ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis: International Bachelor Economics and Business Economics

From Innocence to Inequality: Examining the Child Penalty Across Sectors

Name Student: Bodine Franck
Student ID number: 562770

Supervisor: K.E.A. Hanemaaijer
Second assessor: dr. RA Stans

Date final version: 12th of July 2023

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract: Gender inequality in earnings can almost fully be attributed to having children according to existing literature. The impact of having a child on earnings of women relative to men is called the relative child penalty. In this paper, the relative child penalty and whether it differs across sectors is measured using the event study methodology by Kleven (2019b) on LISS panel data for the Netherlands. Firstly, results show a 30% medium-run relative child penalty in the Netherlands. Next, a significant difference was found across sectors, with relative child penalty levels ranging from 10% to 40%. An attempt at finding out which specific sector characteristics caused this difference was unsuccessful: no clear pattern could be found as sectors that were most similar across characteristics had quite diverging child penalty levels. When focusing on one characteristic at a time by sorting sectors into one of two clusters based on their mean value compared to the overall mean value of the characteristic, diverging child penalty levels might indicate which characteristic carries a larger weight in determining the child penalty level. The results seem to indicate that education level might make the biggest impact, followed by non-standard working hours and then a sector's dominant gender. The average earnings of a sector seemed to be the least important compared to the other characteristics.

Index

1. Introduction

In the first half of the 20th century, the male breadwinner/female caregiver family structure was the norm. Nowadays, a dual-earner structure is more common as gender ideology has changed quite a bit (Kroska & Elman, 2009). While gender inequality has been observed to decrease over time, a gap in earnings between men and women still persists today, also in Western countries (Chantreuil et al., 2021). According to Kleven et al. (2019b), almost the entire remaining gender gap in earnings can be attributed to the presence of children. This is because the earnings of women are impacted more strongly than the earnings of men after having a child. This concept is called the child or motherhood penalty, first coined by Budig and England (2001).

The remaining inequality on earnings for women can only ever be disentangled if the factors affecting it are researched and identified. Therefore, it is essential to know which characteristics of mothers and their employment might influence the child penalty. Quite a lot is known about the child penalty size across countries (Kleven et al., 2019a, Lebedinski et al., 2022; Zhou et al., 2022) and for different gender norms or childcare policies (Andresen and Nix, 2019; Rabaté & Rellstab, 2021; Kleven, 2022). Whereas not much is published about the child penalty size differing between sectors or industries, except for the research done by Fontenay et al. (2021) on the different levels of the child penalty per industry in Belgium. This gap in the literature is socially relevant because by clustering sectors together based on similar characteristics and measuring the child penalty for these different clusters, one can try to distinguish which characteristics might increase gender inequality. In this way, I try to analyze which aspects of a sector 'punish' mothers more in terms of labor market outcomes. As a result, it opens up the possibility to make sector specific policy recommendations in order to accommodate for a smoother transition for parents in continuing their career after having a child.

The child penalty will be estimated across sectors in the Netherlands specifically as it has been found to have a high child penalty level of around 46% after 7 years (Rabaté & Rellstab, 2021) or 47% after ten years (Artmann et al., 2022). This is higher than in most other European countries, the UK and the US. Only Germany and Austria have higher relative child penalty levels (Kleven et al., 2019b; Sieppi & Pehkonen, 2019; Andresen & Nix, 2020). This high level of the child penalty found in the Netherlands is especially puzzling when you take into consideration that the Netherlands ranked third in the EU on the Gender Equality Index in 2022 (EIGE, 2023).

Therefore, the main research question of this thesis will be the following:

Does the relative child penalty differ significantly across sectors in the Netherlands?

In order to find an answer to this question, several sub-questions have to be investigated. First of all, it has to be determined whether a significant child penalty can be measured in the Netherlands by comparing the income trends of mothers and fathers over time before and after having a child. After that, the child penalty should be measured in size across different sectors, to establish whether there is a significant difference across these sectors. Different clusters will be formed based on multiple attributes at once such as average education level, average earnings per sector and the nature of the work being done in the sector. This way, I can compare sectors that have similar characteristics and research whether they also have similar child penalty levels. Afterwards, the relative child penalty levels will be estimated for clusters that only consider one characteristic at a time, by sorting sectors into one of two clusters based on where the sector mean lies compared to the overall mean value of the characteristic. In doing so, I can theorize which attributes might have a bigger impact on the child penalty size.

The paper is structured as follows: section 2 will explain findings from relevant literature closely related to the topic of this thesis, section 3 will cover the data source and the descriptive statistics of the sample, section 4 will describe the methodology used, section 5 will present the results and section 6 will be a conclusion of the findings and their limitations.

2. Underlying theory

Child penalty

The child or motherhood penalty refers to the causal impact of having children on the labor market outcomes of women relative to men and is estimated using the event study approach based around first childbirth also applied in this thesis (Kleven et al., 2019b). There are three components that cause this mechanism of women falling behind men after having a child: the wage rate, the number of hours worked and the labor force participation. In Denmark, these three components were of relatively equal weight in determining the long-run drop in income of around 22% for mothers relative to fathers after having a child (Kleven et al., 2019b), whereas Rabaté and Rellstab (2021) measured a relative long-run child penalty of 46% seven years after childbirth in the Netherlands, which is more than two times as big. They found a penalty of around three times the size for earnings and hours worked compared to labor force participation and wage rates, showing that not all components are of equal weight in the Netherlands. The fact that the relative child penalty is mainly caused by the decrease in hours worked after childbirth, is in line with Dutch labor market trends given that working part-time is very popular for mothers and the likelihood of working part-time increases in the presence of children (Bosch et al., 2008).

The child penalty does not only differ per country, but also between mothers due to a difference in characteristics, such as education level (Anderson et al., 2002; De Quinto et al., 2021). The least-educated mothers bear no penalty whatsoever, while mothers who have a high-school or college education earn ten percent less than childless women with the same education level, this difference becomes even bigger if the number of children rises (Anderson et al., 2002). Contrastingly, De Quinto et al. (2021) found that college-educated mothers in Spain are more likely to work part-time and remain employed than non-college educated mothers, resulting in a significantly smaller decrease in earnings and number of days worked for college-educated women. Thus, it is important to take education level into account when making a distinction between sectors in our analysis to estimate the child penalty per sector cluster. This is because it can affect the child penalty level either directly or by influencing the number of hours worked, which is the main driver of the child penalty in the Netherlands.

Gender norms have also been found to play a role in the size of the child penalty. Andresen and Nix (2019) looked at child penalty differences between heterosexual and lesbian parents, caused by gender norms. They found that both mothers in lesbian couples experience a child penalty around a year after childbirth, where the mother who gives birth suffers from a penalty twice the size of her partner. Yet, this difference is significantly smaller than it is for heterosexual women. Most striking is the fact that the penalties for both mothers in a lesbian couple converge and after 5 years, almost entirely vanishes. They explain this because both partners in lesbian couples are seen to decrease their number of working hours, whereas heterosexual fathers act unlike their counterparts by increasing working hours. Also, more mothers work full time in lesbian couples than in heterosexual couples. Because of this, only heterosexual parents will be used in this analysis, so that the differing gender norms will not impact the results for any of the sector clusters.

Whether there is a significant difference in child penalties for women between industries was researched by Fontenay et al. (2021). They found significant differences between 11 defined industries in Belgium: the estimated child penalty is about twice as large (30%) in health, real estate and hospitality, compared to education and public administration (around 15%). They explain that in public sectors such as education, holidays are well defined and working hours are quite constant, whereas sectors with larger penalties such as health and hospitality suffer from changing schedules and working non-standard hours might be expected from workers. A positive correlation was established between the size of the long-run child penalty per sector and multiple indicators such as having an atypical work schedule, having a poor work-life balance and working irregular hours. This would mean that in sectors where women suffer from a self-reported poorer work-life balance, work more often at night or during the weekends or find themselves facing a different schedule every week suffer from a larger long-run relative child penalty. The paper by Fontenay et al. is the most similar to the research topic of this thesis, where a comparable analysis will follow for the Netherlands. This thesis will add to the analysis done by Fontenay et al. by using a cluster design instead of conducting the methodology on individual sectors, which allows us to research whether heterogeneity in the relative child penalty exists for a subset of different characteristics.

Sector characteristics potentially influencing the child penalty

There is a scarcity of studies on sector differences in relative child penalty levels, but the gender wage gap has been compared across sectors quite often. This allows us to hypothesize which sector characteristics increase gender inequality and thus, might affect the child penalty. The private sector has been found to show more evidence for gender inequality and larger wage gaps than the public sector, likely due to collective bargaining being less common and the lack of information on pay comparability (Miller, 2009; Bergmann et al., 2019). Next to that, sectors or positions within a sector that are male-dominated tend to show higher wage gaps than female-dominated sectors (Bergmann et al., 2019; Shahrabani, 2007).

Shahrabani (2007) compared existing gender wage gaps and predicted wage gaps in her model using four economic sectors: the public sector, services based on high-education, services based on low-education and the products sector. The largest wage gap was found in the high-education service sector, which seems in line with the finding that the child penalty is larger for women who are more educated. In addition, this paper finds evidence for employer discrimination in the male-dominated sectors (products sector and low-services sector), because the gender wage gap would have increased if women had the same characteristics as men in their model. A similar sector classification was followed in this thesis as the one Shahrabani (2007) uses to estimate the child penalty levels for different sector clusters.

Sector characteristics that are closely related to parenthood are also a likely indicator for child penalty divergence between sectors. Flexibility has been found to be an important factor in decreasing the gender pay gap between mothers and childless women, where flexible hours play the most important role (Fuller & Hirsch, 2019). However, in high-income or high-status jobs this demand for flexibility and taking breaks from working are punished, despite being a common demand for parents who just had a child (Bertrand et al., 2010). Flexibility in the form of worker control on work hours are a protective factor over work/life and work/family balance, whilst working non-standard hours have a negative effect on this balance (Arlinghaus et al., 2019). A bad work/life balance has also been found to be a reason for parents to reduce their working hours (Arlinghaus et al., 2019), which is an interesting finding as this reciprocal relationship between work/life balance and working hours will likely influence the child penalty across sectors in the Netherlands. This could indicate the child penalty to be higher for jobs with a worse work/life balance or jobs which are stricter in terms of labor culture. Kumari and Devi (2016) found the worst work/life balance in the healthcare sector, while the balance was the best in the education sector. The highest and lowest child penalty levels estimated by Fontenay et al. (2021) were also found in the respective sectors, indicating that work/life balance might be an important factor for our research.

Considering the findings described above from relevant literature, I would expect the child penalty to be higher in sectors which are:
i) private ii) stricter in necessary qualifications or education level iii) male-dominated or traditionally seen as masculine iv) have a worse work/life balance.

3. Data source

To estimate the size of the child penalty across different sectors, I will use the Longitudinal Internet Studies for the Social Sciences (LISS) panel. The LISS panel data is made up of around 7 500 individuals in 5 000 households, the traditional sample has been randomly drawn in collaboration with Statistics Netherlands (CBS) to achieve a true probability sample.

The LISS panel data archive consists of surveys measuring background variables, which are conducted monthly, and different core studies focusing on more specific topics conducted yearly. I will be using the monthly waves of background variables to retrieve the income of the respondents, their education level as well as their gender. Furthermore, I will use the core study on Family and Household to retrieve the number of children per respondent and at which point in time their first child was born. Lastly, the Work and Schooling core study will be used to retrieve in which one of the fifteen sector categories defined in the survey the respondent works.

I am planning to use a time frame of 7 years to measure the child penalty, so that I can measure the income of both parents 2 years before and up until 5 years after the birth of their first child. Because the available LISS panel data starts at the end of 2007, the first available full year of monthly background variable waves is the year 2008. The most recent core studies on Work and Schooling and Family and Household are from the year 2022. Therefore, to make sure that the parents' income has been observed at least once before or after having a child, I will look at parents who had their first child from 2009 until 2021.

In order to create one dataset containing all the necessary variables, I will merge the datasets by matching them on identical respondent identifiers.

3.1 Variable definitions

Sector

Sector is a categorical variable with 15 groups. The sector defines which sector the respondent was first observed working in. This means that sector data from 2008 is used and if missing, the next available sector data will be used. This is done because taking the most recent sector could bias the results as women are prone to switch to more family-friendly or public sectors after having a child (Kleven et al., 2019b). The original question from the LISS dataset contains 15 categories, but nobody in the sample works in mining, so that sector was dropped out.

The sector 'Other' was not used as the LISS codebook did not give a description of what this sector exactly entailed. Thus, our final sample consists of 13 different sectors. The sectors were then grouped together in clusters to maintain a decent sample size, as there are not enough individuals per sector to obtain an accurate estimate for all 13 sectors separately.

The original sectors were divided into different clusters to create a distinction based on multiple characteristics at once. The sectors were allocated based on descriptive statistics such as average education level, average annual earnings and share of (fe)males per sector (table A.2, figure A.3, figure 3.2.2) as well as the goal/nature of the work. The individual sectors that were most similar across all characteristics were then clustered together. This created five sector clusters: the healthcare, public, high-services, low-services and products sector clusters. The classification is similar to the one used by Shahrabani (2007) in his paper on gender inequality in different economic sectors.

The healthcare sector was the only original sector analyzed on its own, because it contains the same number of observations as the other four clusters and is rather unique in certain characteristics such as having a high share of workers working in the weekend, in the evening or at night (Figure A.4 in the appendix). Fontenay et al. (2021) found that the healthcare sector had the largest child penalty as well as the worst work/life balance out of the sectors they examined for Belgium. Thus, I believe conjoining the healthcare sector into a cluster with education and government services would cause for a higher estimate of the child penalty, mainly driven by the healthcare sector observations. In the interest of making relevant conclusions about which sector characteristics might indicate a larger child penalty for mothers working in the sector, I separated the healthcare sector from the public sector cluster.
The public sector cluster consists of government services, education and healthcare/welfare. The high-services sector cluster consists of financial and business services. The low-services sector cluster consists of catering, retail trade and environmental services. Lastly, the products sector cluster consists of agriculture, industrial production, utilities production, construction and transport/storage.

In further analysis, the sectors have been split repeatedly into 2 categories: feminine or masculine, high or low education, high or low earnings and finally, working standard or non-standard hours. When deciding whether a sector falls under the feminine or masculine category, the share of (fe)males in the sample working in the sector (figure 3.2.2) was considered. The masculine group consists of agriculture, industrial production, utilities production, construction, transport/storage, financial, business services. The feminine group is made up of retail trade, catering, government services, education and healthcare/welfare. Even though the majority working in the financial sector is female in the sample, this sector is mostly male dominated, especially in higher positions as described by Knights and Tullberg (2014). The opposite holds true for retail trade, as this sector shows a great presence of female managers based on Amin and Islam (2014). Therefore, the choice was made to allocate these sectors to the opposite gender instead of the gender of the majority working in these sectors as suggested by the sample.

For the other characteristics, the sectors were allocated based on the mean value for each respective characteristic. If the mean value of an individual sector was above or below the mean value for all sectors considered combined, this determined which of the two clusters the individual sector was allocated to. Table A.2, figure A.3 and figure A.4 show the descriptive statistics per sector for these characteristics (education, average annual earnings and working (non-)standard hours).

Income

Gross monthly earnings were used to obtain the income level. The monthly data was opted for over the annual income available in another dataset (Economic Situation: Income) because it allows us to estimate income more precisely as there are more income observations for the same individuals. The monthly data on income was used to make a more precise estimate of annual income than the annual income variable provided in the Economic Situation: Income dataset.

There are some income outliers as seen in the box plot (figure A.1) in the appendix. This could bias the results, especially because the sample size is not large. The outliers were removed, even though having a much higher income than the rest of the sample does not necessarily have to be incorrect. Because the outliers are observed to be from the same respondent, this seems to corroborate the fact that these outliers do not by definition have to be errors made when storing income in the dataset. However, the two highest income points belonging to the same individual change the results by quite a large margin. Therefore, these income points have been removed as the results cannot be driven by the outliers.

Age (when having a child)

I create two variables, age and age when having a child. In this way, the descriptive statistics at $t-1$ can show the average age of respondents when having a child, which is their age at $t=0$, one year later. Age is the difference between the year in which the survey was taken and the birth year of respondents. Age when having a child is the difference between the birth year of their first child and their own birth year.

Number of children

The number of children is generated by looking at the number of biological children that the respondent has at the end of the event time frame ($t=5$).

Education level

The dataset contains a variable using the 6 education level categories the CBS (2021) uses. The lowest level is primary school and the highest level is WO (university). The CBS often uses a three-scale ladder of education levels ranging from poorly, relatively to highly educated. Following this logic, the six original categories are rearranged to follow the three-scale level system as I find this system to be more indicative.

Table 3.1

Sample Selection Table

Subgroup specified in sample	Number of Individuals in Sample
All survey respondents	31 737
Consistent gender and birth year throughout dataset	30 628
Respondents also in the work and family datasets	15 516
Respondents with children	10 020
Parents with children born between 2009-2021	2 116
Parents who have biological living children	1 573
Parents who were older than 20 when having a child	1 573
Heterosexual respondents	1 295
Most recent sector is known	1 229
Parents with data on income during event time window	1 045

Note: The dataset is cleaned by dropping any observations I cannot use for the analysis. Any respondents who changed in gender or birth year during the data collection over the years have been dropped out due to discontinuity in a time-invariant variable in the dataset. First, all respondents are matched on their identical respondent ids for all three datasets. Next, respondents who are not parents or did not have their first child between 2009 and 2021 are dropped. Furthermore, I will not include parents who adopted their children, because the date of birth of the child does not always match the time the child was adopted into the household. Parents whose child passed away were also dropped, as this event probably greatly affects the labor market choices of parents in an abnormal sense. Teen parents will also not be included in this analysis, as they are outliers in the sense that they most likely have not had a stable income or career before or after the birth of their child and thus, the determinants of the child penalty are different for these parents. Although, no parents left in the sample at this point were teen parents. Parents who are not in a heterosexual relationship will also be dropped out, as gender norms are a main driver of the child penalty, and this would cause a less reliant analysis. Lastly, I need data on the sector the respondent works in as well as their income to perform the analysis. So, parents who did not have data for either of these things had to be removed. Parents with an income of 0 were included, only missing observations were dropped.

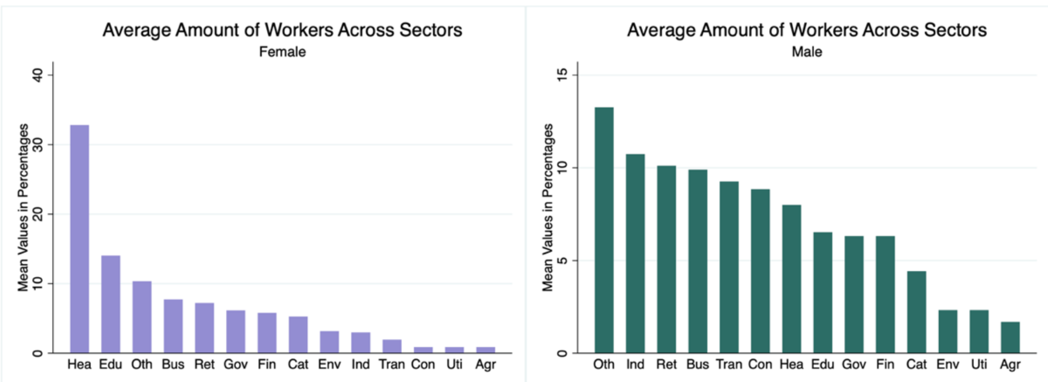
3.2 Descriptive statistics

When analyzing the descriptive statistics table (table 2 in the appendix), I find that the sample is relatively balanced in gender: 55% of the sample is female, whilst 45% is male. The mothers in the sample earn quite a lot less than fathers do; mothers earn 27 880 euros on average, while fathers earn 35 220 euros on average. This means that the average annual income of mothers is 7 340 euros lower. This is especially interesting when considering that the mothers in our sample are relatively more educated. While the majority of the total sample is highly educated, 62% of women are highly educated while 53% of men are in the sample. Only a small portion is poorly educated, there are again relatively more men (10%) in this category than women (4%). However, women in the sample work around 6 hours less a week than men do: the mean value of weekly work hours for men is 39.23 and 33.69 for women. The mean value indicates that a large share of women in the sample work part-time; the OECD makes a distinction between part-time and full-time at the 30-hour threshold, whilst the CBS uses a 35-hour threshold. Using either definition, I conclude that working part-time is rather popular in for women in our sample.

The most interesting finding is the differences across sectors. As seen in figure 3.2.1 below, the relative share of males working in each sector is quite similar, with all values being below 15%. For women, the healthcare/welfare sector is extremely popular, it has about roughly twice the size compared to the second most popular sector. As around 33% of women in our sample work in the healthcare/welfare sector, this might bias the result of the regression analysis done per sector as the other samples will be smaller and thus, less representative.

Figure 3.2.1

Bar Graph of Average Amount of Workers Across Sectors for both Females and Males

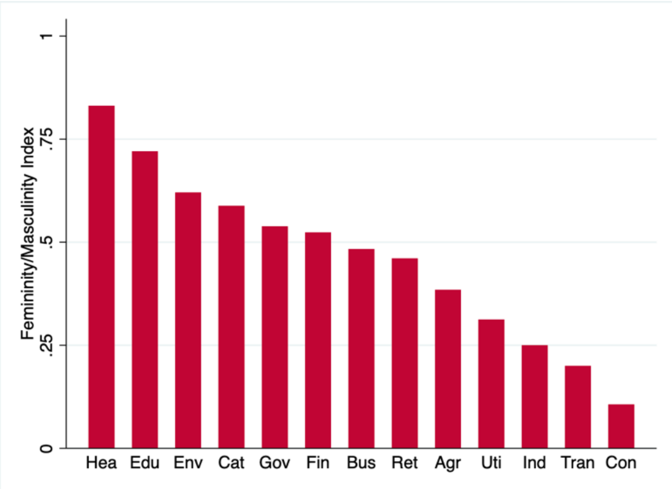


Note: The sector categories were recoded to use the first three letters of their label name to fit them all on the axis and to have a more organized graph. The female and male bar graphs have different scales for the y axis: the female graph reaches from 0 to 40 percent, while the male chart reaches from 0 to 15 percent.

Another interesting finding is that different sectors are more popular across genders, the sectors that are traditionally perceived as more 'masculine' such as transport/storage, construction and industrial production also have more male workers relative to female workers in the sample. This can be seen by looking at these sectors in the femininity/masculinity index in figure 3.2.2 below.

Figure 3.2.2

Femininity/Masculinity Index per Sector



Note: The index shows whether a sector if male- or female-dominated by examining the share of (fe)males working in each sector. The closer to zero, the more masculine the sector is. The closer to one, the more feminine the sector is, as the sex variable takes a value of 0 if male and takes a value of 1 if female.

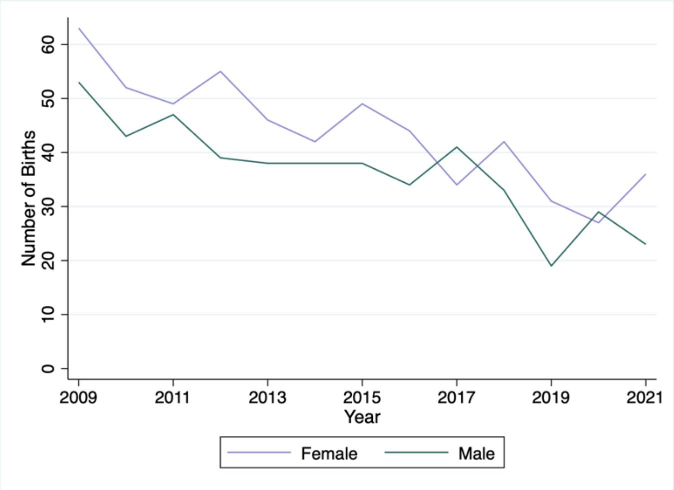
The men in the sample are generally older when they become a first-time parent, as compared to the women in the sample. The average age one has when they have their first child is 30.32 years for females and 32.79 years for males. These results are very much so in line with the average age of first-time mothers and fathers in 2021 in the Netherlands as published by the CBS Statline (2022). In 2021, Dutch mothers were 30.3 years when they had their first child, while Dutch fathers were 32.8 years old.

The average amount of children parents have is also relatively in line with the average amount of children the sample has at the end of the event time frame. In 2021, Dutch mothers had 1.62 children on average, similar to the mean value of 1.60 children for the mothers in our sample (CBS Statline, 2022). Dutch fathers had 1.60 children on average in 2021, which means that mothers had more children than fathers based on Dutch data, which is the other way around for our sample. Even though the mean values are very similar, fathers in the sample had 1.61 children on average.

Figure 3.2.3 below shows how many mothers and fathers became a parent for the first time per year. The number of births is spread quite evenly across gender, considering the fact that we have more females than males in our sample. The number of births differs over the years, but there is not an extreme outlier that would potentially cause problems or bias the results.

Figure 3.2.3

Line Graph of Number of Births of First-Born Children per Year for both Females and Males



Note: The y-axis showing the number of births, ranges from 0 to 60. The x-axis is ordered on all years between 2009-2021. The purple line represents females and the green line represents males.

4. Methodology

To measure the size of the child penalty (across sectors) I will be using the event study method designed by Kleven et al. (2019b). This approach looks at sharp changes in labor market outcomes for mothers relative to fathers around the time of childbirth. The time of birth of the first child will $t = 0$ and all years included in the analysis will be indexed relative to the birth of the first child. This will result in a medium-run event time span of $t - 2$ to $t + 5$. Kleven et al. (2019b) estimate the long-run child penalty at times $t = 10$ or $t = 20$. Since our time frame is smaller, I will end up estimating a medium-run penalty at time $t = 5$. Because of this, the child penalty will be less constant than a more long-run one.

The main assumptions of this method are smoothness and exogeneity. For smoothness, what is meant is that income determinants that are unrelated to the presence of children such as age or education evolve smoothly over time, instead of changing abruptly. Exogeneity relates to the fact that the timing of the birth of a child is exogenous to earnings, this means that your income level will not directly affect whether you will decide to have a child or not. Combining these two factors, when changes to earnings around the event time (childbirth) are sharp, one can relate them to the event. This is because if one were to observe sharp changes in income around childbirth, it would most likely be because of this event rather than being a result of other income determinants as they evolve smoothly over time.

These assumptions are used to estimate the short-run impact. However, the smoothness assumption alone is not sufficient when looking at a longer time frame as income determinants unrelated to children can change considerably over a longer period of time. Thus, I will take the time and life cycle trends into account.

$$Y_{ist}^{gc} = \sum_{j \neq 1} \alpha_j^{gc} \cdot \mathrm{I}[j = t] + \sum_k \beta_k^{gc} \cdot \mathrm{I}[k = age_{is}] + \sum_y \gamma_y^{gc} \cdot \mathrm{I}[y = s] + \varepsilon_{ist}^{gc} \quad (1)$$

The regression above shows the model used to measure the annual income in euros (Y_{ist}^{g}) for individual i working in sector c of gender g in year s and event time t. The three dummies included in the model above are event time, age and year dummies from left to right. The event time dummy has to be omitted at $t = -1$ according to the method because in doing so, the event time coefficients, which estimate the effect of the birth, are all measured relative to the year right before the first childbirth. The other two dummies are used to control for life cycle and other trends over time.

The next step is for us to turn the level effects into percentage effects:

$$P_t^{gc} = \frac{\hat{\alpha}_t^{gc}}{\mathrm{E}[\tilde{Y}_{ist}^{gc} \mid t]} \tag{2}$$

Where $\tilde{Y}_{ist}^{gc}$ shows the predicted outcome when omitting the event time dummies, which gives us the result when a child would not have been had. Thus, P_t^{gc} gives us the effect of having children relative to the counterfactual of not having children.

After running the regression model for mothers and fathers separately and turning it into a relative measure, I want to compare them to get the relative child penalty per sector c:

$$P_t^{c} = \frac{\hat{\alpha}_t^{mc} - \hat{\alpha}_t^{wc}}{\mathrm{E}[\tilde{Y}_{ist}^{wc} \mid t]} \tag{3}$$

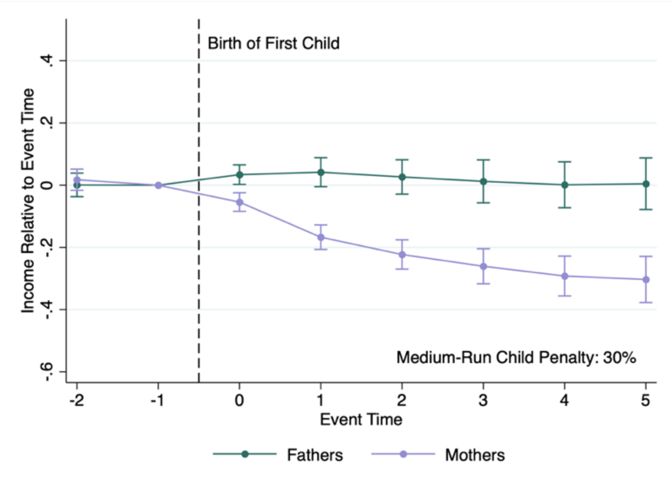
5. **Results**

5.1 Main results

When applying the methodology to the entire sample, I find the presence of a child penalty in our subset of individuals. As seen in figure 5.1 below, the medium-run relative child penalty on income is 30%. Up until five years after birth, almost no effect is witnessed on the income of men in our sample, whereas women start to earn significantly less in a continuously decreasing rate during the time span of five years after childbirth. Thus, I can conclude that there the first hypothesis of the existence of a child penalty holds true.

Figure 5.1

Income Child Penalty of the Entire Sample



Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

The child penalty on earnings consists of three components: work hours, labor force participation and wage rate. To see whether work hours are of larger importance than the other two components in the Netherlands as theorized before, I will look at the decomposition of the child penalty on earnings. Data on net/gross wages was only available for 2018-2022, so therefore the wage rates were calculated using the income of the respondent divided by the number of hours worked. For labor force participation, a dummy variable was created that took a variable of 1 if the respondent worked more than 0 hours that year and 0 otherwise.

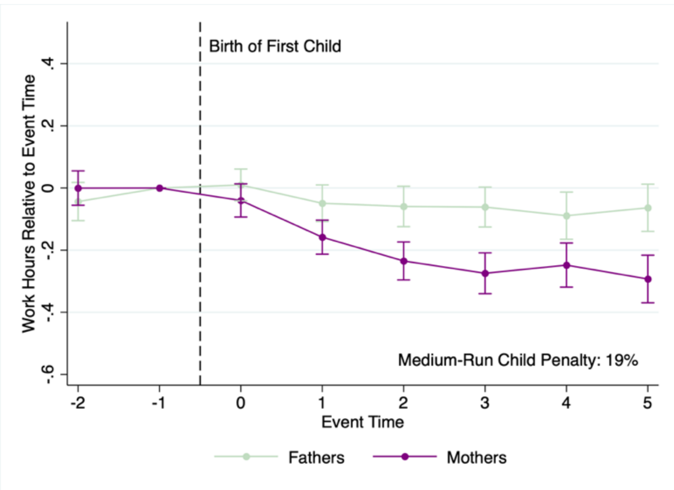
Figure 5.1.2 below shows the medium-run relative child penalty on work hours, with a magnitude of 19%. The figure shows a negative trend in work hours for both mothers as well as fathers, unlike the child penalty on earnings in which there seems to be almost no trend for fathers. Figure 5.1.3 below shows the medium-run relative child penalty on wage rates. The figure shows a stable trend for both men and women, with a very minimal diverging trend in wage rates for the two genders from t=3 onwards. This is somewhat unexpected, given the fact that multiple authors find a strong pattern in reducing wage rates after having a child (Kleven et al., 2019b; Rabaté & Rellstab, 2021). In contrast, Artmann et al. (2022) find that there is hardly a drop in wage rates right after childbirth for women in the Netherlands, with a gradual and relatively small decreasing trend in the years following. They conclude that work hours are the main driver of the relative child penalty on earnings. This flatter trend in the wage rate decomposition is also found by Fontenay et al. (2021), who initially see an increase in wage rates for mothers right after giving birth and a relatively flat decreasing trend in the years after leading up to a relative child penalty on wage rates of 9%.

Figure 5.1.4 below shows a medium-run relative child penalty of 5% on labor force participation, with a slight decreasing trend for both mothers and fathers. The significance of the regression coefficients using labor force participation as dependent variable is somewhat smaller than the significance of the coefficients using earnings or work hours as dependent variable, as seen in table A.3 in the appendix.

The insignificance of the relative child penalty on wage rates in combination with the small magnitude of the relative child penalty on labor force participation, seem to indicate that work hours are of larger importance in determining the relative child penalty on earning in the Netherlands. Still, one should be careful in stating this fact, as there was a slight lack in availability on the number of hours worked compared to the number of observations for income. This in turn also affects the accuracy of the estimates for the relative child penalty on wage rates and labor force participation, as those dependent variables are calculated using the available information on work hours.

Figure 5.1.2

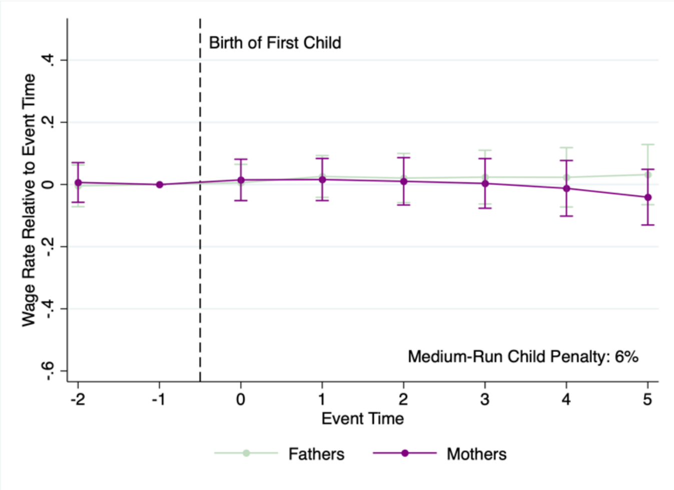
Child Penalty on Work Hours for the Entire Sample



Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

Figure 5.1.3

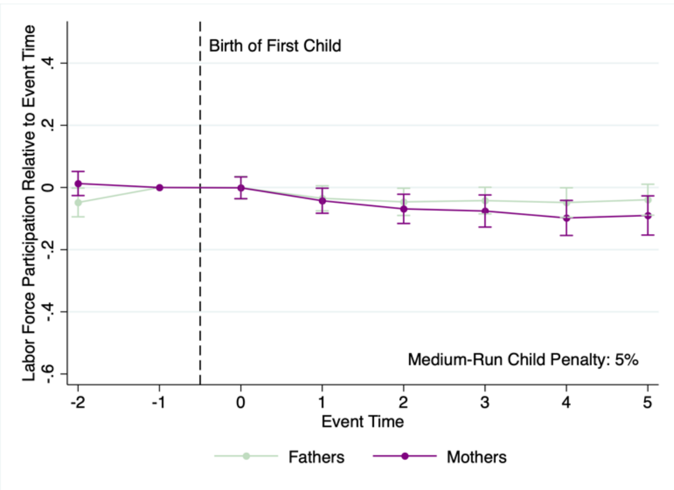
Child Penalty on Wage Rates for the Entire Sample



Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

Figure 5.1.4

Child Penalty on Labor Force Participation for the Entire Sample



Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

5.2 Child penalty on earnings for different clusters

When sectors are clustered together based on multiple characteristics, one can try to find more specific correlations between sector characteristics and the child penalty. Figure 5.2.1 shows the medium-run relative child penalty estimates for five different sector clusters. Most noticeable is the public sector cluster, showing the lowest estimated relative child penalty with a level of only 10%. This seems to confirm the view of the public sector being rather family friendly. The highest relative child penalty (40%) is found in the healthcare sector. These findings are in line with those of Fontenay et al. (2021) who also found the highest child penalty in the health and social work sector, with the lowest penalties estimated in the education and public administration sectors. Even though the healthcare sector in the sample is characterized by being the most feminine and having quite low annual earnings, the estimated relative child penalty is surprisingly high. The sector does have higher education levels as well as a high share of the sample working outside regular office hours or in the weekend, which might contribute more to the level of the child penalty than the femininity or average earnings of the sector.

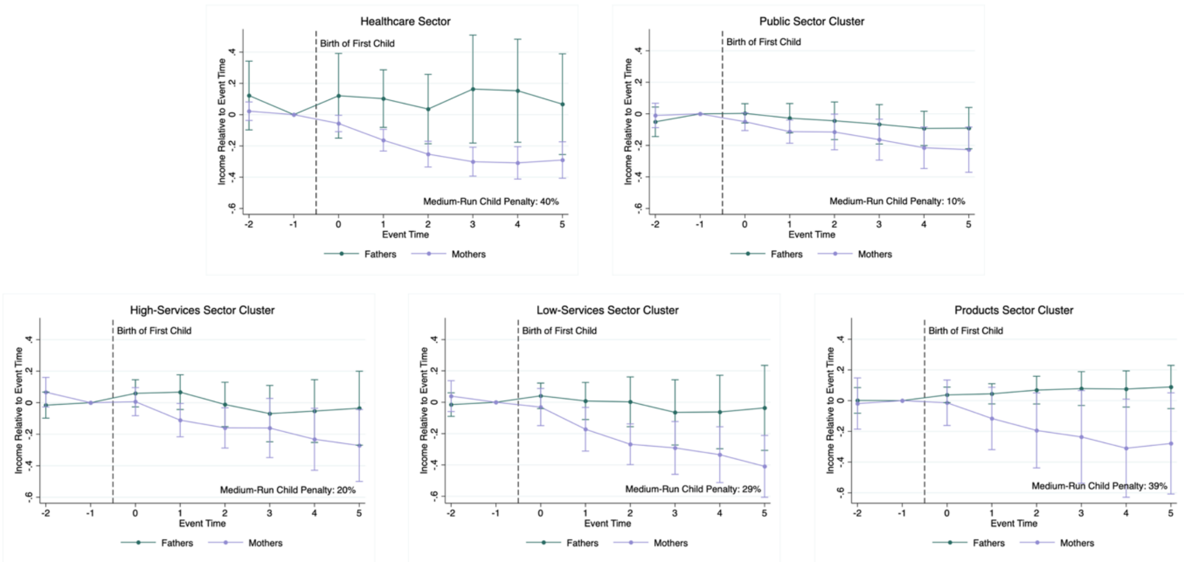
The products sector cluster comes in at a close second, with an estimated 39%. The results for the products sector cluster are in line with those of Shahrabani (2007), who found evidence for employer discrimination in the products and low-services clusters. The low-services sector cluster also has a relatively high medium-run relative child penalty in our sample: 29%. This does indicate that our earlier hypothesis suggesting that a sector characterized by a higher education level would indicate a higher child penalty likely does not hold. As the clusters in our sample that are characterized by a lower education level such as the products or the low-services sector clusters have higher child penalties than those which are characterized by higher education levels (public and high-services sector), with the healthcare sector being an exception. However, as all clusters have a relatively high average education level, it is harder to make a conclusion because there is not a single sector in which the majority of workers in the sample fall in one of the lowest two education levels. Therefore, I can only look at clusters that have lower education levels compared to other clusters. In reality, the difference between clusters might not be big enough to cause a much higher/lower penalty.

Also when looking at average earnings per cluster, I do not seem to find a trend. The clusters with relatively low average annual earnings such as the low-services sector or the healthcare sector have quite different levels of the child penalty. The same holds for clusters with higher earnings such as the products or the high-services clusters.

Thus, I cannot find a clear pattern between the child penalty and multiple characteristics such as education level, femininity/masculinity or average earnings across different clusters. The clusters that have the most similar relative child penalty levels differ quite a lot in characteristics: the products sector is very masculine, has high earnings and a lower average education level, while the healthcare sector is the very feminine, with low average earnings and a higher education level. In contrast, the clusters with the most divergent penalty levels are rather similar in terms of characteristics: the public sector is also quite feminine, has a high education level, but has middle-of the road earnings. This might be due to the fact that there are other elements at play within sectors that I did not include in our analysis such as: individualism/altruism, the availability of information on pay, necessary qualifications or the rank one possesses within a sector. In addition, our sample size was rather small and did not contain as many observations per sector. This is why the confidence intervals are larger for men in the healthcare sector and women in the products sector cluster, as these gender and sector combinations had a low number of observations. Therefore, the results should be approached with some caution.

Figure 5.2.1

Income Child Penalty for all Five Sectors Clusters



Note: From left to right, top to bottom: the healthcare sector, the public sector cluster, the high-services sector cluster, the low-services sector cluster and the products sector cluster. Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

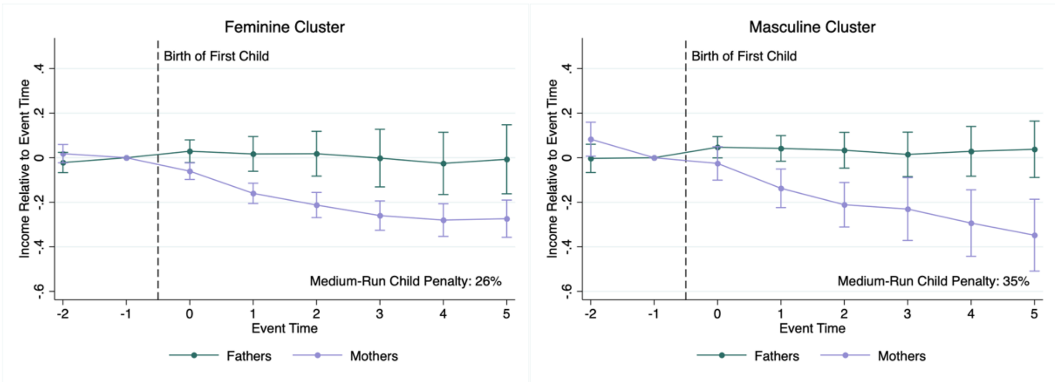
5.3 Heterogeneity between characteristics

The clusters above are made based on multiple characteristics at once, but when focusing on one characteristic (femininity/masculinity) only, one can try to research whether a specific characteristic might be of greater importance in determining child penalty levels across sectors. The findings for the public and products sector clusters seem to indicate that femininity/masculinity might be an important characteristic in determining relative child penalty levels.

When focusing only on femininity/masculinity, I seem to find evidence for the hypothesis that women working in male-dominated sectors or sectors that have been perceived as masculine will suffer from a higher child penalty. Figure 5.3.1 below indicates that mothers suffer from a relative child penalty of 26% 5 years after giving birth in feminine sectors, whereas they suffer from a 35% relative child penalty in masculine sectors. This difference is in line with the findings from Bergmann et al. (2019), who found that there is more gender inequality in male-dominated sectors. Although the difference is not enormous, it is still noteworthy. Still, the specific sectors that make up the feminine and masculine clusters vary considerably in multiple characteristics such as average earnings, education levels and work/life balance within each cluster. Thus, I cannot conclude that a sector that is more popular amongst men will always lead to a higher relative child penalty for women working in this sector.

Figure 5.3.1

Income Child Penalty for both Feminine and Masculine Sectors

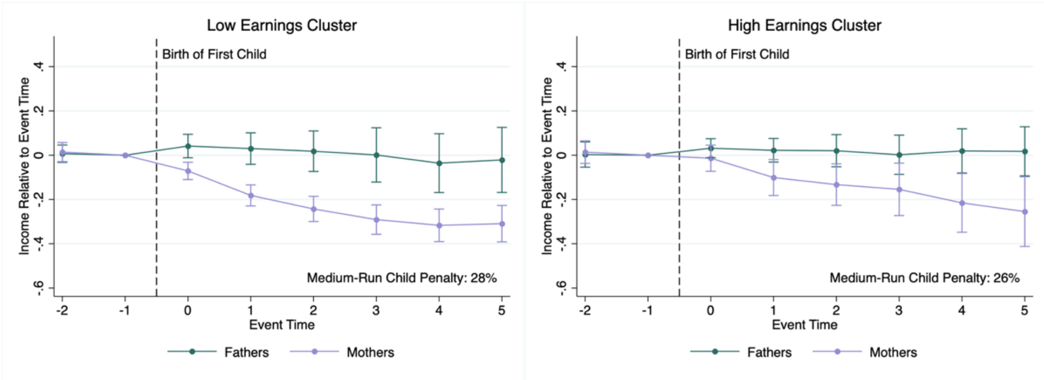


Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

When comparing the relative child penalty in sectors with lower average annual earnings compared to sectors with higher average annual earnings, the relative child penalties lie closer together. Figure 5.3.2 below shows that the relative child penalty is 26% for the sector cluster with average earnings below the sample mean, whereas the estimated relative child penalty is 28% for the sector cluster with higher average earnings. This could be because the average earnings of a sector might not be as important in determining the child penalty level as opposed to other characteristics. Another reason for this small divergence in child penalty levels is probably due to the fact that the average earnings are quite similar across sectors in the sample, as seen in figure A.3.

Figure 5.3.2

Income Child Penalty for both Low and High Earnings



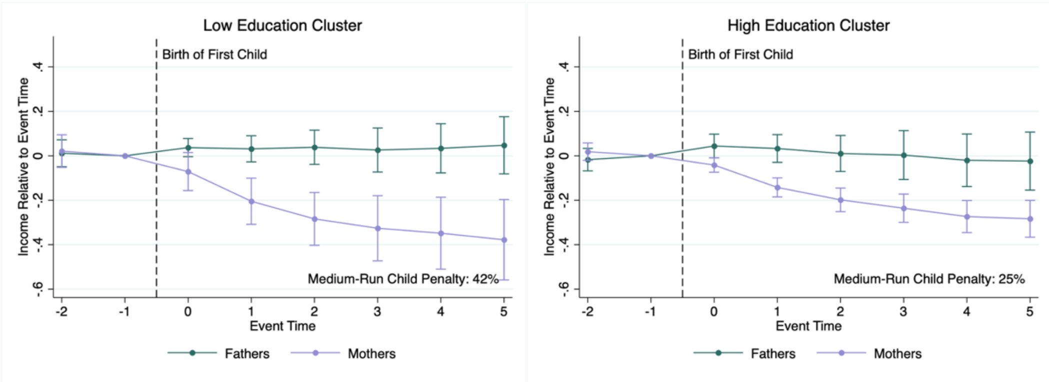
Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

Available literature varies on the effect of education level on the earnings of women. The highest gender wage gap was found in the high-services sector by Shahrabani (2007), which is based on a high-education level. This seems support the finding that the least-educated mothers do not suffer from a motherhood income penalty, whereas college or high-school educated mothers do (Anderson et al., 2002). However, when looking at the effect of childbirth on the earnings of women relative to men, De Quinto et al. (2021) found that the relative child penalty on earnings was lower for college-educated women compared to non-college educated women. In our sample, I found a stark difference in relative child penalty levels when clusters were formed based on education level. Figure 5.3.3 below shows that the low education cluster suffers from a much larger relative child penalty (42%) than the high education cluster (25%). This is likely due to a contrasting change in employment for women with differing education levels. Women who went to college were found to be more likely to stay employed and less likely to decrease the number of days worked than women who did not go to college, resulting in a smaller penalty on earnings (De Quinto et al., 2021). This finding could explain the results found in the sample, as work hours seems to be the most important determinant of the relative child penalty in our sample.

Also, when comparing these results to the five clusters formed before, I find evidence for this hypothesis. Sector clusters with the lowest education level (products sector cluster, followed by the low-services sector cluster) are also seen to have higher relative child penalties than those with higher education levels (high-services sector cluster and especially the public sector cluster). Only the healthcare sector does not seem to follow this pattern, as it has a high education level, but also a high relative child penalty.

Figure 5.3.3

Income Child Penalty for both Low and High Education



Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

A sector characteristic not considered before when forming the five clusters that could be correlated with the child penalty is work/life balance. Bertrand et al. (2010) found that high-income, high-status jobs punish a need for flexibility. This means that in some jobs, it is harder for mothers to take breaks from work or to change their existing working conditions. Mothers have been found to be more likely change to more family friendly jobs and thus, have higher probabilities of working in the public sector after having a child (Kleven et al., 2019b). Fontenay et al. (2021) looked at the correlation between having a poorer work/life balance and the estimated child penalty and established a positive correlation between working irregular hours, having an atypical work schedule and reporting a poor work/life balance and the level of the child penalty. A similar approach was used in this paper, as seen in figure A.5 in the appendix, where the five sector clusters and their respective relative child penalty levels are mapped out against working non-standard hours. The left side shows the share of respondents in the sample who report working in the weekend per cluster, while the right side shows the share of respondents in the sample who report working in the evening or at night. I could not establish a clear positive correlation, as the products and public sector clusters deviate strongly from the estimated responses.

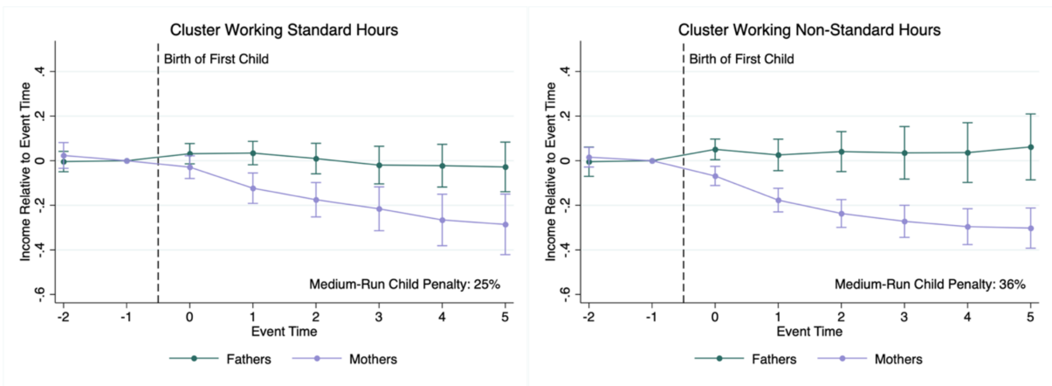
For the products sector cluster, I would expect to see a lower child penalty given the share of the sample given the fraction of respondents working non-standard hours. However, this might be due to the cluster design used in this paper, as the fraction of respondents working non-standard hours varies quite some in this sector as seen in figure A.4 in the appendix. In the transport/storage or agriculture sectors, working non-standard hours is more common whereas it is not (as much) in utilities production or construction.

Therefore, I decided to estimate the relative child penalty on earnings for two separate clusters: one where working non-standard hours (weekend/evening/night) is uncommon and one where it is common. This results in the following medium-run relative child penalties: a 25% relative child penalty for the cluster working standard hours and a 36% relative child penalty for the cluster working non-standard hours. This seems to indicate that this characteristic might be of importance in determining the child penalty level per sector. If one looks at the results for the five sector clusters described earlier, a higher share of the cluster working non-standard hours does seem to lead to a higher relative child penalty. The healthcare sector has an especially high share of individuals working non-standard hours and this sector is also seen to have the highest relative child penalty. For the high-services sector cluster, the share of individuals working non-standard hours is quite low, and this sector cluster has been estimated to have the second lowest child penalty. The low-services sector cluster falls into the non-standard hours category and shows a higher relative child penalty than the low-services sector cluster, which falls into the standard hours category. For the other two sector clusters (public and products), the pattern seems to be less clear. Both clusters are made up of individual sectors that differ in the share of individuals working (non-)standard hours, which causes the individual sectors that make up these clusters to fall into different categories.

Again, I cannot conclude that this characteristic causes a higher child penalty, but these findings seem to indicate that the share of individuals working non-standard hours might carry some weight in determining child penalty levels across sectors.

Figure 5.3.4

Income Child Penalty for Sectors Working Standard Hours or Non-Standard Hours



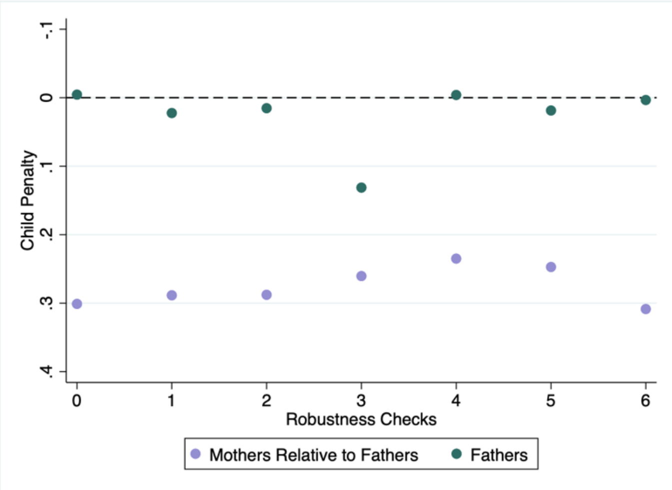
Note: Event coefficients are shown in dots and plotted per event time. The confidence intervals are the vertical lines through the event coefficients. Standard errors are clustered at the individual level. The green line represents fathers, and the purple line represents mothers. The medium-run relative child penalty is measured at t=5. This medium-run child penalty represents the rate at which women are falling behind men.

5.4 Robustness

The robustness of the estimation of the child penalty on earnings for the main sample is presented in figure 5.4 below. Larger sample sizes with different birth time and event time windows (RC1, RC2 and RC6) show a similar relative child penalty level of 29% or 31%. RC3, RC4 and RC5 use different event time windows ranging from t =3 to t = 7, but all use a similar restriction so that the birth time window is restricted to the years before/after which parents will always be observed, leading to a smaller sample size compared to the reference sample. RC3 estimates a longer run penalty at t = 7 of 26%, but this design can be rendered questionable as it finds a negative child penalty for fathers, which might be due to this check having the most restrictive sample. RC4 estimates a shorter run penalty at t = 3 of 24%, which seems logical considering the fact that the relative child penalty still evolves over time. Finally, RC5 uses the same event time window (t = 5) but has a notably smaller sample due to the restriction. It estimates a relative child penalty of 25%. The robustness checks indicate that differing event time and birth time windows slightly change the relative child penalty level, ranging from 24% to 31%. Still, the relative child penalty on earnings for mothers is noteworthy, while it is practically non-existent for fathers.

Figure 5.4

Robustness Checks for Child Penalty on Earnings



Note: The figure compares the results from this thesis to the results from different robustness checks by using different birth time and event time windows. The purple coefficients show the relative child penalty (mothers relative to fathers) on earnings. The green coefficients show the child penalty on earnings for fathers.
RC0: Reference category: the results from this thesis.
RC1: Longer run (relative) child penalty at t = 7, including all possible birth years (2008-2022).
RC2: Longer run (relative) child penalty at t = 7, including the same birth years used in the thesis (2009-2021).
RC3: Longer run (relative) child penalty at t = 7, birth years restricted to always being observed (2010-2015).
RC4: Shorter run (relative) child penalty at t = 3, birth years restricted to always being observed (2010-2019).
RC5: Medium-run (relative) child penalty at t = 5, birth years restricted to always being observed (2010-2017).
RC6: Medium-run (relative) child penalty at t = 5, including all possible birth years (2008-2022).
Sample sizes are larger than the original sample for RC1, RC2 and RC6. Sample sizes are smaller than the original sample for RC3, RC4 and RC5.

6. Conclusion and discussion

Having children almost entirely explains the remaining levels of gender inequality. After having a child, the labor market outcomes of fathers are barely impacted while those of mothers worsen significantly. The difference between men and women in their labor market outcomes years after having their first child is called the long-run relative child penalty. Using the event study methodology from Kleven (2019b) and applying it to panel data on Dutch households, I estimated a medium-run relative child penalty 5 years after the birth of the first child of 30% for the entire sample. This is lower than the child penalty found at t = 5 in the Netherlands by other authors of around 38 or 43% (Rabaté & Rellstab, 2021; Artmann et al., 2022). Still, it proves that a relative child penalty on earnings does exist for our Dutch sample.

After conducting a decomposition of the child penalty in the Netherlands, I found a 19% medium-run relative child penalty on work hours. For wage rates, an insignificant (6%) relative child penalty was found. Lastly, for labor force participation, I found a 5% medium-run relative child penalty. The figure for both the child penalty on work hours as well as labor force participation shows a negative trend for men, whereas the trend for men is flat for both the child penalty on earnings and wage rates. The decrease in wage rates is surprisingly small for women, but other authors have found a flatter trend for wage rates as well (Fontenay et al., 2021; Artmann et al., 2022). The insignificance of the relative child penalty on wage rates and the smaller magnitude of the relative child penalty on labor force participation seem to point to the fact that work hours are of the largest importance in determining the relative child penalty on earnings (Rabaté & Rellstab, 2021). However, significantly less information was available on work hours compared to the available data on earnings. Therefore, this analysis should be approached carefully.

After that, I clustered individual sectors together based on multiple characteristics at once such as femininity/masculinity, average education level, average earnings and followed a similar sector classification such as the one Shahrabani (2007) uses in her paper. This created the following five sectors: the healthcare sector, the public sector cluster, the low-services sector cluster, the high-services sector cluster and the products sector cluster. Results did show different medium-run relative child penalties across sector clusters: a child penalty as low as 10% in the public sector cluster, while it had a level of 40% in the healthcare sector. The high-services sector cluster had the second lowest relative child penalty (20%), followed by the low-services sector cluster (29%) and the products sector cluster (39%). This divergence in child penalty levels did not seem to follow a comprehensible pattern: I could not establish a clear correlation between either education levels or average earnings and the respective relative child penalty level. Clusters with lower education levels seem to suffer from larger child penalty levels, but the healthcare sector does not follow this pattern. For average annual earnings, the pattern seems to be non-existent, as clusters with similar average earnings are the most far apart in child penalty levels. This might be caused by the small sample size or the fact that there are more sector differences at play that I did not consider in my analysis. Concluding, I did find different child penalty levels across sectors, but could not establish a clear reason explaining this difference.

Next, I focused on one characteristic at a time. When sectors are clustered together based on femininity/masculinity, I seem to gather evidence for the thought that women working in sectors that are perceived as being more masculine or attract more male workers suffer from a higher relative child penalty (35%) compared to women who work in feminine sectors (26%). This difference was smaller for the relative child penalties for clusters based average earnings. The sector cluster with earnings higher than the sample mean showed a relative child penalty level of 26%, while this was 28% for the cluster where earnings were lower than the sample mean. This might indicate that the average annual earnings of a sector might be a less important characteristic in determining the overall relative child penalty level. The largest divergence between relative child penalty levels based on one characteristic was found for education level: the cluster consisting of sectors with a lower average education level showed a relative child penalty as high as 42%, while this was only 25% for the cluster consisting of sectors with a higher average education level. This is not in line with literature stating that mothers with a higher education level suffer from a higher income penalty or more gender inequality (Anderson et al, 2002; Shahrabani, 2007). However, in literature where the income penalty of mothers is compared to the income penalty of fathers, college-educated mothers are found to suffer from a lower child penalty than non-college educated mothers do (De Quinto et al., 2021). This is in line with the results found in this thesis. The results for the child penalty across education levels can likely be explained by the finding that women who are college-educated react less strongly to childbirth by decreasing working hours/days than those who are not college-educated (De Quinto et al., 2021). Considering that work hours are the main determinant of the child penalty in the Netherlands, it would make sense to see such a stark difference between education levels.

A characteristic that was not previously considered when forming the five sector clusters was whether individuals in the sector work non-standard hours, which impacts the work/life balance of workers. An attempt at mapping out whether the child penalty level is correlated with working non-standard hours was unsuccessful. Two out of five sectors deviated strongly from the predicted values, which could be due to the cluster design used, as the sectors making up those two sector clusters vary in their shares of respondents working non-standard hours.

Therefore, the same method was used as described before where sectors are clustered together based on the share of workers working non-standard hours compared to the overall average share and then computing the relative child penalty for both clusters. A relative child penalty of 25% was found for the cluster of sectors working standard hours, whereas the penalty was 36% for sectors working non-standard hours. These results are in line with those of Fontenay et al. (2021) who found a positive correlation between child penalty levels and working on the weekends or at night. I cannot conclude that one of the characteristics mentioned above will always cause a higher/lower child penalty level, but I do seem to find evidence that some characteristics might carry different weights when determining the relative child penalty level per sector.

A limitation of this paper was the small sample size and limited number of observations, it would be interesting to repeat this analysis for a larger number of observations to obtain a more accurate estimate and possibly conduct the analysis per individual sector instead of using the cluster design. The cluster design was a limitation in the sense that the sectors within a cluster might still differ in several other characteristics, making it harder to dive into the reason why the sectors differ in terms of child penalties. Additionally, a dataset containing more information on work hours or a variable concerning wage rates would make the decomposition of the relative child penalty for the entire sample more accurate. This would allow us to fully confirm whether work hours are indeed the most important determinant for the child penalty in the Netherlands. Next to that, the decomposition could not be performed per sector cluster or even per characteristic as the number of observations was too small to yield insightful results. It would be interesting to see whether the relative child penalty is driven more by different determinants per sector, or if work hours drive the relative child penalty the most in all sectors.

Additionally, it would be insightful to analyze the difference across sectors for different countries. This is because the same sectors might vary between countries due to differences in collective bargaining or union power and different childcare policies per country. Another suggestion for further research would be to include indicators surrounding flexibility such as whether parents are allowed to change work hours and if so, how easily they can do so. In this way, one could research whether the child penalty decreases in the availability of flexible hours, as flexible hours have been found to be an important factor in decreasing the wage gap between mothers and childless women (Fuller & Hirsch, 2019). Finally, it would be interesting to see how the relative child penalty has evolved over time in the Netherlands, especially after the COVID-19 crisis; working from home became more common which might have impacted the child penalty as working from home has been found to impact the work/life balance (Shirmohammadi et al., 2022).

7. References

Amin, M. M., & Islam, A. (2014). Are There More Female Managers in the Retail Sector? Evidence from Survey Data in Developing Countries. *Journal of Applied Economics*, *17*(2), 213–228. https://doi.org/10.1016/s1514-0326(14)60010-6

Anderson, D. J., Binder, M., & Krause, K. (2002). The Motherhood Wage Penalty: Which Mothers Pay It and Why? *The American Economic Review*, *92*(2), 354–358. https://doi.org/10.1257/000282802320191606

Andresen, M., & Nix, E. (2019). What causes the child penalty? Evidence from same sex couples and policy reforms. *Discussion Papers (No. 902).*

Arlinghaus, A., Bohle, P., Iskra-Golec, I., Jansen, N. W. H., Jay, S. M., & Rotenberg, L. (2019). Working Time Society consensus statements: Evidence-based effects of shift work and non-standard working hours on workers, family and community. *Industrial Health*, *57*(2), 184–200. https://doi.org/10.2486/indhealth.sw-4

Artmann, E., Oosterbeek, H., & Van Der Klaauw, B. (2022). Household specialization and the child penalty in the Netherlands. *Labour Economics*, *78*, 102221. https://doi.org/10.1016/j.labeco.2022.102221

Bergmann, N., Scheele, A., & Sorger, C. (2019). Variations of the same? A sectoral analysis of the gender pay gap in Germany and Austria. *Gender, Work and Organization*, *26*(5), 668–687. https://doi.org/10.1111/gwao.12299

Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, *2*(3), 228–255. https://doi.org/10.1257/app.2.3.228

Bosch, N., Deelen, A., & Euwals, R. (2008). Is Part-time Employment Here To Stay? Evidence from the Dutch Labour Force Survey 1992–2005. *IZA Discussion Paper*, *3367*.

Budig, M. J., & England, P. (2001). The Wage Penalty for Motherhood. *American Sociological Review*, *66*(2), 204. https://doi.org/10.2307/2657415

CBS Statline (2022). *Geboorte; kerncijfers.* https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37422ned/table?ts=1603979261501.

Centraal Bureau voor de Statistiek. (2021). *Standaard Onderwijsindeling 2021*. Centraal Bureau Voor De Statistiek. https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs-en-beroepen/standaard-onderwijsindeling--soi--/standaard-onderwijsindeling-2021

Chantreuil, F., Fourrey, K., Lebon, I., & Rebiere, T. (2021). Magnitude and evolution of gender and race contributions to earnings inequality across US regions. *Research in Economics*, *75*(1), 45–59. https://doi.org/10.1016/j.rie.2020.11.001

De Quinto, A., Hospido, L., & Sanz, C. (2021). The child penalty: evidence from Spain. *Series*, *12*(4), 585–606. https://doi.org/10.1007/s13209-021-00241-9

EIGE (2023), *Gender Equality Index,* Publications Office of the European Union, Luxembourg (https://eige.europa.eu/gender-equality-index/2022).

Fontenay, S., Murphy, T. E., & Tojerow, I. (2021). Child penalties across industries: why job characteristics matter. *Applied Economics Letters*, *30*(4), 488–495. https://doi.org/10.1080/13504851.2021.1994518

Fuller, S., & Hirsh, C. E. (2019). “Family-Friendly” Jobs and Motherhood Pay Penalties: The Impact of Flexible Work Arrangements Across the Educational Spectrum. *Work and Occupations*, *46*(1), 3–44. https://doi.org/10.1177/0730888418771116

Kleven, H. J., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019a). Child Penalties across Countries: Evidence and Explanations. *AEA Papers and Proceedings*, *109*, 122–126. https://doi.org/10.1257/pandp.20191078

Kleven, H. J., Landais, C., & Søgaard, J. E. (2019b). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209. https://doi.org/10.1257/app.20180010

Kleven, H. J. (2022). *The Geography of Child Penalties and Gender Norms: Evidence from the United States*. https://doi.org/10.3386/w30176

Knights, D., & Tullberg, M. (2014). Masculinity in the Financial Sector. In *Oxford University Press eBooks*. https://doi.org/10.1093/oxfordhb/9780199658213.013.023

Kroska, A., & Elman, C. (2009). Change in attitudes about employed mothers: Exposure, interests, and gender ideology discrepancies. *Social Science Research*, *38*(2), 366–382. https://doi.org/10.1016/j.ssresearch.2008.12.004

Kumari, K. T., & Devi, V. R. (2016). Work-Life Conflict of Women Employees in Selected Sectors. *Pacific Business Review International*, *9*(6), 73–80.

Lebedinski, L., Perugini, C., & Vladisavljević, M. (2022). Child penalty in Russia: evidence from an event study. *Review of Economics of the Household*, *21*(1), 173–215. https://doi.org/10.1007/s11150-022-09604-y

Miller, P. D. (2009). The Gender Pay Gap in the US: Does Sector Make a Difference? *Journal of Labor Research*, *30*(1), 52–74. https://doi.org/10.1007/s12122-008-9050-5

Nederlands Jeugdinstituut. (2023). *Cijfers over geboorte*. https://www.nji.nl/cijfers/geboorte#:~:text=In%202021%20heeft%20de%20Nederlandse,gemiddelde%20aantal%20kinderen%201%2C54.

Rabaté, S., & Rellstab, S. (2021). The Child Penalty in the Netherlands and its Determinants. *CPB Discussion Paper*, *424,* CPB Netherlands Bureau for Economic Policy Analysis. https://doi.org/10.34932/trkz-qh66

Shahrabani, S. (2007). Sources of gender wage gap in different economic sectors: the Israeli case. *Applied Economics Letters*, *14*(14), 1019–1022. https://doi.org/10.1080/13504850600706602

Shirmohammadi, M., Au, W. C., & Beigi, M. (2022). Antecedents and Outcomes of Work-Life Balance While Working from Home: A Review of the Research Conducted During the COVID-19 Pandemic. *Human Resource Development Review*, *21*(4), 473–516. https://doi.org/10.1177/15344843221125834

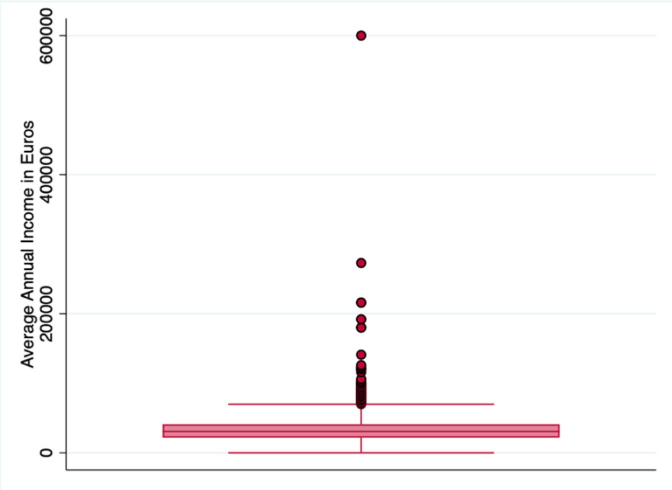
Sieppi, A., & Pehkonen, J. (2019). Parenthood and gender inequality: Population-based evidence on the child penalty in Finland. *Economics Letters*, *182*, 5–9. https://doi.org/10.1016/j.econlet.2019.05.034

Zhou, X., Ye, J., Li, H., & Yu, H. (2022). The rising child penalty in China. *China Economic Review*, *76*, 101869. https://doi.org/10.1016/j.chieco.2022.101869

8. Appendix

Figure A.1

Box Plot of Average Annual Income in Sample



Note: The box plot was made in order to show the presence of possible income outliers in the sample. The highest value of 600 000 euros is more dan double the value of the second highest observation of income of 273 000 euros, both belonging to the same individual (in different years).

Table A.1

The Descriptive Statistics of the Chosen Sample Before Having Their First Child (at t = -1)

	Women	Men	Total
Number of respondents	570	475	1,045
Income in €1 000	27.88 (13.02)	35.22 (13.91)	31.04 (13.88)
Weekly work hours	33.69 (13.01)	39.23 (12.10)	33.25 (13.63)
Age when having a child	30.32 (4.26)	32.79 (5.20)	31.45 (4.87)
Number of children at t = 5	1.60 (0.63)	1.61 (0.59)	1.60 (0.61)
Education level			
Poorly educated	0.04	0.10	0.07
Relatively educated	0.34	0.37	0.35
Highly educated	0.62	0.53	0.58
Sector			
Agriculture	0.01	0.02	0.01
Industrial production	0.03	0.11	0.07
Utilities production	0.01	0.02	0.02
Construction	0.01	0.09	0.04
Retail trade	0.07	0.10	0.09
Catering	0.05	0.04	0.05
Transport/storage	0.02	0.09	0.05
Financial	0.06	0.06	0.06
Business services	0.08	0.10	0.09
Government services	0.06	0.06	0.06

Education	0.14	0.07	0.11
Healthcare/welfare	0.33	0.08	0.22
Environmental services	0.03	0.02	0.03
Other	0.10	0.13	0.12

Note: The mean values of several variables are given above. Standard errors are reported in parentheses, the standard errors of the different categories of the categorical variables are not reported to make the table more organized. I look at the average values of these characteristics of parents at t=-1, right before they have their first child, with the number of children at t=5 being an exception.

Table A.2

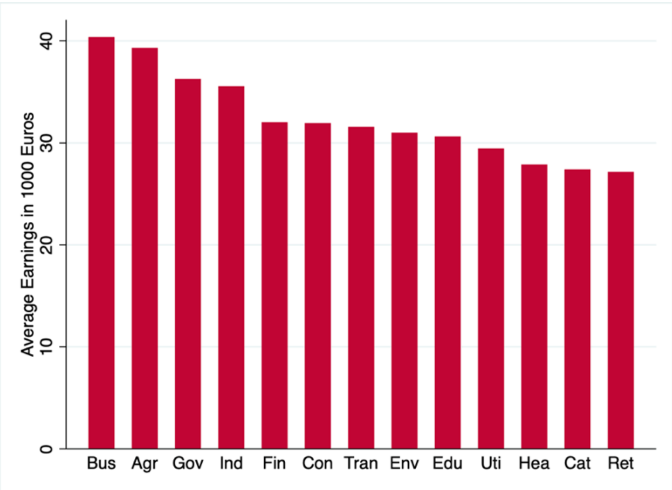
Average Education Level of Respondents per Sector in the Sample

Sector	Poorly Educated	Relatively Educated	Highly Educated
Agriculture	0.00	0.46	0.54
Industrial production	0.07	0.36	0.57
Utilities production	0.06	0.38	0.56
Construction	0.17	0.40	0.43
Retail trade	0.08	0.40	0.52
Catering	0.12	0.45	0.43
Transport/storage	0.20	0.45	0.35
Financial	0.02	0.46	0.52
Business services	0.01	0.22	0.77
Government services	0.05	0.34	0.62
Education	0.02	0.10	0.88
Healthcare/welfare	0.05	0.33	0.62
Environmental services	0.10	0.31	0.59

Note: This table shows the average mean values of respondents per education level per sector. This table was made to group sectors together based on different characteristics such as education level. Mean values are rounded to two decimals.

Figure A.3

Average Annual Earnings per Sector



Note: Average annual earnings per sector ranked from highest to lowest earnings. Earnings is defined per 1000 Euros.

Table A.2

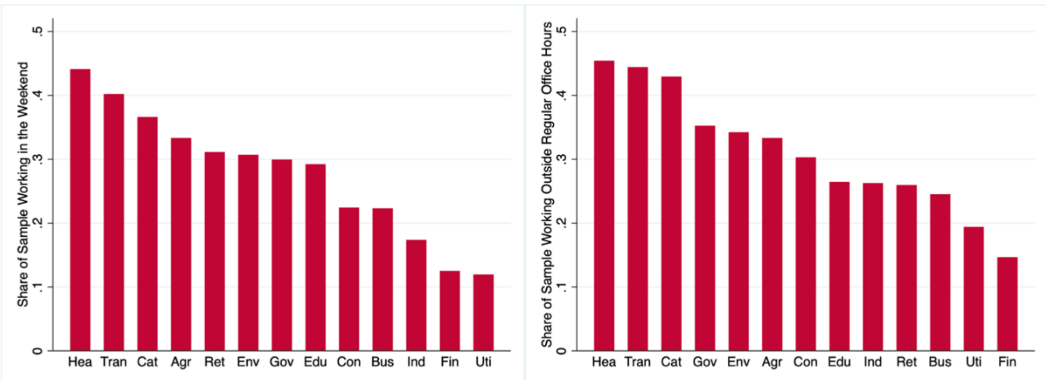
Balance Table between Men and Women in the Chosen Sample (at t = -1)

Factor	Level	Male	Female	P-value	Test
Number of respondents		475	570		
Income in €1 000, median (IQR)		33.6 (26.922, 41.382)	28.036 (20.91, 35.124)	<0.001***	Wilcoxon rank-sum
Weekly work hours, median (IQR)		40 (38, 45)	36 (32, 40)	<0.001***	Wilcoxon rank-sum
Age when having a child, median (IQR)		32 (29, 36)	30 (27, 33)	<0.001***	Wilcoxon rank-sum
Education level	Poorly educated	47 (9.9%)	24 (4.2%)	<0.001***	Fisher's exact
	Relatively educated	174 (36.7%)	192 (33.8%)		
	Highly educated	253 (53.4%)	352 (62.0%)		
Sectors clustered together	Public Sector Cluster	99 (24.0%)	302 (59.1%)	<0.001***	Fisher's exact
	High-Services Sector Cluster	77 (18.7%)	77 (15.1%)		
	Low-Services Sector Cluster	124 (30.1%)	100 (19.6%)		
	Products Sector Cluster	112 (27.2%)	32 (6.3%)		
Masculine/feminine sectors	Masculine	233 (56.6%)	120 (23.5%)	<0.001***	Fisher's exact
	Feminine	179 (43.4%)	391 (76.5%)		

Note: Balance table between men and women. Significance is shown in stars: ***p<0.01; **p<0.05; *p<0.1. The IQR is opted for over the mean, as the continuous variables did not show a normal distribution.

Figure A.4

Fraction of Respondents per Sector Working in the Weekend and Outside Regular Office Hours



Note: The share of respondents who reported working outside regular office hours (after 6pm and before 7am) is calculated by dividing the number of respondents who answered “Yes, I work in shifts.” or “Yes, I (almost) always work in the evening or at night.” or “Yes, I often work outside regular office hours.” by the total number of respondents working in the sector who answered the question. The share of respondents who reported working in the weekend is calculated by dividing the number of respondents who answered “I work during the weekend once every few weeks.” or “I work during the weekend almost every week.” by the total number of respondents working in the sector who answered the question. The y-axis for both bar graphs ranges from 0 to 0.5.

Table A.3

Regression Coefficients for Main Results and Decomposition

Event Time	Earnings		Work Hours		Wage Rates		Labor Force Participation	
	Men	Women	Men	Women	Men	Women	Men	Women
t = -2	26.87	455.08	-1.73	-0.01	-0.07	0.10	-0.05*	0.01
	(649.54)	(451.09)	(1.24)	(0.91)	(0.56)	(0.49)	(0.02)	(0.02)
t = 0	1249	-1609.95***	0.38	-1.36	0.10	0.24	-0.00	-0.00
	(590.83)	(451.95)	(1.02)	(0.93)	(0.54)	(0.56)	(0.02)	(0.02)
t = 1	1616	-5316.37***	-1.95	-5.48***	0.49	0.27	-0.03	-0.04*
	(920.89)	(638.47)	(1.20)	(0.97)	(0.65)	(0.59)	(0.02)	(0.02)
t = 2	1069	-7342.07***	-2.37	-8.28***	0.40	0.18	-0.05*	-0.06**
	(1136.14)	(790.63)	(1.32)	(1.10)	(0.78)	(0.69)	(0.02)	(0.02)
t = 3	510.83	-8953.23***	-2.43	-9.80***	0.47	0.06	-0.04	-0.07**
	(1464.05)	(988.97)	(1.30)	(1.20)	(0.88)	(0.75)	(0.02)	(0.02)
t = 4	55.00	-10443.88***	-3.52*	-8.94***	0.47	-0.24	-0.05*	-0.09***
	(1624.15)	(1169.21)	(1.53)	(1.30)	(0.98)	(0.87)	(0.02)	(0.03)
t = 5	202.00	-11217.62***	-2.53	-10.63***	0.67	-0.80	-0.04	-0.08**
	(1890.40)	(1399.18)	(1.53)	(1.42)	(1.04)	(0.89)	(0.02)	(0.03)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2 376	2 878	1 713	1 961	1 575	1 819	1 713	1 961

Note: The table shows the regression coefficients of the regression from equation (1), presented in the methodology section. The dependent variable differs per column: earnings, work hours, wage rates and labor force participation. FE stands for fixed effects since all models control for age and year fixed effects. Standard errors are clustered at the individual level and shown in parentheses. Significance is shown in stars: *p<0.05; **p<0.01; ***p<0.001.

Table A.4

Regression Coefficients of Women Only for Different Clusters

Event Time	Healthcare	Public	High-Services	Low-Services	Products
t = -2	567.77	-294.68	1954.23	730.66	-599.60
	(773.70)	(1161.28)	(1400.04)	(945.18)	(2718.22)
t = 0	-1679.72*	-1541.51	223.18	-692.35	-441.92
	(795.31)	(886.91)	(1535.48)	(1354.23)	(2443.92)
t = 1	-5194.99***	-3629.58**	-4024.99*	-4135.65*	-3933.89
	(1124.33)	(1211.01)	(1959.96)	(1706.35)	(3422.40)
t = 2	-8284.45**	-3740.87*	-6051.11*	-7100.56***	-6713.44
	(1378.42)	(1878.39)	(2467.64)	(1754.65)	(4310.99)
t = 3	-10382.98***	-5557.64*	-6369.20	-7729.64**	-8019.48
	(1627.83)	(2249.04)	(3792.65)	(2280.12)	(5235.20)
t = 4	-11075.61***	-7442.99**	-9828.21*	-9252.59***	-11055.82*
	(1897.53)	(2325.19)	(4244.93)	(2509.84)	(5816.69)
t = 5	-11009.64***	-8038.51**	-11904.04*	-12202.46***	-9587.83*
	(2258.21)	(2589.65)	(5108.93)	(3002.51)	(5783.40)

Age FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	956	569	413	468	204

Note: The table shows the regression coefficients for women only, the dependent variable is earnings for all columns. FE stands for fixed effects since all models control for age and year fixed effects. Standard errors are clustered at the individual level and shown in parentheses. Significance is shown in stars: *p<0.05; **p<0.01; ***p<0.001.

Table A.5

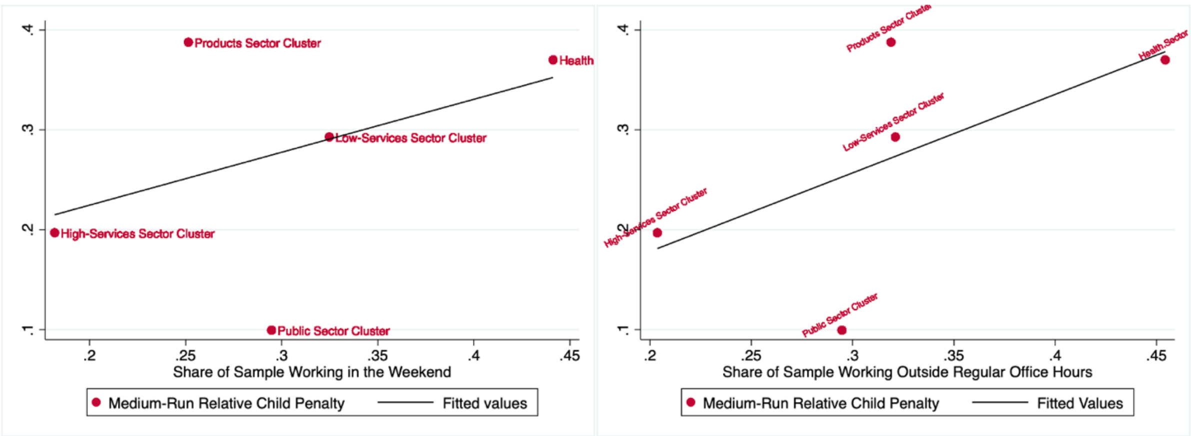
Regression Coefficients of Women Only for Different Characteristics

Event Time	Dominant Gender		Average Earnings		Education Level		Working (Non-)Standard Hours	
	Masculine	Feminine	Low	High	Low	High	Standard	Non-Standard
t = -2	626.23 (962.12)	317.43 (522.81)	344.87 (535.88)	439.42 (820.93)	468.62 (826.63)	515.16 (550.43)	763.47 (841.38)	328.94 (588.24)
t = 0	-755.53 (1157.52)	-1947.70*** (537.44)	-1991.89*** (561.63)	-460.28 (1028.74)	-1826.50 (1126.50)	-1291.03* (511.59)	-1406.12 (860.73)	-1667.37*** (605.46)
t = 1	-4549.97** (1682.36)	-5229.29*** (709.39)	-5459.23*** (728.95)	-3612.84* (1484.06)	-5753.86*** (1491.98)	-4693.62*** (722.91)	-5338.00*** (1102.81)	-4954.85*** (869.46)
t = 2	-6890.71*** (1968.49)	-7076.33*** (896.48)	-7585.54*** (905.82)	-4827.59** (1738.03)	-8406.98*** (1794.49)	-6748.84*** (923.09)	-7299.99*** (1272.15)	-7113.63*** (1077.08)
t = 3	-8149.33** (2785.67)	-8860.61*** (1085.99)	-9494.88*** (1107.33)	-5903.26* (2311.83)	-9810.73*** (2250.53)	-8429.80*** (1159.47)	-9415.53*** (1643.42)	-8257.41*** (1287.51)
t = 4	-10745.28*** (3268.83)	-10117.36*** (1236.52)	-10739.00*** (1271.66)	-8608.83** (2704.35)	-10900.93*** (2585.90)	-10194.92*** (1372.87)	-11512.14*** (1995.04)	-9477.80*** (1483.74)
t = 5	-12459.12** (3957.70)	-10498.57*** (1444.24)	-10938.40*** (1484.32)	-10509.97** (3316.98)	-12162.61*** (2972.30)	-11082.38*** (1655.14)	-11932.44*** (2382.83)	-10309.64*** (1752.52)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	617	1 993	1 858	752	611	1 999	1 167	1 443

Note: The table shows the regression coefficients for women only, the dependent variable is earnings for all columns. FE stands for fixed effects since all models control for age and year fixed effects. Standard errors are clustered at the individual level and shown in parentheses. Significance is shown in stars: *p<0.05; **p<0.01; ***p<0.001.

Figure A.5

Scatterplot of Non-Standard Working Hours and Medium-Run Relative Child Penalty Levels



Note: The relative child penalties per sector cluster are plotted against the share of respondents in the sample who reported working either in the weekend or outside regular office hours (after 6pm and before 7am). The share of respondents who reported working outside regular office hours (after 6pm and before 7am) is calculated by dividing the number of respondents who answered “Yes, I work in shifts.” or “Yes, I (almost) always work in the evening or at night.” or “Yes, I often work outside regular office hours.” by the total number of respondents working in the sector who answered the question. The share of respondents who reported working in the weekend is calculated by dividing the number of respondents who answered “I work during the weekend once every few weeks.” or “I work during the weekend almost every week.” by the total number of respondents working in the sector who answered the question. The fitted values are the estimated responses, shown as a black line.