Children and household earnings

Using an event-time analysis to see how the birth of the first child influences the gross monthly earnings and hours of labor supply of the husband and wife within the household

Master Thesis

J.A. Verschuure
SNR. 2051002

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Econometrics and Mathematical Economics

Tilburg School of Economics and Management
Tilburg University
Supervised by:

Assistant prof. A.Theloudis Tilburg University

June 5, 2023

Executive summary

This thesis was written over a 7-month period at Tilburg University. In this thesis I conduct an event-time analysis to see what is the effect of the birth of a first child on the earnings within the household. To this end, I track individual households over the period before, during and after childbirth. I find that for the husband, the birth of the first child has no impact on monthly earnings. For the wife, I find that the earnings decrease in the years after child birth. When moving to finer granularities, I find that the 'child-penalty' imposed by the birth of the child lies mainly in the period of 4 to 6 years after child birth. I also find that this penalty is mainly a one-time permanent drop in earnings and stays constant after that. After our analysis, I make a recommendation for future research as to why this penalty is imposed on women. I question how long this effect remains and why the view on the traditional household (where the husband is generally financially responsible) remains.

Acknowledgements

First and foremost I am extremely grateful to my supervisor. I would like to thank assistant Prof. Alexandros Theloudis for his invaluable advice, our engaging meetings, interesting discussions and excellent guidance during the process of writing my thesis. I would like to extend my heartfelt appreciation to Alexandros Theloudis for his invaluable contribution to the completion of this dissertation. Alexandros' open-mindedness, willingness to help out and patience have been crucial to the success of this research. His insightful feedback, constructive criticism and suggestions have helped shape my ideas and arguments. Additionally, his generosity in sharing his expertise and resources has been immensely helpful. His patience and positive attitude have made this research journey smoother and more enjoyable. Finally, I would like to thank my friends and family for their endless support and ideas during the process of writing this thesis.

Contents

1 Introduction

Over the last decade, the number of women entering the labor market is consistently increasing. In the twentieth century it was very more common for the wife to stay at home. Women were expected to take care of household chores such as cleaning, cooking and taking care of the children while men would be working a full-time job to provide for the family. But the household dynamics have been steadily changing in the last century. More and more women are entering the labor market which leads to a change in the conservative view that has previously been held about households. The dynamics within a household are changing which leads to a re-allocation of time within a household.

There are many factors at play when it comes to deciding on who is going to work within a household. Will the traditional view of the husband being the provider be adhered to (within a traditional male-female household), will the woman be the one to provide financially or will both members within the household work? And if so, do they both work full-time or part-time? Figure 1 below shows the evolution of the female labor participation rate in the Netherlands for the period 1971-2016. In this 45-year period, female labor participation rates have increased from 14% to 59%. Although not as extreme, similar results can be seen for Germany, the United States and the United Kingdom.

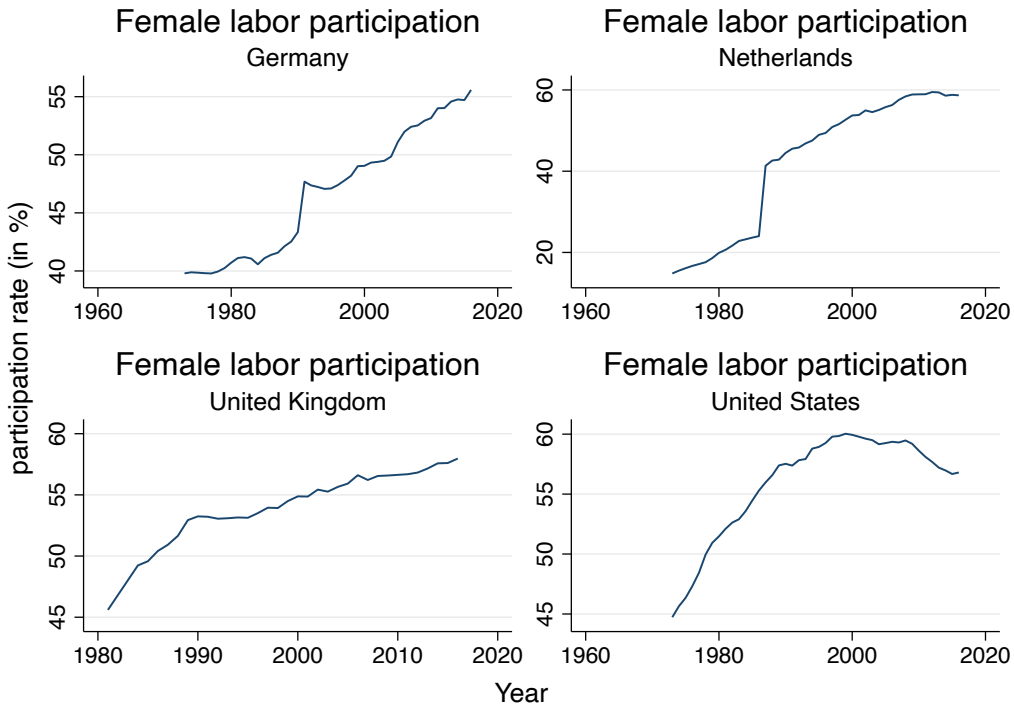

Figure 1: Evolvement of female labor participation in the Netherlands in the period 1971-2016. Important detail is that jump around approximately 1985. This is arguably due to a change in measurement and not an organic increase. Source: www.ourworldindata.org

The labor market is known to favor white men as they tend to have the biggest opportunities on the labor

market when it comes to high-paying jobs (see for example Blau & Kahn (2017)). But also between men and women, the discrimination on the labor market is widely documented. Men are known to be more likely to get a raise, get accepted into leadership positions and often are dominant in executive level functions, as indicated in the quarterly report of Parallelle Finance on gender equality within the S&P500[1]. They are thus more likely to be the financial provider for the family. But with a changing landscape in female labor supply and a more dominant fight for equality between men and women, the once obvious choice for men to go to work and women to take care of the household is changing. In the last few years, governments have increasingly tried to persuade women to enter the labor market (and with succes)[2,3]. They did so by introducing multiple policies to make it more appealing for women to start supplying their labor such as making it easier to take time off (e.g. to take care of the children) or by providing a financial contribution to the cost of childcare[4]. These are all aimed at making it more appealing to enter the labor market and find alternatives for the household chores. In doing so, the government aims to make it easier for women to enter the labor market and thus increase the total labor supply in the economy.

By entering the labor market, there is one question that immediately comes to mind: How many hours of labor will women supply in addition to men supplying labor? Earnings clearly depend on the amount of hours worked and thus are nothing but an extension of a similar analysis. Because of this, I will first focus on hours as the dependent variable. As evidence suggests, women and men work at similar rates (in terms of hours). On the contrary, the birth of the first child seems to disrupt the labor supply as well as the earnings of women. I make an attempt to quantify this effect by conducting an event-time analysis to see how the amount of hours of labor supplied changes in the event of the first child being born for the household. I will track households over time to see how the amount of labor supplied changes for both men and women within the household. I then extend this analysis to substitute hours for earnings to see how earnings change over the period in which the event-time occurs, for both men and women.

A similar study has been conducted by Henrik Kleven, Camille Landais and Jakob Egholt Søgaard in their paper 'Children and Gender Inequality: Evidence from Denmark' which was published in 2019 in the American Economic Journal. Kleven et al. (2019) also conduct an event-time analysis in which they track men and women within the household over time. Rather than starting with hours as dependent variable, they solely look at the earnings of the men and women within the household. Kleven et al. (2019) find that the earnings of men are upwards trending and do not significantly change during the event (the birth of the first child) or in the years after the event. On the other hand, they do find that there is a 'child-penalty' (defined as the amount by which women fall behind men due to having a child) for women which lowers their earnings by roughly 20% over the period 1980-2013.

So whereas the paper written by Kleven et al. (2019) only uses earnings, we will also focus on the hours of labor supplied as dependent variable as well as consider the effects in a different country. A logical forthcoming question then is what a possible explanation might be of a difference in reported earnings for men and women. There could very well be various (different) individual-factors at play that influence the effect on the reported hours (and thus earnings) of labor supplied around (and after) the event of child birth for men compared to

[1] https://parallellefinance.com/wp-content/uploads/2022/09/Q222GLI$_F ull_1$40922.*pdf*
[2] https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werkenden/arbeidsparticipatie-naar-leeftijd-en-geslacht
[3] https://www.rijksoverheid.nl/onderwerpen/vrouwenemancipatie/arbeidsparticipatie-van-vrouwen
[4] see https://www.government.nl/topics/child-benefit/applying-for-child-benefit

women. It does not necessarily mean that the difference can be attributed to discrimination but rather be a personal decision. This is an important consideration to keep in mind that I will address later in the paper.

Following a similar approach as Kleven et al. (2019), I was expecting to find similar results for women and men. Although my approach differs slightly (I will go into specifics on how my approach differs from the approach in Kleven et al. (2019) in section 3), I was able to find similar results. I find that the gross monthly earnings of the husband is increasing over the duration of the event-time period. I hereby compare the earnings of the husband before and after the birth of the first child. This result holds for all three levels of granularity of event-time. For the wife, I find that there is an upward trend in gross monthly earnings until before the birth of the first child and during the year of the birth of the first child. In the years after the birth of the first child, there seems to be a one-time drop in earnings. Again, this result holds for all three levels of granularity of the event-time.

Finally, I give an overview of the section of this paper. Section 1 forms the introduction to our analysis. In section 2, we will dive deeper into the literature written on the topic of earnings for men and women within the household, both in the Netherlands and internationally. In section 3 we will outline our methodology, explain our data and describe the necessary assumptions. In section 4 we will present our results. Section 5 will have recommendations for future research and we conclude in section 6.

2 Literature

The literature previously written on the topic of differences in earnings between men and women in the labor market is very extensive. Despite overall convergence, there is still a large gap in earnings between men and women within the household. Although the degree of pay gap differs between countries, globally the average women earns approximately 20 percent lower than men[5]. Although an easy conclusion would be to assume that this effect is larger for poorer countries and smaller for rich countries (due to various socio-economic factors including a lower degree of emancipation of women), this is in fact not true. As shown in figure 2, the effect is actually quite large for many countries that are considered rich. In the literature written on this

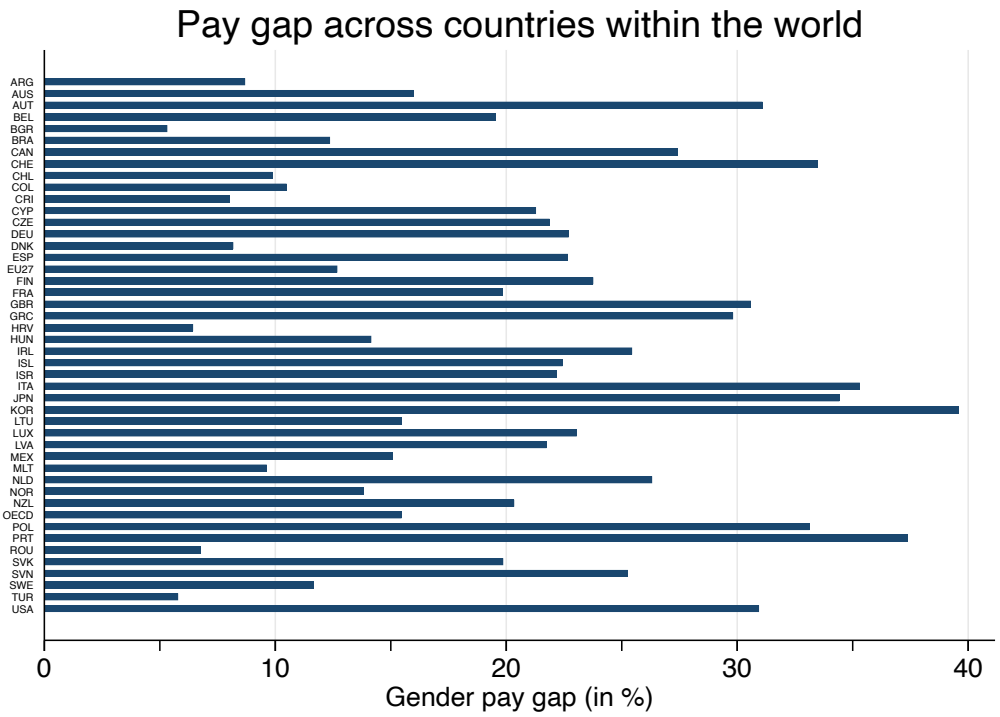

Figure 2: The pay gap in percentages between genders within the World. Source: https://data.oecd.org/earnwage/gender-wage-gap.htm

topic, various angles have been tackled.

While the list of possible explanations for differences in earnings and hours of labor supplied between men and women is extensive, I will limit my scope to explanations within the household. Specifically, I will focus on the impact of children, first on the hours of labor supplied within the household, and then on the earnings of the husband and wife. As mentioned in the introduction, this paper is based on a replication of the study performed by Henrik Kleven, Camille Landais and Jakob Egholt Søgaard and their paper 'Children and Gender Inequality: Evidence from Denmark' which was published in 2019. Their paper focuses on the impact of having children on the gender pay gap and the career progression in Denmark. They use administrative data from 1980-2013 for the full population of Denmark to conduct an event study analysis. Their analysis focuses on the impact of the birth of the first child on the change in earnings, labor partici-

[5]https://news.un.org/en/story/2022/09/1126901

pation, hours worked and wage rates within a household. They use the term 'child-penalty' as the change in earnings for men and women due to the birth of a first child. Specifically, it is the amount that women fall behind men due to having children. They find that the birth of the first child has a negative impact on earnings, labor participation, hours worked and wage rates for women, but not for men. They estimated the child-penalty for women's earnings in the year of the birth, relative to the year prior to child birth, of the first child to be 25 percent compared to men, meaning that the earnings of the women who had a child is approximately 25 percent lower in the year of child birth (relative to the previous year) compared to the earnings of men. In addition to looking at earnings, they also examine the amount of hours worked, the labor participation rates and the wage rates for women who did and did not have a child. They find that the amount of hours worked is approximately 12 percent lower in the year of child birth (relative to the year prior to child birth), the labor participation rate is 11 percent lower (relative to the year prior to child birth) and the wage rates are 10 percent lower (relative to the year prior to child birth) for women who had a child compared men. They then define the 'long-run child-penalty' as the long-lasting effect for women that have a child compared to men. The long-run child-penalty is the estimated difference for a variable between women who had a child and men, 10 years after the event happened. They find the long-run child-penalty for the earnings of women with at least one child to be 19 percent. Next, they find that the long-run child-penalty for hours worked is almost 10 percent, for labor participation is 13 percent and for wage rates is 9 percent, all for women with at least one child compared to men. These results are fairly robust and consistent with their findings when expanding the event-time window to 20 years after child birth.

A similar study was conducted by Angelov, Johansson and Lindahl (2016) in Sweden. They examine the effect of parenthood on the pay gep between women and men within the household. They use data from the Swedish Longitudinal Employer-Employee Data (LEED) register which tracks information on the Swedish labor force and earnings. Similar to the data used by Kleven et al. (2019), their data tracks data on income, labor participation and household characteristics across individuals and time. Their data shows a lot of similarities with the data used by Kleven et al. (2019) as both use data from the Nordic region. Parents in Sweden are allowed to take a maximum of 18 months of full-paid parental leave per child. Men and women can take on any combination of paid leave, as long as the total period does not exceed 18 months per child. Angelov et al.(2016) find that 80 percent of the paid parental leave is taken by women. In addition, they find that 44 percent of all women report to work part-time, whereas only 10 percent of men report to be working part-time. Angelov et al. (2016) find that the gap in earnings between men and women women in the year of child-birth is 135 log points larger than the gap between men and women in the year prior to child-birth. They thus find that the earnings gap is increasing. This effect is decreasing in the years after child-birth as the Swedish women gradually return back to the labor market and start supplying their labor again. They find that the effect never vanishes as 15 years after child-birth, the earnings gap is still 32.2 percent points higher as a result of having a child compared to the gap between men and women before having a child. In addition, they find that the pay gap is larger for parents than for non-parents, even after controlling for various variables such as education and occupation. They find that the effect of child birth on the earnings of the husband is lower than that for the wife. Specifically, they find that after 15 years, the wage gap has increased between men and women with 10 percentage points. Finally, they find that the wage gap continues to grow with each child being born.

This research has been extended to many countries in the paper "Child Penalties Across Countries: Ev-

idence and Explanations" by Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller (2019). They adopt the same methodology as used in the paper by Kleven et al. (2019) to look at the effect of parenthood on the earnings for men and women. Similar to the paper of Kleven et al. (2019), they find that the earnings for men and women evolve similarly before parenthood but start to diverge after parenthood. They find that women experience a large immediate drop in earnings after the birth of the first child. Men, on the other hand, seem to be unaffected in their earnings after the birth of the first child compared to the years prior to child birth. They also find some interesting differences. First, the long-run child-penalty (which is defined in the same way as in Kleven et al. (2019)) differs strongly between countries. They find that the long-run child-penalty for Scandinavian countries is in the range of 21 to 27 percent. For English-speaking countries, they find this range to be between 31 and 44 percent. Finally, for German speaking countries they find this penalty to be as big as 51 to 61 percent. Another interesting difference is that the short-run child-penalty differs substantially between countries. They find that the effect for Swedish women substantially larger than Danish mothers. In addition, Sweden is the only country in their study in which men also are affected by the birth of the first child (although this effect is only small and only in the short-run. In the long run there seem to be no consequences for the earnings of men).

These three papers are all very similar in the essence of their research: they find the extent to which there is a gap in earnings (or hours) between men and women within the household and suggest plausible reasons for this. As previously mentioned, much more has been written on the topic of gender wage discrimination. Bertrand (2011) specifically look at the differences between men and women in leadership positions. In her paper 'New Perspectives on Gender', she provides an overview of the differences for men and women in various economic factors such as labor force participation and earnings. In line with the results of Blau and Kahn (2016), she finds that women are vastly underrepresented in high-paying jobs and male-dominated fields such as finance, law and engineering. In addition, they are more likely to work part-time as opposed to men who generally work full-time. They find that the gender pay gap can only be partially attributed to various human capital factors such as education, age and work experience and is rather persistent across industries.

This dissertation has a similar place in the literature. I conduct an event-time study similar to the Kleven et al. (2019) paper for the households within the Netherlands. My data spans the period of 2009-2021 and is based on survey data, whereas that of Kleven et al. (2019) and Angelov et al. (2016) is based on administrative data. In addition, I will not only look at the earnings of the household during the event-time study, but also at hours of labor supplied by the household. Since earnings can vastly differ among households and occupation, using hours of labor supplied as a leading variable will give me a way to deal with the selection of deciding to work and, to an extent, be a good indicator for the effect of child-birth on earnings for men and women.

3 Data

In this section I will first describe the source of the data and the process of cleaning the data. This includes a step-by-step process of going from the raw data as provided by the LISS panel, to the cleaned data and generating the necessary variables based on my data to conduct the event-time study.

3.1 Data source

The first step to my analysis is deciding on which source data source I will use. To this end, there were two candidates that came to mind: either use the data that is provided by the CBS (Centraal Bureau voor de Statistiek) or the data that is available in the LISS panel. The most logical decision would be to use the data that is gathered by the CBS as it spans a longer period of time (although not the most useful data, it is available starting from the middle of the 20^{st} century whereas the LISS panel only has data available from the beginning of the 21^{st} century) and is considerably more extensive in data. Unfortunately, access to the CBS data is quite expensive whereas the LISS panel is an open-source dataset. Due to this, I have decided on using the LISS panel as my data source.

So what exactly is the LISS panel? The LISS panel ("Longitudinal Internet studies for the Social Sciences") is a Dutch longitudinal study consisting of almost 7500 individuals across 5000 households in the Netherlands that started in October 2007. The LISS panel is based on a true probability sample of households drawn from the population register by Statistics Netherlands and thus an fair representation of the Dutch population. The study gathers data on its core study (which is mostly on household data) as well as various other topics. Participants are asked to conduct a survey, taking approximately 15 to 30 minutes to fill out. The core part of the longitudinal study is repeated yearly and is designed to follow changes in the life course and living conditions of the panel members. The rest of the interview questions are used to gather data on various topics, ranging from data on linguistics to medical sciences.

As mentioned in the introduction, I will be conducting an event-time study to see what the effect is of the birth of the first child on the hours of labor supplied and earnings of the men and women within the household. The goal of this thesis is to analyse the effect of the birth of the first child on the gross monthly earnings for both the husband and wife within the same household. My analysis will therefore have a similar set-up as the paper by Kleven et al. (2019) for Denmark and Angelov, Johansson and Lindahl (2016) for Sweden. To this end, I resort to the data of the LISS panel that contains the necessary information on individual characteristics (earnings, age, hours of labor supplied) and household characteristics (whether married, number of children, when children were born etc.). Specifically, I focus on the data on gross monthly earnings of men and women, the hours worked for both men and women, household structure, whether participants have children (and if so, how many and when were they born), whether participants are married, where they live and how old they are. The data keeps track of both household identifiers and personal identifiers to be able to track participants over time who are within a certain household. As I will discuss later in this section, there are various issues that we must address before going conducting my event-time study. Since the LISS panel is a longitudinal study on a large number of variables and characteristics, a lot of data is available that we are not interested in. To this end, the most important sources of data within the LISS panel are the core studies 1 (background variables) and 5 (Family and Household). Finally, the data of individuals on reported hours of labor supplied will be taken from core study 6 (Work and Schooling).

3.2 Data cleaning and data transformation

As mentioned in the previous section, I will first discuss the cleaning of the data and a few common problems within the data set. As the goal of this thesis is to analyse the effect of the birth of the first child on the reported hours and earnings of men and women within the household, I will first assemble the data on these variables for both men and women within the household for the years prior to child birth, the year of child birth and the years after childbirth. The LISS panel records monthly data on individuals, tracked over multiple years. The LISS panel has quite a few missing observations and incomplete data in the first year. In addition, the last year (2022) has not been fully made public yet. Hence, I will only use data from the period of 2009-2020.

Next to that, my goal is to report the effect of the birth of the first child on the gross monthly earnings for the husband and wife. Since the data is given in waves (each wave representing one month in a given year) and not every participant fills out the survey every month for a whole year, I want to find the average earnings per month for a participant in a year. To do so, I sum the observations per participant that I do have. Then, I take the modal value of the observations of this participant in this year and replace the missing observations with the mode of that year. In case there are multiple modes, I take the maximum of the modes. This way, it is possible to discuss the earnings and hours of labor supplied on an average monthly basis per year. To track the data of individuals and households over time, the LISS panel also uses an unique, individually assigned number to follow the responses of the same participant and same household over time. Because we are only interested in the responses of both the husband and wife within the same household, we first drop all observations who respond to not have a partner. Next, we want to distinguish between the effect on earnings and hours of labor supplied for the husband and wife, hence we need participants to report their gender. We drop the observations of participants who do not report their gender, are of the same gender or the situation in which either household member refrains from specifying their gender.

Further, we drop all observations in which participants use a different household identifier over the years. This indicates that there was a change in the household situation which lead to the household being considered a 'new' or 'different' household than it was previously. We drop these observations because they no longer track the 'same' household (the 'original' household) over time and thus do not give us information on the change in gross monthly earnings for the original household. Following this, we transform the recorded monthly data into average yearly data by taking the mean over all observations within a given household and year. This leaves us with the average reported monthly hours of labor supplied and monthly earnings for both men and women, per year. This leaves us with a total of 13801 observations of 102 households over a time period of 12 years. Aside from the data transformations, there are also a number of data cleaning issues that I will further highlight and address before moving on:

- The household compositions that are not constant over time

This first issue has to do with a change in households. While we track households over a number of years, this does not mean that households cannot change over time. It is possible for individuals to start living together (and thus forming a new household, possibly (but not necessarily) with another participant in the

panel). In addition, a participating household might file for a divorce and thus no longer has the same characteristics as it had before (if both change because, for example, they both re-marry someone else, then there are no new observations for this household completely). This could potentially greatly influence my results as I try to capture the effect of child birth on the gross monthly earnings of both the husband and wife, which will be distorted. To solve this, we only track households for which we have data available in all three time segments (years prior to child birth, year of child birth and years after child birth).

- Missing observations of either the husband or wife within the same household

Similar to not all households consisting of a male and a female, it might be the case that there are missing observations within a given year of either the husband or wife. This means that we cannot track the change in earnings for this year. We deal with this by dropping the observations in which the response of either the husband or wife within a household is missing.

- The birth year of the first child might not be the same for the husband and wife

The birth year of the first child might not be the same year as the birth child of the first child of the other parent due to previously being divorced and/or already having children. This leads to the problem that the amount of children for one parent in the household does not coincide with the reported amount of children for the other parent within the same household. Because we are focusing on the effect of child birth (the same child) on earnings for the husband and wife, we solve this issue by taking the reported birth year of the first child as given by the mother as the actual birth year of the first child.

So what exactly is the data that I am interested in? I previously mentioned that I want to track households over time. In the simple model, I want to see what the effect is of the birth of the first child on the reported hours of labor supplied and the earnings for both the husband and wife within households. In other words: what is the effect of the event on the reported hours of labor supplied and the earnings in the year of the event, and the years after. To this end, it is important to track households over multiple years. Specifically, the goal is to track the earnings and hours of labor supplied of both the husband and wife in the years before the birth of the first child, the year of child birth and the years after the birth of the first child. We can thus define two different variations of time: actual time and event-time. Actual time is, as the name perhaps suggest, the actual year in which the data is recorded in the LISS panel. The actual time will thus be in the period of 2008-2021. On the other hand, event-time is time related to the event itself. This means that event-time is defined in the same way for all households, but not constant for all households. Take for example two households: A and B. Household A I observe in the period of 2012-2016 and household B I observe in the period of 2015-2018. Suppose household A had a child in the year 2015 and household B had their child in the year 2017. This means that the event took place in 2016 for household A but in 2017 for household B. We thus have that the event-time for household A will be -3 for the year 2012, -2 for the year 2013, -1 for the year 2014, 0 for the year 2015 and 1 for the year 2016. For household B, we will have that event-time will be -2 for the year 2015, -1 for the year 2016, 0 for the year 2017 and 1 for the year 2018. We are thus comparing event-time for different households and not calendar time.

Here is also where I observe the largest bottleneck in my analysis: the requirement to observe households before, during and after the event. A large part of the data set contains information on households after

event-time. This is because the data set only spans a period of 12 years. It is thus quite logical that a lot of my observations are of households for which the event (the birth of their first child) happened in the years prior to the start of the LISS panel (or before they started participating). After dealing with this issue and filtering out the observations of households that we do not observe at least one year before the birth of their first child, the year of the birth of their first child and at least one year after the birth of their first child, we are left with only 565 observations of my initial 13801.

3.3 Descriptive statistics

I will first show some descriptive statistics to get a feeling for the data set and possibly issues that I run into. Table 1 shows these descriptive statistics. The statistics are shown of an aggregate level (indicated by the column 'Event-time τ' with row 'All') and spread out based on event-time. All the values are rounded to the nearest integer. There are a few patterns that appear in the data that I will discuss. Looking at the earnings of the husband, there seems to be an upward trend in mean earnings. This is in line with my expectations and in line with the results of the paper written by Kleven et al. (2019). For the wife, there does not seem a upward trend in the years before child-birth compared to the years after child-birth. Looking at the amount of hours of labor supplied, the husband seems to be reporting a consistent amount of hours before, during and after the birth of the first child. The average reported hours is close to a full-time job (defining a full-time job as working 40 hours). For the wife, there seems a downward trend in the amount of hours of labor supplied. Whereas the average hours of labor supplied before the birth of the first child is almost 33 hours, the average hours of labor supplied after the birth of the first child is only 26 hours (both not separately reported in the table). For the husband, the average amount of hours of labor supplied before and after the birth of the first child is respectively 38 and 36 hours. This supports the conservative view that women generally start working less after the birth of the first child and instead spend more time taking care of the children. Another important observation is the number of observations per event-time. The further we are before the event happens (meaning that larger negative values for τ), the less observations we have. This is a logical conclusion as I filtered my original data set on the requirement that observe households at least one year before the birth of the first child, during the year of the birth of the first child and at least one year after the birth of the first child. This means that I have at least three observations for all households in my data set. It is not an unreasonable assumption to assume that households fill in the survey for a consecutive period of time. This immediately makes it clear why there are fewer observations for larger negative values of τ.

As discussed prior, we will first show the effects on the amount of hours of labor supplied in the years aroung the event period. These results can be seen in figure 3. For both the husband and wife, there is a pattern showing a decrease in average hours of labor supplied over the event-time period. The average hours of labor supplied for the husband is around 40 in the years before the birth of the first child whereas it seems to hover around 35 for the years after child birth. For the wife, we see that the average hours of labor supplied in the years prior to the birth of the first child hovers around 33, but this decreases sharply after the birth of the first child to around 27 hours. Following a similar approach as Mazzocco, Ruiz and Yamaguchi (2014), we also plot the average earnings for both the husband and wife over time. Figure 4 show the plotted mean earnings for the husband and wife in the years before the birth of the first child, the year of child birth and the years after the birth of the first child.

There are a few results to mention here. For the husband, we clearly see an upward trend in earnings

Variable	Event-time τ	n	Mean	Standard deviation	Min	Max
Husband - earnings in €	All	532	3220	1179	0	9655
	<-4	17	2385	1084	0	4883
	-3	21	2448	1019	522	4500
	-2	35	2717	963	275	5097
	-1	50	2798	967	940	5382
	0	88	3036	986	1364	6146
	1	71	3115	1140	1388	6601
	2	54	3268	1162	1400	6274
	3	51	3320	1070	963	6000
	>= 4	145	3787	1246	1683	9655
Wife - earnings in €	All	535	2126	1170	0	5100
	<= -4	17	1603	1133	0	4008
	-3	21	2006	975	0	4022
	-2	35	1971	1213	0	4388
	-1	50	2210	1036	0	4421
	0	89	2229	1000	0	4421
	1	72	1985	1142	0	4602
	2	54	2026	1176	0	4285
	3	51	2032	1267	0	4975
	>= 4	146	2288	1292	0	5100
Husband - hours of labor supplied	All	560	37	14	0	80
	<-4	19	37	14	0	60
	-3	23	44	8	0	60
	-2	38	34	16	0	52
	-1	51	38	11	0	50
	0	92	40	11	0	80
	1	74	38	14	0	80
	2	57	37	15	0	60
	3	53	33	19	0	60
	>= 4	153	35	13	0	60
Wife - hours of labor supplied	All	511	28	13	0	60
	<-4	16	30	19	0	55
	-3	19	35	13	0	50
	-2	35	31	15	0	60
	-1	48	34	13	0	55
	0	89	33	12	0	60
	1	69	29	11	0	50
	2	51	26	11	0	40
	3	48	26	13	0	45
	>= 4	136	24	14	0	50

Table 1: Descriptive statistics for the reported earnings and house of labor supplied for both men and women within the household. Statistics are shown on an aggregate level and split out by event-time and rounded to the nearest integer.

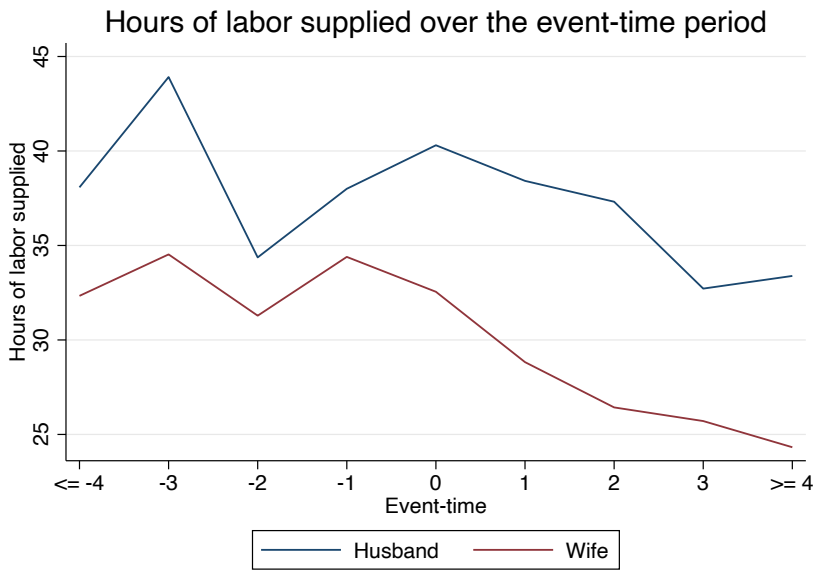

Figure 3: The reported average hours of labor supplied during the event-time period for both the husband and wife

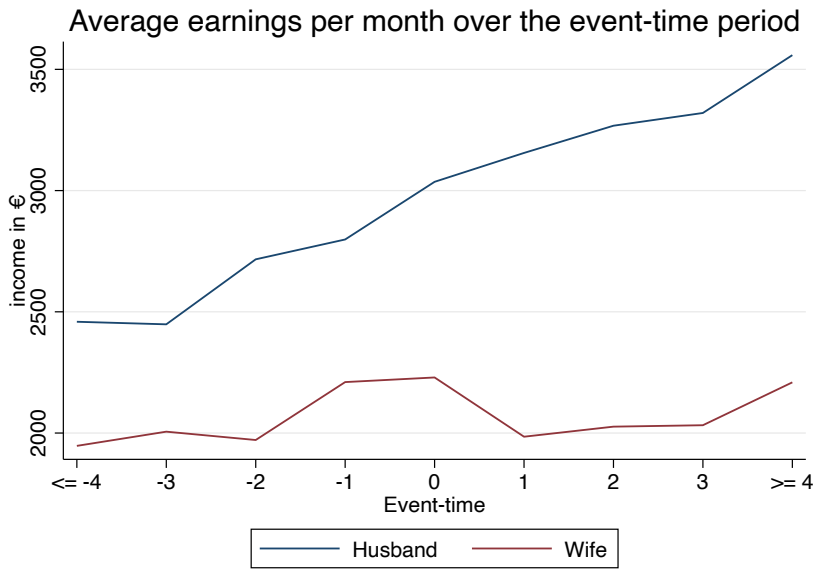

Figure 4: Average earnings in Euro's of the husband and wife over the event period in the household

across event-time periods. Even with a small sample size, this looks representative for the earnings of men based on their position in the labor market and the literature written on it previously. This upward trend can be explained because of the growth in earnings over calendar time. For the women, we see a downward trend except for a strong peak in the year prior to the birth of the first child. An explanation for this could be the small sample size which leads to this peak. Although surprising, there is no reason to assume that this peak would persist when increasing the sample size.

I will now first formally introduce the model and notation. Let Y denote the (log) gross monthly earnings for a participant in the data, i be the household indicator, t the time indicator (in years) and τ denote the event-time (in my analysis this is the birth year of the first child). This means that if a certain household had their first child in 2015, then τ would take on the value -1 for 2014, 0 for 2015, 1 for 2016, 2 for 2017 etc. We then generate dummies for different levels of granularity of event-time within my data (and excluding one year-time dummy to avoid multicollinearity). Letting τ represent the dummy for a chosen granularity, G the set of event-time granularities and Q the set consisting of sets of event-time granularities, we can then define τ as follows:

$$\tau \in G \subset Q \tag{1}$$

I will then repeat my analysis for different granularities. This way it will be possible to see the effect of the event on the reported hours of labor supplied and earnings for both the husband and wife for different periods of time after the event has occurred. More specific, I define Q as the set with different sets of granularities so that event-time $\tau \in$ {before, during, after}, $\tau \in$ {-6/-4, -3/-1, 0, 1/3, 4/6}, or $\tau \in$ {-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5}.

In addition, we introduce dummies for calendar time as to control for time. Let T represent the set of different levels of granularity over which we have data. Next to controlling for year and event-time there is sufficient empirical evidence to conclude that age has a significant effect on earnings too, hence why I will control for age as well. Finally, we include an error term and a constant and we thus end up with model 2. Note that we are not necessarily running the regression based on this model for both the husband and wife. In section 4, I will first determine whether to use log earnings or level earnings, both for the husband and wife.

$$Y_{it} = \alpha + \sum_{t \in T} \delta_t \cdot \mathbb{1}[\text{year} = t] + \sum_{\tau \in G} \beta_\tau \cdot \mathbb{1}[\text{event-time} = \tau] + \theta \cdot age_{it} + \kappa \cdot age_{it}^2 + \epsilon \tag{2}$$

3.4 Methodology

First I will discuss how to estimate the model and specifically, the form of the earnings and the amount of hours of labor supplied. Both the hours worked as well as the earnings of households (either the husband, wife or both) have a number of 0's as reported hours/earnings. The standard approach in the literature is to estimate the earnings in logs. Since the log of 0 is undefined, we have to take a careful approach as to whether the estimate the earnings of the husband and wife in logs (and consider the observations with value 0 as missing) or just use the estimation in levels.

To decide on how to continue, I will first define a dummy variable relating to the reported hours that a participant is working. We define the variable *working* as a dummy variable which is equal to 1 if the reported hours of labor supplied is positive and 0 otherwise. I then run a regression on this variable, using the same independent variables as in my model specification. If I do not see a change in the labor market participation before and after the event-time, then I will use logs rather than levels (meaning the 0's do not impact the decision on whether to enter the labor market or not). If there is a change in labor market participation, then I will continue using levels.

I will now first discuss the estimation method that I use in my data analysis and the necessary assumptions to use these. We are trying to capture the variation that is caused in the amount of hours of labor supplied and in gross monthly earnings for the husband and wife within the household after giving birth to their first child. The main identifying assumption in this analysis the assumption that the timing of child-birth is random and does not influence or relate to any prior intention of the wife to reduce her labor participation.

4 Results

We will now report the results of our analysis based on different levels of chosen granularity in the event-time specification. Before that, we give a bit more descriptive statistics of our remaining dataset and the chosen specification.

4.1 Results on hours

We will first look at the results of the regression to determine whether to use logs or levels for earnings. As mentioned in section 3.4, we run the regression specified in equation 3. As mentioned previously, the log of 0 is not defined. To determine whether the reported zeros make a significant effect, I first regress the independent variables on a dummy variable which is equal to 0 or 1. If the results do not change around child birth, then we can continue and use logs instead of using levels. If the results do change then it would be better to stick to levels anyway. The results can be seen in figure 5.

$$Y \equiv \mathbb{1}[\text{reported hours} > 0] = \alpha + \sum_{t \in T} \delta_t \cdot \mathbb{1}[\text{year} = t] + \sum_{\tau \in G} \beta_\tau \cdot \mathbb{1}[\text{event-time} = \tau] + age_{it} + age_{it}^2 + \epsilon \quad (3)$$

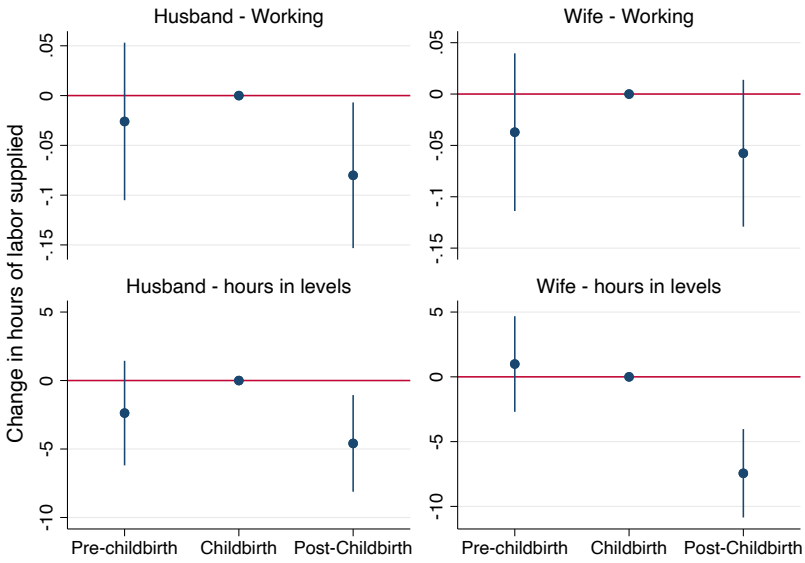

Figure 5: The plotted coefficients for β after running equation 3. I omitted the variable relative to the year of child-birth. Hence the plotted β's are the estimated coefficients relative to the year of child-birth.

I will first discuss the results for the husband. Looking at the top left graph, we see the plotted β's of regression 3 for the husband. From the picture it is clear that there is a significant drop in reported hours of labor supplied by the husband in the years after child birth. From this I conclude that the reported 0's do change the effect around child birth, hence I decide to use levels rather than logs. The same pattern can be seen in the lower left graph. This graph shows the plotted coefficients of regressing the earnings of the husband (in levels) on the independent variables as given in equation 3. This graph shows that there

is a significant drop in reported hours of labor supplied in the years after the birth of the first child. Once again, reinforcing the idea to keep using levels rather than logs. Moving to the top right graph, there are the plotted β's of regression 3 for the wife. The graph shows that there is a decrease in reported hours of labor supplied in the years before the birth of the first child and also in the years after the birth of the first child. Once again, it is clear that there is a change in reported hours and thus the 0's are important to keep in our analysis. This means both graphs have shown that it is a better practice to use levels here, rather than logs. To confirm the analysis, take a look at the bottom left graph. From the graph, one can quickly see that there is a significant drop in reported hours of labor supplied in the years after the birth of the first child. This confirms the conclusion to use levels rather than logs. We thus end up with the following estimating equation for estimating the effect on gross monthly earnings around the event-time period for both the husband and wife:

$$Y_{it} = \alpha + \sum_{t \in T} \delta_t \cdot \mathbb{1}[\text{year} = t] + \sum_{\tau} \beta_\tau \cdot \mathbb{1}[\text{event-time} = \tau] + age_{it} + age_{it}^2 + \epsilon_{it} \tag{4}$$

4.2 Results on earnings

4.2.1 Three event-time dummies

We will start of with analysing the most coarse granularity. This means we will only split up the event-time dummies in three periods: the years prior to child birth, the year of child birth and the years after child birth. Looking at the paper of Kleven et al. (2019) and Angelov et al.(2016), we expect to find two things. First, we expect to find that there is no difference in the pattern of gross monthly earnings between husband and wife in the years prior to childbirth (meaning they are both increasing over time). Secondly, we expect to find no significant decrease in gross earnings for the husband after childbirth, but do find a decrease in gross earnings for the wife. We leave out the event-time at time $\tau = 0$ (meaning the year in which the child was born) to avoid multicollinearity. These results can be found in figure 6 below.

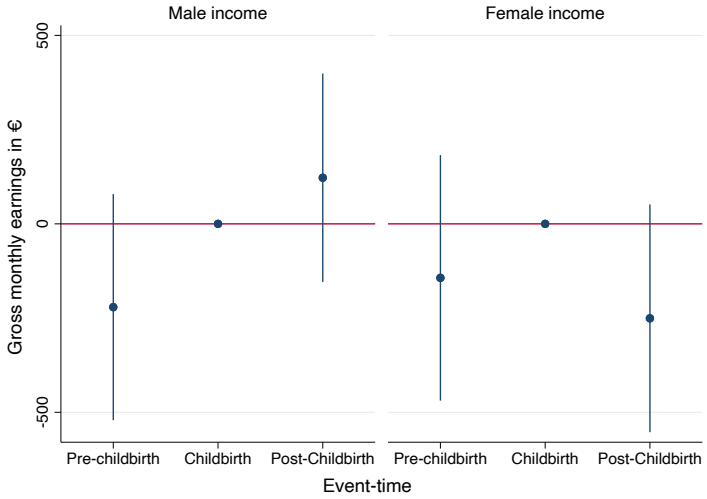

Figure 6: The estimates of β in regression 4 for husband and wife. The estimated coefficients are plotted relative to the year of child birth (event-time zero)

Plotted are the coefficients and the standard errors of the estimated β parameters. Note that these are

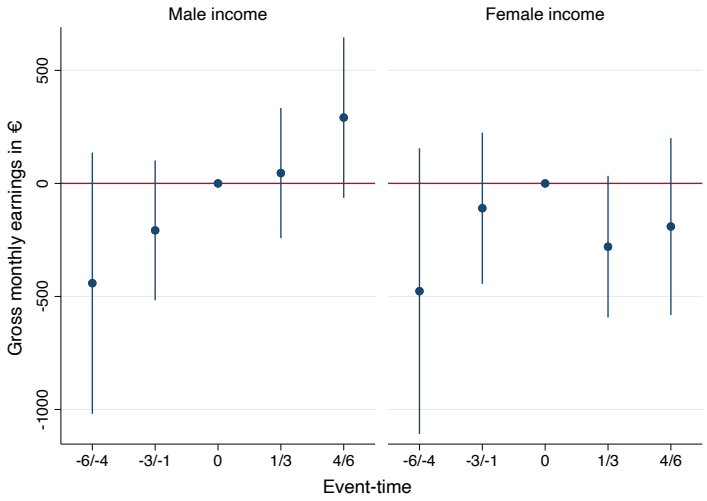

Figure 7: The estimates of β of regression 2 on the birth of the first child on the earnings for a husband and wife for three-year event-time dummies

only the observations of households which we track both before and after the birth of the first child. As mentioned previously, the data set is rather small. This means that I am left with only 565 observations of the original 13801, each corresponding to earnings of a husband and wife within the household, for that specific year. In addition, these observations are not unique households but instead observations per year (and thus multiple observations make up one household). From figure 6 we can see the plotted estimates of the parameters (given by the dot) and the confidence interval for both men and women. We see that for men, the estimated β's are lower before child birth and higher after child birth, which is in line with the expectations following literature previously written. Similarly, the estimated β coefficient for women are lower before and after child birth. This is rather surprising, considering the effects are given relative to the year of child birth. The estimated confidence intervals are large too, which can be attributed to being a consequence of the small sample size. Nonetheless, there is clearly an upward trend for men which is not the case for women.

4.2.2 Three-year event-time dummies

We will now split up the analysis as we did in section 4.2.1 to a bit finer granularity. We group the event-time periods into three-year event-time dummies and regress those on the log earnings of the husband and wife. Just as we did in section 4.2.1, we leave out the event-time dummy corresponding to $\tau = 0$ (the year of child birth) so we can compare the results relative to the year of child birth. We follow the same approach as the Kleven et al. (2014) paper and only keep the event-time dummies that represent the periods before and up to ten-years after child birth. The results of the analysis are found in figure 7. The figure shows results that are in line with our more coarse granularity of event-time. For the husband, we see that there is an upward trend in average earnings in the years before and after the birth of the first child. Again, we see that the earnings of the husband in the first and second three-year period after child birth is higher than the year of child birth. This is in line with the results of our first analysis. For the wife, we see somewhat different results. We see that the estimated earnings for women in the first and second three-year period before the birth of the first child is lower than in the year of child birth. This is similar to the results for

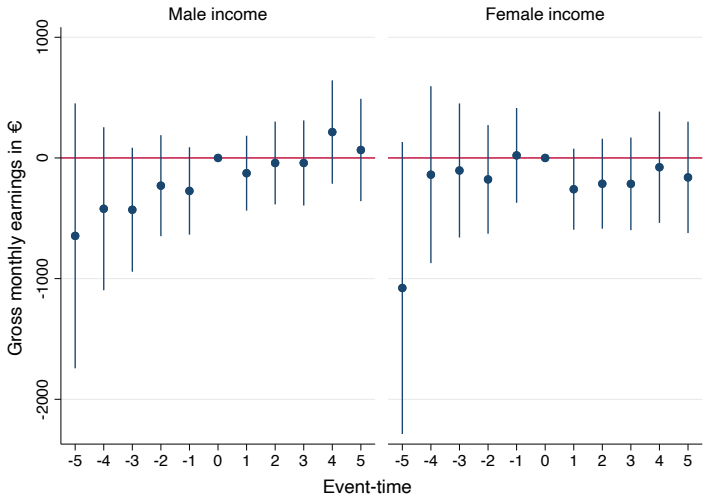

Figure 8: The estimates of β of regression 2 on the first child on the (log) earnings for a husband and wife for one-year event-time dummies

the husband. On the contrary to the results of the husband, there is a drop in estimated earnings in both the first three-year period after the event and the second three-year period. This result is in line with the results of the Kleven et al. (2019) paper as well as the Angelov et al. (2016) paper. Just like we did before, There are rather large standard errors for both the husband and wife which I again attribute to the rather small sample size.

4.2.3 One-year event-time dummies

The study performed by Kleven et al. (2019) uses a very fine granularity, even extending equation 4 to include dummies for age. In addition, they base their results on dummies representing yearly effects (i.e. every dummy represents one year). Although it is desirable to do so as well, the LISS survey has not been around for a very long time. This lead to quite a skewed distribution of observations relative to event time as we have many more participants in the survey that report their earnings after child birth, than before. Because of this, we decide to split up the sample further into one-year event-time dummies. The results can be seen in figure 8 Just as before, we see that the plotted coefficients for the earnings of the husband is steadily increasing over the event-time period. There does seem to be a few years after the birth of the first child (specifically event-time τ = 1-3) in which the earnings is slightly dropping. This could perhaps be attributed to the small sample size. Similar to before, the standard error is rather large again which is another indicator of a small sample size. For the wife within the household, we again see similar results as in section 4.2.2. Although there is an upward trend in earnings for the wife in the years prior to the birth of the first child, the yearly change is substantially smaller than the earnings of the husband. The earnings also seems to suffer from a one-time permanent drop in earnings in the years after the birth of the first child. Comparing τ = 1-3 to $\tau = 0$, the first three years after the event all have near-identical earnings. But these earnings are all lower than the year of the birth of the first child and also lower than the years prior to the birth of the first child. This is again in line with the results of the Kleven et al. (2019) and Angelov et al. (2016) papers, indicating that the results are rather robust across countries. In addition, we see quite large standard errors again (for both the husband and wife). Once again, I will attribute these large standard

errors to the small sample size. In the next section, we will test our results using a different sample selection. There, we will loosen the requirement to observe households before, during and after the year of child birth.

5 Future research

I will now start making recommendations for future research and points to expand on in future research. As mentioned quite extensively in the results section, my results all have quite large confidence intervals which is a consequence of the lack of observations. Replicating my research for the Netherlands with a larger sample size, I expect to get similar estimates but smaller confidence intervals. Additionally, I only had data available from 2008 until 2021. There is reason to believe that major economic events within this time-period influences the earnings of both the husband and wife (not necessarily in the same direction). It would therefore be interesting to replicate the study but with more data from the twentieth century and see whether the results would change. There are various extensions to make to my analysis that I will quickly highlight.

To start, I mentioned that there are various major economic events in a short period of time that influence the earnings of the husband and wife. The period 2008-2021 saw, among others, the fall of Lehman Brothers and the forthcoming crisis and the Covid-19 crisis. These are all major economic events which affect the unemployment of people. The consequence of crises is generally the letting go of employees, mostly the partially employed but also part of the full-time employees. Since women are more likely to work part-time, this could influence the earnings in an uneven way. It would therefore be interesting to extend the research to data in the twentieth century so it would be possible to split up the event-times based on major economic events. This way, it would be possible to account for these and reduce some of the variation due to macroeconomcis variables.

Another common conclusion was that the gender pay gap is especially dominant in male-dominated fields such as law, engineering and finance. If men are generally more likely to work in high-paying fields then they are likely to get paid more for working full-time than a women who also works full-time, but in a field that pays less. It would thus be interesting to split up the sample into data based on profession or field. If women do work in fields that are paying less than the fields men are in, then the difference in pay could be attributed to the field they are working in rather than discrimination between men and women within the same field.

6 Conclusion

I used an event-time study approach to look at the effects of child birth on the evolution of earnings of the husband and wife within the same household. To this end, I assembled data on earnings within the same household from the LISS panel. Due to the nature of the panel, we were left with a rather small sample of observations. Due to this, all my estimates have large standard errors. I thus focus on the estimate within the households and will ignore the large standard errors for now. I found that in the simple model the earnings of the husband were increasing through the period of childbirth. For the wife on the other hand, we found that the income is lower in the year of childbirth relative to the years prior and is also slightly lower after childbirth compared to before. When moving to a finer granularity, we see that this effect is retained, but seems to occur for the wife specifically in the second three-year period after childbirth and not in the first three. This is an interesting result and could call for further analysis. When finally moving to a yearly granularity, we see that the income of the wife is lower in the few years after childbirth compared to before, but it looks like a one-time drop.

Going back to the observations of households that are observed before and after childbirth, our findings are in line with our own expectations as well as the results of Kleven et al. (2019) and Angelov, Johansson and Lindahl (2016). From here on out, it would be useful to find a remedy or explanation for why this gender pay gap remains. After accounting for various variables, we still find that the drop in income for the wife is significant whereas there is no drop for the husband. One explanation is that households remain conservative in the allocation of time. The wife is still the main caretaker of the children and the husband the one providing financially. Although more policies and regulations are being imposed, this conservative view of the household seems to be hard to break through. I end my analysis with the open remark that only time will tell what (and if) this conservative view of the household will change.

References

[AB99] Joseph G. Altonji and Rebecca M. Blank. "Chapter 48-Race and Gender in the Labor Market". In: *Handbook of Labor Economics*. Vol. 3. Elsevier, 1999, pp. 3143–3259.

[ABM18] Jorge M. Agüero, Prashant Bharadwaj, and Emily M. Oster. "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks". In: *Journal of Political Economy* 126.3 (2018), pp. 1057–1101.

[ADS14] Jérôme Adda, Christian Dustmann, and Katrien Stevens. "The Career Costs of Children". In: *Journal of Political Economy* 122.2 (2014), pp. 222–283.

[AJL16] Nikolay Angelov, Per Johansson, and Erica Lindahl. "Parenthood and the Gender Gap in Pay". In: *Journal of Labor Economics* 34.3 (2016), pp. 545–579.

[Ber11] Marianne Bertrand. "Chapter 15-New Perspectives on Gender". In: *Handbook of Labor Economics*. Vol. 4. Elsevier, 2011, pp. 1543–1590.

[BK17] Francine D. Blau and Lawrence M. Kahn. "The Gender Wage Gap: Extent, Trends, and Explanations". In: *Journal of Economic Literature* 55.3 (2017), pp. 789–865.

[Blo+10] David E Bloom et al. "Fertility, female labor force participation, and the demographic dividend". In: *Journal of economic growth* 15.3 (2010), pp. 79–101. DOI: 10.1007/s10887-010-9055-0.

[Fin22] Parallelle Finance. *Global Labor Market Outlook Q2 2022*. https://parallellefinance.com/wp-content/uploads/2022/09/Q222GLI_Full_140922.pdf. Accessed: 08-01-2023. 2022.

[Gol14] Claudia Goldin. "A grand gender convergence: Its last chapter". In: *American Economic Review* 104.4 (2014), pp. 1091–1119. DOI: 10.1257/aer.104.4.1091.

[Hoy] Donald Tomaskovic-Devey Eric Hoyt. *Race, States, and the Mixed Fate of White Men*. Accessed: 02-02-2023. URL: https://www.umass.edu/employmentequity/race-states-and-mixed-fate-white-men.

[Kle+19] Henrik Kleven et al. "Child Penalties Across Countries: Evidence and Explanations". In: *The Quarterly Journal of Economics* 134.4 (2019), pp. 1865–1912.

[LR18] Shelly Lundberg and Elaina Rose. "Parenthood and the earnings of married men and women". In: *Labour Economics* 51 (2018), pp. 70–81. DOI: 10.1016/j.labeco.2017.12.008.

[MRY14] Marta Mazzocco, Isabel Ruiz, and Shoko Yamaguchi. "Labor Supply and Household Dynamics". In: *Journal of Business & Economic Statistics* 32.2 (2014), pp. 221–231.

[Netnda] Government of the Netherlands. *Applying for child benefit*. https://www.government.nl/topics/child-benefit/applying-for-child-benefit. Accessed: 12-01-2023. n.d.

[Netndb] Government of the Netherlands. *Women's labour force participation*. https://www.government.nl/topics/gender-equality/womens-labour-force-participation. Accessed: 18-03-2023. n.d.

[Nie] Marcia Nieuwenhuis. *Witte heteroman nog altijd favoriet op werkvloer*. Accessed: 12-03-2023. URL: https://www.parool.nl/nederland/witte-heteroman-nog-altijd-favoriet-op-werkvloer~b5caa5d6/?referrer=https%3A%2F%2Fwww.google.com%2F.

[OP16] Claudia Olivetti and Barbara Petrongolo. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries". In: *Journal of Economic Perspectives* 30.1 (2016), pp. 205–230.

[Wal98] Jane Waldfogel. "Understanding the 'Family Gap' in Pay for Women with Children". In: *Journal of Economic Perspectives* 12.1 (1998), pp. 137–156. DOI: 10.1257/jep.12.1.137.