N	123,556	62,847	150,816	79,367	150,816	79,367
A.	$\bar{R}^2$	0.023	0.022	0.038	0.041	0.0539	0.0563
	K-P	91.89	95.36	99.46	103.5	99.46	103.5
B.	$\bar{R}^2$	0.038	0.039	0.031	0.034	0.0555	0.0580
	K-P	40.29	41.65	43.73	44.66	40.56	41.91

Notes: The results correspond to pooled regressions following the specification of equation4. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (columns 1 and 3) and who had their first kid in the last 3 years (columns 2 and 4). Year of birth (age 0) is omitted. The dependent variable in columns (1) (2) is a dummy variable for part-time contract, observations before 1996 are omitted since type of contract is not recorded in the MCVL with accuracy. The dependent variable in columns (3) (4) is a dummy variable for employment and is estimated with information from 1994 onwards. Regressions in panel A control for age, education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Area, year sex-area and sex-year fixed effects are included. Regressions in panel B also include education-sex-year and double interaction fixed effects. Robust standard errors in parentheses, clustered at the area level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

At the extensive margin, there is a positive effect on the probability of employment which is driven mostly by lower-skilled mothers. While the effect is not statistically significant after 5 years since childbirth (column 3), it is larger and statistically significant during the first years. Specifically, a 10 percentage point increase in the share of foreign female workers is associated with an increase in the employment probability of 2.1 percentage points after 3 years since childbirth, being significant at the 10% level. The effect increases to 2.8 percentage points for women with up to secondary education. After 5 years since childbirth, the effect is about two-thirds as large but not statistically significant. Unfortunately, the social security records do not contain information on hours of work which limits the analysis of responses at the intensive margin. However, given that contracts are recorded with high accuracy (exact entry and exit day) it is possible to compute days of work. I find that a 10 percentage point increase in the female immigrant share is associated with an extra 0.6 working days per month, about 2.3% of the average.

There are two main takeaways from these results. First, as expected, female labour supply responds more during the first years after childbirth which explains why having accessible childcare and domestic services at that time has the strongest impact. The results are also consistent with the fact that after turning 3, children in Spain could access public childcare (Nollenberger and Rodríguez-Planas, 2015). Second, having a smaller and not significant effect on employment after five years suggests that the reduction in child penalties is not solely driven by changes in the labour supply at the extensive margin, as it has been emphasized in previous research. My findings show that it might also be acting through the intensive margin as well as through the type of jobs that women can retain after childbirth. The latter might be a consequence of enhanced job opportunities during the initial years of motherhood, underscoring the importance of policy interventions that target early provision of household services to families.

The labour supply responses differ to some extent from the ones reported by Farré et al. (2011), who only focus on this margin. These authors find an effect of similar magnitude: namely, a 10 percentage point increase in female migrant share associated with a 2.2 point rise in employment rate. However, the effect corresponds to women with college education. Two main reasons could explain the different findings. First, the empirical design. I use a difference-in-differences type of strategy with treatment intensity, considering also the outcomes of the previous years as reference period (see equation 4). The authors instead use information from 1998 to 2008, practically focusing on the post-shock years only. By including outcomes from the pre-shock period, my aim is to better disentangle general regional characteristics captured by area fixed-effects from the impact of the migration shock. Restricting the analysis to treatment years only might increase the risk of confounding both effects.

Second, and perhaps more important, the samples are different. Since I am interested in analysing child penalties, I focus on parents who had their *first* child in the last three to five years rather than in all women.[24] Later, in Section 6.3 I compare mothers to non-mothers using a matching difference-in-differences design. Thus, when exploring the effects for all women, I build a comparable control. Moreover, I look at a sample of workers with a certain degree of labour force attachment: first, because they had a relationship

[24] Farré et al. (2011) instead look at the effect for women that have *at least* one kid 8 years old or younger, or who are living with an elderly dependent. Therefore, not only the families of interest differ since they potentially consdier families with many kids, but they also consider other types of responsibilities which are beyond the scope of my study. Additionally, these authors compare the labour supply effect to that of families without responsibilities (they include all workers and identify those with responsibilities with an indicator variable).

with the social security which is a requisite to be in the MCVL dataset, and second because my sample is restricted to those workers that had at least one working spell before the shock started, and who remained in the labour force at least 4 years after the shock to make workers pre and post-shock more comparable. Finally, their study is based on cross-sectional LFS data whereas I exploit matched employer-employee longitudinal records. The LFS not only lacks earnings information but may also include workers with little or no labour force attachment. Moreover, it only includes geographic information at the province level. In contrast, exploiting social security records I obtain a panel dataset which allows me: (i) to explore a finer definition of local labour markets based on urban (and rural) areas of municipalities' clusters; and (ii) to control exhaustively for predetermined (pre-birth) workers' characteristics given the panel dimension.

Types of jobs

In a compensating differentials framework, workplace flexibility is likely to be a highly valued amenity for mothers. With improved access to external domestic services, mothers may be able to overcome the *costly flexibility* inherent to the lower-paying jobs where women traditionally sort into after childbirth, relative to men's jobs.[25] To evaluate this channel, I start by looking at the effect the of immigrant-driven domestic services' shock on wages. Exploiting the detailed information on exact days worked, I am able to get a proxy of wages by calculating average daily earnings.[26] Table 3 summarises the results of specification 4 using the log of daily earnings as an outcome and focusing on the coefficients for mothers. The estimations suggest a positive response of native mothers' daily wages to the larger availability of domestic services, concentrated in less-skilled women. The first two columns provide the results for women who became mothers in the last 5 and 3 years, respectively. Columns (3) and (4) control additionally for workers' earnings one and two years before childbirth. According to these results, a 10 percentage point increase in the share of female immigrants specialised in domestic services increased women's daily wages by around 2% to 3% during the first years of motherhood. The effect is positive and larger for lower-skilled women. The higher-skilled mothers also display a positive response, although not statistically significant. The point estimates suggest an effect of around 1.3% during the first three years once I control for pre-birth earnings. In fact, these controls allow me to rule out a response of fathers which is close to zero when previous earnings are accounted for (Appendix Table A.8 for all results).

I complement this evidence by looking at the average wage of the jobs where parents are working. Considering that distinct occupations may be remunerated differently across industries, I calculate the leave-one-out mean of log daily earnings w by 2-digit-sector (s) x occupation (o) x gender (g) and match these averages to workers' based on their

[25] See Goldin (2014).

[26] Unfortunately the social security records do not report information on hourly wages.

Table 3: Effect of immigration on mothers' daily earnings

	ln(Daily earnings)			
	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)
Panel A. All	0.357*	0.460**	0.204*	0.295**
	(0.137)	(0.149)	(0.100)	(0.112)
Panel B. Education				
Secondary	0.471**	0.590***	0.236*	0.344**
	(0.153)	(0.159)	(0.111)	(0.119)
Tertiary	0.199	0.281	0.160	0.228
	(0.176)	(0.194)	(0.132)	(0.145)
N	135,328	70,708	131,811	70,578
A. $\bar{R}^2$	0.193	0.192	0.510	0.531
K-P	103.3	107.9	102.6	107.9
B. $\bar{R}^2$	0.193	0.192	0.509	0.531
K-P	42.87	43.48	42.47	43.53
$\text{Earnings}_{c-1,c-2}$	No	No	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation4. The table only shows the coefficients for women, the coefficients for men are available in appendix (Table A.8) The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. The dependent variable is the logarithm of average daily earnings. Regressions control for age, education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Sex-area and sex-year fixed effects are included. Regressions in panel B also include education-sex-year and double interaction education-year and education-area fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

current job characteristics. $\bar{w}_{\text{-i,osg}} = \frac{1}{n-1}\sum_{j\neq i} w_{\text{osg},j}$ Using these variables as outcomes in the baseline equation (4), I find that mothers living in regions that received larger inflows of domestic workers are more likely to be in higher-paying jobs when classified by occupation-sector remuneration, with a small impact on fathers' types of jobs (Appendix Table 4, Columns (1)-(2)). Similarly, I calculate the average wages at the firm f where the worker is employed: $\bar{w}_{\text{-i,f}}$. Since the MCVL is a 4% sample of social security records, small firms can often be represented by only one or very few workers in the dataset, I avoid these cases by considering firms with at least ten different workers. Columns (3)-(4) show suggestive evidence of mothers in more exposed regions sorting into higher-paying firms compared to men, however the effects are not statistically significant. In part, the lack of precision could be due to the reduced sample size, resulting from the restriction on the number of workers by firm described before. In light of this evidence, it is also important to note that a job-quality channel may take more time to materialise compared to a labour supply response.

Lastly, I evaluate the impact of the immigration-induced shock in domestic services on job-mobility for mothers. To this effect, I construct a finer version of job quality proxied by occupation-sector fixed effects using a movers design, similar to Abowd and Kramarz (1999). Initially proposed to estimate firm-specific pay premiums, the two-way fixed effects AKM estimator relies on the assumption that there is sufficient within-worker

Table 4: Effect of female immigration on jobs: industry, occupation and firm wages

	Occupation x Sector		Firm	
	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)
Mothers	0.254**	0.257**	0.103	0.079
	(0.083)	(0.088)	(0.096)	(0.095)
Fathers	0.110	0.091	0.010	-0.034
	(0.056)	(0.061)	(0.065)	(0.076)
Gap: Fathers-Mothers	-0.145	-0.165	-0.093	-0.113
	(0.109)	(0.114)	(0.108)	(0.119)
N	137,047	71,584	50,876	26,375
$\bar{R}^2$	0.350	0.347	0.444	0.452
KP	99.84	104.0	148.9	150.2
Time-Sex FE	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation4. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. The dependent variable in columns (1)-(2) is the leave-one-out average daily earnings at the 2-digit sector x occupation cell of the workers' job. The dependent variable in columns (1)-(2) is the leave-one-out average daily earnings at the 2-digit sector x occupation cell by sex of workers' job. Cells must have at least 10 different workers. The dependent variable in columns (3)-(4) is the leave-one-out average daily earnings of the firm, restricting to those firms with at least 10 different workers. Regressions control for age, education and pre-determined characteristics before childbirth including cells of experience. Columns (3)-(4) also control 2-digit industry fixed effects. Sex-area and sex-year fixed effects are included. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

variation in terms of the firm they work for. In other words, workers have switched firms or workplaces over time, providing variation that can be used to estimate worker and firm fixed effects. Since I only have limited information on the firm-side given the nature of my dataset –4% sample of workers–, [27] instead of looking at the firm component I estimate a model specifying an individual component and an occupation x sector fixed effect, using the latter as a proxy of job quality. I assume thus that there is sufficient mobility among these cells of jobs. To strengthen its plausibility, I extend the industry classification to the 3-digit-sector and interact it with workers' occupation group. Switching to the full sample of workers, I estimate the following model:

$$w_{it} = \alpha_i + \psi_{J(i,t)} + x'_{it}\beta + \epsilon_{it} \tag{7}$$

In my setting, the function $J(i,t)$ refers to the occupation-sector cell where worker i is employed in year t. It is assumed that the log daily real wage w_{it} of individual i in year t is the sum of an individual component α_i, which can be interpreted as a combination of skills and other attributes that are equally rewarded across jobs, an occupation-sector component $\psi_{J(i,t)}$, a set of time-varying observable characteristics $x'_{it}\beta$, and an error component ϵ_{it}. The vector x_{it} captures life cycle trends by including experience and experience

[27] Given the identification assumption, empirical researchers tend to estimate these models using data for the whole population from matched employer-employee records.

squared which can be obtained with very high precision in my dataset since I observe the exact date of start and end of each contract, interactions of age with educational attainment, age squared, along with a full set of year fixed effects.

After estimating the model, I group the estimated occupation-sector fixed effects in percentiles (p) to construct the variable OS_p and define the following outcomes:

- $Job_{\text{upgrade}} = \mathbf{1}\{OS_{pi,t} > OS_{pi,(\text{pre-birth})}\}$
- $Job_{\text{same}} = \mathbf{1}\{OS_{pi,t} = OS_{pi,(\text{pre-birth})}\}$
- $Job_{\text{downgrade}} = \mathbf{1}\{OS_{pi,t} < OS_{pi,(\text{pre-birth})}\}$

where $Job_{upgrade}$ takes a value of one if the occupation-sector job, OS for worker i at time t, is at a higher percentile than the one of her job one year before childbirth. The other two status: *same* and *downgrade*, are defined analogously. Using these outcomes, I estimate the specification of equation 4, once again restricting the sample to parents of children between 1 and 5 years old in a given year.

As shown in the Appendix Table A.10, the increased availability of domestic services in the local labor market raised the probability that mothers found a better-quality job by 1.7% for mothers of children up to 5 years-old (measured at the interquartile range of exposure, significant at the 10% level). Additionally, it slightly reduced the probability of staying in the same job percentile, with negligible effects on the probability of switching to a worse-quality job. It is interesting to notice that the effects are instead very small and closer to zero during the first 3 years after childbirth (even columns). These findings are intuitive: rather than an immediate upgrade in their career post-childbirth, access to household substitutes might allow mothers to *not downgrade* immediately after childbirth, but instead, maintain their pre-birth job status without large deviations in their ongoing trajectory, eventually getting into a better position. Taken together, these findings suggest that domestic services not only allowed mothers to reintegrate into the labor market but also to achieve better outcomes in the mid-run following childbirth.

6 Robustness and Extensions

6.1 Direct competition in the labour market

The heterogeneity analysis of section 5.2 showed that the reduction on child penalties was concentrated among relatively lower-skilled mothers. While this effect can easily be rationalised, one potential concern is whether the impact actually responds to domestic services consumption or if it is reflecting other types of labour market dynamics. For instance, it could be the case that low-skilled natives are being pushed into better occupations or jobs while immigrants occupy positions at the bottom of the earnings distribution (Amuedo-Dorantes and De la Rica, 2008; Peri and Sparber, 2009). If the effects corresponded to direct labour market competition, one would expect the impact to be concentrated among workers that were more exposed to immigration in the labour market.

Im principle, given that my design is based on *relative* outcomes these alternative adjustment mechanism should not be a matter of concern.[28] In any case, to completely rule out this channel, I perform two tests. First, I exclude non-qualified workers from the sample, who were more likely to be in direct competition with immigrants working in the domestic service sector. Second, considering that during the early 2000s the construction sector was booming and potentially attracting more immigrants to regions with higher dynamism -around one third of immigrants worked in this sector- I exclude the regions that were more exposed to these shocks, namely, I compute the total employment growth in the sector between 1999 and 2007 and exclude those regions at the top 10 percentile of growth rates. As shown in Table 5 the results are robust to these sample modifications.

Table 5: Effect on parenthood penalties: Excluding workers more exposed to overall immigration

	PEP					
	All		No Construction Boom		No Labourers	
	1-5 years	1-3 years	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)	(5)	(6)
Mothers	0.472***	0.549***	0.460***	0.544***	0.461***	0.510***
	(0.127)	(0.141)	(0.130)	(0.143)	(0.127)	(0.138)
Fathers	-0.078	-0.0004	-0.023	0.058	0.0094	0.092
	(0.092)	(0.102)	(0.090)	(0.098)	(0.091)	(0.107)
Gap	-0.551***	-0.549***	-0.482***	-0.486**	-0.451***	-0.418**
	(0.134)	(0.147)	(0.133)	(0.145)	(0.125)	(0.137)
N	150,816	79,367	146,263	76,990	137,259	72,230
$\bar{R}^2$	0.0433	0.0445	0.0438	0.0451	0.0399	0.0407
K-P	99.46	103.5	105.2	108.7	102.4	106.3

Notes: The results correspond to pooled regressions following the specification of equation 4. The dependent variable is the parenthood earnings penalty PEP. A positive effect on women/men implies a reduction in child penalties. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. Columns (3) and (4) exclude workers living in regions more exposed to the construction boom. Columns (5) and (6) exclude labourers i.e. non-qualified workers according to the social security occupational classification. All regressions control for age and education and pre-determined characteristics before childbirth including groups of experience and 2-digit industry fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

[28] In specific, if the effects were capturing direct labour market competition, one could expect an impact on natives regardless of parenthood status. Under this assumption, when constructing a relative outcome (actual earnings with respect to a counterfactual in the absence of kids) these competition effects would partial out.

6.2 Sample Restrictions and Weighting

Table 6 show that the main results are robust to different sample specifications. During the first year and especially the first months after childbirth, women are more likely to stop working as opposed to men. To make sure that this behaviour was not driving the results, the previous analysis was provided for individual's situation after the first year of childbirth, specifically when the first kid was between 1 and 5 years old. Table 6 shows that the key results are robust to whether the year when the first child is born (year 0) is included or not (column 2).

Table 6: Robustness: Impact of female immigration on parenthood penalties

	PEP					
	(1)	(2)	(3)	(4)	(5)	(6)
Mothers	0.472***	0.514***	0.412***	0.438***	0.314**	0.412**
	(0.130)	(0.118)	(0.117)	(0.128)	(0.110)	(0.126)
Fathers	-0.078	-0.0676	-0.127	-0.150	-0.120	-0.036
	(0.092)	(0.0861)	(0.0859)	(0.0859)	(0.112)	(0.092)
Gap: Fathers - Mothers	-0.551***	-0.582***	-0.539***	-0.588***	-0.433**	-0.448**
	(0.134)	(0.128)	(0.125)	(0.153)	(0.133)	(0.128)
N	150,816	187,575	150,816	150,816	199,972	146,503
KP	99.5	106.6	96.43	101.9	0.0449	0.0963
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓	✓
Weight	✓	✓	No	✓	✓	✓
Age FK	1-5	0-5	1-5	1-5	1-5	1-5
Always SS	✓	✓	✓	No	✓	✓
Obs $<$ 1999	✓	✓	✓	✓	No	✓
Earnings$_{c-1,c-2}$	No	No	No	No	No	✓

Notes: The results correspond to pooled regressions following the specification of equation 4. The dependent variable is the parenthood earnings penalty. A positive effect on women/men implies a reduction in child penalties. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. Time SCB refers to the time since childbirth. Always SS refers to the sample based only on observed data in social security records. Column (4) imputes zero earnings for non-employment spells i.e. when an individual is not observed in the MCVL. Earnings$_{c-1,c-2}$ refers to controls for workers' earnings one and two years before childbirth. All regressions control for age and education and pre-determined characteristics before childbirth including experience and 2-digit industry fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

In terms of sample restrictions, the main findings are also robust to the inclusion of non-observed spells in the social security (*Always SS*), which I code as non-employment spells with zero earnings (column 4). Additionally, in the benchmark specification (column 1), I restrict the sample to those individuals that entered the labour market before the migration boom started (*Obs. $<$ 1999*). Column (5) shows that the results hold if I remove this restriction. Next, column (6) shows that the effects remain invariant to controlling for earnings one and two years before childbirth. Further, adding these controls reduces the coefficient of men closer to zero, which nonetheless was non statistically significant in all the other specifications. Finally, I consider four functions that assign higher weights to earlier years after childbirth (Appendix Figure D.1). Table D.1 in the Appendix shows that the results hold for unweighted and different weighted specifications.

6.3 Alternative Counterfactual: Matching Design

The main advantage of my proposed penalty measure is that, by incorporating worker individual heterogeneity in the counterfactual, it simulates individual-specific trajectories in the absence of kids. However, it does so by imposing parametric assumptions on the evolution of earnings. As an alternative, in what follows I analyse the impact of immigration on child penalties by employing a difference-in-differences matching design. In particular, I use propensity score matching on a sample of parents and non-parents to define individual penalties. Note that, unlike my baseline measure, it does not require any functional form assumptions to construct counterfactual earnings. Once more, the results are qualitatively equivalent: an increase in the provision of female domestic workers reduces the child penalties by alleviating the cost for mothers.

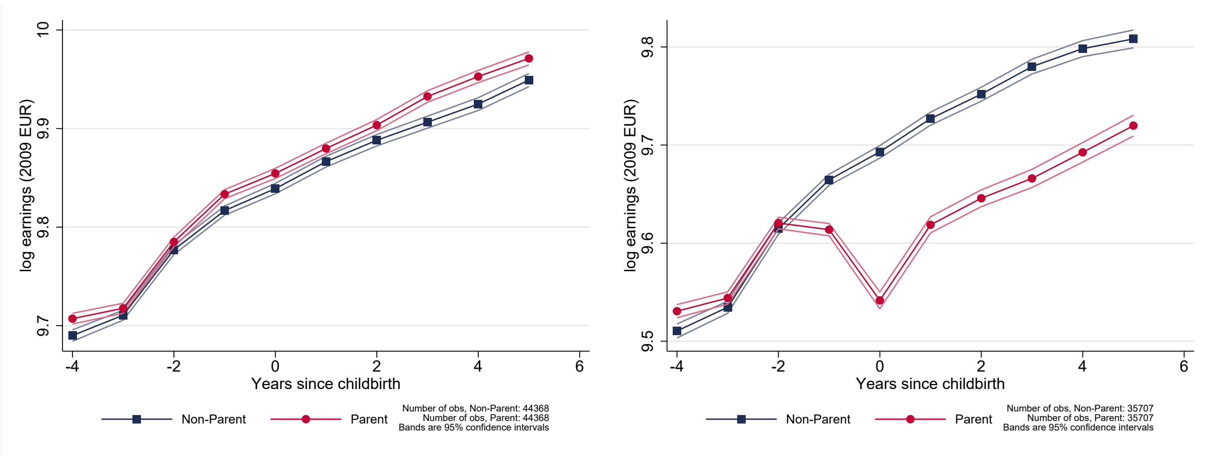
To construct the controls sample, I consider the year before childbirth $t = c - 1$ as the baseline year. I restrict this sample to individuals who are employed at baseline and aged between 19-49.[29] To build a control group for parents, I match exactly within cells of year, gender and autonomous community ending up with 510 cells (15x2x17). I then obtain the propensity score from a probit regression on the probability of becoming a parent for the first time on worker's log earnings at $t = c - 2$ and $t = c - 3$, and age, education and experience at baseline $t = c - 1$. Furthermore, to ensure that the individual counterfactuals remain valid over time, the pool of untreated workers (potential controls) include individuals who have not had a kid at year t and will not become parents within the next six years.

Figure 11 displays the average earnings for the sample of parents (treated) and matched controls (untreated). The left and right panel compare the trajectories for men and women, respectively. As shown in the figures, the matching procedure is able to capture large unconditional penalties for women with almost no deviations for men, consistent with the evidence on child penalties for Spain obtained using Kleven et al. (2019a)'s approach. (See Appendix G and de Quinto et al. (2021)). Additionally, Table A.11 in the Appendix shows that the sample is balanced in terms of pre-determined characteristics.

[29] While these are baseline characteristics to construct a matched control, during the years of parenthood I restrict to individuals who are aged between 25-50 so that the sample is comparable to the one I use for the main analysis.

Figure 11: Propensity Score Matching: Unconditional means of parents earnings

(a) Men (b) Women



Notes: The figures depict the unconditional means of the log of annual earnings for parents compared to the matched controls.

Using the new individual matched-controls I construct the following outcomes:

$$\Delta G^{PSM}_{i,c-1} = \frac{Y_{it} - Y_{i,c-1}}{Y_{i,c-1}} - \frac{Y^{MC}_{it} - Y^{MC}_{i,c-1}}{Y^{MC}_{i,c-1}}$$

$$\Delta G^{I}_{i,c-1} = \frac{Y_{it} - Y_{i,c-1}}{Y_{i,c-1}} - \frac{\widehat{Y_{it}}(0) - Y_{i,c-1}}{Y_{i,c-1}}$$

where ΔG^{PSM}_{c-1} denotes difference in growth earnings Y_{it} with respect to earnings one year before childbirth $(c-1)$ between the observed and matched controls (MC).[30] For ease of interpretation, I compare this outcome both to the original PEP and to the difference in growth between actual and imputed counterfactual ΔG^{I}_{c-1} using instead of the matched controls, the predicted counterfactual earnings from the imputation method: $\widehat{Y_{it}}(0)$.

I use a difference-in-differeces approach as stated in Equation 4 using the outcomes I defined above. I find that the responses go in the same direction using the two new measures as shown in Table 7: namely, a decline in child penalties driven by a positive response of women. Looking at differences in earnings growth from 1 to 3 years after child-birth, the decline in penalties using the matching sample (column 2) is about 72% of the one derived from the imputation counterfactual (column 4) with the growth in earnings for mothers from the first measure being equivalent to an 82% of the latter. Using this metric, fathers reveal considerably smaller but positive gains. Despite not being statistically significant, the point estimates have a very similar magnitude. It is noteworthy that the matching sample is a subset of the baseline one. Since workers are matched on earnings up to three years before the childbirth year, workers with higher labour force attachment are implicitly selected. Finally, taking advantage from the fact

[30] I do not define an earnings ratio as I did before given that matched controls can have non-employment spells over time. This is unlikely to happen when I use the imputation method to define the PEP.

that I have available a comparable sample of parents and observed non-parents, I am able to confirm that the positive response of women is indeed driven by a response of mothers, with no response to the shock of matched-women who have not had a child yet (Appendix Table A.12).

Table 7: Effect of immigration on child penalties: Matching difference-in-differences

	ΔG^{PSM}_{c-1}		ΔG^{I}_{c-1}		PEP	
	1-5 years	1-3 years	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)	(5)	(6)
Mothers	0.413*	0.532**	0.644***	0.630**	0.410***	0.404**
	(0.173)	(0.176)	(0.176)	(0.197)	(0.111)	(0.121)
Fathers	0.0609	0.166	0.0440	0.131	-0.0072	0.0608
	(0.0990)	(0.109)	(0.108)	(0.114)	(0.0741)	(0.0710)
Gap: Men-Women	-0.352*	-0.366*	-0.600**	-0.500*	-0.417***	-0.343**
	(0.165)	(0.185)	(0.194)	(0.219)	(0.119)	(0.129)
N	109,624	60,751	109,624	60,751	109624	60751
$\bar{R}^2$	0.0241	0.0210	0.167	0.177	0.140	0.153
KP	96.64	102.9	96.64	102.9	96.64	102.9
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓	✓
$\text{Earnings}_{c-1,c-2}$	✓	✓	✓	✓	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation4 using different dependent variables. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. Regressions control for age, education and pre-determined characteristics before childbirth including experience, 2-digit industry fixed effects and interactions of each control with sex. All regressions control for baseline earnings one and two years before childbirth. Sex-area and sex-year fixed effects are included. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

6.4 Alternative Approach: Following Workers Over Time

When analysing the previous results, another concern is whether the responses are to some extent driven by population compositional changes resulting from the immigration shock. While I construct the sample imposing labour force attachment restrictions prior to the shock i.e., individuals were in the social security before 1999 and remained in it for at least four years between 2000-2008, I complement the previous findings with an alternative method to assess labour market shocks based on a more traditional individual-level analysis (Autor et al., 2014; Yagan, 2019).

In this approach, instead of defining a sample for each year, I keep the sample of workers fixed over time, following the same individuals before and after the shock. The main advantage is that the method abstracts from effects driven by potential changes in workforce composition. The downside, however, is that with this alternative method I cannot assess the evolution of comparable penalties year by year because as time goes by, workers and children get older which also implies that the time since childbirth is changing.

I restrict the sample based on workers' characteristics as of 1999, at the outset of the immigration boom. I focus on family characteristics at this baseline year, particularly by identifying workers that had at least one kid 3 years old or younger at that time and those who did not. Moreover, I define income groups based on median earnings by gender using pre-determined average income (1995-1998). Similar to Autor et al. (2014), I define the following outcomes:

- Normalised annual earnings/days (by pre-shock average): $y_{it} = \frac{earnings_{it}}{\frac{1}{5}\sum_{95}^{98} earnings_{is}}$
- Normalised cumulative earnings/days (by pre-shock average): $y_{it} = \frac{\sum_{00}^{s} earnings_{it}}{\frac{1}{5}\sum_{95}^{98} earnings_{is}}$

Then, I estimate for each year a specification similar to equation 6 where the dependent variables are the previous outcomes and the *change* in migration shares is the main explanatory variable. I estimate the following regression separately for each year:

$$y_{it} = \alpha + \delta_{1t}\widehat{\Delta M^f}_{r(i99)}^{2007-1999} + \delta_{2t}\widehat{\Delta M^f}_{r(i99)}^{2007-1999} * g_i + \gamma_g + \theta X_{i,\text{pre-shock}} + e_{it}, \tag{8}$$

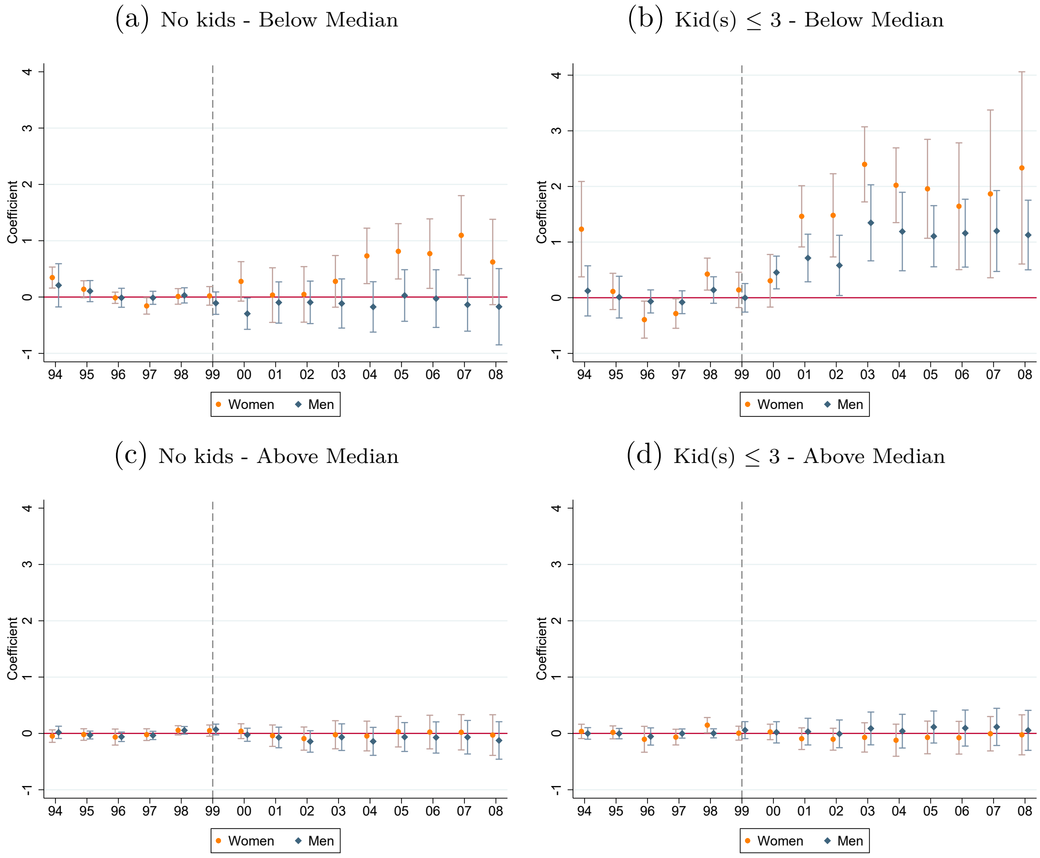
where the migration shock is defined as:

$$\Delta M^f{}_{r(i99)} = \frac{I^f_{r,2007} - I^f_{r,1999}}{(I+N)^f_{r,1999}},$$

and I denotes Immigrant workers, N Natives and f female workers.

Moreover, instead of controlling for characteristics before childbirth, I control for worker characteristics predetermined before the immigration shock including 2-digit level industry and experience as of 1999, average earnings 1995-1998 in addition to birth cohort fixed effects and local labour market population as of 1999. The results are illustrated in Figures 12 and Appendix Figure F.3. The estimates suggest that the female immigration shock generated a positive effect on earnings for women with young children when the boom started. As observed in Figure 12 the effect was considerably larger for mothers of young kids relative to their male counterparts as well as in comparison to other women who did not have young kids at the same time. Notice that this subsample also includes women who might have had kids earlier but were already older in 1999 or that became mothers afterwards, which might explain the positive impact. Moreover, the effect is concentrated in those families who were below the median earnings when the migration wave started, consistent with the previous findings.

Figure 12: Effect on annual earnings by income group
Fixed sample of workers



Notes: Fixed sample of workers restricted to individuals that were between 30 to 49 years old in 1999. Family characteristics are determined at the baseline 1999 year as well as income groups, based on the average earnings between 1995-1999 and their position in the gender-specific income distribution. Robust standard errors clustered at the area level. 95% confidence intervals.

6.5 Alternative Instrument

Taking advantage of the clear change in composition described in section 4.3, in this section I complement the benchmark regressions using an alternative design of the shift-share type of instrument. In particular, I estimate the following first stage equation inspired by Goldsmith-Pinkham et al. (2020), where I exploit the shares of each origin group

$$\tilde{Z}_{rt}^{f} = \alpha + \sum_{o} \delta_o \frac{I_{o,r,1998}^{f}}{I_{o,1998}^{f}} + \varepsilon. \tag{9}$$

The specification incorporates more variation by allowing different first-stage coefficients. However, the identification assumption is stronger. Discussing identification from Bartik instruments Goldsmith-Pinkham et al. (2020) show that using the traditional version of this instrument is equivalent to weighting local shares and use them as instruments, which can be interpreted as exogeneity stemming from these shares. Specifically, the authors show that "the 2SLS estimator with the Bartik instrument is equivalent to a GMM estimator with the local industry shares as instruments and a weight matrix constructed from the national growth rates."

The results from the alternative first-stage regression are provided in Appendix A.2. In terms of relevance, Table A.4 shows the results obtained for the instrument for women immigrants only while Table A.5 is constructed based on both male and female immigrants. Both tables report first stage F-statistics and for each country of origin separately. The preferred specification corresponds to column 4 which excludes two regional outliers based on the highest values of dfbetas from the first stage. The large F-statistic values of 27 for the women-only instrument and 41 for the women and men instrument imply that they also satisfy the relevance condition. As can be observed below, applying this instrument to equation 6 delivers similar results than the ones based on the traditional shift-share instrument used as benchmark.

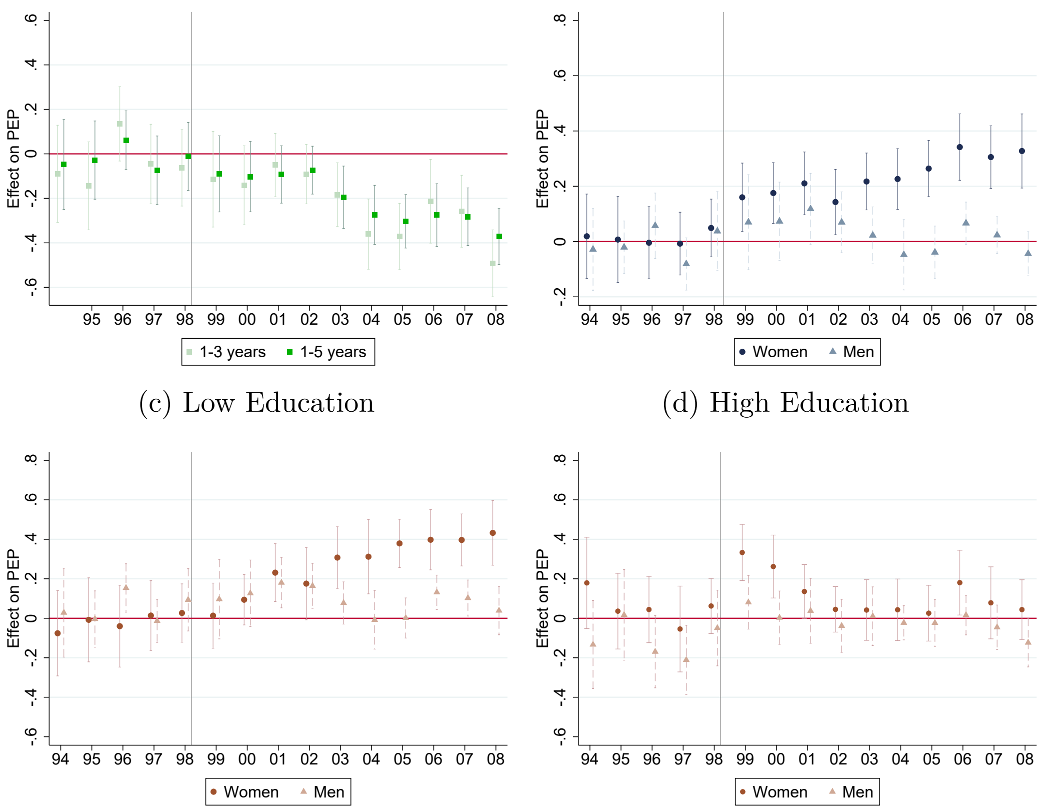
Figure 13: Effect of immigration on PEP
Alternative Instrument

(a) All workers: gender gaps

(b) All workers

(c) Low Education

(d) High Education



Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. *Low education* refers to workers with up to completed secondary education, *high education* denotes college education. Robust standard errors clustered at the area level. 95% confidence intervals. First stage estimated using equation 9.

7 Conclusion

Over the past century, one of the biggest advances experienced in developed countries has been the decline of gender gaps in education and labour force participation. Yet, substantial gaps in the labour market persist. These disparities may vary depending on the specific context under consideration, but the unequal impact of parenthood on men and women has consistently proved to be one of the main drivers. After a surge in the literature documenting this fact, this paper takes a step further by helping understand how to overcome the costs of parenthood. Exploiting the unprecedented supply shock in domestic and childcare services induced by the 2000s' Spanish migration boom, I investigate whether the expansion of more affordable household and childcare services can mitigate gender disparities and alleviate the penalties associated with parenthood in the labour market.

Using rich matched employer-employee longitudinal data, I first develop a novel measure of the so-called child penalty at the individual level. My metric relies on an imputation-based counterfactual of earnings in the absence of kids that takes into account heterogeneous income profiles potentially linked to workers' fertility preferences. Using this penalty as an outcome, I then combine a difference-in-differences strategy with treatment intensity and a shift-share instrument to estimate the causal effect of the domestic services expansion. Complementary, I conduct a set of placebo tests and pre-trends analysis in an event-study type of set up, which show that the shock was not correlated with outcomes before the migration boom started nor driven by pre-existing trends.

This study reveals that the increased availability of household services narrowed the gender gap associated to child penalties driven by a positive response on earnings from mothers, with almost no effect on fathers. Moreover, the impact is persistent and more pronounced for relatively lower-skilled native mothers. Thus, these results suggest that affordable substitutes for household production can help not only to alleviate gender gaps but also to reduce within-gender inequality. Two main mechanisms are driving these results: labour supply and job quality. I show that mothers living in more exposed areas were more likely to be employed after the first years since childbirth, working also more days per year. Furthermore, I find that the expansion of these services allowed mothers to work in higher-paying firms as well as better-remunerated jobs in terms of occupation and sectoral attributes. Finally, considering that my penalty measure is subject to parametric functional form assumptions, I apply a matching design using propensity scores to i) confirm that my findings are robust to alternative designs of the outcome and ii) to test explicitly whether the expansion of these services benefited parents and non-parents alike. I show that only mothers' earnings have a positive and significant response. Altogether, these findings reveal that accessible and affordable substitutes for household production not only allowed for a faster reintegration of mothers into the labour market but also placed them in better career trajectories in terms of the quality of job matches they had over time.

This paper provides evidence on how policies aiming at alleviating household-related time constraints could address gender inequality by allowing mothers to seize opportunities for professional growth, rather than settling for jobs that are more accommodating to caregiving responsibilities and other traditional roles. It opens however the debate to other considerations. First, my empirical findings showcase the potential benefits of low-skilled immigration in recipient countries, however, less is known about the other side of the coin. What would be a sustainable scheme that allows i) natives to afford these services and ii) that ensures adequate workers' protection legislation for potentially underpaid foreign workers who face significant barriers to integrating into the formal labour market. Targeted subsidies and special visa schemes might be options to carefully evaluate. Second, access to external domestic services can allow mothers to overcome the costly flexibility inherent to the less demanding jobs that women traditionally sort into after childbirth. However, it is important to acknowledge the key substitution mechanisms behind, where mothers' housework is replaced by another woman's work. On the one hand, there is a chance that these dynamics perpetuate traditional gender roles. On the other hand, mothers' increasing presence and success in the labour market could be an important first step to reassess these norms. I leave these questions for future venues of research and debate.

References

Abowd, J. M. and F. Kramarz (1999): "Econometric analyses of linked employer–employee data," *Labour economics*, 6, 53–74.

Altonji, J. G. and D. Card (1991): "The effects of immigration on the labor market outcomes of less-skilled natives," in *Immigration, trade, and the labor market*, University of Chicago Press, 201–234.

Amuedo-Dorantes, C. and S. De la Rica (2008): "Does immigration raise natives income? National and regional evidence from Spain," .

Amuedo-Dorantes, C. and S. De La Rica (2011): "Complements or substitutes? Task specialization by gender and nativity in Spain," *Labour Economics*, 18, 697–707.

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.

Autor, D. H., D. Dorn, G. H. Hanson, and J. Song (2014): "Trade adjustment: Worker-level evidence," *The Quarterly Journal of Economics*, 129, 1799–1860.

Baker, M., J. Gruber, and K. Milligan (2008): "Universal child care, maternal labor supply, and family well-being," *Journal of political Economy*, 116, 709–745.

Barone, G. and S. Mocetti (2011): "With a little help from abroad: the effect of low-skilled immigration on the female labour supply," *Labour Economics*, 18, 664–675.

Bensnes, S., I. Huitfeldt, and E. Leuven (2023): "Reconciling Estimates of the Long-Term Earnings Effect of Fertility," Tech. rep., Institute of Labor Economics (IZA).

Bentolila, S., J. J. Dolado, and J. F. Jimeno (2008): "Does immigration affect the Phillips curve? Some evidence for Spain," *European economic review*, 52, 1398–1423.

Berlinski, S. and S. Galiani (2007): "The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment," *Labour Economics*, 14, 665–680.

Bernat, M. J.-S. and M. R. Viruela (2011): "The economic crisis and immigration: Romanian citizens in the ceramic tile district of Castelló (Spain)," *Journal of Urban and Regional Analysis, 3(1), 45e65.*

Berthelon, M., D. Kruger, C. Lauer, L. Tiberti, and C. Zamora (2023a): "Longer School Schedules, Childcare and the Quality of Mothers' Employment," .

Berthelon, M., D. Kruger, and M. Oyarzún (2023b): "School schedules and mothers' employment: evidence from an education reform," *Review of Economics of the Household*, 21, 131–171.

Bertoli, S., J. Fernández-Huertas Moraga, and F. Ortega (2011): "Immigration policies and the Ecuadorian exodus," *The World Bank Economic Review*, 25, 57–76.

Bertrand, M., C. Goldin, and L. F. Katz (2010): "Dynamics of the gender gap for young professionals in the financial and corporate sectors," *American economic journal: applied economics*, 2, 228–255.

Bonhomme, S. and L. Hospido (2017): "The cycle of earnings inequality: evidence from Spanish social security data," *The Economic Journal*, 127, 1244–1278.

Borjas, G. J. (2003): "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *The quarterly journal of economics*, 118, 1335–1374.

Borusyak, K., X. Jaravel, and J. Spiess (2022): "Revisiting Event Study Designs: Robust and Efficient Estimation," *Available at SSRN 2826228*.

Callaway, B., A. Goodman-Bacon, and P. H. Sant'Anna (2021): "Difference-in-differences with a continuous treatment," *arXiv preprint arXiv:2107.02637*.

Callaway, B. and P. H. Sant'Anna (2021): "Difference-in-differences with multiple time periods," *Journal of econometrics*, 225, 200–230.

Card, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.

Card, D., C. Dustmann, and I. Preston (2012): "Immigration, wages, and compositional amenities," *Journal of the European Economic Association*, 10, 78–119.

Card, D. and G. Peri (2016): "Immigration economics by George J. Borjas: a review essay," *Journal of Economic Literature*, 54, 1333–49.

Carrasco, R., J. F. Jimeno, and A. C. Ortega (2008): "The effect of immigration on the labor market performance of native-born workers: some evidence for Spain," *Journal of Population Economics*, 21, 627–648.

Castellanos, M. A. (2020): "Immigration and Native labour market outcomes. Evidence from Spain," *Mimeo, Universidad Carlos III de Madrid*.

Cortés, P. and J. Pan (2019): "When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap," *Journal of Labor Economics*, 37, 351–398.

——— (2020): "Children and the remaining gender gaps in the labor market," Tech. rep., National Bureau of Economic Research.

Cortés, P. and J. Pan (2013): "Outsourcing household production: Foreign domestic workers and native labor supply in Hong Kong," *Journal of Labor Economics*, 31, 327–371.

Cortés, P. and J. Tessada (2011): "Low-skilled immigration and the labor supply of highly skilled women," *American Economic Journal: Applied Economics*, 3, 88–123.

De Chaisemartin, C. and X. d'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–2996.

——— (2023): "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey," *The Econometrics Journal*, 26, C1–C30.

De la Roca, J. and D. Puga (2017): "Learning by working in big cities," *The Review of Economic Studies*, 84, 106–142.

de Quinto, A., L. Hospido, and C. Sanz (2021): "The child penalty: evidence from Spain," *SERIEs*, 12, 585–606.

Dolado, J. J. and P. Vázquez (2008): "Ensayos sobre los efectos económicos de la inmigración en España," .

Dustmann, C. and I. Preston (2012): "Comment: Estimating the effect of immigration on wages," *Journal of the European Economic Association*, 10, 216–223.

Dustmann, C., U. Schönberg, and J. Stuhler (2016): "The impact of immigration: Why do studies reach such different results?" *Journal of Economic Perspectives*, 30, 31–56.

Farré, L., L. González, and F. Ortega (2011): "Immigration, family responsibilities and the labor supply of skilled native women," *The BE Journal of Economic Analysis & Policy*, 11.

Fernández, F. F., M. Hidalgo-Pérez, and S. J. Martín (2010): "Explaining the fall of the skill wage premium in Spain," *Documentos de trabajo (FEDEA)*, 1–58.

Forlani, E., E. Lodigiani, and C. Mendolicchio (2015): "Impact of low-skilled immigration on female labour supply," *The Scandinavian Journal of Economics*, 117, 452–492.

Gelbach, J. B. (2002): "Public schooling for young children and maternal labor supply," *American Economic Review*, 92, 307–322.

Goldin, C. (2006): "The quiet revolution that transformed women's employment, education, and family," *American economic review*, 96, 1–21.

——— (2014): "A grand gender convergence: Its last chapter," *American economic review*, 104, 1091–1119.

Goldman, M. and D. M. Kaplan (2018): "Comparing distributions by multiple testing across quantiles or CDF values," *Journal of Econometrics*, 206, 143–166.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik instruments: What, when, why, and how," *American Economic Review*, 110, 2586–2624.

Gonzalez, L. and F. Ortega (2011): "How do very open economies adjust to large immigration flows? Evidence from Spanish regions," *Labour Economics*, 18, 57–70.

ILO (2013): "Promoting integration for migrant domestic workers in Spain," Tech. rep., International Labor Organization.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," Tech. rep., National Bureau of Economic Research.

Kleven, H. (2022): "The geography of child penalties and gender norms: Evidence from the United States," Tech. rep., National Bureau of Economic Research.

Kleven, H., C. Landais, and G. Leite-Mariante (2023): "The child penalty atlas," Tech. rep., National Bureau of Economic Research.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019a): "Child penalties across countries: Evidence and explanations," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 109, 122–126.

——— (Forthcoming): "Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation," *American Economic Journal: Economic Policy*.

Kleven, H., C. Landais, and J. E. Søgaard (2019b): "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 11, 181–209.

Llull, J. (2021): "Immigration and Gender Differences in the Labor Market," *Journal of Human Capital*, 15, 174–203.

Maza, A., J. Villaverde, and M. Hierro (2013): "Explaining the settlement patterns of foreigners in Spain," *Applied Geography*, 40, 11–20.

Monras, J., J. Vázquez-Grenno, et al. (2018): "Understanding the effects of legalizing undocumented immigrants," *CEPR Discussion Paper No. DP12726*.

Moraga, J. F.-H., A. Ferrer-i Carbonell, and A. Saiz (2019): "Immigrant locations and native residential preferences: Emerging ghettos or new communities?" *Journal of Urban Economics*, 112, 133–151.

Nollenberger, N. and N. Rodríguez-Planas (2015): "Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain," *Labour Economics*, 36, 124–136.

Olivetti, C. and B. Petrongolo (2016): "The evolution of gender gaps in industrialized countries," *Annual review of Economics*, 8, 405–434.

Özgüzel, C. (2021): "The Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession in Spain," .

Peri, G. and C. Sparber (2009): "Task specialization, immigration, and wages," *American Economic Journal: Applied Economics*, 1, 135–169.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2023): "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*.

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.

Yagan, D. (2019): "Employment hysteresis from the great recession," *Journal of Political Economy*, 127, 2505–2558.

A Supplementary Tables

A.1 Descriptive Statistcs

Table A.1: Descriptive Statistics

	Mean (SD)	
	Women	Men
Age	33.81	34.49
	(3.92)	(4.28)
Experience (Before childbirth)	7.54	8.41
	(3.79)	(4.15)
Annual earnings (EUR 2009)	16,631	21,091
	(10,668)	(10,279)
Parenthood Earnings Penalty	-0.20	-0.05
	(0.40)	(0.33)
	Share	
	Women	Men
Education		
Less than secondary	0.33	0.52
Secondary	0.33	0.27
Tertiary	0.34	0.21
Occupation		
Engineers, college graduates & senior managers	0.07	0.07
Technical engineers and graduate assistants	0.10	0.05
Administrative technical managers	0.02	0.04
Non-graduate assistants	0.02	0.03
Administrative officers	0.14	0.09
Subordinates	0.04	0.04
Administrative Assistants	0.33	0.09
First and second class officers	0.05	0.20
Third class officers and technicians	0.11	0.18
Labourers	0.12	0.21
Sector		
Industry & Manufacturing	0.15	0.27
Construction	0.02	0.15
Commerce	0.22	0.17
Services	0.26	0.24
Health, Education, P.A. and Others	0.34	0.16
Observations	60,678	90,138

Notes: Descriptive statistics for the baseline sample. Workers in agriculture and mining are not represented in the analysis given that they do not belong to the general regime of the social security classification.

A.2 First Stage

The following tables show the first stage using the share instrument to predict immigration changes between 1999 and 2007 as described in the robustness section 6.5. Table A.4 shows the results for the instrument for women immigrants only and Table A.5 constructed based on both male and female immigrants. The table reports F-statistics for the first stage and

Table A.2: First stage: shift-share IV for female shock
Results for 2007

	(1) $M^f_{r(i99)t}$	(2) $M^f_{r(i99)t}$	(3) $M^f_{r(i99)t}$	(4) $M^f_{r(i99)t}$
Z_{rt}	0.934***	0.616***	0.860***	0.609***
	(0.132)	(0.0678)	(0.0867)	(0.0657)
N	132	132	127	127
$\bar{R}^2$	0.553	0.649	0.476	0.651
F	50.18	82.41	98.48	85.81
Weights	No	✓	No	✓
Sample	All	All	No outliers	No outliers

Notes: Urban area outliers detected using largest dfbetas in absolute value: Tenerife Sur and Torrevieja. Robust standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table A.3: First stage: shift-share IV for female shock
Pooled regression for all years

	(1) $M^f_{r(i99)t}$	(2) $M^f_{r(i99)t}$	(3) $M^f_{r(i99)t}$	(4) $M^f_{r(i99)t}$	(5) $M^f_{r(i99)t}$	(6) $M^f_{r(i99)t}$	(7) $M^f_{r(i99)t}$	(8) $M^f_{r(i99)t}$
Z_{rt}	1.048***	0.725***	1.005***	0.721***	0.873***	0.644***	0.798***	0.521***
	(0.0539)	(0.0367)	(0.0401)	(0.0367)	(0.0395)	(0.0349)	(0.0483)	(0.0399)
N	1188	1188	1143	1143	1143	1143	1143	1143
$\bar{R}^2$	0.658	0.755	0.615	0.758	0.647	0.788	0.869	0.925
F	378.5	390.5	626.9	385.0	487.3	339.8	273.2	170.7
Weights	No	✓	No	✓	No	✓	No	✓
Sample	All	All	No outliers	No outliers	No outliers	No outliers	No outliers	No outliers
Year FE	No	No	No	No	✓	✓	✓	✓
Area FE	No	No	No	No	No	No	✓	✓

Notes: Urban area outliers detected using largest dfbetas in absolute value: Tenerife Sur and Torrevieja. Robust standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

for each country of origin separately. The preferred specification corresponds to column 4 which excludes two regional outliers based on the highest values of dfbetas from the first stage. The F-statistic of 27 for the women only immigration shock and 41 for the women and men specification proves that this IV version is also relevant.

Table A.4: First Stage for women immigrants instrument
Alternative Instrument

	(1)	(2)	(3)	(4)	(5)	(6)
1 Romania	322.5*	308.9***	322.3**	245.4***	293.2***	251.6***
	(126.9)	(86.99)	(121.6)	(51.55)	(73.00)	(51.56)
2 Morocco	1.448	1.252	1.380	4.954	2.169	4.582
	(2.569)	(1.645)	(2.379)	(2.557)	(2.128)	(2.488)
3 Ecuador	59.13	16.69	62.18	185.7**	56.85	176.9*
	(90.59)	(57.16)	(87.71)	(70.03)	(49.59)	(72.95)
5 Argentina	1.236	1.117	0.426	11.10	6.447	10.39
	(12.77)	(8.177)	(12.32)	(9.411)	(7.640)	(9.085)
6 Colombia	17.50	15.50	17.06	-14.96	-4.109	-16.12
	(33.36)	(28.85)	(32.44)	(22.22)	(24.27)	(21.42)
7 Other	2.395*	1.552**	2.339*	0.787	1.076*	0.845
	(1.073)	(0.575)	(1.034)	(0.655)	(0.469)	(0.615)
Obs.	132	132	132	127	127	127
R2	0.551	0.550	0.538	0.495	0.556	0.486
Weights	NO	Pop 1999	lnPop 1999	NO	Pop 1999	lnPop 1999
Sample	All	All	All	No Outliers	No Outliers	No Outliers
F	6.992	19.594	7.631	27.922	19.667	26.038
F1	6.456	12.610	7.027	22.671	16.130	23.806
F2	0.318	0.579	0.336	3.753	1.039	3.391
F3	0.426	0.085	0.503	7.030	1.314	5.878
F4	0.009	0.019	0.001	1.391	0.712	1.308
F5	0.275	0.288	0.277	0.453	0.029	0.566
F6	4.986	7.269	5.120	1.441	5.270	1.890

Notes: Urban area outliers detected using dfbetas: Tenerife Sur and Torrevieja. The table reports F-statistics for the first stage and for each country of origin separately.
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A.5: First Stage for all immigrants instrument
Alternative Instrument

	(1)	(2)	(3)	(4)	(5)	(6)
1 Romania	347.5*	292.3***	345.9*	293.6***	281.8***	295.7***
	(141.8)	(81.22)	(136.1)	(75.65)	(70.57)	(75.88)
2 Morocco	4.653**	4.397**	4.579**	5.533***	5.108**	5.481***
	(1.592)	(1.636)	(1.568)	(0.997)	(1.732)	(1.031)
3 Ecuador	48.38	1.936	49.24	225.3*	57.64	212.8*
	(125.5)	(88.15)	(125.1)	(98.11)	(75.18)	(98.59)
4 Argentina	-17.66	-5.200	-16.93	1.184	2.561	1.236
	(14.72)	(8.715)	(14.18)	(9.471)	(7.542)	(9.031)
5 Colombia	42.72	24.52	38.91	-22.27	-5.420	-25.27
	(59.70)	(47.87)	(58.24)	(34.42)	(39.91)	(33.41)
6 Other	3.523**	1.865**	3.412**	1.493*	1.213*	1.496*
	(1.240)	(0.712)	(1.213)	(0.690)	(0.543)	(0.654)
Obs.	132	132	132	127	127	127
R2	0.599	0.551	0.585	0.588	0.581	0.579
Weights	NO	Pop 1999	lnPop 1999	NO	Pop 1999	lnPop 1999
Sample	All	All	All	No Outliers	No Outliers	No Outliers
F	11.428	18.478	11.188	41.185	20.179	36.204
F1	6.004	12.948	6.457	15.066	15.948	15.193
F2	8.541	7.221	8.524	30.831	8.703	28.256
F3	0.148	0.000	0.155	5.275	0.588	4.658
F4	1.440	0.356	1.424	0.016	0.115	0.019
F5	0.512	0.262	0.446	0.419	0.018	0.572
F6	8.074	6.866	7.915	4.687	4.987	5.233

Notes: Urban area outliers detected using dfbetas: Tenerife Sur and Torrevieja. The table reports F-statistics for the first stage and for each country of origin separately.
* p<0.05, ** p<0.01, *** p<0.001

A.3 Additional Results

Table A.6: Heterogeneity Analysis: Effect of immigration on PEP by skill group

	Education		Income		Occupation	
	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Mothers	0.504***	0.215	0.567**	0.263	0.543***	0.173
	(0.139)	(0.147)	(0.207)	(0.136)	(0.152)	(0.113)
Fathers	0.0278	0.0478	0.177	-0.057	0.0017	0.110
	(0.0949)	(0.127)	(0.195)	(0.0905)	(0.106)	(0.107)
Gap: Men - Women	-0.476**	-0.167	-0.390*	-0.320*	-0.541***	-0.063
	(0.146)	(0.145)	(0.157)	(0.140)	(0.143)	(0.129)
N	146,263	146,263	146,263	146,263	146,263	146,263
$\bar{R}^2$	0.0642	0.0642	0.0191	0.0191	0.0654	0.0654
KP	46.01	46.01	36.04	36.04	43.30	43.30
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation 4. The dependent variable is the *parenthood earnings penalty.* A positive effect on women/men implies a reduction in child penalties. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. All regressions control for age and education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

Table A.7: Effect of female immigration on PEP by detailed education groups

	(1)	(2)	(3)
	Below Secondary	Secondary	College
Mothers	0.081	1.027***	0.141
	(0.221)	(0.146)	(0.156)
Fathers	-0.057	0.034	0.160
	(0.142)	(0.104)	(0.138)
Gap: Fathers-Mothers	-0.138	-0.993***	0.019
	(0.272)	(0.180)	(0.181)
N	66,494	44,313	40,007
$\bar{R}^2$	0.046	0.050	0.034
KP	62.781	116.185	141.539
Time-Sex FE	✓	✓	✓
Area-Sex FE	✓	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation 4. The dependent variable is the *parenthood earnings penalty.* A positive effect on women/men implies a reduction in child penalties. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. All regressions control for age and education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. I also include an interaction of all these controls with gender. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

Table A.8: Effect of female immigration on daily earnings

	ln(Daily earnings)			
	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)
Panel A. All				
Mothers	0.357*	0.460**	0.204*	0.295**
	(0.137)	(0.149)	(0.100)	(0.112)
Fathers	-0.219*	-0.216*	-0.003	0.0612
	(0.100)	(0.092)	(0.061)	(0.055)
Gap: Men-Women	-0.576***	-0.676***	-0.207	-0.234
	(0.153)	(0.161)	(0.115)	(0.122)
Panel B. Education				
Secondary				
Mothers	0.471**	0.590***	0.236*	0.344**
	(0.153)	(0.159)	(0.111)	(0.119)
Fathers	-0.230	-0.236*	-0.0169	0.0466
	(0.121)	(0.114)	(0.0720)	(0.0647)
Gap: Men-Women	-0.701***	-0.826***	-0.253*	-0.297*
	(0.174)	(0.190)	(0.123)	(0.132)
Tertiary				
Mothers	0.199	0.281	0.160	0.228
	(0.176)	(0.194)	(0.132)	(0.145)
Fathers	-0.186	-0.159	0.0387	0.104
	(0.127)	(0.131)	(0.097)	(0.108)
Gap: Men-Women	-0.385*	-0.440*	-0.122	-0.124
	(0.167)	(0.180)	(0.137)	(0.149)
N	135,328	70,708	131,811	70,578
A. $\bar{R}^2$	0.193	0.192	0.510	0.531
K-P	103.3	107.9	102.6	107.9
B. $\bar{R}^2$	0.193	0.192	0.509	0.531
K-P	42.87	43.48	42.47	43.53
$\text{Earnings}_{c-1,c-2}$	No	No	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation4. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. The dependent variable is the logarithm of average daily earnings. Regressions control for age, education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Sex-area and sex-year fixed effects are included. Regressions in panel B also include education-sex-year and double interaction education-year and education-area fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

Table A.9: Effect of female immigration on differences on earnings with respect to counterfactual: $ln(earnings)_{it} - ln(earnings)_{it(0)}$

	All		No Construction Boom		No Labourers	
	$\Delta ln(ear)$	$\Delta ln(ear)$	$\Delta ln(ear)$	$\Delta ln(ear)$	$\Delta ln(ear)$	$\Delta ln(ear)$
	1-5 yrs	1-3 yrs	1-5 yrs	1-3 yrs	1-5 yrs	1-3 yrs
Mothers	0.632***	0.735***	0.621***	0.719***	0.586***	0.637***
	(0.142)	(0.159)	(0.144)	(0.162)	(0.144)	(0.161)
Fathers	-0.162	-0.0549	-0.0908	0.0251	-0.0295	0.0761
	(0.0920)	(0.104)	(0.0878)	(0.0963)	(0.0841)	(0.101)
Gap: Fathers-Mothers	-0.794***	-0.790***	-0.711***	-0.694***	-0.615***	-0.560***
	(0.158)	(0.169)	(0.158)	(0.169)	(0.148)	(0.162)
N	143073	74823	138802	72616	130833	68444
$\bar{R}^2$	0.0330	0.0324	0.0333	0.0328	0.0296	0.0293
KP	98.33	102.8	104.4	108.4	100.6	104.8
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓	✓

Notes: Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

Table A.10: Effect of female immigration on job quality after childbirth

	Upgrade		Same		Downgrade	
	1-5 years	1-3 years	1-5 years	1-3 years	1-5 years	1-3 years
	(1)	(2)	(3)	(4)	(5)	(6)
Mothers	0.197	0.0883	-0.236	-0.0850	-0.0261	-0.00329
	(0.111)	(0.121)	(0.125)	(0.176)	(0.102)	(0.106)
Fathers	0.271***	0.265***	-0.453***	-0.428***	0.122	0.163
	(0.0719)	(0.0762)	(0.0959)	(0.122)	(0.0949)	(0.100)
Gap: Fathers-Mothers	0.0745	0.177	-0.217	-0.343	0.148	0.166
	(0.126)	(0.139)	(0.172)	(0.220)	(0.150)	(0.164)
N	120,019	63,949	120,019	63,949	120,019	63,949
$\bar{R}^2$	0.0322	0.0261	0.0646	0.0563	0.0310	0.0230
KP	97.56	102.7	97.78	102.7	97.78	102.7
Time-Sex FE	✓	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓	✓

Notes: The results correspond to pooled regressions following the specification of equation4. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. The dependent variables are the probablity of being in a better (Columns (1),(2)), same (Columns (3) (4)) or equal (Columns (5),(6)) type of job with respect to the year before childbirth. Job quality is proxied by occupation-sector estimated fixed effects from a movers design approach, ranked in percentiles. Regressions control for age, education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Sex-area and sex-year fixed effects are included. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

A.4 Matching

Table A.11: Balance test for matching sample: parents and non-parents

	Mother	Non-Mother	Difference
ln(Earnings)$_{c-1}$	9.73	9.70	0.05***
	(0.54)	(0.53)	(34.49)
ln(Earnings)$_{c-1}$	9.72	9.66	-0.01***
	(0.52)	(0.55)	(-5.23)
Experience$_{c-1}$	8.30	7.78	-0.23***
	(3.56)	(4.86)	(-22.32)
Age$_{c-1}$	30.78	30.42	-0.25***
	(3.76)	(5.76)	(-20.84)
N	91,486		
	Father	**Non-Father**	**Difference**
ln(Earnings)$_{c-1}$	9.88	9.85	-0.02***
	(0.45)	(0.49)	(-15.82)
ln(Earnings)$_{c-2}$	9.83	9.81	-0.01***
	(0.49)	(0.51)	(-7.20)
Experience$_{c-1}$	9.14	8.91	-0.13***
	(3.91)	(5.51)	(-12.51)
Age$_{c-1}$	31.47	31.36	-0.15***
	(4.04)	(6.43)	(-12.67)
N	138,294		

Notes: The table displays means and standard deviations in parentheses of baseline variables: i.e., one year before a child is born for parents and matched non-parents using propensity score matching within cells of sex, year and autonomous region. The table restricts to the main sample used in analysis, restricting by age (25-50), up to 5 years after childbirth and the all the sample restrictions described in the analysis.

Table A.12: Effect of female immigration on earnings: parents and non-parents

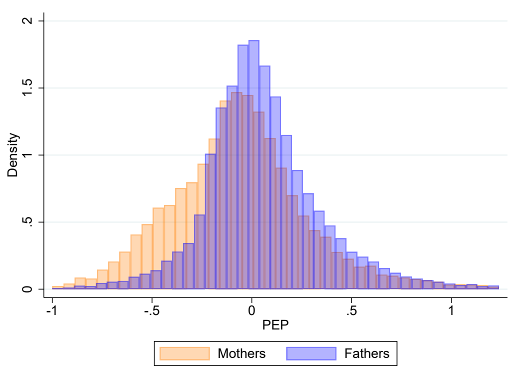
	ln(Earnings)			
	Women		Men	
	(1)	(2)	(3)	(4)
Parent	0.430**	0.309**	-0.315*	-0.136
	(0.140)	(0.104)	(0.127)	(0.084)
Non-Parent	0.0121	-0.0802	-0.125	-0.127
	(0.124)	(0.0911)	(0.124)	(0.091)
Difference	0.418**	0.389**	-0.190	-0.009
	(0.157)	(0.129)	(0.136)	(0.085)
N	80,646	80,646	126,010	126,010
$\bar{R}^2$	0.287	0.540	0.267	0.548
KP	97.87	97.90	91.97	92.00
Time-Sex FE	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓
Earnings$_{c-1,c-2}$	No	✓	No	✓

Notes: The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years (odd columns) and who had their first kid in the last 3 years (even columns). Year of birth (age 0) is omitted. Regressions control for age, education and pre-determined characteristics before childbirth including experience, 2-digit industry fixed effects and interactions of each control with sex. All regressions control for baseline earnings one and two years before childbirth. Sex-area and sex-year fixed effects are included. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

B Comparing Distributions of Parenthood Earnings for Mothers and Fathers

Figure B.1: Parenthood Penalty Distribution

(a) Employed parents



(b) All parents

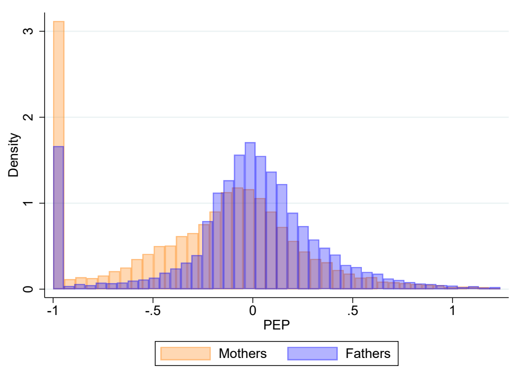


Figure B.1 shows the distribution of the *parenthood penalty* for fathers and mothers, with Panel a) displaying the distribution for employed parents and Panel b) for both employed and non-employed. In addition, Figure B.2 shows the empirical cumulative distribution function (CDF) of the *parenthood penalty* for fathers and mothers. The thick horizontal line shows the range where CDF equality was rejected at the 10% familywise error rate (FWER) level based on the Goldman and Kaplan (2018) methodology. The results confirm that the parenthood earnings penalty are higher for fathers than for mothers over the whole range of the empirical distribution. The test is similar to the two-sample Kolmogorov-Smirov (KS) test, however instead of comparing a single distributional equality it evaluates the equality of distributions point by point. Additionally, Table B.1 provides results for the KS test.

Table B.1: Kolmogorov-Smirnov Test

Smaller group	All parents		Employed	
	D	p-value	D	p-value
Fathers	0.0000	0.999	0.0000	0.999
Mothers	-0.1962	0.000	-0.1637	0.000
Combined K-S	0.1962	0.000	0.1637	0.000

Notes: The first line tests the hypothesis that the distribution of fathers has *smaller* values than for mothers, the largest difference between the distribution functions is 0.000. The second line tests the hypothesis that the distribution of fathers contains *larger* values than for mothers. The largest difference between the distribution function is 0.1962. The last line provides the results for the combined test.

Intuition behind the test

Following Goldman and Kaplan (2018) let $F(\cdot)$ be the first group's CDF; $G(\cdot)$ is the second group's CDF. Estimated CDFs $\hat{F}(\cdot)$ and $\hat{G}(\cdot)$ are computed from iid samples. The Kolmogorov–Smirnov (KS) goodness-of-fit (GOF) null hypotehsis is:

$$\text{(GOF) } H_0 : F(r) = G(r) \text{ for all } r, \tag{10}$$

indicating identical CDFs. If H_0 is true, then $\hat{F}(\cdot)$ and $\hat{G}(\cdot)$ should be *close* to each other; if not, the test rejects. Where the distance at r is defined as:

$$bD(r) \equiv |\hat{F}(r) - \hat{G}(r)|. \tag{11}$$

To compare the entire functions, KS looks across all r to find the biggest gap, $\max_r bD(r)$. Instead of 10, null hypotheses can be defined to show where two distributions differ. For each possible value r of the outcome variable, define:

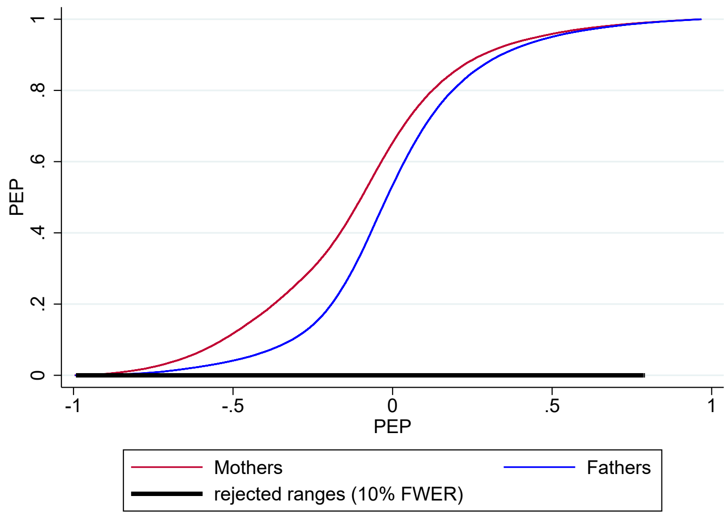
$$H_{0r} : F(r) = G(r). \tag{12}$$

The GOF null hypothesis could be rewritten as:

$$\text{(GOF again) } H_0 : \text{all } H_{0r} \text{ are true.} \tag{13}$$

Whereas the GOF test only distinguishes whether all H0r are true or at least one is false, the Goldman and Kaplan (2018) method looks specifically which H0r are true and which are false.

Figure B.2: Empirical CDF of parenthood earnings penalty



Notes: The figure shows the CDF of the parenthood earnings penalty for fathers and mothers. The sample restrictions are the same imposed for the baseline estimations analysed in section 5. It considers parents who are between 25 to 50 years old and had their first kid over the last five years. The thick horizontal line shows the range where CDF equality was rejected at the 10% FWER level.

C Regional estimates

Constructing regional cells of median outcomes by sex, let $\Delta PEPrt^k = PEP_{rt} - PEP_{r,1998}$. Using long difference of outcome with respect to 1998, I estimate for each year:

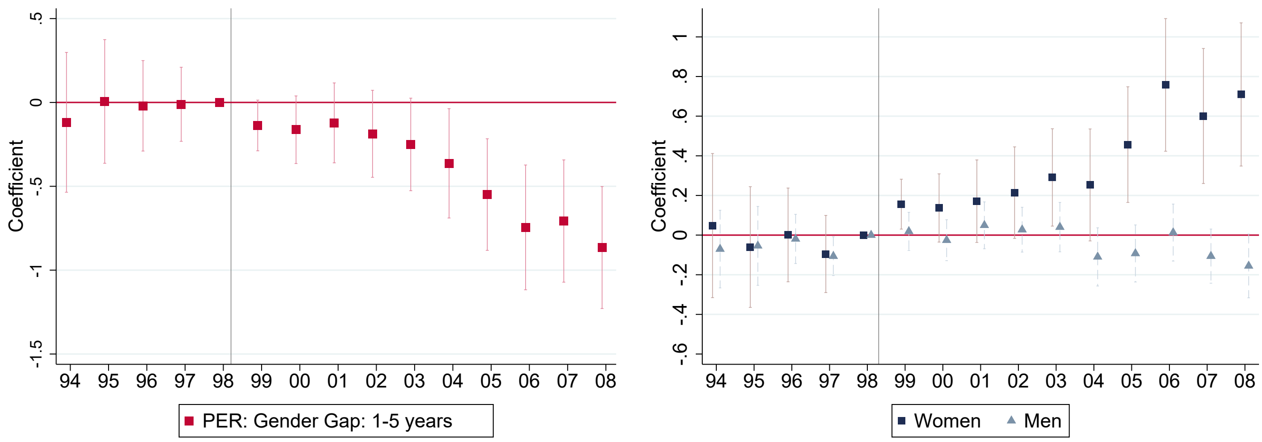
$$\begin{aligned}\Delta PEP_{rt}^k = \alpha &+ \beta_{1t}\Delta M^f{}_{r(i99)t=2007} \\ &+ \beta_{2t}\Delta M^f{}_{r(i99)t=2007} * g_r + \gamma g_r + \theta \bar{X}_{(pre-birth)_{grt}} \\ &+ e_{it},\end{aligned} \tag{14}$$

where $\bar{X}_{grt}$ includes the average of: age, experience (before birth), age of first child and share of workers with up to secondary education by region, gender and year.

Figure C.1: Regional effect on PEP: FD equation 14

(a) FD eq. 14: Median PEP

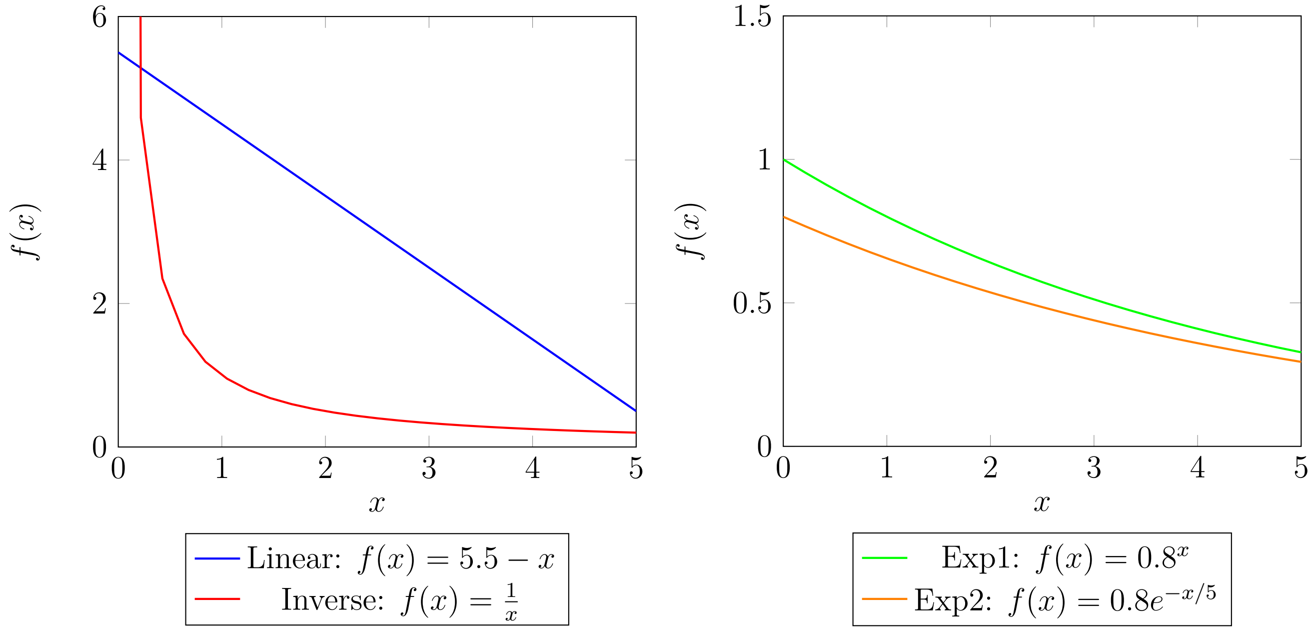
(b) FD eq. 14: Median PEP



Notes: Regional estimates. The outcome is the median of PEP at the area level. Estimates are weigthed by cell size.

D Robustness: Weighting functions

Figure D.1: Weigthing functions



Notes: For each function f(x), x denotes time since childbirth

Table D.1: Impact of female immigration on PEP: Weighted specifications

	PER				
	(1)	(2)	(3)	(4)	(5)
Mothers	0.472***	0.412***	0.481***	0.446***	0.443***
	(0.127)	(0.117)	(0.127)	(0.122)	(0.121)
Fathers	-0.0787	-0.127	-0.0970	-0.102	-0.104
	(0.0916)	(0.0859)	(0.0906)	(0.0877)	(0.0874)
Gap: Men - Women	-0.551***	-0.539***	-0.578***	-0.548***	-0.547***
	(0.134)	(0.125)	(0.134)	(0.129)	(0.129)
N	150,816	150,816	150,816	150,816	150,816
$\bar{R}^2$	0.0433	0.0423	0.0436	0.0429	0.0428
KP	99.46	96.43	99.68	98.15	97.97
Time-Sex FE	✓	✓	✓	✓	✓
Area-Sex FE	✓	✓	✓	✓	✓
Weight	Linear	No	Inverse	Exp1	Exp2

Notes: The results correspond to pooled regressions following the specification of equation 4. The dependent variable is the parenthood earnings penalty. A positive effect on women/men implies a reduction in child penalties. The sample is restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Year of birth (age 0) is omitted. All regressions control for age and education and pre-determined characteristics before childbirth including cells of experience and 2-digit industry fixed effects. Robust standard errors in parentheses, clustered at the area level. * p<0.05, ** p<0.01, *** p<0.001.

E Event-Study Design

In this section I estimate an alternative to specification 6 based on an event-study design approach:

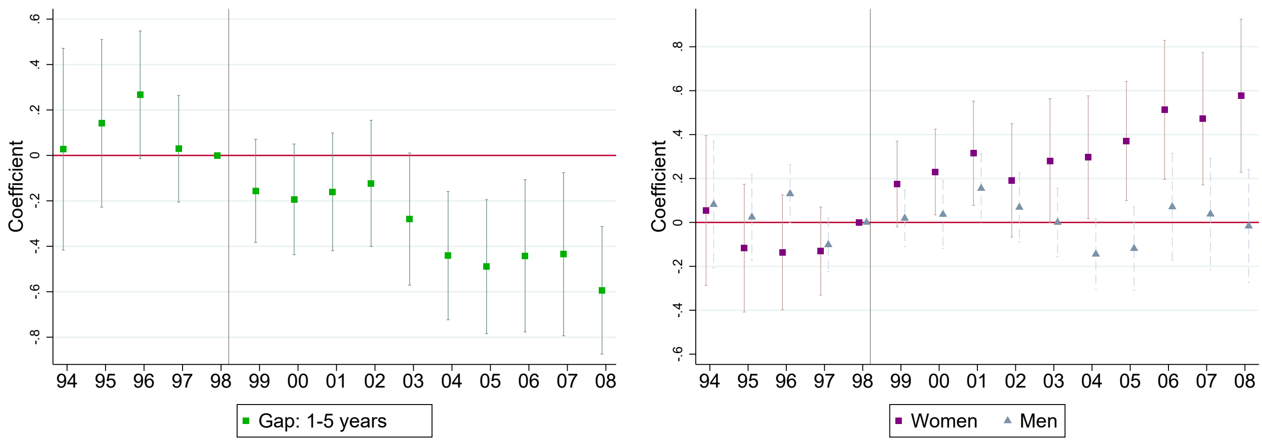
$$
\begin{aligned}
PER^k_{it} = \alpha + & \sum_{s \neq 1998} \beta_{1t} {M^f}_{r(i99)} I\{s = t\} \\
& + \sum_{s \neq 1998} \beta_{2t} {M^f}_{r(i99)} I\{s = t\} \cdot g_i + \gamma_g + \theta X_{i(pre-birth)} \\
& + \rho_{gt} + \delta_{gr} + e_{it},
\end{aligned} \tag{15}
$$

where the migration rates are instrumented via an IV shift-share strategy.

Figure E.4a plots the coefficients β_{2t} which captures the effect on the PEP gender gap. Figure E.4b plots the effect on women captured by β_{1t} and the effect on men captured by $\beta_{1t} + \beta_{2t}$.

Figure E.1: Effect of immigration on PEP based on event-study design (equation 15)

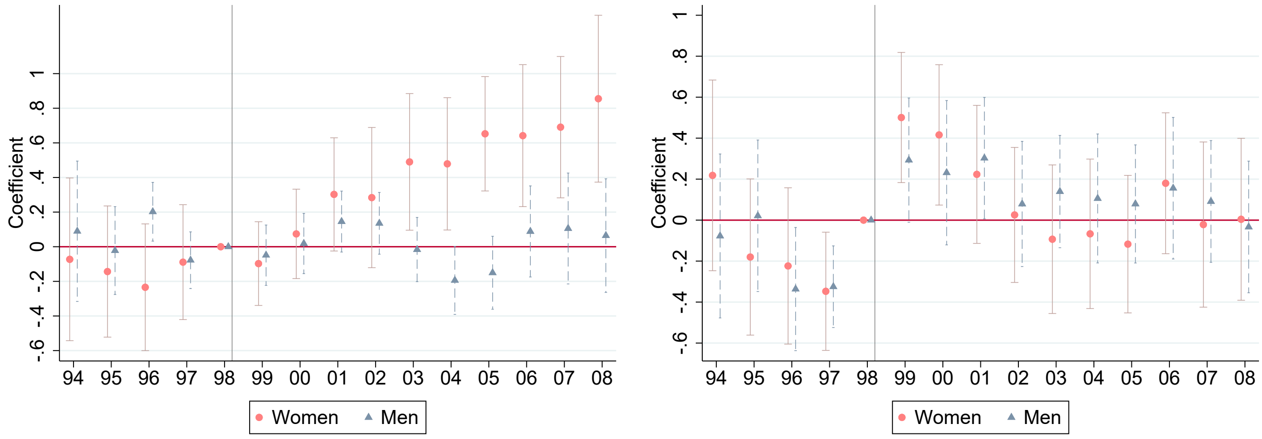
(a) Gender Gap (b) Effect by sex



Notes: First kid in the last 5 years. Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

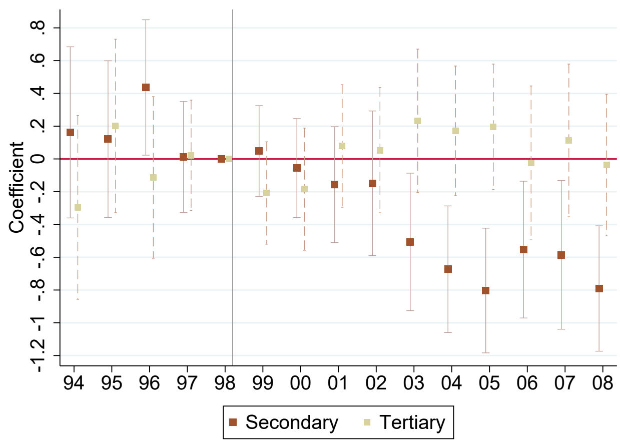
Figure E.2: Effect of immigration on PEP by education and sex

(a) Secondary (b) Tertiary



Notes: First kid in the last 5 years. Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

Figure E.3: Effect on PEP Gender Gap by education.

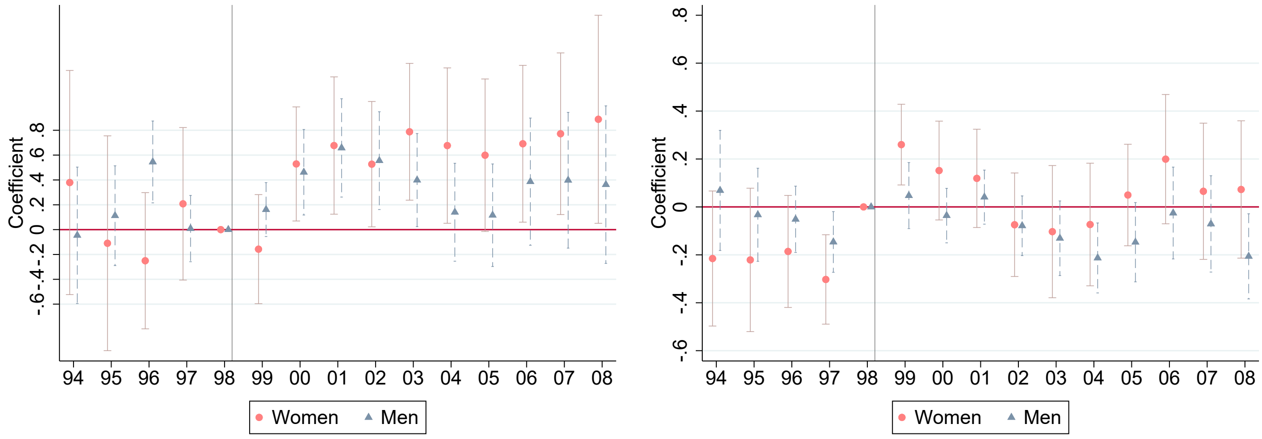


Notes: First kid in the last 5 years. Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

Figure E.4: Effect of immigration on PEP by income group and sex

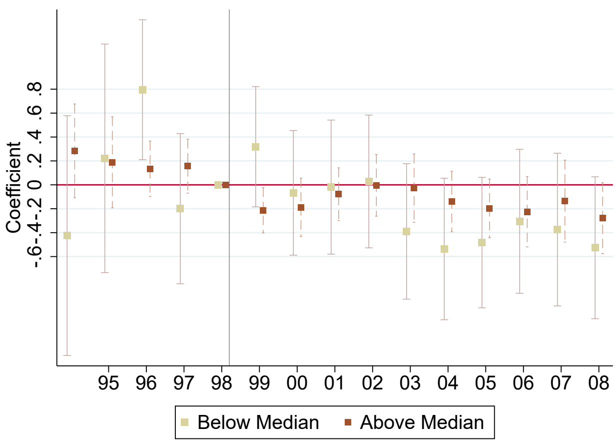
(a) Below Median

(b) Above Median



Notes: First kid in the last 5 years. Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

Figure E.5: Effect of immigration on PEP: Gender Gap by median income.



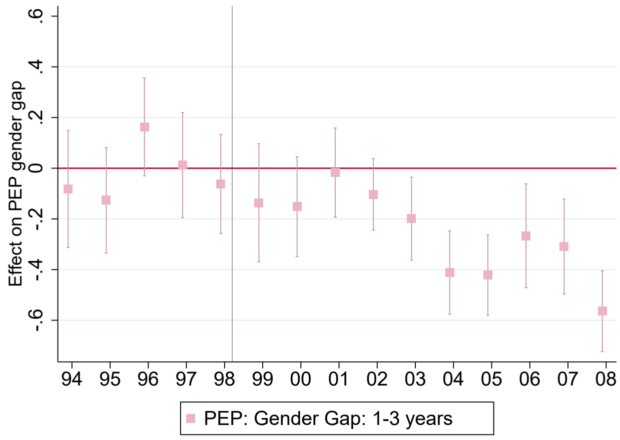
Notes: First kid in the last 5 years. Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 5 years. Robust standard errors clustered at the area level. 95% confidence intervals.

F Additional Results

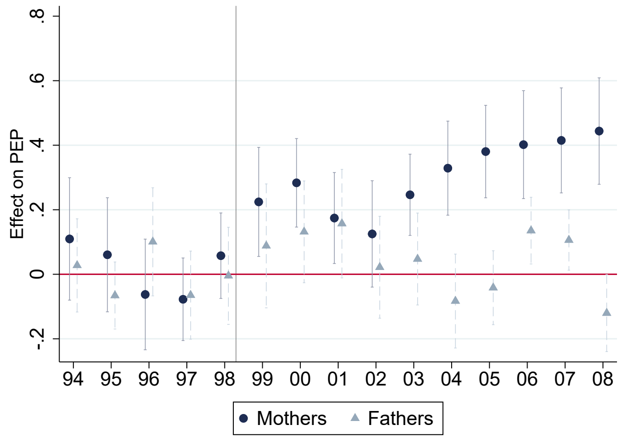
F.1 Effect of immigration on PEP for 1-3 years after childbirth

Figure F.1: Effect of immigration on PEP
First kid in the last 3 years

(a) Impact on gender gap



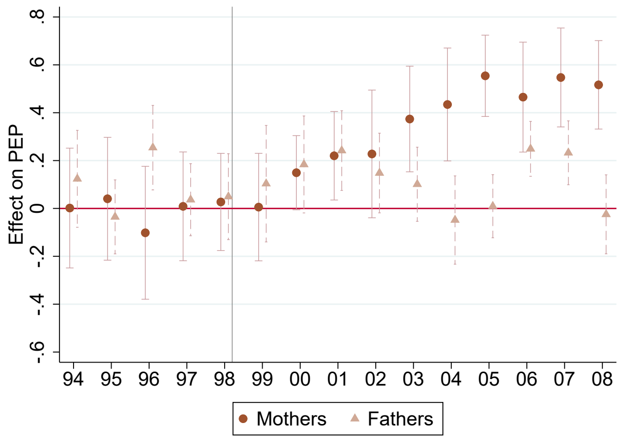
(b) Impact by sex



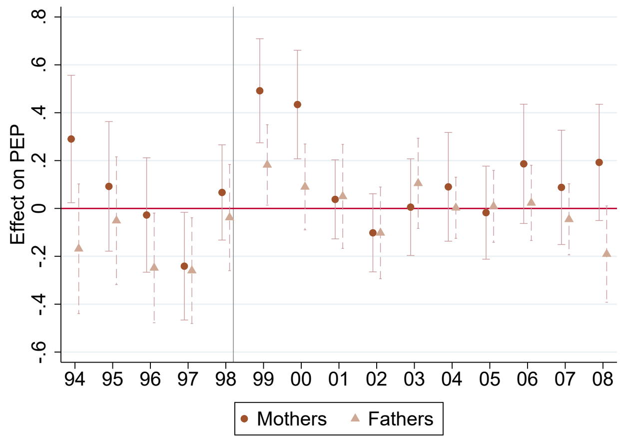
Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 3 years. Robust standard errors clustered at the area level. 95% confidence intervals.

Figure F.2: Effect of immigration on PEP by education level
First kid in the last 3 years

(a) Secondary



(b) Tertiary

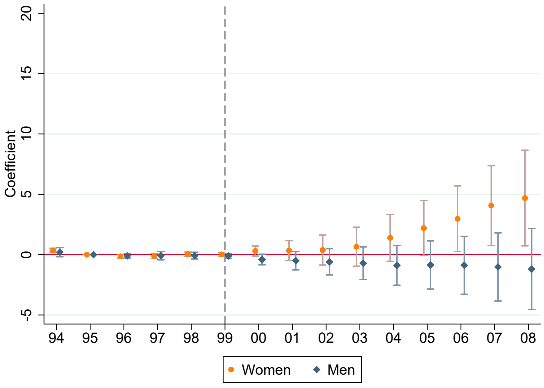


Notes: Sample restricted to parents between 25 to 50 years old in a given year who had their first kid in the last 3 years. Robust standard errors clustered at the area level. 95% confidence intervals.

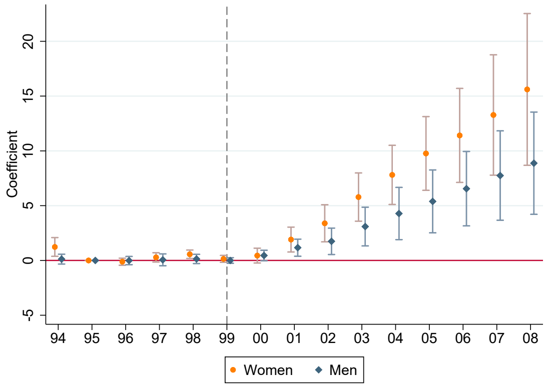
F.2 Following workers over time

Figure F.3: Effect on cumulative earnings by income group at baseline

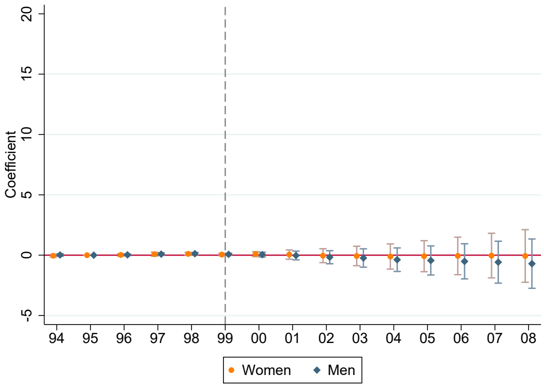
(a) No kids - Below Median



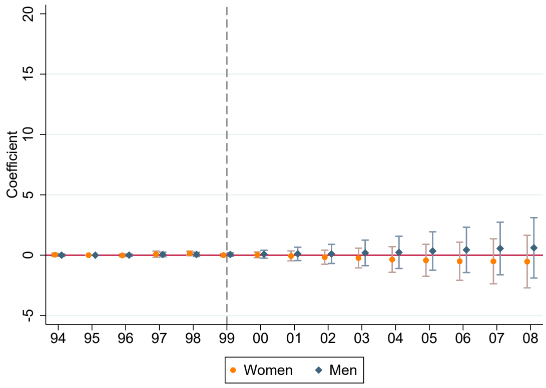
(b) Kid(s) younger than 3 at baseline - Below Median



(c) No kids - Above Median



(d) Kid(s) younger than 3 at baseline- Above Median



Notes: Fixed sample of workers restricted to individuals that were between 30 to 49 years old in 1999. Family characteristics are determined at the baseline 1999 year as well as income groups, based on the average earnings between 1995-1999 and their position in the gender-specific income distribution. Robust standard errors clustered at the area level. 95% confidence intervals.

G Evidence of child penalties in Spain: Aggregate estimates

This section estimates the traditional child penalty average effect by following Kleven et al. (2019a) and de Quinto et al. (2021). Particularly I estimate the following event-study regression:

$$Y_{ist}^{g}=\sum_{j}\alpha_{j}^{g}I\left(j=age_{is}\right)+\sum_{k}\beta_{k}^{g}I(k=s)+\sum_{l\neq-1}\delta_{l}^{g}I(l=t)+\varepsilon_{it}^{g}$$

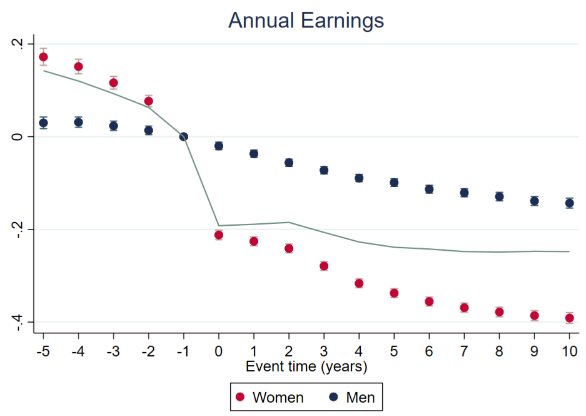
where Y_{ist}^{g} represents the outcome of interest for individual i of gender g at calendar time s at event time t. The event time dummy corresponding to $t=-1$ is excluded, so that the event time coefficients capture the impact of parenthood relative to the year preceding the first childbirth. The inclusion of age FE controls nonparametrically for latent lifecycle trends. The year and month FEs control for business cycle effects.

In a second step, the estimated level effects are converted into percentage figures, calculated as

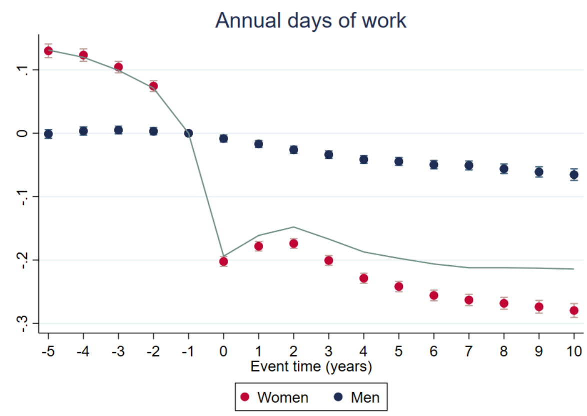
$$P_{t}^{g}=\hat{\delta}_{t}^{g}/E\left[\tilde{Y}_{ist}^{g}\mid t\right],$$

where $\tilde{Y}_{ist}^{g}$ is the predicted labour income net of the event time dummies, that is, the counterfactual in the hypothetical case of not having children:

$$\tilde{Y}_{ist}^{g}=\sum_{j}\hat{\alpha}_{j}^{g}I\left\{j=age_{is}\right\}+\sum_{k}\hat{\beta}_{k}^{g}I\{k=s\}$$



(a) Year of (child) birth 1994-2008



(b) Year of (child) birth 1994-2008

Notes: Panel a) displays the child penalties in earnings. Panel b) displays the child penalties in days of work. The sample is restricted to children that were born between 1994 and 2008.

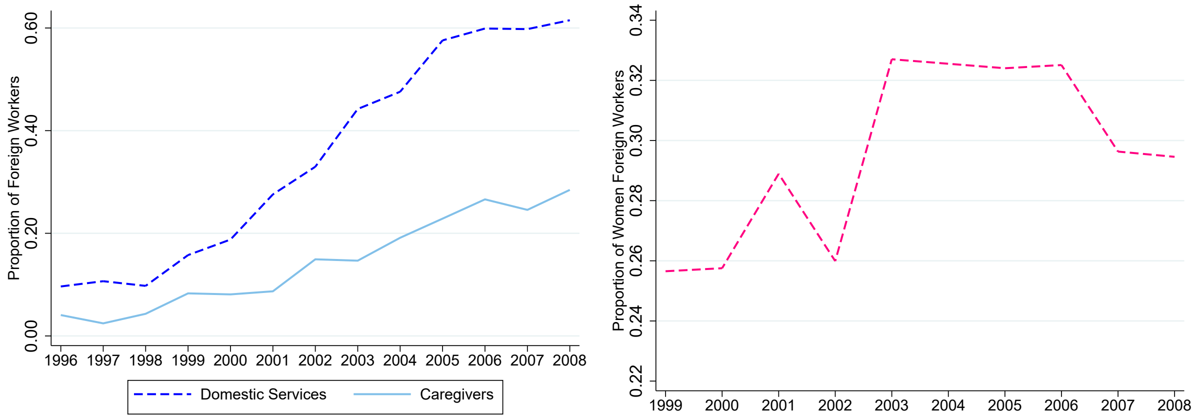
H Immigrant Population

H.1 Foreign-born domestic workers

Figure H.1: Proportion of immigrants in the household services sector

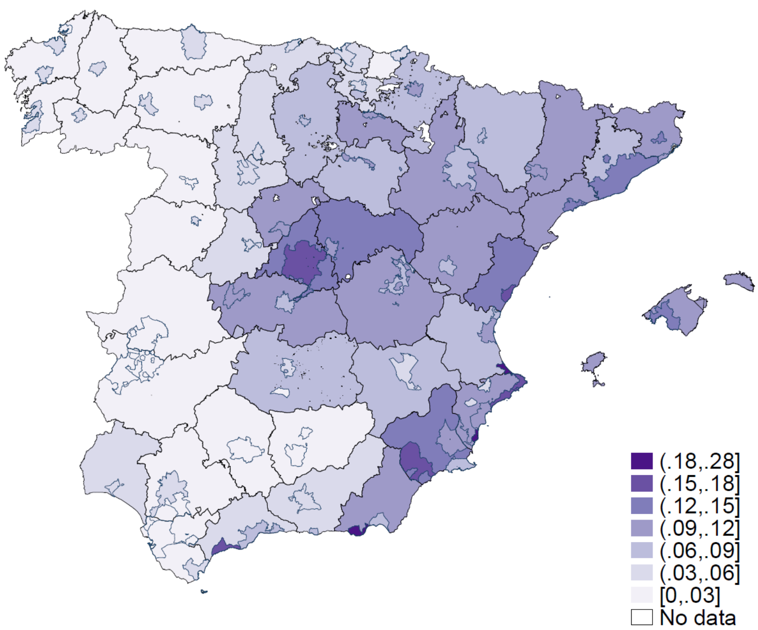
(a) Immigrants in household services
(Share of total workers in the sector)

(b) Female immigrants in household services



Notes: Panel a) displays the proportion of immigrants in the domestic and caregiving as share of all workers in those sectors. Panel b) displays the proportion of female immigrants employed in household services as share of all female immigrant workers in Spain. Source: Labour Force Survey.

Figure H.2: Share of female immigrants from selected countries, 2007



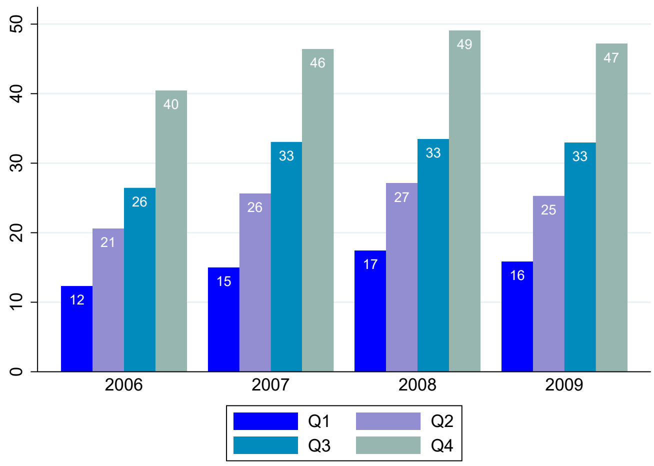
Notes: Selected countries include: Ecuador, Colombia, Romania, Dominican Republic, Poland, Peru, Morocco, Bolivia, Argentina, Bulgaria, Paraguay and Brazil. Urban areas within provinces are outlined in the map. Source: Municipal Registry, Peninsula and Balearic Islands.

Table H.1: Top 10 origin countries in household services

2003			2007		
Rank	Country	Share	Rank	Country	Share
1	Ecuador	0.26	1	Romania	0.17
2	Colombia	0.25	2	Ecuador	0.17
3	Romania	0.11	3	Bolivia	0.14
4	Dominican Republic	0.04	4	Colombia	0.07
5	Poland	0.04	5	Paraguay	0.05
6	Peru	0.03	6	Morocco	0.04
7	Morocco	0.02	7	Dominican Republic	0.04
8	Bolivia	0.02	8	Argentina	0.04
9	Argentina	0.02	9	Peru	0.03
10	Bulgaria	0.02	10	Brazil	0.03

Source: Labour Force Survey.

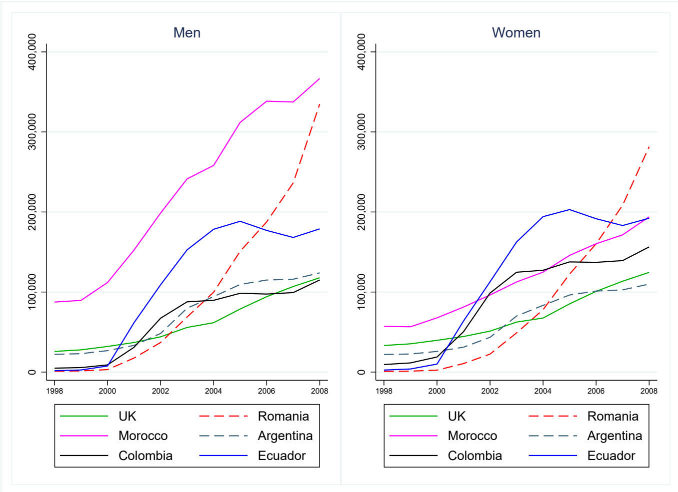
Figure H.3: Share of Households that spent in Domestic Services (%)



Notes: The figure shows the share of households that spent a positive ammount in househol services. The categories correspond to expenditure quartiles. Earlier years are not included due to a change in methodology. Source: Household's Expediture Survey

H.2 Immigration trends in Spain

Figure H.4: Evolution of foreign-born population from top origin countries by gender



Notes: Source: Local Registry (Padrón Continuo Municipal).

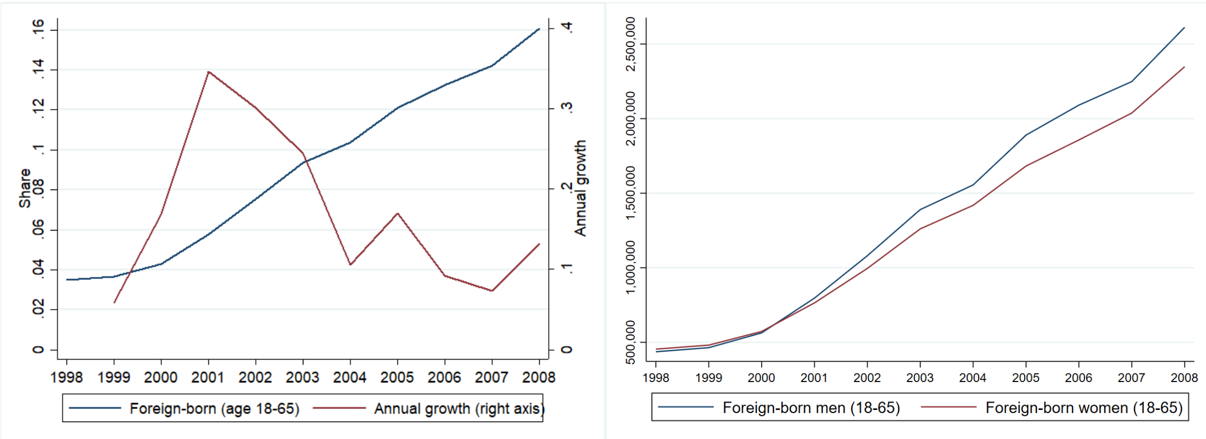
Table H.2: Foreign-born population in Spain aged 18-65 by country of origin

	1999		2004		2007	
Origin	*Thousands*	*Share (%)*	*Thousands*	*Share (%)*	*Thousands*	*Share (%)*
France	116	12.1	145	4.9	170	4.0
Italy	15	1.6	34	1.1	56	1.3
Portugal	41	4.3	57	1.9	92	2.1
UK	63	6.6	129	4.3	220	5.1
Germany	99	10.3	139	4.6	165	3.8
Other CEE	19	1.9	376	12.6	770	17.9
Other Europe	96	10.0	139	4.7	176	4.1
Morocco	155	16.3	396	13.3	523	12.2
Other Africa	48	5.0	156	5.2	216	5.0
USA	12	1.2	19	0.6	21	0.5
Cuba	18	1.9	50	1.7	62	1.5
Argentina	46	4.8	178	5.9	219	5.1
Venezuela	40	4.2	84	2.8	112	2.6
Mexico or Canada	13	1.3	30	1.0	37	0.8
Other CAC	29	3.0	72	2.4	116	2.7
Other South America	97	10.1	844	28.3	1132	26.3
Asia & Oceania	51	5.3	140	4.7	213	4.9

Source: Local Registry. Own Elaboration

Figure H.5: Foreign-born population aged 18-65 in Spain

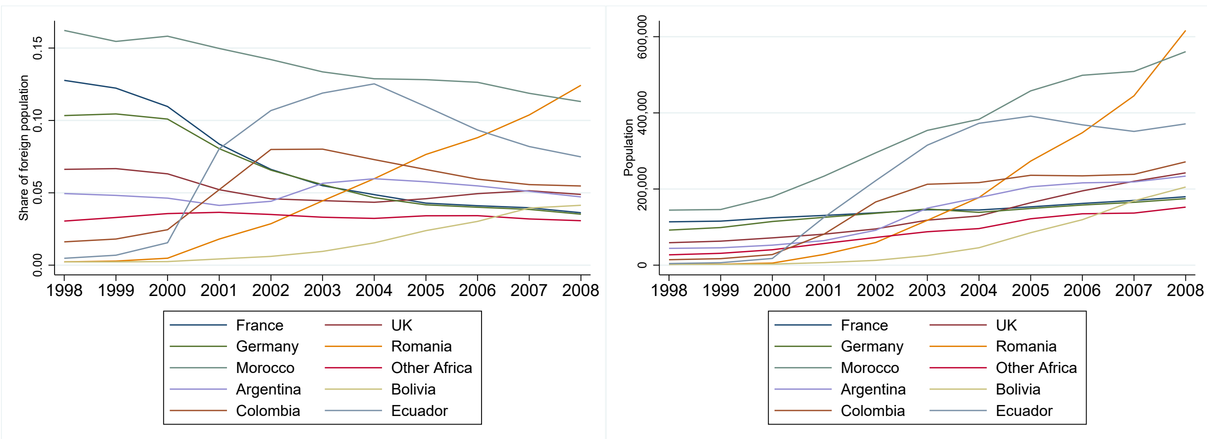
(a) Shares (b) Stocks



Notes: Source: Local Registry.

Figure H.6: Foreign-born population by country of origin

(a) Shares (b) Stocks



Source: Local Registry