Série des Documents de Travail

n° 2019-17

Child Penalties and Financial Incentives: Exploiting Variation along the Wage distribution

Pierre PORA[1]
Lionel WILNER[2]

[1] Insee-Crest. E-mail : pierre.pora@insee.fr
[2] Insee-Crest. E-mail : lionel.wilner@insee.fr

Child Penalties and Financial Incentives: Exploiting Variation along the Wage Distribution*

Pierre Pora† Lionel Wilner‡

September 27, 2019

Abstract

We relate women's labor earnings losses due to motherhood to their pre-childbirth rank in the distribution of hourly wages. Using French administrative data, we show that these "child penalties" decrease steeply along the distribution; by contrast, the related hourly wage losses are fairly homogeneous. Low-wage mothers opt out of the labor market or reduce their working hours more frequently; the magnitude of such responses is consistently monotonic along the distribution. This empirical evidence highlights the relevance of financial incentives and suggests that child penalties arise from decisions based on specialization gains rather than on gender differences in preferences or on gender norms.

Keywords: Gender pay gap, child penalties, labor supply, difference-in-difference, wage distribution.

JEL Classification: J13, J16, J22, J31.

*We thank Martin Andresen, Thomas Breda, Alex Bryson, Olivier Godechot, Libertad González, Dominique Goux, Erica Lindahl, Dominique Meurs, Sébastien Roux, Andreas Steinhauer, Milena Suarez-Castillo, Grégory Verdugo and Josef Zweimüller, as well as attendees at CASD-IAB Worskshop "Advances in Social Sciences Using Administrative and Survey Data" (Paris, 2019), AFSE (Paris, 2018), EALE (Uppsala, 2019), ESEM (Cologne, 2018), ESPE (Bath, 2019), JMA (Bordeaux, 2018), JMS (Paris, 2018), LAGV (Aix-en-Provence, 2019) and at Insee seminar for useful suggestions. All errors and opinions are ours.

†Insee-Crest. Corresponding author. Address: 88 avenue Verdier, 92120 Montrouge, France. Phone: (+33)187695944. Email: pierre.pora@insee.fr

‡Insee-Crest. Email: lionel.wilner@insee.fr

1 Introduction

Recent research has highlighted that women's earnings losses due to motherhood, referred to as child penalties, have become the main driver of gender inequality in the labor market in developed countries (Kleven, Landais, and Søgaard, forthcoming; Juhn and McCue, 2017). From this perspective, the key question therefore concerns the channels that generate such child penalties. On the supply side of the labor market, two plausible settings may entail such earnings losses. On the one hand, the arrival of a child is likely to emphasize within-household *specialization* in the labor market *versus* housework ("home production"): by increasing the need for home production, and especially childcare, the arrival of a child would make women's comparative advantage in home production more salient and thus lead to decreases in female labor supply and labor earnings. On the other hand, households can change their time allocation when they have children, either because women have an intrinsically stronger preference for childcare (*preferences*) or because households penalize choices that deviate from traditional gender roles that are followed by their peers (*norms*).

Disentangling the influences of specialization, preferences and norms is crucial to the design of public policies that aim at promoting gender equality because different responses to policy instruments are expected depending on the prevailing channel. For instance, the usual policy tools that target work disincentives for mothers either by reforming the allocation of family-related benefits or by expanding the provision of external childcare are likely to reduce child penalties and thus gender inequality in the labor market if the specialization channel dominates. Conversely, if the preferences and norms channel dominates, reforms that aim at changing the way households perceive gender roles, such as the introduction of paternity leaves or/and parental leaves with mandatory splitting, may have a greater impact.

This paper addresses the above issues by contrasting the effect of the arrival of a child across groups of workers who face different trade-offs with respect to the labor market *versus* home production and childcare. Under the specialization hy-

pothesis, the responses of the respective workers to the arrival of a child should be highly heterogeneous, exhibiting a pattern that the preferences and norms channel is less likely to account for. Specifically, the specialization hypothesis predicts that women with the largest returns on time spent in the labor market are much less likely to reduce their labor supply due to children. The financial incentives related to the labor market *versus* home production and childcare trade-off are therefore driven by the opportunity cost of time, a key parameter in the time allocation problem faced by households. We approximate this cost by relying on potential hourly wages, which we measure *before* the arrival of a child. As a result, in this paper, we document the heterogeneity of the consequences of childbirth along the distribution of pre-childbirth wages.

We consider the short-run (one-year) to medium-run (five-year) impacts on several labor outcomes: total labor earnings, hourly wages, and labor supply at both extensive and intensive margins of employment. Our empirical strategy embeds a difference-in-difference method within a nonparametric ranking of individuals along the hourly wage distribution *à la* Guvenen et al. (2016); the latter framework aims precisely at depicting heterogeneity in individual labor market trajectories along the wage distribution. Our treatment group consists of parents with n children, while our control group contains parents with exactly $n-1$ children. We provide additional evidence to support the validity of our approach. First, the identification of child penalties is likely affected by neither measurement errors nor potential endogeneity of fertility decisions with respect to labor outcomes. Second, we show that the correlation between child penalties and hourly wages reflects rather a causal mechanism than heterogeneity in norms and preferences. To rule out the possibility of this heterogeneity being a confounding factor, we rely on survey data and find that preferences towards childcare exhibit only limited heterogeneity along the wage distribution. Furthermore, our results are robust to controlling for childcare-preferences driven human capital accumulation decisions that may affect pre-childbirth hourly wages.

We apply this method to French administrative data, namely, the DADS panel,

a comprehensive linked employer-employee dataset[1] that covers the period from 2005 to 2015 (a stable period in terms of policy changes) and contains information on individuals' labor earnings and paid hours. This panel is merged with the census data from the EDP, including longitudinal birth and marriage records at the individual level. Due to the richness of the dataset, we are able to consider the above-mentioned control and treatment groups at specific locations of the hourly wage distribution.

Our main results are as follows: (i) high-wage women experience much smaller labor earnings losses due to childbirth than do their lower-paid counterparts and (ii) are much less likely to interrupt their careers or reduce their paid hours; (iii) importantly, the magnitudes of the latter two effects exhibit monotonic behavior along the hourly wage distribution, and (iv) hourly wage losses appear rather homogeneous along the pre-childbirth hourly wage distribution. We relate the observed monotonic patterns to the increasing opportunity cost of time spent outside the workforce and/or increasing returns on experience along the wage distribution. These results strongly suggest that the specialization channel is at play: mothers, whose financial incentives to remain in the workforce are the strongest due to a high hourly wage and hence to a high cost of career interruption, are very unlikely to opt out of working or reduce their working hours. Conversely, those with much larger work disincentives because current hourly wages make child benefits that compensate for a reduction in labor supply worth considering (additionally, e.g., at the minimum wage level career interruptions do not significantly affect future career prospects) are likely to leave the workforce or at least reduce their labor supply at the intensive margin.

Overall, by stressing the importance of the specialization channel as opposed to the role of norms and preferences, these results suggest that public policies aimed at increasing incentives for women to remain employed after childbirth are instrumental in reducing gender inequality in the labor market. The results draw special attention to financial incentives generated by parental leave allowances; e.g., those studied by Piketty (2005); Lequien (2012); Joseph et al. (2013) provide

[1] Filling out the DADS form is a mandatory part of the process of paying payroll taxes.

parents who interrupt their careers or reduce their working hours to take care of their children with a fixed income that does not depend on their pre-childbirth hourly wage. By reducing the opportunity cost of career interruptions for low-wage mothers, such incentives may lead part of them to get stuck between low labor force participation and low hourly wages. Our results additionally imply that increasing the provision of external childcare and reducing its cost are likely to be efficient in reducing gender gaps in the labor market.

Literature This paper is primarily related to the vast literature devoted to the impact of fertility decisions on labor market outcomes. Childbirths tighten time constraints and shift women's labor supply and labor market outcomes, which helps explain a substantial share of the gender pay gap as shown by, e.g., the seminal contributions on the "motherhood penalty" by Waldfogel (1995, 1997, 1998). Recent empirical evidence suggests that motherhood not only explains a large part of the gender gap in labor earnings but also accounts for a growing share of this gap in developed countries (Kleven, Landais, and Søgaard, forthcoming). More generally, childbirths have been shown to explain a significant share of the aggregate gender gap, though there is no consensus on the exact share or whether this contribution is increasing over time (Bertrand, Goldin, and Katz, 2010; Wilner, 2016; Adda, Dustmann, and Stevens, 2017; Juhn and McCue, 2017; Kleven, Landais, and Søgaard, forthcoming).

Given these findings, identifying the channels that generate such child penalties is a key issue. While the existence of some employers that practice discrimination against mothers cannot be dismissed easily, the most likely channels involve the supply side of the labor market. It is plausible that the most prominent contribution to child penalties stems from children-induced career interruptions and adjustments in labor supply, which results in human capital depreciation (Meurs, Pailhé, and Ponthieux, 2010; Ejrnæs and Kunze, 2013; Adda, Dustmann, and Stevens, 2017). Other channels involve reduction in work effort (Becker, 1985; Hersch and Stratton, 1997) and mothers having a strong preference for time flexibility (Goldin, 2014), which in turn generates compensating wage differentials or

leads mothers to work in family-friendly firms that are likely to exert monopsony power (Coudin, Maillard, and Tô, 2018).

As to the causes of such decisions, two views can be contrasted. The first builds on the model of time allocation proposed by Becker (1981), based on the comparative advantage between the labor market and home production, i.e., on specialization. The second view, related to preferences and norms, refers to the identity model of Akerlof and Kranton (2000) and suggests that childbirth enhances the perception of oneself and her spouse as belonging to one gender or another, which distorts households' time allocation decisions in the sense that is compatible with gender-specific prescriptions.

Disentangling these two channels is an empirical task. The first strategy requires a structural model of fertility and labor supply similar to that of Adda, Dustmann, and Stevens (2017) even if the identity channel is absent from the researchers' framework. The second relies on policy changes that affect exogeneously either the labor market vs. home production trade-off (to identify the specialization channel) or gender identity (to identify the norms and preferences channel). In Austria, Kleven et al. (2019a) take advantage of various parental leave reforms and childcare expansions, and provide evidence that such reforms and expansions do not lead to substantial changes in the long-run consequences of women's fertility decisions; the authors conclude that gender norms and preferences are the prominent channel. The latter strategy contrasts child penalties across groups of individuals exposed either to heterogeneous labor market vs. home production trade-off or to different gender prescriptions. On the one hand, Angelov, Johansson, and Lindahl (2016) connect the impact of childbirth on labor earnings to within-couple pre-childbirth gender gap, and stress the specialization channel. Following Goldin (2014), Bütikofer, Jensen, and Salvanes (2018) compare child penalties across occupations among top earners. Their analysis quantifies the contribution of nonlinear wage structures to child penalties. On the other hand, Steinhauer (2018), Nix and Andresen (2019) and Kleven et al. (2019b) rely either on heterogeneity across linguistic groups in Switzerland, on differences between same-sex and heterosexual couples or on cross-country comparisons to emphasize

the role of gender norms.

All of the above empirical strategies rely on an evaluation of the causal impact of parenthood on labor outcomes, which requires overcoming the issue of endogeneity of fertility decisions. For instance, among the few papers that attempt this, Lundberg and Rose (2000) rely on twin sisters, and Miller (2011) exploits biological fertility shocks as an indicator of age at first birth. However, Kleven, Landais, and Søgaard (forthcoming) emphasizes that, empirically, correcting for potential endogeneity does not make too much of a difference: the causal effect of the third childbirth, estimated by using sex-mix instruments, does not differ much from an OLS estimate based on an event-study approach. In this paper, we rely to some extent on this result to advocate for our difference-in-difference strategy. We develop additional tests, especially one based on the study of the impact of job displacement on fertility decisions by Huttunen and Kellokumpu (2016). This enables us to show that endogeneous fertility decisions likely do not affect our empirical strategy.

By contrasting child penalties among individuals characterized by their pre-childbirth ranks in the wage distribution, this paper is also related to the few studies that have investigated the distributional impact of the arrival of a child. A small body of sociological literature has been devoted to this question, following Budig and Hodges (2010). Due to methodological issues regarding the interpretation of quantile regression coefficients, it, however, remains difficult to identify the main lessons from this literature (see Killewald and Bearak, 2014; Budig and Hodges, 2014; England et al., 2016). Among economists, Ejrnæs and Kunze (2013) rely on policy changes in Germany to estimate the impact on wages due to one additional year spent outside the labor market due to childbirth; the researchers observe that such wage losses are far more substantial for the most highly skilled mothers. While this question is close to but different from the impact of childbirth *per se*, these findings can easily be reconciled with ours.

Lastly, this paper is relevant to the analysis of heterogeneity of the gender pay gap along the wage distribution (e.g., Albrecht, Björklund, and Vroman, 2003; Arulampalam, Booth, and Bryan, 2007; Gobillon, Meurs, and Roux, 2015). In

particular, Fortin, Bell, and Böhm (2017) points out that vertical segregation, i.e., women being underrepresented at the very top of the distribution, can account for a large share of the aggregate gender gap in earnings. Our results suggest that while child penalties may well contribute to this underrepresentation at the top, it is not the main explanation: child penalties are, if anything, smaller at the top of the distribution. Albrecht, Thoursie, and Vroman (2015) argue that the generosity of the Nordic parental leave system causes employers to place fewer women in top positions (cf. statistical discrimination, Phelps, 1972). In a self-confirming belief equilibrium where women opt for family-friendly jobs, both vertical segregation and motherhood penalties prevail.[2] A somewhat similar argument is provided by Datta Gupta, Smith, and Verner (2008), who suggest that generous parental leave policies may result in a welfare state-based glass ceiling in Nordic countries.

The rest of the paper is organized as follows. The next section presents our data and the institutional setting. In section 3, we describe our empirical approach. Section 4 presents our results; section 5 discusses the validity of our identification strategy, and section 6 concludes the paper.

2 Data and institutional background

2.1 Data

Our analysis is based on a large panel of French salaried employees, namely, the longitudinal version of the *Déclarations Annuelles de Données Sociales* (DADS). By law,[3] French firms have to fill out the DADS form – an annual form that is the analogue of the W-2 form in the US – for every employee subject to payroll taxes. As of 2002, the panel contains information on individuals born on January 2-5, April 1-4, July 1-4 and October 1-4; these (more or less) first four days of each quarter correspond to the birthdays of individuals for whom we obtain census records in addition to labor market characteristics (see *infra*). This panel

[2]In practice, the Swedish glass ceiling tends to be higher at the top and to increase with age, which is consistent with the previous argument. Moreover, one half of the gender pay gap is present before the first childbirth as if the two explanations were equally important.

[3]The absence of DADS as well as incorrect or missing answers are punished with fines.

is therefore a representative sample of the French salaried population at rate 4.4%. Because of the comprehensiveness of the panel with respect to individuals' careers, the data is of exceptional quality and has low measurement error in comparison with survey data, in addition to a large sample size and no top-coding.

The database contains detailed information about gross and net wages, days worked, paid hours, other job characteristics (the beginning, duration and end of a period of employment, seniority, and part-time employment), firm characteristics (industry, size, and region) and individual characteristics (age and gender). We are also able to recover the numbers of male and female employees at each firm by resorting to the cross-sectional version of the DADS to this end and using the linked employer-employee dataset (LEED) dimension. Our main variables of interest are (i) net real annual labor earnings defined as the sum of all salaried earnings over all employers, (ii) time worked, measured as the number of paid hours as well as the number of days worked, and (iii) hourly wages defined as the ratio of annual earnings and time worked. In Appendix A, we provide some further details on the measurement of earnings and time worked. The main point is that, with few exceptions, (i) maternity leave allowances paid by social security are not included in our measure of earnings; (ii) the duration of maternity leave in days corresponds to a positive number of days worked; (iii) the number of hours worked during the maternity leave is equal to 0, and (iv) the number of hours worked (resp., hourly wages) is overestimated (resp., underestimated) for workers that are not paid by the hour in years in which they take maternity leave.

Individuals are identified by their NIR, a 13-digit social security-like number that enables the researcher to merge the DADS panel with *Échantillon démographique permanent*. The latter is a longitudinal version of the census that includes births and marriage registers as of 1968. However, information on childbirth is missing before 2002 for individuals born in January, April or July. For this reason, we consider first individuals born on October 1-4. Additionally, some childbirth-related data is available in administrative birth registers for individuals born October 2-3; however, it was incomplete during the 1990s (for details, see Wilner, 2016): as a

result, for these individuals we rely on the census rather than birth records.[4] Finally, partial data on education is available in this dataset (see Charnoz, Coudin, and Gaini, 2011) that indicates the highest degree obtained at the end of studies.

Our working sample is composed of salaried male and female employees in the private sector with the exclusion of agricultural workers and household employees. We restrict our analysis to individuals aged 20 to 60 living in metropolitan France between 2005 and 2015.

The empirical analysis described in Section 3 requires selecting individuals with a strong attachment to the labor market. We specify that these individuals be employed in the private sector for at least two years between $t-5$ and $t-2$ in addition to being present in $t-1$.[5] To deal with individuals with very low labor participation, an individual is considered employed at t if her paid hours exceed 1/8 of the annual duration of work (1,820 hours as of year 2002), if her total employment duration exceeds 45 days per year and if her hourly wage exceeds 90% of the minimum wage. We also winsorize labor earnings at the quantile of order 0.99999 to avoid outliers. We exclude individuals for which one observation has the ratio of net labor earnings to gross labor earnings less than (resp., greater than) 1/100 of (resp., 100). Our working sample has approximately 1.9 million individuals-years of observations, corresponding to nearly 270,000 workers. Appendix C provides summary statistics on the sample selection process both in terms of labor outcomes and in terms of fertility decisions.

2.2 Institutional background

Family-friendly policies in France have a long-lasting history (see Rosental, 2010) that dates back at least to pro-natalist concerns during the interwar period (Huss, 1990). These policies rely on (i) tax cuts, especially the *quotient familial* introduced in 1945, whereby the income tax rate depends on the number of children

[4] Appendix B explains how we recover such data, the quality of which is comparable with that of individuals born October 1 or 4 for whom birth records are available.

[5] The core results of this paper rely on years t from 2005 to 2015. As a result, because data are only available from 2002, the inclusion condition is slightly stronger for years 2005 and 2006. However, dropping these years and focusing only on years 2007 to 2015 does not change our estimates, as shown in Figures F.9 and F.10.

in a household, (ii) various child benefits, and (iii) some other welfare benefits, such as bonuses included in retirement pensions that depend on realized fertility, or housing allowances. In France, income is taxed jointly within households; this scheme is the source of strong incentives towards within-household specialization.

Maternity leaves were created in 1909; they were first unpaid, and subsequently became fully covered up to some threshold for all salaried workers by social insurance from 1970 onwards. Since 1980, the arrival of the first two children granted a woman a 16-week maternity leave consisting of 6 weeks before childbirth and 10 weeks after. Starting from the arrival of the third child, the total duration becomes 26 weeks (8+18), and maternity leave duration may increase to 46 weeks in the case of multiple births. Maternity leaves also have a minimum duration of 8 weeks, consisting of 2 weeks before childbirth and 6 weeks after.

Paternity leaves came in force in 2002 in addition to birth leaves that amounted to 3 consecutive days following childbirth. Such a leave grants a father an 11-day leave that is fully covered, up to some threshold, by social insurance. Its duration can increase to 18 days in the case of multiple births but always includes weekends and public holidays. The idea of extending that duration has recently attracted some attention; the French government has asked for an internal *ex ante* evaluation, but no decision has been made yet.

In addition to the above leaves, there are various parental allowances that were merged in 2004 into the *Prestation d'Accueil du Jeune Enfant* (PAJE). It comprises a one-shot means-tested bonus at childbirth (*prime de naissance*), monthly means-tested benefits (*allocations familiales*), a childcare subsidy (*Complément libre choix du Mode de Garde* (CMG)), and some child benefits granted when parents interrupt their careers or work part-time (previously *Complément Libre Choix d'Activité* (CLCA) and now *Prestation Partagée d'Éducation de l'enfant* (PreParE)).

These child benefits date back to 1985 and appeared with the creation of *Allocation Parentale d'Éducation* (APE) initially restricted to mothers of 3 or more children. APE was extended to mothers of 2 children in 1994, and was replaced by the CLCA in 2004, becoming effective with the first childbirth and providing

a fixed not-means-tested amount for the maximum duration of 6 months. Lastly, CLCA was replaced in 2015 by PreParE that introduced incentives to split the leave between parents; it amounted to approximately €400 per month in the case of career interruption and to nearly €200 in the case of 80% part-time work. Several papers have shown that these benefits induce mothers to reduce their labor supply (Choné, Le Blanc, and Robert-Bobée, 2004; Piketty, 2005; Lequien, 2012; Joseph et al., 2013).

In contrast, other policies favor participation in the labor force by decreasing the cost of childcare; an example of such a policy is CMG that is not means-tested, and entails payroll tax cuts or income tax credits.[6] It is not straightforward to determine the exact scheme of financial incentives provided by such childcare subsidies because they depend on numerous dimensions (the type of childcare chosen among day nurseries, child-minder and nannies,[7] family structure and geographic location) but always depend on earnings in a way that makes mothers at the bottom of the wage distribution more likely to stop or reduce their activity (see, e.g., Givord and Marbot, 2015).

Considering the labor supply, the current family insurance scheme therefore provides contradictory incentives: on the one hand, PreParE should reduce labor supply after childbirth; on the other hand, CMG should preserve it. Determining which effect dominates is an empirical task; the answer to this question depends on the location in the wage distribution. Mothers at the top of the wage distribution will not be particularly responsive to PreParE since career interruption or, more likely, part-time employment is particularly costly for them. In contrast, the combination of PreParE benefits (€200) with a reduction of childcare expenditures is worth considering for mothers earning low wages: e.g., at the minimum wage (slightly above €1,200 per month), a switch to 80% part-time work means a wage reduction of approximately €240, and hence the net monetary loss of €40 only.

[6] A typical tax credit amounts to 50% of childcare expenditures up to some threshold that depends on the type of chosen daycare. The annual threshold is €2,300 for childcare providers or wet nurses, but it may be as high as €13,500 (€16,500 in the first year) for nannies employed at home.

[7] This very choice itself depends on parents' earnings; affluent households are more likely to opt for nannies, while poor households more often choose child-minders or day nurseries, though there is variation in this respect.

Hence, the current system including family allowances and childcare subsidies is more likely to make the "mommy track" all the more attractive if the mother is at the bottom of the wage distribution.

Lastly, other welfare benefits including bonuses included in pensions and housing allowances also depend on the number of children. Other family-friendly policies may be available within firms; e.g., employers may provide childcare services to employees. These firm-specific family policies can be subject to further tax reductions or credits, such as the *Crédit d'impôt famille* created in 2004.[8]

3 Empirical analysis

Our main outcome of interest is total annual labor earnings of individual i during year t; we denote such earnings by $\tilde{y}_{it}$. We decompose them into four components: d_{it} is a dummy variable for participation; $\tilde{x}_{it}$ represents the employment duration in days, and is between 0 and 360[9]; $\tilde{h}_{it}$ denotes the average number of paid hours per day during year t, and lastly $\tilde{w}_{it}$ is the average hourly wages of individual i during year t. Hence,

$$\tilde{y}_{it} = d_{it}\tilde{x}_{it}\tilde{h}_{it}\tilde{w}_{it} \tag{1}$$

3.1 Normalization

Providing estimates of the causal effect of childbirth by comparing parents and non-parents requires netting out other lifecycle effects as confounding factors; e.g., the number of childbirths an individual has experienced is a nondecreasing function of age. We choose to net out lifecycle and business cycle effects only; many other factors that determine labor outcomes could be adjusted in response to fertility decisions, and hence should be taken into account as part of child penalties instead of being controlled for. As a result, the first step of our empirical framework derived from that of Guvenen et al. (2016) consists of normalizing earnings and each of

[8] To the best of our knowledge, to date there has been no empirical evaluation of the effect of such policies.

[9] The number of days in a year is capped at 360 in DADS.

earnings' components with respect to age, cohort and period. Let $\tilde{z}$ denote either labor earnings or one of its components with the exception of the participation dummy. We start by regressing the logarithm of $\tilde{z}_{it}$ on a set of cohort (year of birth), age and period dummies. We estimate the following pooled cross-sectional regression:

$$\log(\tilde{z}_{it}) = \sum_c \lambda_c^z \mathbb{1}_{cohort_i=c} + \sum_a \mu_a^z \mathbb{1}_{age_{it}=a} + \sum_T \nu_T^z \mathbb{1}_{t=T} + \epsilon_{it}^z \tag{2}$$

The identification of age-period-cohort (APC) models can be achieved at the cost of normalizations.[10] In this particular paper, the choice of normalization is insignificant, given that we rely on the sum $\hat{\lambda}+\hat{\mu}+\hat{\nu}$ and never use these components separately.

Previous estimates enable us to define the normalized component z_{it} as

$$z_{it} = \frac{\tilde{z}_{it}}{\exp(\hat{\lambda}^z_{cohort_i} + \hat{\mu}^z_{age_{it}} + \hat{\nu}^z_t)} \tag{4}$$

An accounting decomposition similar to that of (1) is used for normalized earnings:

$$y_{it} = d_{it} x_{it} h_{it} w_{it} \tag{5}$$

[10]The major challenge in the simultaneous identification of λ, μ and ν stems from collinearity between age, cohort and period: age is equal to the current period less the year of birth. Several solutions have been explored in the sociological literature; e.g., Mason et al. (1973) assume that any two ages, periods or cohorts have the same effect, in addition to removing one dummy in each dimension. Deaton and Paxson (1994) and Deaton (1997) suggest a transformation of period effects to meet two requirements: (i) that time effects sum to zero, and (ii) that they are orthogonal to a time trend so that age and cohort effects capture growth while year dummies account for cyclical fluctuations (or business cycle effects) that average to zero over the long run. Hence, the parameters of the model (λ, μ, ν) are identified, provided that $\lambda_{\underline{c}} = 0$ and $\sum_{t=1}^{T} \nu_t (t-1) = 0$. The corresponding transformation of time dummies $d_T = \mathbb{1}_{t=T}$ is written as follows:

$$d_T^* = d_T - [(T-1)d_2 - (T-2)d_1] \tag{3}$$

where $d_1^* = d_2^* = 0$. In practice, it is convenient to include all age dummies, all cohort dummies but the first, and all transformed dummies d_T^* but d_1^* and d_2^* in the regression.

3.2 Ranks in the hourly wage distribution

Our empirical strategy embeds a difference-in-difference setting within a framework that aims at modeling heterogeneity in the consequences of childbirth along the hourly wage distribution. To this end, we rely on comparisons both within groups of workers with similar hourly wages and across these groups. Hence, our analysis relies on the definition of such groups based on a measure of recent hourly wages:

$$W_{i,t-1} = \frac{\sum_{\tau=t-5}^{t-1} d_{i\tau}\tilde{w}_{i\tau}}{\sum_{\tau=t-5}^{t-1} d_{i\tau}\exp(\hat{\lambda}^{w}_{cohort_i} + \hat{\mu}^{w}_{age_{i\tau}} + \hat{\nu}^{w}_{\tau})} \tag{6}$$

We compute this measure for individuals who participate in year $t-1$ and at least twice between years $t-5$ and $t-2$ (i.e., provided that $d_{i,t-1}\sum_{\tau=t-5}^{t-1} d_{i\tau} \geq 3$). Within each age × year cell, we rank workers according to their recent wages $W_{i,t-1}$. We use this ranking to create 20 cells: P0-P5, P5-P10, ..., P90-P95 and P95-P100. Hence, we assume that workers within each age × year × recent wage cell are, if not identical, at least *ex ante* similar with respect to their hourly wage levels before year t. Ranks are not conditional on gender: within these cells, men and women have approximately the same recent wages. As a result, women are more (resp., less) numerous at the bottom (resp., top) of the distribution, which merely reflects the existence of a gender gap in hourly wages (see Appendix C).

This depiction of heterogeneity along the wage distribution yields estimates that are conceptually different from those of a quantile regression approach. Such an approach would be based, by definition, on the rank in the (potential) outcome distribution, which is not what we want here. There is no particular reason for ranks in the wage distribution to coincide with ranks in the labor supply distribution unless strong assumptions are made. However, both approaches provide complementary insights: e.g., Albrecht, Thoursie, and Vroman (2015) observe that the effect of parental leave is higher at the top of the distribution.

3.3 Difference-in-difference strategy

Our estimates of the consequences of childbirth are based on a difference-in-difference approach. The endogeneity of fertility decisions is often regarded as

a key issue, but recent empirical results suggest that it is not a particularly significant problem (Kleven, Landais, and Søgaard, forthcoming). We discuss the plausibility of the assumption that fertility decisions are exogeneous, and devise additional tests of its validity in Section 5.

We define N treatments, where the nth treatment consists of experiencing the nth childbirth during year t. Our control group for the nth childbirth is composed of individuals of the same gender with $n-1$ children and who never had the nth child. Due to the omission ("right-censoring") of unknown but relevant data on fertility decisions taken after 2015, individuals belonging to the nth control group may experience the nth childbirth after 2015; we address this issue *infra*. In practice, we restrict our attention to the first three childbirths that represent 96% of childbirths. Year $t-1$ is regarded as the reference year; by construction, all individuals participate in the labor market during year $t-1$.

Due to multiple treatments, the same individual may be considered several times in our estimation, though at different dates, either as a member of a treated group or a control group. Proper inference has to take this issue into account; we therefore cluster standard errors at the individual level (Bertrand, Duflo, and Mullainathan, 2004).

This difference-in-difference approach is embedded in our ranking along the hourly wage distribution. Our control groups are therefore restricted to individuals with the same rank in the recent hourly wage distribution as our treated individuals. Moreover, the effect of childbirth is allowed to vary along that distribution of recent wages.

We now propose two distinct implementations of this approach.

3.3.1 Accounting framework

First, we rely on an accounting framework to provide estimates of effects of childbirth on labor market outcomes and labor supply. Our estimate of the impact of the nth childbirth on earnings k years after childbirth for individuals of gender g at rank r in the recent wage distribution is written as

$$\beta_{g,r}^{y,n,k} = \underbrace{\log\left(\frac{\mathbb{E}[y_{i,t+k}|b_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}{\mathbb{E}[y_{i,t-1}|b_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}\right)}_{\text{Treated}} - \underbrace{\log\left(\frac{\mathbb{E}[y_{i,t+k}|c_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}{\mathbb{E}[y_{i,t-1}|c_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}\right)}_{\text{Control}} \tag{7}$$

where b_{it}^n is a dummy for experiencing the nth childbirth during year t, c_{it}^n is a dummy for belonging to the nth control group at time t, i.e., having $n-1$ children at time t but never experiencing the nth childbirth according to the data, and $\mathcal{T}_k$ is the set of time periods for which $t-3$ to $t+k$ are observed in the data.

Considering the causal impact of childbirth $\beta_{g,r}^{y,n,k}$ being identified on a subset of time periods that depends on k, we assume that treatment effects are time-homogeneous, i.e., that having a k-year-old nth child bears the same consequences if the child was born in 2005 as it does if he or she was born in 2015. We discuss and assess the plausibility of this assumption, among others, in Section 5. Importantly, considering $k < -1$ allows us to verify that trends are parallel before childbirth.

The overall impact of childbirth on the gender gap in pay can be obtained directly as the difference between the impact on men's labor outcomes, and that on women's labor outcomes, both computed by the difference-in-difference method. It is thus written ass (omitting indices y, k, r, z for clarity)

$$\beta_{\text{gap}} = \beta_f - \beta_m \tag{8}$$

Decomposition (9) states that average normalized earnings growth can be represented as a sum of its four components, and a selection term due to the fact that individuals who participate in the labor market in year $t+k$ may not have the exact same past earnings $y_{i,t-1}$ as those who do not participate:

$$
\begin{aligned}
\underbrace{\log\left(\frac{\mathbb{E}\left[y_{i,t+k}\right]}{\mathbb{E}[y_{i,t-1}]}\right)}_{\text{Labor earnings changes}} &= \underbrace{\log\left(\mathbb{P}(d_{i,t+k}=1)\right)}_{\text{Participation}} \\
&+ \underbrace{\log\left(\frac{\mathbb{E}[y_{i,t-1}|d_{i,t+k}=1]}{\mathbb{E}[y_{i,t-1}]}\right)}_{\text{Selection}} \\
&+ \underbrace{\log\left(\frac{\mathbb{E}\left[x_{i,t+k}h_{i,t-1}w_{i,t-1}|d_{i,t+k}=1\right]}{\mathbb{E}\left[x_{i,t-1}h_{i,t-1}w_{i,t-1}|d_{i,t+k}=1\right]}\right)}_{\text{Employment Duration Changes}} \\
&+ \underbrace{\log\left(\frac{\mathbb{E}\left[x_{i,t+k}h_{i,t+k}w_{i,t-1}|d_{i,t+k}=1\right]}{\mathbb{E}\left[x_{i,t+k}h_{i,t-1}w_{i,t-1}|d_{i,t+k}=1\right]}\right)}_{\text{Hours-per-day Changes}} \\
&+ \underbrace{\log\left(\frac{\mathbb{E}\left[x_{i,t+k}h_{i,t+k}w_{i,t+k}|d_{i,t+k}=1\right]}{\mathbb{E}\left[x_{i,t+k}h_{i,t+k}w_{i,t-1}|d_{i,t+k}=1\right]}\right)}_{\text{Hourly Wage Growth}}
\end{aligned}
\tag{9}
$$

In Appendix D, we clarify the interpretation of this decomposition, showing that it can be rewritten in terms of expected values of changes in labor outcomes, up to some reweighting. This decomposition of labor earnings growth allows us to consider separately each component of the impact of childbirth on earnings; we write it as $\beta^y = \beta^s + \beta^d + \beta^x + \beta^h + \beta^w$, where β^s stands for the selection term, and the four other terms correspond to each component of labor earnings (for readability, we omit all other unnecessary indices). This decomposition is made in an accounting sense. A causal decomposition would require precise modeling of causal links between components of labor earnings, such as labor supply being provided on the basis of the wage rate offered in the labor market, and hourly wages depending on past labor supply, e.g., through human capital accumulation; however, such modeling is beyond the scope of this paper.

3.3.2 Regression framework

Second, we implement an alternate specification of the same approach. Here, the decomposition of earnings is performed at the individual rather than the aggregate level, and we are able to control for additional covariates to obtain a better sense of channels that lead to child penalties. Our estimate of the impact of the nth childbirth on earnings k years after childbirth for individuals of gender g at rank r in the recent wage distribution is now written as

$$\begin{aligned}\theta_{g,r}^{y,n,k} = & \underbrace{\mathbb{E}[\log(y_{i,t+k}) - \log(y_{i,t-1}) | b_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}_{\text{Treated}} \\ & - \underbrace{\mathbb{E}[\log(y_{i,t+k}) - \log(y_{i,t-1}) | c_{it}^n = 1, r_{it} = r, g_i = g, t \in \mathcal{T}_k]}_{\text{Control}}\end{aligned} \tag{10}$$

Given a component z[11] of decomposition (5), we consider its growth between $t-1$ and $t+k$, determined as $\delta^k z_{it} = \log(z_{i,t+k}) - \log(z_{i,t-1})$ and defined for all individuals working in the private sector in year $t+k$. We estimate the following $(k+1)$-difference regression by OLS for each component z, gender g, rank r and arbitrary duration k (we omit indices g, k, r, z for clarity):

$$\delta^k z_{it} = \alpha + \sum_n \gamma^n (b_{it}^n + c_{it}^n) + \sum_n \theta^n b_{it}^n + \zeta X_{it} + u_{it} \tag{11}$$

where X_{it} is a vector of either invariant or time-varying covariates, and u_{it} is an idiosyncratic error term.[12]

Parameters of interest, namely, treatment effects θ, tell us how parents' out-

[11] When focusing on participation, we use $d_{i,t+k}$ as the outcome. By construction, $d_{i,t-1} = 1$; hence, $d_{i,t+k}$ accounts for changes in labor supply at the extensive margin.

[12] An alternate specification of the same regression is

$$\delta^k z_{it} = (\alpha + \alpha_{\text{gap}} g_i) + \sum_n (\gamma^n + \gamma_{\text{gap}}^n g_i)(b_{it}^n + c_{it}^n) + \sum_n (\theta^n + \theta_{\text{gap}}^n g_i) b_{it}^n + (\zeta X_{it} + \zeta_{\text{gap}} X_{it} g_i) X_{it} + u_{it}$$

Here, θ^n corresponds to the impact of childbirth on fathers' labor outcomes, while θ_{gap}^n gives us information on how mothers' outcomes shift with respect to those of fathers. Once again, the comparison holds for individuals with similar recent hourly wages, as we have already controlled for the gender-based divergence among nonparents; hence, it measures directly how childbirths contribute to the gender gap in labor outcomes. In this respect, θ_{gap}^n results from a triple-difference estimation.

comes change k years after childbirth with respect to parents' siblings, i.e., non-parents of the same gender and with similar hourly wages. Interestingly, our approach enables us to determine non-parametrically the impact of childbirth that varies along the entire recent wage distribution.

4 Results

4.1 Heterogeneous consequences of childbirth

First, we assess the consequences of childbirth on labor outcomes of men and women by relying on the accounting framework. Our estimates of the impact of the first three childbirths on individuals' total labor earnings are shown in Figure 1 for women and in Figure 2 for men. We plot those estimates for $t+k \in \{t-3, ..., t+5\}$ with the exception of $t-1$ since it is the reference year (our estimates are hence all equal to zero for that year).

Mothers experience large earnings losses after childbirth relative to women who earned similar hourly wages a few years before. All components contribute to these losses: after the arrival of a child, mothers are more likely to leave employment, work fewer days, work fewer hours per day and earn lower hourly wages than are women belonging to our control groups. Nevertheless, in the short to medium run, labor supply decisions seem to be driving these large earnings losses. Moreover, the impact of childbirth on women's labor outcomes increases in magnitude with the rank of the child. This empirical evidence is consistent with previous findings in the literature.

More interestingly, children-related earnings losses display substantial heterogeneity: low-wage women experience far larger earnings losses than do high-wage women. At the very bottom of the distribution, women's earnings losses amount to 70 log-points the year women first give birth, and 37 log-points one year after childbirth, and remain at 47 log-points 5 years after the arrival of a child.[13] In contrast, women ranked in the top 5% of the hourly wage distribution experience

[13] By definition and by law, year t includes a mixture of both maternity leave and employment periods, as discussed in Subsection 2.1.

earnings losses of 22 log-points, 9 log-points and less than 5 log-points, respectively. The main result is that child penalties decrease along the wage distribution as pre-childbirth hourly wage increases.

The decomposition of annual earnings growth into each of its components helps clarify the channels that contribute the most to this pattern. Previous heterogeneity is primarily driven by labor supply decisions at the extensive margin: childbirth reduces by 15 log-points (resp., 60 and 85 log-points) the probability that women are employed one year after the arrival of their first (resp., the second and the third) child at the bottom of the distribution but does not actually reduce this probability in the top 5% of the distribution. Once again, a monotonic behavior of labor supply responses is observed in the ranking within the hourly wage distribution; this striking monotony suggests that financial incentives matter, and gives some credit to the specialization channel as opposed to the preferences and norms channel.

Conversely, while hourly wage losses exhibit a U-shaped pattern along the distribution during the year of childbirth that may be driven by some problems in the measurement of hours during maternity leaves for workers that are not paid by the hour, who are more numerous in the upper part of the hourly wage distribution (see Subsection 2.1 and Appendix A), the pattern of motherhood wage penalties appears much more homogeneous one to five years later; the penalties amount to approximately 5 log-points for the first child, and even less for subsequent children.[14]

A nice feature of this approach is that it enables us to verify that trends of the treated and control groups before treatment are parallel, an assumption upon which the difference-in-difference methodology rests. Under this assumption, there should be no difference between treated and control groups before $t-1$. Formally, this assumption is rejected by the data: we observe small differences between groups' earnings in years $t-3$ and $t-2$ with respect to year $t-1$. The difference is slightly positive (resp., negative) when considering the arrival of the first (resp., the second) child: mothers had slightly slower (resp., faster) earnings growth than

[14] Additionally, this U-shaped pattern may be due to the institutional setting: the maternity leave compensation scheme involves various thresholds and depends on its duration.

did non-mothers (resp., mothers of one child) prior to the first (resp., the second) childbirth. However, these differences are less than 10 log-points, which is not much in comparison with earnings differences after childbirth (up to 130 log-points). More importantly, these differences vary little along the wage distribution, which is reassuring as far as the identification of heterogeneity of the impact of childbirth on women's labor outcomes is concerned. In Section 5, we nevertheless discuss the credibility of the parallel trend assumption *post-treatment*, which is crucial to our identification strategy.

When it comes to men, our estimates suggest that childbirths increase labor earnings slightly, especially through higher participation and hourly wages. The increase in participation is slightly more pronounced for fathers at the top of the wage distribution.

Additionally, Figure 3 displays our estimates of the impact of the first childbirth on the gender gap in labor earnings, participation, hours worked and hourly wages for individuals who belong to either the very bottom or the very top of the distribution. These are merely triple-difference estimates, i.e., the effect estimated for women minus that estimated for men. In particular, the figure makes it very clear that (i) the gender gap in earnings widens much more at the bottom than at the top of the wage distribution, and (ii) this pattern is nearly entirely driven by differences in the impact on participation rather than on other components of labor earnings. However, it remains difficult to assess whether these differences persist in the long run, as standard errors become large past the first seven years after the arrival of the first child due to the small sample size.[15]

4.2 Motherhood penalties and fatherhood premia five years after childbirth

We now focus solely on the impact on hourly wages five years after childbirth and rely here on our difference-in-difference approach to compare hourly wage growth

[15] In Appendix E, we investigate the impact on hourly wages in the longer term; we observe that it is plausible for hourly wage penalties to be persistent along the entire distribution, even though high-wage women may catch-up some of their wage losses.

between years $t-1$ and $t+5$ for individuals who experienced childbirth during year t with that of individuals belonging to the appropriate control group. Our choice to focus on year $t+5$ stems from the measure of hourly wages at time t being biased due to the incorrect measurement of hours worked during maternity leaves, which results in spurious patterns in the data that we do not want to confound with the effect of childbirth (see Subsection 2.1 and Appendix A). Figures 4 and 5 display our estimates for women and men.

The arrival of the first child has a negative and significant impact on women's hourly wages five years after her birth that is approximately -5 log-points for the largest part of the distribution (Model 1). The difference in the effect along a large part of the distribution is not significant: the effect is mostly homogeneous. However, both ends of the distribution represent exceptions: the consequences of childbirth are slightly less harsh for women that earned either low or high wages before childbirth. Controlling for horizontal segregation (by occupation, industry and firm composition) does not alter the estimates significantly (Model 2). However, controlling for experience, mobility and career interruptions lowers the effect (Model 3), which confirms that post-birth labor supply decisions are a key driver of motherhood penalties, implying, e.g., less human capital accumulation. Having a second child does not lead to statistically significant motherhood penalties compared to those faced by mothers of one child, but the confidence intervals are large and we cannot reject the hypothesis that it generates economically significant wage losses.

In the upper half of the distribution, fathers experience faster hourly wage growth after the first childbirth than do their childless counterparts (Model 1); this fatherhood premium amounts to 6 log-points at the very top of the distribution. Controlling for horizontal segregation (Model 2) as well as for experience accumulation and job mobility (Model 3) attenuates the estimates, which suggests that faster human capital accumulation due to increased labor supply may be at play. The second childbirth does not generate significant fatherhood premia with the exception of the very top of the distribution; once again, the confidence intervals are large, thus we cannot reject economically significant effects.

The consequences of childbirths on the gender pay gap are estimated by triple difference and displayed in Figure 6. Consistently with previous results, the first childbirth has a significant and negative impact on the gender pay gap: it widens this hourly wage differential for all workers but those belonging to the lowest part of the distribution. Our estimates suggest that the first childbirth leads to a larger gap among top-earners, of up to 8 log-points five years after birth, than among workers with lower wages, for whom the effect amounts to 5 log-points only.

We also document how the effect of covariates varies along the distribution: Figure 7 displays our estimates of coefficients related to actual experience,[16] career interruptions,[17] job mobility[18] and firm composition[19] in Model 3. Hourly wage growth is much more positively (resp., negatively) correlated with experience (resp., career interruptions) at the top of the wage distribution, which is consistent with Dustmann and Meghir (2005) who observe that returns to experience are higher among skilled workers.

These results are consistent with Ejrnæs and Kunze (2013), who observe that one additional year spent outside the labor market is much more detrimental to the most highly skilled mothers, and with Adda, Dustmann, and Stevens (2017) who show that human capital depreciation due to time spent outside the labor market is much more pronounced in abstract occupations. Our estimates should be interpreted with caution since, in contrast to Ejrnæs and Kunze (2013) who rely on policy changes to identify the effects of exogeneous decisions, here experience and past career interruptions reflect past labor supply decisions that were made based on expected future wages. Nevertheless, our estimates indeed convey some information. First, this substantial heterogeneity helps rationalize why low-earning women do not encounter larger hourly wage penalties while being more

[16] As opposed to *potential* experience. Experience is computed as the sum of hours worked between years t and $t+5$ divided by the median duration of work for individuals employed full time over one year without interruptions (namely, 1820 hours).

[17] Career interruptions are proxied by a dummy variable for spending at least one year between t and $t+5$ outside employment in the private sector.

[18] Job mobility is measured by a dummy variable for having different main employers at times $t-1$ and $t+5$. The main employer is the firm that pays the worker the highest labor earnings during a year.

[19] Firm composition is measured by the share of women working part-time among employees of the firm that employed individual i at time $t-1$.

likely to reduce labor supply upon the arrival of a child.[20] Second, it concurs with the argument based on the opportunity cost of career interruptions (and its heterogeneity along the distribution), which is key to understanding mothers' labor supply decisions. Forward-looking mothers base such decisions not only on their current wages but also on their expected future wages that are more contingent on current labor supply decisions at the top of the distribution. Moreover, these decisions depend also on the financial incentives provided by maternity leave allowances and childcare subsidies. Overall, the heterogeneity at stake here appears to be mostly consistent with mothers adjusting their labor supply according to financial incentives.

Finally, switching from one firm to another coincides with hourly wage gains of workers at the bottom of the distribution and with negative wage growth at the top. Workers of firms with high shares of women working part-time tend to experience slower hourly wage growth, which might indicate that the sorting dimension investigated by Card, Cardoso, and Kline (2016) or by Coudin, Maillard, and Tô (2018) affects hourly wages not only in levels but also in terms of progression.

5 Threats to identification

In this section, we address various factors that could affect the empirical validity of our identification strategy. These factors stem first from the fact that our treated and control groups are defined based on realized fertility observed in 2015, which creates a right-censoring issue; we also investigate other sources of measurement error. The second issue has to do with the endogeneity of fertility decision with respect to potential labor outcomes. The last concern is the nonrandom assignment to pre-childbirth wage groups based on expected children-related labor supply decisions; this last source of endogeneity would not affect our estimates of the heterogeneity of child penalties *per se* but may rather affect their causal interpretation, and thus puts in question our ability to disentangle the specialization

[20] Another institutional explanation for such heterogeneity is merely the high level of the minimum wage in France; as a result, hourly wage losses at the bottom of the distribution can never be too large.

channel from the influence of preferences and norms.

5.1 Right-censoring and measurement error

Our definition of control and treatment groups, despite being practical, raises some issues. First, due to right-censoring, individuals in our control group are not of the same age as those in our treatment group. Second, and for the same reason, our treatment effect estimate corresponds to the difference in labor market outcomes between parents with k children and individuals with $k-1$ children over the lifetime; this is true for old cohorts, but our estimate for younger cohorts might be spuriously affected by selection bias, namely, differences between parents of k children who experience childbirths quite early and parents who eventually have k children but do so later in life. Third, the definition of our treatment as experiencing the kth childbirth during year t might be blurred by the timing of labor supply decisions mainly because women are entitled a maternity leave that begins several weeks before childbirth and ends several months after. Choosing year $t-1$ as a reference for labor market outcomes is therefore likely to lead to biases with respect to childbirths that occur in the very beginning of the year since part of the childbirth effect might already have happened.

We address all three issues by providing several estimations based on alternative definitions of control and/or treatment groups:

1. We define our nth control group as individuals that experience $n-1$ childbirths according to the data, as of age randomly drawn from the empirical distribution of age at the nth childbirth within education × cohort cells. This allows us to assess robustness with respect to the age difference between control and treatment groups (see Figures F.1 and F.2).

2. We restrict our analysis to individuals born in 1975 or before: such individuals are most likely to have made all of their fertility decisions by year 2015 (see Figures F.3 and F.4).

3. We define our nth control group as individuals who have n children according to the data as of time t and do not experience any childbirths between $t-1$

and $t+k$ (see Figures F.5 and F.6). This strategy is closer to that of Kleven, Landais, and Søgaard (forthcoming) in that it relies on the timing of the nth childbirth among those who indeed have n children.

4. We restrict our nth treatment group to individuals who experience the nth childbirth during the second quarter, i.e., between April and June: their maternity leaves do not begin before January and do not end after December (see Figures F.7 and F.8).

Our findings prove robust to these alternative definitions.

Additionally, our measure of the causal impact of childbirth rests on the assumption that treatment effects are time-homogeneous, i.e., that childbirths occurring in 2005 have the same causal impact as those that occurred in 2015 if they are considered after the passage of the same amount of time. This assumption justifies reliance of our estimates on a time-varying window: the impact of childbirth at time t is estimated on all childbirths from 2005 to 2015, while our estimate of the impact of childbirths at time $t+5$ only relies on childbirths that occurred before 2011. The credibility of this assumption rests on our choice to focus on a stable period in terms of family policy changes: the PAJE reform took place in 2004, and the only other change in parental leave rules, which was a slight change in the incentives to split parental leave between parents, happened in 2015 and thus only applied to a small part of the sample. Nevertheless, we replicate our analysis while restricting it to childbirths during 2007-2010; hence, this compositional change does not distort our estimates of dynamic treatment effects: all treatment effects for all durations of time to childbirth are computed on the very same sample. Additionally, by choosing 2007 as the beginning of the estimation timespan, we can ignore the fact that the selection condition in our sample is harsher for childbirths in 2005-2006 due to left-censoring issues in the data. Figures F.9 and F.10 display our estimates and show that our approach is completely robust with respect to these concerns.

5.2 Selection into treatment

5.2.1 Comparing treatment effects between subpopulations

In the presence of heterogeneous treatment effects and under the parallel trend assumptions, i.e., assuming that the difference in average *actual* outcomes for the control group between the pre-treatment period and the post-treatment period is equal to the difference in average *potential* outcomes in the absence of treatment for the treated group, this strategy yields the average treatment effect on the treated (ATT). As a result, the estimates provided by such an approach within two subpopulations may differ due to two distinct channels: (i) in the labor market, low-wage women may experience more detrimental consequences of having had children, and (ii) high-wage women may suffer from the same detrimental consequences of fertility; however, among them, those who face the largest career costs would choose not to have children. Ruling out the second channel requires either assuming that treatment effects are homogeneous within groups or that treatment effects be mean-independent of actual fertility decisions within groups. In both cases, our approach would yield average treatment effects (ATE) on the whole population, which would ensure that the comparison across wage groups is relevant.

Each of these assumptions is neither testable nor plausible. However, assuming that the distribution of treatment effects is constant across wages groups (i.e., as if treatment effects were heterogeneous but independent of the pre-childbirth rank in the wage distribution), selection into treatment would imply that high-wage women are much less likely to experience childbirth than their low-wage counterparts. We can therefore assess the plausibility of this assumption by computing the probability of giving birth to the nth child during year t among those eligible, i.e., those who already have $n-1$ children in year $t-1$, along the entire recent wage distribution. Figure 8 shows that among women, the probability of giving birth does not significantly vary along the wage distribution, and that high-wage women are, if anything, in fact more likely to have children than are their low-wage counterparts. As a result, this kind of selection into treatment does not seem to

be a significant problem here.

5.2.2 Endogeneity of fertility decisions

A second kind of selection into treatment could also affect our results. Here, the challenge would be not merely to compare ATTs across subgroups rather than ATEs but related to the very possibility of identifying ATTs (including those on the whole population). This could be possible if individuals made their fertility decisions based on an unobserved shock common to both potential treated and untreated labor outcomes, and not on the difference between potential treated and untreated labor outcomes, as we investigated before. For instance, if women expected large earnings losses, especially due to dismissals or cuts of the number of paid hours, to occur in the near future, women would be more likely to have children. This kind of endogeneity has long been investigated by the maternal labor supply literature that has resorted to using various instruments to identify the consequences of exogeneous fertility decisions (see, e.g., Rosenzweig and Wolpin, 1980; Korenman and Neumark, 1992; Angrist and Evans, 1998). This possibility leads to a violation of the parallel trend assumption post-treatment[21] and in our case it would lead us to inflate the detrimental consequences of children.

In the absence of plausible exogeneous shocks to fertility decisions, there is no simple way of quantifying this potential source of bias. However, a recent empirical study by Kleven, Landais, and Søgaard (forthcoming) investigates this issue and observes that, for the third childbirth, child penalties estimated through simple event studies do not differ from those obtained by using a sex-mix instrument. Additionally, if high-wage women respond to expected future shocks to their labor outcomes the same way as low-wage women do, this source of bias would be constant along the distribution, and would not affect our claim that child penalties are larger at the bottom of the wage distribution than at the top.

In addition to these arguments, we provide direct evidence that plausible sources of negative shocks to labor outcomes do not trigger problematic fertility

[21] The parallel trend assumption's validity before treatment is not sufficient for ruling out this possibility.

responses. First, we estimate how macro-level shocks in the labor market affect fertility decisions. We explore the question of whether the Great Recession has altered the probability of having children. Figure 9 suggests that it has not.

Next, we document the effect of macro-shocks in greater detail. Within the population of eligible individuals, i.e., those with exactly $n-1$ children at $t-1$, we estimate the probability of birth of their nth children at time t along the business cycle based on a linear probability model:

$$b_{it}^{n} = \eta^{n}\{\log(GDP_t) - \log(GDP_{t-1})\} + \kappa_{age_{i,t}}^{n} + \pi^{n}t + \xi_{it} \tag{12}$$

where all coefficients are indexed by rank in the recent wage distribution and gender. Coefficients η^n account for the sensitivity of fertility decisions to macro-level shocks. An endogeneity problem would arise if those coefficients were estimated to be significantly negative, especially at the bottom of the wage distribution. According to Figure 10, with very few exceptions, this is not the case.

Second, we ask whether micro-level shocks tend to generate such fertility responses. We build on Huttunen and Kellokumpu (2016), who show that job displacement triggers negative fertility responses (as opposed to positive fertility responses that would be problematic here). We rely on the linked employer-employee nature of our data, which we have not exploited thus far, to identify plausible mass layoff episodes, as frequently done in the job displacement literature (e.g., Jacobson, LaLonde, and Sullivan, 1993). Namely, we assume that individual i is subject to a firm-level shock f_{it} at time t if more than 25% of individuals having the same main employer[22] as that of i at time $t-1$, but who are not individual i herself, leave the firm at time t. First, within each eligible subpopulation, these firm-level shocks indeed correlate with job losses. We estimate a linear model for the probability l_{it} of being jobless at time t,

$$l_{it} = \rho^{n} f_{it} + \sigma_{age_{it},t}^{n} + \upsilon_{it}, \tag{13}$$

[22]The main employer of an individual is defined as the firm that pays that individual the largest earnings during a given year.

where all coefficients depend on gender and the rank in the wage distribution. Figure 11 displays the estimates of coefficients ρ^n, and shows that it is plausible that exogeneous firm-level shocks indeed are felt at the individual level. Second, we estimate the probability of having the nth child at time t,

$$b_{it}^n = \phi^n f_{it} + \psi_{age_{it},t}^n + \omega_{it} \tag{14}$$

Figure 12 displays the corresponding estimates of ϕ^n. In most cases, we cannot reject the null hypothesis that ϕ^n coefficients are equal to 0, which suggests that firm-level employment shocks do not trigger positive fertility responses that would render our estimates of child penalties meaningless. Of course, there could be other kinds of unobserved shocks upon which individuals may base their fertility decisions in ways that could bias our estimated child penalties. Nevertheless, the finding that neither macro-level nor firm-level negative shocks trigger positive fertility responses supports the credibility of our identification of the impact of having children on labor outcomes.

5.3 Endogeneity of pre-childbirth wages

Even with properly identified child penalties along the pre-childbirth wage distribution, there could still be some doubts as to the random assignment of individuals to wage groups. The ideal design for documenting specialization would be to assign individuals exogeneously to various incentives to remain in the labor force, i.e., to varying wage levels, and to contrast child penalties across those groups. However, pre-childbirth wages reflect individuals' decisions that likely depend on their preferences or on the gender norms they are exposed to. Namely, when considering forward-looking individuals who differ in their unobserved work-family preferences, an intertemporal human capital accumulation model with two sectors *à la* Becker (1981) predicts that those with stronger preferences for family over career will invest less in acquisition of labor market-valued skills, and will therefore earn lower hourly wages prior to childbirth. These individuals may, e.g., choose to work in firms that tend to not only comparatively support work-family balance

but also to pay lower wages (Card, Cardoso, and Kline, 2016; Coudin, Maillard, and Tô, 2018). This would generate some reverse causality bias that would prevent interpreting the variation of child penalties with the rank in the wage distribution as a causal relationship.

We assess the extent to which that channel is likely to explain our results (i) by resorting to survey data in which individuals report their preferences towards childcare, in which we observe that the correlation between preferences and hourly wages is quite limited; (ii) by interacting our difference-in-difference method not only with the rank in the recent wage distribution but also with other variables that proxy human capital investment and work-family preferences before childbirth.

In 2010, the French Labor Force Survey (LFS) was complemented with a module devoted to work-family balance. In this module, individuals with children aged 3 or less were asked what was the ideal childcare solution to them for children as old as their youngest child. In Figure 13, we display the share of individuals declaring children should be taken care of by their parents along the hourly wage distribution. First, the vast majority of parents do not view parents-provided childcare as the best childcare solution for children aged the same age as their youngest child: over 2/3 of parents favor external childcare, provided either in formal or informal settings (e.g. by grandparents). Second, we find limited heterogeneity along the wage distribution: the share of women (men) that view parents as the ideal childcare solution varies between 14% and 32% (23% and 48%). Importantly the pattern is non-monotonic, in contrast with our finding as to the profile of labor supply responses to childbirth along the hourly wage distribution, which suggests preferences may not be the main driver of these labor supply decisions. Additionally, the pattern for men does not match the pattern for women. This descriptive evidence is affected by selection bias, given that hourly wages in the LFS can only be computed for individuals who are salaried employees; Appendix G provides additional evidence that addresses this issue. These additional results cannot reject the null hypothesis that parents with the most conservative views regarding childcare do not differ from the others in terms of their (potential) hourly wages.

We then show that additional sources of heterogeneity, that correspond to

past human capital decisions that could plausibly both affect pre-childbirth hourly wages, and stem from childcare-related preferences do not drive our results. To this end, we consider education, measured by the highest degree obtained at the end of studies, as an 8-level variable, as well as the rank in the distribution of labor supply at the intensive margin, conditionally on age and year from $t-5$ to $t-1$. We also rank individuals according to the share of females working part-time for the main employer of each individual at time $t-1$, which leads us to consider another 20 firm-composition-related groups. Ultimately, we estimate (omitting indices g,k and z for clarity):

$$\delta^k z_{it} = (\bar{\alpha}_{r_{it}}+\bar{\alpha}_m M_{it})+\sum_n(\bar{\gamma}^n_{r_{it}}+\bar{\gamma}^n_m M_{it})(b^n_{it}+c^n_{it})+\sum_n(\bar{\theta}^n_{r_{it}}+\bar{\theta}^n_m M_{it})b^n_{it}+(\bar{\zeta}_{r_{it}}+\bar{\zeta}_m M_{it})X_{it}+u_{it} \tag{15}$$

Here, heterogeneity in $\bar{\beta}^n_r$ stems from variations along the wage distribution of childbirth-related changes in z *within* groups of individuals with similar M, i.e., within groups of individuals with similar education, labor force attachment and firm composition as measured at time $t-1$. While this certainly is not sufficient for capturing all the variation that arises from different human capital investment and sorting due to, e.g., heterogeneous work-family preferences, finding substantial heterogeneity in $\bar{\theta}^n_r$ along the wage distribution, i.e., with respect to rank r, can be viewed as evidence that suggests that contemporary hourly wages are a financial incentive that drives much of childbirth-related labor decisions.

We operationalize the approach presented in (15) by considering the probability of remaining in employment one year after childbirth as the outcome. Figure 14 reports the coefficients that depict heterogeneity along the recent wage distribution, first in the case of labor supply decisions being allowed to vary depending on recent hourly wages only (the first and second panels), and next in the case of them potentially also differing depending on education, recent paid hours and firm composition (the third and fourth panels).[23]

If labor supply decisions at time $t+1$ are allowed to be contingent not only on recent wages but also on education, past paid hours and firm composition, we

[23]The coefficients related to heterogeneity along the education dimension, past labor supply decisions and firm composition are shown in Appendix H.

still observe substantial remaining heterogeneity along the distribution. Within groups of women with similar education, similar paid hours between $t-5$ and $t-1$ and similar employers with respect to female labor supply at time $t-1$, those at the bottom of the distribution are 10 (resp., 29 and 25) log-points less likely to remain in employment one year after the arrival of the first (resp., the second and the third) child than are those ranked at the top of the distribution. Hence, differences in education and full-time status as well as differential sorting account for a rather low share of heterogeneity in labor supply responses to childbirth at the extensive margin. While this contribution is non-negligible at approximately one third of the initial estimate, this supports the idea that the phenomenon of mothers at the bottom of the wage distribution being more likely to interrupt their careers is not primarily driven by unobserved preferences.

6 Conclusion

This study of gender differences in career progression focuses particularly on the heterogeneity of the effect of having children on parents' labor outcomes along the wage distribution. Childbirths have a large and negative impact on mothers' labor earnings but are associated with slightly faster labor earnings growth in the case of high-earning fathers. In the short to medium run, this is primarily the result of labor supply decisions, and not so much the consequence of hourly wage losses. Moreover, this effect is heterogeneous along the hourly wage distribution: low-wage mothers are more likely to leave the labor market or to reduce their paid hours. Overall, this striking monotony along the wage distribution is consistent with the specialization channel and with the financial incentives provided by the French family insurance scheme. While intrinsic preferences towards family and career are certainly significant, we show that the observed patterns are likely to be driven by the opportunity cost of career interruptions.

In contrast, the effect of childbirth on mothers' hourly wages is rather homogeneous along the hourly wage distribution (being slightly larger, if varying at all, for mothers with high pre-childbirth wages). Top male earners may also experience a

slight fatherhood premium. In the medium run, the gender gap in hourly wages tends widen more among top earners than among low earners. As a result, the consequences of childbirths contribute to some extent to women's underrepresentation at the top of the distribution, and hence to glass ceiling. This empirical finding may seem at odds with the fact that career interruptions and labor supply reductions among mothers are less frequent at the top of the distribution. However, we show that both returns to experience and hourly wage losses due to career interruptions are presumably also larger, which explains both why mothers with high pre-childbirth wages would be more reluctant to spend time outside the workforce and why them spending less time outside the workforce has still more significant consequences on their hourly wages. This is consistent both with the view that better-compensated occupations have more nonlinear pay structures (Goldin, 2014) so that balancing career and family concerns is more difficult and leads to larger motherhood penalties (Bütikofer, Jensen, and Salvanes, 2018), and with higher human capital depreciation due to time spent outside the labor force for high-skilled workers (Adda, Dustmann, and Stevens, 2017).

That children-related labor supply decisions of mothers are seemingly driven by economic gains of within-household specialization and financial incentives suggests that reshaping the design of parental leave allowances (Piketty, 2005; Lequien, 2012; Joseph et al., 2013) including childcare subsidies (Givord and Marbot, 2015) will have first-order consequences on mothers' career progression and therefore on the gender pay gap. Since current parental leave allowances are mainly lump-sum transfers that do not depend on the hourly wage, mothers with low (potential) wages are more likely to be pulled out of the workforce, and therefore to experience lower wage growth; the sticky floor is consistent with this financial scheme in this regard.

Another policy instrument related to childbirth includes extended and mandatory paternity leaves: the French ministry in charge of promoting gender equality has recently asked for an *ex ante* evaluation on this topic. A recent working paper by Nix and Andresen (2019) considering a broad Norwegian reform that expanded access to childcare suggests that subsidized early childcare is more efficient at re-

ducing the child penalty than are paternity leaves, which would nevertheless be useful in balancing childcare within the household according to Pailhé, Solaz, and Tô (2018).

References

Adda, J., C. Dustmann, and K. Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy* 125:293–337.

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

Albrecht, J., A. Björklund, and S. Vroman. 2003. "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics* 41:89–114.

Albrecht, J., P.S. Thoursie, and S. Vroman. 2015. "Parental Leave and the Glass Ceiling in Sweden." In *Gender Convergence in the Labor Market*. Emerald Publishing Ltd, vol. 41 of *Research in Labor Economics*, pp. 89–114.

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

Angrist, J.D., and W.N. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *The American Economic Review* 88:450–477.

Arulampalam, W., A.L. Booth, and M.L. Bryan. 2007. "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution." *Industrial and Labor Relations Review* 60:163–186.

Becker, G. 1981. *A Treatise on the Family*. Cambridge: Harvard University Press.

Becker, G.S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics* 3:33–58.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119:249–275.

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* 2:228–55.

Budig, M.J., and M.J. Hodges. 2010. "Differences in Disadvantage: Variation in the Motherhood Penalty across White Women's Earnings Distribution." *American Sociological Review* 75:705–728.

—. 2014. "Statistical Models and Empirical Evidence for Differences in the Motherhood Penalty across the Earnings Distribution." *American Sociological Review* 79:358–364.

Bütikofer, A., S. Jensen, and K.G. Salvanes. 2018. "The role of parenthood on the gender gap among top earners." *European Economic Review* 109:103–123.

Card, D., A.R. Cardoso, and P. Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *The Quarterly Journal of Economics* 131:633–686.

Charnoz, P., E. Coudin, and M. Gaini. 2011. "Changes in the French Wage Distribution 1976-2004: Inequalities within and between Education and Experience Groups." Working paper INSEE.

Choné, P., D. Le Blanc, and I. Robert-Bobée. 2004. "Offre de travail féminine et garde des jeunes enfants." *Économie et Prévision* 162:23–50.

Coudin, E., S. Maillard, and M. Tô. 2018. "Family, firms and the gender wage gap in France." IFS Working Papers No. W18/01, Institute for Fiscal Studies.

Datta Gupta, N., N. Smith, and M. Verner. 2008. "The impact of Nordic countries' family friendly policies on employment, wages, and children." *Review of Economics of the Household* 6:65–89.

Deaton, A.S. 1997. "Econometric Issues for Survey Data." In T. W. Bank, ed. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The Johns Hopkins University Press, pp. 63–132.

Deaton, A.S., and C. Paxson. 1994. "Saving, Growth, and Aging in Taiwan." In *Studies in the Economics of Aging*. National Bureau of Economic Research, Inc, NBER Chapters, pp. 331–362.

Dustmann, C., and C. Meghir. 2005. "Wages, Experience and Seniority." *The Review of Economic Studies* 72:77–108.

Ejrnæs, M., and A. Kunze. 2013. "Work and Wage Dynamics around Childbirth." *The Scandinavian Journal of Economics* 115:856–877.

England, P., J. Bearak, M.J. Budig, and M.J. Hodges. 2016. "Do Highly Paid, Highly Skilled Women Experience the Largest Motherhood Penalty?" *American Sociological Review* 81:1161–1189.

Fortin, N.M., B. Bell, and M. Böhm. 2017. "Top earnings inequality and the gender pay gap: Canada, Sweden, and the United Kingdom." *Labour Economics* 47:107 – 123.

Givord, P., and C. Marbot. 2015. "Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies." *Labour Economics* 36:99 – 111.

Gobillon, L., D. Meurs, and S. Roux. 2015. "Estimating Gender Differences in Access to Jobs." *Journal of Labor Economics* 33:317–363.

Goldin, C. 2014. "A Grand Gender Convergence: Its Last Chapter." *The American Economic Review* 104:1091–1119.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song. 2016. "What do data on millions of US workers reveal about life-cycle earnings risk?" mimeo.

Hersch, J., and L.S. Stratton. 1997. "Housework, Fixed Effects, and Wages of Married Workers." *The Journal of Human Resources* 32:285–307.

Huss, M.M. 1990. "Pronatalism in the Inter-War Period in France." *Journal of Contemporary History* 25:39–68.

Huttunen, K., and J. Kellokumpu. 2016. "The Effect of Job Displacement on Couples' Fertility Decisions." *Journal of Labor Economics* 34:403–442.

Jacobson, L.S., R.J. LaLonde, and D.G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review* 83:685–709.

Joseph, O., A. Pailhé, I. Recotillet, and A. Solaz. 2013. "The economic impact of taking short parental leave: Evaluation of a French reform." *Labour Economics* 25:63 – 75.

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

Killewald, A., and J. Bearak. 2014. "Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression." *American Sociological Review* 79:350–357.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimuller. 2019a. "The Impact of Family Policies on the Dynamics of Gender Inequality." mimeo.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. 2019b. "Child Penalties Across Countries: Evidence and Explanations." NBER Working Papers No. 25524, National Bureau of Economic Research, Inc, Feb.

Kleven, H., C. Landais, and J.E. Søgaard. forthcoming. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, pp. .

Korenman, S., and D. Neumark. 1992. "Marriage, Motherhood, and Wages." *The Journal of Human Resources* 27:233–255.

Lequien, L. 2012. "The Impact of Parental Leave Duration on Later Wages." *Annals of Economics and Statistics*, pp. 267–285.

Lundberg, S., and E. Rose. 2000. "Parenthood and the earnings of married men and women." *Labour Economics* 7:689–710.

Mason, K.O., W.M. Mason, H.H. Winsborough, and W.K. Poole. 1973. "Some Methodological Issues in Cohort Analysis of Archival Data." *American Sociological Review* 38:242–258.

Meurs, D., A. Pailhé, and S. Ponthieux. 2010. "Child-related career interruptions and the gender wage gap in France." *Annals of Economics and Statistics*, pp. 15–46.

Miller, A.R. 2011. "The effects of motherhood timing on career path." *Journal of Population Economics* 24:1071–1100.

Nix, E., and M.E. Andresen. 2019. "What Causes the Child Penalty? Evidence from Same Sex Couples and Policy Reforms." Discussion Papers No. 902, Statistics Norway.

Pailhé, A., A. Solaz, and M. Tô. 2018. "Can Daddies Learn How to Change Nappies? Evidence from a Short Paternity Leave Policy." Working paper, Ined.

Phelps, E. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62:659–61.

Piketty, T. 2005. "L'impact de l'allocation parentale d'éducation sur l'activité féminine et la fécondité en France, 1982-2002." *Les Cahiers de l'INED*, pp. 79–109.

Rosental, P.A. 2010. "Politique familiale et natalité en France : un siècle de mutations d'une question sociétale." *Santé, Société et Solidarité* 9:17–25.

Rosenzweig, M.R., and K.I. Wolpin. 1980. "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models." *Journal of Political Economy* 88:328–348.

Steinhauer, A. 2018. "Working Moms, Childlessness, and Female Identity." Sciences Po publications No. 79, Sciences Po, May.

Waldfogel, J. 1997. "The Effect of Children on Women's Wages." *American Sociological Review* 62:209–217.

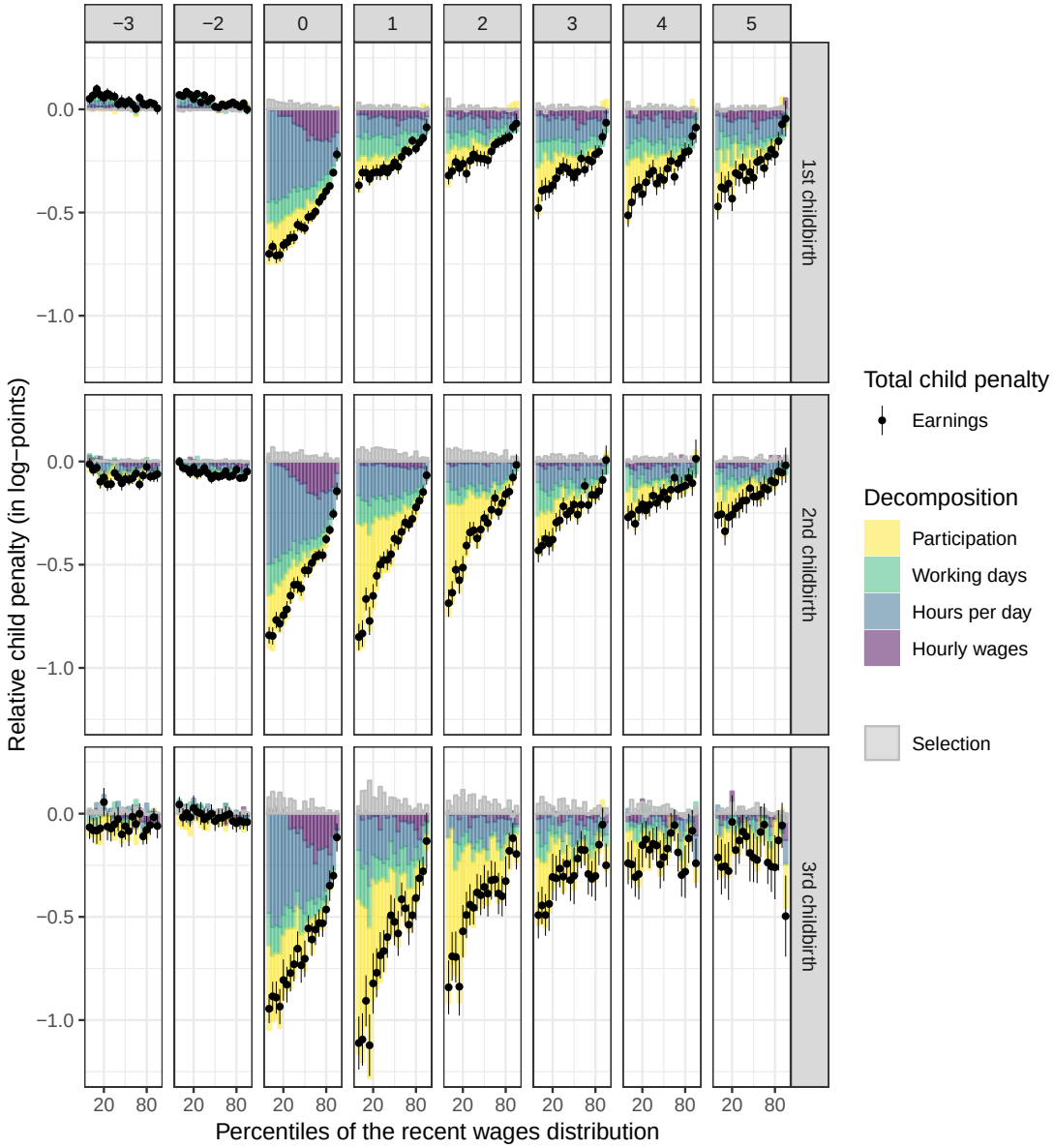
—. 1995. "The Price of Motherhood: Family Status and Women's Pay in Young British Cohort." *Oxford Economic Papers* 47:584–610.

—. 1998. “Understanding the ”Family Gap” in Pay for Women with Children.” *The Journal of Economic Perspectives* 12:137–156.

Wilner, L. 2016. “Worker-firm matching and the parenthood pay gap: Evidence from linked employer-employee data.” *Journal of Population Economics* 29:991–1023.

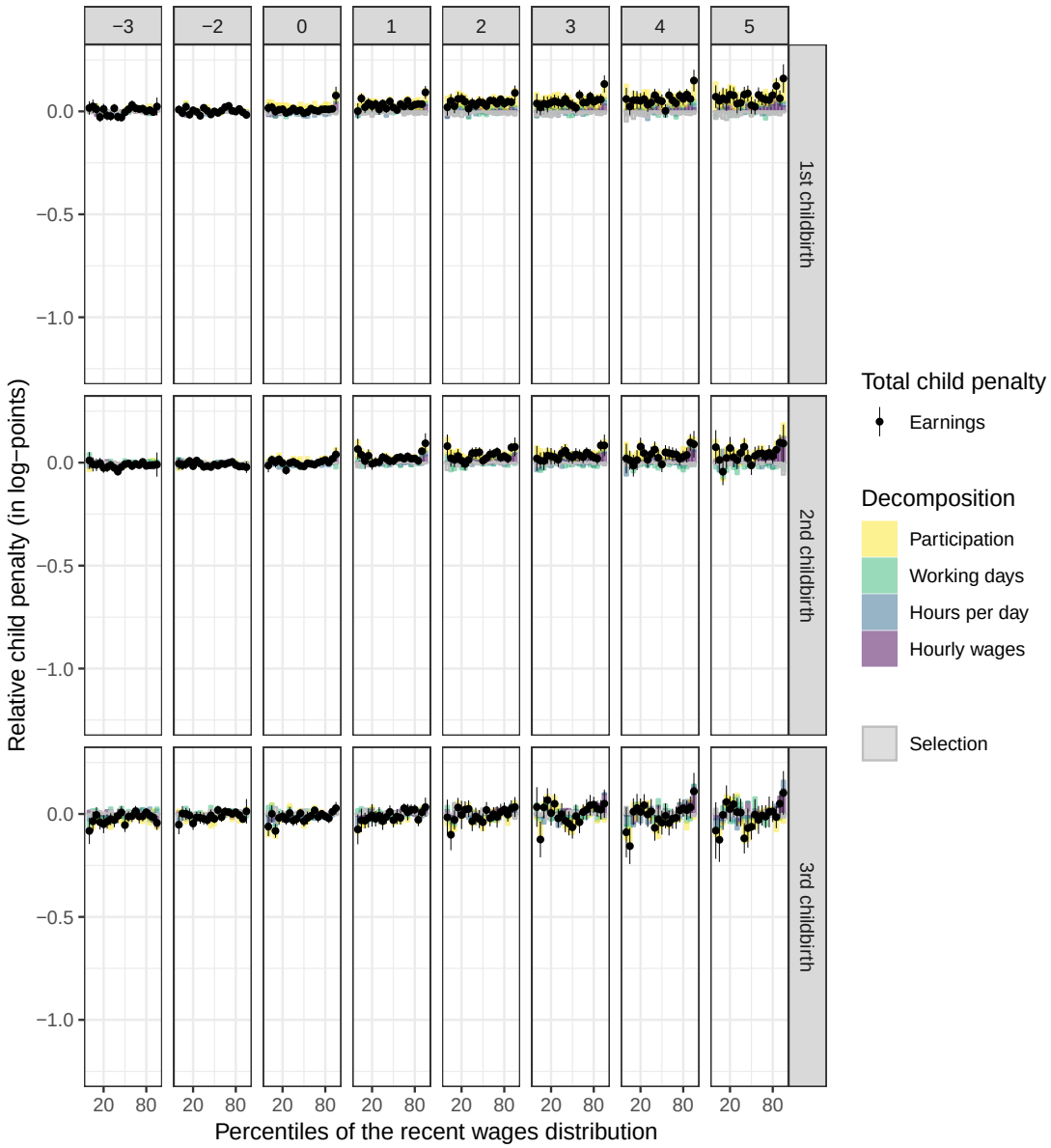
Figures

Figure 1 – Consequences of childbirth for women's labor outcomes



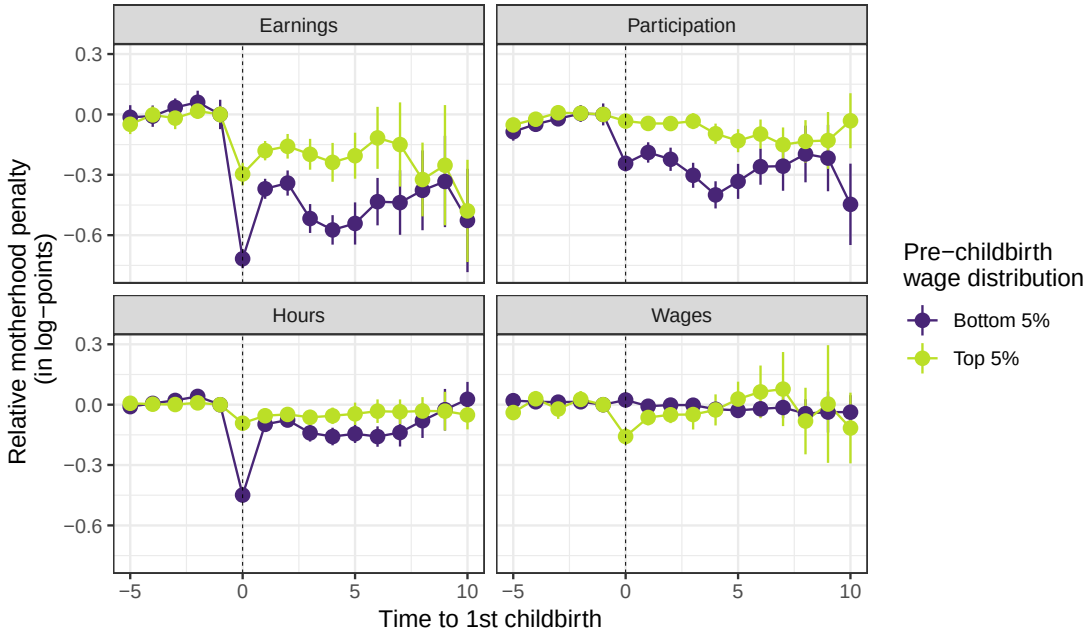
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see Equation (7)) for various values of time-to-childbirth expressed in years. Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure 2 – Consequences of childbirth for men's labor outcomes



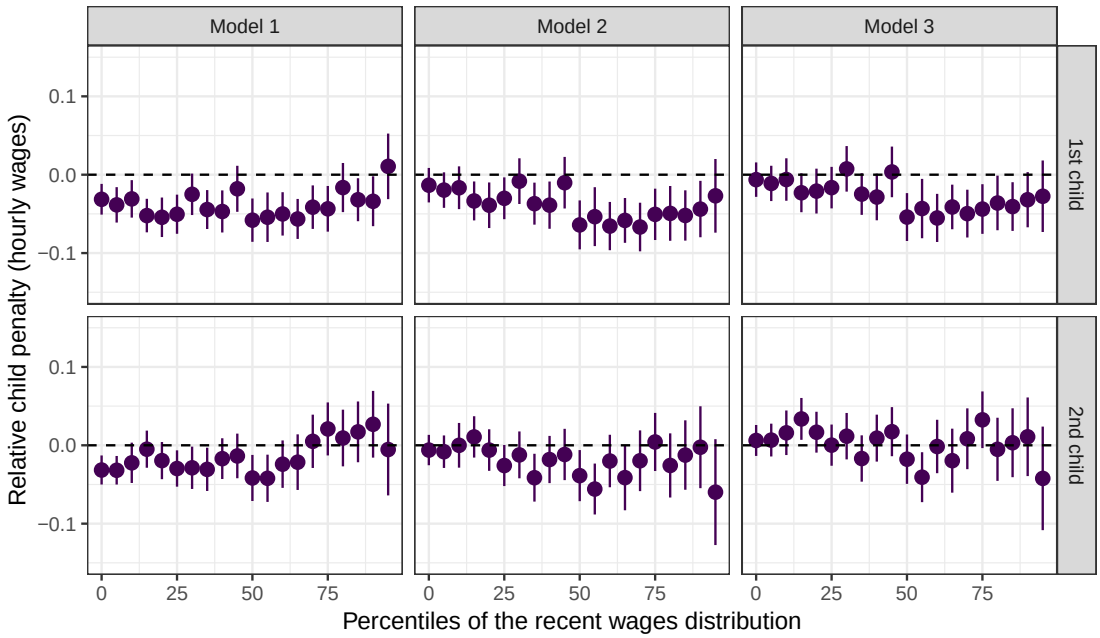
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see Equation (7)) for various values of time-to-childbirth expressed in years. Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure 3 – Impact of first childbirth on the gender gap in earnings and labor outcomes



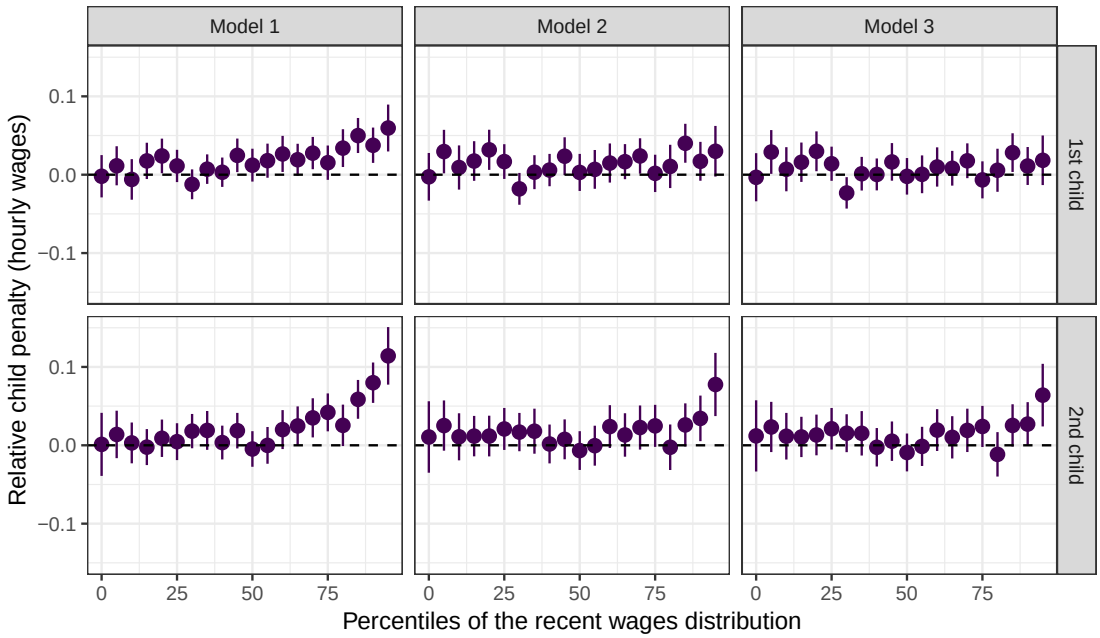
Estimates of the impact of first childbirth on the gender gap in pay, obtained by the difference-in-difference-in-difference method (see Equation (8)). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure 4 – Medium-run impact of childbirth on hourly wages (women)



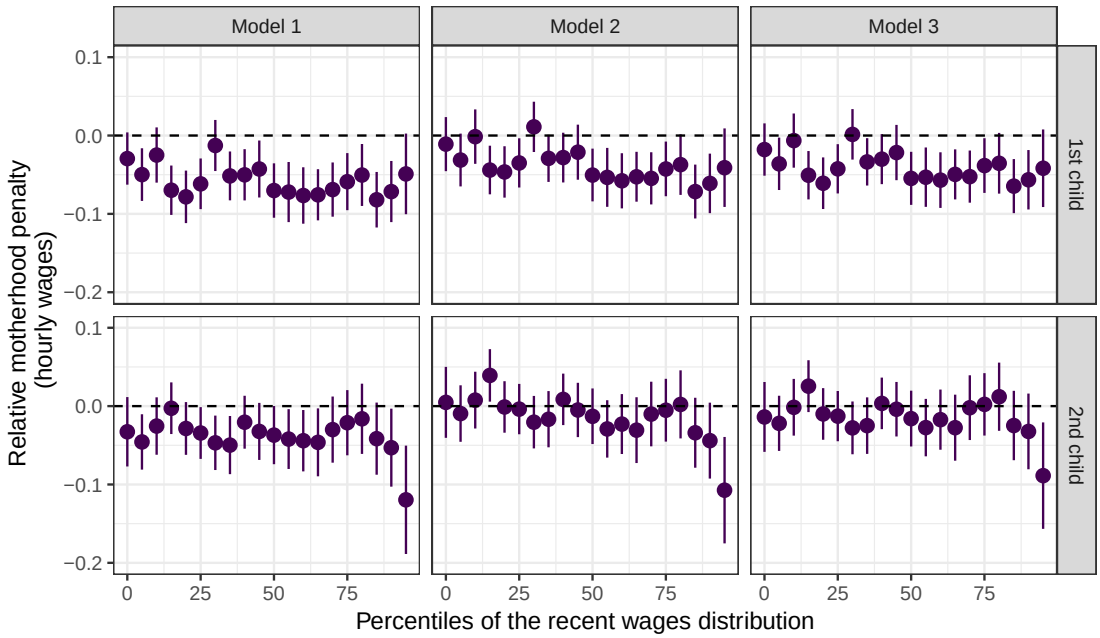
Estimates of coefficients related to childbirth for women in the hourly wage growth model that interacts a double-difference setting with gender and rank in the recent wage distribution (11). The outcome is a (logarithmic) hourly wage growth between times $t-1$ and $t+5$. Model 1 includes no controls. Model 2 controls for year, age, industry, firm composition (share of women working part-time) and 1-digit occupation within each gender × recent wage cell. Model 3 includes all of these controls as well as experience between times t and $t+5$, a dummy variable for having spent at least one year outside private sector employment, and having changed firms between $t-1$ and $t+5$. Standard errors are clustered at the individual level. The sample includes individuals up to age 55 at time t.

Figure 5 – Medium-run impact of childbirth on hourly wages (men)



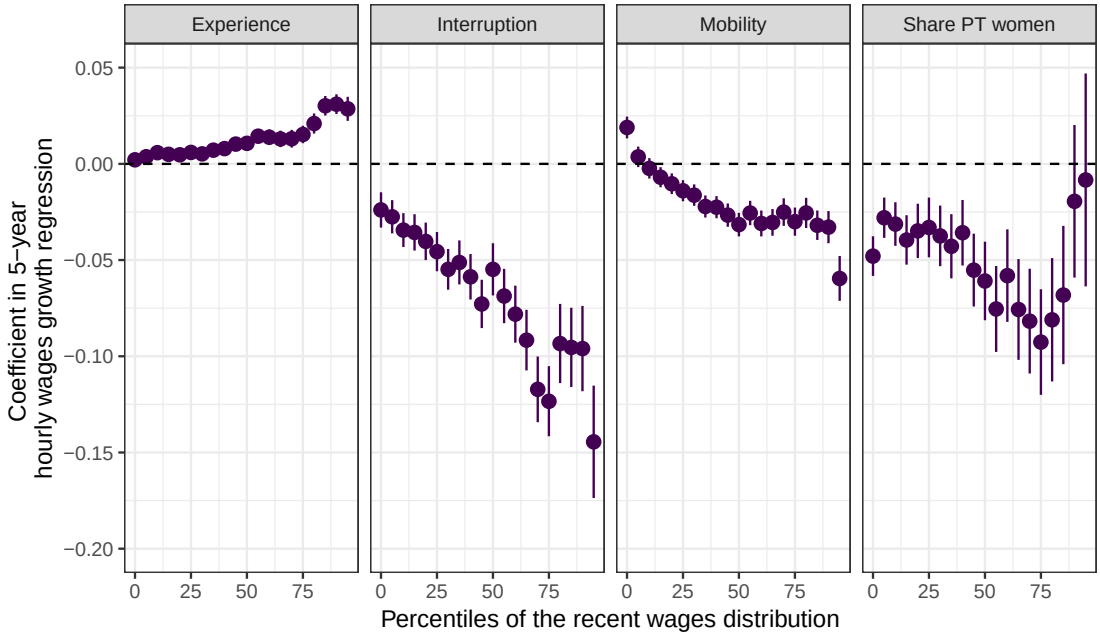
Estimates of coefficients related to childbirth for men in the hourly wage growth model that interacts a double-difference setting with gender and rank in the recent wage distribution (11). The outcome is a (logarithmic) hourly wage growth between times $t-1$ and $t+5$. Model 1 includes no controls. Model 2 controls for year, age, industry, firm composition (share of women working part-time) and 1-digit occupation within each gender × recent wage cell. Model 3 includes all of these controls as well as experience between times t and $t+5$, a dummy variable for having spent at least one year outside private sector employment, and having changed firms between times $t-1$ and $t+5$. Standard errors are clustered at the individual level. The sample includes individuals up to age 55 at time t.

Figure 6 – Medium-run impact of childbirth on hourly wages (gender gap)



Estimates of coefficients related to childbirth for women (with men considered as a reference) in the hourly wage growth model that interacts a double-difference setting with gender and rank in the recent wage distribution (12). The outcome is a (logarithmic) hourly wage growth between times $t-1$ and $t+5$. Model 1 includes no controls. Model 2 controls for year, age, industry, firm composition (share of women working part-time) and 1-digit occupation within each recent wage cell. Model 3 includes all of these controls as well as experience between times t and $t+5$, a dummy variable for having spent at least one year outside private sector employment, and having changed firms between times $t-1$ and $t+5$. Standard errors are clustered at the individual level. The sample includes individuals up to age 55 at time t.

Figure 7 – Heterogeneity in returns to experience, career interruptions, firm composition and between-firm mobility



Estimates of coefficients related to experience and career interruptions in the hourly wage growth model that interacts a double-difference setting with gender and rank in the recent wage distribution (12). The outcome is a (logarithmic) hourly wage growth between times $t-1$ and $t+5$. The model controls for year, age, industry, firm composition (share of women working part-time), 1-digit occupation, experience between times t and $t+5$, a dummy for having spent at least one year outside private sector employment, and having changed firms between times $t-1$ and $t+5$. Standard errors are clustered at the individual level. The sample includes individuals up to age 55 at time t.

Figure 8 – Probability of having children

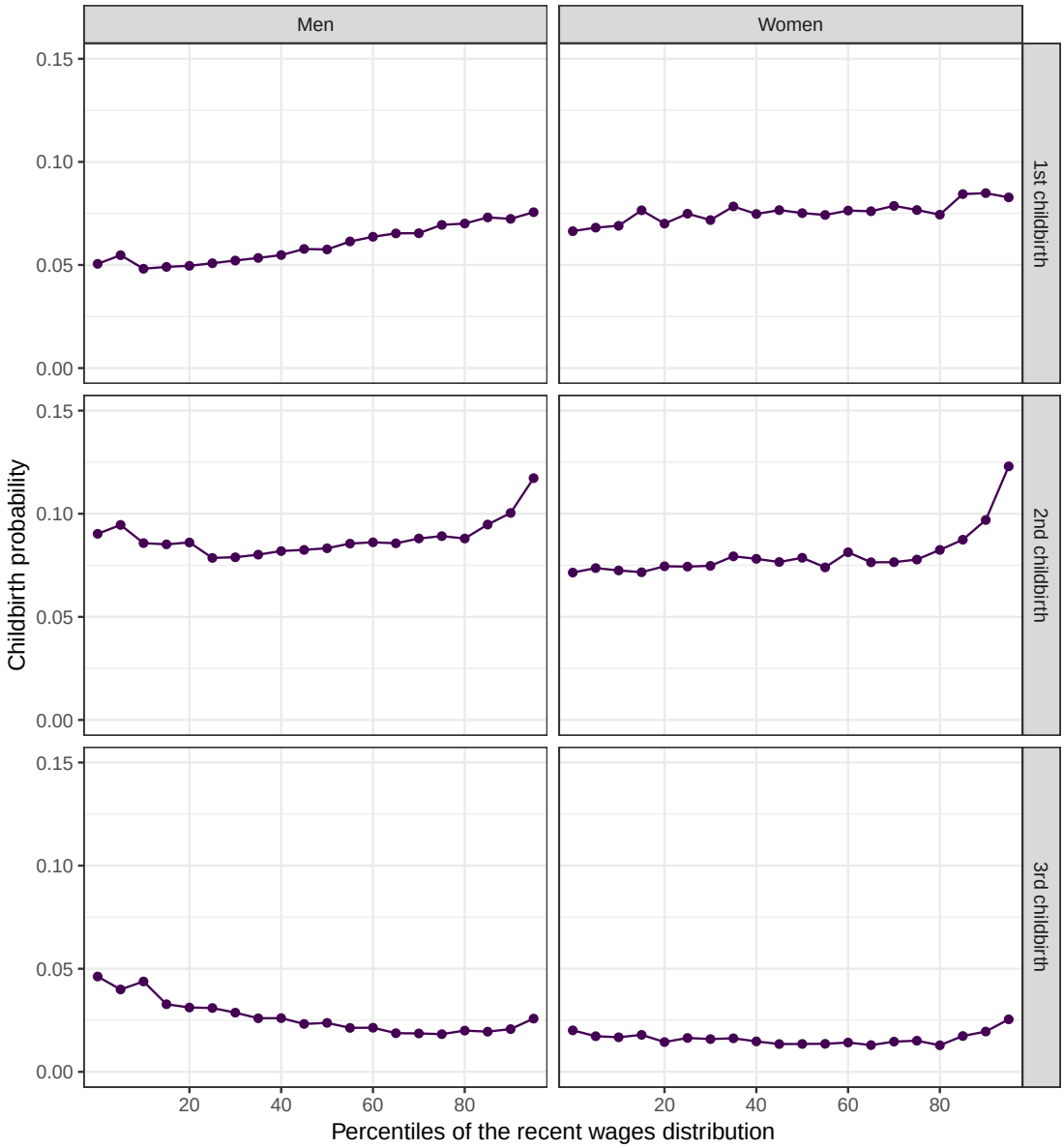


Figure 9 – Probability of having children (by subperiod)

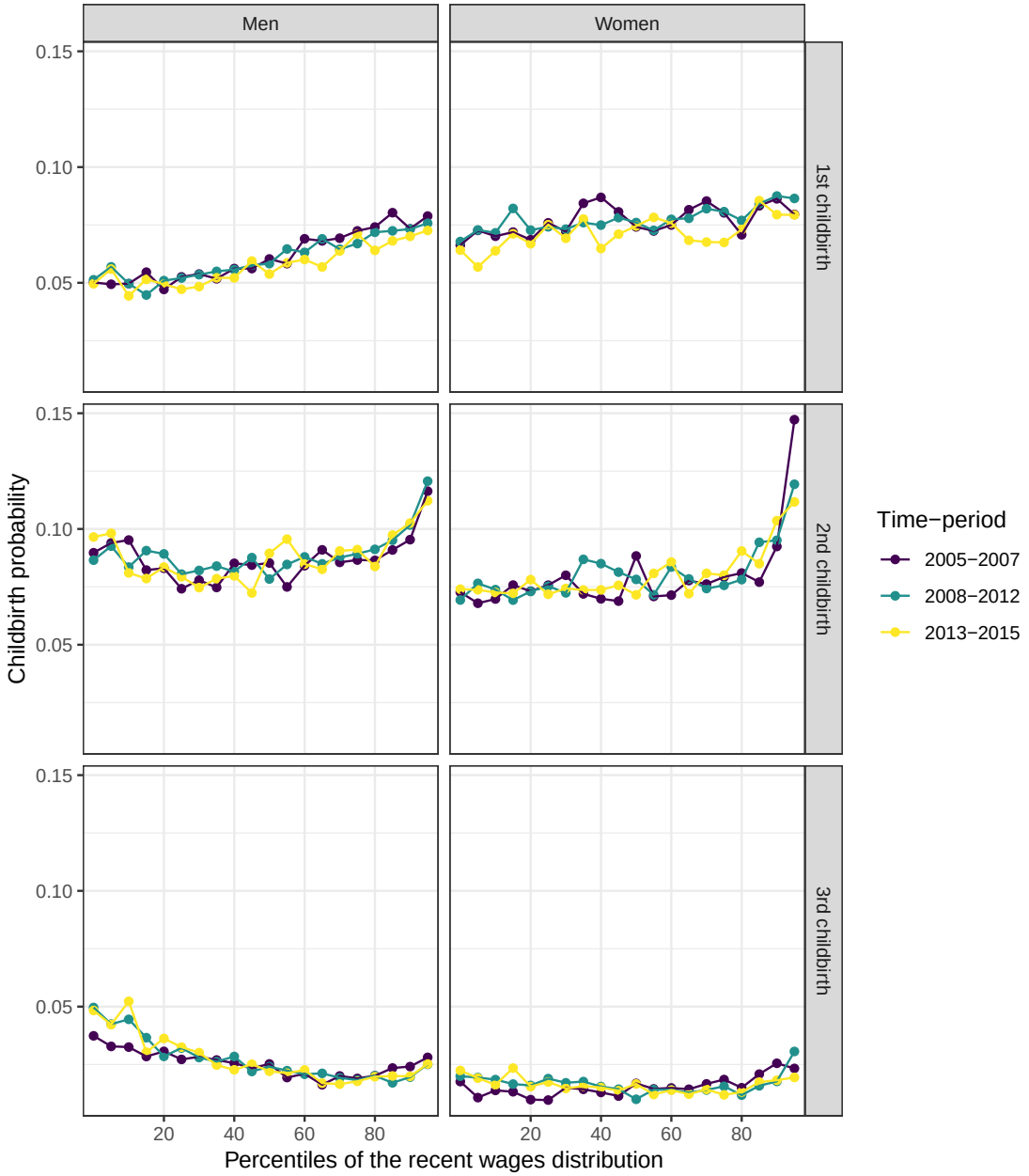
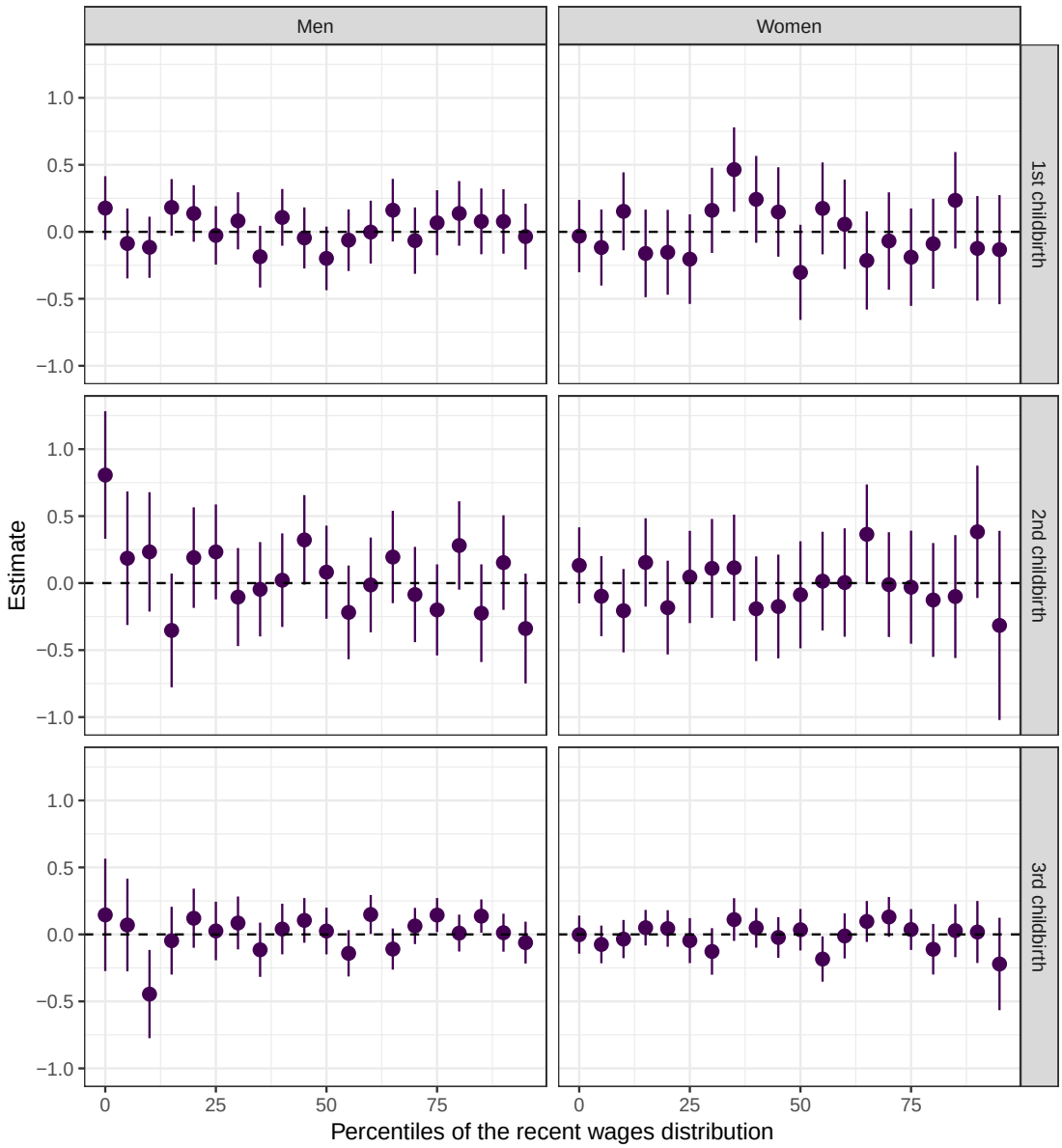
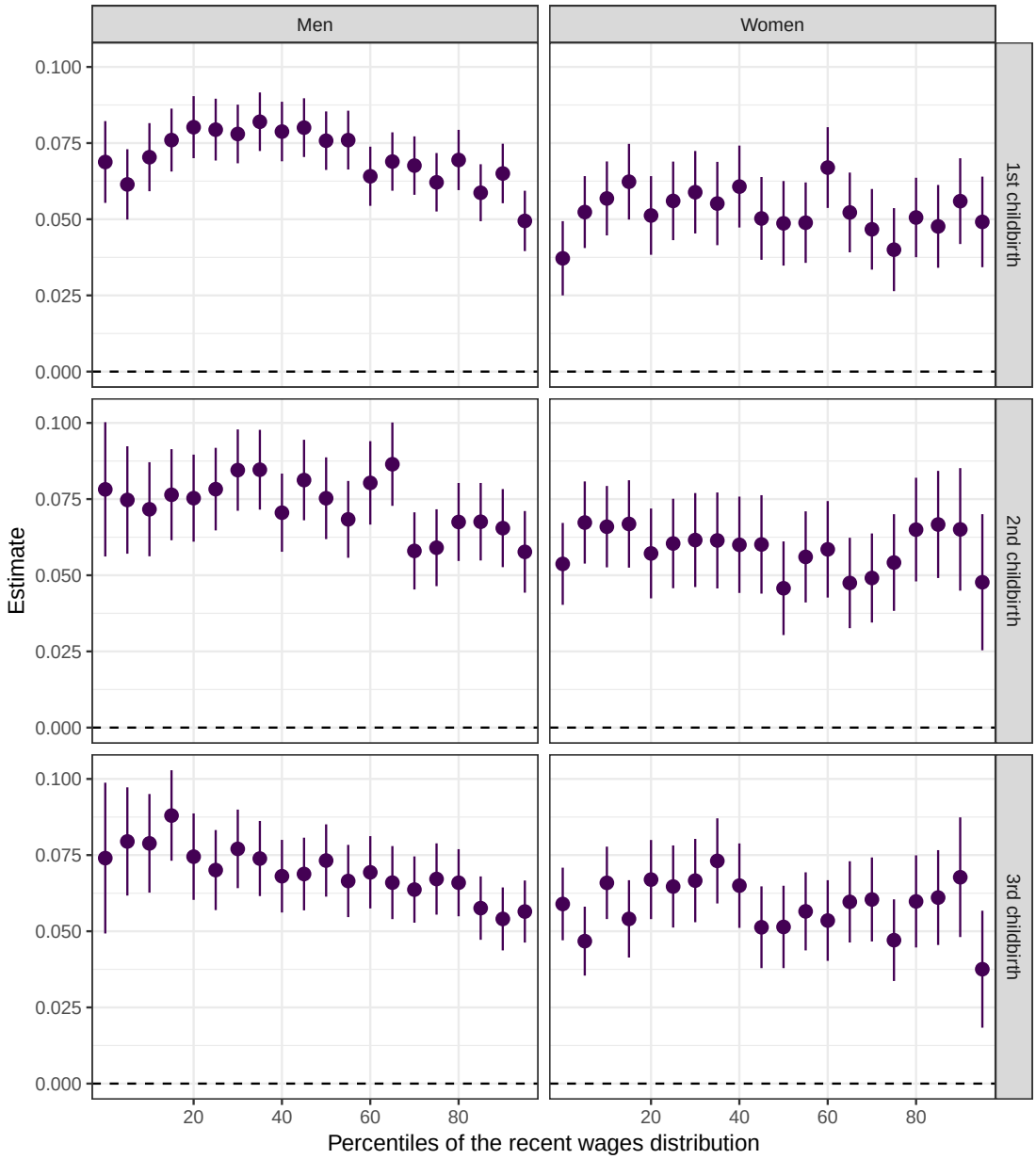


Figure 10 – Probability of having children (sensitivity to the business cycle)



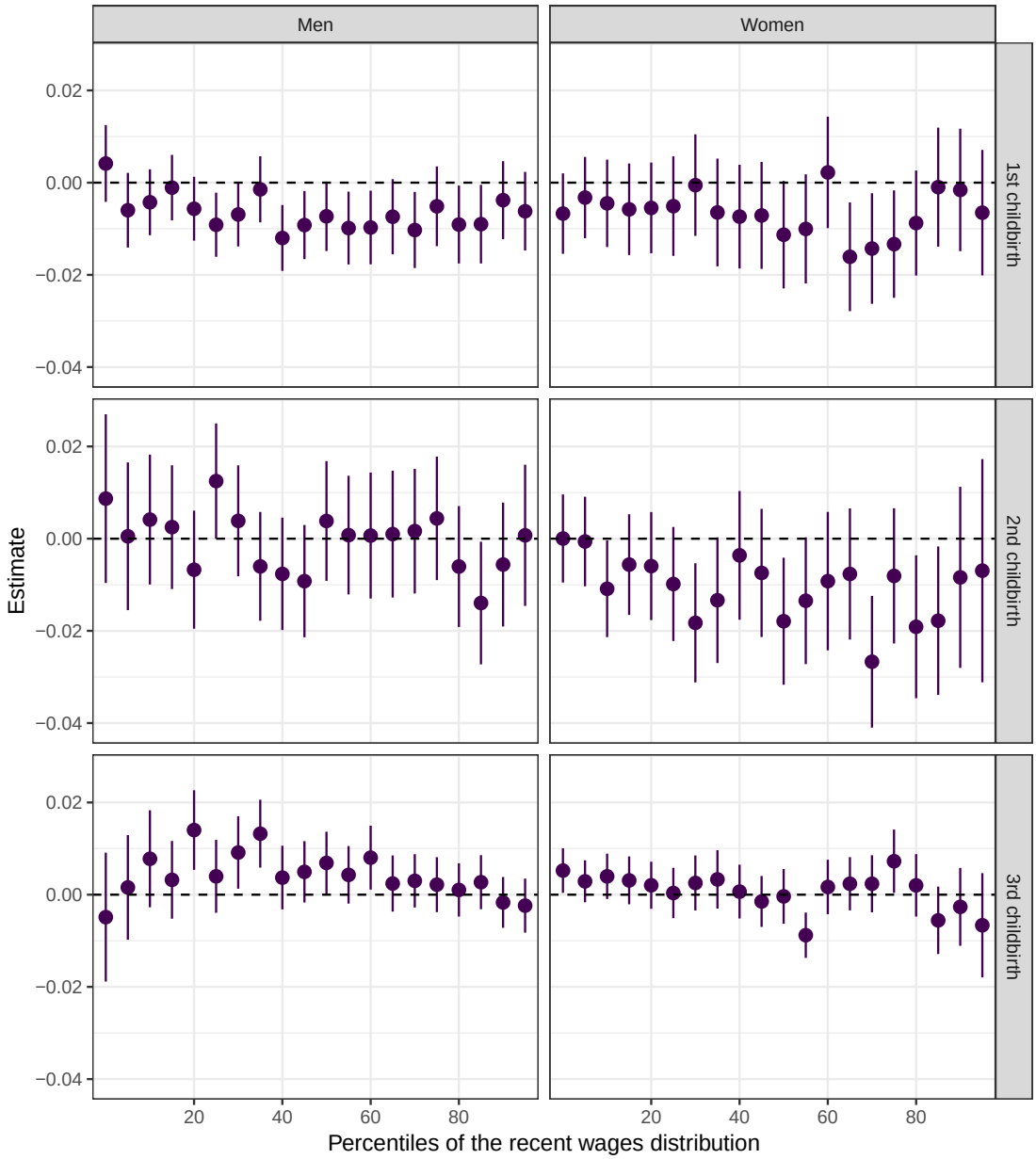
Estimates of coefficients related to log-GDP growth between times $t-1$ and t in a linear probability model with rank in the recent wage distribution × age fixed effects (12). The outcome is a dummy variable for having the nth childbirth at time t. Standard errors are clustered at the individual level. The sample includes individuals up to age 60 at time t.

Figure 11 – Probability of losing one's job (sensitivity to firm-level shocks)



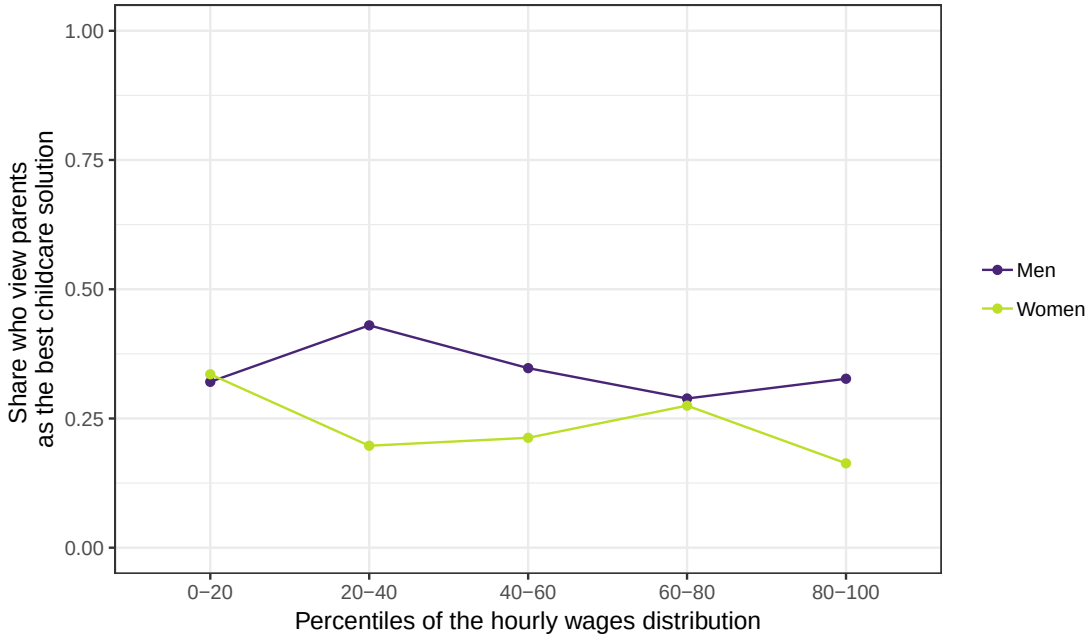
Estimates of coefficients related to firm-level shocks in a linear probability model with rank in the recent wage distribution × age × year fixed effects (13). The outcome is a dummy variable for being jobless at time t. Standard errors are clustered at the individual level. The sample includes individuals up to age 60 at time t.

Figure 12 – Probability of having children (sensitivity to firm-level shocks)



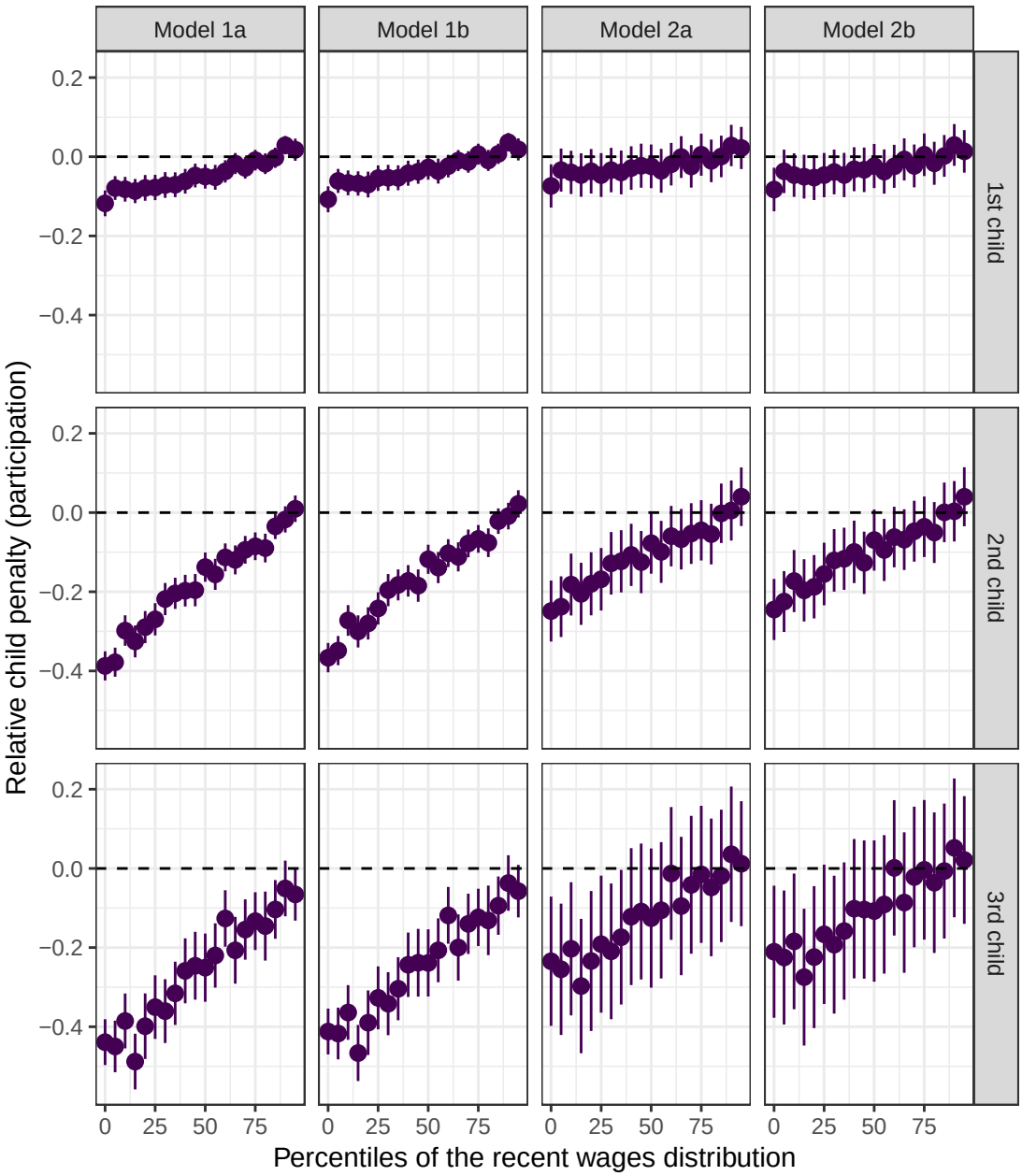
Estimates of coefficients related to firm-level shocks in a linear probability model with rank in the recent wage distribution × age × year fixed effects (14). The outcome is a dummy variable for having the nth childbirth at time t. Standard errors are clustered at the individual level. The sample includes individuals up to age 60 at time t.

Figure 13 – Childcare preferences and hourly wages



Share of individuals who declare parents-provided childcare is the best childcare solution for children aged the same age as their youngest child. Ranks in the hourly wages distribution are conditional on the age of the youngest child. The sample includes individuals with children aged 3 or less.

Figure 14 – Heterogeneity in the probability of remaining in employment one year after childbirth: allowing for additional sources of heterogeneity



Estimates of coefficients related to childbirth for women in a linear probability model that interacts a double-difference setting with gender and rank in the recent wage distribution (15). The outcome is a dummy variable for participating in the labor market at time $t+1$. In Model 1, the difference-in-difference is only interacted with the recent wage distribution; in Model 2, it is also interacted with education, rank in the distribution of recent paid hours and rank in the distribution of firm composition. Models 1a and 2a include no controls; models 1b and 2b control for year, age, industry and 1-digit occupation within each cell. Standard errors are clustered at the individual level. The sample includes individuals up to age 59 at time t.

For Online Publication - Appendix

A Earnings and working time measures

A.1 Earnings

Our measure of labor earnings relies on net annual earnings. This measure aggregates all wages paid to an individual, including performance pay and bonuses, paid vacations, in-kind benefits, the share of severance payments that exceeds the legal minimum, and early retirement benefits (to the extent that these benefits exceed an amount approximately equal to the minimum wage) but excludes stock-options. Social security contributions, public pension schemes, unemployment benefits and other contributions including two flat-rate taxes on earned income (CSG and CRDS) are subtracted to this amount to compute our measure of net annual earnings. In that sense, we measure earnings before income taxes but after some transfers.

Maternity leave allowances are paid by the Social Security administration, and as such are not part of our measure of earnings. They may, however, be paid through the employer (*subrogation*): in this setting, the employer pays the employee the equivalent of maternity leave allowances during her maternity leave, and is later reimbursed by the Social Security administration. The employer subsequently subtracts the maternity leave allowances that the employer advanced from the measure of earnings. Because the reimbursement occurs after the maternity leave itself, the decline in earnings may occur a few weeks after the maternity leave. Because we consider annual earnings, this problem is restricted to childbirths that occur at the end of the calendar year. Our results are, however, very robust to considering only childbirths that occur in the 2nd quarter of the year that are immune to this issue (see Subsection 5.1).

Lastly, in some firms the employer may be bound by collective agreement to complement earnings during maternity or sick leaves in addition to Social Security-

provided allowances. This complement is part of labor earnings as measured by the DADS.

A.2 Days

In the DADS dataset, days worked refer to the duration during which an employee is part of the workforce of a firm within a given year. As a result, maternity and sick leaves, or paid vacations are part of this measure of days, whereas a period of unemployment between two distinct employment spells is not. Additionally, this measure of days is capped at 360.

A.3 Hours

In our dataset, hours worked refer to hours for which the worker is paid according to the labor contract. The data on hours is reported by employers when they fill out payroll tax forms. Before making the data available, Insee performs three checks:

- the total number of hours for a given individual × employer × year observation should not exceed an industry-specific threshold of 2,500 hours per year in a small subset of industries (mostly manufacturing industries, transportation, hotels and restaurants), and 2,200 hours per year in the rest of the private sector;
- the implied hourly wages should exceed 80% of the minimum wage;
- the total number of hours should be positive, with the exception of a narrow subset of occupations (mostly journalists and salespersons) working on a fixed-price basis.

If one of these conditions does not hold, Insee ascribes hours to the observation to make the hourly wage consistent within narrow cells defined by 4-digit occupation, full-time or part-time status, age and gender.

As to workers whose compensation does not depend on the time worked, but who do not belong to one of the above-mentioned occupations, i.e., typically man-

agers ("forfait-jour"), employers provide the number of days only. A number of hours is ascribed to these observations based on the legal duration of work for full-time workers, the number of work days, and the implied hourly wages.

Because during a maternity leave, an employee is not paid by her employer for any hours worked but is instead paid by the Social Security Administration (and possibly receives a complementary payment from her employer), hours worked during a maternity leave are equal to 0. Workers who are not paid by the hour are an exception to this rule because their hours are imputed based on their days worked, which do not vary during maternity leaves. As a result, the DADS dataset overestimates hours worked – and underestimates hourly wages – for such workers during years when they give birth to children. In general, these are qualified workers that belong to the upper part of the hourly wage distribution, so the decomposition of earnings penalties into hours and hourly wages may be biased at the top of the distribution for the specific year workers take maternity leaves.

B Childbirth imputation

We combine data obtained from administrative birth records with census data to deal with the incompleteness of the former for individuals born October 2 and 3 in our dataset. Specifically, (part of) birth records are missing for these individuals between 1982 and 1997. Our strategy is to take information from the censuses of 1990 and 1999 to fill the gap.

For each individual in our sample, our data provides us with

- the years of birth of the 1st to the 12th children appearing in birth records as of 1967;
- the years of birth of the 1st to the 12th children as declared in the 1990 census;
- the years of birth of the 1st to the 12th children as declared in the 1999 census.

Information from birth records has been available since 1967 only, which results in left-censoring. However, because we are mostly interested in individuals giving birth between 2005 and 2015, we do not try to deal with this issue. Our goal is to fill the gap in administrative records between 1982 and 1997 for half of the sampled individuals, which increases our sample size substantially.

For each individual i belonging to the incomplete half of the sample, we impute first the year of first childbirth according to the following principles:

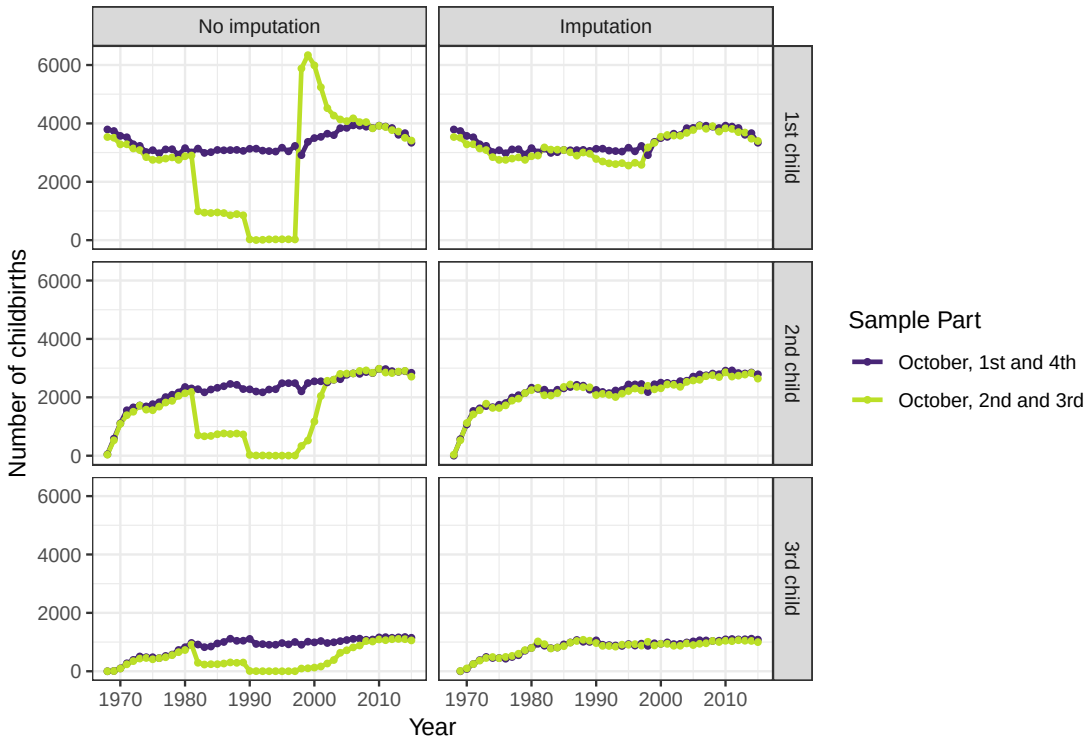
- if the first childbirth in birth records occurs before 1982, we regard it as the first childbirth;
- else,
  - if the earliest of years of childbirth she declared in the 1990 census is after 1982, we consider the earliest of these years and the year of the first childbirth as it appears in birth records as the year of the first childbirth;

- else,
    - if the earliest of years of childbirth she declared in the 1999 census is after 1982, we consider the earliest of these years and the year of the first childbirth as it appears in birth records as the year of the first childbirth;
    - else,
        - if birth records indicate that she has children, we consider the year of the first childbirth in birth records as the year of first childbirth;
        - else, we assume that she has no children.

We then consider the nth childbirth with $n > 1$ as the minimum of years of childbirth within both birth records and censuses among years of birth that follow the computed year of the $n-1$th childbirth.

This approach does not take multiple births into account; more generally, it does not account for individuals who experience more than one childbirth per year. Despite this *caveat*, our approach matches the historical pattern in the complete half of the sample quite well. Figure B.1 plots the number of childbirths by rank of childbirth for each year since 1968 for both parts of the sample, relying on birth records only (left panel) and on our approach (right panel). While we still slightly underestimate first childbirths that occur in the beginning of the 1980s or in the late 1990s in the incomplete half of the sample, our approach matches reasonably well the patterns observed in the complete half of the sample; this is especially true for the 2005-2015 period on which our analysis focuses.

Figure B.1 – Imputation of childbirths for individuals born October 2 and 3



C Summary statistics

Table 1 provides several statistics for the selection process. First, censoring of observations with low numbers of paid hours or low employment duration is illustrated. Second, the restriction to individuals for whom data are available for two years between $t-5$ and $t-2$ in addition to years $t-1$ and t is applied. As expected, both steps increase average hourly wages within a given gender, age group and industry. The selection is harsher for women than it is for men, as women are more likely to experience career interruptions. Censoring reduces the share of younger workers slightly, which is consistent with entry into the workforce through shorter and non-full-time employment spells; selection has the same effect for the same reason. Censoring reduces the share of workers in the service industry who are more likely to have short employment spells and to work part-time. Selection also reduces the share of service industry workers among men and the share of trade industry workers among women, as these individuals have less stable employment histories than those of their counterparts working in other industries.

Both within our base sample (after censoring) and within our selected sample, the gender gap in hourly wages is larger among older workers than among their younger counterparts.

Figure C.1 displays the number of childbirths both in the raw EDP dataset and in our final sample.[24] Because we focus on childbirths that occur after individuals have experienced rather stable employment for several years in a row, and because our data only covers salaried employment in the private sector, numerous childbirths are not included in our final sample: we disregard between one third (the third childbirth) and one half (the first childbirth) of women who experienced childbirth between 2005 and 2015. These proportions amount to 50% and 60% for men during the same period.

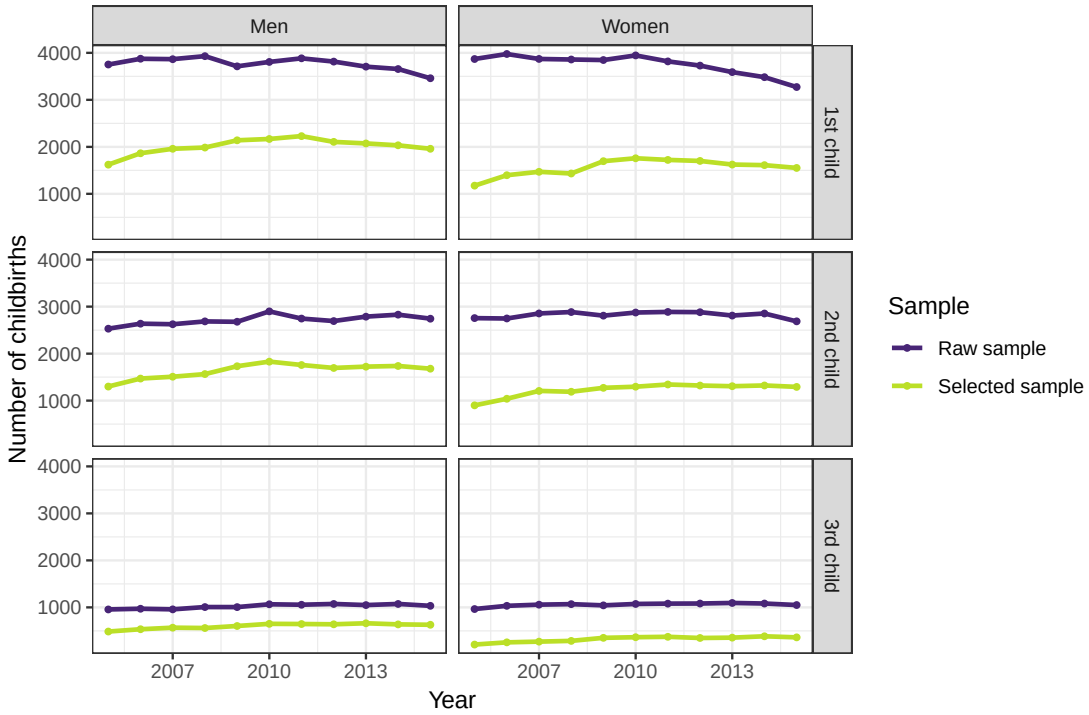
[24]The raw EDP dataset itself is not perfectly representative of all childbirths that occur in France because it only provides information on fertility of individuals that have appeared at least once in labor market data, the sample of which has varied over time.

Table 1 – Sample selection

	Base sample				Censoring				Final sample			
N_{it}	2 851 895				2 568 211				1 886 171			
N_i	362 637				334 811				267 632			
	Women		Men		Women		Men		Women		Men	
N_{it}	1 264 722		1 587 173		1 107 859		1 460 352		795 133		1 091 038	
N_i	169 670		192 967		153 268		181 543		118 260		149 372	
	Frequency (in %)	Average hourly wages (2015 €)	Frequency (in %)	Average hourly wages (2015 €)	Frequency (in %)	Average hourly wages (2015 €)	Frequency (in %)	Average hourly wages (2015 €)	Frequency (in %)	Average hourly wages (2015 €)	Frequency (in %)	Average hourly wages (2015 €)
Age												
23-24	6,8	9,4	6	9,9	6,1	9,4	5,5	9,9	4,4	9,3	4,1	10
25-29	16	10,6	15,4	11,7	15,7	10,6	15	11,5	14,8	10,7	14,2	11,7
30-34	14,8	11,8	15,3	13,7	14,8	11,9	15,4	13,4	14,6	12,2	15,3	13,6
35-39	14,3	12,5	15	15,1	14,4	12,5	15,1	14,8	14,4	12,9	15,3	15
40-44	14,2	12,8	14,4	16,2	14,5	12,8	14,7	15,9	14,8	13,1	15,1	16,1
45-49	13,5	12,9	13,2	16,9	13,7	12,9	13,4	16,6	14,4	13,2	14	16,9
50-54	11,7	13	11,7	17,3	12	13,1	11,9	17,3	12,8	13,3	12,4	17,5
55-59	8,7	13,6	9	18,5	8,8	13,4	9	18,4	9,8	13,6	9,5	18,6
Industry												
Construction	1,8	12,7	11,8	13,4	1,9	12,7	12,3	13,1	2,1	13,3	12,6	13,7
Manufacturing	12,5	12,8	23,7	15,6	13,5	12,8	25	15,4	14,3	13,2	26,3	15,7
Services	66	12	48,9	14,8	64,2	12,1	46,6	14,9	63,2	12,7	44,8	15,7
Trade	19,7	10,7	15,6	13,8	20,3	10,7	16,1	13,6	20,4	11,2	16,2	14,1

Base sample includes all individuals aged 20 to 60 that have positive employment in the private sector at time t. Censoring excludes individuals that that work less than 45 days a year, less than 1/8 of the legal duration a week, or paid less than 90% of the minimum hourly wage. Final sample includes only individual that are over this threhold at time t, $t-1$ and at least twice between $t-5$ and $t-5$. Figures for the final sample are computed at time $t-1$.

Figure C.1 – Consequences of sample selection with respect to childbirths



D Accounting decomposition

The log-change in total labor earnings between time $t-1$ and time $t+k$ is written as

$$\Delta y_{t+k} = \log\left(\mathbb{E}[y_{i,t+k}]\right) - \log\left(\mathbb{E}[y_{i,t-1}]\right) \tag{16}$$

Δy_{t+k} can also be rewritten as

$$\Delta y_{t+k} = \log\left(\frac{\mathbb{E}[\frac{y_{i,t+k}}{y_{i,t-1}} y_{i,t-1}]}{\mathbb{E}[y_{i,t-1}]}\right) \tag{17}$$

This formulation is particularly relevant here since we require that all individuals be employed at time $t-1$, so $y_{i,t-1} > 0$. Hence, Δy_{t+k} is simply the log-average of individual changes $y_{i,t+k}/y_{i,t-1}$ weighted by initial earnings $y_{i,t-1}$.

Next, we use an accounting decomposition of labor earnings at the individual level. First, using the law of iterated expectations yields

$$\begin{aligned}\mathbb{E}[y_{i,t+k}] =& \mathbb{P}(d_{i,t+k} = 0)\mathbb{E}[y_{i,t+k}|d_{i,t+k} = 0] \\ &+ \mathbb{P}(d_{i,t+k} = 1)\mathbb{E}[y_{i,t+k}|d_{i,t+k} = 1]\end{aligned} \tag{18}$$

Since $d_{i,t+k} = 0 \Rightarrow y_{i,t+k} = 0$, the first term vanishes:

$$\begin{aligned}\Delta y_{t+k} =& \log(\mathbb{P}(d_{i,t+k} = 1)) + \log(\mathbb{E}[y_{i,t+k}|d_{i,t+k} = 1]) - \log(\mathbb{E}[y_{i,t-1}]) \\ =& \underbrace{\log(\mathbb{P}(d_{i,t+k} = 1))}_{\text{Participation}} + \underbrace{\log(\mathbb{E}[y_{i,t-1}|d_{i,t+k} = 1]) - \log(\mathbb{E}[y_{i,t-1}])}_{\text{Selection}} \\ &+ \underbrace{\log(\mathbb{E}[y_{i,t+k}|d_{i,t+k} = 1]) - \log(\mathbb{E}[y_{i,t-1}|d_{i,t+k} = 1])}_{\Delta y_{t+k}^{\text{Participants}}}\end{aligned} \tag{19}$$

We are thus left with the decomposition of the latter term $\Delta y_{t+k}^{\text{Participants}}$; for these participants, all components of labor earnings – days, hours and hourly wages –

are observed in the data. Then,

$$\Delta y_{t+k}^{\text{Participants}} = \underbrace{\log\left(\frac{\mathbb{E}\left[\frac{w_{i,t+k}}{w_{i,t-1}} x_{i,t+k} h_{i,t+k} w_{i,t-1} | d_{i,t+k} = 1\right]}{\mathbb{E}\left[x_{i,t+k} h_{i,t+k} w_{i,t-1} | d_{i,t+k} = 1\right]}\right)}_{\text{Hourly wages growth}} + \log\left(\frac{\mathbb{E}\left[x_{i,t+k} h_{i,t+k} w_{i,t-1} | d_{i,t+k} = 1\right]}{\mathbb{E}[x_{i,t-1} h_{i,t-1} w_{i,t-1} | d_{i,t+k} = 1]}\right) \tag{20}$$

We continue to perform similar substitutions in the second term with respect to the two remaining components (hours and days). It follows that

$$\underbrace{\Delta y_{t+k}}_{\text{Labor earnings changes}} = \underbrace{\log\left(\mathbb{P}(d_{i,t+k} = 1)\right)}_{\text{Participation}} + \underbrace{\log\left(\frac{\mathbb{E}[y_{i,t-1} | d_{i,t+k} = 1]}{\mathbb{E}[y_{i,t-1}]}\right)}_{\text{Selection}} + \underbrace{\log\left(\frac{\mathbb{E}\left[\frac{x_{i,t+k}}{x_{i,t-1}} x_{i,t-1} h_{i,t-1} w_{i,t-1} | d_{i,t+k} = 1\right]}{\mathbb{E}\left[x_{i,t-1} h_{i,t-1} w_{i,t-1} | d_{i,t+k} = 1\right]}\right)}_{\text{Changes in Days Worked}} + \underbrace{\log\left(\frac{\mathbb{E}\left[\frac{h_{i,t+k}}{h_{i,t-1}} x_{i,t+k} h_{i,t-1} w_{i,t-1} | d_{i,t+k} = 1\right]}{\mathbb{E}\left[x_{i,t+k} h_{i,t-1} w_{i,t-1} | d_{i,t+k} = 1\right]}\right)}_{\text{Changes in Hours Per Day}} + \underbrace{\log\left(\frac{\mathbb{E}\left[\frac{w_{i,t+k}}{w_{i,t-1}} x_{i,t+k} h_{i,t+k} w_{i,t-1} | d_{i,t+k} = 1\right]}{\mathbb{E}\left[x_{i,t+k} h_{i,t+k} w_{i,t-1} | d_{i,t+k} = 1\right]}\right)}_{\text{Hourly Wage Growth}} \tag{21}$$

This accounting identity clarifies that the (reweighted) log-average of individual earnings' changes can be decomposed into the sum of (reweighted) log-average of individual changes for each component, and a selection term.

E Additional results

E.1 Long-term child penalties

Our results suggest that the arrival of a child indeed results in a short-run shift in hourly wages for mothers at the top of the wage distribution, after which catching-up may occur but is not sufficient for women to recover in comparison with their male counterparts (Figures 1 and 2). We further investigate the impact of children on hourly wages in the longer run to better understand how the arrival of a child impacts the wage rate past the first few years after childbirth. The motivation here entails the possibility that mothers recover at least partly from the negative hourly wage shock that childbirth leads to. It could also be that there is in fact no recovery from childbirths, i.e., that the latter are a source of some permanent hourly wage shift, and that afterwards men and women experience similar career progressions, i.e., have parallel wage growth trajectories. The last possibility corresponds to men and women having diverging hourly wage levels due to the arrival of children so that in the long run the child penalty increases over time, e.g., because mothers spend less time in the labor market and thus experience slower wage growth than do their male counterparts.

We now focus on the gender gap in hourly wage growth among parents of children aged 6 or above and compare it to the gap that prevails among nonparents. After the age of 6, most children attend school; thus, time constraints become less stringent. We resort to the same methodology as before, and our control group is composed of nonparents; however, the analogue of the time dimension in a standard difference-in-difference approach is replaced here by the gender dimension. Figure E.1 displays the results of an OLS estimation, the outcome being 1-year hourly wage growth.[25] We resort to this strategy because there are significantly fewer observations for which we know both pre-childbirth wage levels and long-run hourly wages: typically, our previous approach yields only 10-year child penalties for individuals who experienced a childbirth in 2005; the small sample size leads

[25]The gender gap in hourly wage growth among nonparents, our control group, is discussed in Appendix E.2.

to large standard errors and renders the results uninterpretable (see Figure 3).

In the lowest part of the distribution, the gender gap is similar or slightly larger among parents, and amounts to approximately 0.3 log-points per year. Among individuals belonging to the highest half of the distribution, the sign is reversed, which suggests that some of high-achieving mothers may recover some part of their children-related wage losses (Model 1). These results suggest therefore the possibility of long-run effects of childbirths on their mothers' career progressions at the bottom of the distribution. Controlling for horizontal segregation somehow lowers the estimates (Model 2), even though the difference is not statistically significant. This empirical evidence is consistent with the arrival of a child resulting in gender differentials in sorting across firms (Coudin, Maillard, and Tô, 2018), which is a plausible explanation for this observed gap. Surprisingly, controlling further for both experience and job mobility widens this gap (Model 3).

In summary, even if childbirths do not result in large short-run reductions in hourly wages at the bottom of the distribution, which stems most likely from the minimum wage (and more generally from the institutional setting), they may lead to a sticky floor pattern in the long run. In contrast, at the top of the distribution, mothers may well experience short-run shifts in hourly wages; however, several years after their last child is born, they experience slightly better career progression than do their male counterparts, which may enable mothers to partly catch-up with men.

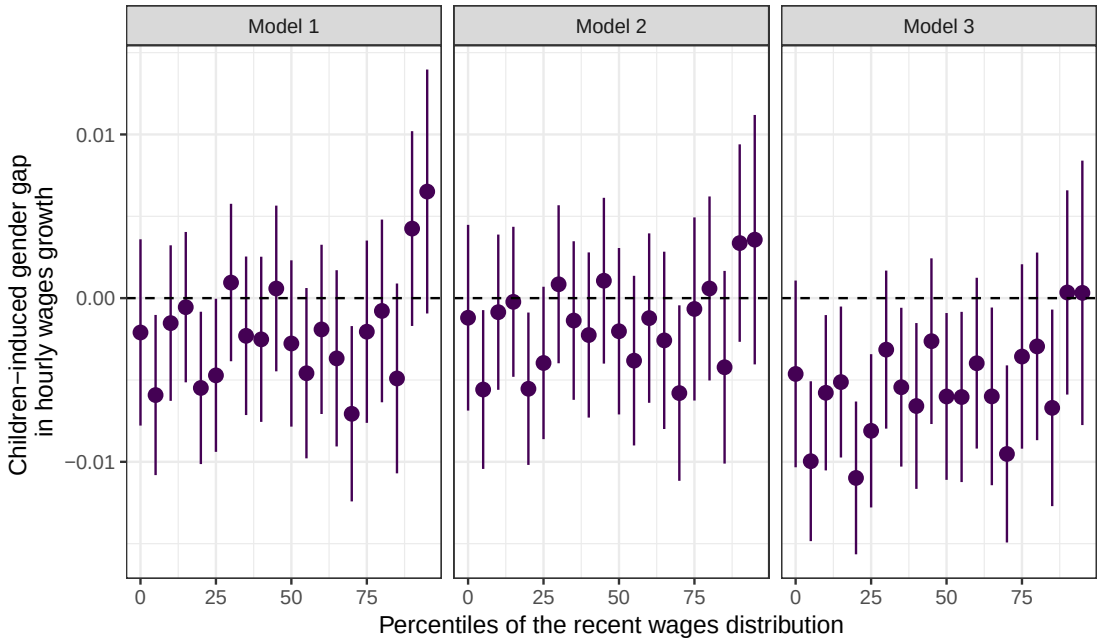
E.2 Career progression among nonparents

We show that men and women still experience different career progressions even if they do not have children. We focus on individuals with no observed child in the data; such individuals could experience childbirth after 2015. We estimate gender differences in hourly wage growth along the entire recent wage distribution, as Figure E.2 shows.

We obtain a U-shaped pattern: at both ends of the distribution, women experience slower hourly wage growth than do their male counterparts, while among median workers gender differences are not statistically significant (Model 1). Though

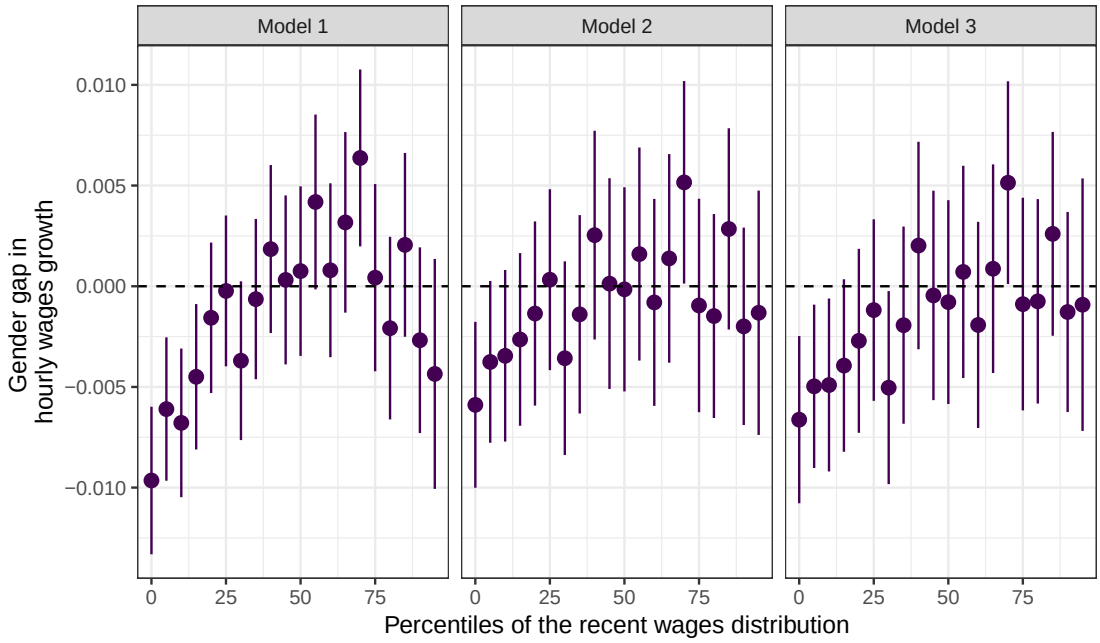
the difference is not large (approximately 1 log-point), such differentials accumulate over time and are the source of substantial wage gaps. If we control for horizontal segregation (Model 2) as well as for experience and job mobility (Model 3), the difference shrinks at both ends of the distribution. As far as low earners are concerned, the gap remains barely significant even in the full control specification: the sticky floor effect is not merely the consequence of children only but instead causes the respective women to be progressively left behind in the bottom of the distribution.

Figure E.1 – Children-induced gender differences in hourly wage growth among parents of older children



Estimates of coefficients related to gender (a female dummy variable) × having all children aged above 6, relative to gender × never having children, interacted with location in the recent wage distribution, in an hourly wage growth model. The outcome is a (logarithmic) hourly wage growth between times $t-1$ and t. Model 1 includes no controls. Model 2 controls for year, age, industry, firm composition (share of women working part-time) and 1-digit occupation within recent wage cell. Model 3 includes all of these controls as well as experience between t and $t+5$ and having changed firms between times $t-1$ and t. Standard errors are clustered at the individual level. The sample includes individuals up to age 60 at time t.

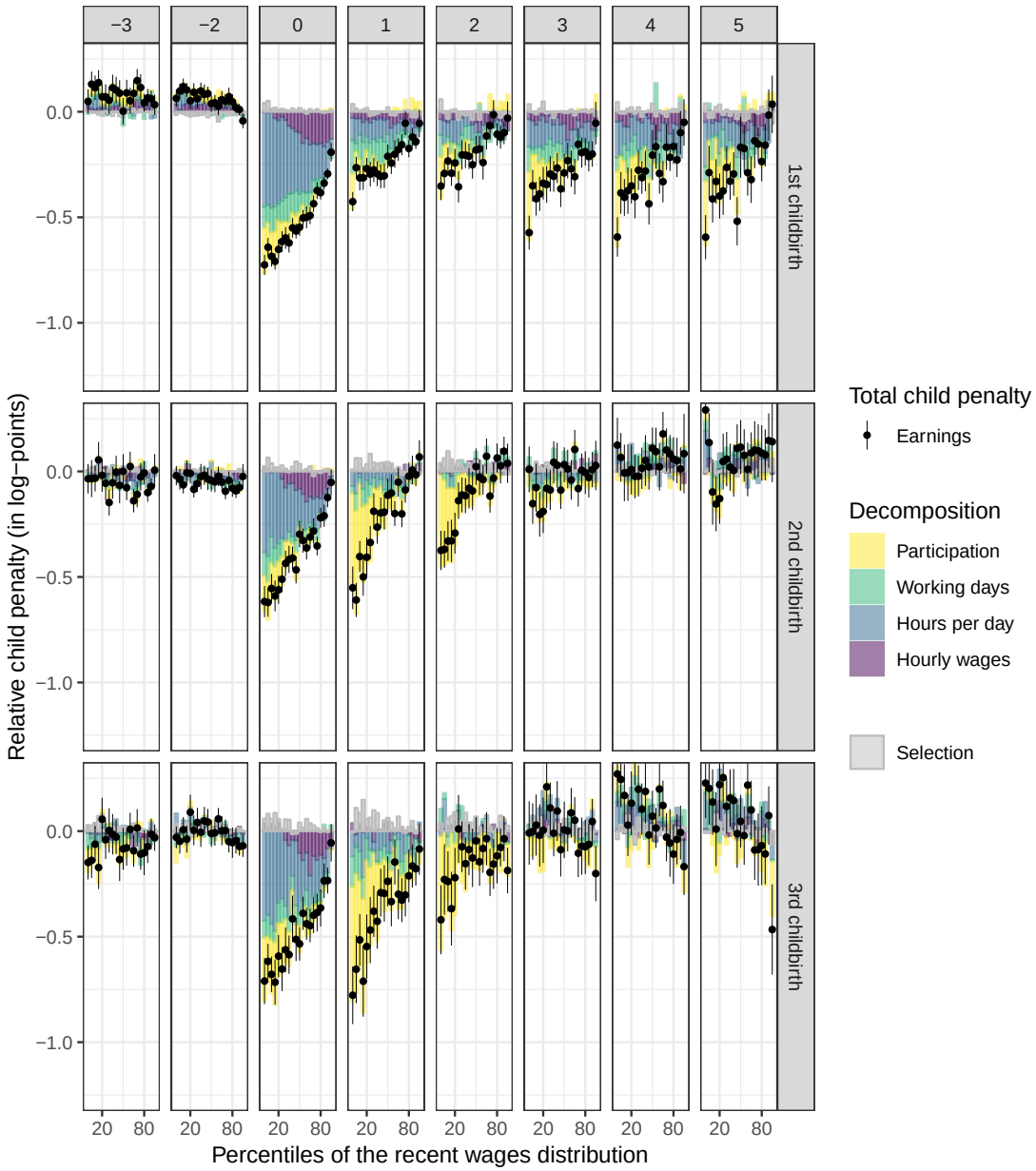
Figure E.2 – Gender differences in hourly wage growth among nonparents



Estimates of coefficients related to gender (a female dummy variable) in the hourly wage growth model that interacts gender and rank in the recent wage distribution. The outcome is a (logarithmic) hourly wage growth between times $t-1$ and t. Model 1 includes no controls. Model 2 controls for year, age, industry, firm composition (share of women working part-time) and 1-digit occupation within recent wage cell. Model 3 includes all of these controls as well as experience between times t and $t+5$ and having changed firms between times $t-1$ and t. Standard errors are clustered at the individual level. The sample includes individuals up to age 60 at time t.

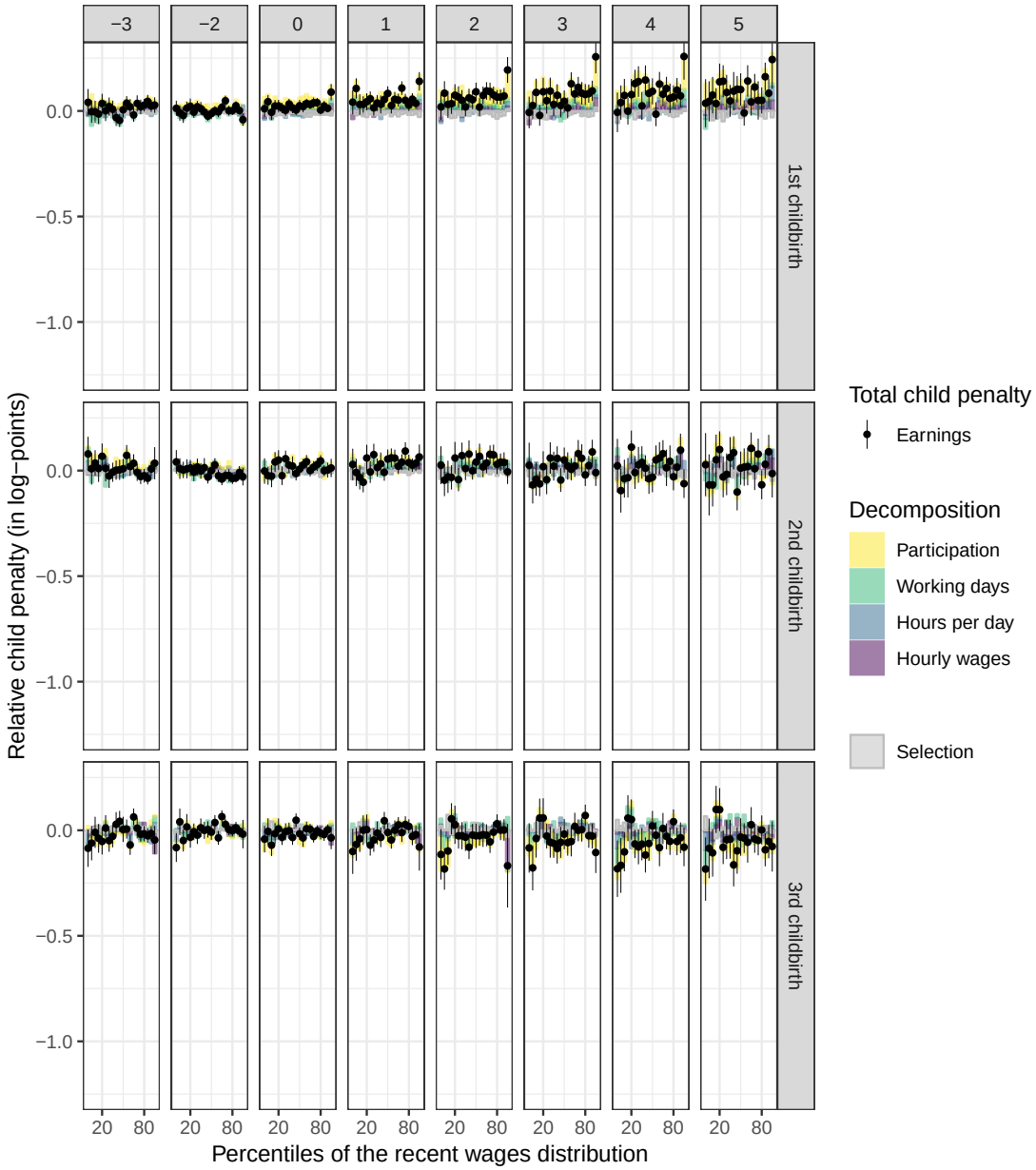
F Robustness checks

Figure F.1 – Consequences of childbirth for women's labor outcomes: a comparison with a control group determined at the imputed age of childbirth



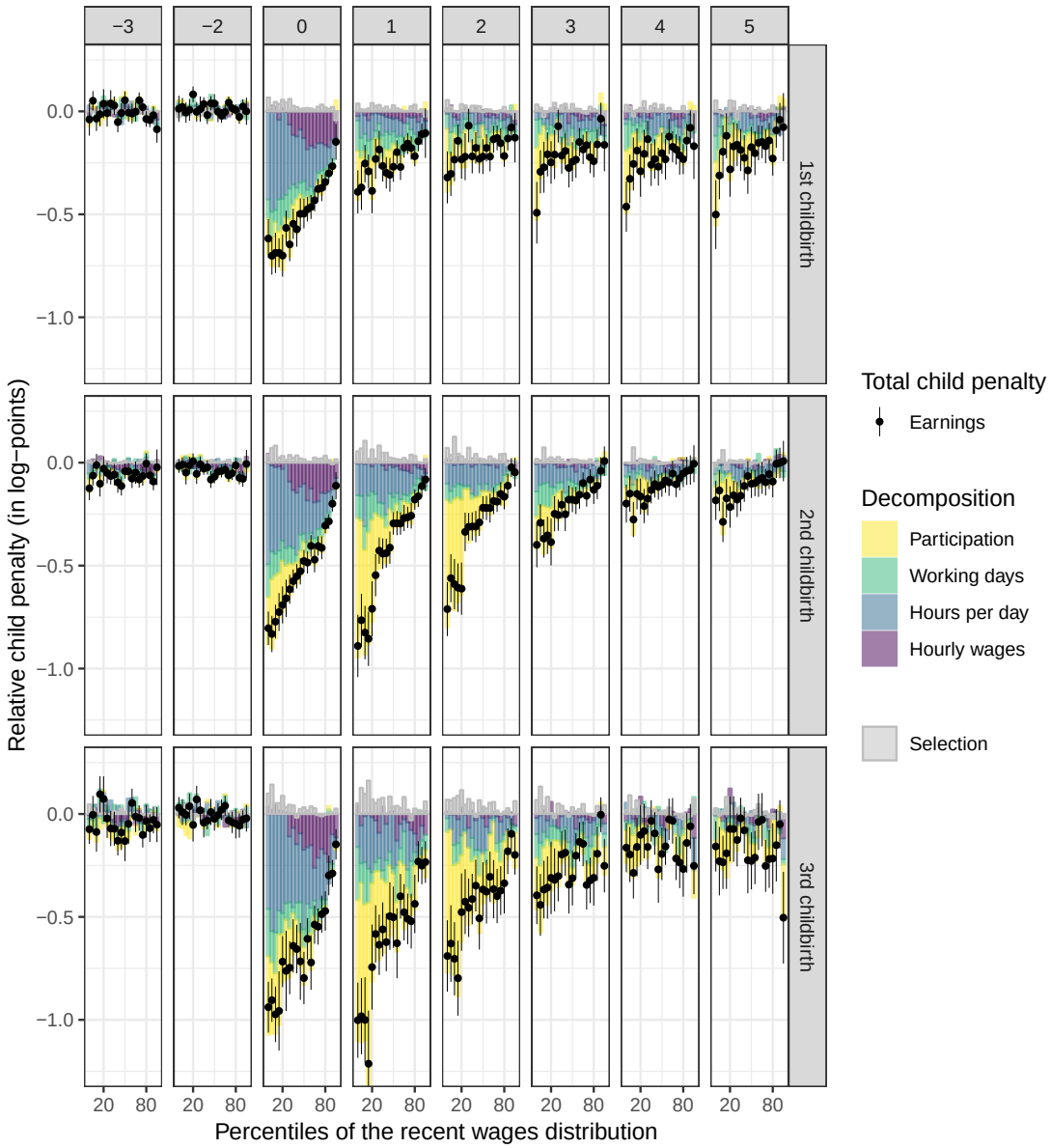
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The control group is determined at age randomly drawn from the distribution of age at the nth childbirth within gender × cohort × education cells (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.2 – Consequences of childbirth for men's labor outcomes: a comparison with a control group determined at the imputed age of childbirth



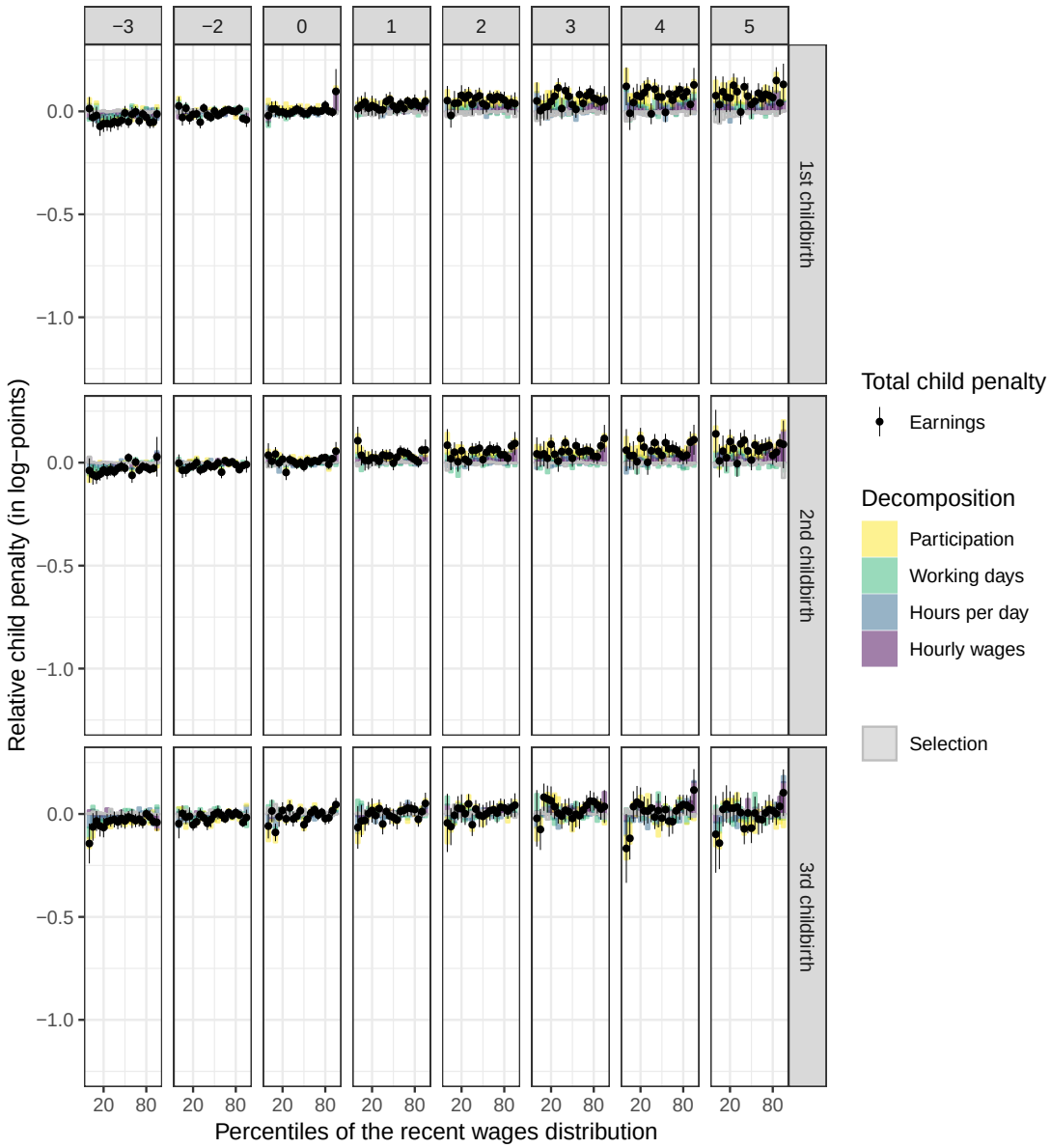
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The control group is determined at age randomly drawn from the distribution of age at the nth childbirth within gender × cohort × education cells (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.3 – Consequences of childbirth for women's labor outcomes: restriction to older cohorts that have made complete fertility decisions



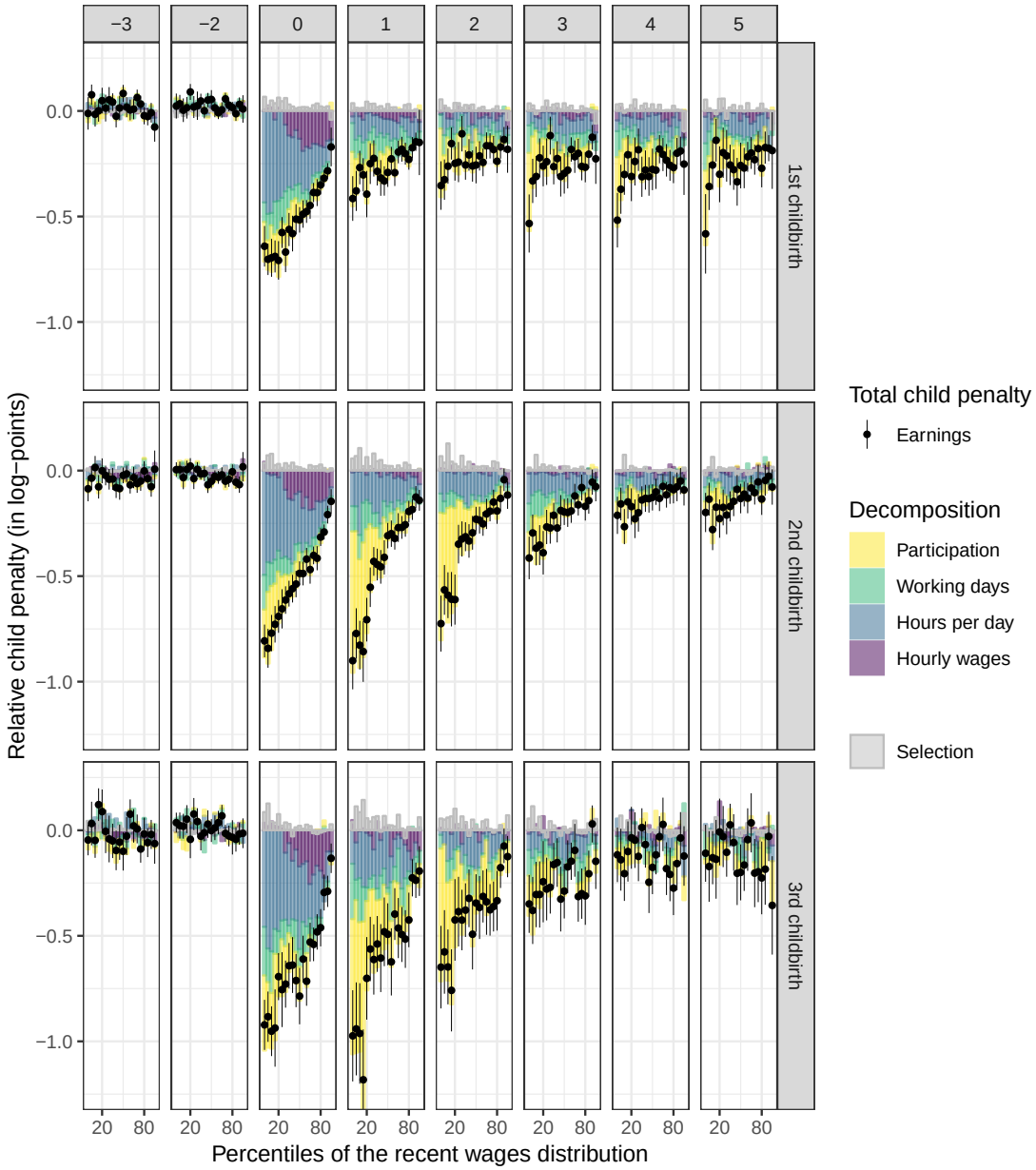
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The sample is restricted to individuals born in 1975 or earlier (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.4 – Consequences of childbirth for men's labor outcomes: restriction to older cohorts that have made complete fertility decisions



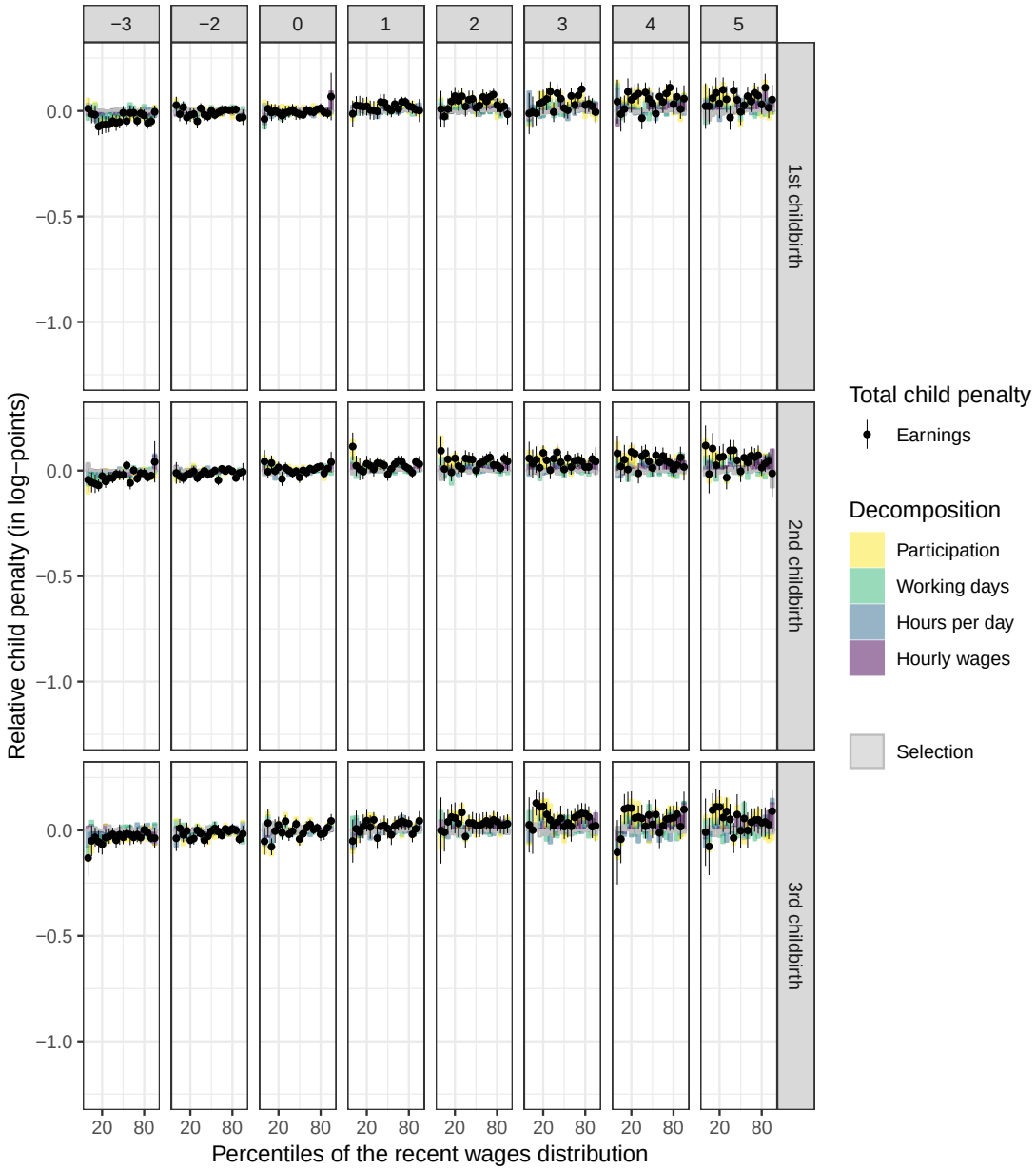
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The sample is restricted to individuals born in 1975 or earlier (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.5 – Consequences of childbirth for women's labor outcomes: identification based on the timing of the kth childbirth



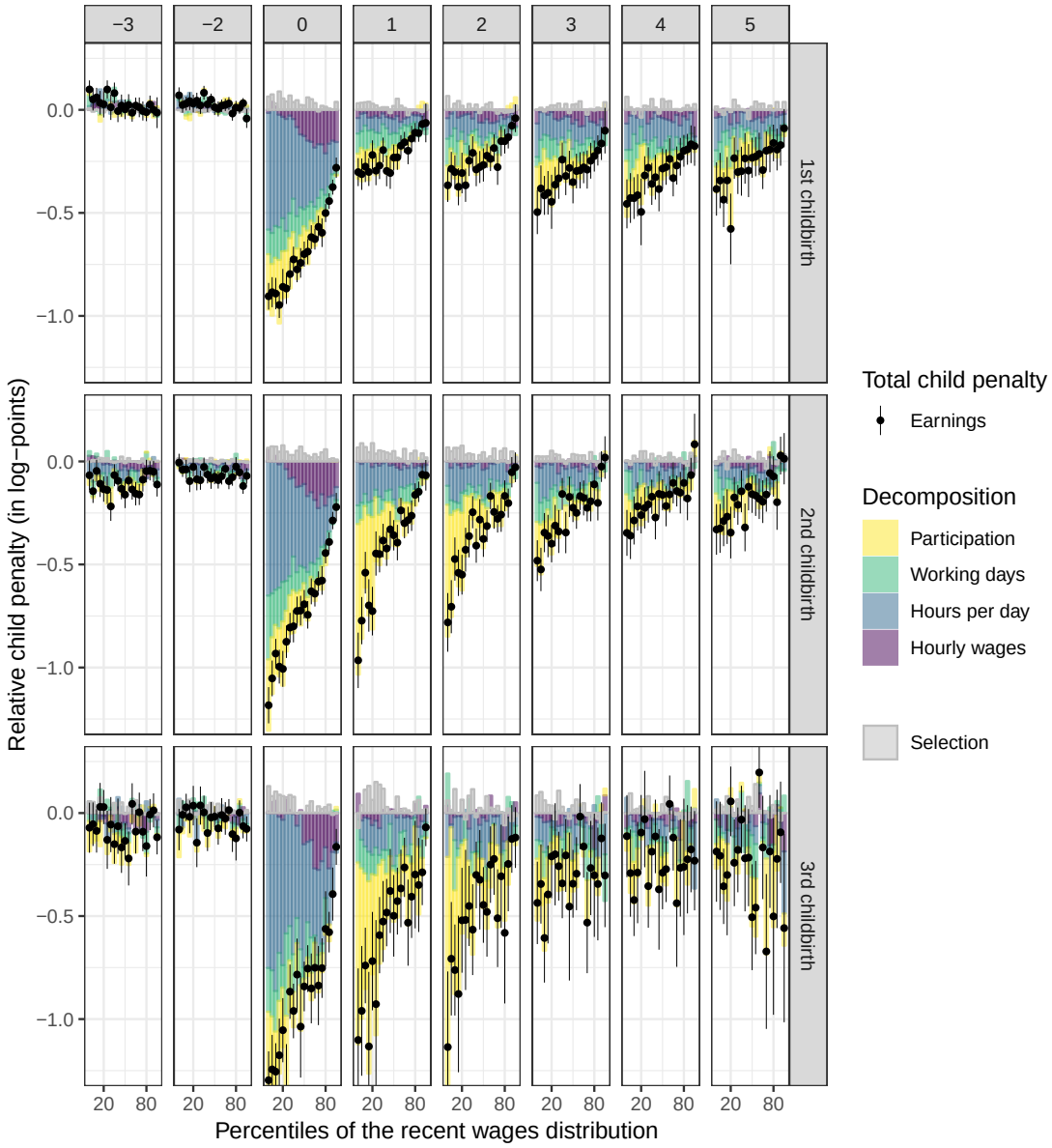
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The control group includes individuals with exactly n children in 2015 and who do not experience childbirth between $t-1$ and $t+k$ (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.6 – Consequences of childbirth for men's labor outcomes: identification based on the timing of the kth childbirth



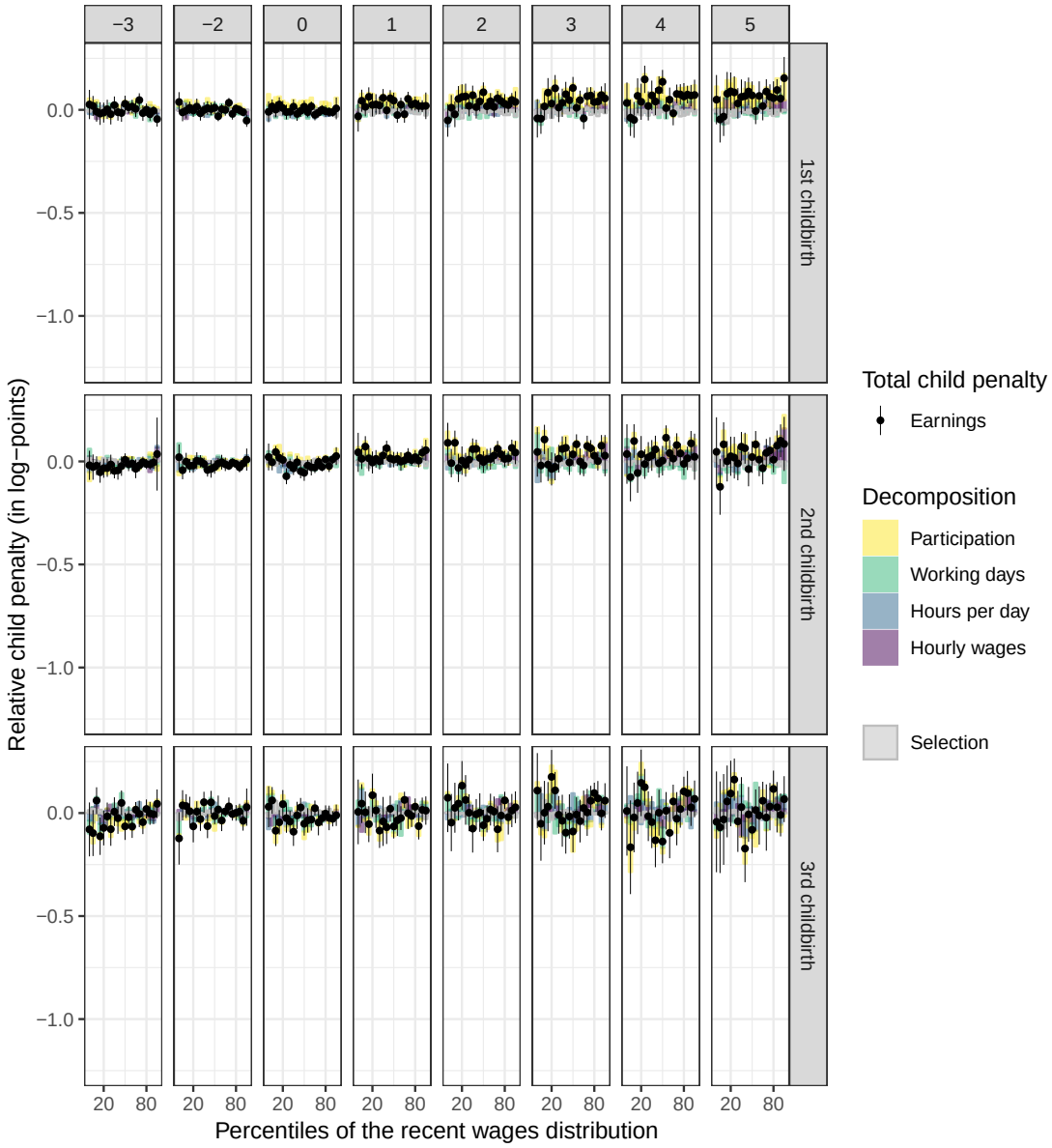
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The control group includes individuals with exactly n children in 2015 and who do not experience childbirth between $t-1$ and $t+k$ (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.7 – Consequences of childbirth for women's labor outcomes: restriction to childbirths in the second quarter



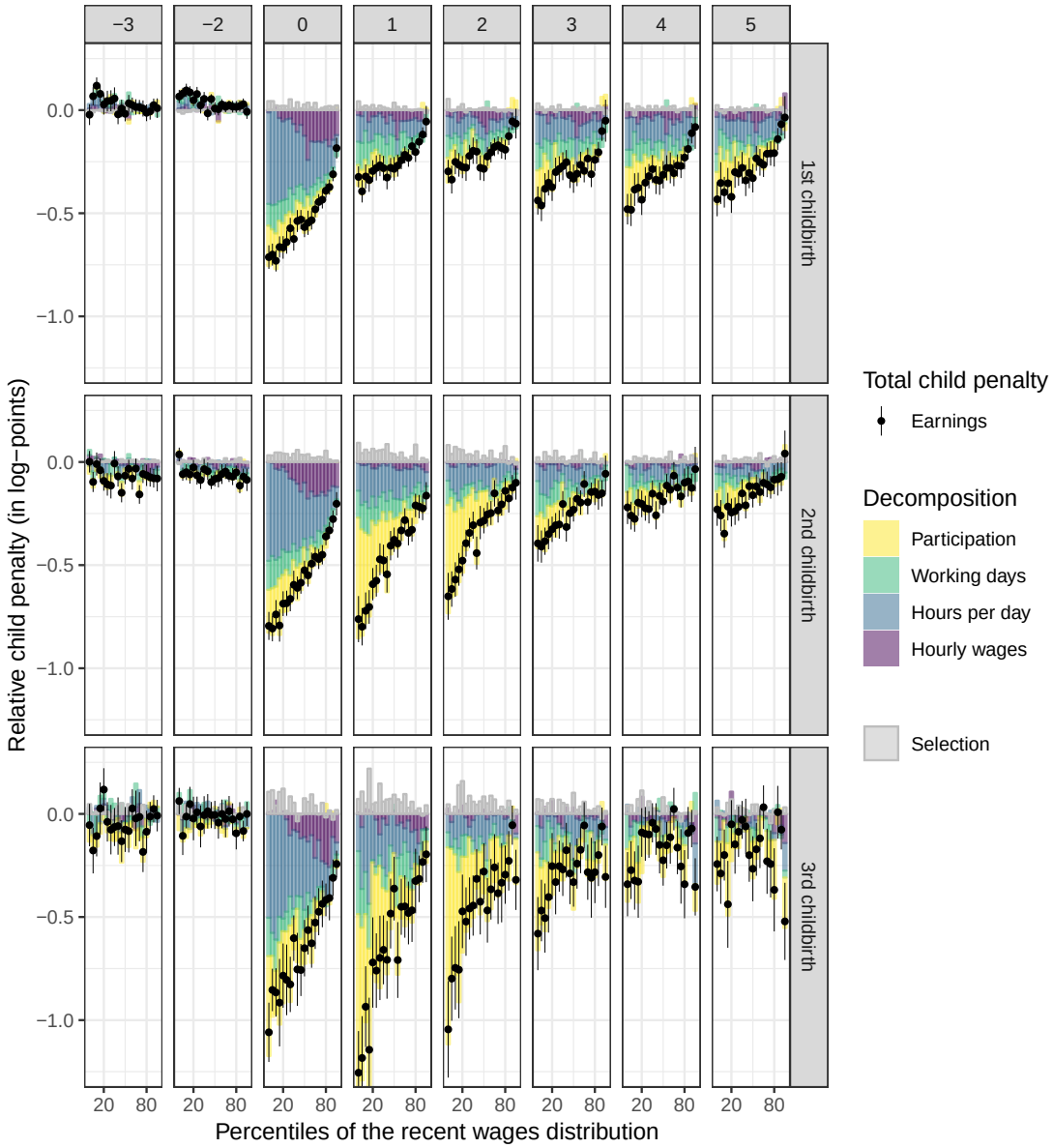
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The treated group is restricted to individuals that experience the nth childbirth during the second quarter of year t (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.8 – Consequences of childbirth for men's labor outcomes: restriction to childbirths in the second quarter



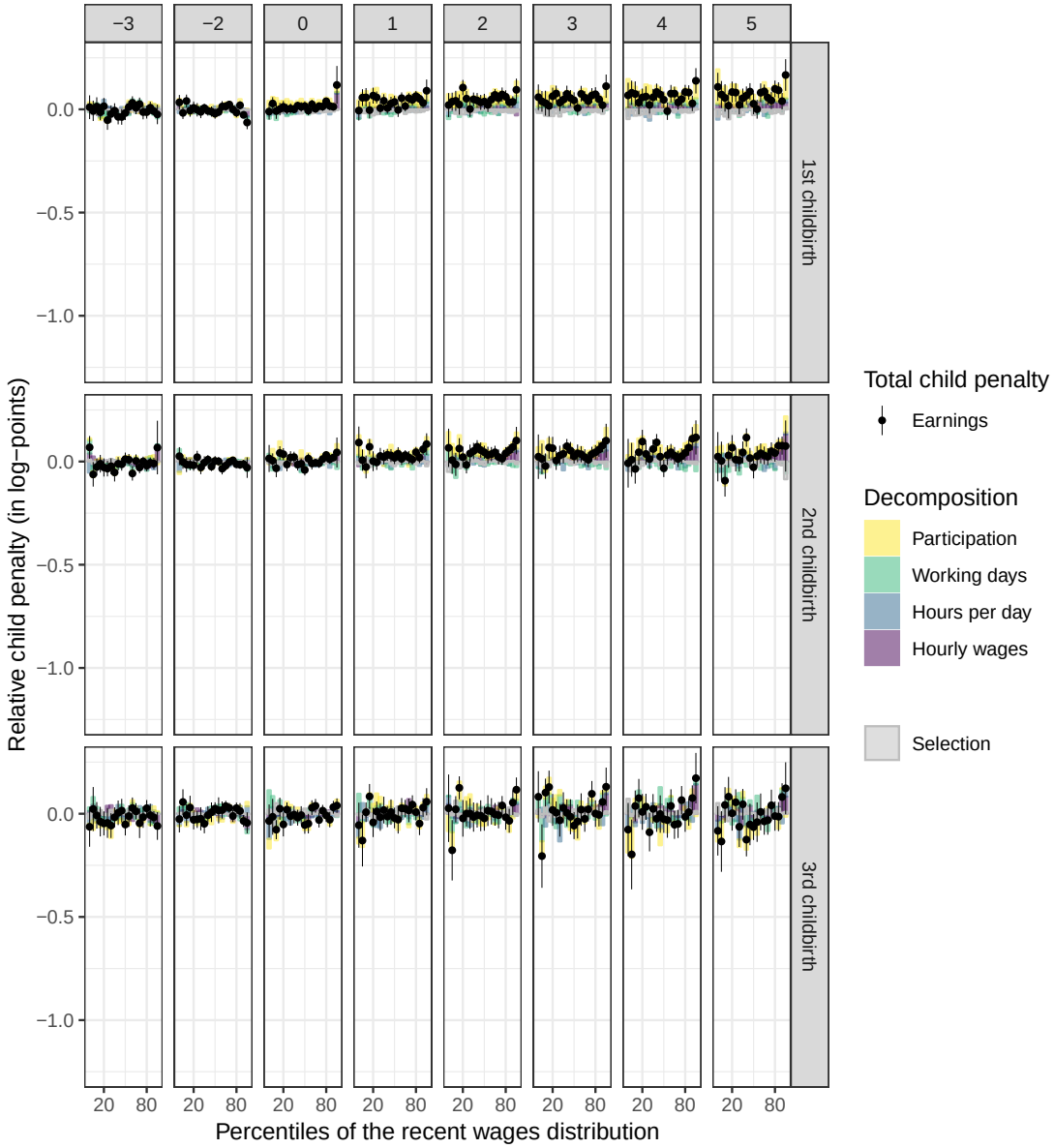
Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The treated group is restricted to individuals that experience the nth childbirth during the second quarter of year t (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.9 – Consequences of childbirth for women's labor outcomes: restriction to childbirths in 2007-2010



Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The treatment time is restricted to years 2007 to 2010 (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

Figure F.10 – Consequences of childbirth for men's labor outcomes: restriction to childbirths in 2007-2010



Each panel displays the estimates of child penalties obtained by the difference-in-difference method (see 7) for various values of time-to-childbirth expressed in years. The treatment time is restricted to years 2007 to 2010 (see 5.1). Bootstrapped standard errors using 100 replications are clustered at the individual level.

G Childcare preferences

We rely on the complemented LFS data to estimate the relation between preferences towards childcare and hourly wages. To this end, we estimate a log-hourly wage equation in which we introduce a dummy variable which indicates that the individual views childcare provided by parents as the best childcare solution. Table 2 displays our results. Consistent with Figure 13, we find a small negative correlation between preferences and hourly wages among those for which information regarding hourly wages is available in the LFS, i.e. those who were in salaried employment at the time they were surveyed. This correlation is not significant at usual levels so that we cannot reject the null hypothesis that childcare preferences are not correlated with hourly wages in this selected sample. Additionally, this correlation reverses for men, and shrinks substantially for women when education is accounted for.

Relying on observed hourly wages only, this result may yet be misleading due selection bias. Specifically, women with the lowest potential hourly wages and the largest taste for childcare are more likely to leave the labor force, which may lead to understate the correlation between childcare preferences and potential hourly wages. We therefore estimate a Heckman sample selection model. We take advantage of one question of the survey that asked respondent who were either out of the labor force or worked part-time whether their labor supply decision was the result of problems regarding childcare. We therefore rely on this dummy variable, assuming that it is equal to 0 for all individuals working full-time; the exclusion restriction is that childcare issues affect the participation decision but do not directly affect hourly wages. It turns out that women who hold the most conservative views regarding childcare or who face problems regarding the access to external childcare are more likely to be outside the labor force. However, correcting for sample selection does not change substantially our estimates of the relation between hourly wages and childcare preferences. As a result, we do not find significant effects of childcare preferences on potential hourly wages.

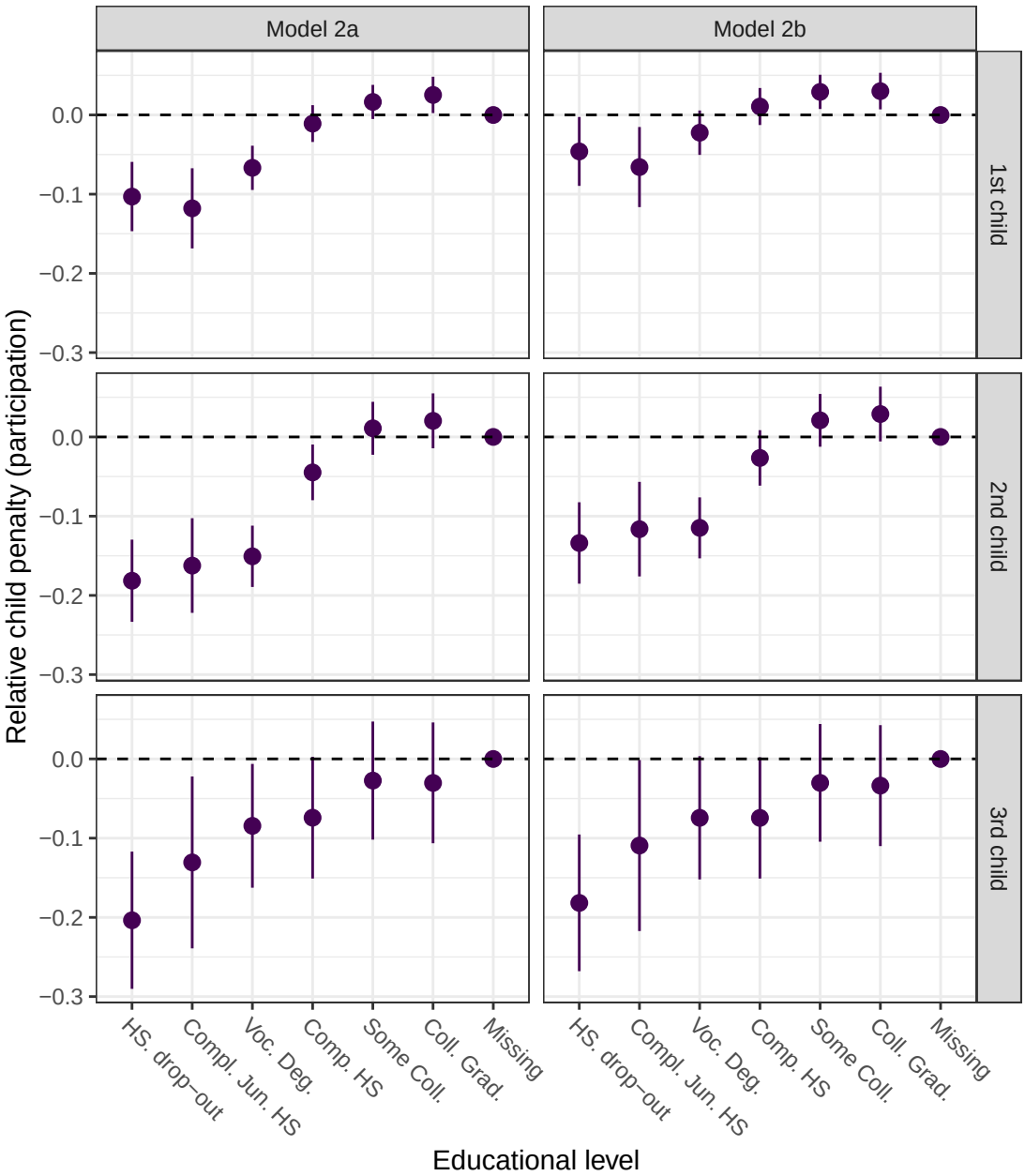
Table 2 – Correlation between childcare preferences and hourly wages

	(1)	(2)	(3)	(4)
Log hourly wage				
Parents	-0.055	0.011	-0.055	0.010
	(0.035)	(0.031)	(0.035)	(0.031)
Parents × Female	-0.032	-0.037	-0.030	-0.033
	(0.058)	(0.049)	(0.061)	(0.049)
Selection				
Parents			0.16	0.18
			(0.098)	(0.098)
Parents × Female			-0.52	-0.52
			(0.14)	(0.14)
Childcare issues			0.15	0.075
			(0.41)	(0.38)
Childcare issues × Female			-1.00	-0.92
			(0.43)	(0.41)
Correlation between residuals				
$\hat{\rho}$			-0.015	-0.034
			(0.15)	(0.089)
Additional controls				
Age youngest child	✓	✓	✓	✓
Age	✓	✓	✓	✓
Education		✓		✓
Observations	1151	1151	2550	2550

The sample is restricted to individuals with children aged 3 or less. All covariates are further interacted with a female dummy. Standard errors in parentheses.

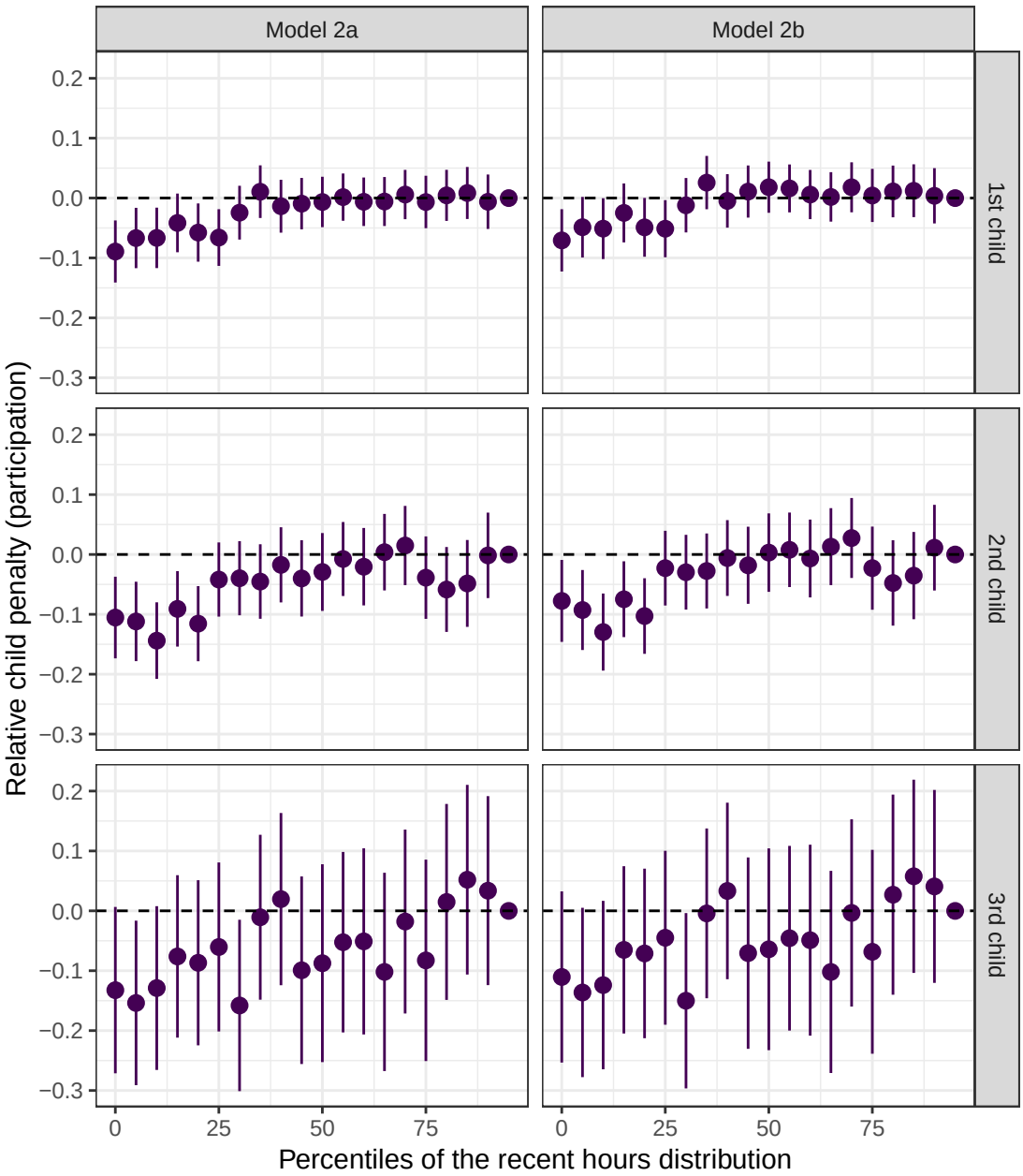
H Other sources of heterogeneity

Figure H.1 – Heterogeneity in the probability of remaining employed one year after childbirth: education



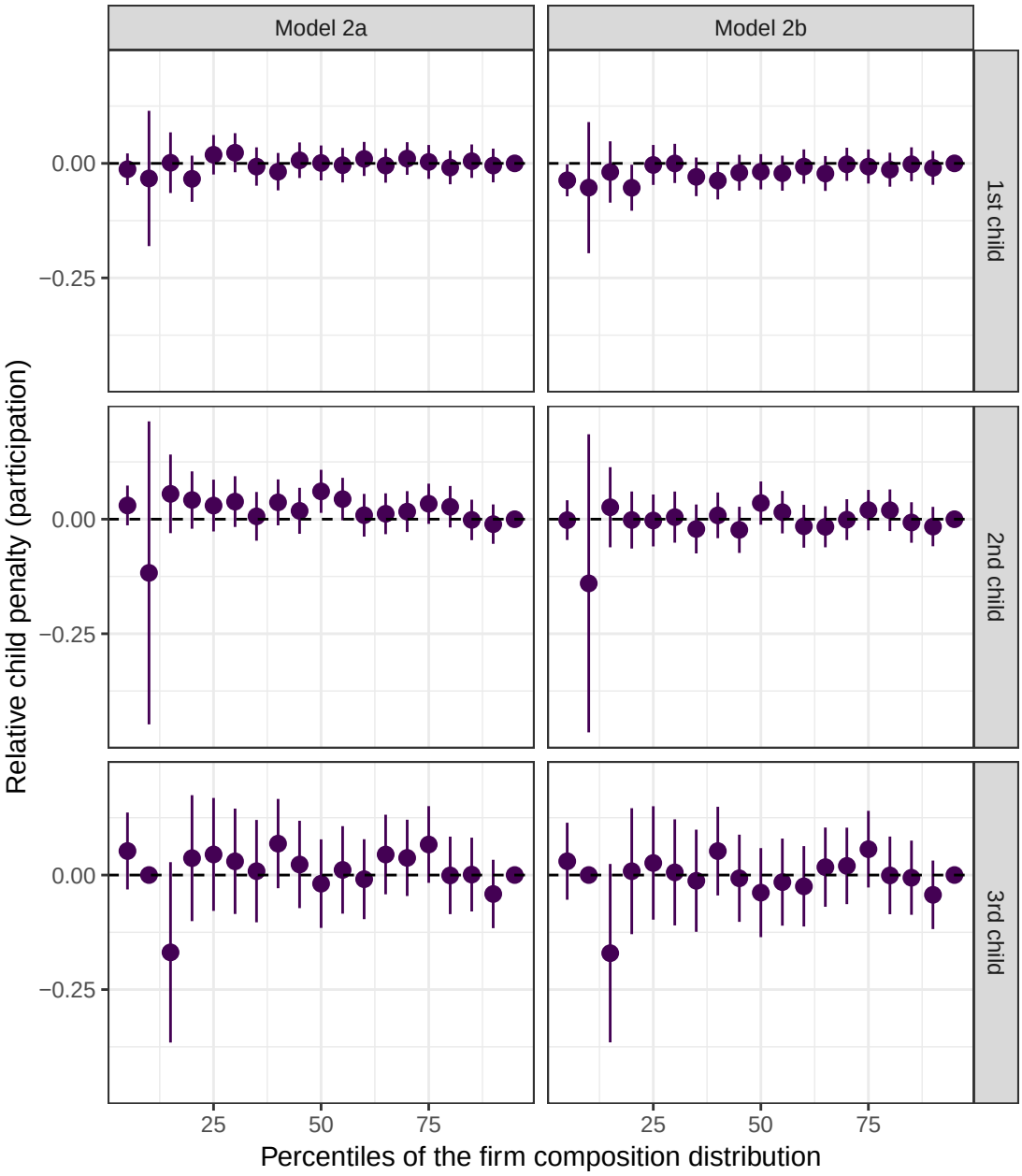
Estimates of coefficients related to childbirth for women in a linear probability model that interacts a double-difference setting with gender and rank in the recent wage distribution, education, rank in the distribution of recent paid hours and rank in the distribution of firm composition (15). Individuals with missing education data are considered as a reference. The outcome is a dummy variable for participating in the labor market at time $t+1$. Model 2a includes no controls. Model 2b controls for year, age, industry and 1-digit occupation within each cell. Standard errors are clustered at the individual level. The sample includes individuals up to age 59 at time t.

Figure H.2 – Heterogeneity in the probability of remaining employed one year after childbirth: past labor supply decisions



Estimates of coefficients related to childbirth for women in a linear probability model that interacts a double-difference setting with gender and rank in the recent wage distribution, education, rank in the distribution of recent paid hours and rank in the distribution of firm composition (15). Individuals who belong to the highest percentile group according to recent paid hours are considered as a reference. The outcome is a dummy variable for participating in the labor market at time $t+1$. Model 2a includes no controls. Model 2b controls for year, age, industry and 1-digit occupation within each cell. Standard errors are clustered at the individual level. The sample includes individuals up to age 59 at time t.

Figure H.3 – Heterogeneity in the probability of remaining employed one year after childbirth: firm composition



Estimates of coefficients related to childbirth for women in a linear probability model that interacts a double-difference setting with gender and rank in the recent wage distribution, education, rank in the distribution of recent paid hours and rank in the distribution of firm composition (15). Individuals whose main employer at time $t-1$ was in the highest percentile group in terms of the share of women working part-time among its employees are considered as a reference. The outcome is a dummy variable for participating in the labor market at time $t+1$. Model 2a includes no controls. Model 2b controls for year, age, industry and 1-digit occupation within each cell. Standard errors are clustered at the individual level. The sample includes individuals up to age 59 at time t.