Bachelor Thesis BSc^2 Econometrics & Economics

Health, labor and motherhood

Investigating career and health effects of motherhood in Germany.

Name student: Laura Zwiers

Student ID number: 468008

Date: July 1st, 2021

Supervisor: Dr. Pilar García-Gómez

Second Assessor: Dr. Mikhail Zhelonkin

Abstract

Over the past 50 years, gender wage inequalities have decreased substantially. However, women still earn less than men, and a large reason for this is the persistent child penalty, which implies that having a child affects women's careers more than men's careers. Potential explanations for child penalties are biological factors and gender norms, and family policy schemes are found to have different effects on penalties. Besides, previous research suggests different mechanisms through which mothers' employment decisions could affect maternal health, and that different characteristics and policies play a role in this framework. This thesis studies child penalties and maternal health in Germany, which is a country suggested to have traditional gender norms and relatively generous parental leave benefits. Using an event study regression, Germany's child penalty is estimated to be around 67% nine years after childbirth. Findings suggest that this penalty is mainly caused by women working less hours after giving birth, and that having parental leave reserved for fathers reduces penalties. An event study investigating maternal health does not find any effect of childbirth on health, but results from Lasso regressions suggest that working less relates to lower long-term health satisfaction, and that some pre-birth characteristics could function as predictors of health. However, reforms in parental leave benefits are not found to affect health outcomes of mothers.

Contents

1 Introduction

Until the 1950s, the marriage bar in the United States made it common practice to fire women when they got married, or to not hire married women (Goldin, 1988). The Netherlands had similar policies in place, where women working for the government were fired right after getting married. Such policies of gender discrimination were reflected in wages, with women in the US receiving wages equal to only 60% of the wages males received in the year 1960. Since then, the gender pay gap has decreased substantially, with the average wage of women equal to 80% of the wage of men in 2011 (Blau & Kahn, 2017).

Despite this improvement, gender wage inequalities remain persistent. Research has sought for explanations in human capital differences and discrimination (Altonji & Blank, 1999), but with women becoming equally educated as men and many anti-discriminatory policies in place, it seems that something else is causing the wage gap. Kleven, Landais, and Søgaard (2019) argue that one of the main drivers of the gender wage gap is the fact that having children affects women's careers much more than men's careers. This is referred to as the child penalty.

Child penalties differ across countries, with the height of penalties often being related to a country's gender norms and family policies. Gender norms are argued to cause child penalties, as they typically make mothers take on the main care-taking tasks within a family (Blair-Loy, 2009; Townsend, 2010). Besides, Kleven, Landais, Posch, Steinhauer, and Zweimüller (2019) show that women's labor market decisions are driven by what they experience in their childhood and Becker (1985) argues that the choice to work in more family-friendly jobs after giving birth might go at the cost of higher wages. With respect to family policies, Andresen and Nix (2019) find that parental leave does not necessarily reduce penalties, while affordable childcare does.

The first part of this thesis studies child penalties in Germany, which is an interesting country because of the rich data availability, as well as the country's gender norms and family policies. Specifically, gender norms in Germany are suggested to be relatively traditional (Kleven, Landais, Posch, et al., 2019), and at the same time the country has a rich history of parental leave programs. Paid parental leave for both parents was introduced in 1986, but payments were far lower than actual wages, which made that fathers often did not use it. A reform was introduced in 2007, which shortened the leave period from 36 to 14 months and increased payments. Moreover, two months of leave are reserved specifically for fathers (Geisler & Kreyenfeld, 2012).

Child penalties in Germany are estimated under different conditions, in order to potentially get insights into the mechanisms behind them. The first research question therefore is

> *To what extent do child penalties exist in Germany and how do different mechanisms affect these?*

This research question is answered by using an event study around the birth of the first child, as proposed by Kleven, Landais, Posch, et al. (2019). Defining the child penalty as the percentage by which mothers fall behind fathers in terms of wages, it is estimated to be around 67% nine years after childbirth. Results also suggest that this is mainly due to German women starting to work less after giving birth and that penalties decreased after the 2007 parental leave reform.

The second part of this thesis focuses on maternal health, and whether this is related to career effects of motherhood. Some research suggests that working mothers experience worse health due to the levels of stress and fatigue that are caused by combining work with childcare (Cunningham-Burley, Backett-Milburn, & Kemmer, 2006). On the other hand, unemployment is often related to worse health outcomes (Brand, 2015). Hence, it is of interest to look into health outcomes of both working as non-working mothers, leading to the second research question:

> *Does having a child affect health and how do career effects of motherhood relate to long-term health outcomes?*

Different characteristics of mothers may be related to health outcomes as well as labor market decisions. Identifying which characteristics relate to bad health outcomes creates the possibility of targeting certain health-improving policies towards those who need them most. Besides, it is of interest to see whether large income losses are related to similar characteristics as large declines (or increases) in health. One sub-question of the second research question therefore is the following:

1. *Which characteristics of mothers relate to bad health outcomes of mothers, and are these characteristics different for mothers with different labor force participation?*

As family policies have often been linked to career effects of motherhood, it is interesting to investigate whether reforms affect health outcomes. Some existing research suggests that the introduction of paid maternity leave programs improves mothers' long-term well-being (Avendano, Berkman, Brugiavini, & Pasini, 2015; Bütikofer, Riise, & Skira, 2021), and this thesis assesses whether the same holds for German women. Hence the second sub-question is

2. *How does paid parental leave affect long-term health outcomes of mothers?*

The second research question and its sub-questions are answered using the Least Absolute Shrinkage Operator (Lasso) and its Bayesian counterpart. This method has the ability to outperform simpler regression methods, by having higher predictive power and yielding models that are easy to interpret. Thereby, using Lasso creates the possibility of identifying which mothers may suffer from bad health and how parental leave affects health, which is valuable for policy makers.

Childbirth in itself is not found to affect health, but results from Lasso regressions suggest that some mechanisms are of interest. For instance, a decrease in annual work hours is related to worse long-term health satisfaction, but a decrease in earnings is not. Some pre-birth characteristics, such as age and educational level, are also found to relate to health outcomes. I do not find any significant effect of the parental leave reform on maternal health outcomes.

This research contributes to existing literature in several ways. First, it is the first paper to examine both the career effects of motherhood as maternal health in Germany. The health effects of family policy reforms in Germany have also not been studied extensively yet. Moreover, by using Lasso regressions for investigating maternal health, new insights into the topic might be provided.

In the remainder of this thesis, I first discuss findings of earlier research on the topic of child penalties and maternal health, as well as research on relevant econometric methods. Afterwards, the dataset is introduced and the steps taken in sample selection and pre-processing are described. In the methodology, the econometric methods used are explained, after which results are presented. Finally, the conclusion provides answers to research questions and an extensive discussion of them.

2 Literature Review

2.1 Child penalties

The most common way of investigating the effect of becoming a mother on wages is by estimating the child penalty, or family gap, which is the negative effect that motherhood has on wages (Felfe, 2006). Specifically, Kleven, Landais, and Søgaard (2019) define the child penalty as the percentage by which women fall behind men due to having children. Waldfogel (1997) investigates penalties in the United States and finds that, even after controlling for part-time employment, mothers with one child experience a penalty of four percent, and those with two children even of twelve percent. Budig and England (2001) estimate the wage penalty to be around five percent after controlling for experience. Moreover, they find that penalties are higher for married women compared to unmarried women. When linking educational level to child penalties, Anderson, Binder, and Krause (2003) argue that penalties do not exist for the least-educated mothers, while those with a high-school or college diploma earn 10 percent less per child compared to non-mothers.

Compared to the United States, family gaps in the United Kingdom and Germany are larger (Gangl & Ziefle, 2009). In fact, while wage penalties in the United States are estimated to be around seven percent, those in Britain in Germany are found to be nine and 15 percent, respectively. Kleven, Landais, Posch, et al. (2019) compare six different countries and show that long-run child penalties are between 21 and 27% in Sweden and Denmark, while these are 31 to 44% in the United States and the United Kingdom. For Germany and Austria, long-run penalties are even estimated to be as high as 51 to 61 percent. Using the same estimation methodology, Rabaté and Rellstab (2021) calculate the Dutch child penalty to be 46% seven years after childbirth.

2.1.1 Potential explanations

Since mothers are the ones bearing children, they almost always spend some time out of work around childbirth, due to the physical impact that childbearing brings and the fact that only women have the option to breastfeed. Human capital theory (Becker, 1985) suggests that loss of labor experience due to childbearing is likely to explain most of the wage gap. Adding to this, Anderson et al. (2003) argues that loss of experience has different effects for mothers of different educational levels and specifically, that higher-educated mothers suffer from larger income drops due to time out of work. The findings of Gangl and Ziefle (2009) underline this possible explanation by linking family policies to child penalties. Specifically, they argue that child penalties in Germany are higher than those in the United States and Great-Britain because of the paid parental leave scheme Germany has.

The idea that childbearing is the main driver behind child penalties is contradicted by Kleven, Landais, and Søgaard (2020). When comparing wage penalties between biological and adoptive families in Denmark, they find that long-run child penalties are equally large for adoptive mothers as they are for biological mothers. Investigating wage penalties in same-sex couples in Norway, Andresen and Nix (2019) find that there exists a short-run penalty for the biological mother, but that the penalty is equally large for both mothers two years after childbirth. Within five years after childbirth, the wage penalty even disappears for these couples. This suggests that child penalties are, at least partly, explained by preferences with respect to the mother working.

Preferences with respect to working mothers are often driven by gender norms. Research suggest that many people still think that mothers should, above all, care for their children, while a good father is often argued to also be a hard-working employee (Blair-Loy, 2009; Townsend, 2010). This belief is confirmed by Cuddy, Fiske, and Glick (2004), who find that people evaluate women to be less competent when they are known to be mothers. Besides, they find evidence for statistical discrimination against mothers in hiring decisions. Results from a laboratory experiment by Correll, Benard, and Paik (2007) suggest that undergraduate students judge mothers to be less competent and less committed than non-mothers. Moreover, an audit study among actual employers shows that mothers receive less callbacks than non-mothers with similar résumés. In addition, more traditional gender norms within a country relate to higher child penalties (Kleven, Landais, Posch, et al., 2019) and women with more traditional views on gender roles experience larger income losses after childbirth (Feldhoff, 2021; Rabaté & Rellstab, 2021).

Aside from gender norms, childhood experiences affect mothers' preferences as well. Fernández, Fogli, and Olivetti (2004) argue that the labor market history of a man's mother affects the decision of his wife with respect to working after childbirth. On the other hand, Kleven, Landais, and Søgaard (2019) suggest that the labor market history of women's own mothers is more influential. It can thus be argued that predominant beliefs and preferences within one's social environment affect preferences with respect to being a working mother. Besides, the tendency of mothers to bear the main childcare tasks could even lead to child penalties for women that stay employed, as they might opt for jobs that are more "child-friendly" at the cost of higher wages (Becker, 1985). Becker's "work-effort" hypothesis even argues that mothers may be less productive at work because they are saving energy for childcare after work.

Gender norms in Germany are relatively traditional, with data from the International Social Study Program (ISSP) suggesting that around 35% of Germans believe that women with children under school age should stay at home (ISSP Research Group, 2016). Compared to several other countries, this is a particularly high percentage (Kleven, Landais, Posch, et al., 2019). On the other hand, Germany's family policy system protects women against job loss after giving birth and allows them to spend time with their children before returning to the labor force. A downside of this system is that it might cause employers to discriminate against potential mothers upfront, leaving them with lower wages anyway (Gangl & Ziefle, 2009). Besides, Andresen and Nix (2019) argue that paternity leave take-up is low and does not reduce child penalties in Norway, while more affordable childcare does. In the Netherlands, on the other hand, a reform making childcare more affordable only led to a slight decrease in penalties (Rabaté & Rellstab, 2021). Penalties in different countries thus react differently to family policy reforms.

2.2 Health of working mothers

Short- and long-run maternal health has extensively been studied. Research on short-run health often focuses on postpartum depressions, as well as physical discomfort that comes with giving birth and caring for a baby. In the longer run, health effects of motherhood can be linked to labor market behavior of mothers. Klumb and Lampert (2004) evaluate findings of a set of papers investigating the relationship between employment and health for females and find that marital and family status affect the magnitude of this relationship in several cases. For instance, Waldron, Weiss, and Hughes (1998) conclude that health benefits of employment decrease as the number of young children in the household increases. This is in line with findings from a survey study by Cunningham-Burley et al. (2006). They argue that, as working mothers often still take on the role of main caretaker, they suffer from higher levels of stress and fatigue. Moreover, Grice et al. (2007) find that women whose jobs and family life interfere with each other report lower mental health.

On the other hand, unemployment is often suggested to cause a decline in both physical as mental health status (see Brand (2015) for a review). Explanations for this are mostly sought in the fact that the economic consequences of unemployment cause stress, and that job loss also means a loss of social contact. However, results are often contested due to possible selection bias. That is, it is not unlikely that people that lose their jobs already have worse health upfront.

The idea that poor health leads to job loss, and not vice versa, is argued for in several papers. Schmitz (2011) investigates a sample of German individuals and, although the unemployed are less healthy on all aspects, argues that these people lose their jobs because of their health status. This is in line with García-Gómez, Jones, and Rice (2010), who argue that health status is a determinant of employment transitions. Olesen, Butterworth, Leach, Kelaher, and Pirkis (2013) focus on mental health and conclude that this aspect of health is both a cause as a consequence of job loss.

This reversed causality is of interest in the study of maternal health, as women often experience worse health after giving birth. For instance, more than half of mothers experience health problems one year after childbirth and this makes them feel guilty for not living up to their own expectations of motherhood (Saurel-Cubizolles, Romito, Lelong, & Ancel, 2000). Moreover, mothers are at higher risk of breast cancer compared to non-mothers (Lambe et al., 1994). These negative health effects of motherhood could lead to selection out of the labor force. However, there is also evidence that women whose mental health decreases are more likely to return to work, potentially because they enjoy their jobs too much (García-Gómez et al., 2010).

2.2.1 The role of family policies

As discussed in the introduction, Germany has a rich history of parental leave benefits and a reform was introduced in 2007. This is of interest because many researchers have looked into the relationship between family policies and health outcomes. Avendano et al. (2015) examine whether maternity leave benefits affect women's later-life mental health, and find that depression occurs much less often for women who received extensive benefits after the birth of their first child. In the short run, physical health also increases with the length of maternity leave (Dagher, McGovern, & Dowd, 2014). Recent research by Bütikofer et al. (2021) examines the impact of the introduction of paid maternity leave in Norway and finds that it improved a range of maternal health outcomes. For instance, mothers benefiting from the reform reported healthier BMI levels, better mental health and were more likely to engage in health-promoting behaviors. Finally, the paper finds that returns are diminishing to leave length, hence extending the leave might not have much impact.

2.3 Econometric methods

This section discusses the literature underlying the different econometric methods used for estimation in this thesis. Technical details of these methods are discussed in Section 4.

2.3.1 Child penalty estimation

Ideally, the impact of children on individuals' careers would be estimated through a randomized experiment. Since it is not possible to randomly assign fertility to individuals, different methods for estimation have been employed. Most of the papers discussed in Section 2.1 use fairly similar methods for estimating child penalties, of which the most standard is fixed-effects regression with added control variables (Anderson et al., 2003; Budig & England, 2001; Gangl & Ziefle, 2009; Waldfogel, 1997). In addition, both Waldfogel (1997) and Anderson et al. (2003) use simple log regression models in their estimation procedure. A completely different approach is to estimate penalties using matching, as is done by Feldhoff (2021) for German mothers.

Kleven, Landais, and Søgaard (2019) use a different method, namely the event study, and this method is also used in this thesis. In event studies, the time of treatment is known to the researcher and treatment cannot be reversed. Treatment is defined as the birth of the first child in this research, and it is obvious that one cannot reverse this event. The advantage of using event studies is that it creates the possibility of estimating dynamic effects over time.

2.3.2 Estimating maternal health

When looking into health effects of motherhood, researchers often opt for performing estimation using simple Ordinary Least Squares (OLS) regressions (Grice et al., 2007; Waldron et al., 1998). Some research even omits the use of econometric models and conducts qualitative research instead (Cunningham-Burley et al., 2006). OLS is a perfectly suitable method for estimating these effects, but it has its downsides (Tibshirani, 1996). Firstly, the prediction accuracy of OLS is not very high because of high variance, especially when many predictors are included. Secondly, because OLS often results in a high number of significant predictors, interpreting results can become difficult.

Two standard techniques for overcoming the downsides of OLS are subset selection and ridge regression. The idea behind subset selection, as described in Miller (2002), is that including a large number of features in the model always increases variance, but does not always decrease bias. Subset selection therefore involves finding a set of features that do not add much to the predictive power of the model and to exclude these. Ridge regression, as introduced by Hoerl and Kennard (1970), decreases all coefficients by a certain factor, which is a form of regularization. By shrinking coefficients, variance is reduced, but no features are eliminated.

Both subset selection and ridge regression only solve one of the two issues of OLS. Subset selection yields models that are easy to interpret, but it does not necessarily increase predictive power. Shrinking coefficients in ridge regression makes for lower variance and higher stability, but since no coefficients are set to zero, interpreting results is still difficult. A method that retains the good features of both models but overcomes the downsides is Lasso regression (Tibshirani, 1996). In Lasso regression, all coefficients are shrunk by the same absolute term, resulting in lower coefficients and hence lower variance. It also decreases coefficients to zero, which is useful for interpretation.

Ordinary Lasso represents a so-called “frequentist” approach, which assumes that estimates are fixed and that they describe the true distribution of the data. Bayesian statistics, which originates from the well-known Bayes’ theorem (Bayes, 1763), proposes to model parameters as random variables. Doing this for Lasso coefficients allows for calculating confidence intervals, while there yet has to be found an good way of doing this in ordinary Lasso (Casella, Ghosh, Gill, & Kyung, 2010). Tibshirani (1996) introduced the idea that Lasso coefficients could be interpreted as Bayesian estimates, and Park and Casella (2008) derive a full Bayesian Lasso specification. Bayesian Lasso results are similar to those from ordinary Lasso, but they are useful in improving stability of estimates.

3 Data

3.1 Data source and sample selection

The data used in this thesis is a representative, multi-cohort survey from the German Socio-Economic Panel (SOEP) (Socio-Economic Panel (SOEP), 2021). The data come from a longitudinal survey study of individuals living in German households conducted between 1984 and 2019. The dataset includes information on labor supply, income, childbirth and health and can be requested through the Research Data Center of the German SOEP. The steps taken in sample selection are different for different parts of this thesis, which are discussed below.

Child penalty estimation Kleven, Landais, Posch, et al. (2019) propose to estimate child penalties using a sample of individuals who have their first child between the age of 20 and 45 and who are observed at least five years prior to childbirth and ten years after. In order to not end up with a too small sample, the latter restriction is relaxed. That is, individuals are included if they are observed at least eight times in the period between five years prior to childbirth and ten years after. Moreover, labor earnings should be observed for the year prior to childbirth as this is considered the baseline year, and at least once between that year and five years before. Imposing this set of restrictions leads to a sample of 3960 individuals, of which 2189 are females and 1771 are males.

Health effects of motherhood For the Lasso regressions concerning health and earnings effects of motherhood, sample selection is different, as it suffices to have only one baseline and one "long-run" observation for each individual. That is, it is only needed to have information for the baseline year and for a certain number of years after childbirth. Specifically, I have chosen to perform the regressions both for females observed five years after childbirth as for females that are observed in the eighth year after giving birth. After matching observations from the baseline year to these two "end" years and deleting missing values, two relatively small samples remain. In particular, 756 females are observed both the year before childbirth and five years after, and 599 females are observed eight years after childbirth.

A downside of these new samples is that some health-related variables are not available for relatively many individuals in them, and deleting observations for which these variables are missing would result in even smaller samples. Luckily, there are little missing values reported for self-rated health satisfaction, which makes it possible to study health based on this outcome variable.

Policy reform evaluation For estimating the effects of the German policy reform in 2007 on health outcomes, the sub-sample described above is reduced to only include females who give birth within a relatively small number of years around the time of the reform. More specifically, the control group consists of women having their first child between 2000 and 2006, and the treatment sample includes women who gave birth between 2007 and 2012.

In order to not end up with a too small sample, all women for whom there is information at the year before childbirth and either five, six or seven years after are included. In the final estimation procedure, the outcomes for the last available year are used. The resulting number of observations in the control group is 239, and there are 124 females in the treated sample.

3.2 Variable definitions

The variables used in this thesis are defined as in the list below. Definitions are based on the definitions by the Socio-Economic Panel (SOEP) (2021).

Survey year	Year in which a survey is conducted.
Labor earnings	Wages and salaries from all employment, including bonuses, profit-sharing and overtime. Earnings are reported for the year prior to the survey year.
Age	Age of individual at the time of survey.
Event time	Time (in years) with respect to birth of first child.
Birth year	Year in which the first child of an individual is born, corresponds to event time 0.
Age birth	Age at which the individual got their first child.
Employment	Binary indicator of whether someone is employed (1) or not (0).
Work hours	Total number of hours worked, reported for the year prior to the survey year.
Employment level	Discrete variable indicating whether an individual worked full-time (1), part-time (2) or did not work (3) in the previous year.
Frequency of sports	Indicates how often a person exercises. Measured as a discrete variable with four levels, with 1 indicating no exercise and 4 exercise on a weekly basis.
Health satisfaction	Subjective health satisfaction of the individual at the time of survey on a scale from 0 to 10.
Marital status	Discrete variable indicating the marital status of an individual. The levels are Married or living with partner (1), single (2), widowed (3), divorced (4) or separated (5).
Years of education	Total number of years of education completed at the time of survey.

3.3 Data pre-processing

Several variables, including labor earnings, are reported in survey year s for the year $s-1$, and hence pre-processing is needed to link these variables to the correct year. The pre-processing step taken to account for this involves creating new numeric variables for event time, age and survey year, which are one lower at each observation. This is done after the first part of sample selection, hence labor earnings are actually observed between six years before and nine years after childbirth.

For the study of health effects, pre-processing involves changing the data from long to wide format. This implies that, instead of having one observation per individual per year, the data frame is reshaped such that there is one row per individual included. Besides, some new variables are created, which indicate changes in relevant variables between the base and the end observation. The base year corresponds to the year prior to childbirth, and the end year is the year for which the observed health outcome is used in estimation. The generated variables are the following:

ΔLabor earnings	Earnings in the end year minus earnings in the baseline year.
ΔWork hours	Work hours in the end year minus work hours in the baseline year.
Separated	Dummy indicating whether someone separated from thier spouse after childbirth. This variable is equal to 1 if an individual is married in the baseline year and not married in the end year.
Stop full time	Dummy indicating whether someone quit working full-time and started working part-time or not working between the baseline year and the end year.

4 Methods

4.1 Child penalty estimation

The impact of having a child on labor market outcomes is estimated using an event study for both men and women. Using notation from Kleven, Landais, and Søgaard (2019), the outcome of interest is denoted Y^g_{ist} and represents the gross labor earnings of individual i in year s, excluding taxes or transfers. The gender g is denoted w for females and m for males. Moreover, t is the event time and $t = 0$ in the year of childbirth. The resulting event study regression is

$$Y^g_{ist} = \sum_{j \neq -1} \alpha^g_j \cdot \mathbf{I}[j = t] + \sum_k \beta^g_k \cdot \mathbf{I}[k = age_{is}] + \sum_y \gamma^g_y \cdot \mathbf{I}[y = s] + \nu^g_{ist} \tag{1}$$

Here, the indicator functions $\mathbf{I}$ equal 1 if the condition in the brackets is true, and 0 otherwise. The second term in the regression controls for age, which is controlled for since women tend to be younger than men when having their first child. The third term controls for the year of childbirth. Note that coefficients for these controls are gender- and age- or year-specific, with the subscript k relating to age and y to year. The event dummy for t = -1 is excluded from the summation, such that event time coefficients measure the impact of children relative to the year just before the child is born. To measure the extent to which women fall behind men due to having children, the child penalty at event time t is estimated as

$$P_t \equiv \frac{\hat{\alpha}_t^m - \hat{\alpha}_t^w}{E[\tilde{Y}_{ist}^w|t]} \tag{2}$$

The expectation in the denominator of this formula represents the predicted income of females in absence of an event. This means that it is the expectation of Y_{ist}^w given t, calculated as

$$\tilde{Y}_{ist}^g \equiv \sum_k \hat{\beta}_k^g \cdot \mathbf{I}[k = age_{is}] + \sum_y \hat{\gamma}_y^g \cdot \mathbf{I}[y = s] \tag{3}$$

Next to simply calculating numerical child penalties, earnings of males and females relative to $t = -1$ are plotted to visualize differences in how these evolve over time. These plots show $\hat{\alpha}_j^g$ as a percentage of the outcome in absence of children, based on Equation (3).

The above-mentioned estimation procedure is repeated conditional on employment and for sub-samples based on birth year. Moreover, the regressions are run with dependent variables other than labor earnings. Specifically, the event study is also performed for annual work hours, as well as health satisfaction.

4.2 Lasso and Bayesian Lasso regressions

After estimating German child penalties, I continue by investigating health and labor effects of motherhood using Lasso regressions. Regression studies in this section involve different combinations of dependent and independent variables. By using (Bayesian) Lasso, the final models might be sparse, but it is important to include many different possible predictors in the estimation procedure. In the remainder of this section, the dependent variable is always denoted H, while the set of predictors is captured in the matrix $\mathbf{X}$.

The different outcome variables H used in the regression studies are ΔLabor earnings and Health satisfaction for different t. The predictors included in the $\mathbf{X}$ matrix include a large set of variables from the lists given in Section 3. The exact set of variables included in the regressions differ per dependent variable and sub-sample and specifics are discussed in the results section. Individuals' age at the time of childbirth and the birth year are always included to control for similar effects as in the event study of Equation (1). Finally, the number of observations in the sample is always denoted by n.

4.2.1 Lasso regression

The goal of Lasso regression is to minimize the sum of squares $\sum_{i=1}^{n}(H_i - \beta_0 - \sum_j x_{ij}\beta_{1,j})^2$ subject to the constraint $\sum_j |\beta_{1,j}| \leq s$. This implies that the coefficients can be calculated as

$$\hat{\beta}_0, \hat{\beta}_1^{Lasso} = \underset{\beta}{\text{argmin}}\{\sum_{i=1}^{N}(H_i - \beta_0 - \sum_j x_{ij}\beta_{1,j})^2 + \lambda\sum_j |\beta_{1,j}|\} \tag{4}$$

Note that the intercept term, β_0, is generally not penalized. The tuning parameter λ determines the strength of the penalty and can be determined by different methods, which are discussed in Section 4.2.3. If $\lambda = 0$, Lasso regression is equivalent to OLS regression and higher values for λ imply a higher amount of shrinkage. This means that increasing λ results in a higher number of coefficients being set to zero, thereby increasing bias, but decreasing variance.

4.2.2 The Bayesian Lasso

In Bayesian statistics, the parameter of interest is assumed to be following some prior density before observing the data, which represents prior beliefs about the value of the parameter. After observing the data, the posterior distribution is calculated to be a combination of the prior distribution and the observed likelihood. When Tibshirani (1996) introduced Lasso regression, he already suggested that $\hat{\beta}_1^{Lasso}$ can be viewed as the mode of the posterior distribution of β_1, with the $\beta_{1,j}$ following an independent and identical Laplace distribution. The posterior mode of the distribution is defined as the value that appears most often. Hence, denoting y to be the dependent variable of interest, the Lasso estimates can be viewed as

$$\hat{\beta}_1^{Lasso} = \underset{\beta_1}{\text{argmax}}\{p(\beta_1|y, \sigma^2, \tau)\} \tag{5}$$

Where the distribution of β_1 and the likelihood function $p(y|\beta_0, \beta_1, \sigma^2)$ are defined as

$$\begin{aligned} p(y|\beta_0\beta_1, \sigma^2) &= \mathcal{N}(y|\beta_0\mathbf{1}_n + \mathbf{X}\beta_1, \sigma^2\mathbf{I}_n), \\ p(\beta_1|\tau) &= \big(\frac{\tau}{2}\big)^p\exp(-\tau||\beta||_1) \end{aligned} \tag{6}$$

In the Bayesian Lasso, the penalty term is provided by a Laplace prior distribution, as suggested by Park and Casella (2008). This prior makes that the penalty parameter also needs to be estimated from the data. This is done using the following equation:

$$\pi(\beta_1|\sigma^2) = \prod_j \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_{1,j}|/\sqrt{\sigma^2}} \tag{7}$$

The following Bayesian Lasso regression model, in the notation of Hans (2009), can be written as in the equation below.

$$\begin{aligned} p(H|\beta_0, \beta_1, \sigma^2, \tau) &= \mathcal{N}(H|\beta_0 \mathbf{1}_n + \mathbf{X}\beta_1, \sigma^2 \mathbf{I}_n), \\ p(\beta_1|\tau, \sigma^2) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(\tau\sigma^{-1}||\beta||_1) \end{aligned} \tag{8}$$

The second line of Equation (8) is the prior, and this prior differs from the one in Equation (6). This is done to ensure that, for fixed values of τ and σ^2, the mode of this prior equals the Lasso estimate with $\lambda = 2\tau\sigma$.

While ordinary Lasso returns a set of non-zero coefficients and does not return the ones that are shrunk to zero, Bayesian Lasso returns coefficient estimates for all covariates included in the estimation procedure. To decide which of the coefficients should be set to zero, posterior probabilities are investigated. That is, the posterior probability that each coefficient is smaller than zero is estimated. If this probability is close to 0.5, it means that the estimate is centered around zero and hence should be set to zero.

4.2.3 Choice of tuning parameter

Finding an appropriate value for the tuning parameter is important because it leads to the correct level of parsimony of the model. For ordinary Lasso, a widely-used technique for finding λ is by using 10-fold cross-validation, as was also suggested by Tibshirani (1996). Other techniques include, amongst others, minimization of information criteria such as the Bayesian Information Criterion.

The idea of cross-validation is to find a value of λ that minimizes the Mean Squared Error (MSE) of predictions. In practice, performing a k-fold cross-validation implies splitting the data into a training and test set k times, where the test set has size equal to $\frac{1}{k}$-th of the full data. For each fold, the values in the test set are predicted using the data in the training set for different values of λ. The MSE is calculated as $\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$, with y_i and $\tilde{y}_i$ the observed and predicted values of the outcome variable, respectively.

The use of Bayesian Lasso creates the possibility of using Bayesian methods for choosing λ. Two of these are empirical Bayes through Marginal Maximum Likelihood and the use of an appropriate hyperprior (Park & Casella, 2008). In this research, the second option is employed and a gamma prior is chosen on λ^2. Thus, we have that

$$\pi\Big(\frac{\lambda^2}{2}\Big) \sim \text{Gamma}(a, b) \tag{9}$$

Park and Casella (2008) suggest choosing a and b such that the Gamma distribution is relatively flat. In this research, the shape parameter a is set to 1 and the scale parameter b to 0.1, which not only yields a relatively flat distribution, but also a high posterior probability near the Maximum Likelihood Estimate of λ (Casella et al., 2010).

The tuning parameter of Bayesian Lasso differs from the one used in ordinary Lasso, partly because ordinary Lasso standardizes covariates during computation. Park and Casella (2008) argue that the 10-fold cross-validation choice of λ in ordinary Lasso is often unstable, and suggest that Bayesian Lasso could aid in improving stability. In fact, by choosing the tuning parameter in ordinary Lasso such that coefficient estimates resemble those of Bayesian Lasso, stability increases. This specific tuning parameter is chosen by first looking at the number of non-zero coefficients for different values of λ and subsequently choosing the value of λ such that this number is as close as possible to the number of non-zero coefficients found by the Bayesian Lasso. If several values of λ result in the same number of coefficients, the option with the lowest Bayesian Information Criterion is chosen. Finally, the Lasso regression is performed again with the newly chosen tuning parameter.

4.3 Policy reform evaluation

For estimating the effect of the 2007 German family policy reform on health satisfaction of mothers, a method based on the Neyman-Rubin method is used (Rubin, 1974; Splawa-Neyman, Dabrowska, & Speed, 1990). This method estimates the Average Treatment Effect (ATE), for which an underlying assumption is randomness of treatment. Therefore, it is assumed that childbirth is not planned around policy interventions, especially not when the estimation sample includes mothers that give birth within a small time frame before and after policy implementation. Although the number of years around the policy implementation included in the sample is quite large due to small sample size, I still hold on to this assumption for simplicity.

The most basic ATE estimator represents the difference between average outcomes of the treatment and control groups, and can thus be written as $\widehat{ATE}_{baseline} = \overline{H_t} - \overline{H_c}$, with $\overline{H}_t$ and $\overline{H}_c$ the average outcomes of the treatment and control group, respectively. By adjusting for regression coefficients, variance of the ATE estimator can be reduced. This yields the adjusted ATE estimator

$$\widehat{ATE}_{adjusted} = [\overline{H_t} - (\overline{\mathbf{x}}_t - \overline{\mathbf{x}})^T \hat{\beta}_t] - [\overline{H_c} - (\overline{\mathbf{x}}_c - \overline{\mathbf{x}})^T \hat{\beta}_c] \tag{10}$$

$\hat{\beta}_t$ and $\hat{\beta}_c$ represent the coefficient estimates for the regression of health outcomes on the coefficients in $\mathbf{X}$. The $(\overline{\mathbf{x}}_t - \overline{\mathbf{x}})$ terms represent the amount of fluctuation between predictors of the sub-sample and the full sample. Bloniarz, Liu, Zhang, Sekhon, and Yu (2016) argue that using OLS estimates for the β parameters in this formula may lead to overfitting, especially when the number of covariates is large and the sample is small. With the number of mothers in the study sample small, it makes sense to employ this method in this thesis. Therefore, the coefficients used in the calculation of $\widehat{ATE}_{adjusted}$ are the $\hat{\beta}^{Lasso}$ for both the treatment and control group.

To find confidence intervals for the ATE estimate, Bloniarz et al. (2016) discuss that, under a set of conditions, it holds that

$$\sqrt{n}(\widehat{ATE}_{Lasso} - ATE) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \tag{11}$$

Where

$$\hat{\sigma}^2_{Lasso} = \frac{n}{n_t}\hat{\sigma}^2_{e^t} + \frac{n}{n_c}\hat{\sigma}^2_{e^c} \tag{12}$$

Here, n_t and n_c represent the number of observations in both sub-samples. $\hat{\sigma}^2_{e^t}$ and $\hat{\sigma}^2_{e^c}$ are estimates for the population variances of both subgroups. An explanation of how these are calculated is given in Appendix A.

It should be noted that there are cases in which the Lasso-adjusted ATE estimate does not outperform the OLS-adjusted estimate. Specifically, if the Lasso regressions yield a different set of covariates for the treatment and control groups, OLS-based ATE estimation might be better. Moreover, trends in maternal health might still have some sort of effect on ATE estimates after controlling. Therefore, the same procedure is performed on a placebo intervention taking place in 2000. Sample selection for this procedure is similar to what is discussed in Section 3, with the exception that the treatment sample now includes mothers who gave birth between 2000 and 2006, and the control group consists of mothers who had their first child between 1995 and 2000.

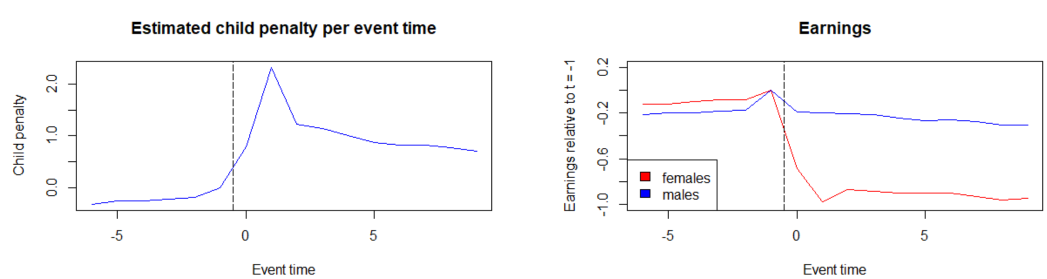
5 Results

5.1 Career effects of motherhood

This section gives an insight into career effects of motherhood by first presenting child penalty estimates for Germany. In order to see which characteristics of mothers relate to losses in earnings, results for Lasso regressions with ΔLabor earnings as the dependent variable are also presented.

5.1.1 Child penalties

Child penalties are estimated from $t = -6$ until $t = 9$, unconditional on employment. Figure 1(a) shows that the German child penalty peaks at $t = 1$ and subsequently decreases, with the long-run penalty at $t = 9$ equal to 0.671. When looking at earnings relative to $t = -1$, Figure 1(b) shows a clear difference between men and women. In the year after childbirth, German females earn 91.6% less compared to $t = -1$, and this percentage does not decrease much over time. For males, on the other hand, only a slight income decrease is visible. One thing that stands out in Figure 1(b) is the small peak in earnings at $t = -1$, which can potentially be explained by people advancing in their careers prior to having a child.



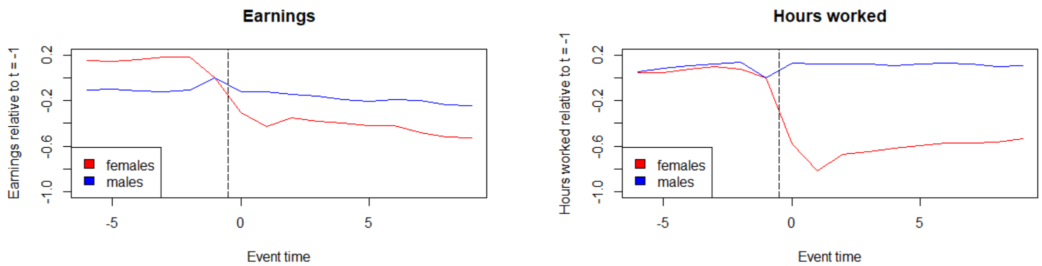
(a) Child penalty estimates

(b) Earnings with respect to $t = -1$

Figure 1. Estimation results for the full sample

[1] Figure (b) shows event time coefficients $\hat{\alpha}_j^g$ estimated from Equation (1), as a percentage of the counterfactual outcome in absence of children for males and females separately.

Conditional on employment, the child penalty is substantially smaller, with the estimated long-run penalty equal to 0.201. This implies that quite a large share of the total penalty arises from women exiting the labor force after giving birth. This is in line with estimation results when annual work hours are taken as the dependent variable. The child penalty in work hours is equal to 1.186, and hence the decrease in work hours for females compared to males is very large. Figure 2(a) clearly shows that earnings of males and females develop much more similarly when conditioning on employment. In Figure 2(b) it can be seen that there is a clear minimum in annual work hours directly after childbirth for females, while annual work hours stay relatively constant for males.



(a) Conditional on employment *(b)* Hours worked

Figure 2. Earnings and hours worked with respect to $t = -1$

[1] The figure shows event time coefficients $\hat{\alpha}_j^g$ estimated from Equation (1), as a percentage of the counterfactual outcome in absence of children for males and females separately.

5.1.2 Change in labor earnings

This section presents results from (Bayesian) Lasso regressions with ΔLabor earnings as the dependent variable. These regressions are performed in order to get insights into which characteristics relate to losses in earnings. Most variables presented in Section 3 are included in **X**, with the exception of some labor-related variables. Throughout the results section, presented Lasso results are found by first estimating Lasso coefficients β_{CV} using the tuning parameter found through 10-fold cross-validation (λ_{CV}). If Bayesian Lasso estimates differ much from these coefficients, ordinary Lasso is performed again, but with tuning parameter λ_B. This yields coefficient estimates β_B.

Tables 1 and 2 show coefficient estimates for both changes in labor earnings between the baseline and $t = 5$ as changes between the baseline and $t = 8$. It should be noted that Lasso coefficients using λ_{CV} resemble Bayesian coefficients most, and hence coefficients for λ_B are not shown.

For the change in earnings between $t = -1$ and $t = 5$, coefficients for age at childbirth, baseline health satisfaction and baseline frequency of sports are all estimated to be negative, meaning that higher values for these variables imply a larger income drop after childbirth. All other coefficients in Table 1 are positive. For instance, an extra year of education is estimated to decrease the income loss by around 518 euros, and a later birth year by 275 euros. Since baseline employment level is positively related to ΔLabor earnings, women who do not work full-time at the baseline are less likely to experience sharp income losses. This seems quite logical, as women working full-time before giving birth are likely to have a higher income at the baseline and thus can lose more of it. With marital status having a positive coefficient, it can be argued that non-married women experience lower income losses, potentially due to not having a partner to financially rely on after childbirth.

For the longer run, results are similar in sign, but different in magnitude. For instance, the birth year has a larger effect in the longer run, while baseline health satisfaction is estimated to affect long-run changes in earnings less. The procedure is also performed using ΔWork hours as dependent variable, yielding very similar results. This can be seen in Appendix B.

Table 1

Lasso and Bayesian Lasso coefficient estimates for the full sample of females.

Dependent variable: ΔLabor earnings between $t = -1$ and $t = 5$.
Number of observations: 738

	Lasso	Bayesian Lasso	
	β_{CV}	β	95% CI
Intercept	-5.591×10^5	-3.670×10^5	[-5.296×10^5, -2.049×10^5]
Age birth	-530.186	-697.250	[-930.692, -384.064]
Birth year	274.603	182.218	[101.157, 262.816]
Years of education	517.750	531.276	[151.325, 914.428]
Baseline employment level	7.416×10^3	6.538×10^3	[4.872×10^3, 8.293×10^3]
Marital status	2.315×10^3	1.990×10^3	[7.178×10^2, 3.272×10^3]
Baseline health satisfaction	-302.639	-422.141	[-1036.816, 125.861]
Baseline frequency of sports	-790.912	-726.479	[-1678.245, 189.157]
λ	13.689	5.109	[2.657, 8.052]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

Table 2

Lasso and Bayesian Lasso coefficient estimates for the full sample of females.

Dependent variable: ΔLabor earnings between $t = -1$ and $t = 8$.
Number of observations: 575

	Lasso	Bayesian Lasso	
	β_{CV}	β	95% CI
Intercept	-7.546×10^5	-6.354×10^5	[-7.605×10^5, -5.106×10^5]
Age birth	-619.138	-761.108	[-1022.611, -436.178]
Birth year	374.012	317.809	[252.090, 378.085]
Years of education	561.701	571.988	[149.567, 967.861]
Baseline employment level	6.831×10^3	6.194×10^3	[4.464×10^3, 7.941×10^3]
Marital status	2.161×10^3	1.854×10^3	[4.492×10^2, 3.149×10^3]
Baseline health satisfaction	-30.016	-198.816	[-823.710, 364.519]
Baseline frequency of sports	-667.926	-707.037	[-1.694×10^3, 268.423]
λ	68.360	4.964	[2.582, 7.878]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

5.2 Health effects of motherhood

After estimating career effects of motherhood, I continue by investigating maternal health. This section therefore first presents some insights into results from the event study with health satisfaction as the dependent variable, after which Lasso and Bayesian Lasso estimates for long-term health satisfaction are shown.

5.2.1 Health satisfaction around childbirth

Figure 3 shows that health satisfaction stays relatively constant around childbirth for both males and females. Looking at the estimated values of α_j^w, as in Equation (1), it can be argued that the event time dummies have no dynamic effect on health satisfaction. In fact, while α_j^w changes sign after $t = -1$ in the regressions where labor earnings or annual work hours are taken as dependent variable, it stays relatively constant over time in this framework and does not change sign. Therefore, there is no reason to study child penalties in terms of health satisfaction any further.

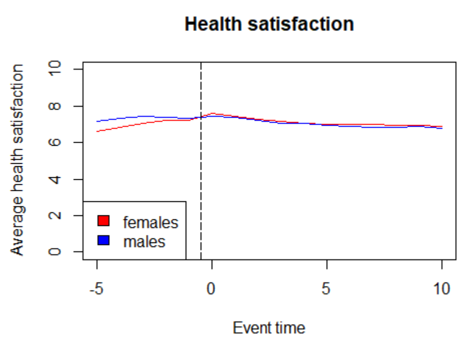


Figure 3. Average health satisfaction for males and females at each event time

5.2.2 Health satisfaction after five years

Table 3 shows coefficient estimates for Lasso and Bayesian Lasso regressions on the full sample of females, with the dependent variable being individuals' health satisfaction at $t = 5$. Most coefficients are estimated non-zero, with the only predictors from $\mathbf{X}$ that are shrunk to zero being marital status at the baseline and the dummy indicating separation from one's spouse after childbirth.

The table shows that mothers who gave birth at a later age or in a later year are estimated to be less satisfied with their health. Education, baseline health satisfaction, quitting a full-time job and baseline frequency of sports all relate to higher health satisfaction. The coefficient for baseline employment level is negative, hence working more at the baseline is considered to positively affect health. The coefficients for ΔLabor earnings and ΔWork hours have opposite signs. In particular, a positive change in labor earnings relates to lower health satisfaction, while an increase in annual work hours is related to higher health satisfaction. The magnitude of the latter two coefficients seems small, but is actually quite large looking at changes in income that occur. For instance, a women who loses 10,000 euros of yearly income is estimated to grade her health 0.112 points lower, while one extra year of education is considered to increase health satisfaction by only 0.033 points.

Table 3
Lasso and Bayesian Lasso coefficient estimates for the full sample of females at $t = 5$.

Dependent variable: Health satisfaction at $t = 5$
Number of observations: 756

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.255	14.053	20.486	[17.807, 22.742]
Age birth	-	-0.002	-0.014	[-0.041, 0.018]
Birth year	-0.554×10^{-3}	-0.005	-0.008	[-0.009, -0.007]
Years of education	0.033	0.039	0.047	[0.002, 0.094]
Baseline employment level	-0.037	-0.064	-0.114	[-0.367, 0.141]
Baseline health satisfaction	0.344	0.354	0.347	[0.272, 0.410]
ΔLabor earnings	-0.876×10^{-5}	-0.112×10^{-4}	-0.140×10^{-4}	[-0.251×10^{-4}, -0.343×10^{-5}]
ΔWork hours	0.177×10^{-3}	2.491×10^{-4}	2.833×10^{-4}	[0.839×10^{-4}, 4.821×10^{-4}]
Stop full time	0.169	0.238	0.185	[-0.148, 0.578]
Baseline frequency of sports	0.039	0.049	0.055	[-0.059, 0.168]
λ	0.028	0.012	6.610	[3.993, 9.625]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

The regression is repeated for two sub-samples of females: those who work full-time in the baseline year and those who work part-time or do not work at this time. For these regressions, the variables stop full time and baseline employment level are omitted from $\mathbf{X}$. Table 4 shows estimated coefficients β_B and β for both sub-groups of females. Full estimation results, including β_{CV} and confidence intervals are shown in Tables C.2 and C.3 in Appendix C.

Table 4 shows some differences between the two sub-groups. First of all, the average health satisfaction is substantially higher for those working full-time before giving birth, while less coefficients are estimated non-zero for this group. For women not working full-time at the baseline, estimation results are quite similar to those of the full sample. The only differences are that separation from one's spouse is estimated to positively affect health satisfaction, and that baseline frequency of sports is set to zero. For women working full-time at the baseline, the variable separated is estimated to negatively affect health, and baseline frequency of sports is considered a positive predictor of health satisfaction. For these women, however, the coefficients for age at the time of birth and years of education are shrunk to zero.

Table 4
Lasso and Bayesian Lasso coefficient estimates for sub-samples of women who work full-time or not full-time in the baseline year.

	Women not working full-time		Women working full-time	
	β^{Lasso}	$\beta^{Bayesian}$	β^{Lasso}	$\beta^{Bayesian}$
Intercept	32.111	50.792	5.046	2.752
Age birth	-0.002	-0.017	-	-
Birth year	-0.014	-0.024	-	0.001
Years of education	0.093	0.113	-	-
Separated	1.426	0.326	-	-0.151
Baseline health satisfaction	0.366	0.388	0.289	0.333
ΔLabor earnings	-0.711×10^{-6}	-0.208×10^{-4}	-0.239×10^{-5}	-0.134×10^{-4}
ΔWork hours	1.805×10^{-4}	3.972×10^{-4}	0.167×10^{-4}	2.479×10^{-4}
Baseline frequency of sports	-	-	0.020	0.076
λ	0.053	6.064	0.082	6.282
Number of observations	224	224	502	502
Average health satisfaction	6.849	6.849	7.234	7.234

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient to be zero.

5.2.3 Health satisfaction after eight years

When performing Lasso and Bayesian Lasso regressions for the full sample of females, but with health satisfaction at $t = 8$ as the dependent variable and all other time-related variables linking to $t = 8$, some results become different from those found in Table 3. Most strikingly, Table 5 shows that ordinary Lasso shrinks many coefficients to zero. Moreover, a change in annual work hours is not estimated to affect health at $t = 8$, and the impact of ΔLabor earnings is estimated much smaller compared to $t = 5$. On the other hand, marital status at the baseline is found to affect longer-run health satisfaction. Finally, it is interesting to see that Bayesian Lasso estimates the coefficient for baseline employment level to have an opposite sign compared to Table 3, with working part-time or not working before giving birth being positively related to health satisfaction eight years after giving birth.

Table 5
Lasso and Bayesian Lasso coefficient estimates for the full sample of females at $t = 8$.

Dependent variable: Health satisfaction at $t = 8$
Number of observations: 599

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.458	5.263	11.544	[5.808, 14.713]
Age birth	-	-0.003	-0.032	[-0.066, 0.010]
Birth year	-	-	-0.003	[-0.005, -0.001]
Years of education	-	-	0.020	[-0.030, 0.075]
Baseline employment level	-	-	0.106	[0.149, 0.400]
Marital status	-	-0.015	-0.076	[-0.258, 0.090]
Baseline health satisfaction	0.224	0.254	0.288	[0.204, 0.367]
ΔLabor earnings	-	-0.236×10^{-5}	-0.794×10^{-5}	[-0.217×10^{-4}, -0.606×10^{-5}]
Baseline frequency of sports	-	0.024	0.094	[-0.030, 0.217]
λ	0.135	0.077	6.660	[3.983, 9.752]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

5.2.4 Relating career and health effects

Lasso and Bayesian Lasso results suggest that different characteristics of mothers can be linked to different levels of health satisfaction, and that health and career effects of motherhood could potentially be related. Looking at pre-birth characteristics, results suggest that older mothers experience lower health satisfaction, and also lose more income after giving birth. The negative health effect for these mothers could be explained by the fact that being older simply brings some more health problems, but other mechanisms may play a role as well. Education is positively related to both health satisfaction as labor earnings, which is in line with many earlier findings.

It is not surprising that baseline health and frequency of sports are both positively related to long-run health satisfaction, but it is interesting to see that these variables are suggested to relate to losses in labor earnings. Moreover, a negative change in earnings, as well as quitting a full-time job, are found to relate to higher health satisfaction in the long run. On the other hand, a decrease in annual work hours is estimated to negatively affect health and working full-time at the baseline relates to better health. Further research is needed to find out which mechanisms cause these seemingly contradictory findings, but it could be that women's choice of working in more family-friendly jobs relates to this. Specifically, working the same number of hours, but at a different type of job can cause less stress, but also means receiving a lower wage.

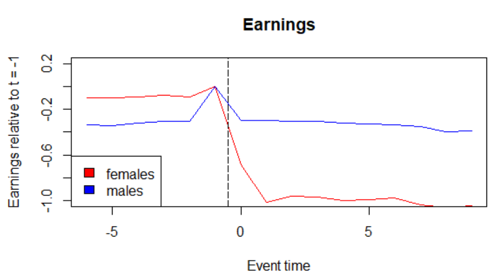
A final interesting characteristic relating to health outcomes is marital status. For women working part-time, or not working, at the baseline, separation from one's spouse is estimated to positively affect health satisfaction and being married at the baseline is related to higher health satisfaction after eight years, as well as higher labor earnings. The latter could be explained by the fact that non-married mothers cannot rely on a spouse to financially provide for the family, and hence they cannot lose too much income themselves.

5.3 Effect of parental leave reform

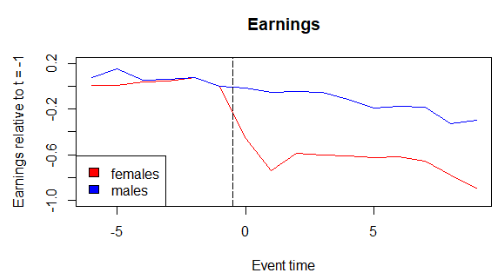
As discussed in the introduction, a reform in Germany's family policy system was introduced in 2007. This reform increased payments during the parental leave period and introduced two months of leave specifically reserved for fathers, thereby making it more attractive for both parents to take on leave. This section presents a comparison of child penalties between before and after the reform, as well as an estimate for the effect of this reform on long-term maternal health satisfaction.

5.3.1 Effect on child penalties

Child penalty estimation is performed separately for births before and after the reform. This leads to a sub-sample of 1713 parents having their first child before 2007, and one of 698 parents who benefited from the reform. Figure 4 shows that earnings develop differently between the two sub-samples. Specifically, the effect of the two "daddy-months" is clearly visible in Figure 4(b), as earnings of fathers also decrease after childbirth. This is reflected in long-term child penalties, with the penalty for the sub-sample having their first child before the reform equal to 0.662 at $t = 9$, while it is estimated to be 0.556 for the other sub-sample.



(a) Before 2007



(b) After 2007

Figure 4. Earnings with respect to $t = -1$

[1] The figure shows event time coefficients $\hat{\alpha}_j^g$ estimated from Equation (1), as a percentage of the counterfactual outcome in absence of children for males and females separately. [2] Figure (a) shows results from a sample of 1713 parents and Figure (b) of 698 parents.

5.3.2 Health effects

For calculating the Lasso-adjusted value of $\widehat{ATE}$, Lasso regressions are run first for both the treatment and control group, with the treatment group being the group of mothers who gave birth in 2007 or later. The estimation procedure is similar to the one discussed in Section 5.2.2, and full estimation results can be found in Appendix D.

Lasso regression results show that the set of non-zero coefficients differs between the treatment and control group. This means that in this research, using OLS-adjusted ATE estimates might be better than using the Lasso-adjusted ones and hence both options are investigated. The found estimates are quite close, with $\widehat{ATE}_{Lasso}$ = -0.229 and $\widehat{ATE}_{OLS}$ = -0.339 and both variances between 14 and 15. These estimates suggest that the policy reform negatively affected long-term maternal health. However, Figure 5 shows that zero is included in the confidence intervals of both estimates, and hence there is insufficient evidence to provide proof for this claim. From these results, it can thus be concluded that the reform had no significant impact on long-term health satisfaction of mothers.

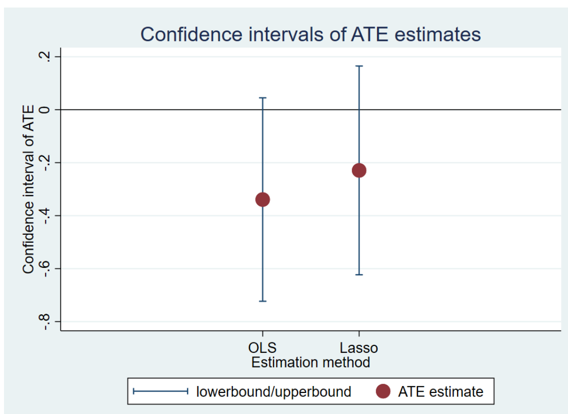


Figure 5. ATE estimates and 95% confidence intervals

Despite the insignificant estimates, it remains interesting to look at ATE estimates for a placebo intervention taking place seven years before the actual reform. Appendix E shows that there are large differences in which coefficients are shrunk to zero between the placebo treatment and control groups. It therefore suffices to only estimate the OLS-adjusted ATE for this intervention. This yields $\widehat{ATE}_{OLS}$ = 0.077, hence the estimate has opposite sign than for the actual intervention. However, with an estimated variance equal to 12.594, the 95% confidence interval of this estimate ranges from -0.224 to 0.377, implying no conclusions can be drawn from this result either.

6 Conclusion

The first part of this thesis studies child penalties in Germany, with the aim of answering the first research question:

> *To what extent do child penalties exist in Germany and how do different mechanisms affect these?*

This question can be answered by first recognizing that child penalties in Germany are of relatively large size and that women lose significantly more income after having a child compared to men. Results suggest that the main reason for this lies in the fact that women start working less or quit work after giving birth, with child penalties conditional on employment much smaller, while annual work hours decrease much more for mothers compared to fathers. Looking at the 2007 parental leave reform, results show that penalties decreased after it. Specifically, having leave time reserved for fathers seems to make them more likely to lose income after childbirth as well.

The second part of the thesis investigates long-term health satisfaction of mothers in order to answer the second research question:

> *Does having a child affect health and how do career effects of motherhood relate to long-term health outcomes?*

Results from the event study regarding health effects of having a child suggest no significant changes in health satisfaction around childbirth. However, some interesting results are found when looking at characteristics relating to long-term health satisfaction. Younger, better-educated mothers with higher health upfront are suggested to be less likely to suffer from bad health in the long run. Moreover, working full-time prior to giving birth and not decreasing annual work hours is related to higher health satisfaction. At the same time, an increase in income is not necessarily good for mothers' health. In fact, several characteristics that relate to better health are also related to losses in earnings.

The second sub-question focuses on Germany's parental leave reform in 2007 and how this affected mothers' long-term health. Two different ATE estimates are calculated in order get insight into the effect of the reform, but results are inconclusive. That is, there is no significant evidence in support of any effect of the reform on long-term health satisfaction.

Overall, the second research question can be answered by arguing that the health effect of motherhood remains vague, but that career effects of motherhood and long-term health satisfaction are somewhat related. The precise mechanism behind this relation remains unidentified. That is, losses in earnings positively relate to health, but working less relates to declines in health. Moreover, potential career effects of parental leave seem to not affect long-term health satisfaction. In order to formulate a more conclusive answer to this research question, further research is still needed.

6.1 Policy implications

The results have some interesting implications for policy makers. First of all, German policy makers should be aware of the fact that the gender wage gap is much larger in Germany compared to many other countries. Especially in terms of annual work hours, women experience large career effects. It might therefore be sensible to think about policies that incentivize more mothers to stay employed, or policies that create more equality between mothers and fathers. The latter has already partly been achieved through the parental leave reform that reserved some leave specifically for fathers. However, to keep more mothers in the labor force, different policies might be needed.

Results with respect to health effects of motherhood are not fully conclusive on all aspects, but still yield some interesting policy implications. First of all, with results suggesting that older and less-educated mothers suffer from worse health, it might be useful to provide more health screenings or health-supporting programs for these mothers. If further research suggests a wider set of characteristics relating to worse health, the targeting of these policies can be made more precise. Moreover, since a decrease in annual work hours is estimated to negatively affect health, incentivizing mothers to work might not only decrease child penalties, but also help improve health.

6.2 Limitations

A main limitation of this research is that the available data from the German SOEP is not completely suitable for it. For the study of child penalties, the data is the same as the data used in Kleven, Landais, Posch, et al. (2019), and hence it is a good fit for this. Due to slightly different steps taken in sample selection and estimating penalties until nine years after childbirth instead of 10 years, results in this thesis differ slightly from those found in the original paper, but conclusions are very similar. Having more variables on individuals' characteristics included in the data could potentially have given more insight into mechanisms behind penalties.

For the study of health effects, the data does not provide all information needed to get very meaningful results. For instance, the majority of observations have missing values for many health-related variables. Therefore, only subjective health satisfaction could be used as a dependent variable in the Lasso regressions, while it would have been interesting to also have more objective measures of health included. Moreover, including a wider variety of predictors in the regressions would have given a better insight into which mothers are likely to suffer from bad health.

A final note on the limitations caused by data availability concerns the evaluation of the 2007 parental leave reform in Germany. Because of the small sample size, the estimation sample includes mothers who give birth in a number of years around the reform. Hence, the assumption of randomness of treatment is not likely to hold. A larger sample of mothers would have made it easier to study the causal effect of the reform on long-term health satisfaction. Moreover, having more observations and variables could have helped in filtering out the effect of the 2007-2008 economic crisis on fertility decisions, labor force participation and health outcomes.

Aside from data availability, there are also limitations relating to the methods used. One disadvantage of using Lasso regression is that it does not perform well in group selection. That is, if there is a group of variables with high pairwise correlations, Lasso is likely to randomly select the coefficient of only one of these variables to be non-zero (Zou & Hastie, 2005). Since the regression studies in this thesis included some predictors that are likely to be related - such as changes in earnings and changes in annual work hours - this might have caused slightly incorrect results in some cases. Moreover, for the Bayesian Lasso the specified priors used are the same in each specification. While used priors are based on suggestions from earlier research, and might not affect results substantially, trying different priors could lead to more robust results.

6.3 Suggestions for further research

A quite obvious suggestion for further research is to repeat the study on a different dataset, which includes a wider variety of predictors of health and income, as well as more objective health outcome variables. For instance, it would be interesting to have information on whether individuals had certain health issues prior to childbirth, or to include indicators of individuals' personal beliefs with respect to gender norms. Examples of such indicators could be reliogisity or migration history of individuals (Rabaté & Rellstab, 2021). Moreover, looking at child penalties and health outcomes of mothers in the longer run, like 20 years after childbirth, could potentially be very insightful.

Another suggestion is to use different methods for investigating the effect of motherhood on health and the effect of the parental leave reform on outcomes. Specifically, it would be interesting to compare health outcomes between mothers and non-mothers instead of using the event study employed in this thesis. Besides, with a larger sample of mothers giving birth around the implementation of the reform, one could use a regression discontinuity design for estimating the effect of the policy reform on health. This might reduce the bias caused by including females that give birth long before or after the reform in the sample.

I would also suggest to perform the research on data from another country. Germany's child penalties are very high, and even the highest of all countries studied by Kleven, Landais, Posch, et al. (2019). The height of the child penalty and conservative gender norms in Germany may have a large impact on results with respect to maternal health. It would be very interesting to see whether different factors relate to maternal health in countries with more equal gender norms and lower child penalties, such as Scandinavian countries. Studying maternal health in these countries would also create the opportunity of looking into the effect of different family policy schemes. For instance, these countries have different parental leave benefit options, but also have policies in place that make childcare more affordable. Investigating whether these policies have significant health benefits for mothers would be very valuable for policy makers around the globe.

Finally, in the future it would be interesting to investigate how the current Covid-19 pandemic affects labor force participation and health outcomes in the long run. Several papers argue that time constraints in families caused by the closing of schools and childcare facilities have made mothers become more engaged in caring tasks (e.g. Biroli et al. (2020), Sevilla and Smith (2020)). Alon, Coskun, Doepke, Koll, and Tertilt (2021) even find that lockdown measures have caused a drop in female labor force participation and Boring and Moroni (2021) conducted a study among French women concluding that gender norms have become more traditional due to lockdowns. It is of interest to see whether these trends cause higher child penalties in a few years, and whether health outcomes of mothers are also affected by this.

References

Alon, T., Coskun, S., Doepke, M., Koll, D., & Tertilt, M. (2021). *From mancession to shecession: Women's employment in regular and pandemic recessions* (Tech. Rep.). National Bureau of Economic Research.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, *3*, 3143–3259.

Anderson, D. J., Binder, M., & Krause, K. (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *ILR Review*, *56*(2), 273–294.

Andresen, M. E., & Nix, E. (2019). *What causes the child penalty? evidence from same sex couples and policy reforms* (Tech. Rep.). Discussion Papers.

Avendano, M., Berkman, L. F., Brugiavini, A., & Pasini, G. (2015). The long-run effect of maternity leave benefits on mental health: evidence from European countries. *Social Science & Medicine*, *132*, 45–53.

Bayes, T. (1763). Lii. an essay towards solving a problem in the doctrine of chances. *Philosophical transactions of the Royal Society of London*(53), 370–418.

Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, *3*(1, Part 2), S33–S58.

Biroli, P., Bosworth, S., Della Giusta, M., Di Girolamo, A., Jaworska, S., & Vollen, J. (2020). Family life in lockdown.

Blair-Loy, M. (2009). *Competing devotions: Career and family among women executives.* Harvard University Press.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, *55*(3), 789–865.

Bloniarz, A., Liu, H., Zhang, C. H., Sekhon, J. S., & Yu, B. (2016). Lasso adjustments of treatment effect estimates in randomized experiments. *Proceedings of the National Academy of Sciences*, *113*(27), 7383–7390.

Boring, A., & Moroni, G. (2021). Turning back the clock: Beliefs in gender norms during lockdown.

Brand, J. E. (2015). The far-reaching impact of job loss and unemployment. *Annual review of sociology*, *41*, 359–375.

Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American Sociological*

Review, 204–225.

Bütikofer, A., Riise, J., & Skira, M. M. (2021). The impact of paid maternity leave on maternal health. *American Economic Journal: Economic Policy*, *13*(1), 67–105.

Casella, G., Ghosh, M., Gill, J., & Kyung, M. (2010). Penalized regression, standard errors, and Bayesian Lassos. *Bayesian analysis*, *5*(2), 369–411.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American journal of sociology*, *112*(5), 1297–1338.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues*, *60*(4), 701–718.

Cunningham-Burley, S., Backett-Milburn, K., & Kemmer, D. (2006). Constructing health and sickness in the context of motherhood and paid work. *Sociology of health & illness*, *28*(4), 385–409.

Dagher, R. K., McGovern, P. M., & Dowd, B. E. (2014). Maternity leave duration and postpartum mental and physical health: implications for leave policies. *Journal of Health Politics, Policy and Law*, *39*(2), 369–416.

Feldhoff, C. H. (2021). *The child penalty: Implications of parenthood on labour market outcomes for men and women in Germany* (Tech. Rep.). SOEPpapers on Multidisciplinary Panel Data Research.

Felfe, C. (2006). *The child penalty–a compensating wage differential?* CEPS.

Fernández, R., Fogli, A., & Olivetti, C. (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, *119*(4), 1249–1299.

Gangl, M., & Ziefle, A. (2009). Motherhood, labor force behavior, and women's careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, *46*(2), 341–369.

García-Gómez, P., Jones, A. M., & Rice, N. (2010). Health effects on labour market exits and entries. *Labour Economics*, *17*(1), 62–76.

Geisler, E., & Kreyenfeld, M. (2012). How policy matters: Germany's parental leave benefit reform and fathers' behavior. *Rostock: Max Planck Institute for Demographic Research*.

Goldin, C. (1988). Marriage bars: Discrimination against married women workers, 1920's to 1950's. *NBER working paper*(w2747).

Grice, M. M., Feda, D., McGovern, P., Alexander, B. H., McCaffrey, D., & Ukestad, L. (2007). Giving birth and returning to work: the impact of work–family conflict on women's health after childbirth. *Annals of Epidemiology*, *17*(10), 791–798.

Hans, C. (2009). Bayesian lasso regression. *Biometrika*, *96*(4), 835–845.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

ISSP Research Group. (2016). *International Social Survey Programme: Family and Changing Gender Roles IV - ISSP 2012.* GESIS Data Archive, Cologne. Data file Version 4.0.0.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019). Child penalties across countries: Evidence and explanations. In *AEA papers and proceedings* (Vol. 109, pp. 122–26).

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209.

Kleven, H., Landais, C., & Søgaard, J. E. (2020). *Does biology drive child penalties? evidence from biological and adoptive families* (Tech. Rep.). National Bureau of Economic Research.

Klumb, P. L., & Lampert, T. (2004). Women, work, and well-being 1950–2000:: a review and methodological critique. *Social science & medicine*, *58*(6), 1007–1024.

Lambe, M., Hsieh, C. C., Trichopoulos, D., Ekbom, A., Pavia, M., & Adami, H. O. (1994). Transient increase in the risk of breast cancer after giving birth. *New England Journal of Medicine*, *331*(1), 5–9.

Miller, A. (2002). *Subset selection in regression.* CRC Press.

Olesen, S. C., Butterworth, P., Leach, L. S., Kelaher, M., & Pirkis, J. (2013). Mental health affects future employment as job loss affects mental health: findings from a longitudinal population study. *BMC psychiatry*, *13*(1), 1–9.

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681–686.

Rabaté, S., & Rellstab, S. (2021). *The child penalty in the Netherlands and its determinants* (Tech. Rep.). CPB Discussion Paper.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Saurel-Cubizolles, M. J., Romito, P., Lelong, N., & Ancel, P. Y. (2000). Women's health after childbirth: a longitudinal study in France and Italy. *BJOG: An International Journal of Obstetrics & Gynaecology*, *107*(10), 1202–1209.

Schmitz, H. (2011). Why are the unemployed in worse health? the causal effect of unemployment on health. *Labour economics*, *18*(1), 71–78.

Sevilla, A., & Smith, S. (2020). Baby steps: the gender division of childcare during the COVID-19 pandemic. *Oxford Review of Economic Policy*, *36*(Supplement_1), S169–S186.

Socio-Economic Panel (SOEP). (2021). *Data for years 1984-2019, SOEP-Core v36, EU Edition.* doi: 10.5684/soep.core.v36eu

Splawa-Neyman, J., Dabrowska, D. M., & Speed, T. (1990). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 465–472.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Townsend, N. (2010). *Package deal: Marriage, work and fatherhood in men's lives.* Temple University Press.

Waldfogel, J. (1997). The effect of children on women's wages. *American sociological review*, 209–217.

Waldron, I., Weiss, C. C., & Hughes, M. E. (1998). Interacting effects of multiple roles on women's health. *Journal of health and social behavior*, 216–236.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301–320.

Appendices

A Calculation of population variance

The population variances used in the calculation of $\hat{\sigma}^2_{Lasso}$ are defined as the following:

$$\hat{\sigma}^2_{e^t} = \frac{1}{n_t - df^{(t)}} \sum_{i \in T} (H_i - \overline{H}_t - (x_i - \overline{x}_t)^T \hat{\beta}_t^{Lasso})^2 \tag{A.1}$$

$$\hat{\sigma}^2_{e^c} = \frac{1}{n_c - df^{(c)}} \sum_{i \in C} (H_i - \overline{H}_c - (x_i - \overline{x}_c)^T \hat{\beta}_c^{Lasso})^2 \tag{A.2}$$

Where n_t and n_c are the number of observations in the treatment and control groups, and T and C represent the full sets of observations in these groups. The degrees of freedom $df^{(t)}$ and $df^{(c)}$ are calculated as

$$\begin{aligned} df^{(t)} &= ||\hat{\beta}_t^{Lasso}||_0 + 1 \\ df^{(c)} &= ||\hat{\beta}_c^{Lasso}||_0 + 1 \end{aligned} \tag{A.3}$$

With the $||\hat{\beta}_t^{Lasso}||_0$ representing the L0-norm, which is equal to the number of non-zero elements in the vector of coefficient estimates.

B Lasso results for changes in work hours

Table B.1

Lasso and Bayesian Lasso coefficient estimates for the full sample of females.

Dependent variable: ΔWork hours between $t = -1$ and $t = 5$.
Number of observations: 738

	Lasso	Bayesian Lasso	
	β_{CV}	β	95% CI
Intercept	-2.895×10^4	-8.334×10^3	$[-1.989\times10^4, -2.337\times10^3]$
Age birth	-4.476	-12.651	[-31.575, 6.015]
Birth year	13.282	3.152	[-0.047, 9.204]
Years of education	14.614	19.475	[0.223, 42.207]
Baseline employment level	886.589	845.695	[738.227, 954.824]
Marital status	169.532	159.075	[83.860, 239.382]
Separated	237.607	110.634	[-176.855, 454.412]
Baseline health satisfaction	-15.152	-25.794	[-55.146, 49.677]
λ	4.603	4.300	[2.249, 6.922]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.

[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

C Lasso results for sub-samples of females

Table C.2

Lasso and Bayesian Lasso coefficient estimates for the sample of females not working full-time in the baseline year.

Dependent variable: Health satisfaction at $t = 5$
Number of observations: 224

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	15.506	32.111	50.792	[45.902, 56.249]
Age birth	-	-0.002	-0.017	[-0.075, 0.041]
Birth year	-0.006	-0.014	-0.024	[-0.026, -0.021]
Years of education	0.078	0.093	0.113	[0.022, 0.196]
Separated	1.150	1.426	0.326	[-0.460, 1.541]
Baseline health satisfaction	0.347	0.366	0.388	[0.261, 0.512]
ΔLabor earnings	-	-0.711×10^{-6}	-0.208×10^{-4}	$[-0.561\times10^{-4}, 0.146\times10^{-4}]$
ΔWork hours	1.431×10^{-4}	1.805×10^{-4}	3.972×10^{-4}	$[0.236\times10^{-4}, 7.642\times10^{-4}]$
λ	0.084	0.053	6.064	[3.343, 9.057]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.

[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

Table C.3

Lasso and Bayesian Lasso coefficient estimates for the sample of females working full-time in the baseline year.

Dependent variable: Health satisfaction at $t = 5$
Number of observations: 502

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.265	5.046	2.752	[-0.033, 0.034]
Birth year	-	-	0.001	[0.000, 0.002]
Separated	-	-	-0.151	[-0.460, 1.541]
Baseline health satisfaction	0.267	0.289	0.333	[0.256, 0.405]
ΔLabor earnings	-	-2.39×10^{-6}	-0.134×10^{-4}	[-0.024×10^{-4}, -2.56×10^{-6}]
ΔWork hours	-	0.167×10^{-4}	2.479×10^{-4}	[0.263×10^{-4}, 0.463×10^{-3}]
Baseline frequency of sports	-	0.020	0.076	[-0.037, 0.199]
λ	0.120	0.082	6.282	[0.678, 4.597]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.

D Lasso results for treatment and control groups

Table D.1

Lasso and Bayesian Lasso coefficient estimates for the treatment group.

Dependent variable: Health satisfaction at $t = 5$, $t = 6$ or $t = 7$
Number of observations: 124

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.040	4.464	3.083	[0.655, 5.869]
Years of education	-	0.012	0.079	[-0.021, 0.180]
Baseline employment level	-	-0.025	-0.441	[-1.158, 0.164]
Marital status	-	-	-0.215	[-0.593, 0.112]
Baseline health satisfaction	0.256	0.314	0.461	[0.288, 0.632]
ΔWork hours	-	-	2.794×10^{-4}	[-1.507×10^{-4}, 7.491×10^{-4}]
Baseline frequency of sports	-	-	0.152	[-0.096, 0.422]
λ	0.363	0.275	5.349	[2.987, 8.334]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.
[3] For each individual, the information of the latest available year (out of 5, 6 and 7) is used for estimation.

Table D.2

Lasso and Bayesian Lasso coefficient estimates for the control group.

Dependent variable: Health satisfaction at $t = 5$, $t = 6$ or $t = 7$
Number of observations: 239

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.110	4.962	4.946	[3.185, 7.161]
Years of education	-	-	-0.015	[-0.094, 0.060]
Baseline health satisfaction	0.274	0.292	0.319	[0.210, 0.428]
ΔLabor earnings	-	-0.152×10^{-5}	-0.108×10^{-5}	[-0.297×10^{-5}, -0.824×10^{-5}]
ΔWork hours	-	-	0.919×10^{-4}	[-2.262×10^{-4}, 4.139×10^{-4}]
Baseline frequency of sports	-	-	0.025	[-0.149, 0.191]
λ	0.115	0.079	6.127	[3.419, 9.164]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.
[3] For each individual, the information of the latest available year (out of 5, 6 and 7) is used for estimation.

E Lasso results for placebo treatment and control groups

Table E.1

Lasso and Bayesian Lasso coefficient estimates for the placebo treatment group.

Dependent variable: Health satisfaction at $t = 5$, $t = 6$ or $t = 7$
Number of observations: 239

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.110	4.962	4.946	[3.185, 7.161]
Baseline employment level	-	-	0.056	[-0.290, 0.449]
Baseline health satisfaction	0.274	0.292	0.319	[0.210, 0.428]
ΔLabor earnings	-	-0.152×10^{-5}	-0.108×10^{-4}	[-0.297×10^{-4}, 0.824×10^{-5}]
ΔWork hours	-	-	0.919×10^{-4}	[-2.262×10^{-4}, 4.193×10^{-4}]
Baseline frequency of sports	-	-	0.025	[-0.149, 0.191]
λ	0.115	0.079	6.127	[3.419, 9.164]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.
[3] For each individual, the information of the latest available year (out of 5, 6 and 7) is used for estimation.

Table E.2
Lasso and Bayesian Lasso coefficient estimates for the placebo control group.

Dependent variable: Health satisfaction at $t = 5$, $t = 6$ or $t = 7$
Number of observations: 285

	Lasso		Bayesian Lasso	
	β_{CV}	β_B	β	95% CI
Intercept	5.539	5.612	5.530	[3.793, 7.084]
Age birth	-	-0.009	-0.034	[-0.094, 0.018]
Years of education	-	-	0.035	[-0.039, 0.111]
Baseline employment level	-0.019	-0.071	-0.101	[-0.452, 0.234]
Marital status	-	0.010	0.115	[-0.112, 0.345]
Baseline health satisfaction	0.210	0.237	0.285	[0.172, 0.397]
ΔLabor earnings	-0.531×10^{-5}	-0.779×10^{-5}	-0.148×10^{-4}	[-0.371×10^{-4}, -0.905×10^{-5}]
ΔWork hours	-0.209×10^{-4}	-0.336×10^{-4}	0.554×10^{-4}	[-4.293×10^{-4}, 3.042×10^{-4}]
Baseline frequency of sports	-	-0.005	-0.078	[-0.274, 0.095]
λ	0.147	0.092	6.116	[3.511, 9.196]

[1] Only coefficients that are estimated non-zero are shown, a - indicates that the model estimates the coefficient non-zero.
[2] The 95% CI is the equal-tailed 95% confidence interval of Bayesian Lasso estimates.
[3] For each individual, the information of the latest available year (out of 5, 6 and 7) is used for estimation.

F Programming code

The programming code used to get the results shown in this thesis is attached in a zip-file. This zip-file contains two main folders: Child penalties and Health effects. The former includes all files used for estimating child penalties, and the latter has all files used for estimation of health effects and treatment effects. This section discusses the different files.

F.1 Child penalties

Estimation of child penalties using event studies is done in R. Before actually performing the event study, some sample selection is done. The folder 'Data cleaning' contains the three files that are used for this:

- *preprocess_data*: This file performs some pre-processing on the data. Specifically, it moves up year-specific variables such that they correspond to the year for which labor earnings are reported.
- *sample_selection*: This file creates the main estimation sample of individuals, according to the sample selection criteria discussed in Section 3.1.

- *select_sample*: This file introduces a function that can be called when a sub-sample of observations is needed for estimation of penalties.

The folder 'Event study' contains the files that actually calculate child penalties and relative earnings. This folder consists of six files:

- *samples_eventstudy*: This file can be seen as the 'main' file of the event study. The file calls different function to calculate results for different sub-groups of the population.
- *eventstudy*: This file introduces the function that is called when child penalties and relative earnings should be calculated. This function is then called in the *samples_eventstudy* file.
- *eventstudy_workhours*: This file introduces the function that calculates the child penalty in terms of work hours.
- *eventstudy_health*: File that runs the event study regressions with health satisfaction as the dependent variable.
- *eventstudy_graphs*: File that introduces different functions that create plots. These functions are all called separately in the *samples_eventstudy* file.
- *avg_health_satisfaction*: This file shows the code needed to create Figure 3.

F.2 Health effects

The study of health effects is performed in Stata. I chose to switch programming languages since Stata provides user-friendly packages for Bayesian Lasso estimation, with features that are more suitable for this thesis than the R packages doing the same. This folder contains four sub-folders introducing a different part of the estimation procedure. The first one is 'Data cleaning', which includes the following files:

- *Create dataframe*: File that uses the original data (in long format) for creating a data frame in wide format that can be used in further steps. This new data frame has one row per individual, with information of different survey years attached to it.
- *Complete dataframe t5*: This file creates the data that can be used for estimating health effects five years after childbirth. This involves deletion of observations with missing values, as well as calculation of new useful variables.

- *Complete dataframe t8*: This file is similar to the previous one, except that it creates this frame for eight years after childbirth.

- *Complete dataframe for ATE*: This file performs similar steps as the two previous files, but it also deletes all observations for which the first child was born before the start of the sample period. Moreover, the file makes sure that the end observation included is always the latest available year (between five, six en seven years after childbirth).

- *Complete dataframe ATE placebo*: File that is similar as the former one, except that it creates a sample around the placebo intervention of 2000.

The folder 'Lasso regressions t5' contains different files that perform (Bayesian) Lasso regressions for outcomes at time $t = 5$. The three main files in this folder are the following:

- *lasso_t5_health_satisfaction*: This file performs Lasso and Bayesian Lasso regressions for the full sample of females using health satisfaction at $t = 5$ as dependent variable. The file also contains lines of code that select the value of λ such that the estimates β_B can be obtained.

- *lasso_t5_delta_earnings*: File that performs similar regressions as the previous one, but with the change in earnings as dependent variable.

- *lasso_t5_delta_hours*: File that performs similar regressions as the previous one, but with the change in annual work hours as dependent variable.

This folder also contains files that perform regressions similar as in *lasso_t5_health_satisfaction*, but for different sub-samples of the population. In the folder 'Lasso regressions t8', similar files are also included, but with the outcome variables for $t = 8$ included.

The final folder is 'ATE estimation', which stores the files used for calculation of treatment effects. Specifically, the following files are included:

- *treatment_control_lasso*: File that runs Lasso and Bayesian Lasso regressions for both the treatment and control group, and finds the best values of λ.

- *treatment_control_lasso_placebo*: Similar as the former file, except that it works around the placebo intervention.

- *lasso_ATE*: This file calculates the Lasso-adjusted ATE based on the β_B coefficient estimates.

- *variance_ATE*: This file calculates the variance of the Lasso-adjusted ATE.

- *OLS_ATE*: File that calculates the OLS-adjusted ATE.
- *OLS_ATE_placebo*: File that calculates the OLS-adjusted ATE for the placebo intervention.
- *variance_ols_ATE*: File that calculates the variance of the OLS-adjusted ATE.
- *ATE plots*: File that creates the boxplot of ATE estimates.