

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

The Role of Dutch Grannies in Mothers's Labor Force Participation: An Instrumental Variables Approach

MSc Economics of Markets and Organizations

Gencay Işıl Başbüyük

Student ID: 576409



Supervisor: D. Sisak

Second reader: G. Moroni

22 November 2021

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

Netherlands is one of the few countries in Europe with very little public provision of the childcare system, where the childcare system is characterized by high prices and low coverage in rural regions. Grandparents caring for grandchildren is widespread along with high rates of maternal labor force participation. The unique characteristics of the Netherlands, raise the question of whether regular grandparental childcare has a direct influence on maternal labor supply and if so, to what extent? Using multiple waves from Longitudinal Internet studies for Social Sciences core panel, to deal with the endogeneity of grandparental childcare, we empirically attempt to take advantage of the panel data structure by employing a fixed effects instrumental variables approach and an instrumental variables approach on a single wave. As an instrument we use whether the grandparent is alive to exogenously predict the use of grandparental childcare. We find no direct effect of regular grandparental childcare on maternal labor supply. Based on our conclusions, as an extension we study the direct effect of grandchild care on the number of hours worked weekly for mothers that are employed using an instrumental variables approach. Although we find that mothers who receive such assistance work a greater number of hours in average, our results are not robust and must be interpreted with caution. We argue that the availability of part-time opportunities in the Netherlands explains our results.

Keywords: Grandparental childcare, Maternal labor force participation, Maternal grandmother, Endogeneity, Unobserved preferences, Part-time opportunities

Table of Contents

Abstract	2
1. Introduction	4
2. Literature Review	8
3. Endogeneity of Grandparental Childcare	14
4. Research Methodology	16
4.1. <i>Fixed Effects with Instrumental Variables</i>	16
5. Data Description	19
5.1. <i>Endogenous variable: grandparental childcare</i>	20
5.2. <i>Outcome variable: mother's labor force participation (MFLP)</i>	20
5.3. <i>Instrumental variable: maternal grandmother life status</i>	20
5.4. <i>Explanatory variables</i>	20
6. Descriptive Statistics	22
7. Results	23
7.1. Fixed Effects	23
7.2. Fixed Effects with Instrumental Variables	25
7.3. Instrumental Variables (MLFP)	28
7.3.1. <i>Ordinary Least Squares (MLFP)</i>	28
7.3.2. <i>Instrumental Variables (MFLP)</i>	29
7.4. Instrumental Variables (Logarithm of Hours Worked Weekly)	33
7.4.1 <i>Ordinary Least Squares (Logarithm Hours Worked Weekly)</i>	35
7.4.2 <i>Instrumental Variables (Logarithm Hours Worked Weekly)</i>	36
8. Robustness Checks	39
8.1 <i>Geographical Proximity</i>	39
8.2 <i>Grandparents in Need of Assistance</i>	39
8.3 <i>Bivariate Probit Model</i>	40
8.4 <i>Instrumental variables: 13th Wave, 2020</i>	40
9. Discussion and Conclusion	41
10. References	44
11. Appendices	51

1. Introduction

As the primary caregiver of children, the availability of childcare is crucial for maternal women's decision of labor force participation. Economic models who look into models of female labor supply predict childcare costs as the key variable in women's decision to participate (Arpino et al., 2014). Traditionally these models investigating formal childcare find that there is an inverse relationship between childcare costs and female labor force participation (Kimmel, 1998). Although, these studies provide valuable information and policy suggestions for governments and organizations for their childcare policies regarding childcare costs, affordability and availability, the effect of 'costless' or in other words informal care have been often overlooked in the literature.

Although overlooked, many families make use of informal childcare, specifically care that is provided by maternal grandparents. Regardless of how widespread it is, little is known about grandparental childcare's effect on women's working decisions and few studies have empirically investigated grandparental childcare's effect on maternal labor force participation (Aassve et al, 2012; Arpino et al., 2014; Compton & Pollak, 2013; Dimova & Wolff, 2007; Dimova & Wolff, 2011; Posadas & Vidal-Fernandez, 2013; Zammaro, 2011). Due to the presence of strong family ties (Zuanna, 2001) and other institutional characteristics, grandparental childcare that is provided regularly is relatively more common in Mediterranean countries and Southern Europe relative to Western Europe, making such regions a popular field of research (Arpino et al., 2014; Hank and Bueber, 2009). In several countries in the former regions, scholars have found positive causal links between grandparental childcare and MFLP (Aasve et al., 2012; Arpino et al, 2014).

Among the countries in these regions, given the contrast the Kingdom of the Netherlands has with regards to the childcare system and attitudes towards childcare, investigating the causal effect of grandparental childcare on maternal labor force participation is of great interest in the Dutch population. The Dutch market for childcare stands out as one of the few countries with very little to public provision of childcare (Noally & Visser, 2009; OECD, 2020). Several OECD countries, mainly Nordic countries such as Denmark, Iceland, Norway and Sweden provide early childhood care through large-scale publicly operated and/or publicly subsidized systems. Countries such as Germany, Italy and Spain allow for private not-for profit provision of childcare next to public provision. Anglo-Saxon countries on the other hand allow for private profit childcare centers to operate next to public and not-for-profit centers. According to OECD, those countries which provide early childhood care mostly publicly, because the government has direct provision, authorities have discretion over the fees charged which enables them to keep the cost of childcare relatively low compared to that of countries with market-based systems. However, because direct provision is expensive for the government, the availability of formal childcare is dependent on the level and efficiency

of the public funding which in some countries, for example in Germany results in shortage of the available formal childcare places (OECD, 2020).

In contrast, the Kingdom of Netherlands, similar to Australia, Ireland and the United Kingdom, rely on market-based systems. Public childcare support in these countries mostly consist of cash benefits or tax concessions often on an income-tested basis. The advantage of the market-based system is such that formal childcare facilities can expand in line with demand resulting with high availability of formal childcare, however the reliance on the market suggests that the government has less control over fees and hence net costs can be very high for child bearing families (OECD, 2020).

In the present context, the topic of this paper is of particular policy relevance. Between the years of 2013 and 2019, around 26,000 parents were falsely identified as fraudsters who were accused of making false childcare benefit claims, which led them to being stripped of their benefits and ordered to repay said benefits in full (Amaro, 2021). Fines amounted to tens of thousands of Euros. The scandal plunged very low-income families into crisis, resulting in many losing their personal possessions, jobs and marriages and even being evicted from their homes (Wang, 2021). According to the Bernard van Leer Foundation, after such a systemic failure, people are more aware and more demanding towards the Dutch childcare system where working women express the crucial need for a reset in the childcare system in the Netherlands. Families are expressing that the system is very complex and it must be over hauled (Bernard van Leer Foundation, 2021). For most families childcare is way too expensive and the allowances are not sufficient enough to cover expenses (Robin, 2021).

The Social and Economic Council of the Netherlands has drawn up suggestions for reforming the childcare sector in the Netherlands as a part of the next governments strategy. The reforms are planned to boost the number of people in work as parents would no longer fear losing childcare benefits if their income is increased, another proposition is to simplifying the benefit system and setting a fixed maximum price for childcare, some political parties have even called out the government to make pre-school childcare free (Robin, 2021).

Such a direction of reforms, although would incentivize labor force participation and ease mother's work and life responsibilities, before these reforms are put into effect, it is important to understand and explore the importance of grandparental childcare and see where it stands among the Dutch population. We aim to raise awareness of the importance of devoting more effort to have a more comprehensive understanding of informal childcare services, particularly the service provided by grandparents and integrate them into the discussion of reforming the current childcare system. Is grandparental childcare a reliable tool for parents in the Netherlands and to what extent should public policies incorporate grandparents into the childcare system? Was the policy of introducing payments to grandparents for regular childcare efficient in

2007 (OECD, 2011 as cited in Posadas & Vidal Fernandez, 2013). By initiating to explore the importance of grandparental childcare in the Dutch population, we aim to take a step towards answering such questions.

When the Netherlands is considered in detail given that there is little public provision, it is clear that the trade-off described between availability and costs do reflect on the childcare system. Childcare is very expensive in the Netherlands. While it comes in 10th place among the OECD countries, the Netherlands ranks 4th in Europe after Ireland, United Kingdom and Switzerland (Kelleher, 2019). Low income families struggle with childcare affordability. Supporting data from the European Union Statistics show some families cannot afford childcare in Europe and specifically for Netherlands this number is relatively high (OECD, 2020). Additionally, although childcare is available to a great extent in urbanized regions, evidence suggests that due to market dynamics, private providers relocate to urbanized areas which can lead to insufficient coverage and lack of available childcare facilities in poorer, less profitable areas (Noally & Visser, 2009).

Despite the lack of available childcare facilities and a short paid parental leave consisting of 16 weeks (Seveno, 2021), one feature of the labor market conditions of the Netherlands that could partially make up for this problem is the availability of part-time opportunities. Although a persistent concern for part-time opportunities is job quality, security, lower pay and fewer opportunities for training and promotion, relative to other countries in Europe the Netherlands is doing well with respect to ensuring good quality part-time jobs (Frey, 2019). Part-time workers enjoy high levels of employment protection and have more rights than part-time countries in most countries (Morel, 2007). For this reason, highly skilled and less skilled workers work part-time in the Netherlands as there are well-paid prestigious part-time jobs available. In fact, 26.5% of the Netherlands' highly educated workers (those with tertiary education) work part-time making up the highest share among European OECD countries (Frey, 2019). In fact, according to several scholars the high female labor force participation rate is driven due to the abundance and prevalence of part-time opportunities (Saraceno, 2008) Such institutional features also led the model of division of work and care to change where the traditional male breadwinner with a stay-at-home wife transitioned into a one-and-a-half earner model where the man tends to work full-time and the women works part-time and does most of the childcare (Kremer 2007; Plantenga 1996; Pott-Butter 1993; Verweij and Reiman, 2016). Due to such transition and the policies targeted towards increasing part-time work, Dutch parents are more often employed part-time to parents in other Western countries: in 2008, 75% of Dutch working mothers were employed part-time compared to 41% across Europe (Keuzenkamp and Steenvoorden, 2008 as cited in Geurts et al., 2014). Women's transition to part-time work is especially common after they become parents. Among all part-time workers, 75.2% are mothers aged between 35-39. After age thirty, the share of the female population working full-time declines with age (Frey, 2019).

Given the availability and the acceptance of such opportunities, one possibility of mothers is to decrease working hours and switch to part-time working to balance work and care needs of their children. Another possibility, especially for those mothers who wish to stay employed full-time is substituting a relatively easier source for formal childcare i.e. the use of grandparental childcare (Zamarro, 2011). Indeed, studies show that relative to other countries in Western Europe Dutch mothers do receive frequent help from their mothers where across Europe the Netherlands follows Southern Europe, specifically Greece, Italy and Spain in having the highest proportion of grandmothers who help their daughters with childcare Zamarro (2011).

The unique characteristics of the Dutch childcare systems and the widespread use of grandparental childcare raises the question of whether a causal link between maternal labor force participation and grandparental childcare exists among the Dutch population. Few studies have found a positive association but the direct influence of grandparental childcare on maternal labor force participation has, however not been investigated thoroughly in the literature (Aasve et al., 2012; Svensson-Dianellou et al., 2010; Zamarro, 2011). In this respect our main research objective is to fill in this gap by answering the following research question: *How does regular childcare provided by grandparents in the Netherlands affect mother's decision to participate in the labor force?* Due to issues of endogeneity and unobserved preferences related to maternal employment and childcare decisions, standard ordinary least squares (OLS) produces biased results (Arpino et al., 2014). Following Arpino et al. (2014) we will use the potential outcome framework to explain why an instrumental variables approach should be adopted. As an extension, by taking advantage of the panel data structure we will combine an instrumental variable approach with a fixed effects model so as to have further control over-time invariant unobserved preferences and individual heterogeneity among mothers.

In our study we use the Longitudinal Internet studies for the Social Sciences (LISS) core studies panel, specifically the Family and Household study, where a nationally representative sample of 7,500 individuals and 5,000 households complete online surveys every year. We conduct our main analysis, fixed effects with instrumental variable (FEIV) using the last 5 waves i.e. from year 2015 to 2020 on a sample of mothers who have at least one child under the age of 10. As our instrument we use information on whether the maternal grandmother is alive to exogenously predict the probability of using grandparental childcare. Although we aren't able to answer our research question due to the presence of weak instruments using our main methodology, both in order to observe what caused our instruments to be weak and to further investigate our research question, we employed an instrumental variable approach without the addition of fixed effects on a single wave (wave 10th, year 2017) of the panel. We found that a possible reason for our instruments to become weak was indeed the addition of the fixed effects approach. Regardless, we were unable to find a direct influence of grandparental childcare on the maternal labor force participation.

On the basis that the availability of part-time opportunities might be encouraging mothers to stay in the labor force by only decreasing their work hours or incentivizing them to enter regardless of their childcare responsibilities, as an extension to the literature, we investigated whether grandparental childcare has a direct influence on the number of hours worked weekly for working women. Although our findings must be interpreted with caution, we found that mothers who make use of grandparental childcare on average tend to significantly work a greater number of hours.

2. Literature Review

Half a century ago, women in their childbearing age had the role of giving birth, staying at home and taking care of their children. External childcare outside of the nuclear family was uncommon and so were high rates of female labor force participation (Geurts et al., 2014; Klein 1975). In contrast, in contemporary times, there has been a significant increase in female labor force participation. For the case of the Netherlands, from 1990's to 2006, women's labor force participation increased from 55% to 70% and for women aged in their maternal years i.e. between the ages 25 and 39, it increased from 67% to 82% (Thévenon, 2013).

The increased labor-force participation among Dutch women, meant that they would now have to find a healthy balance between their work and care responsibilities. As mentioned, part-time opportunities are relatively more common in the Netherlands compared to other Western European countries. While 23% of Dutch male workers are employed part-time, this number rises up to 75% for Dutch female workers compared to an average of 41% across Europe (Keuzenkamp & Steenvoorden, 2008). Common strategies for Dutch families include maximizing the time spent at home by making use of these part-time opportunities and working at jobs which have flexible working hours to increase their time available for childcare (Carriero et al., 2009; Cloin and Hermans 2006). Supporting this idea, Frey (2019) claims that Dutch women's transition to part-time work is especially common after they become parents. Even after most schooling is completed, and within the main child-rearing years, the share of the female population who are employed part-time reaches 44% among 40- to 44-year-olds. In this group of women working part-time, 66.9% are mothers of children aged 14-years-old or younger. 75.2% of part-time workers in the 35-39 age group are mothers (Frey, 2019). Such figures reflect the one-and-a-half earner model where Dutch mothers seem to prefer to stay in the labor force but manage to do so by decreasing their working hours which explains why the Netherlands has one of the highest female labor force participation in the Netherlands among OECD countries (Frey, 2019)

Although the part-time strategy provides considerable ease, as more mothers have paid jobs outside the home, external childcare is more often needed. Those who do wish to stay in the labor force, or in the

context of the Netherlands those who wish to keep their regular working hours have to opt for formal childcare or informal care for their younger children.

Formal childcare, due to the market-based childcare system, the cost of formal childcare is very expensive and although largely available in urbanized regions there is an insufficient coverage in more rural regions (OECD, 2020). Even in cases when it is not free of charge, grandparental care is less expensive than formal childcare where three out of four Dutch parents do not pay their parents for child-care activities (Portegijs et al., 2006). In this respect, for some working women, grandparental care might represent the most or only affordable or available option. It might also be that it is the most convenient with respect to financial concerns or the location of the grandparent's residence as a decrease in travel time between grandparents and their adult children is likely to increase the use of grandparental childcare (Harms, 2008).

Other factors that explain the frequent use of grandparental childcare is the attitudes and opinions of the Dutch culture towards formal childcare and family relationships. According to Portegijs (2006), Dutch mothers consider grandparents to be the best child-care providers. Wheelock and Jones (2002) put forward that parents that rely on grandparental childcare usually do so, not because of financial reasons or issues related to availability but because they trust their own parents more than formal child-minders and share the belief that grandparents are capable of providing the safest and the most emotionally nurturing environment that benefits the development of their small children, whereas some hold the opinion that formal childcare can even be potentially harmful (Thomese & Kiefbroer, 2013; Treas and Widmer 2000).

Choice of childcare is a joint decision between children's parents and the grandparents, the ability and willingness of the grandparent to take care of their grandchildren is also a factor that contributes to the extent to how much a mother resorts to grandparental childcare. Grandparents that have the opportunity can be motivated to provide childcare for many reasons along with having many constraints to do so. Studies that investigate motivation to provide childcare find that affection towards the middle generation or the grandchild, the desire to support, secure and improve her own child's labor market position are important factors for grandmothers in their decision process (Thiele & Whelan, 2006). Reciprocity is another significant motivator as some believe that their efforts of providing childcare will be repaid by the parent at a later stage when the grandparent is in need of health assistance or support (Silverstein & Giarusso, 2010). Additional to willingness, the grandparents must have sufficient health and also must not have conflicting commitments such as paid work or taking care of their grandchildren from other children (Thomese & Kiefbroer, 2013).

Grandparents are more likely to provide childcare if their daughter is a single working parent, less likely to provide care in larger families and less likely if they live far away from their grandchildren (Hank & Buber, 2009; Igel & Syzdik, 2011; Thomese & Kiefbroer, 2013). Due to the strongly and rapidly ageing workforce of the Dutch population (NL Times, 2021; Remery et al., 2003) from all of these factors,

specifically for the case of the Netherlands, the grandparent's employment status is an especially significant factor in the decision to provide childcare as grandparents time for caring is restricted when employed (Del Boca, Pasqua and Pronzato, 2008).

Alongside an increase in elderly labor force participation, an increase in divorce rates, could have changed the needs of mother and also opportunities for grandparents to provide childcare. Between 1996 to 2006, the proportion of divorced parents with children under 18 years old have increased by 10% (Kalmijn, 2013). It is often observed that the grandparents are more likely to provide childcare when the parent is single rather than when the parent has a partner (McLanahan & Sandefur, 1994), and interestingly between 1995 and 2008, the number of single-parent families in the Netherlands increased by 30% (Statistics Netherlands, 2008) suggesting higher usage of grandparental childcare in the Netherlands.

Regardless of what the underlying reason is, statistical figures do show that grandparental childcare is common in Europe and even more common in the Netherlands. Hank and Buber (2009), using the Survey of Health Ageing and Retirement of Europe 2004 (SHARE) data show that within Europe 49% of grandfathers and 58% of grandmothers with grandchildren aged 15 or younger provide some kind of care for a grandchild. A study investigating the international variation of preferences to informal/grandparental childcare shows that among European countries, the highest prevalence of help provided by grandparents is found in Sweden, France, the Netherlands and Denmark, particularly in the latter two countries with around 57% of grandparents looking after a grandchild in the past 12 months (Glaser et al., 2010). Another supporting study finds that the use of grandparental childcare is particularly high in the Netherlands and Hungary (Aassve, 2012).

Although it is common to make use of grandparental childcare in the Netherlands, it must be noted that intensity and the frequency of informal care is much less relative to the childcare provided by grandparents in Southern Europe. While around 57% of the grandparents provide help in childcare activities compared to the 40% of those in Southern Europe (Glaser et al, 2010), only 2 percent of grandparents provide regular/daily childcare in the Netherlands (as opposed to the 20% in Europe) (Geurts et al, 2014). Instead of looking after their grandchildren on a regular basis, the childcare provided is much more dispersed and is provided when needed. Dykstra (2012) states that such differences are related to the availability of public policy arrangements and the extent to which mothers are employed full-time or part-time (Herlofson, 2011).

In the context of the Netherlands, the availability of part-time opportunities is expected to influence the intensity of care received from the grandparent. Bordone et al. (2017), using SHARE finds a positive association between the share of female part-time jobs and weekly grandparental childcare. Among the countries in the study, the Netherlands shows the highest proportion of women working part-time (74.7%) and the highest proportion of weekly grandparental childcare (42.1%) but a very low rate of

grandparents who are providing regular care indicating that Dutch families have a preference towards informal childcare but simply spend more time at home and hence are less in need of their mothers to step in for help regularly.

Given the institutional features of the current childcare system characterized with high prices and lack of coverage, the positive cultural attitudes and opinions toward informal childcare and the extent to how widespread it is one would expect the use of grandparental childcare to affect maternal labor force participation positively. On the other hand, due to the Dutch mother's preference to work part-time and their tendency to reduce working hours instead of exiting the labor force and several empirical studies showing that high participation rates in the Netherlands have been driven by the availability of part-time jobs (Saraceno, 2008) it is also expected that regular grandparental childcare does not have any direct influence on the working decisions of mothers.

While demographic and sociological studies regarding grandparental childcare and how it relates to the employment decisions are very abundant in the literature, empirical studies are not as common. We believe this is mainly due to the endogenous nature of grandparental childcare and the lack of data availability for research purposes.

The issue of unobserved heterogeneity between mothers' preferences and grandparental characteristics have been dealt with several ways in the literature. Del Boca (2002), adopts an intention-to-treat (ITT) analysis where she uses the key variable 'having at least one grandparent' as a randomization device and later analyzes individuals according to the group they were originally assigned regardless of the treatment they received thereby allowing her to draw unbiased conclusions regarding the effectiveness of an intervention. In this context, according to Arpino et al. (2014), ITT estimates can be downward biased because while only individuals with mothers alive are able to use childcare, not all individuals with alive mothers are able to actually use childcare.

Similar to Del Boca (2002), Ogawa and Ermisch (1996) make use of an ITT-like approach in their study investigating female labor force participation and intergenerational co-residence and nuclear structured families in Japan. The analysis is an intermediate between an ITT analysis and using indicators for alive grandparents and an instrumental variable using actual help received by grandparents. Their findings pointed towards a positive relationship between intergenerational co-residence and the labor supply of young women reflecting the childcare role played by the maternal and paternal grandparents. Using the European Social Survey, Abendroth et al. (2012), studied the relationship between intergenerational co residence and the working hours of mothers by employing a Heckman selection model but wasn't able to find a significant effect of having at least one grandparent living in the household on the number of working hours of employed mothers.

Relating to intergenerational co-residence, other scholars have studied geographical proximity to grandparents and its relation to maternal labor force participation using several methods. Garcia-Moran and Kuehn (2012), build a general equilibrium model of residence choice, fertility and female labor force participation. Through a simulation approach, they find that having grandparents reside close increases the probability of their daughter participating in the labor force. Additional to this finding, through counterfactual simulations they also show that without childcare provided by the grandparent there would be fewer women participating in the labor market.

Supporting Garcia-Moran and Kuehn's (2012) findings, Compton and Pollak (2013) find that close geographical proximity to maternal or paternal grandmothers has a significantly positive effect on the labor force participation of women with young children. To deal with the endogeneity of grandparental childcare they adopt a bivariate probit approach where they instrument whether the grandparents reside in close proximity or not. They argue that proximity is a good instrumental variable for childcare arrangements as the positive effect of proximity on labor force does not affect labor force participation other than through child-care arrangements, except for certain exceptions such as those that have to reside close to their mothers due to the mother's needs of assistance related to health and old age.

Hank (2007) points out that family values is a very significant factor that possibly determines the proximity between the mothers and the daughters which in turn might be associated with unobserved preferences about female labor force participation. To alleviate these concerns Compton and Pollak (2014) implement falsification tests where they show no clear evidence that family values are associated with female's employment decisions. Nevertheless, mothers may proactively choose to live in close proximity to their mothers anticipating the need or desire to use grandparental childcare in the future. Supporting this view Leopold (2012) finds that young adults who have their own children tend to stay closer to their parental home. As these concerns cannot be ruled out with the falsification tests of Compton and Pollak (2013), geographical proximity of grandmother's might not be a credible instrument. This will be put to test in the robustness checks section of this paper.

Several scholars, given that geographical proximity might not be a valid instrument, have adopted different instruments. Aassve et al. (2012) use a bivariate probit model to investigate the causal relationship between grandparental childcare and maternal labor force participation using information on whether the maternal grandmother is alive in addition to the number of siblings of the respondent. According to Arpino et al. (2014), the second instrument could be problematic if the size of the respondent's family is correlated to unobserved preferences for childcare and labor force participation decisions.

Additional to a bivariate probit model, in the light of the potential outcome framework, the two-stage-least-squares (2SLS) has often been employed by scholars the literature due to ease of reading and interpretation. Arpino et al. (2014), investigates the relationship between grandparental childcare and

maternal labor force participation in Italy by employing a 2SLS approach on a sample of mothers with partners that have at least one child under the age of 14, he instruments grandparental childcare with four instruments where each instrument is binary and indicates whether the four grandparents are alive or not. They find that the effect of grandparental childcare is statistically significant and economically relevant and that this effect is stronger for less educated mothers who have younger children and live in northern or central Italy. Using the exact same methodology, but with only one instrument i.e. whether the maternal grandmother is alive or not, Posadas and Vidal Fernandez (2013) make the first attempt to measure the effect of grandparental childcare on maternal labor force participation in the United States with a large sample of mothers that have children under the age of 14.

Dimova and Wolf (2011) dealt with the endogeneity of grandparental childcare using a bivariate probit model and also a three-stage-least-squares (3SLS) approach, where they jointly estimated three equations for mothers working decisions and the probabilities to receive time and money transfers from grandmothers by allowing the error terms in the three models to be correlated. The authors exclude some variables related to grandparents' characteristics from the mothers' labor supply equation using them implicitly as instrumental variables. For example, they exclude grandparents' age and health status that might well be factors influencing the daughter's decision to work or not. Additional to the 3SLS model they employ a fixed-effects analysis to control for unobserved household-level characteristics. They conduct this study using data on grandparental childcare from ten European countries, hence the effect they find represents an average effect from people residing in the ten countries included in the study. For Europe as a whole, they find a strong positive effect of grandchild care on the labor force participation and the degree of labor market involvement of the mother. However, Aassve et al. (2012) points out, that this approach might be problematic due to the possibility of the estimated average representing the country-specific effects poorly.

To our knowledge there has been only one study that empirically investigates the causal relationship between grandparental childcare and maternal labor force participation specifically in the country of the Netherlands. Aassve et al. (2012) adopts a bivariate probit model approach, and tests two different instruments to tackle the endogeneity of grandparental childcare among 7 other European countries including the Netherlands. He instruments whether the grandmother of the grandchildren is alive and the number of siblings the individual has. Although he finds a very significant and positive association by using a probit model, when the model is corrected for endogeneity, it is seen that the probit approach has an upward bias and that grandparental childcare does not have a causal effect on maternal labor force participation in the Netherlands. Because we were able to find only a single paper investigating the causal effect of grandparental childcare empirically in the Netherlands, we believe there is a gap that needs to be filled in the literature and that several other methods using other datasets must be made use of in order to

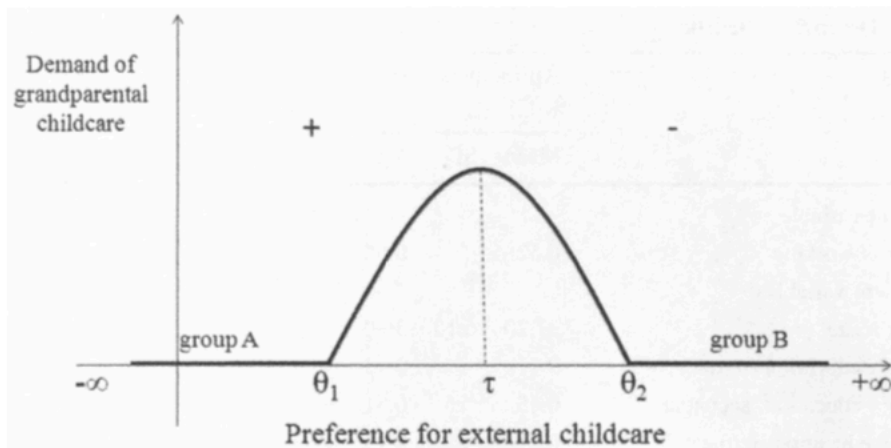
understand the effect of regular grandparental childcare on maternal labor force participation. Following the information, we have gathered through the literature, *we hypothesize that regular childcare provided by grandparents in the Netherlands does not have a positive direct influence on the mother's decision to work.*

Our analysis will contribute to the literature by combining an instrumental variable approach and a fixed effects model to account for the unobserved heterogeneity. In this respect our study contributes to the literature by not only accounting for the possible endogeneity of grandparental childcare, but also the unobserved heterogeneity among mothers. While the two-stage least squares (2SLS) approach enables us to create a false experiment environment and eliminates the problem of endogeneity, with the addition of the fixed effects we will attempt to have further consistent estimates that are more robust to biases.

3. The Endogeneity of Grandparental Childcare

According to the theoretical framework formulated by Arpino et al. (2014), the endogeneity of grandparental childcare can be explained by the degree of the 'externalness' of childcare preferred by the mother. The stronger the mother prefers taking care of her own child herself, the higher the minimum wage she would be willing to accept to participate in the labor market. In contrast, mothers who have a higher preference for participating in the labor market might have a stronger preference for more external types of childcare. If it was assumed that there is a positive association between preference for external childcare and mother's willingness to participate in the labor market, we would have a spectrum with two distinct group of mothers, where group A is composed by mother's who prefer to provide care of their solely their own and group B which has a preference toward the most external type of childcare i.e. formal childcare facilities. In between the two extremes, present are the mothers who are willing to accept help in childcare from grandparents as it is external to the mother, but internalized to a degree as the care is provided by the own mother of the individual (Arpino et al., 2014).

Figure 1 Theoretical Framework for the Demand of Grandparental Childcare (Arpino et al, 2014)



According to Arpino et al. (2014) up to a certain threshold T , the preference for external types of childcare is positively associated with the probability to accept grandparental childcare. For lower levels of T this relationship is now negative, because for higher levels of preference for externalness, mothers start preferring formal childcare facilities to grandparental childcare.

In this context, the non-linear effect of grandparental childcare will produce an overall upward or downward bias, and the direction of this bias for each individual will depend on the individual's unobserved preference regarding informal childcare. As discussed in the literature review, mothers can have many different reasons for opting for grandparental childcare instead of formal childcare.

Another potential source of bias is the unobserved characteristics of the grandparents. Even if the mother prefers grandparental childcare, some may not be able to receive it given that it is a joint decision that mothers and grandparents have to make together. Grandparental childcare depends on grandparent's willingness and ability to do so. In this respect, the grandparent's employment status is of great importance. Grandparents, might not be willing to retire early for childcare purposes or not they might not be willing to forgo their leisure time for childcare purposes. Such grandparents, on the basis that they are not traditional, are more likely to have female children that are less traditional as well. Such individuals more likely to be participating in the labor market, introducing a downward bias to the relationship between grandparental childcare and mother labor force participation. Additionally, although they are willing to provide help, they might not have the ability to do so. Elder parents, might be too old to provide childcare and might not have sufficient health. There might be cases in which the mother drops out of the labor force to provide informal care to her parents.

Given such unobservable factors, if one were to adopt an ordinary least squares (OLS) estimation, the estimation is expected to produce biased results casting doubt on their causal interpretation. To overcome this problem of endogeneity, one of the tools we will be using a 2SLS approach.

However, while the 2SLS approach helps overcome the problem of endogeneity by creating a false experiment environment as explained further below, the use of such instruments may also create additional problems as the strict assumptions of the IV method may lead to the possibility of the creation of certain biases.

One example for such a situation would be a correlation between the death of the maternal grandparents and mothers labor force participation through channels other than the availability of grandparental childcare. Lang and Zagorsky (2001) find that parental death is correlated both with variables such as lower health or education which is also related to poorer labor force participation decisions. If the grandparents are alive because they have access to high-quality healthcare and are socio-economically advantaged, it is likely to be the case, their daughters are well-educated (Schwartz and Mare, 2005), increasing the probability that they enter or stay in the labor market (Mulligan and Rubinstein, 2005; Posadas

and Vidal-Fernandez, 2012). In this case IV estimates are expected to overestimate the effect of grandparental childcare on MLFP. On the other hand, grandparents might still be alive, because they are younger and had children earlier in life, and come from a disadvantaged socio-economic background, thereby affecting their daughters labor market decisions. In contrast to the former scenario, IV estimates would underestimate such an effect (Posadas & Vidal-Fernandez, 2012).

Another source of bias could originate from the health conditions and status of the maternal grandparents. Posadas and Vidal-Fernandez (2012) argue that grandmothers or grandfathers who passed away are likely to have spouses that are also older in age as well. In this case, the middle-generation might have to provide informal care to take care of her sick widowed mother her father, thereby causing an upward bias on the effect of grandparental childcare of MLFP. Additional to this issue, grandparents who have passed away, might have passed away due to a hereditary illness and also could have transmitted the disease along to her daughter preventing her from entering or staying in the labor market similarly causing an overestimation in the IV estimates (Posadas and Vidal Fernandez (2012).

As discussed above although the 2SLS method is a good way to deal with the endogeneity and is more reliable with respect to the estimates produced by an OLS approach, certain biases are still at play. The following section describes how we will incorporate a fixed effects (FE) approach to an instrumental variable approach to further reduce biases that are caused by unobserved preferences so as to reach more consistent estimates.

4. Research Methodology

4.1. Fixed effects regression with instrumental variables

We employ a binary treatment, *gcc* i.e. grandparental childcare (=1 for mothers who use grandparental childcare on a frequent basis, =0 otherwise), and a binary outcome *mlfp* i.e. maternal labor force participation (=1 for mothers who are employed, =0 otherwise). Each individual has two potential outcomes: *mlfp₁*, her employment status if she was treated i.e. if she makes use of frequent grandparental childcare and *mlfp₀*, if she is in the control group i.e. she doesn't make use of grandparental childcare. If there weren't unobserved preferences with respect to mother employment and childcare decisions, the following model could have been used to estimate the effect of interest:

$$mlfp = \alpha + \delta gcc + \sum_k \beta_k X_k + e$$

where $mlfp = mlfp * gcc + mlfp_0 * (1 - gcc)$ is the observed outcome and α δ and β represent parameters to be estimated. δ is our parameter of interest and e represents the error term. Due to the unobserved preferences, the error term might be correlated with our outcome variable *mlfp*, causing the OLS estimates to suffer from bias. To deal with this problem of endogeneity, and as an extension to Arpino et al. (2014)

we will combine the potential outcome framework (Neyman 1923; Rubin 1974) and a fixed effects model by instrumenting for our endogenous variable *gcc* so as to estimate the causal effect of grandparental childcare on maternal labor force participation.

As our instrumental variable, by arguing that the death of a grandparent is an exogenous event, we will use whether the maternal grandmother is still alive or not. The paternal grandparents and the maternal grandfather are not included because the LISS panel does not include information on whether the paternal grandparents are alive, and on the basis that grandparental childcare is strongly gendered, with grandmothers typically being involved most often and most intensely (Koslowski, 2009). Studies conducted through the Survey of Health and Retirement in Europe show that (SHARE) grandchildren are mostly cared by maternal grandmothers, and then by maternal grandfathers, followed by paternal grandmothers and paternal grandfathers (Posadas & Vidal-Fernandez, 2012). Consequently, we have one instrument: *gmomalive* (=1 if the grandmother is alive, =0 otherwise). The chosen instruments have to affect the likelihood of using grandparental childcare (relevance) without having a direct influence on the mother's employment status (exclusion restriction). These instruments must only operate on *mlfp* only through their effect on the treatment, *gcc*. *gcc* should not share common causes with *mlfp* (exchangeability) (Angrist et al. 1996). In this respect, the instrumental variables approach can be considered as a randomization device that creates a false experiment environment i.e. the instrument, conditional on the control variables, assigns units to either the treatment or the control groups independently of the outcome. Given providing grandparental childcare is a joint decision, we will first assume that grandparents are willing and able to supply care when and later consider the bias due to unobserved preferences regarding this joint decision.

We will use a standard 2SLS estimator despite the fact that both the endogenous and the dependent variables are binary as the interpretation of the results is more simple and that tests on the instrumental variables can be easily implemented through user-written commands of Stata. To confirm endogeneity, we will make use of the endogeneity test implemented by the user-written command *xtivreg2* of Stata. The endogeneity test reports test statistics that are robust to various violations of homoskedasticity. The null hypothesis is that the specified endogenous regressors can actually be treated as exogenous.

We will consider the first stage F-statistic, Kleibergen-Paap rk Wald F test and the Kleibergen-Paap rk LM statistic to test the validity of our instruments. Bound et al. (1995) and Staiger and Stock (1997) support the use of the first stage F-statistics. They show that there is a close relation between the first stage F-statistic and the bias of the IV estimator. It must be noted that we are correcting for heteroskedasticity so that our estimates are robust the presence of outliers. According to Baum, if the robust option is specified, the reported weak-instrument tests statistics is a Wald F statistic based on the Kleibergen-Paap rk statistic, and the Cragg Donald Wald F statistic is no longer valid. Kleibergen – Paap rk Wald F-statistic should be used instead because the robust analog is a sensible choice and clearly superior to the use of the former in the

presence of heteroskedasticity (Baum, 2007). Baum (2007), further states that in the special case of one endogenous regressor, the Cragg–Donald and Kleibergen–Paap statistics reduce to familiar statistics available from an OLS estimation of the single reduced-form equation with an appropriate choice of variance-covariance estimator. In the case of a single endogenous regressor, Staiger and Stock (1997) a first-stage F statistic larger than 10 is sufficient for an instrument to be considered relevant. Additional to the rule of thumb, Stock and Yogo (2005), proposes that the weak ID test critical values proposed by Stock and Yogo (2005) are also valid for the Kleibergen – Paap rk Wald F-statistic. The critical values proposed by Stock and Yogo are presented in the Appendix A.

Additional to relevance, we conduct a test of under identification. The under-identification test, tests whether the (excluded) instruments are not correlated with the endogenous regressor. To see if this is the case for our model, we will use the rk LM statistics proposed by Kleibergen and Paap (2006) which is an appropriate robust statistic for testing the rank of a matrix when any sort of robust covariance estimator is used.

As an extension to Arpino et al. (2014) and Posadas et al. (2012), we combine our 2SLS approach with a fixed-effects approach. The reason is as follows: In contrast to the data used by Arpino et al. (2012) uses to investigate the causal effect of grandparental childcare on MFLP of Italian mothers, the LISS panel data provides a relatively small sub-sample of interest. Given that 2SLS might produce biased results due to a small sample size, we believe taking advantage of the panel data structure alongside our 2SLS approach will enable us to have more control over time-invariant unobserved preferences and thus control for individual heterogeneity among mothers, so as to have more consistent estimators and investigate whether grandparental childcare significantly influences MFLP in the Dutch population.

For our instrumental variables fixed effects analysis, to control for the unobserved individual characteristics, we will assume that the error term u_{it} has two components. The first component, $\alpha_i(\beta_0)$ will accommodate individual characteristics that do not vary over time. α_i will be allowed to correlate with the observable regressors $gcc(gmomalive)$ and X_k . The second component, the error term e_{it} , is assumed to be uncorrelated with all observations for our explanatory variables, varies across time and across individuals, and is assumed to be independent and identically distributed (Woolridge, 2009). Given the panel structure of the data, because there are multiple observations in time for individuals, autocorrelation between observations is allowed. However as not correcting for autocorrelation would increase the likelihood that our coefficient estimates are further than the correct population value (Yobero, 2016), we will we will cluster standard errors at the individual level so as to obtain heteroskedasticity and autocorrelation-consistent standard errors.

In this respect, the following equations will be simultaneously estimated to carry out the FEIV method. Given our choice of model, we make comparisons within individuals by taking the average of the differences across individuals. The within transformation is as follows:

$$(gcc_{i,t} - \overline{gcc_i}) = \beta_1(gmomalive_{i,t} - \overline{gmomalive_i}) + \sum_k \beta_k (X_{k,i,t} - \overline{X_{k,i}}) + (v_{i,t} - \overline{v_i})$$

$$(mlfp_{i,t} - \overline{mlfp_i}) = \delta(gcc_{i,t} - \overline{gcc_i}) + \sum_k \beta_k (X_{k,i,t} - \overline{X_{k,i}}) + (e_{i,t} - \overline{e_i})$$

5. Data Description

The LISS Core Panel consists of nationally representative sample of 5,000 households, comprising approximately 7,500 individuals. It is a longitudinal study repeated yearly and is designed to follow changes in the life course and living conditions of the panel members. The study covers 8 questionnaires, on a wide range of topics i.e. health, politics and values, religion and ethnicity, social integration and leisure, personality, economic situation, work and schooling and family household. It has been in full operation since 2007, and the last available wave is for the year 2021.

For this study the two main questionnaires that will be used within the LISS core panel is the Family and Household and Work and Schooling. The questionnaires are exceptionally suited for the purpose of this study because since the first wave in 2007, it contains detailed information about childcare arrangements for the child's first five years of life i.e. informal care by relatives, formal childcare facilities such as day-care centers, pre-schooling and also contains detailed information on level of education the individual has completed, details of current and past employment e.g. number of contracted hours worked, number of actual hours worked and other information regarding employment characteristics. Furthermore, the study provides background information on the mother's income, household size, the urban character of residence, proxies regarding health status of relatives and information on distance between the grandparents and the child's parent

For our study, we will create a sub-sample using information on approximately 5000 Dutch mothers, who have at least one child that is younger than 10 years old. In choosing the age of 10, we followed the American Academy of Pediatrics as they state that until about age 11 or 12 most children require supervision and are not able to be left on their own (Schor, 1999). Although single mothers might have large differences in terms labor force participation, household income characteristic's and the frequency in which they make use of grandparental childcare, we decided to include them in order to get the whole picture on the effect of grandparental childcare on maternal labor force participation in the Netherlands.

Our main analysis will be conducted with the last 5 waves i.e. 8-13th (2015-2020). Waves 1 to 7 will be excluded because with given the variables created it is not possible to trace the number of alive children

the individual currently has, as the survey question is asked in a way that the answer might include the number of children who might have passed away as well i.e. “how many children have you had in total?”

For our second analysis which is further clarified in the results section of this paper, we will make use of the 10th wave (2017). Our reasoning of choice is explained by the 10th wave being one of the most recent wave that provides the sub-sample with the greatest number of observations i.e. 508.

5.1 Endogenous variable: grandparental childcare (*gcc*)

In the data set, the respondents were first asked the following question: “For your living-at-home children born from (year) onwards, do you make regular use, that is at least once a week, of the following childcare options?”. The (year) specified is so that the children of the individuals taking the survey are not older than approximately 4 years old for each wave. For those respondents who have responded that they do, a follow up question we used to construct our endogenous variable, *gcc*, was asked. The question is: “Who is generally this unpaid child sitter?” Possible options consisted of the following: 1 own parent, 2 partner’s parents, 3 neighbor, 4 friend acquaintance, 5 brother or sister, 6 other own child, someone else.

We construct our variable *gcc*, as a binary variable where the variable *gcc* equals 1 given that own parents i.e. the maternal grandparents provides grandchild care, and 0 for all other options.

5.2 Outcome variable: mother’s labor force participation (*mlfp*)

Respondent was asked whether they have paid work, and a binary variable indicating mother’s employment status was constructed where *mlfp* equals 1 if the mother’s employed and 0 if not.

5.3. Instrumental variable: Maternal grandmother life status (*gmomalive*)

Respondents were asked whether their mothers were still alive. Using the provided answers *gmomalive* was created where it equaled 1 if the grandmother was alive and 0 if not.

5.4. Explanatory variables

Other variables which are likely to affect mother’s labor supply decisions and hence should be controlled for are age and education of the mother. Due to our small sample size and to see differences between different age groups in terms of their labor force participation we created 4 dummy variables representing 4 different age categories i.e. 16-24, 25-34, 35-44, 45-54 and chose the mothers aged between 16-24 as the reference category, as that specific age group would have a lower probability of participating in the labor force. To control for differences in the education levels of the mother we created 2 dummy variables i.e. secondary and tertiary education. The base category is those who have completed only primary education or have not started or not yet completed any education. We chose not include information on the

level of education completed by father/partner. We followed the literature in doing so as the papers either did not include such a variable (Posadas & Vidal-Fernandez, 2012) or the ones that did were not able to find any significant differences in maternal labor force participation and the level of education completed by the father/partner (Arpino et al., 2014).

Additionally, we can only reach the individual's partners background information through the household box of the survey of the LISS panel, where the individual, before joining the panel completes a survey on the background characteristics of the members in her household and updates it every month for the changes. Unfortunately, as we are only able to trace partner's information using a variable that indicates the household that individual belongs in, it is complicated to trace the individual's information for households that have more than one older aged male. Additionally, as not all individuals have filled out the survey, the inclusion of the father's characteristics, leads to a considerable increase in the number of missing observations hence decreasing the statistical power of our analysis.

We then included two variables summarizing the couple's fertility history; a dummy variable indicating whether the couple has at least 1 child between the ages of 0-4 (the 4th age is not included) and the number of children younger than 10 years of age. Even though compulsory education starts at 5 years of age in the Dutch education system, according to the European Commission in the Netherlands, nearly all children attend primary school from the age of 4 (European Commission, 2021; Expatica, 2021; Herwig, 2021). Hence by constructing this dummy variable we are able to distinguish them from mothers who have the opportunity to make use of free schooling services as with the exception of private schools, most Dutch primary and secondary schools are free (Ministry of Education, Culture and Science, 2014; Ministry of General Affairs, 2021).

Ideally, we would have liked to use the region of residence of the individuals so as to construct a set of dummy variables for indicators of the childcare system as well to control for the difference in family practices and norms prevailing in different regions but vary within the same country (Jappens and Bavel, 2012). Unfortunately, due to identification and privacy reasons, the LISS panel does not provide this variable of interest. However, it includes a variable that measures the degree of urbanization with respect to the density per km² of the surrounding address the respondent resides in. The categories vary from 1-Extremely urban to 5-Not urban. The availability, quality and cost of formal childcare varies across regions in the Netherlands with respect to how urbanized the regions are (Noally & Visser, 2009). Furthermore, childcare choices may be related to family practices which vary across urban and suburban regions (Smith, 2006). For this reason, although we do not have the mentioned indicators, using the variable measuring degree of urbanization we will control for very urbanized areas and very suburban areas.

Because single mothers are expected to provide for themselves and partly for their children, it is expected that they are more likely to participate in the labor force and also are more likely to make use of

grandparental childcare on the basis that they potentially have less help in assistance with taking care of kids. In order to control for these possible differences, we create a dummy variable that indicates whether the mother is single.

6. Descriptive Statistics

Table 1 Summary Statistics

Variables	No. of Obs.	Mean	St. Dev.	Value=1	Min	Max
Maternal Labor Force Participation	2,484	0.730	0.444	1,813	0	1
Age	2,484	36.79	5.540		18	56
16-24	2,484	0.0105	0.102	26	0	1
25-34	2,484	0.343	0.475	852	0	1
35-44	2,484	0.567	0.496	1,409	0	1
45-56	2,484	0.0793	0.270	197	0	1
No education or primary education	2,484	0.0193	0.138	48	0	1
Secondary education	2,484	0.122	0.327	303	0	1
Tertiary education	2,484	0.859	0.348	2,133	0	1
At least one child between 0-4	2,484	0.365	0.481	906	0	1
Number of children under 10	2,484	1.528	0.640		1	5
Very urban	2,484	0.138	0.345	342	0	1
Not urban	2,484	0.191	0.393	474	0	1
Single mother	2,484	0.133	0.340	331	0	1
Grandparental childcare	2,484	0.153	0.360	381	0	1
Grandparental help	2,484	0.369	0.483	723	0	1
Grandmother alive	2,484	0.915	0.279	2,273	0	1
2020	2,484	0.166	0.372	412	0	1
2019	2,484	0.138	0.345	343	0	1
2018	2,484	0.162	0.369	403	0	1
2017	2,484	0.205	0.403	508	0	1
2016	2,484	0.151	0.358	375	0	1
2015	2,484	0.178	0.383	443	0	1

Table 1 shows that on average Dutch mothers who have at least one child under the age of 10 tend to participate in the labor force as the mean of the binary variable is 0.73. If each observation is considered as

a separate individual, 1,813 mothers work out of 2,484 making up 73% of our total sample. It is seen that the most common age to have a child under the age of 10 is between the ages 34-44. Among the sub-population we see 86% of mothers have completed tertiary education and those who haven't completed tertiary education remain as outliers indicating that the literacy rate between mothers in the Netherlands is very high.

The average of our dummy variable indicating whether the mother has a child between the ages 0 to 4 is 0.37 and the number of such mothers in this sub-sample amount to 906 people out of 2484. We believe that the abundance of such mothers in our sub-sample provides an adequate composition for our analysis as those mothers are the ones that are expected to make use of grandparental childcare the most.

Table 1 also illustrates that most of the sub-sample chooses not to reside in very urbanized regions as the mean of the dummy variable is only 0.14 percent. Similarly, relative to married mothers or mothers with partners, single mothers are in the minority where the number of single mothers make up 13% of the sample.

Supporting the literature, we see that the number of grandparents who provide help i.e. several times in the past three months are almost twice the number of those who provide regular childcare. Another interesting point worth mentioning is the relatively high proportion of alive grandmothers and in contrast, the low proportion of individuals who receive or make use of regular childcare. Such statistics tell us that although there are a lot of grandmothers that are in good health and could provide childcare, a very small number of mothers receive regular grandparental childcare. Considering the high maternal labor force participation rate, our descriptive statistics already give us some hints regarding the existence of a causal relationship between grandparental childcare and maternal labor force participation.

7. Results

7.1 Fixed Effects

Based on our discussion of the endogeneity of grandparental childcare we start by presenting our fixed effects results without instrumental variables so as to later compare them with the results of our FEIV approach. Before we present our results, in order to see whether the model is correctly specified i.e. whether a fixed effects or a random effects model is more appropriate we employ a Hausman test. Following the recommendations of Wooldridge (2020), we conducted the Hausman test with an additional option i.e. `sigmamore` using the statistical software Stata. According to Wooldridge (2020), when a regression contains some variables that don't change over time such as education, the Hausman test without the `sigmamore` option reports an incorrect number for the degrees of freedom and the `sigmamore` option enables Stata to compute the correct degrees of freedom. Using the option, with a p-value of .0000 we reject the null

hypothesis that the difference in the coefficients are not systematic confirming the appropriateness of a fixed effects approach. The null hypothesis is also rejected without using the sigmamore option.

Table 2 Fixed effects estimates of the effect of grandparental childcare on mother’s labor force participation

Dependent variable	Maternal Labor Force Participation
Grandparental Childcare	0.0077 (0.0197)
Ages: 25-35	-0.0614 (0.0638)
Ages: 35-44	-0.0589 (0.0697)
Ages: 45-54	-0.0856 (0.0786)
Secondary education	-0.1175 (0.2640)
Tertiary education	0.2961 (0.3026)
At least one child between the ages of 0-4	0.0157 (0.0204)
Number of children under 10	-0.0353** (0.0163)
Single mother	0.0657 (0.0595)
Very urban	0.0195 (0.0444)
Not urban	0.0497 (0.0393)
Constant	0.5773* (0.2995)
Observations	2,484
Number of groups	904
R2	0.027

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2 presents the estimates of a fixed effects model where the maternal labor force participation of the individual is regressed on grandparental childcare. Our results fail to find a significant relationship between grandparental childcare and the probability of the mother to have paid work. Similarly, our results highlight no significant differences between the level of education mothers have completed and the urban characteristic of their region of residence. As it becomes harder to juggle work and life responsibilities, reasonably one would expect a significantly negative relationship between the number of children/having at least one child between the ages of 0-4 and the mother’s employment status. Our results do not indicate

such a relationship; however, the number of children under the age of 10 is negatively and significantly associated with labor force participation.

7.2 Fixed Effects with Instrumental Variables

Because of the potential endogeneity of grandparental childcare, we have employed a 2SLS approach alongside our fixed effects. To see if the chosen method is valid, similarly, we will first check whether the fixed effects is more suitable compared to a random effects model when combined with an instrumental variables approach, later we will attempt to verify that grandparental childcare is indeed endogenous and make sure that the fundamental assumptions of the IV approach are fulfilled.

After running random effects instrumental variables and FEIV regression, we conducted the Hausman test. With a p-value of 0.0019 we are able to reject the null that the errors are not systematic and conclude that combining our instrumental variable approach with a fixed effects model is more appropriate.

To confirm the endogeneity of grandparental childcare, we will use the endogeneity test implemented by `ivreg2`, a user written Stata command (Baum et al., 2010). Unlike the Durbin-Wu-Hausman test statistic, this endogeneity test reports statistics that are robust to various violations of conditional homoskedasticity (Baum et al., 2010). With a p-value of 0.246, we are unable to reject the null hypothesis that grandparental childcare is exogeneous. In our case, we believe that the reason we fail to reject the test of endogeneity is the strength of our instrument and that the precision of our IV estimates are poor i.e. the standard error of the IV estimator is large.

We believe that the following arguments support our previous claim for the following reasons. For our analysis to be consistent the fundamental assumptions of the instrumental variables approach must be fulfilled i.e. relevance, exclusion restriction and monotonicity. As the death of a grandparent is an considered an exogenous event, the assumption of monotonicity is satisfied by the nature of the instrument. For the assumption of exclusion restriction to be satisfied, having the maternal grandmother alive should not influence maternal labor force participation through other channels than childcare provision. As previously discussed in the literature review, although there can be several exceptions to how the grandmother's health and longevity status could indirectly influence maternal labor force participation, for the sake of our instrumental variables approach we will assume that there is no direct influence and will later conduct a robustness check where we eliminate parents or parents-in-law that are in need of assistance so as to eliminate possible channels between labor force participation and having a parent alive.

Even when it is considered that monotonicity and exclusion restriction are fulfilled, as it can be seen from the first stage regression (Column 1) in Table 3 our instrument is unable to fulfill the assumption of relevance which is crucial for our FEIV approach to be consistent. Contrary to the literature that supports that Dutch grandmothers are very involved with taking care of their children, we fail to find a significant

relationship between having a maternal grandmother alive and regular grandparental childcare in the Dutch population. This is further indicated by the tests on the validity of our instruments.

As presented in Table 4, we have a Kleibergen – Paap rk Wald F-statistic of 0.235, which is considerably below the threshold required for a 10% maximal IV size i.e. 16.38 and also below the rule of thumb of proposed by Staiger and Stock (1997) confirming that our instrument is not valid and does not have a sufficient relevance to explain grandparental childcare.

Additional to relevance, our instrument also fails to pass the test of under identification with a p-value 0.628 for the rK statistic and hence we are unable to reject the null hypothesis that the instruments are not correlated with our endogenous variable and cannot conclude that our instruments are completely relevant.

According to the literature on weak-instrument-robust inference, the test-statistic proposed by Anderson and Rubin (1949) and Stock and Wright LM statistics are employed to specifically deal with weak instruments. According to the National Bureau of Economic Research (2018), such inference techniques avoid effective coverage regardless of instrument strength and avoids throwing away application with valid instruments just because they are weak. In both cases, the null hypothesis tested is that the coefficient of the endogenous regressors in the structural equation are jointly equal to zero and that overidentifying restrictions are valid and both tests are robust to the presence of weak instruments (Baum et al., 2010). As the literature has a general consensus on Anderson-Rubin statistic being more efficient in just-identified models with one endogenous variable with a single instrument (Andrews et al., 2019; Baum, 2007) we use the Anderson-Rubin test statistic to test whether the coefficient of grandparental childcare on maternal labor force participation is significantly greater than 0. Unfortunately, with a p-value of 0.223 we are unable to reject the null that the coefficient of grandparental childcare is equal to 0 and hence unable to make a conclusion using the FEIV approach. Regardless Table 3 and 4 presents the results of our estimation and test on endogeneity and validity on the instruments.

Due to a weak first-stage, a FEIV approach cannot fix the problem of endogeneity and is inefficient as well as our fixed effects results presented on Table 2. According to Milner et al. (2017), combining these two analytic models are appropriate if researchers are interested in examining and improving causality, however that reducing potential bias and improving causal inferences comes at a cost of lower power and wider confidence intervals than traditional random effect models.

On the basis that the combination of a FE and IV model might have also led to invalid instruments and lower power for our coefficients, to see whether this is actually the case and to observe the differences in between regressions, given IV is more efficient in dealing with endogeneity, we will compare and contrast our FEIV results with the results of only an IV approach in the next section.

Table 3 Instrumental variables fixed effects regression estimates of the effect of grandparental childcare on mother's labor force participation

Dependent variable:	(1 st Stage) Grandparental Childcare	(2 nd Stage) Maternal Labor Force Participation
Grandmother alive	-0.0184 (0.0379)	
Ages: 25-35	0.0710 (0.174)	-0.264 (0.723)
Ages: 35-44	0.0195 (0.177)	-0.119 (0.547)
Ages: 45-54	-0.00690 (0.180)	-0.0734 (0.515)
Secondary education	0.0473 (0.0392)	-0.248 (0.380)
Tertiary education	-0.105 (0.103)	0.586 (0.716)
At least one child between the ages of 0-4	0.114*** (0.0271)	-0.303 (0.707)
Number of children under 10	0.0355* (0.0183)	-0.134 (0.227)
Single mother	0.00424 (0.0596)	0.0542 (0.180)
Very urban	0.0450 (0.0339)	-0.104 (0.277)
Not urban	-0.0691* (0.0364)	0.243 (0.435)
Grandparental Childcare		2.802 (6.139)
Observations	2,161	2,161
R-squared		-12.609
Number of groups	581	581

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4 Test of Endogeneity and the Validity of the Instruments

	Relevant statistic	P-value
Wu-Hausman test statistic	1.345	0.246
Kleibergen-Paap rk Wald F test	0.235	
Kleibergen-Paap rk LM statistic	0.234	0.628
Anderson-Rubin Wald test(1,580)	1.49	0.223

7.3 Instrumental Variables (Outcome variable: MLFP)

For this section of our analysis, on the basis that it is the most recent wave that also has the highest number of observations i.e. 508, we will be using the 10th wave of the LISS Core Study: Family and Household. We will be using the same explanatory variables as in the previous section. The descriptive statistics for this sub-sample can be found in Appendix B. The equations that are estimated are as follows:

$$gcc = \beta_0 + \beta_1 gmomalive + \sum_k \beta_k X_k + e_1$$

$$mlfp = \alpha_0 + \delta gcc + \sum_k \beta_k X_k + e_2$$

7.3.1 Ordinary Least Squares (MLFP)

Table 5 OLS estimates of the effect of grandparental childcare on mother's labor force participation

Dependent variable	Maternal Labor Force Participation
Grandparental Childcare	0.1446*** (0.0516)
Ages: 25-35	0.1686 (0.1517)
Ages: 35-44	0.2323 (0.1528)
Ages: 45-54	0.0710 (0.1740)
Secondary education	-0.0523 (0.1746)
Tertiary education	0.1898 (0.1673)
At least one child between the ages of 0-4	-0.0037 (0.0489)
Number of children under 10	-0.0473 (0.0308)
Single mother	-0.0271 (0.0564)
Very urban	0.0672 (0.0489)
Not urban	0.0151 (0.0553)
Constant	0.4199** (0.2067)
Observations	508
R2	0.080

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We start by presenting OLS results to be able to compare them with our IV approach. Our OLS results show a highly positive and significant relationship between grandparental childcare and the probability of the mother to have paid work. Mothers who make use of grandparental childcare are on average 15 percentage points more likely to work. We find no significant differences in labor force participation between different age groups and different levels of education. Our results highlight no significant difference between our other measures included in this regression.

7.3.2 Instrumental Variables (MLFP)

Table 6 IV estimates of the effect of grandparental childcare on maternal labor force participation(wave 10)

Dependent variable:	(1) Grandparental Childcare	(2) Maternal Labor Force Participation
Grandmother alive	0.117*** (0.0273)	
Ages: 25-35	-0.0851 (0.139)	0.101 (0.265)
Ages: 35-44	-0.0900 (0.138)	0.152 (0.266)
Ages: 45-54	-0.148 (0.138)	-0.0902 (0.296)
Secondary education	0.128** (0.0537)	0.0695 (0.207)
Tertiary education	0.173*** (0.0481)	0.369* (0.211)
At least one child between the ages of 0-4	0.195*** (0.0430)	0.211 (0.140)
Number of children under 10	0.0147 (0.0222)	-0.0356 (0.0406)
Single mother	-0.00110 (0.0401)	-0.0234 (0.0692)
Very urban	0.0353 (0.0439)	0.103 (0.0712)
Not urban	0.0125 (0.0408)	0.0217 (0.0732)
Grandparental Childcare		-0.956 (0.641)
Constant	-0.133 (0.129)	0.391 (0.257)
Observations	508	508
R-squared		-0.602

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 7 Test of Endogeneity and the Validity of the Instruments

	Relevant statistic	P-value
Wu-Hausman F-test (F(1,496))	3.925	0.048
Kleibergen-Paap rk Wald F test	18.328	0.000
Kleibergen-Paap rk LM statistic	13.617	0.000

Contrary to our IVFE analysis, using the Wu-Hausman test with a p-value of 0.048, we are able to reject the null hypothesis that grandparental childcare is exogenous confirming our initial predictions of endogeneity. This indicates that the OLS is indeed biased and that the IV method will produce more efficient results that are closer to the true mean of the population. As indicated in the first column, we are able to confirm relevance. Our first-stage results show that having a maternal grandmother alive is strongly correlated with grandparental childcare and is statistically significant at the 1% level (t-statistic: 4.28). Children that have a maternal grandmother who is alive are on average 12 percentage points more likely to receive grandparental childcare. Our first-stage results also provide information on how grandparental childcare differs across mothers in certain aspects.

Although we don't see any significant differences between different age groups, we see that those who have completed university or secondary school are significantly more likely to make use of grandparental childcare relative to those the refence group, respectively on the 5% and the 1% significance level. The size of the coefficients is respectively 0.13 and 0.17, meaning that on average there is a 0.13- and 0.17-degree difference in the probability of using grandparental childcare between mothers who have completed secondary and tertiary education and those who have not completed any education. One possible explanation for such results is that mothers who have only completed primary education or have not yet started or completed any education are less likely to be employed relative to those who have completed their education with higher degree. Those who are employed might have a harder time juggling work and childcare activities which would explain the higher use of grandparental childcare. Column 2, presenting the second-stage regression, supports this idea as those who have completed tertiary education are significantly more likely to work relative to both those who have finished high school and our reference group.

Our first-stage regression fails to find a positive relationship between the number of children under 10 and grandparental childcare, this is explained by the fact that parents start enrolling their kids to facilities such as pre-school and elementary school in which they no longer need assistance from grandparents during regular work hours. As expected we find a highly positive and statistically significant relationship (at the

1% significance level) between grandparental childcare and mothers who have at least one kid between the ages of 0-3. On average this group of mothers have a 0.2-degree difference with mother's that do not have such younger kids with regards to grandparental childcare. We don't find any differences in grandparental childcare with regards to the urban characteristics of place of residence or the individual's marital status.

Our first stage F-statistic is equivalent to 18.328 which is both above the rule of thumb proposed by Staiger and Stock (1997) and the weak ID test critical value proposed by Stock and Yogo (2005) i.e. 16.38 for a 10% maximal IV size. Since we only have one endogenous regressor and a single instrument we are not able to conduct a test of overidentifying restrictions, however with a p-value of the Kleibergen-Paap rK 0.0002 we reject the null that the model is under-identified and conclude that in contrary to our previous analysis our instrument is completely relevant.

Unlike our OLS results we fail to find any significant causal relationship between grandparental childcare and maternal labor force participation. Indicating that OLS had an upward bias while estimating the size and the significance of the coefficient. Additionally, with an average of 0.37-degree difference to the reference group, we find a significant difference in labor force participation with respect to the level of education mothers have completed where mothers who have completed their university education have a higher probability of working than mothers who have not. We are unable to find any other differences between mothers with respect to their labor force participation.

Given that by removing the fixed approach, endogeneity of grandparental childcare is confirmed and our instruments are valid, we can now conclude that in the year 2018, the available supply of grandparental childcare had no causal effect on maternal labor force participation. However, as this conclusion cannot be extended to our FEIV approach, comparing our IVFE and IV approaches is necessary to see what has caused the IVFE instruments to be invalid. Table 8 shows the first-stage estimates of our two approaches.

The reason that our methodology is inefficient and not appropriate in our IVFE approach was that the assumption of relevance wasn't fulfilled. We can see that although having a grandmother alive is statistically and significantly correlated with grandparental childcare at the 1% level on our IV approach, this correlation becomes insignificant after the addition of multiple waves. This indicates that the reason for our instruments to lose validity is indeed the addition of the fixed effects approach. This claim is supported by the other big difference between the first stage regressions in the two models. While mothers who have completed university education and high school education are respectively more likely to make use of grandparental childcare when an IV approach is adopted, in our IVFE approach completing any level of education seems to lose its significance.

We conclude that the fixed effects approach weakens our first stage and hence renders our IVFE approach to be invalid. This finding can have two interpretations. Combining an IV approach with fixed

effects, reduces bias caused by the unobserved preferences and improves causality. In this case, by using FE, obtaining an insignificant coefficient for having a grandmother alive could mean that there is truly no relationship between having a grandmother alive and regular grandparental childcare.

However, introducing the addition of FE also introduces a trade-off: it improves causal inferences and reduces potential bias of the estimated effect size with the cost of lower statistical power and wider confidence intervals. As it can be seen below on Table 8 and 9, although differences in size are not large for our first stage estimations, the addition of a FE effects approach leads to larger standard errors and wider confidence intervals for our second-stage regressions. Such results may also indicate that FE decreases statistical power to such an extent that we observe insignificant coefficients. Although we are not able to reach a conclusion regarding this trade-off, given that grandparental childcare is insignificant in explaining maternal labor force participation in every econometric methodology we employed except for OLS, we make the conclusion that grandparental childcare does not have a causal effect on maternal labor force participation.

Table 8 IVFE and IV estimations of our first-stage regressions for grandparental childcare's effect on maternal labor force participation

Dependent variable: Grandparental Childcare	Coefficient	IVFE		Coefficient	IV	
		Robust Std. Err.	Confidence Intervals		Robust Std. Err.	Confidence Intervals
Grandmother alive	-0.018	0.038	-0.093 - 0.056	0.117***	0.027	0.063 - 0.170
Ages: 25-35	0.071	0.174	-0.270 - 0.412	-0.085	0.139	-0.358 - 0.187
Ages: 35-44	0.019	0.177	-0.328 - 0.367	-0.090	0.138	-0.361 - 0.181
Ages: 45-54	-0.007	0.180	-0.360 - 0.346	-0.148	0.138	-0.419 - 0.124
Secondary Education	0.047	0.039	-0.030 - 0.124	0.128**	0.054	0.023 - 0.234
Tertiary Education	-0.105	0.103	-0.306 - 0.097	0.173***	0.048	0.079 - 0.268
At least one child between the ages of 0-4	0.114***	0.027	0.061 - 0.168	0.195***	0.043	0.110 - 0.279
Number of children under 10	0.035*	0.018	-0.001 - 0.071	0.015	0.022	-0.029 - 0.058
Single mother	0.004	0.060	-0.113 - 0.121	-0.001	0.040	-0.080 - 0.078
Extremely urban	0.045	0.034	-0.021 - 0.111	0.035	0.044	-0.051 - 0.122
Not urban	-0.069*	0.036	-0.141 - 0.002	0.012	0.041	-0.068 - 0.093
Observations	2,161			508		
Number of groups	581					

Table 9 shows the IVFE and IV estimations of our second-stage regressions for grandparental childcare's effect on maternal labor force participation

Dependent variable: Maternal Labor Force Participation	IVFE			IV		
	Coefficient	Robust Std. Err.	Confidence Intervals	Coefficient	Robust Std. Err.	Confidence Intervals
Grandparental Childcare	2.802	6.156	-9.264 - 14.867	-0.956	0.641	-2.212 - 0.301
Ages: 25-35	-0.264	0.725	-1.685 - 1.156	0.101	0.265	-0.418 - 0.619
Ages: 35-44	-0.119	0.549	-1.195 - 0.957	0.152	0.266	-0.370 - 0.673
Ages: 45-54	-0.073	0.517	-1.086 - 0.940	-0.090	0.296	-0.670 - 0.489
Secondary Education	-0.248	0.381	-0.994 - 0.498	0.069	0.207	-0.337 - 0.476
Tertiary Education	0.586	0.718	-0.822 - 1.993	0.369*	0.211	-0.044 - 0.781
At least one child between the ages of 0-4	-0.303	0.709	-1.694 - 1.087	0.211	0.140	-0.063 - 0.484
Number of children under 10	-0.134	0.228	-0.580 - 0.312	-0.036	0.041	-0.115 - 0.044
Single mother	0.054	0.181	-0.300 - 0.408	-0.023	0.069	-0.159 - 0.112
Extremely urban	-0.104	0.278	-0.649 - 0.441	0.103	0.071	-0.037 - 0.242
Not urban	0.243	0.437	-0.613 - 1.098	0.022	0.073	-0.122 - 0.165
Constant	0.268	0.838	-1.375 - 1.911	0.391	0.257	-0.113 - 0.895
Observations	2,484			508		
Number of groups	904					

A possible explanation for this result could be tied to the abundance of the part-time opportunities in the Netherlands. As discussed in the literature, if females are indeed switching to part-time working instead of exiting or not entering the labor force for childcare purposes, potentially, although grandparental childcare does not have a causal effect on maternal labor force participation, it could have a causal effect on the number of weekly hours of working mothers. To investigate whether this is the case, in the next section, we will conduct an analysis where we examine the causal effect of grandparental childcare on the number of weekly hours worked on women with positive working hours.

7.4 Instrumental Variables (Dependent variable: Logarithm of Hours Worked Weekly)

Due to reasons of endogeneity and the inefficiency caused by the addition of the fixed effects approach, similar to the previous section we will be using the IV approach on the 10th wave of the LISS panel. For this analysis, in order to gather information on the details of how many hours the women in the previous sample work, we will also be using the Work and Schooling questionnaire of 2017 alongside the

Family and Household questionnaire. As the number of respondents that have filled out both questionnaires is smaller, unfortunately the number of observations, with the inclusion of 0 for non-working mothers, drops down to 364. When non-working mothers are also excluded out of the sample our number of observations drops down to 264. Table 10 presents our summary statistics.

Table 10 Summary statistic for the variables used in this sub-section of the study

Variables	No. of Obs.	Mean	St. Dev	Value =1	Min	Max
Maternal labor force participation	264	1	0	264	1	1
Contractual hours worked weekly	249	26.80	8.155		0	44
Actual hours worked weekly	264	28.25	9.023		3	60
Grandparental childcare	264	0.163	0.370	43	0	1
Grandparental help	264	0.496	0.501	131	0	1
Grandmother alive	264	0.894	0.309	236	0	1
Single mother	264	0.125	0.331	33	0	1
Logarithm of actual hours worked	264	3.270	0.430		1.099	4.094
16-24	264	0.0114	0.106	3	0	1
25-35	264	0.330	0.471	87	0	1
35-44	264	0.602	0.490	159	0	1
45-max	264	0.0568	0.232	15	0	1
No or primary education	264	0.0114	0.106	3	0	1
Secondary education	264	0.0947	0.293	25	0	1
Tertiary education	264	0.894	0.309	236	0	1
Very urban	264	0.152	0.359	40	0	1
Not urban	264	.188461	.39183	43	0	1
Number of children under 10	264	.1575758	.66551		1	3
At least one child between the ages 0-4	264	.371212	.48404	98		

The mean of the contractual hours worked and actual hours worked respectively is 26.8 and 28.3. In the Netherlands, if an individual works more than 12 hours but less than 35 they are considered as a part-time worker (Brown, 2021). For our data set, the mean values for hours worked weekly are well below 35 indicating that on average Dutch mothers prefer to work part-time over full-time. In fact, in our current sample of mothers, 183 out of 264 i.e. approximately 70%, work part-time. Such findings are in line with the literature and do in fact explain why we weren't able to find a direct effect of grandparental childcare on maternal labor force participation. The idea is simply that mothers spend more time at home for care activities and only receive help from grandparents instead of providing regular childcare. Further supporting

our predictions, while 50% of the mothers in our sample receive help from their grandparents only 16% receive regular childcare. Nevertheless, we believe that grandparental childcare might be a possible solution for mothers to increase their working hours or keep their regular working hours the same without reducing it for childcare purposes.

To see whether that is the case, we will again consider the same binary treatment *gcc* childcare (=1 for mothers who use grandparental childcare on a frequent basis, =0 otherwise). For our dependent variable, both in order to create the necessary sub-sample and also to deal with the skewed distribution caused by the number of zeros for mothers that are not employed, rather than using actual hours worked weekly, we will take the natural logarithm of the outcome variable, denoted by $\ln(hours)$. In this case, because the logarithm of 0 is undefined, due to the nature of the outcome variable and the properties of the statistical software the software used, Stata denotes mothers who are not employed as missing observations and removes it from the regression analysis enabling us to reach the characteristics of the desired sample. We use the actual hours worked as it is a more realistic measure and because mothers with young children may be on paid leave at the time they are taking the survey.

The following equations will be estimated to carry out the 2SLS method:

$$gcc = \beta_0 + \beta_1 gmomalive + \sum_k \beta_k X_k + e_1$$

$$\ln(hours) = \alpha_0 + \delta gcc + \sum_k \beta_k X_k + e_2$$

7.4.1 Ordinary Least Squares (Log (Hours Worked Weekly))

Below we present our OLS results for comparison purposes. Our OLS results fail to find a positively significant relationship between grandparental childcare and the number of weekly hours worked weekly. Our results highlight that working mothers who have at least one kid between the ages of 0-4 (not including 4) are significantly more likely to work for greater number of hours during the week at the 10% significance level. Since exponential of .0987 \approx 1.103735, number of hours worked weekly are 10% higher for such mothers compared to mothers who do not have kids between the ages of 0-4. This is counterintuitive given that one would expect to mothers to work less as younger kids require more attention and care taking. We find no significant differences between the age categories, the level of education, marital/relationship status and region of residence for working mothers.

Table 11 OLS estimates of the effect of grandparental childcare on the number of hours worked weekly

Dependent variable:	Log (Number of Hours Worked Weekly)
Grandparental Childcare	0.0082 (0.0511)
Ages: 25-35	-0.0828 (0.1230)
Ages: 35-44	-0.0971 (0.1240)
Ages: 45-54	-0.0529 (0.1892)
Secondary education	0.0092 (0.3098)
Tertiary education	0.2309 (0.2978)
At least one child between the ages of 0-4	0.0987* (0.0567)
Number of children under 10	-0.0335 (0.0369)
Single mother	0.0717 (0.0851)
Very urban	0.0709 (0.0651)
Not urban	-0.1023 (0.0807)
Constant	3.1651*** (0.2073)
Observations	264
R2	0.068

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.4.2 Instrumental Variables: Log (Hours Worked Weekly)

Below we present our 2SLS estimates. Our first stage regression shows that having a grandmother alive is positively and significantly associated with receiving grandparental childcare at the 1% significance level where those children who have alive grandmothers have a higher probability of receiving grandparental childcare by a .12-degree difference confirming a strong first-stage (relevance). Additionally, with a t-statistic of 4.6, relative to those who do not, working mothers who have children between the ages of 0-4, are more likely to receive help in childcare by their grandparents by a 0.3-degree difference.

Table 12 Two-stage least squares (2SLS) estimates of the effect of grandparental childcare on number of weekly hours worked for the sample of employed mothers

Dependent variable:	1 st Stage Grandparental Childcare	2 nd Stage Log(Number of Hours Worked Weekly)
Grandmother alive	0.123*** (0.0386)	
Ages: 25-35	-0.193 (0.238)	0.187 (0.401)
Ages: 35-44	-0.190 (0.239)	0.190 (0.410)
Ages: 45-54	-0.223 (0.241)	0.335 (0.448)
Secondary education	0.214 (0.146)	-0.322 (0.430)
Tertiary education	0.207 (0.138)	-0.107 (0.415)
At least one child between the ages of 0-4	0.280*** (0.0604)	-0.394 (0.314)
Number of children under 10	0.0107 (0.0341)	-0.0515 (0.0703)
Single mother	0.0145 (0.0629)	0.0456 (0.146)
Very urban	0.0636 (0.0655)	-0.0291 (0.141)
Not urban	0.0755 (0.0607)	-0.215 (0.158)
Grandparental childcare		1.783* (1.046)
Constant	-0.107 (0.233)	3.177*** (0.381)
Observations	264	264
R-squared		-1.894

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 13 presents the tests on endogeneity and validity of the instruments. With a p-value of 0.038, we reject the null of the Wu-Hausman test and confirm endogeneity. Given a Kleinbergen-Paap rk F statistic of 10.007, according to Staiger and Stock's (1997) rule of thumb, the relevance assumption is confirmed. However, Stock and Yogo (2005) proposes that the rule of thumb isn't sufficient and proposes cutoff values presented in Appendix A. According to these cut-off values we aren't able to reject the null hypothesis that our instrument is weak.

Table 13 Test of Endogeneity and the Validity of the Instruments

	Relevant Statistic	P-value
Wu-Hausman F-test (F(1,251))	4.374	0.038
Kleibergen-Paap rk Wald F test	10.077	
Kleibergen-Paap rk LM statistic	7.733	0.005
Anderson-Rubin Wald test(1,252)	4.36	0.038

In order to make sure that we conduct weak-instrument-robust inference we will be using the test-statistic proposed by Anderson and Rubin (1949) on the basis that it is the most efficient in just-identified models. With a p-value of 0.034 we are able to reject the null that the coefficient of grandparental childcare is equal to 0 indicating a positive effect of grandparental childcare on the working hours of working women.

Although we can confirm that the coefficient is significantly greater than 0, because our instruments are weak, bias is present in the size of our coefficient of grandparental childcare and the confidence intervals presented in our second stage regression do not necessarily reflect the actual coverage.

To mitigate this issue, following the literature we will construct a new confidence interval using the method of test inversion. Theoretically explained by Andrews et al. (2019), test inversion using weak instruments implies testable restrictions on the distribution of the data regardless of the strength of the instruments in just-identified models (Andrews et al., 2019; Baum, 2007) Using the user-written weakiv (Finlay, 2013) command implemented by Stata, we present our coverage-corrected confidence set and size-corrected p-value based on the Anderson-Rubin test statistic. The coverage corrected confidence sets are also robust to heteroskedasticity.

Table 14 shows coverage-corrected confidence sets and p-value for linear IV

Test	Relevant Statistic	P-value	Confidence level	Confidence Set
Anderson-Rubin	4.47	0.0345	95%	[.167073,]
Wald	2.90	0.0884	95%	[-.267874, 3.833305]

The Wald statistic and the confidence interval presented in the last row is based on our original estimation and is not robust to weak instruments. Table 14 shows that the coverage-corrected interval only

includes positive values enabling us to conclude that on mothers who receive regular grandparental childcare on average tend to work a greater number of hours indicating that the OLS estimates have a downward bias in estimating the size of the coefficient of our endogenous variable.

8. Robustness Checks

8.1 Geographical proximity between mother and grandparent

Garcia-Moran and Kuehn (2012) and Compton and Pollak (2013) find that close geographical proximity to mothers or mothers-in-law has a substantial positive effect on the labor force participation of mothers with young children. In this respect, to see if our conclusions extend to different instruments, as a robustness check we will redefine our instrument. The instrument will now equal to 1 if the grandparent of interest resides geographically close i.e. if the distance is at most 16km. In choosing the distance regarding what classifies as residing close to mother, we followed Arpino et al. (2014).

We attempted making use of this instrument in both of our two analysis i.e. maternal labor force participation and the number of weekly hours worked as the dependent variable however, unfortunately, as shown on Appendix C, in both cases we were not able to reject the Wu-Hausman test of endogeneity, weren't able to prove relevance and weren't able to conduct weak-instrument robust inference due to the sizes our Anderson-Rubin statistics and hence were not able to test our findings with a different instrument for our endogenous variable.

8.2 Grandparents in need of Health Assistance

In an attempt to eliminate possible channels between our instrument and our outcome variables, we exclude individuals that are providing informal care to their parents or parents-in-law as it might affect the individuals working decisions. Doing so we are able to decrease potential bias of our IV method by strengthening the assumption of exclusion restriction.

To construct our new sample, we used the answers to the following question on the Work and Schooling questionnaire: 'To whom are you providing informal care?'. If the individual responded that she provided informal care to either her parents or her parents in law she was removed from the sample.

The estimations results are presented in Appendix D. It is seen that out of 508, 16 mothers were providing some type of informal care to their elders i.e. 3.2% of the total population. With the exclusion of these individuals, our F statistic still passes the required threshold value and our p-values are sufficiently small to reject tests of exogeneity and under identification (Appendix E). In accordance with our previous results, we do not observe any causal relationship between grandparental childcare and maternal labor force participation and additionally no significant differences between the regression results of our main analysis.

If we repeat the analysis with the logarithm of number of hours worked weekly, as shown in Appendix F we are unable to reject the null that grandparental childcare is endogenous (p-value: 0.056). Even if it were assumed that the variable is endogenous, we can't reject the null hypothesis that the coefficient of grandparental childcare is significantly different than 0 (Appendix G). Nevertheless, because we weren't able to support our previous conclusion in our original analysis, our conclusion regarding grandparental childcare's effect on number of hours worked must be interpreted cautiously our econometric approach possibly suffers from bias.

8.3 Bivariate Probit Model

The bivariate probit model is a maximum likelihood model used to test binary endogenous covariates when the dependent variable is also binary. According to Nichols (2011), while IV is a consistent estimator of an average effect of treatment, it is biased and small performance may be inferior to a correctly specified maximum likelihood model where bivariate probit outperforms IV in sample sizes smaller than 5000. Moreover, Nicholas (2011) puts forward the power of any instrumental variables' strategy might be very low. He states that too often researchers make a significant coefficient insignificant by instrumenting and then conclude the true effect is zero. Additionally, bivariate probit can be used to estimate unconditional causal effects and/or effects on treated, whereas 2SLS only promises local average causal effects (Nichols, 2011). Appendix H presents the bivariate probit estimates of grandparental childcare on maternal labor force participation

Although we have chosen to present our bivariate probit results for the sake of completeness, we will not present the marginal effects of grandparental childcare on maternal labor force participation as the effect is not statistically significant. However, our results are consistent with our IV estimates and further support the finding that there isn't a causal relationship between maternal labor force participation and grandparental childcare.

8.4 Instrumental variables: 13th wave, year 2020

As a final robustness check we will be employing our IV analysis on the 13th wave (2020) of the LISS panel as it is the most recent available year. Appendix I presents the IV estimates of grandparental childcare on maternal labor force participation.

Appendix J shows that our model passes all tests for employing an instrumental variables approach. In accordance with our results for the year 2017 we see that grandparental childcare is not significantly associated with maternal labor force participation. In contrast with the year 2017, in 2020 we see observe that being a young mother, those who are in the age group 16-25 are less likely to participate in the labor force and also less likely to make use of grandparental childcare. There are no differences in the use of

grandparental childcare on mothers who have completed different levels of education and, the number of children the mother has under the age of 10 is negatively associated with maternal labor force participation.

We also repeated the analysis with our dependent variable as the logarithm of number of hours worked weekly, but was not able to make any conclusions due to the presence of weak instruments and inability to conduct a weak-instrument-robust-inference.

9. Discussion and Conclusion

Unfortunately, due to the weak instruments and the inefficiency of our FEIV methodology we weren't able to answer our research question using our main methodology of choice. In order to observe whether the FEIV approach became inefficient due to the addition of FE, we employed an IV approach on one wave in the LISS panel. We concluded that the addition of the fixed effects approach is the reason the weak first-stage. Although the appropriateness of an IV approach is confirmed using a single wave, our IV approach, further robustness checks and a FE model shows that regular grandparental childcare does not directly influence maternal labor force participation in the Netherlands. We believe the reasoning behind our findings, is the high labor force participation rates driven by the availability of part-time opportunities and mothers having the chance to rely on regular grandparental childcare less frequently.

In this respect, as an extension to the literature, we investigated the effect of regular grandparental childcare on the working hours of employed mothers using an IV approach. Due to the limitations of our methodology we had to conduct weak-instrument-robust inference and made use of the Anderson-Rubin test statistic to ensure correct coverage of the confidence intervals regardless of the strength of the instrument. Our findings pointed towards the fact that the use of grandparental childcare had a positive effect on the number of working hours of women, where those who make use of grandparental childcare on average worked a greater number of hours.

It must be noted that without the exclusion of mothers who don't work positive hours, even though the analysis we conducted on the same wave, because the number of individuals who responded to both questionnaires (Family and Household, and Work and Schooling) are different, our sample size dropped from 508 to 355. We note that such a large drop that is not in our control might have led to sample selection bias causing non-randomness in our sample. In this sense although we believe this finding is a valuable addition to the literature, because we could not support our findings using different methodological approaches via robustness checks we believe that the effect of grandparental childcare on the number of working hours for working mothers need to be further verified and further investigated in the literature possibly using different methods, different samples with higher numbers of observations and a sample that does not suffer from sample selection bias.

Although our findings are contradictory to qualitative studies in the literature it must be noted that most available literature focusing on this topic specifically for the Netherlands is composed of studies that have been conducted around approximately a decade ago, hence it is also possible that values and opinions along with the institutional structure of the childcare system may have changed, not to mention the significant increase in the elder labor force participation rates and the shift in the statutory retirement age in the Netherlands from 65 to 2013 to 65 and 9 months in 2017 (Atav et al., 2019). All of these factors separately are plausible and possible explanations for both the structure of our data, whereby the number of alive grandmothers who provide childcare is small, as well as our statistically insignificant results for our IV approach. Additionally, although most empirical studies conducted on this topic do find positive associations, such studies do not investigate a possible causal relationship.

It is worth mentioning that after observing the difference in the number of individuals who receive only help from grandparents versus those who receive regular childcare we repeated our IV analysis with our dependent variable as maternal labor force participation with the newly constructed endogenous variable *help_grandparents* (a binary variable if the grandmother or the grandfather had helped the mother several times in childcare activities in the past three months) and both of our approaches showed that there is no causal effect of receiving help from grandparents several times in the past few months. This finding is reasonable as receiving a small amount of help several times is not expected to influence a mother's decision to work or not. Further studies could explore whether less frequent help influences the number of working hours of working mothers.

Although we conclude that grandparental childcare does not have a casual effect on the mother's decision to work in the country of Netherlands, we believe that further studies should explore, compare and contrast their findings with the ones of our own by correcting for the limitations we had in this study. Recommendations would be to repeat the study with a data set that provides a higher number of observations for each wave as more controls could then be incorporated, and hence the effect of grandparental childcare on maternal labor force participation would be further isolated. This would enable the researcher to overcome some of our shortcomings by also controlling for specific characteristics for example, net household income, region of residence and other institutional characteristics of the childcare facilities specific to that region such as fees and availability, as well as ages of the grandparents and their working status. Using a large data set, ideally one could incorporate more instruments into the analysis, such as an instrument for each grandparent.

Finally, as policy recommendations, given that we find no causal effect of grandparental childcare on the decision to work or not we believe that an effective strategy to incentivize the female workforce to work would be to increase the availability of formal childcare in the Netherlands both in terms of costs and coverage as well as providing a substantial increase in the quality of the service provided. Given the positive

association between grandparental childcare and maternal labor force participation and the possible causal relationship between grandparental childcare and the number of hours worked in the Netherlands, we believe increasing trust in a setting where parents face information asymmetry might lower the minimum wage the mother would accept to forgo providing internal childcare. According to Frey (2019), among other OECD countries, the Netherlands is the least likely to have two full-time working partners, and as care responsibilities mainly drive part-time working decisions for females, we believe such policies will challenge the gendered nature of part-time work and the one-and-a-half earner model in the Netherlands (Frey, 2019) Even if the portion of grandparents providing regular grandparental childcare is not a significant amount compared to the whole population, such a policy of increasing trust and quality would also incentivize working grandparents to stay in the labor force who would otherwise leave to take care of their grandchildren regardless of whether they are receiving transfers for the childcare that they provide.

10. References

- Aassve, A., Arpino, B., & Goisis, A. (2012). Grandparenting and mothers' labour force participation: A comparative analysis using the Generations and Gender Survey. *Demographic Research*, 27, 53–84. <https://doi.org/10.4054/demres.2012.27.3>
- Abendroth, A. K., van der Lippe, T., & Maas, I. (2012). Social support and the working hours of employed mothers in Europe. *Social Science Research*, 41(3), 581–597. <https://doi.org/10.1016/j.ssresearch.2011.12.008>
- Amaro, S. (2021, January 15). *Dutch government resigns after childcare benefits scandal*. CNBC. <https://www.cnn.com/2021/01/15/dutch-government-resigns-after-childcare-benefits-scandal.html>
- Anderson T, Rubin H. 1949. Estimators for the parameters of a single equation in a complete set of stochastic equations. *Annals of Mathematical Statistics* 21:570–582
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, 11(1), 727–753. <https://doi.org/10.1146/annurev-economics-080218-025643>
- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434), 444–455.
- Arpino, B., Pronzato, C. D., & Tavares, L. P. (2014a). The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach. *European Journal of Population*, 30(4), 369–390. <https://doi.org/10.1007/s10680-014-9319-8>
- Atav, T., Jongen, E. L., & Rabaté, S. (2019). The effects of the increase in the retirement age in the Netherlands. *The effects of the increase in the retirement age in the Netherlands*.
- Baum, C. F., Schaffer, M. E., & Stillman, S. (2007). Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. *The Stata Journal: Promoting Communications on Statistics and Stata*, 7(4), 465–506. <https://doi.org/10.1177/1536867x0800700402>
- Baum, C. (2007, September). Instrumental Variables: Overview and advances [PowerPoint slides]. Boston College and DIW Berlin. <http://fmwww.bc.edu/repec/usug2007/baumUKSUG2007.pdf>
- Baum, C.F., Schaffer, M.E., Stillman, S. 2010. ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. <http://ideas.repec.org/c/boc/bocode/s425401.html>
- Bernard van Leer Foundation (2021) *The Netherlands is ready for a reset of the childcare system*. Retrieved from: <https://bernardvanleer.org/blog/the-netherlands-is-ready-for-a-reset-of-the-childcare-system/>

- Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90(430), 443-450.
- Bordone, V., Arpino, B., & Aassve, A. (2017). Patterns of grandparental child care across Europe: the role of the policy context and working mothers' need. *Ageing & Society*, 37(4), 845-873.
- Brown, E. (2021, March 10). *Vacation, free time and working hours in the Netherlands: all you need to know*. DutchReview. <https://dutchreview.com/expat/work/vacation-free-time-working-hours-in-the-netherlands/>
- Carriero, R., Ghysels, J., & Van Klaveren, C. (2009). Do parents coordinate their work schedules? A comparison of Dutch, Flemish, and Italian dual-earner households. *European Sociological Review*, 25(5), 603-617.
- Cloin, M., & Hermans, B. (2006). Unpaid work and the combination of work and care. *W. Portegijs, B. Hermans & V. Lalta (eds.). Emancipation Monitor* , 100-140.
- Compton, J., & Pollak, R. A. (2014a). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics*, 79, 72–90. <https://doi.org/10.1016/j.jue.2013.03.007>
- Dalla Zuanna, G. (2001). The banquet of Aeolus. *Demographic Research*, 4, 133–162. <https://doi.org/10.4054/demres.2001.4.5>
- Davidson, R., & MacKinnon, J. G. (2014). Confidence sets based on inverting Anderson–Rubin tests. *The Econometrics Journal*, 17(2), S39–S58. <https://doi.org/10.1111/ectj.12015>
- Del Boca, D. (2002). The effect of child care and part time opportunities on participation and fertility decisions in Italy. *Journal of population economics*, 15(3), 549-573.
- Del Boca, D., Pasqua, S., & Pronzato, C. (2008). Motherhood and market work decisions in institutional context: a European perspective. *Oxford Economic Papers*, 61(suppl_1), i147-i171.
- Dimova, R., & Wolff, F. C. (2007). Grandchild Care Transfers by Ageing Immigrants in France: Intra-household Allocation and Labour Market Implications. *European Journal of Population / Revue Européenne de Démographie*, 24(3), 315–340. <https://doi.org/10.1007/s10680-007-9122-x>
- Dimova, R., & Wolff, F. C. (2011). Do downward private transfers enhance maternal labor supply? Evidence from around Europe. *Journal of Population Economics*, 24(3), 911–933. <https://doi.org/10.1007/s00148-010-0305-0>
- Dykstra, P. A. (2012). Families in alle staten?[Families across time and societies]. *Mens & Maatschappij*, 87(2), 174-96.
- Frey, V. (2019). *Part-time and partly equal: gender and work in the Netherlands*. OECD, Paris, France.

- Finlay, K., Magnusson, L.M., Schaffer, M.E. 2013. weakiv: Weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models. <http://ideas.repec.org/c/boc/bocode/s457684.html>
- Garcia-Moran, E., & Kuehn, Z. (2013). With Strings Attached: Grandparent-Provided Child Care and Female Labor Market Outcomes. *SSRN Electronic Journal*. Published. <https://doi.org/10.2139/ssrn.2365716>
- Geurts, T., van Tilburg, T., Poortman, A. R., & Dykstra, P. A. (2014). Child care by grandparents: changes between 1992 and 2006. *Ageing and Society*, 35(6), 1318–1334. <https://doi.org/10.1017/s0144686x14000270>
- Glaser, K., Montserrat, E. R., Waginger, U., Price, D., Stuchbury, R., & Tinker, A. (2010). Grandparenting in Europe. *London: Grandparents Plus*.
- Habhoubi, E. E. H. (2012). The effect of caregiving on employment and mental health examined for midlife European citizens. *Erasmus School of Economics*.
- Hank, K. (2007). Proximity and contacts between older parents and their children: A European comparison. *Journal of Marriage and Family*, 69(1), 157-173.
- Hank, K., & Buber, I. (2009). Grandparents Caring for their Grandchildren. *Journal of Family Issues*, 30(1), 53–73. <https://doi.org/10.1177/0192513x08322627>
- Harms, L. 2008. *Overwegend onderweg: De leefsituatie en de mobiliteit van Nederlanders [Predominantly on the Way: Living Arrangement and Mobility of Dutch People]*. Netherlands Institute for Social Research, The Hague. [Google Scholar](#)
- Herlofson, K., Hagestad, G., Slagsvold, B., & Sørensen, A. M. (2011). Intergenerational family responsibility and solidarity in Europe. *Multilinks Project. Norwegian Social Research (NOVA)*. (104 Multilinks Project).
- Herwig, K. (2021, September 8) *Primary School Starts at Four – And Here’s Why I Love That*. Amsterdam Mamas. <https://amsterdam-mamas.nl/articles/primary-school-starts-four-%E2%80%93-and-heres-why-i-love>
- Igel, C., & Szydlik, M. (2011). Grandchild care and welfare state arrangements in Europe. *Journal of European Social Policy*, 21, 210-22
- Jappens, M., & Van Bavel, J. (2012). Regional family norms and child care by grandparents in Europe. *Demographic Research*, 27, 85-120.
- Kalmijn, M. (2013). Long-term effects of divorce on parent–child relationships: Within-family comparisons of fathers and mothers. *European sociological review*, 29(5), 888-898.

- Kelleher, L. (2019, July 8). *Childcare costs here among the most expensive in Europe*. Independent. <https://www.independent.ie/business/personal-finance/latest-news/childcare-costs-here-among-the-most-expensive-in-europe-38290772.html>
- Keuzenkamp, S., & Steenvoorden, EH (2008). Labor participation and working hours compared internationally. *Part-time country in the Netherlands. Women and part-time work* (pp. 46-64). Social and Cultural Planning Office.
- Kimmel, J. (1998). Child Care Costs as a Barrier to Employment for Single and Married Mothers. *Review of Economics and Statistics*, 80(2), 287–299. <https://doi.org/10.1162/003465398557384>
- Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics*, 133(1), 97-126.
- Klein, D. (1975). Women in the labor force: The middle years. *Monthly Labor Review*, 98(11), 10-16.
- Koslowski, A. S. (2009). Grandparents and the care of their grandchildren. In *Fertility, living arrangements, care and mobility* (pp. 171-190). Springer, Dordrecht.
- Kremer, M. (2007) *How Welfare States Care: Culture, Gender and Parenting in Europe*, Amsterdam: Amsterdam University Press.
- Lang, K., & Zagorsky, J. L. (2001). Does growing up with a parent absent really hurt?. *Journal of human Resources*, 253-273.
- Leopold, T., Geissler, F., & Pink, S. (2012). How far do children move? Spatial distances after leaving the parental home. *Social Science Research*, 41(4), 991-1002.
- LISS Panel Data*. (2007–2019). [Dataset]. <https://www.lissdata.nl/>
- McLanahan, S., & Sandefur, G. (1994). *Growing up with a single parent. What hurts, what helps*. Harvard University Press, 79 Garden Street, Cambridge, MA 02138.
- Milner, A., Aitken, Z., Kavanagh, A., LaMontagne, A. D., Pega, F., & Petrie, D. (2017). Combining fixed effects and instrumental variable approaches for estimating the effect of psychosocial job quality on mental health: evidence from 13 waves of a nationally representative cohort study. *Journal of Public Health*, 40(2), 426–434. <https://doi.org/10.1093/pubmed/fox070>
- Ministry of General Affairs. (2021, June 25). *Costs of primary school*. Primary Education | Government.NL. <https://www.government.nl/topics/primary-education/costs-of-primary-school>
- Ministry of Education Culture and Science. (2014, December 16). *Secondary education: fees and other educational costs*. Secondary Education | Government.NL. <https://www.government.nl/topics/secondary-education/secondary-education-fees-and-other-educational-costs>

- Morel, N. (2007) 'From subsidiarity to "free choice": child- and elder-care policy reforms in France, Belgium, Germany and the Netherlands', *Social Policy & Administration*, 41 (6): 618–37.
- Mulligan, C. B., & Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3), 1061-1110.
- National Bureau of Economic Research. (2018, July 22). Robust Inference with Weak Instruments [PowerPoint slides]. https://www.nber.org/sites/default/files/2020-12/robustinference_openissues.pdf
- Neyman, J. S. (1923). On the application of probability theory to agricultural experiments. essay on principles. section 9.(1990), 5, 465-480). *Annals of Agricultural Sciences*, 10, 1-51.
- Nichols, A. (2011, July). Causal inference for binary regression. In *Stata Conference Chicago (version June 14, 2011)*.
- NL Times. (2021, August 25). *Over 300,000 people 65+ continue in the workforce*. <https://nltimes.nl/2021/08/25/300000-people-65-continue-workforce>
- Noally, J., & Visser S. (2009). The Impact of Market Forces on Child Care Provision: Insights from the 2005 Child Care Act in the Netherlands. *Journal of Social Policy*, 38(3), 477–498. <https://doi.org/10.1017/s0047279409003109>
- OECD (2020), "Is Childcare Affordable?" Policy Brief on Employment, Labour and Social Affairs, OECD, Paris, [oe.cd/childcare-brief-2020](https://www.oecd.org/childcare-brief-2020)
- Ogawa, N., & Ermisch, J. F. (1996). Family structure, home time demands, and the employment patterns of Japanese married women. *Journal of Labor Economics*, 14(4), 677-702
- Plantenga, J. (1996) 'For women only? The rise of part-time work in the Netherlands', *Social Politics*, 3 (1): 57–71. Illinois: University of Illinois
- Plantenga, J., Siegel, M. (2004). European child care strategies, Position Paper for the conference child care in a Changing World.
- Portegijs, W., Cloin, M., Ooms, I. and Eggink, E. (2006) Hoe het werkt met kinderen: Moeders over kinderopvang en werk [How it Works with Children: Mothers About Daycare and Work]. Netherlands Institute for Social Research, The Hague
- Posadas, J., & Vidal-Fernandez, M. (2013a). Grandparents' Childcare and Female Labor Force Participation. *IZA Journal of Labor Policy*, 2(1), 14. <https://doi.org/10.1186/2193-9004-2-14>
- Pott-Buter, H.A. (1993) Facts and Fairy Tales about Female Labour, Family and Fertility: A Seven Country Comparison, 1850–1900, Amsterdam: Amsterdam University Press.
- Robin. (2021, June 10). *Make childcare accessible, affordable and less complex: SER*. DutchNews.NL. <https://www.dutchnews.nl/news/2021/06/make-childcare-accessible-affordable-and-less-complex-ser/>

- Remery, C. R., Henkens, K. H., Schippers, J. S., & Ekamper, P. E. (2003). *Managing an Aging Workforce and a Tight Labor Market: Views Held by Dutch Employers on JSTOR*. JSTOR. https://www.jstor.org/stable/40230809?seq=1&cid=pdf-reference#references_tab_contents
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701. <https://doi.org/10.1037/h0037350>
- Saraceno, C. (2008). Introduction: Intergenerational relations in families - A micro-macro perspective. In: Saraceno, C. (ed.). *Families, ageing and social policy. Intergenerational solidarity in European welfare states*. Cheltenham: Edward Elgar: 1-19.
- Schor, E. L. (1999). *Caring for Your School Age Child: Ages 5-12 (Child Care S)*. Bantam. Rev Sub edition.
- Seveno, V. S. (2021, April 23). *Parents in the Netherlands to receive at least nine weeks of paid leave*. IamExpat. <https://www.iamexpat.nl/expat-info/dutch-expat-news/parents-netherlands-receive-least-nine-weeks-paid-leave>
- Silverstein, M. & Giarusso, R. (2010) ‘Aging and family life: A decade review’, *Journal of Marriage and Family*, vol. 72, no. 5, pp. 1039–1058. [Crossref], [Web of Science®],
- Smith, K. (2006). *Rural Families Choose Home-Based Child Care for Their Preschool-Aged Children*. Policy Brief Number 3. *Carsey Institute*.
- Splawa-Neyman, J., Dabrowska, D. M., & Speed, T. P. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5(4). <https://doi.org/10.1214/ss/1177012031>
- Staiger, D. and J.H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica* 65(3), 557-586.
- Statistics Netherlands. (2008, September 5). *Nearly half a million single parents in the Netherlands*. <https://www.cbs.nl/en-gb/news/2008/36/nearly-half-a-million-single-parents-in-the-netherlands>
- Stock, J.H. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression. In D.W.K. Andrews and J.H. Stock (Eds.), *Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg*, 80-108. New York: Cambridge University Press.
- Svensson-Dianellou, A., Smith, P. K., & Mestheneos, E. (2010). Family Help By Greek Grandparents. *Journal of Intergenerational Relationships*, 8(3), 249–263. <https://doi.org/10.1080/15350770.2010.498737>
- Thévenon, O. (2013), "Drivers of Female Labour Force Participation in the OECD", *OECD Social, Employment and Migration Working Papers*, No. 145, OECD Publishing, Paris, <https://doi.org/10.1787/5k46civrngms6-en>.

- Thiele, D. M., & Whelan, T. A. (2006). The Nature and Dimensions of the Grandparent Role. *Marriage & Family Review*, 40(1), 93–108. https://doi.org/10.1300/j002v40n01_06
- Thomese, F., & Liefbroer, A. C. (2013). Child Care and Child Births: The Role of Grandparents in the Netherlands. *Journal of Marriage and Family*, 75(2), 403–421. <https://doi.org/10.1111/jomf.12005>
- Treas, J., & Widmer, E. D. (2000). Married women's employment over the life course: Attitudes in cross-national perspective. *Social forces*, 78(4), 1409-1436.
- Verweij, M., & Reimann, M. (2016). The part-time story: Dutch couples at the life-course transition to parenthood. In *Couples' Transitions to Parenthood*. Edward Elgar Publishing.
- Wang, V. (2021, May 10). *The Childcare Benefits Scandal in the Netherlands*. The Borgen Project. <https://borgenproject.org/the-childcare-benefits-scandal/>
- Wheelock, J., & Jones, K. (2002). 'Grandparents are the next best thing': informal childcare for working parents in urban Britain. *Journal of social policy*, 31(3), 441-463.
- Woolridge, Jeffrey M. (2020, March 26). *Fixed effects or random effects for panel data?* [Stata forum post]. Statalist. <https://www.statalist.org/forums/forum/general-stata-discussion/general/1543001-should-i-use-fixed-effects-or-random-effects-for-my-paneldata>
- Wooldridge, Jeffrey M. 2009. *Introductory econometrics*. Mason, OH: South Western, Cengage Learning
- Yobero, C. (2016, June 6). *Methods for Detecting and Resolving Heteroskedasticity*. RStudio. https://rstudio-pubs-static.s3.amazonaws.com/187387_3ca34c107405427db0e0f01252b3fbdb.html
- Zamarro, G. (2011). Family Labor Participation and Child Care Decisions: The Role of Grannies. *SSRN Electronic Journal*. Published. <https://doi.org/10.2139/ssrn.1758938>

11. Appendices

Appendix A - Stock-Yogo (2005) weak ID test critical values

10% maximal IV size	16.38
15% maximal IV size	8.96
20% maximal IV size	6.66
25% maximal IV size	5.53

Appendix B – Summary statistics for the 10th wave (year 2017)

Variables	No. of Obs.	Mean	St. dev.	Value = 1	Min	Max
Grandparental childcare	508	0.146	0.353	74	0	1
Maternal Labor Force Participation	508	0.724	0.447	368	0	1
Grandmother Alive	508	0.909	0.287	462	0	1
Very urban	508	0.150	0.357	76	0	1
Not urban	508	0.165	0.372	84	0	1
Single Mother	508	0.156	0.363	79	0	1
Age 16-24	508	0.0177	0.132	9	0	1
Age 25-34	508	0.344	0.476	175	0	1
Age 35-44	508	0.571	0.495	290	0	1
Age 45-56	508	0.0669	0.250	34	0	1
Secondary education	508	0.136	0.343	69	0	1
Tertiary education	508	0.844	0.363	429	0	1
No education or primary education	508	0.0197	0.139	10	0	1
At least one child between the ages 0-4	508	.36023	.48054	183	0	1
Number of children under 10	508	1.5255	0.6774		1	5

Appendix C- Tests of endogeneity and validity of the instruments for IV regression where the outcome variable is maternal labor force participation

Dependent variable:	Maternal Labor Force Participation	Log (Number of Hours Worked Weekly)
P-value for Wu-Hausman F-test statistic	0.8943	0.5867
Kleibergen-Paap rk Wald F test statistic	0.03	0.16
P-value for Anderson-Rubin Wald test(1,252) statistic	0.9107	0.5841

Appendix D – IV estimates of grandparental childcare on maternal labor force participation for the sub-sample excluding individuals who provide assistance to parents or parents-in-law

Dependent variable:	Grandparental childcare	Maternal Labor Force Participation
Grandmother alive	0.125*** (0.0298)	
Ages: 25-35	-0.0891 (0.139)	0.104 (0.268)
Ages: 35-44	-0.0915 (0.138)	0.140 (0.269)
Ages: 45-54	-0.137 (0.139)	-0.0616 (0.297)
Secondary education	0.127** (0.0548)	0.0721 (0.208)
Tertiary education	0.178*** (0.0490)	0.382* (0.211)
At least one child between the ages of 0-3	0.199*** (0.0438)	0.226 (0.139)
Number of children under 10	0.0174 (0.0228)	-0.0311 (0.0420)
Single mother	0.00515 (0.0417)	0.00763 (0.0702)
Extremely urban	0.0368 (0.0453)	0.103 (0.0739)
Not urban	0.0112 (0.0418)	0.00806 (0.0753)
Grandparental childcare		-0.982 (0.616)
Constant	-0.147 (0.131)	0.376 (0.259)
Observations	492	492
R-squared		-0.654

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix E – Tests on endogeneity and instrument validity for the regression in Appendix D

	Statistic	P-value
Wu-Hausman F-test (F(1,419))	4.68316	0.0310
Kleibergen-Paap rk Wald F test	17.616	
Kleibergen-Paap rk LM statistic	13.127	0.0003

Appendix F– IV estimates of grandparental childcare on the logarithm of number of hours worked weekly for the sub-sample of working mothers who do not provide assistance to their parents or parents-in-law

Dependent variable	Grandparental Childcare	Log (Number of Hours Worked Weekly)
Grandmother alive	0.134*** (0.0422)	
Ages: 25-35	-0.200 (0.238)	0.155 (0.358)
Ages: 35-44	-0.188 (0.239)	0.150 (0.361)
Ages: 45-54	-0.216 (0.242)	0.265 (0.402)
Secondary education	0.210 (0.148)	-0.267 (0.415)
Tertiary education	0.213 (0.139)	-0.0655 (0.402)
At least one child between the ages of 0-3	0.287*** (0.0617)	-0.361 (0.304)
Number of children under 10	0.0126 (0.0358)	-0.0420 (0.0676)
Single mother	0.0271 (0.0660)	0.0326 (0.148)
Very urban	0.0751 (0.0698)	-0.0412 (0.144)
Not urban	0.0784 (0.0635)	-0.228 (0.154)
Grandparental childcare		1.589 (0.985)
Constant	-0.126 (0.235)	3.181*** (0.351)
Observations	252	252
R-squared		-1.561

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Appendix G - Tests on endogeneity and instrument validity for the regression in Appendix F

	Relevant Statistic	P-value
Wu-Hausman F-test (F(1,239))	3.68437	0.0561
Kleibergen-Paap rk Wald F test	10.076	
Kleibergen-Paap rk LM statistic	7.725	0.005
Anderson-Rubin Wald test	3.66	0.053

Appendix H – Bivariate probit estimates for grandparental childcare on maternal labor force participation

Dependent variable:	(1 st Stage)	(2 nd Stage)
	Grandparental Childcare	Maternal Labor Force Participation
Grandparental Childcare		0.9709 (0.7203)
Ages: 25-35	-0.6490 (0.5629)	0.4967 (0.3887)
Ages: 35-44	-0.7266 (0.5646)	0.7019* (0.3961)
Ages: 45-54	-5.6511*** (0.5774)	0.2813 (0.4596)
Secondary education	5.1610*** (0.5313)	-0.2134 (0.4430)
Tertiary education	5.4908*** (0.4779)	0.4247 (0.4445)
At least one child between the ages of 0-3	0.8252*** (0.1748)	-0.0921 (0.2166)
Number of children under 10	0.0996 (0.1088)	-0.1459 (0.0912)
Single mother	-0.0372 (0.2261)	-0.0985 (0.1708)
Extremely urban	0.1999 (0.1936)	0.2393 (0.1758)
Not urban	0.0799 (0.2061)	0.0460 (0.1673)
Grandmother alive	1.1095*** (0.4152)	
athrho	-0.2741 (0.4777)	
Constant	-7.4412*** (0.9150)	-0.2089 (0.5377)
Observations	508	508
R2		

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Appendix I- IV estimates of grandparental childcare on maternal labor force participation on wave 13 (year 2020)

Dependent variable:	(1) Grandparental childcare	(2) Maternal Labor Force Participation
Grandmother alive	0.0948*** (0.0231)	
Ages: 25-35	0.265*** (0.0529)	0.424* (0.231)
Ages: 35-44	0.201*** (0.0602)	0.496** (0.198)
Ages: 45-54	0.198*** (0.0568)	0.566*** (0.202)
Secondary education	-0.135 (0.140)	0.129 (0.244)
Tertiary education	-0.0644 (0.136)	0.483** (0.213)
At least one child between the ages of 0-3	0.211*** (0.0418)	-0.134 (0.169)
Number of children under 10	0.0241 (0.0285)	-0.109** (0.0440)
Single mother	-0.00183 (0.0412)	-0.0147 (0.0643)
Very urban	0.0349 (0.0468)	0.0811 (0.0709)
Not urban	0.0358 (0.0396)	0.0185 (0.0633)
Grandparental childcare		0.708 (0.740)
Constant	-0.248 (0.157)	-0.0650 (0.275)
Observations	412	412
R-squared		-0.030

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Appendix J - Tests on endogeneity and instrument validity for the regression in Appendix I

	Relevant Statistic	P-value
Wu-Hausman F-test (F(1,399))	.552818	0.4576
Kleibergen-Paap rk Wald F test	16.829	
Kleibergen-Paap rk LM statistic	11.891	0.0006

Appendix K – List of abbreviations

2SLS	Two-stage-least-squares
3SLS	Three-stage-least-squares
FE	Fixed Effects
FEIV	Fixed Effects Instrumental Variables
ITT	Intention-to-treat
IV	Instrumental Variables
LISS	Longitudinal Internet Studies for Social Sciences
OLS	Ordinary Least Squares
MLFP	Maternal Labor Force Participation
SHARE	Survey of Health and Retirement in Europe