

Gender norms and the child penalty: evidence from the Dutch bible belt

Sara Rellstab

To cite this article: Sara Rellstab (2024) Gender norms and the child penalty: evidence from the Dutch bible belt, Applied Economics, 56:45, 5428-5441, DOI: 10.1080/00036846.2023.2257030

To link to this article: <https://doi.org/10.1080/00036846.2023.2257030>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 10 Oct 2023.



[Submit your article to this journal](#)



Article views: 895



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

Gender norms and the child penalty: evidence from the Dutch bible belt

Sara Rellstab ^{a,b}

^aErasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands; ^bUniversità della Svizzera Italiana, Lugano, Switzerland

ABSTRACT

There are substantial gender gaps in earnings once children are born in many developed countries, despite similar education levels of men and women and accessible childcare facilities. I examine whether gender norms are a driver of women's higher labour market costs of having children using Dutch administrative data. Exploiting large local variation in gender norms in the Netherlands, I compare parents from the Dutch bible belt, where gender norms are on average less egalitarian, with parents from other regions in the Netherlands, where gender norms are more egalitarian. My findings show that having children leads to an about 30% larger earnings decrease for women in the bible belt, mainly driven by a larger reduction in working hours. I rule out that differences in pre-child parental characteristics and institutions such as childcare availability explain this result and argue that different gender norms in these two areas likely explain this finding.

KEYWORDS

Child penalty; parenthood; gender norms; division of labour in the household

JEL CLASSIFICATION

J13; J16; J22



I. Introduction

Women's earnings drop importantly after the first child, while men's earnings are largely unaffected (for example, Kleven et al. 2019). The persistence of this negative earning effects of having children (or 'child penalty') for women is remarkable for two reasons. First, in the context of the Netherlands, young women are on average better educated than young men.¹ This advantageous starting position for women's careers implies that it is not clear a priori who in the couple should invest less in their career to take care of the children. Second, the increasing availability of childcare would give room for both parents to keep pursuing their careers with modest child penalties. In this article, I examine whether non-egalitarian gender norms can explain why large child penalties persist exploiting local variation of gender norms in the Netherlands.


Kleven et al. (2019) show that there is a correlation between gender norms and the child penalty on a country-level. However, examining the causal effect of gender norms on the child penalty is difficult for two reasons. First, norms often co-

evolve with institutions, and hence separating norms from institutional factors like family policies is not straightforward. Second, norms change slowly over time, which implies that quasi-experimental designs can rarely be used to study norms. As a result, I take a different route to establish the impact of gender norms on the child penalty.

I compare groups that live in a similar institutional framework and who are similar in baseline observables, but who differ starkly in held gender norms. The Netherlands represent an ideal setting to compare regions with different gender norms, as there is large regional variation on the municipal level. The so-called 'bible belt', stretching from the South West to the North East of the country, consists of municipalities with a high concentration of orthodox protestants. The location of these communities can be traced back to the reformation and is unrelated to modern policies and administrative entities such as provinces. While parents-to-be in the bible belt have similar access to childcare and the same legal institutions, held beliefs on gender are less egalitarian among orthodox protestants than in the rest of the Netherlands. Comparing

CONTACT Sara Rellstab  sara.rellstab@protonmail.com  Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, DR Rotterdam 3000, The Netherlands

¹In 2021, among the 25 to 45 years old, 57% of women have tertiary education in the Netherlands, compared to 49% of men (CBS Statline 2021).

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/00036846.2023.2257030>.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

child penalties between the bible belt and the rest of the country can thus be informative of the role of gender norms without the confounding effects of institutional differences.

The bible belt is defined by municipal voter shares for the political party representing the Dutch orthodox protestants, the *SGP* (*Staatskunding Gereformeerde Partij*). Combining Dutch administrative micro data with voting outcomes at municipal level, I divide Dutch first-time parents of 2008–2013 into two groups: parents residing inside or outside the bible belt in the year before their first child. Based on this classification, I estimate the child penalty for both groups of parents. When matching parents on educational and demographic characteristics in a robustness test, I show that the difference in child penalty between the two areas is not a product of these factors. I argue that prevailing gender norms are the most likely driver of the difference in child penalty between parents inside and outside the bible belt, and I show evidence that other factors are not likely to explain my results.

The impact of having children on labour market outcomes of their parents has been studied for many developed countries (for example, Andresen and Nix 2022; Angelov, Johansson, and Lindahl 2016; Bütikofer, Jensen, and Salvanes 2018; Cortés and Pan 2020; Kleven et al. 2019; Rabaté and Rellstab 2022; Rosenbaum 2021). All these studies find a large and persistent decline in labour market outcomes of mothers compared to fathers when having the first child. In the Netherlands, mothers earn around 46% less than predicted absent child birth, which is higher than in the Nordic countries like Denmark or Norway but lower than German-speaking countries like Germany and Austria (Rabaté and Rellstab 2022).

The persistence of the child penalty has led to research on its explanations. A group of study shows that biological differences between men and women (such as giving birth or breastfeeding) can only limitedly explain the existence of the child penalty. Comparing parents with biological and adoptive children, Rosenbaum (2021) finds that there is also a child penalty for adoptive mothers,

and Kleven et al. (2020b) shows that adoptive mothers have a similarly high child penalty than biological mothers in Denmark. This suggests that biological factors cannot explain the child penalty of women. In addition, female same-sex couples have a considerably lower child penalty than different-sex couples (Adema, Rabaté, and Rellstab 2020; Andresen and Nix 2022; Moberg 2016; Rabaté and Rellstab 2022; Rosenbaum 2021). This implies that when a woman gives birth to a child, a high child penalty is not an unavoidable consequence.

A second reason for the child penalty may be insufficient family policies. However, even though family policies are a necessary condition to reduce the child penalty, their existence alone does not guarantee a reduction in the child penalty, and evidence on the immediate effectiveness of these policies is mixed (see, for example, Andresen and Nix 2022, 2022; Bettendorf, Jongen, and Muller 2015; Dehos and Paul 2021; Kleven et al. 2020; Krapf, Roth, and Slotwinski 2020; Mari, Cutuli, and Pre-Print 2018; Olivetti and Petrongolo 2017; Patnaik 2019; Rabaté and Rellstab 2022). While studies evaluating the introduction of family policies find a small or negligible effect of these policies on the child penalty, countries with a long tradition of generous family policies, like the Nordic countries, have relatively low child penalties compared to countries with less generous or more recent family policies (Kleven et al. 2019; Olivetti and Petrongolo 2017). This seems contradictory but may be reconciled if such policies mostly have long-run effects by changing norms.

The limited short-term potential of family policies to decrease the child penalty and the small explanatory power of biology calls for additional explanations. Gender norms are a good candidate for explaining the persistence in child penalties, because various studies have shown that gender norms can alter labour market outcomes (Akerlof and Kranton 2000; Bertrand 2011; Bursztyn, Fujiwara, and Pallais 2017; Cardoso and Morin 2018; Farré and Vella 2013; Fernández 2007; Fernandez and Fogli 2009; Fortin 2015, 2005; Hicks, Santacreu-Vasut, and Shoham 2015; Vella 1994).²

²Preferences may represent an alternative explanation for the child penalty. However, preferences are likely to be shaped by norms and hence cannot be disentangled from norms.

Exploring the role of gender norms in the Netherlands may be promising, as Rabaté and Rellstab (2022) found a correlation between religiosity, less egalitarian gender norms, and high child penalties in Dutch data.

So far, the role played by culture and norms for the child penalty has only been causally investigated by three other studies, and these studies mostly focus on the child penalty in the probability of having paid work. First, Boelmann et al. (2021) use the German reunification and the resulting migration flows to show that exposure to culture of origin from East Germany leads to a larger post-birth female labour supply at the extensive margin; and that there is adaptation of West German mothers once exposed to East German culture. Second, Steinhauer (2018) shows that around the Swiss French-German language border, the employment of mothers with German-speaking origins is lower than for mothers with French-speaking origins, where working mothers are seen more favourably than in the German-speaking culture. Finally, Cavapozzi et al. (2021) show that peers' gender norms influence a mothers labour supply in the UK.

I contribute to the existing literature by extending the correlational analysis of Rabaté and Rellstab (2022) to a more causal framework taking advantage of the Dutch bible belt. Furthermore, I add evidence to the findings of Boelmann et al. (2021), Steinhauer (2018) and Cavapozzi et al. (2021) using a different setting with local variation in gender norms stemming from the religious landscape of the Netherlands. I compare spatially concentrated communities that share pre-child characteristics and the same set of family policies but exhibit large differences in elicited gender norms with respect to parenthood. These internal variations in gender norms originating from different religious or other communities can be found in many countries. Therefore, my results can potentially be extrapolated to many contexts. Furthermore, combining administrative data from the universe of Dutch first-time parents with municipal information on voting and infrastructure, I am able to follow recent cohorts of parents for a long period of time and observe their earnings, employment, FTE and wage rate. Hence, the findings inform not only about the impact of norms on the child penalty for the probability of

having paid work, but also for other margins of labour supply in the context of a country where part-time work is widely available.

I find that the child penalty of women is around 10% points (or about 30%) larger in the bible belt in the Netherlands compared to other Dutch municipalities. This difference is mainly driven by an equally large child penalty difference in hours worked (10–12% points). I then provide additional evidence to support the gender norm interpretation of my findings and to rule out alternative explanations. First, I use a re-weighting strategy to ensure comparability between parents-to-be inside and outside the bible belt in education and demographics. Second, I argue why other competing explanations for the difference in child penalty between mothers inside and outside the bible belt such as differential childcare availability do not drive my results.

II. Background

The Dutch bible belt & gender norms

The bible belt in the Netherlands is a geographical area spanning from the South West of the country up to the North East, and it is characterized by a high concentration of Dutch orthodox protestants. The origins of orthodox protestants can be traced back to the reformation, but only in the 1960s to 1980s, the perception of a 'bible belt' emerged as the rest of the country became more secular. Communities in the bible belt belong to different denominations where the local congregation plays an important role, but they share a set of convictions (Ruijs 2013). Governmental institutions are shared inside and outside of the bible belt, as they are mostly determined at the national level.

The Dutch bible belt can be located and proxied by the voter share of the *Staatskundig Gereformeerde Partij* (SGP) in parliamentary elections, a conservative Christian party that represents the orthodox protestant community politically (see Figure A1 for a geographical representation of voter shares for the SGP).

The Dutch orthodox protestant community is a religious and cultural minority in the Netherlands, for whom religion plays an important role in daily life. One pillar of orthodox protestant belief is

predestination. Part of the community therefore rejects (preventive) vaccination, which has led to epidemics in the past (Ruijs 2013). Prevailing gender norms are more non-egalitarian in the bible belt than elsewhere in the Netherlands (Gielen and Zwiers 2018). The LISS panel survey³ provides direct evidence that *SGP* voters have less egalitarian elicited gender norms than voters for other Christian parties and the rest of the Netherlands. Appendix Table A1 shows that there is a considerably higher share of *SGP* voters agreeing to statements that advocate for a clear division of labour where the father is the breadwinner and the mother the unpaid care giver compared to voters of other Christian parties and the rest.

Family policies in the Netherlands

Family policies in the Netherlands are set at the national level and are accessible to all citizens independently of province or municipality, thus inside as well as outside of the bible belt. The main goal of Dutch family policies is to facilitate and incentivize the combination of having children and paid work. These include universal childcare, parental leave, a short paternity leave of two days, maternity leave, and tax incentives for second earners. Appendix A1.1.3 describes these policies in detail.

III. Data

I use administrative micro data from Statistics Netherlands (CBS) containing information about parent–children linkages, basic demographic information, employment (1999–2018), earnings (1999–2018), the death registry, address information, and partners; and transform it into a yearly panel. The address information is linked to parliamentary election outcomes by municipality (Kiesraad 2019) which serve as a basis for the identification of bible belt municipalities (see A2 for details on the data).

Analysis sample and main outcomes

The analysis includes different-sex legal parents of all first-born children in the Netherlands between 2008 and 2013 who are between 20 and 45 years old at birth, and who live together at birth. I then follow the labour market path for this cohort of parents five years before and after the birth of their first child. The panel is balanced in event-time.

I analyse four labour market outcomes: unconditional earnings, the probability of paid work, conditional full-time equivalent (FTE), and the annual wage rate. The probability of paid work includes employment and self-employment. Earnings are defined as income from work, both from employment or self-employment. The conditional full-time equivalent (FTE) is a measure of part-time work, and it is defined as working hours normalized to the reference full-time hours. The full-time reference hours are defined by sector or firm by Statistics Netherlands and are only available from 2001 to 2016 for the employed.⁴ The annual wage rate per FTE is constructed by dividing yearly earnings from employment by the FTE, and thus available for employees of the parent cohorts 2008–2012.

Identification of bible belt residents

Religion is not registered in the Dutch administrative data at the individual level. Instead, I identify municipalities with a large voter share for the orthodox protestant party *SGP* from election data (Kiesraad 2019). I classify municipalities with an *SGP* voter share above 4% as belonging to the bible belt.⁵ All municipalities with an *SGP* vote share below 0.5% in all election years are classified as outside of the bible belt. Municipalities with a vote share between 0.5 and 4%, or a vote share that crosses the threshold on either side in some elections but not in others are excluded from the analysis. The spatial distribution of this bible belt classification is depicted in Figure 1.

³The LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata is a representative sample of Dutch individuals who participate in monthly Internet survey on household background information and a yearly longitudinal survey on different topics.

⁴Therefore, this outcome is limited to employed parents of the cohorts 2008–2012.

⁵This corresponds to the 90th percentile of parents in my sample. This threshold is a compromise to the trade-off between i) having enough parents in the bible belt to estimate the child penalty and high external validity; and ii) choosing 'extreme enough' municipalities to have large differences in gender norms to detect a meaningful effect. Mechanically, the results are sensitive to threshold: the higher the threshold, the higher the share of parents with non-egalitarian gender norms, and hence the higher the difference of the child penalty between the regions inside and outside the bible belt.

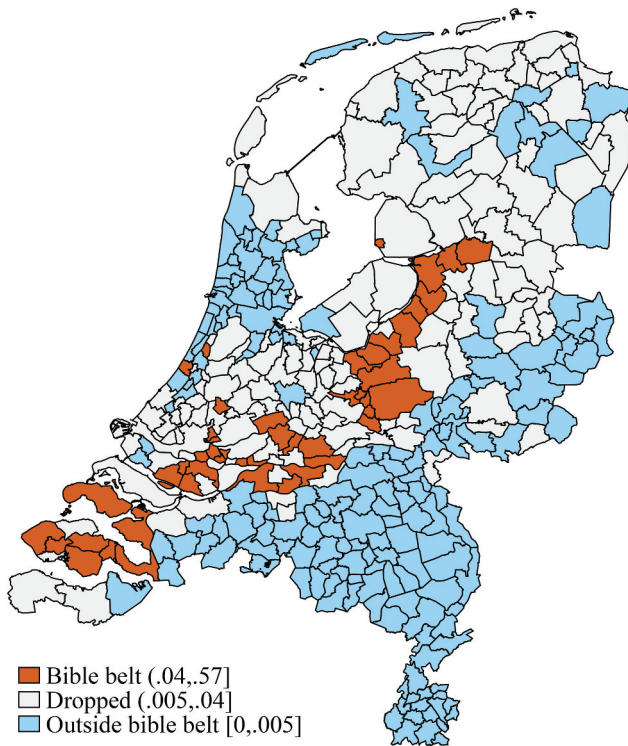


Figure 1. Classification of municipalities. Classification of municipalities for the analysis as bible belt (*SGP* voter share $\geq 4\%$), outside bible belt (*SGP* voter share $< 0.5\%$), and dropped (*SGP* voter share between 0.5% and 4% , or classification according to the *SGP* voter share is not stable in the observation period).

I determine parental bible belt status with the municipality of residence one year before the first child. About 14% of or 80,000 (86% of or 500,000) parents are inside (outside) the bible belt in the sample.

The *SGP* voter share is only a proxy to identify orthodox protestants. There may be orthodox protestants that do not vote for the *SGP*, and many parents are classified as living in the bible belt while they are not orthodox protestant and share little of these values. This is also reflected in the relatively low average *SGP* voter share in the bible belt. Yet, the *SGP* voter share is a useful proxy as it identifies areas where gender norms are less egalitarian than elsewhere.

IV. Empirical strategy

I estimate the child penalty using the method by Kleven et al. (2019b). First, this method requires

estimating Equation 1, where y_{it} is the labour market outcome in levels, measured in calendar year t for parent i (with N the total number of parents). To obtain a separate child penalty estimate for fathers and mothers inside and outside the bible belt, I estimate Equation 1 by gender and by bible belt status separately. The main explanatory variables of interest are a set of event time indicators ranging from 5 years before birth of the first child up to five years after with baseline at $eventtime = -1$, one calendar year before child birth. The remaining control variables are age and year fixed effects. I choose this set of controls to obtain results comparable to Kleven et al. (2019b) and other studies employing this method. Standard errors are clustered at the individual level to take into account the panel structure of the data.⁶

$$\begin{aligned}
 y_{it} = & \sum_{q=-5}^5 \alpha_q 1[event\ time_{it} = q] \\
 & + \sum_{k=15}^{50} \beta_k 1[age_{it} = k] \\
 & + \sum_{t=2003}^{2018} \gamma_t 1[time_{it} = t] + v_{it}
 \end{aligned} \tag{1}$$

$$\tilde{y}^q = \frac{1}{N} \sum_i \left[\sum_{k=15}^{50} \hat{\beta}_k 1[age_{it} = k] + \sum_{t=2003}^{2018} \hat{\gamma}_t 1[time_{it} = t] \right] \tag{2}$$

$$CP_q = \frac{\hat{\alpha}_q}{\tilde{y}^q} \tag{3}$$

Second, I estimate the child penalty CP_q (Equation 3) by dividing the relative change in the outcome due to the child ($\hat{\alpha}_q$) by the pre-birth trend for fathers and mother inside and outside the bible belt separately.⁷ The pre-birth trend is defined in Equation 2 as the average predicted labour market outcome at event time q in absence of a child. Hence, the counterfactual is calculated based on parents-to-be, and not on a control group. The child penalty represents the percentage change in the

⁶The problems discovered with event study models by the recent advances in the econometrics of event studies (see e.g. Sun and Abraham 2021) are not present in the model below, because it is not a two-way fixed effects model.

⁷In the literature, the child penalty is sometimes defined as mothers' relative to fathers' earnings instead of pre-birth trends. Since the child penalty for fathers is close to zero, I compare women to their pre-birth trend in this study.

outcome due to the child and is a cumulative effect of all children a parent has up to that point in time. To determine whether the child penalties inside and outside the bible belt are statistically different from each other, I bootstrap the difference in child penalty with 100 replications.

To estimate the effect of having a child on labour market outcomes, Kleven et al. (2019b) relies on two identifying assumptions: 1) a sharp discontinuity at the arrival of the child; and 2) no sharp discontinuities in other determinants of income at child birth that are unrelated to the child.⁸ However, the main interest here is to compare child penalties between parents inside and outside the bible belt. Similar to a difference-in-difference framework, this implies that trends in outcomes should be parallel absent treatment. This can be tested only in the periods before child birth, where labour market outcomes

should evolve similarly for both parents in the bible belt and outside. This assumption holds descriptively (Figure 2).

V. Results

Descriptive evidence

Figure 2 shows average trends in earnings, the probability of paid work, the FTE and the annual wage rate for parents inside and outside the bible belt from five years before their first child up to five years after. By gender, parents of the two groups are on very similar labour market trajectories before birth. After birth of the first child, all women experience a dip in earnings, but the decrease is larger for women in the bible belt. By five years after the birth of the first child, this difference in annual earnings is about 8000€ per year. This larger decrease in earnings for mothers

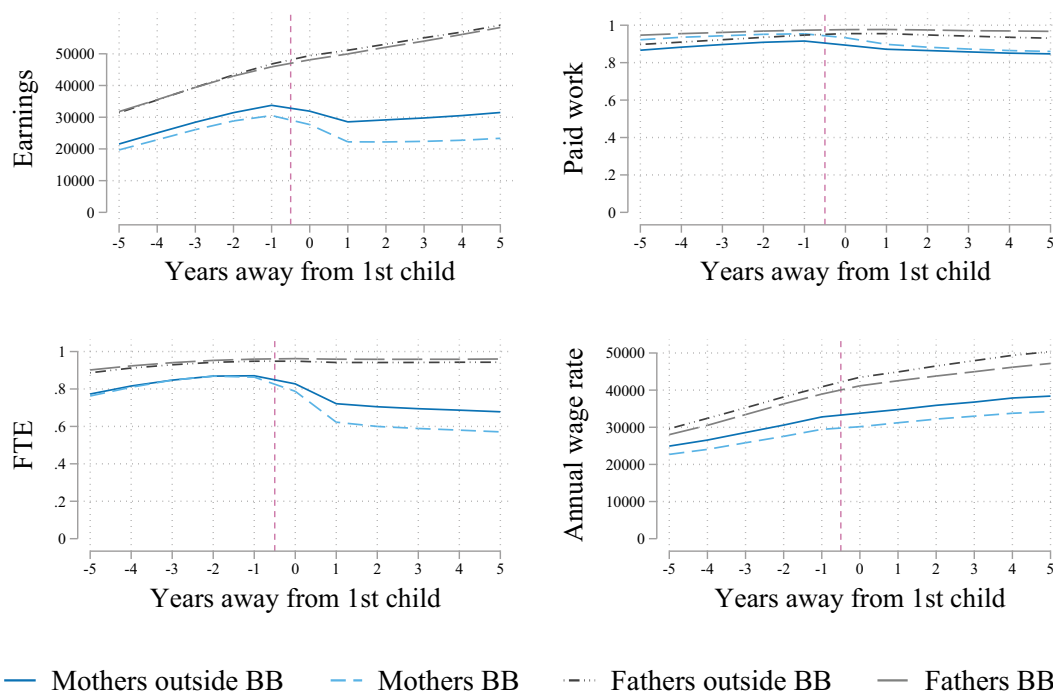


Figure 2. Average labour market outcomes inside and outside the bible belt before and after having children. Yearly average earnings and participation for first-time parent cohorts 2008–2013, and yearly average FTE and the wage rate for parent cohorts 2008–2012.

⁸Kleven et al. (2019b) provide evidence that the event study framework controlling for age and time fixed effects gives similar results as methods where the arrival of a child is instrumented (using the sex-composition of the first two children as an instrument for the third), or with a control group consisting of childless individuals.

in the bible belt is driven by a larger decrease in working hours, as there is no difference between mothers inside and outside the bible belt for the probability of having any paid work; and because the difference in wage rate between the two groups does not change importantly before and after birth. For men inside and outside the bible belt, there is no impact of having children on earnings, and both are on similar earnings trajectories in the observation period. A balancing table comparing parents in the bible belt with parents outside (Table A3) shows that there are imbalances in levels, but these differences should not matter for my identification strategy as long as trends in outcomes are similar.

Child penalties inside and outside the bible belt

Figure 3 shows the child penalties for men and women inside and outside the bible belt (CP_q from Equation 3). This measure indicates how labour market outcomes evolve differently from what the counterfactual (i.e. not having a baby) would have projected. Event times -5 to -1 depict the pre-

trends, which should be similar between women (or men) inside and outside of the bible belt. The labour market effects of having a child by group can be inferred from event time zero onwards.

The top left panel of Figure 3 indicates that the growth of earnings of mothers-to-be is relatively stable before the child is born, with a slight decrease in the year of the pregnancy (event time -1). There is no difference in earnings development between mothers-to-be in the bible belt and outside. After the child is born, there is a decrease in mother's earnings that grows over time for both groups. The drop is larger for women in the bible belt from year 1 onwards. When the first child is five, the difference between both groups of women is 10% points. In relative terms, this means that the child penalty of women from the bible belt is about 30% larger than outside. Men's earnings in the bible belt grow more after having a child than before, and five years after the first child they have a 'child premium' of about 5%. However, this seems to be a continuation of a pre-birth trend,

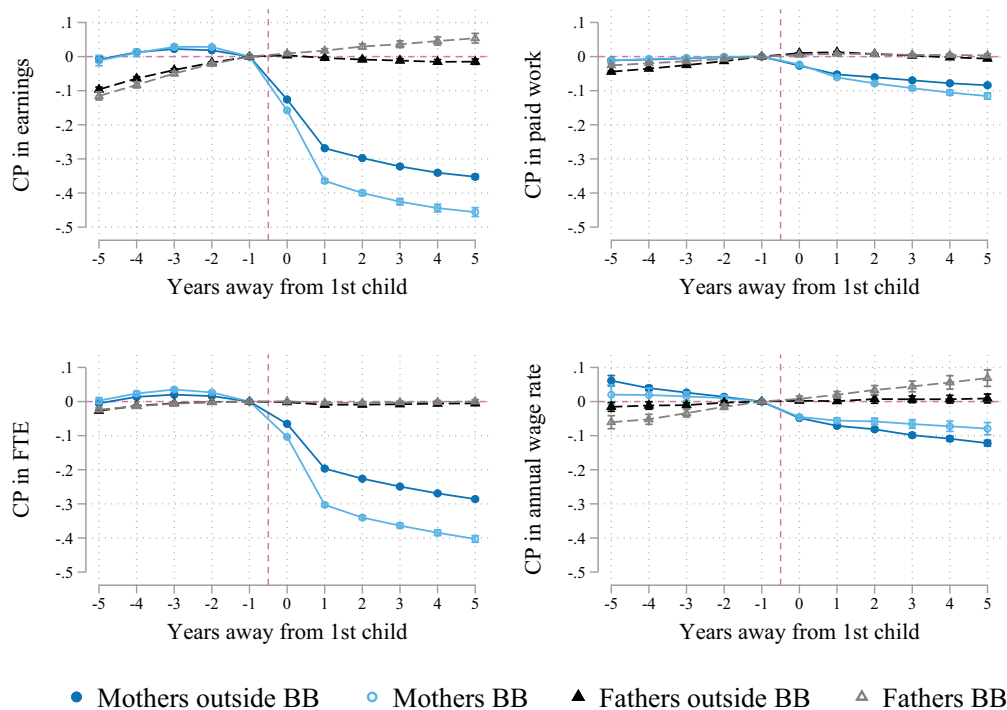


Figure 3. Child penalties (CP) inside and outside the bible belt. Child penalties (CP) with 95% confidence intervals by event time q for parents inside and outside the bible belt (BB), where CP_q measures the % change in the outcome due to the child compared to the pre-birth trend (see Equation 3). Four labour market outcomes are depicted: unconditional earnings, the probability of having any paid work, full-time equivalent (FTE), and annual wage rate. Solid lines refer to women, men are represented with dashed lines.

and hence difficult to attribute to the birth of the child only. Outside the bible belt, there is a small earnings drop for fathers of around 1%.

The reduction in unconditional earnings of mothers is partly due to exits from the labour market (top right panel in Figure 3), but this is a relatively small contributor to the overall child penalty. For the probability of having any paid work, pre-trends are again similar for women inside and outside the bible belt. Once the child is born, participation drops for women in the bible belt by 12%, and outside by 8%, implying that the drop is 4% points larger in the bible belt. For men in both areas, the probability of having paid work is not affected by child birth.

For the FTE, pre-birth trends for women inside and outside the bible belt are again very similar and slightly decreasing in the year of the pregnancy. After giving birth, the FTE decreases importantly for both groups of women inside and outside the bible belt (40% and 29%, respectively, five years after), so the drop in the bible belt is about 11% points (or 37%) larger five years after the child is born. The difference in child penalty in FTE corresponds to approximately 0.1 FTE, or half a day of paid work per week. For men, there is no significant change in FTE before and after birth.

Finally, the growth of the wage rate slows down for both women in and outside the bible belt in the five years before the child is born, but slightly more so for women outside the bible belt. Five years after the child is born, the wage rate drops for women in the bible belt by 8%, and outside the bible belt by 12% five years after the first child. This is the only outcome where women in the bible belt have a slightly lower child penalty than women outside the bible belt. However, the difference between the groups is relatively small and compared to the overall child penalty in earnings, the child penalty in wage rate is not a very important contributor to the overall child penalty. For men outside the bible belt, the annual wage rate is not affected by having children in the observation window. For men in the bible belt, the annual wage rate grows slightly over time, resulting in a 7% child premium. Again, this cannot be separated from a pre-birth trend, and it is therefore not attributable to the child.

Figure 3 does not convey whether the differences in child penalty between parents in the bible belt and outside are statistically significant. Table 1 assesses whether these differences in child penalty are statistically significant using a bootstrap.⁹ For example, the table shows that five years after birth of the first child, the child penalty in earnings of women in the bible belt is 10.3% points larger than the child penalty in earnings of women outside the bible belt. This difference is statistically significant. In general, the differences in child penalty between the bible belt and outside five years after the birth of the first child are all statistically significant, even pre-children. As the sample is large, very small differences may become statistically significant. However, these pre-birth effects are very small and mostly below one percentage point, and hence not economically relevant. For men, the bootstrap results show small and persistent child premiums in the bible belt (even if subtracting the small pre-trends in differences). This finding may be consistent with a stronger breadwinner norm for fathers in the bible belt.

Since the bible belt indicator in the model is at the municipal level, the results have to be interpreted as an average effect in the bible belt. If the conjecture is true that SGP voter families have less egalitarian gender norms and hence a higher child penalty, the results imply that SGP voter families experience an even higher child penalty than reported in Figure 3 for the bible belt, and that non-SGP voting families have a lower child penalty than reported for the bible belt in Figure 3. While I cannot implement this analysis by voting behaviour on individual level due to unavailability of data, the average difference between the bible belt and outside is informative as well for the role of norms on average.

Increasing comparability of parents inside and outside the bible belt with matching

Parents-to-be in the bible belt are meaningfully different from parents in the bible belt with respect

⁹The bootstrapped estimates may slightly differ from the estimates reported in Figure 3, as the bootstrap estimate is only asymptotically equivalent to the estimate from Equation 3.

Table 1. Differences in child penalties (in percentage points) between parents inside and outside the bible belt by event time q (5 years before to 5 years after the first child).

q	Women				Men			
	Earnings	Paid work	FTE	Wage rate	Earnings	Paid work	FTE	Wage rate
-5	-0.002*** (0.001)	.000*** (.000)	0.008*** (0.000)	-0.042*** (0.002)	-0.019*** (0.001)	.019** (.000)	0.003*** (0.000)	-0.045*** (0.001)
-4	-0.002*** (0.001)	.002*** (.000)	0.010*** (0.000)	-0.021*** (0.001)	-0.019*** (0.000)	.015** (.000)	-0.001*** (0.000)	-0.039*** (0.001)
-3	0.006*** (0.000)	.001*** (.000)	0.015** (0.000)	-0.012*** (0.001)	-0.010*** (0.000)	.011** (.000)	-0.001*** (0.000)	-0.024*** (0.001)
-2	0.010*** (0.000)	.001*** (.000)	0.011** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)	.007*** (.000)	-0.001*** (0.000)	-0.013*** (0.000)
-1		<i>baseline</i>				<i>baseline</i>		
0	-0.032*** (0.000)	.003*** (.000)	-0.038*** (0.000)	0.004*** (0.000)	0.006*** (0.000)	-.005*** (.000)	0.002*** (0.000)	0.005*** (0.000)
1	-0.096*** (0.000)	-.009*** (.000)	-0.106*** (0.000)	0.015** (0.000)	0.022** (0.000)	-.005*** (.000)	0.006*** (0.000)	0.018** (0.001)
2	-0.102*** (0.000)	-.019*** (.000)	-0.114*** (0.000)	0.023** (0.001)	0.039** (0.000)	.000*** (.000)	0.006*** (0.000)	0.026** (0.001)
3	-0.103*** (0.000)	-.024*** (.000)	-0.114*** (0.000)	0.033** (0.001)	0.049** (0.001)	.002*** (.000)	0.006*** (0.000)	0.037** (0.001)
4	-0.103*** (0.000)	-.028*** (.000)	-0.115*** (0.000)	0.037** (0.001)	0.062* (0.001)	.007*** (.000)	0.005*** (0.000)	0.049** (0.001)
5	-0.103*** (0.000)	-.032*** (.000)	-0.117*** (0.000)	0.043** (0.001)	0.070* (0.001)	.010*** (.000)	0.006*** (0.000)	0.060* (0.002)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ based on 100 bootstrap replications. Depicts the child penalty in the bible minus the child penalty in outside the bible belt by event time q (or time away from the first child). A negative number means a larger child penalty in the bible belt than outside, and the unit of the numbers is percentage points.

to education, age, and some labour market outcomes one year before the child (Table A3). While the identification strategy relies on parallel trends in the outcomes only and does not require balance in observables, these imbalances may hide an omitted factor relevant for the trends in unobserved counterfactual outcomes. To exclude this threat to identification, I use a coarsened exact matching (CEM) procedure (King and Nielsen 2019) as a robustness test. The CEM reweighting procedure achieves similar goals as propensity score matching by approaching the distribution of observables of parents outside the bible belt to the parents inside the bible belt pre-children, but it can be preferable with a large data set. I reweight the sample such that on couple level, parents outside the bible belt are similar to parents in the bible belt on the following dimensions: five education categories for fathers and mothers at $q - 1$, five evenly spaced age groups at $q - 1$ for fathers and mothers, and marital status at $q - 1$.

The main trade-off with CEM is that with more matching categories the internal validity increases, but the less likely it is to find matches for all treated, leading to a decrease in external validity. I choose relatively few matching criteria, implying that I find

matches for all parents in the bible belt except for 13 couples. Hence, the sacrifice in external validity to increase internal validity is relatively small.

The CEM weighting eliminates almost all meaningful differences between the two groups of parents before having children (Table A4). Figure 4 shows the child penalty results when using the CEM weights. The comparison to the main results is displayed in Figures A2 and A3 for mothers and fathers separately, as the estimates are so similar that the comparison is not visible in a combined figure. This implies that child penalties outside the bible belt are the same when these parents have the same characteristics in terms of education, age, and marital status one year before the first child as parents in the bible belt. Hence, controlling for the imbalances in observables does not alter the difference in child penalties inside and outside the bible belt as reported in Figure 3.

Competing explanations for the difference in child penalty between the bible belt and elsewhere

The difference in child penalty between the mothers inside and outside the bible belt can only

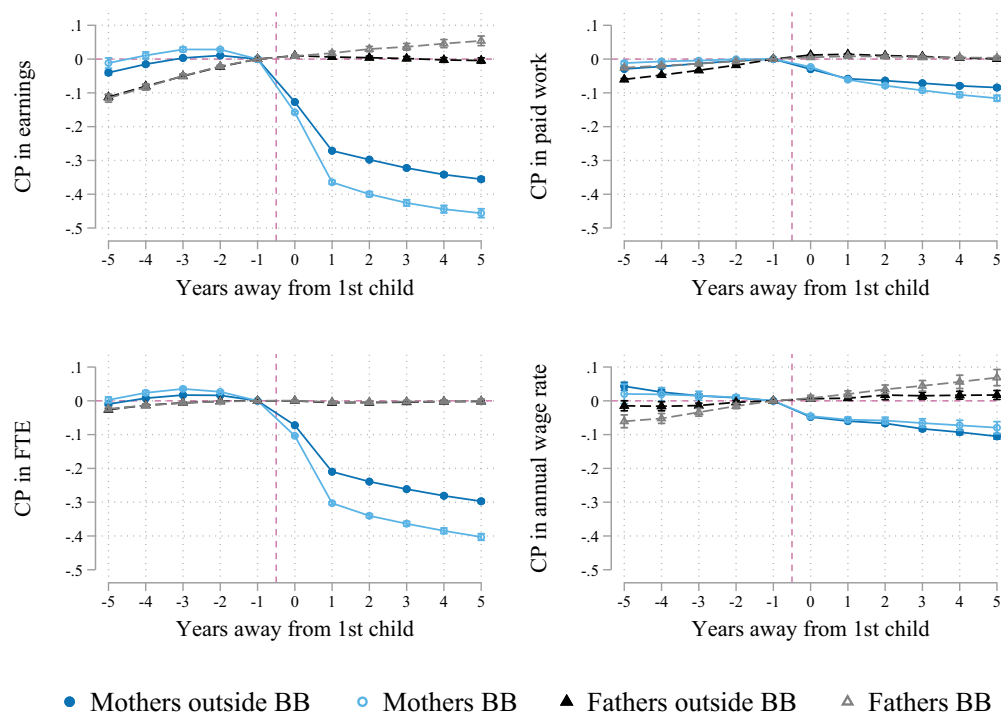


Figure 4. Child penalties (CP) inside and outside the bible belt: CEM weighted. Weighted child penalties (CP) with 95% confidence intervals by event time q for parents inside and outside the bible belt (BB), where CP_q measures the % change in the outcome due to the child compared to the pre-birth trend (see Equation 3). Four labour market outcomes are depicted: unconditional earnings, the probability of having any paid work, full-time equivalent (FTE), and annual wage rate. Solid lines refer to women, men are represented with dashed lines.

be attributed to gender norms if there is no other difference in institutions or residents' characteristics that may explain my findings.

Municipal fixed effects

First, I test whether unknown municipal-level confounders drive the differences in child penalty between the bible belt and elsewhere by including municipal fixed effects in Equation 1. The results are robust to the inclusion of municipality fixed effects, indicating that unobserved municipality characteristics are not driving my results.

Total fertility

Parents in the bible belt have 0.12 children more on average than parents outside the bible belt five years after birth (see A3), and having more children leads to a higher child penalty (Sieppi and Pehkonen 2019). However, the difference in the number of children between women inside and outside the bible belt only arises from three years

after the birth of the first child, whereas the child penalty difference of 10% points already manifests one year after birth of the first child and is stable thereafter (Figure A5). Therefore, it is unlikely that the difference in child penalty is driven by a higher fertility in the bible belt. One may argue that the higher total fertility in the bible belt may alter expectations of first time parents about the number of children they are going to have. This may in turn affect the division of labour in the household at the birth of the first child and the child penalty. However, it is difficult to separate expectations from gender norms, as gender norms are linked to expectations about women's behaviour as mothers.

Divorce rates

Couples in the bible belt are more likely to remain married. This could contribute to the higher child penalty in the bible belt, because a high likelihood of not divorcing may make it

economically less necessary that both parents work. However, when I restrict my sample to parents who stay together in the whole observation period, I find almost identical results for the child penalties (Figure A6). The higher probability of remaining married in the bible belt hence does not drive the difference in child penalty. Similar to total expected fertility, the expectation about the likelihood of remaining married cannot be separated from gender norms.

Access to formal childcare

The national universal childcare policy foresees that childcare is accessible everywhere in the Netherlands, and hence access should be similar both in the bible belt and outside. Data on use shows that childcare is more frequently used outside the bible belt (5.2 children with childcare per new-born baby in 2015) than in the bible belt (3.6 children with childcare per new-born baby) (CBS 2016). However, use is determined both by supply and demand, and the higher use outside the bible belt may be reflected by a higher demand instead of higher accessibility. Two types of evidence point in this direction. First, the average distance to the next childcare facility in the bible belt is 1.1 km, which is reachable by bike in less than 10 minutes. Whereas this distance is 0.2 km smaller outside the bible belt, accessibility in the bible belt is not compromised. Second, while there is unfortunately no data on waiting lists by municipality, waiting lists are mainly long regions and municipalities outside the bible belt (Aan de Wiel 2021; B&A groep 2012). Hence, while childcare use outside the bible belt is more common, it is easily available in the bible belt.

Migrants' gender norms

In the Netherlands, migrants from a Turkish/Moroccan background hold less egalitarian gender norms on average than Dutch natives (Khoudja 2018), and one may hypothesize that this influences the difference in child penalty between the bible belt and elsewhere. However, this is not consistent with the data, as there are as many migrants with a Moroccan and fewer migrants with a Turkish background in the bible belt than elsewhere.

Progressive cities

One may be concerned that the lower child penalty outside the bible belt is mainly driven by a lower child penalty in the big cities that tend to be more progressive than rural areas. To alleviate this concern, I estimate the child penalty when excluding all parents living in cities with more than 100,000 inhabitants and find a similar difference in child penalty between the bible belt and elsewhere (Figure A7).

Sorting on norms

There may be sorting on norms, such that people agreeing with the norms in the bible belt are more likely to stay in the bible belt, whereas people who disagree move somewhere with norms that are more in accordance with their own values. However, even if sorting on norms explains the difference in child penalties, this is no threat to identification of this study, as the aim is to compare parents with more egalitarian gender norms to parents with less egalitarian gender norms; and not parents who have their origins in certain geographical areas. If anything, sorting on norms makes this comparison cleaner.

Sectors

Parents in the bible belt may work in different sectors with more or less generous employers regarding parental leave. This is not supported by the data, as mothers-to-be work approximately in the same sectors or under the same collective labour agreements inside and outside the bible belt (Table A5).

Employer behaviour

Employers may treat parents inside or outside the bible belt differently. While I cannot exclude this, the difference in employer behaviour would have to be unrelated to gender norms to represent a competing explanation for my findings, which is unlikely.

Together, this suggest that the difference in child penalty between parents in the bible belt and outside is unlikely to be solely explained by the factors discussed in this section. Since the difference in gender norms is the most striking one between the bible belt and areas outside, it is the most likely explanation for the differences in child penalty. Additional robustness checks on data and modelling decisions can be found in Appendix A1.6.

VI. Discussion

A recent literature shows that women face substantial child penalties in labour market outcomes after the first child, while men do not. The literature to date has not found evidence that factors such as biology and the absence of appropriate family policies are the main drivers of the child penalty for mothers. In this article, I explore the role of gender norms as a driver of the child penalty for mothers. In the Netherlands, gender norms are less egalitarian in bible belt municipalities, where the concentration of orthodox protestants is high, while national institutions such as childcare availability are similar in every municipality. I find that becoming a mother in the bible belt results in a 10% points or about 30% larger child penalty than elsewhere in the Netherlands. Compared to the effects of family policies on the child penalty found by other studies, a 30% difference is very large. However, the effect may still be plausible, as its main driver is a change of about 0.1 FTE – or half a day of work per week.

Although I cannot explicitly link gender norms and the child penalty, I argue that norms are the most likely explanation for the higher child penalty in the bible belt, because I can exclude that differences in education and demographic characteristics do not influence the difference in child penalty; and because other examined potential explanation such as a differential access to childcare is not supported by the data.

My findings suggest that the correlation between religiosity and the child penalty found in Rabaté and Rellstab (2022) also persists in a more causal framework of analysis. Moreover, they align well with the findings in the literature studying the effect of gender norms on mothers' labour supply in other contexts. Steinhauer (2018) finds a 15–25% lower probability of working for mothers in the German-speaking part of the Swiss language border, where prevailing norms are less in favour of working mothers compared to the French speaking counterparts who share the same institutional setting. Boelmann et al. (2021) report that East German mothers in East–West crossborder commuting zones are about 6% points more likely to be employed than their West German peers seven years after the first child. Moreover, the child penalty literature on

same-sex mothers, for whom gender norms are less informative for the division of labour in the household, shows considerably lower child penalties than for different-sex parents (Adema, Rabaté, and Rellstab 2020; Andresen and Nix 2022; Moberg 2016; Rosenbaum 2021).

The finding that gender norms influence the child penalty implies that if policy makers want to reduce the child penalty, changing gender norms may be necessary to achieve this. It is difficult to implement policies that target norms of adults, because norms are transmitted by the example that parents and others set when we are growing up (Farré and Vella 2013; Fernández, Fogli, and Olivetti 2004; Fogli and Veldkamp 2011; McGinn, Ruiz Castro, and Lingo 2019; Olivetti, Patacchini, and Zenou 2020). While schools and childcare could be used to reduce gender stereotypes learned in childhood, such a policy may not be feasible in the Netherlands. Since the beginning of the twentieth century, Dutch schools have much freedom in their curriculum and teaching method. This is the result of a historical political compromise, and it is not conceivable to change this rule. However, trying to reduce gender stereotyping in childcare and schools may be an option for other countries. Finally, everyone who has children in their social environment has the option to encourage these children to engage in gender role non-conforming activities if they want to.

Acknowledgments

I am grateful to Statistics Netherlands (CBS) for providing access to non-public microdata to conduct this study and to CentERdata for giving me access to the LISS panel data. I would like to thank García-Gómez, Pieter Bakx, Eddy van Doorslaer, Fabrizio Mazzonna, Anne Boring, Aline Bütikofer, Simon Rabaté, Andreas Steinhauer, Martin Halla, Patricia Funk, Esmée Zwiers, Paola Profeta, Johanna Rickne, Niccolò Gatti, five anonymous referees, the health economics group of ESE and ESPHM at Erasmus University Rotterdam, and the economics department of the University of Lugano for helpful comments. The author has no conflict of interest and no funding to declare.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCIDSara Rellstab  <http://orcid.org/0000-0002-7920-7391>**References**

- Aan de Wiel, E. 2021. "Wachtlijsten bij de kinderopvang: hoe vind je toch een plekje?" <https://www.nu.nl/kind-gezin/6122657/wachtlijsten-bij-de-kinderopvang-hoe-vind-je-toch-een-plekje.html>.
- Adema, Y., S. Rabaté, and S. Rellstab. 2020. "Inkomen moeders halveert bijna na komst kinderen." *Econome nvakblad ESB*.
- Akerlof, G. A., and R. E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* CXV (3): 715–753. <https://doi.org/10.1162/003355300554881>.
- Andresen, M. E., and E. Nix. 2022. "Can the Child Penalty Be Reduced? Evaluating Multiple Policy Interventions." *Discussion Papers* 983.
- Andresen, M. E., and E. Nix. 2022. "What causes the child penalty? Evidence from adopting and same-sex couples." *Journal of Labor Economics* 40 (4): 971–1004. <https://doi.org/10.1086/718565>.
- Angelov, N., P. Johansson, and E. Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3): 545–579. <https://doi.org/10.1086/684851>.
- B&A groep. 2012. "Wachtlijsten en wachttijden kinderdagverblijven en buitenschoolse opvang - 8 e meting - rapportage." *Technical Report Projectnummer: 35551*.
- Bertrand, M. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, 1543–1590. Vol. 4. Elsevier B.V. [https://doi.org/10.1016/S0169-7218\(11\)02415-4](https://doi.org/10.1016/S0169-7218(11)02415-4).
- Bettendorf, L. J. H., E. L. W. Jongen, and P. Muller. 2015. "Childcare Subsidies and Labour Supply — Evidence from a Large Dutch Reform." *Labour Economics* 36:112–123. <https://doi.org/10.1016/j.labeco.2015.03.007>.
- Boelmann, B., A. Raute, and U. Schönberg. 2021. "Wind of Change? Cultural Determinants of Maternal Labor Supply." *CESifo Working Paper*.
- Bursztyn, L., T. Fujiwara, and A. Pallais. 2017. "Acting wife? Marriage Market Incentives and Labor Market Investments." *American Economic Review* 107 (11): 3288–3319. <https://doi.org/10.1257/aer.20170029>.
- Bütikofer, A., S. Jensen, and K. G. Salvanes. 2018. "The Role of Parenthood on the Gender Gap Among Top Earners." *European Economic Review* 109 (262675): 103–123. <https://doi.org/10.1016/j.eurocorev.2018.05.008>.
- Cardoso, A. R., and L.-P. Morin. 2018. "Can Economic Pressure Overcome Social Norms? The Case of Female Labor Force Participation." *Barcelona GSE Working Paper, no. 1051*.
- Cavapozzi, D., M. Francesconi, and C. Nicoletti. 2021. "The Impact of Gender Role Norms on mothers' Labor Supply." *Journal of Economic Behavior and Organization* 186:113–134. <https://doi.org/10.1016/j.jebo.2021.03.033>.
- CBS. 2016. "Kinderopvang per gemeente in Nederland, 2015." <https://www.cbs.nl/nl-nl/maatwerk/2016/41/kinderopvang-per-gemeente-in-nederland-2015>.
- CBS Statline. 2021. "Bevolking; hoogstbehaald onderwijsniveau en onderwijsrichting." <https://opendata.cbs.nl/#/CBS/nl/dataset/82816NED/table?ts=1631169850245>.
- Cortés, P., and J. Pan. 2020. *Children and the Remaining Gender Gaps in the Labor Market*. National Bureau of Economic Research. No. w27980.
- Dehos, F., and M. Paul. 2021. "The Effects of After-School Programs on Maternal Employment." *Journal of Human Resources* 58 (5): 1644–1678. <https://doi.org/10.3368/jhr.58.5.0120-10651R1>.
- Farré, L., and F. Vella. 2013. "The Intergenerational Transmission of Gender Role Attitudes and Its Implications for Female Labour Force Participation." *Economica* 80 (318): 219–247. <https://doi.org/10.1111/ecca.12008>.
- Fernández, R. 2007. "Women, Work, and Culture." *Journal of the European Economic Association* 5 (2–3): 305–332. <https://doi.org/10.1162/jeea.2007.5.2-3.305>.
- Fernandez, R., and A. Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1 (1): 146–177. <https://doi.org/10.1257/mac.1.1.146>.
- Fernández, R., A. Fogli, and C. Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *The Quarterly Journal of Economics* 119 (4): 1249–1299. <https://doi.org/10.1162/0033553042476224>.
- Fogli, A., and L. Veldkamp. 2011. "Nature or Nurture? Learning and the Geography of Female Labor Force Participation." *Econometrica* 79 (4): 1103–1138.
- Fortin, N. M. 2005. "Gender Role Attitudes and the Labour-Market Outcomes of Women Across OECD Countries." *Oxford Review of Economic Policy* 21 (3): 416–438. <https://doi.org/10.1093/oxrep/gri024>.
- Fortin, N. M. 2015. "Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery." *Annals of Economics and Statistics* 117/118:379–401. <https://doi.org/10.15609/annaeconstat2009.117-118.379>.
- Gielen, A., and E. Zwieters. 2018. "Biology and the Gender Gap in Educational Performance: The Role of Prenatal Testosterone in Test Scores." *IZA Discussion Paper* 11936. <https://doi.org/10.2139/ssrn.3285836>.
- Hicks, D. L., E. Santacreu-Vasut, and A. Shoham. 2015. "Does Mother Tongue Make for Women's Work? Linguistics, Household Labor, and Gender Identity." *Journal of Economic Behavior and Organization* 110:19–44. <https://doi.org/10.1016/j.jebo.2014.11.010>.
- Khoudja, Y. 2018. "Women's labor market participation across ethnic groups." PhD thesis, Utrecht University.
- Kiesraad. 2019. "Databank Verkiezingsuitslagen." <https://www.verkiezingsuitslagen.nl/>.

- King, G., and R. Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27 (4): 435–454. <https://doi.org/10.1017/pan.2019.11>.
- Kleven, H., C. Landais, J. Posch, and A. Steinhauer. 2020. "Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation*."
- Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. 2019. "Child Penalties Across Countries: Evidence and Explanations." *AEA Papers and Proceedings*, 2014 Broadway, Suite 305, Nashville, TN 37203, 122–126. Vol. 109. American Economic Association.
- Kleven, H., C. Landais, and J. E. Sogaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209. 10. <https://doi.org/10.1257/app.20180010>.
- Kleven, H., C. Landais, and J. E. Sogaard. 2020. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3593260>.
- Krapf, M., A. Roth, and M. Slotwinski. 2020. "The Effect of Childcare on Pa-Rental Earnings Trajectories." *CESifo Working Paper*.
- Mari, G., G. Cutuli, and D. U. Pre-Print. 2018. "Do Parental Leaves Make the Motherhood Wage Penalty Worse? Assessing Two Decades of German Reforms." 1–51.
- McGinn, K. L., M. Ruiz Castro, and E. L. Lingo. 2019. "Learning from Mum: Cross-National Evidence Linking Maternal Employment and Adult Children's Outcomes." *Work, Employment and Society* 33 (3): 374–400. <https://doi.org/10.1177/0950017018760167>.
- Moberg, Y. 2016. "Does the Gender Composition in Couples Matter for the Division of Labor After Child- Birth?" *IFAU Working Paper Series*.
- Olivetti, C., E. Patacchini, and Y. Zenou. 2020. "Mothers, Peers, and Gender-Role Identity." *Journal of the European Economic Association* 18 (1): 266–301. <https://doi.org/10.1093/jeea/jvy050>.
- Olivetti, C., and B. Petrongolo. 2017. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives* 31 (1): 205–230. <https://doi.org/10.1257/jep.31.1.205>.
- Patnaik, A. 2019. "Reserving Time for Daddy: The Consequences of Fathers' Quotas." *Journal of Labor Economics* 37 (4): 1009–1059. <https://doi.org/10.1086/703115>.
- Rabaté, S., and S. Rellstab. 2022. "What Determines the Child Penalty in the Netherlands? The Role of Policy and Norms." *De Economist* 170 (2): 195–229. <https://doi.org/10.1007/s10645-022-09403-x>.
- Rosenbaum, P. 2021. "Pregnancy or Motherhood Cost? A Comparison of the Child Penalty for Adopting and Biological Parents." *Applied Economics* 53 (29): 3408–3422. <https://doi.org/10.1080/00036846.2021.1881431>.
- Ruijs, W. 2013. "Acceptance of Vaccination Among Orthodox Protestants in the Netherlands."
- Sieppi, A., and J. Pehkonen. 2019. "Parenthood and Gender Inequality: Population-Based Evidence on the Child Penalty in Finland." *Economics Letters* 182:5–9. <https://doi.org/10.1016/j.econlet.2019.05.034>.
- Steinhauer, A. 2018. "Working Moms, Childlessness, and Female Identity." *LIEPP Working Paper* 79.
- Sun, L., and S. Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–199. <https://doi.org/10.1016/j.jeconom.2020.09.006>.
- Vella, F. 1994. "Gender Roles and Human Capital Investment: The Relationship Between Traditional Attitudes and Female Labour Market Performance." *Economica* 61 (242): 191–211. <https://doi.org/10.2307/2554957>.