ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis: Economie en Bedrijfseconomie

Assessing The Role of Religiosity in The Child Penalty in The Netherlands

Name Student: Sandra Bloem

Student ID Number: 477361

Supervisor: Sara Rellstab

Second Assessor: Lisa Voois

Date Final Version: August 5th 2020

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

Abstract: Previous research showed that the main reason for the persistent gender inequality in the labor market is having children. This inequality for women relative to men caused by childbirth is called the child penalty. In this paper I study whether there is a child penalty in the Netherlands, and most importantly if and how religiosity impacts the child penalty. I make use of the LISS panel data and the event study methodology proposed by Kleven et al. (2019a). I conclude that there is a child penalty in the Netherlands, but I find no evidence that religiosity influences the child penalty in any way.

Index

1. Introduction

Even though an increasing amount of attention has been put on mitigating gender inequality the past decade, the gender gap in different labor market outcomes remains substantial (Sieppi & Pehkonen, 2019). A survey held in the United States in 2019 revealed that 45% of the respondents believed that sexism was one of the major reasons the gender pay gap still exists (Gebhardt, 2019). However, recent studies found that the leading cause of gender disparities on the work floor is not the differences in human capital or discrimination. The main contribution to the gender pay gap is having children (Kleven, Landais, & Søgaard, 2019). This concept is called the child penalty, defined as "the gap in labor market outcomes of women relative to men caused by childbearing" (Sieppi & Pehkonen, 2019, p. 5).

The child penalty has been studied in many countries and settings. It is present in every country studied so far, although differences in shape and magnitude remain. Kleven, Landais, Posch, Steinhauer and Zweimüller (2019) found that differences in countries exist in both the short run and the long run, with a long-run child penalty in earnings of only 21% in Denmark, compared to 61% in Germany. The long-run child penalty in the Netherlands is 39% (Adema, Folmer, Vlekke, Rabaté, & Visser, 2019). Moreover, multiple studies agreed that gender norms play a significant part in the magnitude of the child penalty (Kleven et al., 2019a; Kleven et al., 2019b; Andresen & Nix, 2019). Kleven et al. (2019a) found that the division of labor of a woman's parents when she was growing up is correlated to the child penalty she will endure when she becomes a mother herself.

Although gender norms have been identified as a main driver of the child penalty, it needs more precise analysis to determine its influence. In order to mitigate gender inequality, it is necessary to get to the root of the problem. Understanding why the child penalty still exists, what its drivers are and how they affect this is essential for the government. In order for policy interventions to work, we must comprehend why women work and/or earn less when becoming mothers.

Therefore, I want to focus on a good indicator of gender norms: religiosity. A study by Guetto, Luijkx and Scherer (2015) showed that, among other things, "religiosity is a good predictor of traditional attitudes toward gender roles" (p. 166). It was also found that religion has a significant negative effect on the labor supply of women, especially when children are involved (Lehrer, 1995; Maneschiöld & Haraldsonn, 2007).

The connection between religiosity and attitude toward gender roles of working women, and the connection between gender norms and the child penalty leads me to believe that religiosity might influence the child penalty. Hence, in this paper I study the following:

"How does religiosity affect the child penalty in the Netherlands?"

A representative sample observed between 2008 and 2019 is retrieved from the LISS panel data. I use the event study methodology proposed by Kleven et al. (2019a) to calculate the child penalty. The Centrality of Religiosity Scale (S. Huber & O. Huber, 2012) measures religiousness, looking at people's believe in God, church attendance, and prayer frequency. Child penalties are compared between non-religious men (women) and religious men (women).

Two main points will add to the novelty of this study. First of all, not much is known about the driving factors of the child penalty. Specifically, the relationship between religiosity and the child penalty has never been studied as of now (to my knowledge). Hence, researching this could provide more insight to the predictors of the child penalty. This does not just contribute to scientific literature, it is also socially relevant. Having a good understanding of the reasons behind people's choices facilitates policymaking, which will further improve economic outcomes. Secondly, it is important to study the child penalty and its potential drivers in different countries. Aforementioned, Kleven et al. (2019b) found substantial differences in the child penalty between countries. This means that people's reaction to parenthood differs across countries. But on top of that it could indicate that there is a difference in drivers between countries. The Netherlands is not well studied in this field.

The main results indicate that a child penalty exists in the Netherlands, but religiosity does not affect the child penalty significantly. This holds true for both men and women. Together with the finding of Kleven et al. (2019a) of an intergenerational transmission of child penalties, the results could indicate that female gender norms regarding labor are mainly influenced by environmental factors. The government could use this information to improve gender egalitarian views. However, this needs to be studied further before making any firm conclusions. The results of this paper are merely an indicator of this idea, not evidence.

The paper is set up as follows. Section 2 provides a literature review of the current findings and states the hypotheses. Section 3 describes the data. Section 4 explains the methodology of the study. Section 5 presents the results and section 6 concludes.

2. Literature Review

Child Penalty

Even though the concept is called a penalty, it is not by definition a punishment. The reason for the lower income could be due to voluntary changes in labor work. The child penalty simply indicates the effect of having children on income of women versus men.

The main approach to estimate the child penalty is the event study model proposed by Kleven et al. (2019a), which is used to illustrate the child penalty in various countries. Adema et al. (2019) studied the Netherlands and found a long-run child penalty of 39%. However, the child penalty differs between countries. For example, Kleven et al. (2019b) found long-run child penalties of 21% in Denmark, 44% in the UK and 61% in Germany. Therefore it's important to study the child penalty in multiple countries.

The mechanisms of the child penalty differ as well between countries. Earnings is made up of three components: hourly wage rate, number of hours worked (intensive margin), and labor force participation (extensive margin). In Denmark these measures are approximately equally important for the child penalty (Kleven et al., 2019a). Adema et al. (2019) concluded that the reduction of hours worked is the greatest mechanism of the child penalty in the Netherlands. This may be because working part-time is extremely prevalent in the Netherlands and most Dutch women want to work part-time (Booth & Van Ours, 2013).

There are some variables that have an influence on the child penalty in the Netherlands (Adema et al., 2019). First, highly educated people face a smaller child penalty than non-highly educated people, with long-run child penalties of 30% and 52%, respectively. Second, the amount of children someone has is positively correlated with the child penalty in the long run. Mothers with multiple children decrease the amount of labor work they do every time a new child is born. Hence, mothers with more children face a bigger child penalty over time.

The main driver of the child penalty, researched so far, is gender norms. Kleven et al. (2019b) found considerable differences in the long-run child penalties between countries, as said before. They came to the conclusion that the long-run child penalties don't respond well to public policies. However, they did find a correlational relationship between child penalty and traditional gender norms on a national level. The population prefers, on average, more traditional gender roles in countries that experience higher child penalties. Andresen and Nix (2019) studied four different mechanisms for the child penalty by looking at the effects for hetero couples versus same-sex couples in Norway. Female preference to

family over career and gender norms were identified as the greatest potential sources of the child penalty. Kleven et al. (2019a) examined whether the relative labor force participation in a household is transmitted through generations, specifically from parents to daughters. They indeed showed this. A woman who grew up in a home where the mother worked less compared to the father will face a bigger child penalty herself when she has children. They say this is in line with the idea that a female's gender norms are formed during her childhood. In contrast, Fernández, Fogli and Olivetti (2004) argued that the labor force participation of the *mother-in-law* is positively correlated with the female labor force participation, instead of the labor force participation of the biological mother. Hence, they suggest that female gender norms are not formed during her childhood.

Thus, gender norms are generally seen as a big influence on child penalty, but why and how this happens is still unclear. Therefore, I am going to examine a predictor of gender norms in order to possibly uncover a new mechanism: religiosity.

Religiosity

According to the CBS (Schmeets, 2018), roughly 50% of the Dutch population was religious in 2017. Over time the Netherlands became less religious (Schmeets, 2018). In the late nineties 60% was religious. However, religiosity within someone's life cycle remains relatively stable (Guetto et al., 2015). A remarkable observation is that women are on average more religious than men in the Netherlands (Kregting, Scheepers, Vermeer, & Hermans, 2019). They find three main determinants for this religious gender gap. First, they claim that women experience more health impairments than men which leads to more self-depreciation and insecurity, driving religiosity.[1] Second, women have higher levels of personality traits that strengthen religiosity, with the most important being conscientiousness. At last, women have a higher level of femininity and a lower level of masculinity than men, enlarging the religious gender gap. Being feminine relates to wanting to take care of others and being thoughtful. These values are generally shared in religion, hence why feminine people are more likely to follow a religion. Masculinity is associated with wanting to be self-reliant and wanting to be in charge. This conflicts with religion, which generally has the notion that God is sovereign and that we as people need Him. Masculine people are thus less likely to follow a religion. It is important to note that masculinity and femininity are not restricted to one gender. Women can be very masculine, just like men can be very feminine. However, in the Netherlands women are more feminine and less masculine than men, according to Kregting et al. (2019).

---

[1] They derive the claim that Dutch women have more health impairments than Dutch men from Knoops, Verweij, Merens and Pulles (2016).

Religiosity is correlated with multiple measures that influence the child penalty. For example, religiosity and education are negatively correlated. Arias-Vazquez (2012) found that education has a negative effect on religiosity in the US. H'madoun (2010) observed a negative correlation between religiosity and education for women.

In addition, various studies reveal a positive relation between religiosity and fertility (e.g. Philipov & Berghammer, 2007). Berghammer (2012) examined the relationship between church attendance and childbearing in the Netherlands. She found that the level of church attendance is positively correlated with future fertility, but not the other way around. This suggests that religiosity, measured as church attendance in this case, is not influenced by having children, but religious people will have more children than non-religious people, on average.

Lastly, various studies state that religiosity is negatively correlated with female labor participation (Guetto et al., 2015; Heineck, 2004; H'madoun, 2010).

The most important connection is that "religiosity is found to be a good predictor of traditional attitudes toward gender roles" (Guetto et al., 2015, p.166). In addition, Seguino and Lovinsky (2009) concluded that religiosity is strongly connected to unequal gender norms and strongly correlated with gender inequality outcomes.

Thus, I first hypothesize that a child penalty on earnings exists. Consequently, I hypothesize that a higher level of religiosity results in a higher child penalty.

3. Data

Data Source and Sample Selection

To examine the child penalty in the Netherlands I use data from the Longitudinal Internet Studies for the Social sciences (LISS) panel. This is administered by CentERdata (Tilburg University, The Netherlands). A true probability sample of households was drawn from a population register by Statistics Netherlands. It consists of approximately 7,500 individuals in 5,000 households. The survey about background variables, such as gender and civil status, is updated monthly. Additionally, the LISS Core Study, comprising multiple topics, is carried out every year. This is a longitudinal survey that repeatedly measures the same set of variables on a yearly basis. Data is available from 2007 onwards.[2] I make use of the demographics data, the Family and Household data, and the Religion and Ethnicity data.

Even though the panel is meant to be representative for the Dutch-speaking population in the Netherlands, some people are underrepresented. In analysis of the representativeness of the LISS panel by Van Der Laan (2009) and by Knoef and De Vos (2009) it became evident that mostly older people, people with no internet access, first-generation immigrants and people in single-person households were underrepresented. Nevertheless, I assume this sample to be representative of the research population. The reason is that the underrepresented groups are not of great relevance in the study sample. Single-person households are excluded from the sample, since only people becoming parents rather recently – entailing a household of at least two people – are researched. Moreover, older people are not likely to become parents. People in the Netherlands are on average around the age of 30 when entering parenthood (CBS StatLine, 2019). Hence, older people are generally not needed in this study.[3] Likewise, people without internet access are relatively old, so they aren't essential in this study. The only underrepresented group that could lead to a lower external validity is immigrants, but the results are likely to be representative for native Dutch first-time parents.

I retrieve yearly data about parents whose first child was born between 2008 and 2017. Data on the individual level (parent) are collected from 2 years before the parent's first childbirth to 5 years after their first childbirth. I choose this time frame instead of a larger window in order to obtain a relatively decent sample size.

---

[2] More information about the LISS panel can be found at www.lissdata.nl.

[3] Also, the chance of getting pregnant naturally from age 45 onwards is very rare (ACOG, n.d.).

Only biological children are identified and recognized as children in this study. Adoptive children are not necessarily adopted at birth, which means that income of adoptive parents does not react to the birth of the child, but rather the arrival of the child. Moreover, being a stepparent is different than being the actual parent of a child. These complicate the issue too much.

A few respondents were removed from the sample in order to eliminate as many outliers as possible. First, parents that changed gender at some point in time are removed.[4] Second, parents who were in a same-sex relationship at one point are excluded. Same-sex parents are subject to the same gender norms as their partner, and therefore they are not informative for this analysis. Third, parents with at least one child deceased after being born, are dropped. The impact of this on parents is so great and unpredictable, which could possibly lead to outliers. Finally, they should be at least 20 years old when they first become parents. This is to eliminate teen parents. After all restrictions a sample of 789 parents is left. This is a small sample size, especially when compared to the sample sizes of the other child penalty studies, who mainly used population-based data. This increases the chance of a type II error, also known as a false negative. You fail to reject the null hypothesis, while it is false in reality. Another problem would be that visible differences in for example effects are not significant, as is presented in this paper later on. Effects/coefficients that are right below the significance point are plausible to be significant in reality. The small sample size might be the reason for the statistical insignificance. In this case you could talk about a suspicion of an effect, but it is not proven.

Variables

Religiosity

Every parent will be categorized as either religious or non-religious. To measure religiosity, a modification of the Centrality of Religiosity Scale (CRS) is used (S. Huber & O. Huber, 2012). This scale measures religiosity based on five categories: intellect, ideology, public practice, private practice, and experience. Due to the limited survey questions in the LISS panel about religion, only three categories are used in this study. These are ideology (your idea of God), public practice (church attendance), and private practice (prayer frequency). See appendix A for an overview of the survey questions and the possible answers. The answers to these three questions can be put in item scores from 1 to 5, with a lower number indicating lower intensity of religiosity. To calculate someone's CRS score, their item scores are averaged. This creates a range of the CRS score between 1.0 and 5.0. Then someone's CRS score that is available closest to the year of childbirth is taken as his CRS score and is

---

[4] Additionally, the birth year of parents was not always constant. The changes were usually enormous, which indicates that these were most likely errors in the data. Parents with an inconstant year of birth are excluded as well.

fixed across time. Hence, religiosity is a time-invariant variable. Thereafter, people are divided in two categories according to their CRS score, following the thresholds described and justified by S. Huber and O. Huber (2012). People with a CRS score between 1.0 and 2.0 are considered non-religious. People with a score above 2.0 to 5.0 are considered religious.[5]

I choose to measure religiosity based on the CRS instead of the self-reported religiosity in the survey. For religiosity to be a predictor of traditional gender norms, it is important that religion is actively present in someone's life. Being a member of a religion or church is not equivalent to sincerely practicing your religion, believing in God, and implementing rules, ideas and ways of thinking in your life. With the CRS the level of religiosity is based on the centrality of religion in your active life.

Income

Income is measured as annual gross earnings in euros. This is calculated by averaging someone's monthly gross income per year times 12. When gross monthly income was not specified, the imputed gross monthly income was used. It is worth noting that the imputed income is not self-reported, but calculated based on self-reported net income. However, by only using the self-reported gross income the sample becomes too small. There are a few disadvantages to using income reported in the monthly survey instead of the data from the detailed annual income survey. These mainly comprise the fact that the survey is filled in by one household member for the entire household, and the matter that some people reported an income of zero while stating that they have paid employment. However, these data are superior in that it lists monthly income, which is more accurate than annual income, and in that this income measure is observed more often. It follows that this measurement of income is the best choice.

As previously mentioned, income encompasses three categories: labor force participation, number of hours worked, and the hourly wage rate. It would be very interesting to examine all outcome variables. However, this is beyond the scope of this study. Nonetheless, the earnings variable is unconditional on being employed. Therefore it implicitly describes labor force participation; it is comprised in income.

One point worth noting is that there are multiple income outliers in the sample. These could change the results, especially because of the small sample size. However, I did not remove any outliers in this study, because extreme income values are not per se false. It could be a recording error, where someone typed the wrong income by accident in the survey. Another possibility is that the person changed jobs and now has a leading role in a company, for example, earning much more. In short, in this setting it is impossible

---

[5] S. Huber and O. Huber (2012) categorized people into three groups instead of two; they also had a highly-religious group. However, due to the small sample size available to me, I combine religious and highly-religious people into one group.

to know the reason for the outlier. And since I'm looking at change in income over time, I think it's inappropriate to remove these outliers, as they could possibly be influenced by the children effect. In this case the risk is too big to remove outliers of income, as it could lead to unethical data manipulation.

Age

Parent's age in years is measured as the calendar year minus the birth year.

Number of children

The number of children a parent has is measured as the number of children someone has 5 years after their first child was born. As previously mentioned, this only includes biological children. A parent in the sample can have step kids, but these children don't count as the number of children he/she has.

Education

Education is a variable indicating the highest level of education someone achieved before becoming a parent. It is a categorical variable with six groups, ranging from primary school (1) to university (6). These six categories are subdivided in three groups, indicating a more general education level, like CBS uses (2017). These are poorly educated, moderately educated, and highly educated. I will be using education with these 3 categories because of the small sample size and for easier comparison.

The number of children and education level of a parent are not necessary for the analysis, but it is beneficial to see if there are differences in these variables over religiosity. Why this is beneficial and the actual comparison are discussed in the next subsection.

Descriptive Statistics

Only 31% of the sample is religious. Schmeets (2018) reported that in 2017 49% of the Dutch population was religious. As explained above the measure of being religious in this study is more strict. This most likely explains the difference. Females are relatively more religious than men in the sample. This is in line with the observations of Kregting et al. (2019) and Schmeets (2018).

Table 1: Descriptive statistics of parent characteristics at first childbirth $(t-1)$.

	All		Men		Women	
	Men	Women	Non-Religious	Religious	Non-Religious	Religious
N individuals	358	431	261	97	282	149
Religious	0.271	0.346	0	1	0	1
	(0.445)	(0.476)	(0)	(0)	(0)	(0)
Earnings (€ 1000)	33.778	26.280	34.460	31.496	28.033	22.657
	(12.383)	(12.508)	(12.441)	(12.057)	(12.609)	(11.555)
Age when becoming a parent	32.550	30.123	32.797	31.887	30.248	29.886
	(5.300)	(4.271)	(5.246)	(5.414)	(4.079)	(4.619)
Year entering parenthood						
2008	0.089	0.095	0.084	0.103	0.096	0.094
2009	0.117	0.125	0.115	0.124	0.103	0.168
2010	0.115	0.114	0.126	0.082	0.117	0.107
2011	0.128	0.109	0.123	0.144	0.113	0.101
2012	0.103	0.114	0.084	0.155	0.110	0.121
2013	0.106	0.100	0.111	0.093	0.110	0.081
2014	0.106	0.111	0.115	0.082	0.128	0.081
2015	0.089	0.104	0.069	0.144	0.085	0.141
2016	0.087	0.077	0.111	0.021	0.082	0.067
2017	0.059	0.051	0.061	0.052	0.057	0.040
Number of children (t = 5)	1.514	1.485	1.502	1.546	1.479	1.497
	(0.574)	(0.558)	(0.573)	(0.578)	(0.567)	(0.541)
Number of children (t = 5)						
1 child	0.525	0.545	0.536	0.495	0.557	0.523
2 children	0.436	0.425	0.425	0.464	0.408	0.456
3 children	0.039	0.030	0.038	0.041	0.035	0.020
Education Level						
Poorly educated	0.120	0.056	0.119	0.125	0.057	0.054
Moderately educated	0.356	0.357	0.375	0.302	0.349	0.372
Highly educated	0.524	0.587	0.506	0.573	0.594	0.574

Note: The mean and standard deviation (in parentheses) are given. Certain standard deviations have been left out to make the table more concise. All variables are time-invariant, except for income.

12

Men are on average 32.55 years old when becoming a parent, while women are 30.12 years: a difference of almost 2.5 years. Religious people are slightly younger when entering adulthood, compared to non-religious people. For men the difference is 0.91 years, for women it is 0.36 years. This is in accordance with results the CBS generated (2019).

The mean of a certain year entering parenthood declares the fraction of the sample that have become parents for the first time in that year. For example, in 2009 11.5% of non-religious men have become first-time parents. The birth years of the babies are quite evenly spread across gender, religiousness, and time. There are some differences but nothing that would potentially cause problems. There is not a year with a sudden (relative) increase or drop in the number of babies born.

The number of children a parent has can affect the child penalty in the long run (Adema et al., 2019; Kleven et al., 2019a). As mentioned in section 2, various studies found a relationship between religiosity and fertility. A detected difference in child penalty by religiosity might merely encompass the difference caused by fertility, instead a difference caused by religiosity itself. Therefore it's useful to see if there are any differences between religious and non-religious people. In the sample there are no significant differences. Non-religious men (women) have on average 1.50 children 5 years after the first child was born (1.48), while religious men have 1.55 children (1.50). When looking at the fraction of people per number of children, you can see that there are no distinct differences between each level of fertility. For example, 42.5% of non-religious men have 2 children, while 46.4% of religious men have 2 children. This holds true for women as well. Hence, this could indicate that any differences in the child penalty between non-religious people and religious people is not because of a difference in fertility.[6]

The same problem goes for education. In table 1 you can see that the fraction of religious men that are highly educated is slightly bigger in comparison to the fraction of non-religious men. In this sample religious men are slightly more educated than non-religious men. This should be kept in mind. There is no visible difference between non-religious women and religious women. Thus, any difference in the child penalty between non-religious women and religious women cannot be explained by level of education.

---

[6] There is a possibility that a difference might appear when looking at even more years past childbirth, for example at $t = 10$. However, the fact that there is no difference the first five years is sufficient when looking at the penalty till 5 years post childbirth.

Average yearly income over event time by gender and religiosity is displayed in figure 1. Religious women have a lower average income than non-religious women. This is consistent with previous literature (e.g. Heineck, 2004). However, what is interesting is the difference between non-religious men and religious men. This could have multiple reasons.[7] Religious men might have relatively more jobs in non-profit companies with the aim to improve social justice. They choose helping others over income. Religiousness is linked to prosocial behavior (Einolf, 2011). Another possibility is that the religious group contains more immigrants, which would be consistent with the population (Schmeets, 2018). Being an immigrant is negatively correlated with income (Jongen, Bolhaar, Van Elk, Koot, & Van Vuuren, 2019). It is worth noting that the difference between non-religious women and religious women is bigger than the difference between non-religious men and religious men. One year before childbirth, income of non-religious men is 2.96 thousand euros higher than that of religious men (9.4% higher). Income of non-religious women is 5.38 thousand euros higher than that of religious women (23.7% higher).

With the short time window it is difficult to see and check if income moves parallel between men and women before childbirth. However, given these results, you can see that income trends of non-religious men and non-religious women are somewhat parallel, as well as of religious men and religious women, whose trends are a bit more parallel. After childbirth it is visible that men's earnings keep rising, while income of women of both religious groups stagnates and even slightly diminishes.
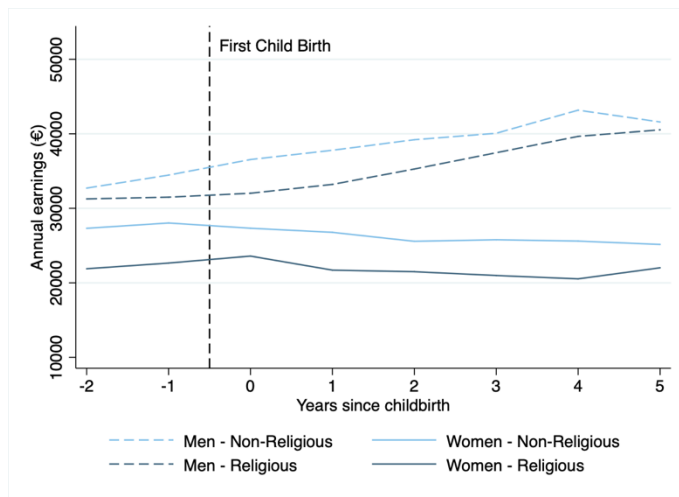

Figure 1: Evolution of income for non-religious and religious fathers and mothers by year away from birth.

---

[7] Studying this is beyond the scope of this study, but it would be very insightful.

4. Methodology

Analysis

To analyze the child penalty I follow the event study approach, identified by Kleven et al. (2019a). The year of birth of the first child is set at event time $t = 0$, and all years are indexed relative to that year. I look at a time window of $t - 2$ to $t + 5$. Using the unbalanced panel data described previously, I first estimate the following model separately for men and women, which gives two regressions. Then I estimate the model separately for non-religious men and women, and religious men and women, which gives four regressions.

$$Y_{ist}^{gr} = \sum_{j \neq -1} \alpha_j^{gr} \cdot I[j = t] + \sum_k \beta_k^{gr} \cdot I[k = age_{is}] + \sum_y \gamma_y^{gr} \cdot I[y = s] + \varepsilon_{ist}^{gr} \tag{1}$$

Y_{ist}^{gr} specifies earnings in euros for individual i of gender g with religiosity r in calendar year s at event time t. This is unconditional on employment. The right side of the equation contains these three terms respectively: event time dummies, age dummies, and year dummies. I omit the event time dummy at $t = -1$, so that the event time coefficients measure the impact of having children relative to the year before the first childbirth. The age dummies and year dummies account nonparametrically for lifecycle trends and time trends, respectively. A significance level of 5 percent is used, and standard errors are clustered on an individual level.

To estimate the percentage effects equation (2) is performed. $\tilde{Y}_{ist}^{gr}$ is the predicted outcome without children,[8] i.e., the predicted outcome someone would have had had they not had any children.

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

Hence, P_t^{gr} measures the effect in year t of having children as a percentage of the estimated income had he/she not had any children.

---

[8] $\tilde{Y}_{ist}^{gr} = \sum_k \hat{\beta}_k^{gr} \cdot I[k = age_{is}] + \sum_y \hat{\gamma}_y^{gr} \cdot I[y = s]$

By comparing men and women I estimate the impact becoming a parent has on women relative to men: the relative child penalty. It is calculated as follows per religiosity level and t.

$$P_t^r = \frac{\hat{\alpha}_t^{mr} - \hat{\alpha}_t^{wr}}{E[\tilde{Y}_{ist}^{wr}|t]} \tag{3}$$

The long-run child penalty is calculated as the average child penalty over event time $t = 4,5$. The more well-known studies use a larger time window and therefore they calculate the long-run child penalty at a later point in time post childbirth (e.g. $t = 10$). They generally conclude that the child penalty is reasonably constant at that time. In this study it is difficult to assess the level of constancy due to the small time window. Therefore I take the average at $t = 4,5$ to slightly smooth the gender-specific child penalties. These are still changing at that time. Hence, I believe looking at the average at $t = 4,5$ is more credible than looking at just $t = 5$. However, the concept of long-run penalty in this paper should be taken lightly. A medium-run penalty might be a more appropriate term for this assessment.

About The Event Study Methodology

Kleven et al. (2019a) describe two main assumptions/conditions for this event study to identify the impact of children. First, for the short-run effect to be identified smoothness is assumed. This means that income determinants that are independent of children should change smoothly over time.[9] Hence, the average non-child earnings around the event time are smooth as well. Income before childbirth should have no sharp changes. Then, the timing of the childbirth is exogenous to income. When income exhibits sharp changes right at the event or right after, and the income determinants unrelated to children are smooth, than the changes in income can be assigned to the event.

There are two main differences in regard to the long-run impact. First, it encompasses the impact of total fertility, instead of just the first-born child. Second, the smoothness assumption is not sufficient for the long-run impact to hold. Over a longer period of time the non-child determinants could still exhibit major changes. Therefore, it is necessary to control nonparametrically for lifecycle- and time trends with age and year. When looking at the long-run child penalty, the men are essentially used as a control group for women. When their income pre-trends, controlled nonparametrically for age and year, are parallel the long-run impact of women relative to men can be identified. In their paper Kleven et al. (2019a) conducted two analyses to test the accuracy of their event study methodology and concluded that the event study as described above in the equations is valid to estimate the causal effects of children on income.

---

[9] Examples are someone's age, preferences, and competence.

16

5. Results

All the figures present the percentage effects of becoming a parent on income relative to one year before childbirth, equation (2), over time. The 95% confidence intervals are based on standard errors clustered at the individual level. Table 2 with regression coefficients from equation (1) can be found on page 20.

General Child Penalty

First, I look at the general child penalty between men and women, displayed in figure 2. You can see that childbirth barely has any effect on the father's income, while it lowers women's earnings. This negative effect is significant in all years after childbirth. The long-run relative child penalty is 29%. Given these results, I can accept the first hypothesis that a child penalty on earnings exists.
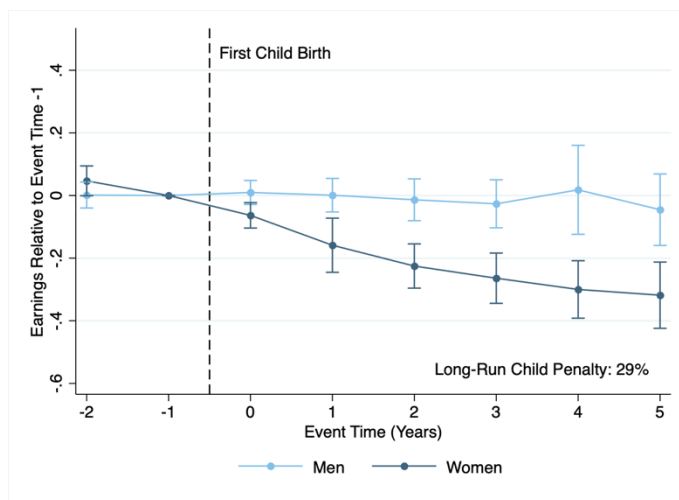

Figure 2: The child penalty on income for men and women.

Differences In Religiosity

Next, I look at the income trajectories of non-religious men and religious men. This is displayed in figure 3. The impact of children on their income is really similar, but some differences occur further away from birth. The difference is mainly noticeable in 3-5 years post childbirth. However, the differences are not significant. The reason might be the small sample size. With a much bigger sample size the effect could be significant. Another possible reason is the absence of a significant effect in reality. I'm hesitant to hint about a suggestion for an effect, since the progression of religious men is very gradual and still relatively small. There were also a few income outliers for men in 4 and 5 years post childbirth. In conclusion, there are some differences visible between non-religious men and religious men, but these are small and not significant. Hence, I do not find evidence that religiosity affects the impact of becoming a parent on income for men. Important to note is that I could not reject the null hypothesis that there isn't any difference, but that is not the same as accepting this null hypothesis. I did not find evidence, but that doesn't mean that there is none.
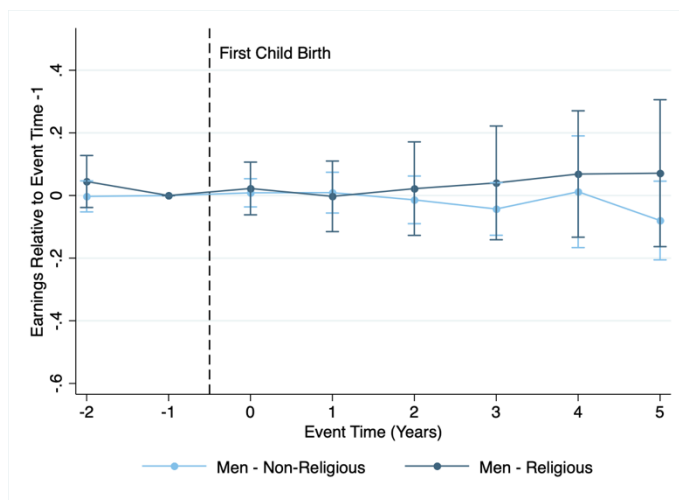

Figure 3: The child penalty on income for non-religious men and religious men.

Then I look at the non-religious women and religious women, shown in figure 4. Their child penalties are extremely similar. The only noticeable differences are in the year the child is born, and 5 years post childbirth, but these are not significant. This could be because of the small sample size, or just because there is no significant difference in reality. But, by and whole the magnitude and the shape are almost identical. Thus I did not find any evidence of an effect of religiosity on the child penalty for women.
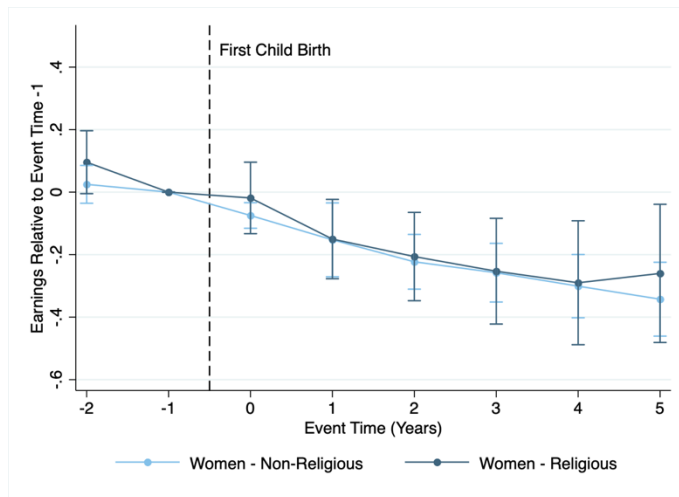

Figure 4: The child penalty on income for non-religious women and religious women.

Overall Result

Finally, I look at the child penalties over event time for all four groups together in figure 5.[10] Here you can see again that the difference between religious women and non-religious women is really small. For men the income trajectories differ more, but it is not significant. The long-run child penalty for non-religious women relative to non-religious men is 28%, while this is 36% for religious people. Again, these differences can't be interpreted as a significant difference between religious people and non-religious people. Thus, the second hypothesis, which stated that higher religiosity results in a higher child penalty, can't be proven. This doesn't mean that there is no difference in effect, but it simply means that I did not find evidence to prove that a difference does exist.

---

[10] In the figure no confidence intervals are shown. This is to make it more comprehensible.
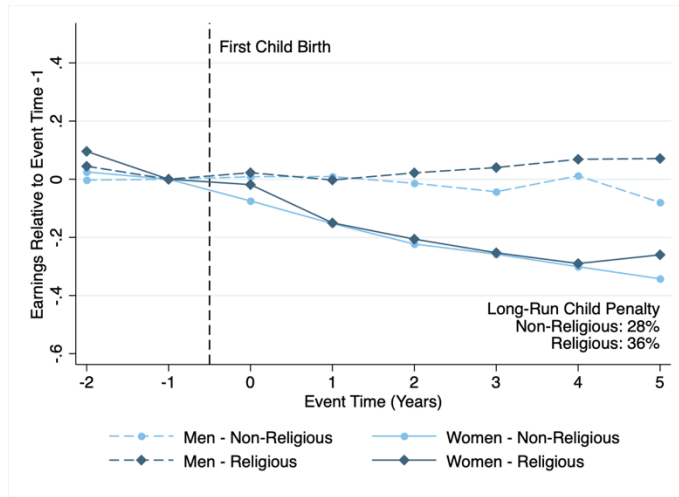
Figure 5: The child penalty on income for non-religious men and women and religious man and women.

Table 2: The child penalty on income for different groups.

	All parents		Men		Women	
	Men	Women	Non-Religious	Religious	Non-Religious	Religious
–2 years	48.052	1145.365	–93.450	1338.756	665.325	1914.799
	(682.729)	(595.267)	(830.307)	(1266.903)	(823.494)	(1026.053)
0 years	340.519	–1760.692**	310.239	702.771	–2206.440***	–445.984
	(686.128)	(574.265)	(832.114)	(1345.658)	(616.185)	(1402.885)
1 year	35.966	–4716.806***	343.418	–88.489	–4823.463*	–3836.753*
	(999.679)	(1310.332)	(1244.773)	(1913.936)	(1904.166)	(1653.896)
2 years	–528.980	–7003.943***	–554.244	758.112	–7334.131***	–5580.609**
	(1318.651)	(1121.386)	(1543.877)	(2628.463)	(1469.850)	(1951.791)
3 years	–1072.337	–8631.175***	–1802.035	1449.811	–8956.412***	–7108.804**
	(1582.686)	(1338.510)	(1793.082)	(3337.589)	(1662.521)	(2424.595)
4 years	748.748	–10201.051***	509.864	2546.339	–11013.410***	–8386.482**
	(3002.073)	(1593.505)	(3883.520)	(3819.974)	(1890.716)	(2924.935)
5 years	–1958.858	–11232.060***	–3608.442	2695.538	–13102.126***	–7736.573*
	(2520.029)	(1905.872)	(2897.335)	(4529.970)	(2308.164)	(3356.178)
Observations	1890	2265	1389	501	1475	790
F–test	$F_{(6, 357)} = 0.97$	$F_{(6, 430)} = 8.26$	$F_{(6, 260)} = 1.60$	$F_{(6, 196)} = 0.40$	$F_{(6, 281)} = 7.41$	$F_{(6, 148)} = 3.94$
	Prob > F = 0.443	Prob > F = 0.000	Prob > F = 0.148	Prob > F = 0.881	Prob > F = 0.000	Prob > F = 0.001

Note: the table displays part of the regression coefficients of equation (1). The dependent variable is income in all models. Controls in the regression are age fixed effects and year fixed effects. The F-test tests for a joint significance for all post-birth event times ($t \geq 0$). The standard errors are in parentheses and are clustered at individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6. Conclusion

Previous studies showed that most of the gender inequality in the labor market can be attributed to having children. Becoming a parent generally won't affect men's labor market outcomes, but those of women decrease significantly. This disparity is called the child penalty. Researchers also showed that gender norms affect the magnitude of the child penalty substantially. Using LISS panel data and an event study methodology, I examined the child penalty in the Netherlands and the effect religiosity has on the child penalty, as a predictor of traditional gender norms.

First, I found a large child penalty in earnings in the Netherlands, with a long run child penalty of 29%. This is somewhat similar to the research of Adema et al. (2019), who also calculated the child penalty in the Netherlands. The child penalty they measured was a bit higher and the effect was more abrupt, whereas the child penalty in this paper over event time increased more gradually.

Second, I found no evidence that religiosity significantly affects the child penalty. Income relative to the year before childbirth moved almost identical for non-religious women and religious women, with a few minor differences. Religious men and non-religious men have the same income trajectories, but the relative income of religious men increased slightly after 3 years post childbirth. However, the difference between religious men and non-religious men is non-significant. The small sample size could be the reason for this insignificance. Further implications are therefore only about women.

In conclusion, I did not find evidence that religiosity affects the child penalty in any way. This could imply that being religious is not a good predictor for having traditional gender and family norms at all. A more plausible explanation is that in the Netherlands specifically, and during this time, religiousness is not a good predictor for gender norms regarding work.

Traditionally religiosity is a good indicator of gender norms. Most religions heavily promote traditional social norms, especially focusing on women and mothers. However, in the Netherlands there has been a huge decline for decades in religiosity. This turnover to secularization made non-religious people have more egalitarian gender views (Thijs, Te Grotenhuis, Scheepers, & Van Den Brink, 2019). Thijs et al. (2019) found that women who are surrounded by a higher level of secularization have more egalitarian gender norms in the Netherlands. I combine my findings in this paper with the findings of Thijs et al. (2019) and of Kleven et al. (2019a) that child penalties are transmitted through generations via daughters. Together it could indicate that female gender norms regarding labor are barely influenced by personal beliefs and preference, but are mainly formed by the conditions of the woman's own surroundings and what is deemed as normal.

21

I want to emphasize that this is all speculation. More research should be done regarding this to identify the predictors of female gender norms regarding labor. Assuming this is true, the government could influence the environmental effects to improve egalitarian gender norms, thereby improving the child penalty.

Suggestions for further research include the replication of this study with more precise data, and to include other measures of labor market outcomes, such as labor force participation, the number of hours worked, and the hourly wage rate. It would also be interesting to compare this to other countries. The effect of religiosity might be a lot stronger in countries that are overall more religious. Additionally, by comparing the effect now to the effect previously, we could see if there is a time trend relating to religiosity and gender. Lastly, more extensive research to other possible predictors of gender norms is highly relevant to not only the scientific field, but also for society.

Bibliography

ACOG. (n.d.). *Having a baby after age 35: How aging affects fertility and pregnancy*. Retrieved from https://www.acog.org/patient-resources/faqs/pregnancy/having-a-baby-after-age-35-how-aging-affects-fertility-and-pregnancy

Adema, Y., Folmer, K., Vlekke, M., Rabaté, S., & Visser, D. (2019). Arbeidsparticipatie, gewerkte uren en economische zelfstandigheid van vrouwen. In *CPB notitie*. Den Haag.

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

Arias-Vazquez, F. J. (2012). A note on the effect of education on religiosity. *Economics Letters, 117*(3), 895-897.

Berghammer, C. (2012). Church attendance and childbearing: Evidence from a Dutch panel study, 1987–2005. *Population Studies, 66*(2), 197-212.

Booth, A. L., & Van Ours, J. C. (2013). Part-time jobs: What women want?. *Journal of Population Economics, 26*(1), 263-283.

CBS. (2017). *Partners steeds vaker beiden hoogopgeleid*. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2017/45/partners-steeds-vaker-beiden-hoogopgeleid

CBS. (2018). *Inkomen naar leeftijd, branche en opleiding, 2007-2015*. Retrieved from https://www.cbs.nl/nl-nl/maatwerk/2018/18/inkomen-naar-leeftijd-branche-en-opleiding-2007-2015

CBS StatLine. (2019). *Geboorte; kerncijfers*. Retrieved from https://opendata.cbs.nl/#/CBS/nl/dataset/37422ned/table?dl=D02F

Einolf, C. J. (2011). The link between religion and helping others: The role of values, ideas, and language. *Sociology of Religion, 72*(4), 435–455.

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.

Gebhardt, J. (2019). *On equal pay 2019, lack of awareness persists*. Retrieved from https://www.surveymonkey.com/curiosity/equal-pay-day-2019/

Guetto, R., Luijkx, R., & Scherer, S. (2015). Religiosity, gender attitudes and women's labour market participation and fertility decisions in Europe. *Acta Sociologica, 58*(2), 155-172.

Heineck, G. (2004). Does religion influence the labor supply of married women in Germany?. *Journal of Socio-Economics, 33*(3), 307–328.

H'madoun, M. (2010). *Religion and labor force participation of women*. UA, Faculty of Applied
     Economics.

Huber, S., & Huber, O. W. (2012). The centrality of religiosity scale (CRS). *Religions, 3*(3), 710-724.

Jongen, E., Bolhaar, J., Van Elk, R., Koot, P., & van Vuuren, D. (2019). Inkomensongelijkheid naar
     migratieachtergrond. In *CPB Policy Brief*. Den Haag.

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

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller Josef. (2019b). Child penalties across
     countries: Evidence and explanations. *AEA Papers and Proceedings, 109*, 122–126.

Knoef, M., & De Vos, K. (2009). The representativeness of LISS, an online probability panel.

Knoops, K., Verweij, G., Merens, A. & Pulles, I. (2016). Gezondheid. In W. Portegijs & M. van den
     Brakel (Eds.), *Emancipatiemonitor 2016* (pp. 214-244). Den Haag: Sociaal en Cultureel
     Planbureau / Centraal Bureau voor de Statistiek.

Kregting, J., Scheepers, P., Vermeer, P., & Hermans, C. (2019). Why Dutch women are still more
     religious than Dutch men: Explaining the persistent religious gender gap in the Netherlands
     using a multifactorial approach. *Review of Religious Research, 61*(2), 81-108.

Lehrer, E. L. (1995). The effects of religion on the labor supply of married women. *Social Science
     Research, 24*(3), 281-301.

Maneschiöld, P., & Haraldsonn, B. (2007). Religious norms and labour supply of married women in
     Sweden. *Finnish Economic Papers, 20*(1), 41-56.

Philipov, D., & Berghammer, C. (2007). Religion and fertility ideals, intentions and behaviour: A
     comparative study of European countries. *Vienna yearbook of population research*, 271-305.

Seguino, S., & Lovinsky, J. (2009). The impact of religiosity on gender attitudes and outcomes.
     *UNRISD background paper, United Nations Research Institute for Social Development,
     Geneva.*

Sieppi, A., & Pehkonen, J. (2019). Parenthood and gender inequality: Population-based evidence on
     the child penalty in Finland. *Economics Letters, 182*, 5-9.

Schmeets, H. (2018). Wie is religieus, en wie niet? *CBS, Statistische trends.*

Thijs, P., Te Grotenhuis, M., Scheepers, P., & van den Brink, M. (2019). The rise in support for
     gender egalitarianism in the Netherlands, 1979-2006: The roles of educational expansion,
     secularization, and female labor force participation. *Sex Roles, 81*(9-10), 594–609.

Van Der Laan, J., & Netherlands. Centraal Bureau voor de Statistiek. (2009). *Representativity of the
     LISS panel* (Ser. Discussion paper, 09041). Statistics Netherlands.

Appendix

A: Measurement Of Religiosity

As explained in section 3, religiosity is measured similar to the Centrality of Religiosity Scale (CRS) (S. Huber & O. Huber, 2012). The respondents are asked three multiple choice questions regarding their religiosity. These are about their ideology (idea of God), public practice (church attendance), and private practice (prayer). Their answers can be put on an item score from 1 to 5. Here are the three questions reported, and the corresponding item score that belongs to the answers.

Ideology: Which of the following statements best describes your idea of God?

Public practice: Aside from special occasions such as weddings and funerals, how often do you attend religious gatherings nowadays?

Private practice: Aside from when you attend religious gatherings, how often do you pray?

Table A1: Recoding of ideology answers into 5 five score levels.

Ideology	5 levels
A) I believe without any doubt that God exists	5
B) I believe in God, although I have my doubts	4
C) At some moments I do believe in God, at other moments I don't	3
D) I do not believe in a God that is personally concerned with each of us, but I do believe in a higher power	2
E) I do not know if God exists, and I do not believe that we have any way of knowing	
F) I do not believe in God	1

Table A2: Recoding of public practice answers into 5 five score levels.

Public practice		5 levels
A)	Every day	5
B)	More than once a week	
C)	Once a week	4
D)	At least once a month	3
E)	Once or a few times a year	2
F)	Never	1

Table A3: Recoding of private practice answers into 5 five score levels.

Private practice		5 levels
A)	Every day	5
B)	More than once a week	4
C)	Once a week	3
D)	At least once a month	
E)	Once or a few times a year	2
F)	Never	1