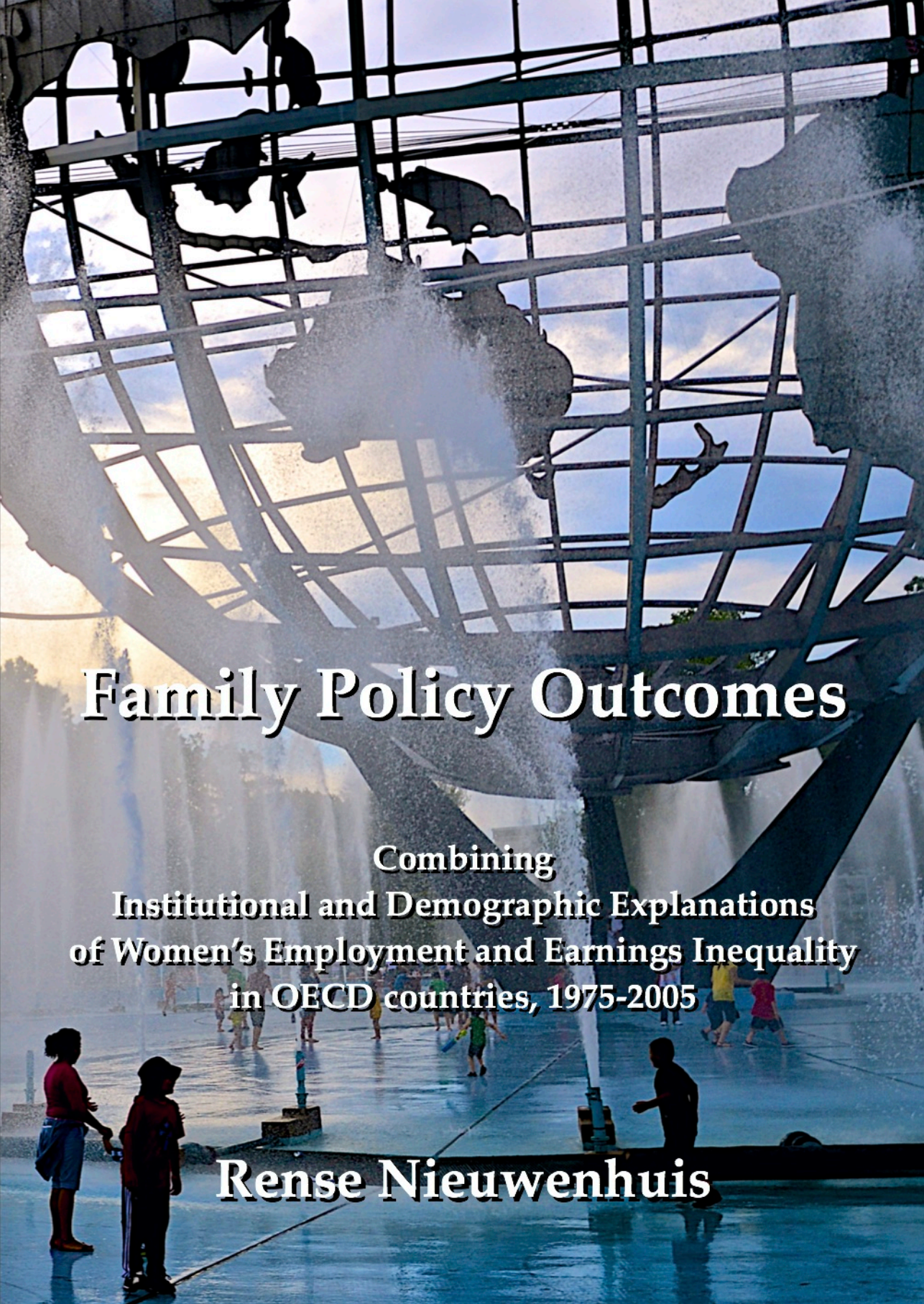


Family Policy Outcomes

Combining
Institutional and Demographic Explanations
of Women's Employment and Earnings Inequality
in OECD countries, 1975-2005

Rense Nieuwenhuis

FAMILY POLICY OUTCOMES

RENSE NIEUWENHUIS

Combining Institutional and Demographic Explanations of Women's Employment and Earnings Inequality in OECD Countries, 1975-2005

January 2014

Graduation Committee:

Chair	Prof. Dr. K. I. van Oudenhoven-van der Zee
Promotor	Prof. Dr. A. Need
Assistant Promotor	Dr. H. van der Kolk
Members	Prof. Dr. C.W.A.M. Aarts (University of Twente)
	Prof. Dr. S.A.H. Denters (University of Twente)
	Prof. Dr. A.H. Gauthier (University of Calgary)
	Prof. Dr. J.C. Gornick (City University of New York)
	Prof. Dr. N.D. de Graaf (University of Oxford)
	Prof. Dr. Ir. A.G. van der Lippe (Utrecht University)

Colofon

Rense Nieuwenhuis: *Family Policy Outcomes,* Combining Institutional and Demographic Explanations of Women's Employment and Earnings Inequality in OECD Countries, 1975-2005, ©January 2014

Cover: Unisphere in Flushing Meadows; Queens, New York City.
Cover Photography: Rense Nieuwenhuis
Typeset in LaTeXusing the ClassicThesis package.
ISBN: 978-90-365-3582-3
DOI: 10.3990/1.9789036535823
Website: www.rensenieuwenhuis.nl/family-policy-outcomes
This website contains (links to) an electronic version of this dissertation, data, publications, contact information, and more.
Printed by CPI Wöhrmann Print Service - Zutphen

Family Policy Outcomes

Combining Institutional and Demographic Explanations of Women's Employment and Earnings Inequality in OECD Countries, 1975-2005

DISSERTATION

to obtain
the degree of doctor at the University of Twente
on the authority of the rector magnificus
Prof. Dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Friday, January 10, 2014 at 16.45

by

Rense Robijn Nieuwenhuis

born July 9, 1981
in Amsterdam, The Netherlands

This dissertation has been approved by:

Promotor: Prof. Dr. A. Need
Assistant Promotor: Dr. H. van der Kolk

ACKNOWLEDGMENTS

Doing a PhD represents several years of supervised training, developing oneself to become a researcher capable of independently contributing to, and participating in, a scientific discipline. Contributing to a scientific discipline means that a PhD candidate is supervised to create scientific products of the highest possible quality, and does so in an increasingly independent manner. Participating in a scientific discipline entails presenting these scientific products to others, frequently discussing these with colleagues, and collaborating with representatives of that discipline. This means that doing a PhD is by no means a companionless endeavor.

Ariana Need and Henk van der Kolk have always formed a great team with the single goal of supervising me to do the best possible research and to grow as an academic. Already in Nijmegen, I learned that Ariana was a supervisor who would inspire me to do innovative research, support me in difficult times, and would always be focused on theory-testing, empirical science. While I was still focused on finishing my dissertation, Ariana was already preparing me for future steps in my career.

Henk helped me to better understand my sometimes complex arguments and to express myself more clearly. With his Socratic method he asked me one question after the other, until I learned what the weakness in my arguments was. His commitment to science made him prepare our meetings meticulously, and his enthusiasm for research led to many original insights.

As a team, Ariana and Henk never had their doors closed and were sure to notice when mine was for too long. Their comments were complementary, and always showed a clear direction for improvement. I could decide upon my own direction for research, but never felt unsupervised. I looked forward to our joint meetings because of their enjoyable and inspirational atmosphere.

My two paranymphs represent my peers, from whom I learned so many things about both their work and my own work, as well as about life as a PhD. In Twente, Wouter is the PhD with whom I discussed these topics most. He has this remarkably clear way of expressing his view on what constitutes good research (not to mention his remarkable sharp sense of humour). We often discussed papers and presentations, after seminars or at one of the several conferences we attended together. I was often impressed by how well he boiled his remarks down to the essentials. We shared the highs, lows, and editorial rejections that are part of being a PhD, and always had a good time. Laurie is so passionate about doing research for a better world, it is impossible not to be inspired by her. Laurie invited Emmy and me to a wonderful day in Queens, where the photo on the cover of my dissertation was taken. During my time in New York we often had lunch or a Venti Soy Chai Tea Latte to discuss our shared research interests and future (now current) research projects, and her comments and questions about my work were spot on.

The department of Public Administration has become too large to thank everyone personally. My roommates Chris, Wouter, Annemieke, Judith, and other fellow PhD candidates Mariecke, Cherelle, Ben, Maurits, Sedef, Ann-Kristin, and Kira made every day more cheerful. Jörgen was my friend and go-to person to informally try out an idea (and for some well-earned procrastination). Ann introduced me to the people of LIS, and Minna introduced me to the world of social policy researchers. Ringo and Marieke taught me how to teach. As trainees, Rob and Mariëlle contributed to the data I used. The energy and ideas of Carolien, Adrie, Mirjan, Veronica, Sawitri, Bengü, Pieter-Jan, Guus, Martin, Kostas & Roula, Ria, Manon, and so many more wonderful people, have helped me during my PhD. Last - but most certainly not least - Annette was invaluably helpful in everything practical and beyond.

In the board of the PhD Network of the University of Twente (P-NUT), I learned the trade of running this organisation from Silja, Josine, Juan, and Giovane. After my term as president, I was very happy to find Nana willing and very capable to take over my responsibili-

ties. With a great team - also including Victor, Björn, Raja, Bijoy, Juan-Carlos, Harmen, Adithya, David, Joana, Mohammad, Febriyani, Burcu, Janne, Mihaela, Anja, Rong, and Jonathan - we professionalised P-NUT, started the tradition of annual PhD days, and voiced our concerns about the studentification of PhD candidates. As president I learned to appreciate the value of a team with widely diverse personalities, made dear friends, and learned how much I enjoy being part of an international community.

The Institute for Innovation and Governance Studies (IGS, headed by Kees Aarts), and the Netherlands Institute of Government (NIG, currently headed by Bas Denters) have facilitated my research project. In addition, I was fortunate to be able to work with, and learn from, scholars from other institutes.

The LIS Datacenter has offices in Luxembourg and in New York. Janet Gornick hosted me in New York, and advised me while I was working on the LIS Chapters of this dissertation. Our meetings and her advice were always very productive, fascinating, and motivating. Janet has strongly supported my work and career opportunities while I worked in New York, continuing this support after I returned home. I met many other wonderful people in New York. Caroline, who helped me cope with visa-stress, an office, and kindly invited me to have dinner with her lovely family followed by a night out in Red Hook. Sarah also welcomed me to a home-cooked dinner, and we often discussed social policies in "Europe, the country".

Teresa Munzi hosted me for a two-week stay in Luxembourg and co-authored a paper with me and Janet. Caroline assisted with everything logistical and Thierry facilitated me with everything binary. Paul provided very important insights in LIS data and my project. Marco made a welcome lunch of really good gnocchi to which Piotr added slightly scary but delicious mushrooms, and this and the rest of the team (Lindsay, Jörg, Carmen, Païvi) made me feel welcome.

Iga Sikorska collaborated with me at the Warsaw School of Economics. I remember this as a very intense and productive week. The project turned out much more complex than anticipated, but on the

final day, the pieces came together. A week later we won the 'IGS Best Paper Award' based on this work.

Manfred te Grotenhuis generously sent me to Vienna, and together with Ben Pelzer we developed, and published about, the influence.ME software. I think fondly of the people of the Research Master Social and Cultural Sciences: Anja, Michael, Rik & Sonja, Marijn, Kim, Marieke, Marloes, Anke, Wieteke and Roza. In our shared and slightly cramped office we were aspiring great things to come from our group (although, I must admit, a baby was a bit beyond our expectations). Before starting my PhD I worked as a research assistant with Mark Levels in Nijmegen. I thoroughly enjoyed this collaboration and learned a great deal from our successful projects, our un(der)reported one, and our conversations about the philosophy of science. Looking back, I regard my time and collaborations in Nijmegen as the best possible preparation for starting my PhD.

To conclude, no sociologist should maintain that their successes are solely their own. This dissertation would never have materialised without me having been raised by my loving parents Jan and Corrie. They have always stimulated me to think critically and independently, and to make my own decisions. Moreover, they have always supported me when I reconsidered and made new decisions. My parents, as well as Koos, Joke, Kjeld, Linda, Mariëlle, Isabel, Katja, Niels, Tijmen, and Ilya, are a very warm family to me.

The final words are, of course, for my dearest Emmy. Our personal and professional lives have seen growth because we complement each other so well. Your energy and curiosity continue to inspire me. I cannot wait to discover and share the adventures ahead of us.

Rense Nieuwenhuis
Stockholm, November 2013

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

QUESTIONS

1

BACKGROUND AND RESEARCH QUESTIONS

Over the past decades, women's employment in OECD countries has increased (Charles, 2011). This resulted in women having higher earnings and some have argued that this has contributed to the increasing earnings inequality between households (Esping-Andersen, 2007, 2009; McCall & Percheski, 2010).

Explanations of why women's employment was higher in some countries than in others, and explanations of trends towards higher rates of women's employment have been based either on women's institutional context, such as family policies, or on women's demographic characteristics such as motherhood, educational level, or marriage (Bernhardt, 1993; Pettit & Hook, 2005; Van der Lippe & Van Dijk, 2002). These institutional and demographic explanations of women's employment have by and large been tested separately, to the point that the distinction between these two types of explanation has been referred to as polarised (Pettit & Hook, 2005, p. 780). In this dissertation we argue that institutional and demographic explanations are not mutually exclusive and that women's employment can best be explained by a combination of institutional and demographic determinants.

This combination of institutional and demographic determinants is also used to explain the extent to which rising women's earnings have affected earnings inequality between households. A demographic explanation suggests that if women's earnings are positively correlated to those of their spouse, this contributes to a larger inequality between households. On the other hand, if earnings inequality among women is low, women's earnings attenuate between-household inequality. We examine how changes in these aspects of women's earnings have af-

fected inequality between households, and use the family-policy context to explain differences between countries in the degree to which women's earnings have affected inequality between households.

The central question answered in this dissertation is:

CENTRAL RESEARCH QUESTION To what extent can (a.) women's employment and (b.) the contribution of women's earnings to inequality between households in OECD countries between 1975 and 2005 be explained by a *combination* of institutional and demographic factors?

1.1 THE FIRST EMPIRICAL REGULARITY: RISING WOMEN'S EMPLOYMENT

Increasing rates of women's employment in OECD countries over recent decades are illustrated in Figure 1.1. This Figure shows an overall trend towards higher female labour force participation rates between 1975 and 2005, and also reveals that this trend varies across countries. For instance, throughout the period covered by Figure 1.1, female labour force participation has been lower in Southern European countries such as Greece, Spain, and Italy than in Nordic countries as Sweden, Finland, and Denmark. Although a trend towards higher female labour force participation was observed in each of the countries represented in Figure 1.1, this positive trend was much stronger in Ireland, Luxembourg, and The Netherlands, than in the Nordic countries.

This variation in female labour force participation rates across countries, and within countries over time, warrants explanation. Two types of explanation are generally given for this variation. The first of these explanations highlights the importance of the institutional context, including family policies that affect women's employment-related decisions. The second type of explanation highlights the importance of women's demographic background, such as being a mother, educational level, and marital status. These institutional and demographic

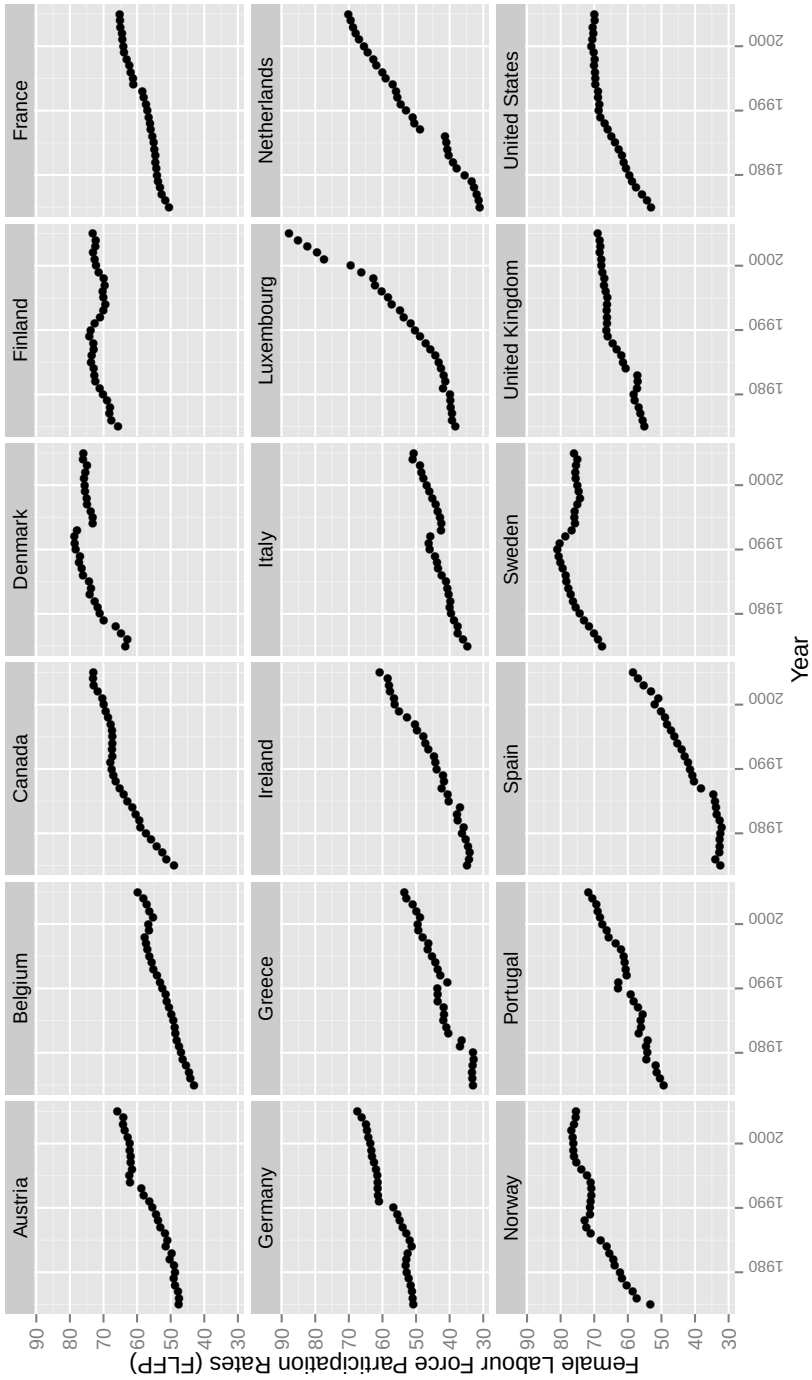


Figure 1.1: Female Labour Force Participation Rates in 18 OECD countries, 1975-2005.
Source: The Comparative Family Policy Database (Gauthier, 2010).

explanations of women's employment, as well as the research testing these explanations, are introduced below.

1.2 INSTITUTIONAL EXPLANATIONS OF RISING WOMEN'S EMPLOYMENT

The first strand of literature explains trends and cross-national variation in the women's employment rates, by considering how the institutional context facilitates women's employment in some countries and impedes women's employment in others. Reconciliation policies, a type of family policy, are regarded as particularly important in helping women to combine motherhood and employment. This may not be true for all types of family policy. In this section, we contrast reconciliation policies with another type of family policy: financial support policies for families with children. We hypothesise about how these two types of family policy affect women's employment differently, and test these hypotheses throughout this dissertation.

1.2.1 Two types of Family Policies

Family policies include a wide range of social policies that aim to support families in various phases of their lives (OECD, 2011). In this dissertation, we differentiate between two categories of family policy: *reconciliation* policies and *financial support* policies.[1] In doing so, we follow Thévenon, who distinguished between family policies minimising the "*indirect cost arising from the incidence of children on the parents' work-life balance and on the aggregate level of employment*" and family policies reducing the "*direct monetary cost of raising children*" (2012, p. 855, also see: Gauthier, 1996).

1 A category of family policies that is beyond the scope of this dissertation is formed by family policies aimed at "*helping parents to have the number of children they desire*" (OECD, 2011, p. 11). This relates to an additional goal of family policies distinguished by Thévenon (2011), raising fertility rates, but also includes policies on contraceptives and induced abortion (Levels, Need, Nieuwenhuis, Sluiter, & Ultee, 2012; Levels, 2011; Rahmqvist, 2006).

Reconciliation policies are family policies that facilitate families in combining family care and employment (OECD, 2011; Thévenon, 2011). Examples of such policies include maternity leave, parental leave, childcare leave, (public) childcare services, and continued pay during leave. Reconciliation policies facilitate combining work and family. Leave policies provide time for care-giving with the guarantee of being able to return to employment afterwards (Gornick & Meyers, 2003). Full or partial compensation of wages during leave provides the opportunity for families to actually take up the available leave. As the employment of women is negatively affected by having (young) children, and because fathers' take-up of leave is substantially lower than that of mothers (Gornick & Meyers, 2003; OECD, 2001), reconciliation policies are argued to positively affect the employment of mothers.

Figure 1.2 on Page 8 shows the availability of paid leave in 18 OECD countries, between 1975 and 2005. The values on the y-axis represent the total number of weeks of combined maternity, parental, and childcare leave that can be taken up with full replacement of wages. (The index used to make Figure 1.2 is also used in Chapters 4 and 6, where the construction of this index is described in more detail.) On average, and in the vast majority of OECD countries, the availability of paid leave increased over time. Substantial differences exist between countries, both in terms of the availability of paid leave to mothers and in terms of how this availability changed over time.

The second category of family policies is financial support policies. Financial support policies provide financial means to families with children (OECD, 2011; Thévenon, 2011). Examples of such policies include family allowances and tax benefits for families with children. These policies have long been criticised as negatively affecting the employment of mothers (Dingeldey, 2001; Schwarz, 2012). Initially, the concern was that, for instance, family allowances would increase women's dependency on their husbands (Gauthier, 1996; ILO, 1924). More recently it was suggested that such financial support may be a "*disincentive*" (Schwarz, 2012, p. 19) to the employment of women (Dingeldey, 2001) and particularly of mothers in the case of financial support policies for families with children.

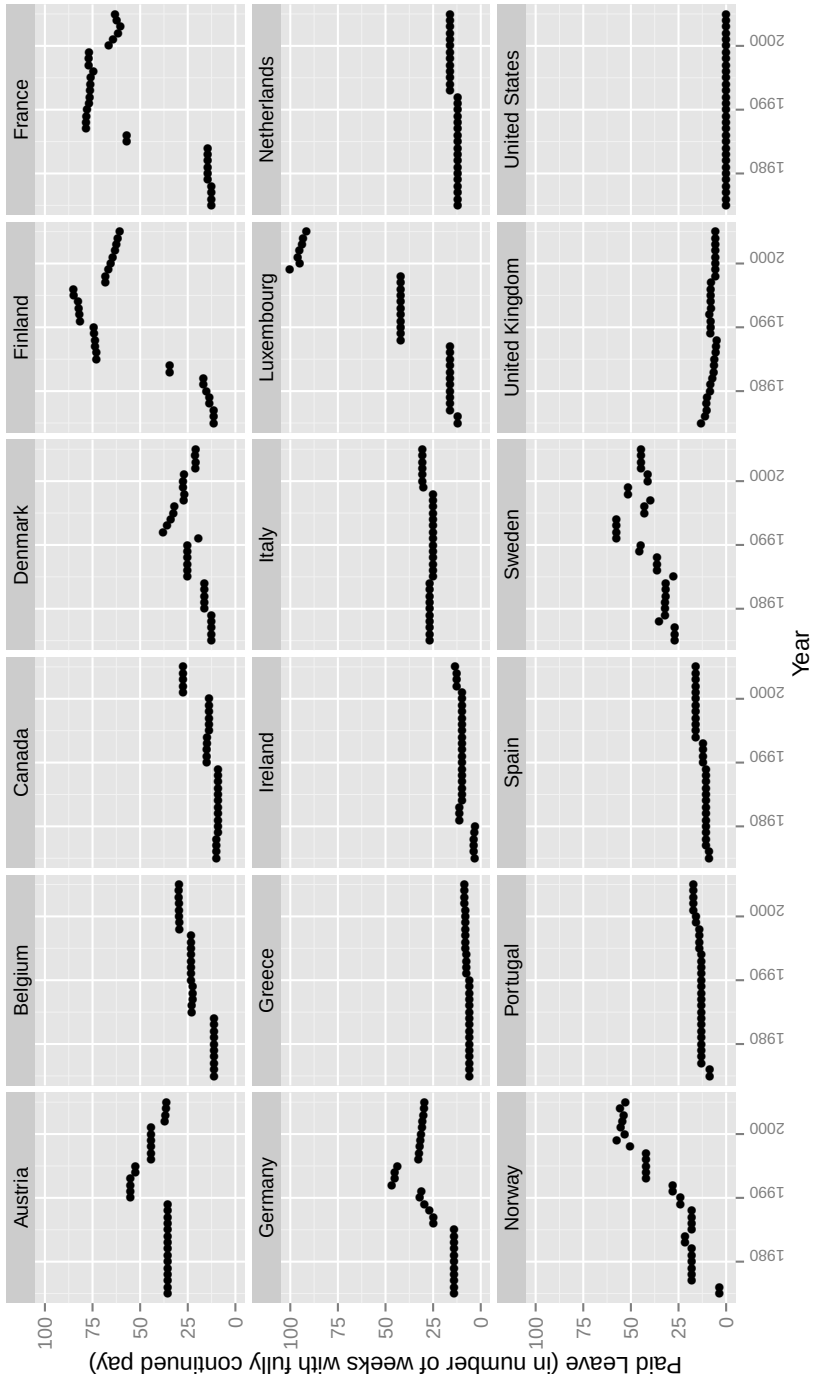


Figure 1.2: Paid Leave in 18 OECD countries, 1975-2005
Source: The Comparative Family Policy Database (Gauthier, 2010).

In Figure 1.3 on Page 10 the expenditure of governments on family allowances is represented as a percentage of Gross Domestic Product (GDP) for 18 OECD countries from 1975 to 2005. Government expenditure on family allowances is used as an indicator of financial support policies for families in Chapters 4 and 6 of this dissertation. Substantial variation exists across countries in both the levels of expenditure, and the trends in expenditure over time. In some countries, such as Canada, Denmark, and Ireland, there is a trend towards higher expenditures, whereas a negative trend is observed in countries including Belgium, France, and the Netherlands.

The trends in paid leave (Figure 1.2) and expenditure on family allowances (Figure 1.3) are in line with a general pattern in which OECD governments have increased the provision of (and spending on) different 'in kind' family policies (cf. Vandenbroucke & Vleminckx, 2011), such as many of the reconciliation policies described above, and have (on average) shown a stable pattern of spending on various cash transfers such as the financial support policies described above (OECD, 2011).

Throughout this dissertation we examine how reconciliation policies and financial support policies, two categories of family policies, have affected women's employment and earnings inequality.

1.3 DEMOGRAPHIC EXPLANATIONS OF RISING WOMEN'S EMPLOYMENT

For a long time, the majority of studies on women's employment ignored the influence of the institutional context (Bernhardt, 1993), but identified a variety of demographic determinants of women's decisions related to employment. These demographic explanations will be shown to be important for explaining the outcomes of family policies. Demographic determinants of women's employment include, but are not limited to, education, marriage, and motherhood. Education has been found to be positively associated with women's employment (e.g. Mincer, 1974), while marriage has been found to negatively af-

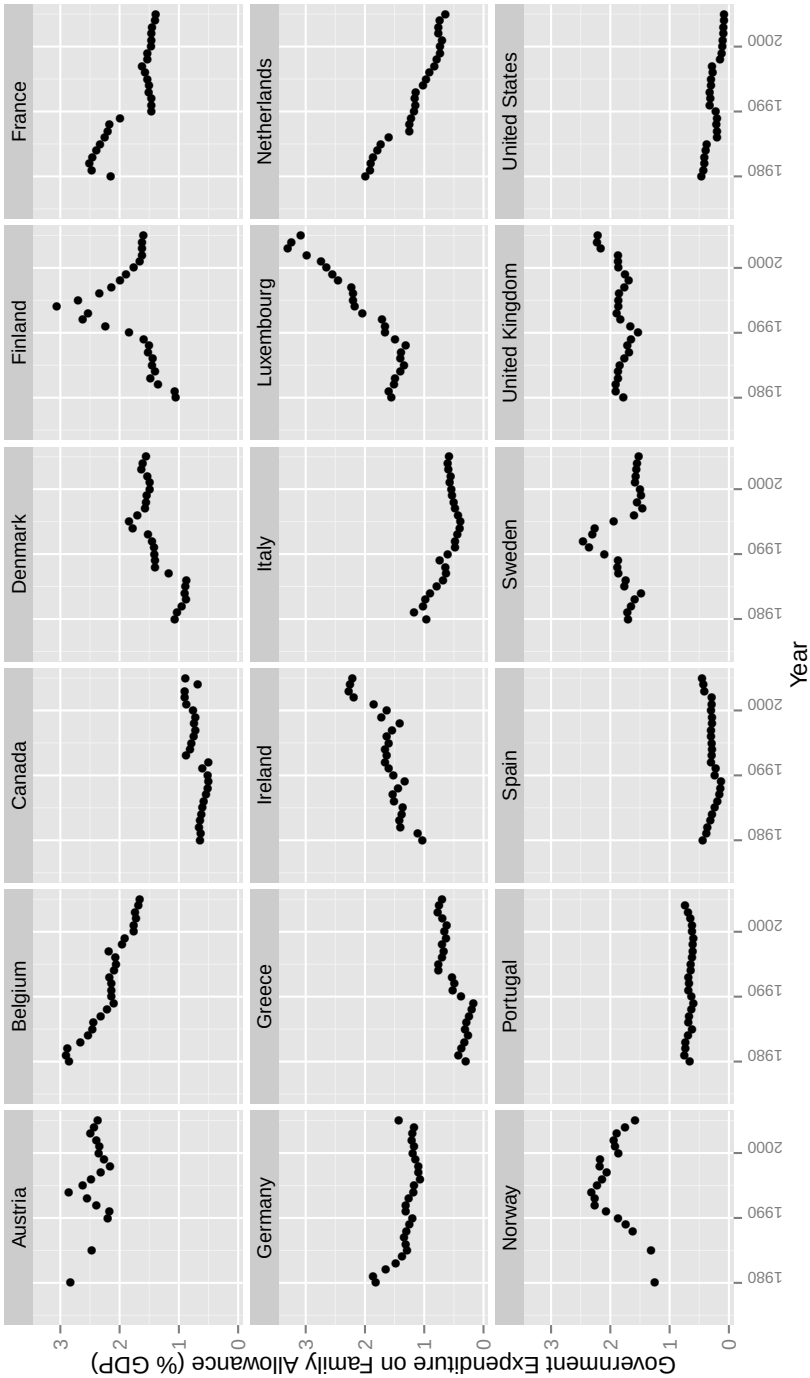


Figure 1.3: Family Allowance Expenditure in 18 OECD countries, 1975-2005
Source: The Comparative Family Policy Database (Gauthier, 2010).

fect women's likelihood of being in employment. (e.g. Becker, 1965, 1985, 1991; Grindstaff, 1988; Mincer, 1958). Motherhood has been identified as being negatively associated with both employment (Blake, 1965; Myrdal, 1941; Myrdal & Klein, 1956) and wages (Waldfogel, 1997; Wellington, 1993).

Explanations of individual women's employment generally treat the decision to seek paid employment as the outcome of an evaluation of the woman's interest in employment and the costs involved with employment. These interests and costs are not exclusively monetary. They also include factors like time spent at work, stress resulting from the practical difficulties of combining employment and motherhood, human capital development, contact with colleagues, and financial independence (Becker, 1991; Bernhardt, 1993; Brewster & Rindfuss, 2000). In single-person households, the likelihood of a person seeking employment is expected to rise with increased investment in human capital, such as a higher level of education, as this investment results in a stronger interest in employment (Del Boca & Locatelli, 2006; Del Boca, Pasqua, & Pronzato, 2009; Pettit & Hook, 2005).

A key insight in new home economics theory is that decisions regarding (women's) employment are often taken in coupled households. Becker (1965; 1991) has argued that it is more efficient for members of a shared household to specialise between home production and economic production. Even when both members can expect equal returns from participation on the labour market, Becker argues that the household is more productive when either one of the members enters the labour market, while the other specialises in home production. When one of the members of the household can expect higher returns from paid employment, for instance as a result of higher investment in human capital, it will be this person who participates on the labour market while the other specialises in home production. Since men, on average and traditionally, had a higher level of education, were married to younger women, had more experience in the labour market, and could expect higher wages in the labour market, this theory was often used to explain why employment rates among men were higher

than those of women, and these explanations were tested successfully (Becker, 1985).

The most important factor in explaining women's employment has been found to be the presence of children in the household (Van der Lippe & Van Dijk, 2002). In new home economics theory it is argued that the presence of children increases the need for home production, which further increases the gain to be had from specialisation. Using the arguments detailed above, it is hypothesised that the presence of children in the household will predominantly limit the employment rates of women. The difficulties of combining the responsibilities involved with raising children and those involved with paid employment lead to the expectation that women are less likely to be employed when they are a mother, (Bernhardt, 1993; Brewster & Rindfuss, 2000; Cramer, 1980; Stycos & Weller, 1967). This incompatibility of roles (Myrdal & Klein, 1956) increases the costs of employment: a mother may evaluate employment as a less valuable option compared to the arrangements she needs to make to combine the responsibilities that result from being a mother with those involved with employment (cf. Sweet, 1981). In other words, mothers have fewer opportunities for employment than women without children.

Demographic explanations of women's employment therefore read that a woman is more likely to be employed if she is highly educated, single, and without children. Trends towards higher female labour force participation rates are then explained by women having fewer children and having their first child at a later age, by women being less likely to be married (and, again, at a later age), and more likely to have higher levels of education. These demographic, person-level, explanations also provide a foundation for understanding how institutional contexts affect women's employment related decisions, and how these decisions are affected by the interplay between women's demographic background and institutional context. This is discussed in the next section, where we combine institutional and demographic explanations of women's employment.

1.4 COMBINING INSTITUTIONAL AND DEMOGRAPHIC EXPLANATIONS OF RISING WOMEN'S EMPLOYMENT

Institutional explanations have predominantly been invoked to help understand differences in women's employment *between* countries, or trends in their employment, based on inter country differences in the context in which women make their employment-related decisions. Demographic explanations have predominantly been invoked to understand differences in employment rates *within* between women with a different demographic background within a single country. To combine these institutional and demographic explanations of women's employment, we formulate a single rational choice theory that is based on new home economics (Becker, 1965, 1991)). Applications of rational choice theory have often been predisposed to formulating explanations based on "social structural determinants" (Hechter & Kanazawa, 1997, p. 193), and have paid considerably less attention to personal motivational factors. In order to combine institutional and demographic explanations of women's employment in a single theory, we use the concepts of *opportunities* to consider the social structural determinants of women's employment and *interests* consider all the reasons which may motivate women to seek employment.

We formulate and test our theory combining institutional and demographic explanations of women's employment throughout this dissertation. An introduction to the theory is illustrated in Figure 1.4. The schematic in Figure 1.4 is limited to two demographic determinants (motherhood and education) and two institutional determinants (reconciliation policies and financial support policies).[2]

In our discussion of the demographic explanations of women's employment, in Section 1.3, it was argued that new home economics sug-

2 For clarity, several determinants of women's employment such as living in a coupled household and the labour market structure, are not represented in this Figure. Also, several arrows in Figure 1.4 represent interaction hypotheses, but not all constitutive terms are shown here (e.g. the direct effect of family policies on employment). In later Chapters these effects are discussed theoretically and estimated in the statistical models used for testing the hypotheses.

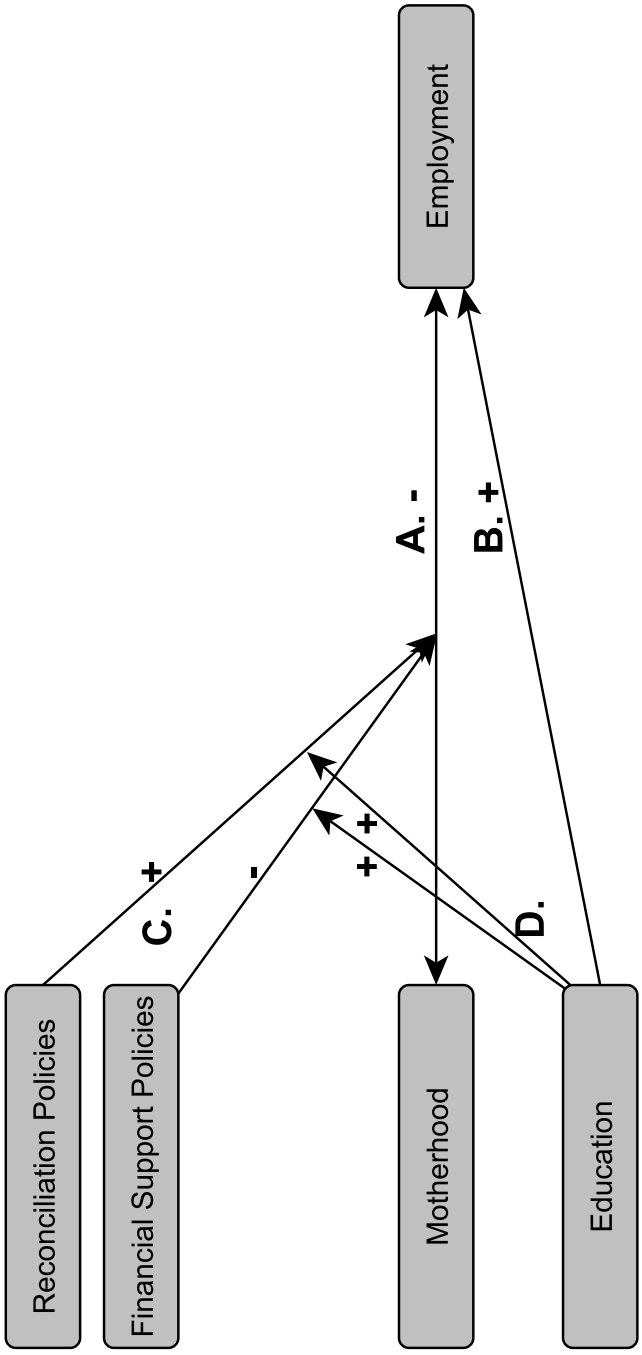


Figure 1.4: Theoretical Model Combining Institutional and Demographic Explanations of Women's Employment

gests that mothers are less likely to be employed than women without children, because they have fewer opportunities. The negative association between motherhood and employment, the '*motherhood-employment gap*', is represented in Figure 1.4 by arrow A. Arrow B. represents the expectation that more educated women are more likely to be employed, both because they may have better opportunities (e.g. because of larger human capital investments) and because they may have a stronger interest in employment. These interests refer to the aforementioned personal motivational factors (Hechter & Kanazawa, 1997).

The demographic explanations for women's employment, as formulated above, are not sensitive to (aspects of) the societal context in which employment-related decisions are made, and therefore cannot explain how family policies affect these decisions. However, using the concept of opportunities, it is possible to explain how family policies affect the decision process of individual women regarding employment.

Reconciliation policies provide opportunities for women to continue to be employed after becoming a mother. As such, reconciliation policies counter the reduced opportunities for employment of mothers compared to women without children. It is thus expected that reconciliation policies reduce the size of the motherhood-employment gap. Financial support policies provide financial means or tax benefits to families with children, effectively reducing the (relative) value of the monetary returns of mothers' employment (Apps & Rees, 2004). Financial support policies provide the opportunity not to be employed to families with children, and therefore to mothers. It is thus expected that financial support policies increase the size of the motherhood-employment gap. These expectations regarding reconciliation policies and financial support policies are represented by the arrows labelled (C.) in Figure 1.4. The *combination* of institutional and demographic explanations of women's employment is more informative than simultaneous reference to determinants of these different strands of explanation. Already in the expectation of how family policies affect the

size of the motherhood-employment gap, the *interaction* between both institutional and demographic explanations was present.

The final step in combining institutional and demographic explanations pertains to how the effect of family policies on the motherhood-employment gap is moderated by women's educational level. This is represented in Figure 1.4 by the arrows labelled (D.). We argue that mere opportunities provided by family policies do not have consequences for those without an interest (in this case: in employment), or conversely that mere interests to act have no consequences without the opportunities to do to so (cf. De Graaf, Need, & Ultee, 2000; Hedström, 2005; Ultee & Luijkx, 1998). This can be applied here, by considering the opportunities provided by family policies in combination with the interest of women (and particularly mothers) in employment. We thus improve upon applications of rational choice theory that are solely based on socio-cultural determinants (Hechter & Kanazawa, 1997), referred to as determining opportunities here, by introducing an additional assumption about interaction between opportunities and interests. Based on the assumption that women with a higher level of education have stronger a interest in employment, this allows differentiation of the outcomes of reconciliation policies and financial support policies between mothers with higher and lower levels of education.

While it is not realistic to assume that mothers are entirely without opportunity for employment, it is realistic to argue that their opportunities are more limited than those of women without children. Similarly, less educated women are not without interest in employment, but it can be assumed that less educated women have a weaker interest in employment than those with higher levels of education.

Women with limited opportunities and limited interest are least likely to be employed, and women with extensive opportunities and a strong interest in employment are most likely to be employed. Women with extensive opportunities but a weak interest in employment are less likely to be employed than women with both extensive opportunities and a strong interest in employment. The opportunities provided by reconciliation policies are therefore expected to have most impact on those with the strongest interest in employment. In other words, it

is expected that reconciliation policies have more impact on reducing the motherhood-employment gap among more educated women than among those with lower levels of education. We have argued that financial support policies provide the opportunity for mothers not to be employed. Based on this argument, we expect these opportunities to have the least effect on those with a strong interest in employment: the more educated.

1.5 EMPIRICAL TESTS OF INSTITUTIONAL EXPLANATIONS OF WOMEN'S EMPLOYMENT

Institutional explanations of women's employment have been tested using both country-level data and person-level data. Country-level data pertain to measurements of policy arrangements at the country-level and to measurements of the countries' population, such as the female labour force participation rate and the total fertility rate. Person-level data pertain to measurements of individuals, such as employment, motherhood, and educational level. Each type of study has their respective advantages and disadvantages, which are discussed below.

1.5.1 The Warning of an Aggregation Paradox

Hypotheses on how family policies have affected women's employment have commonly been tested using country-level data. This use of predominantly country-level data has resulted in interesting and relevant findings, but looking at how fertility and (women's) employment were related to each other revealed a paradox.

Using data aggregated to the country-level, Sundström and Stafford (1992) observed a positive association between female employment rates and total fertility rates across 22 OECD countries in 1988 (also see: Bernhardt, 1993). This suggested that countries with high fertility rates were also the countries with high rates of women's employment. Ahn and Mira (2002) presented an even more enigmatic empirical regular-

ity: whereas the cross-sectional correlation between the female labour force participation rate and the total fertility rate in 22 OECD countries had been negative prior to 1985, it reversed to become a positive correlation after 1985.

This finding garnered a substantial amount of attention (Adserà, 2004; Del Boca et al., 2009; Engelhardt, Kögel, & Prskawetz, 2004; Engelhardt & Prskawetz, 2005; Kögel, 2004; Rocha & Fuster, 2006). It is illustrated in Figure 1.5. In panel A, the association between total fertility rates (TFR) and female labour force participation rates (FLFP) in 22 OECD countries is shown to be negative in 1975. Panel B shows that this correlation was positive in 2000. Panel C shows the trend of this correlation over time, turning from negative to positive around 1985.

The positive country-level correlation between fertility rates and female labour force participation has given rise to unwarranted interpretations. These are discussed in the next section.

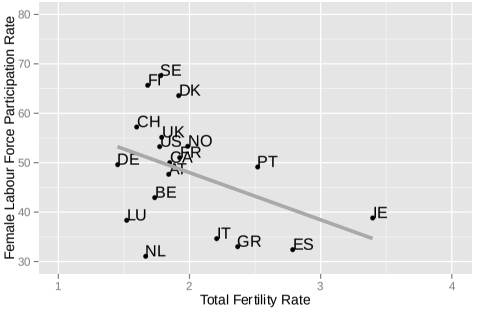
1.5.2 Unwarranted Interpretations of a Country-Level Correlation

The reversal of the correlation between total fertility rates and female labour force participation rates has been interpreted as an indication that motherhood and employment became more compatible over time. For instance, Yasuoka (2012, p. 658) wrote: "*Compatibility between childcare and working is the reason that the negative correlation between fertility and female labour participation is weakened and changes to a positive correlation.*". Similarly, Diprete, Morgan, Engelhardt, & Pacalova (2003, p.442-443): "*. . . the reversal in the cross-sectional correlation between female labour force participation and fertility from negative to positive suggests that the incompatibility between work and child rearing may be weakening across industrialised societies . . .*".

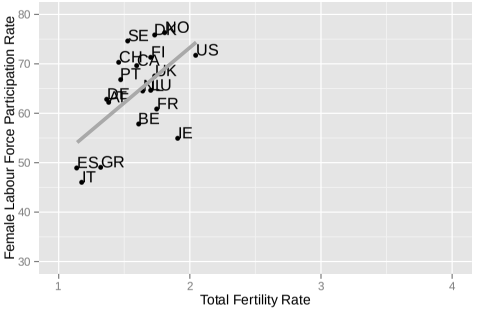
Brewster and Rindfuss (2000) interpret the positive country-level correlation as indicating that conditional on a marked rise of female labour force participation in OECD countries, fertility declined in some countries more than in others. They conclude that "*This comparison suggests that, in some countries [...] women have found ways to combine work*

Figure 1.5: Cross-Country Correlation Between Total Fertility Rates (TFR) and Female Labour Force Participation Rates (FLFP), 1975-1999
Source: Authors' calculations on the Comparative Family Policy Database (Gauthier, 2010).

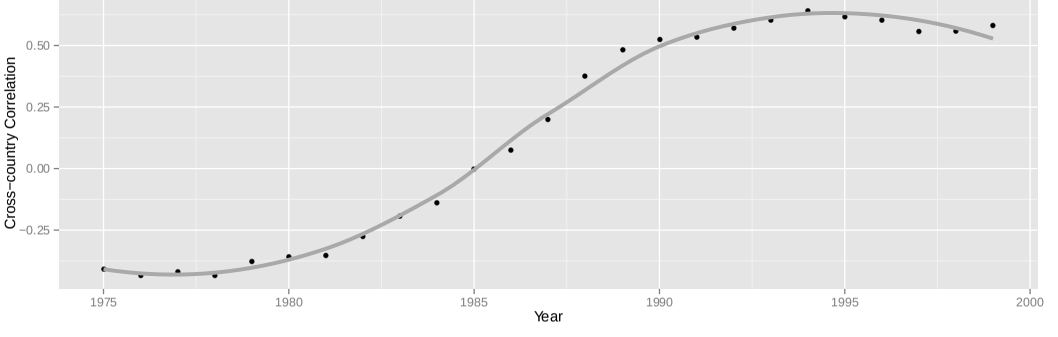
(a) Association between Total Fertility Rate and Female labour Force Participation Rates in 1975



(b) Association between Total Fertility Rates and Female labour Force Participation Rates in 1999



(c) Cross-country Correlation



and child rearing, and in other countries they have not." (p. 279; Also see: Rindfuss and Brewster, 1996, and Sleebos, 2003).

Jaumotte (2003) discussed concerns that increased female labour force participation rates may decrease fertility rates, as employment and fertility were historically found to be negatively correlated over time. She counters this concern by arguing: "*[...] the cross-sectional evidence and recent time-series evidence for some countries do not support such concerns and points to the role of work-family reconciliation policies in avoiding this trade-off*" (p. 72). A similar argument on family policies having reduced the negative association between motherhood and employment was made by Rindfuss, Guzzo and Philip (2003).

Other studies attempted to interpret the positive country-level correlation between fertility rates and female labour force participation rates without referring to the underlying person-level correlation. (Aassve & Lappegård, 2008, p. 68) refer to a number of studies at the country level in which it is hypothesised "*that countries facilitating social policies that make female employment and childrearing more compatible, both experience higher female labour market participation and higher fertility.*". Similar interpretations were giving by Daly (2000), Esping-Andersen (2003), and Stier, Lewin-Epstein, and Braun (2001).

The suggestion that the positive country-level correlation between female labour force participation rates and total fertility rates indicates that motherhood and employment became more compatible over time has been countered as unwarranted (Gauthier, 2007; Kögel, 2004). This interpretation is unwarranted for three reasons. Firstly, it runs the risk of committing an ecological fallacy: a person-level (micro) interpretation based on a country-level (macro) correlation. It could very well be that the person-level association between motherhood and employment became less negative over time in OECD countries, whether caused by changing institutional contexts or not, but this cannot be inferred from the country-level correlation.[3]

3 Such a reversal of the sign of an association after aggregation of data is not uncommon in the social sciences. Already in 1903, Yule formalised that when two or more contingency tables are aggregated, an association between two characteristics can be found on one level of association, but appears to be absent on the other (Yule, 1903, Very early observations of such aggregation paradoxes include Pearson (1896), Galton (1896), Pearson, Lee, & Bramley-Moore (1899), and Fawcett & Lee (1902)). A famous example of an aggre-

Secondly, if the positive country-level correlation did apply at the person-level, thus would suggest that in OECD countries mothers were more likely to be employed than women without children. This would be unprecedented.

Finally, it has been shown that the correlation at the level of the country between total fertility rates and female labour force participation rates turning positive, came from "*country-heterogeneity in the magnitude of the negative time-series association between fertility and female employment*" (Kögel, 2004, p. 45, also see: Engelhardt and Prskawetz (2005) and Engelhardt et al. (2004)). Countries showed different patterns over time in both fertility and employment, which is best illustrated in Figure 1.5 by comparing the different trajectories of the Nordic countries and the Southern-European countries from 1975 (in Panel A) to 1999 (in Panel B). In 1975, the Nordic countries showed relatively low fertility and high female labour force participation, whereas the Southern-European countries showed relatively high fertility and low employment. In 1999, the fertility in the Nordic countries was similar to that in 1975, but employment had risen. In the same period, the Southern-European countries showed a decline in fertility and a rise in employment that was similar to that observed in the Nordic countries. As a consequence, in 1999 the Southern European countries had both lower fertility rates and lower female employment rates than the Nordic countries, resulting in the positive country-level correlation. Thus, variation in the country-level trends caused the country-level correlation to turn positive, which does not imply necessarily that the person-level correlation changed.

---

gation paradox in discrimination research pertains to the sex bias in graduate admission at Berkeley University (Bickel, Hammel, & O'Connell, 1975). Later studies on similar paradoxes in two-by-two contingency tables (e.g. motherhood crossed by employment) found that in contrast with the association in each of the underlying contingency tables, the association in the aggregated table can be stronger, weaker, absent, or of a different direction (Blyth, 1972; Gheng, 1992; Good & Mittal, 1987; Mittal, 1991; Pearl, 2009; Saari, 1995; Samuels, 1993; Simpson, 1951; Wagner, 1982; Yule, 1903).). Also, it was shown that aggregation paradoxes can occur with various measures of association such as the odds ratio, correlation coefficient, and regression parameters (Messick & Van de Geer, 1981).

1.5.3 Country-Level Tests of Institutional Explanations of Women's Employment

The aggregation paradox discussed in the previous section does not imply that country-level data cannot be used to test institutional explanations of women's employment, but it does mean that country-level data is not informative about the person-level association between motherhood and employment. Moreover, institutional explanations of country-variation or trends in women's employment have typically been tested using country-level data as inter-country comparable person-level data covering a long period of time were not available.

Jaumotte (2003) found women's total labour force participation rate to be higher where there was high government expenditure on childcare support, as a reconciliation policy. An OECD (2011) report showed that expensive childcare arrangements in a country were associated with lower rates of women's full-time employment. The results regarding parental leave were mixed. Schwarz (2012) found parental leave to be positively associated with female labour force participation, while an OECD report (2011) found longer periods of leave to be associated with lower rates of female employment. Using country-level data Jaumotte (2003) presented a weakly positive (correlation of .05) linear association between the duration of paid leave and the female labour force participation rate in a country.

Research on the outcomes of financial support policies for families is limited compared to that on reconciliation policies. Scholars have argued that joint taxation of members of a household reduces women's employment rates in a country (Apps & Rees, 2004; Schwarz, 2012; Thévenon, 2011; Thévenon & Luci, 2012). Jaumotte (2003) found that tax disincentives for second earners reduce women's employment rates, but she also found that financial benefits for families with children reduced women's (part-time) employment rates.

1.5.4 Person-level Tests of Institutional Explanations of Women's Employment

Studies using person-level data have also examined the outcomes of family policies. We distinguish between three strategies used for testing institutional explanations of women's employment using person-level data.

Firstly, studies covering a single country have surveyed individual women about their employment history and regressed the individual women's employment decisions on their reported expenditure on childcare (Blau & Robins, 1988, 1989, 1991; Cleveland, Gunderson, & Hyatt, 1996; Connelly, 1992; Heckman, 1974; Ribar, 1992, 1995). This approach led to the conclusion that high costs for childcare are negatively associated with women's employment. Studying women's intentions about employment Presser and Baldwin (1980) found that a "*substantial minority*" (p.1202) of mothers reported that they would seek employment (or work more hours) if childcare availability improved. The outcomes of parental leave were studied using person-level data in which (the duration of) leave take-up was measured. It was found that women in the United States and Sweden had lower wages after taking leave from work, and that this decline in wage was stronger the longer the leave (Albrecht & Edin, 1999; Jacobsen & Levin, 1995; Mincer & Ofek, 1982). In Germany, it was found that an extensive duration of leave reduced the likelihood of women's return to employment after childbirth (Gorlich & De Grip, 2008; Ondrich, Spiess, & Yang, 1996).

Secondly, studies on the outcomes of family policies on women's employment based on person-level data have compared person-level data from a limited number (typically two or three) of countries. The countries in this type of study are purposely selected to represent widely different or highly similar family policy arrangements. Differences between the institutional contexts of the selected countries are described in detail, and it is hypothesised how person-level factors affect women's employment differently in these countries. Regression models are estimated separately for each country in the analysis. For instance, Charles, Buchmann, Halebsky, Powers, and Smith (2001) describe maternal employment as being negatively affected by a variety

of cultural and organisational constraints in Switzerland, more so than in the United States. Using separate regression models on person-level data from these two countries, it was found that the negative associations between being married and employment and between having young children and employment were stronger in Switzerland than in the United States, controlled for a variety of person-level factors. Waldfogel, Rönsen, and Sundström (1999) compared the United States, Britain and Japan and concluded that expansions of family leave in all three countries stimulated mothers' return to employment after childbirth, and had the strongest effect in Japan. Comparing Norway and Sweden using separate regression models based on person-level data for each country, Rönsen & Sundström (1996) found that in both countries women were more likely to return to employment after giving birth, and did so earlier if they were entitled to paid leave with job security. In addition, women were found to return to employment faster in Sweden than in Norway. Comparing Germany and the United States, Grunow, Hofmeister, & Buchholz (2006) found that the institutional context in Germany provided strong incentives for women to exit the labour market during the "*active family phase*" (p. 122, also see: Drobnič, Blossfeld, & Rohwer (1999), Drobnič (2000), and Schober (2013). Gustafsson & Wetzels (1996) found that mothers are more likely to enter the labour market in Sweden than in Germany or Great Britain, which was attributed to the generous family policy arrangements in Sweden. On the other hand, Evertsson and Grunow (2012) found that the accumulated duration of taken up leave in Sweden negatively affected upwards career mobility, in contrast to Germany. Gornick and Jacobs (1998) presented separate analyses of person-level data in seven countries, arranged by a welfare state typology. Comparing differences in outcomes per welfare state type, the authors found employment in the public sector to vary widely between welfare state types, but not to explain variation in the magnitude of the gender wage gap.

Thirdly, studies on the outcomes of family policies on women's employment, based on person-level data, have compared a cross-section of several countries. This most closely resembles the analytical strategy employed by studies using country-level data, as a measurement

of women's employment is regressed on indicators of the social policy context. These indicators are either welfare state typologies, scales and indices, or direct indicators of family policies. The distinction from country-level studies is that the dependent variable is measured on the person-level.

For instance, Stier et al. (2001), describe how different types of welfare states affected women's employment in 12 industrialised countries. The authors concluded that in welfare state types that facilitate the employment of mothers, employment continuity around the time of childbirth is highest and wage penalties resulting from employment discontinuities are smaller. Gornick, Meyers and Ross (1997; 1998) developed an index representing family policies (using various indicators on leave, job protection, and childcare), finding that in the 14 countries studied more extensive family policies reduced the employment penalty for mothers of young children. Person-level data from 22 countries were used by Mandel & Semyonov (2006), who used a 'Welfare State Intervention Index' considering fully paid maternity leave, daycare facilities, and a large public service sector. It was found that a higher score on this Welfare State Intervention Index was associated with higher rates of women's employment, but was negatively associated with women in "*powerful and desirable positions*" (p. 1910). Del Boca et al. (2009) regressed women's employment in seven countries on measures of institutional context, finding that childcare availability stimulates women's employment, family allowances reduce women's employment, and brief periods of leave increase employment while long periods of leave reduce employment. Pettit and Hook (2005) used person-level data to estimate the degree to which motherhood negatively affected the likelihood of women's employment in a cross-section of 19 countries, and interacted the effect of motherhood on employment with indicators of the institutional context: measures on family policies and labour market structure. They found that the gap in employment between mothers and women without children was smaller in countries providing childcare and parental leave (although very long childcare leave negatively affected mothers' employment), but was not affected by labour market characteristics.

In this section we have discussed person-level studies of how family policies affect women's employment, and in the previous section we discussed country-level studies. Next, we discuss the advantages and disadvantages of person-level and country-level studies of how family policies affect women's employment.

1.5.5 Country-level and Person-level Tests of Institutional Explanations: Advantages and Disadvantages

An advantage of country-level studies is the relative ease of covering many countries and / or a long period of time. For instance, the work of Jaumotte (2003) described above was based on 17 OECD countries from 1985 to 1999. Kögel (2004) covered 21 OECD countries from 1960 to 2000, Schwarz (2012) 21 OECD countries from 1979 to 2002, and Semyonov (1980) a cross-section of 61 countries in the 1970s. Typically, such country-level data were used to regress country-level measures of women's employment on indicators of family policy and the labour market. This approach, however, has two limitations. Firstly, country-level associations do not necessarily imply analogous person-level associations, as was discussed extensively in the section on the unwarranted interpretation of country-level correlations. For instance, if high rates of female labour market participation are associated with high rates of fertility (cf. Figure 1.5, Panel B, on Page 19), one cannot infer who participates more frequently on the labour market: mothers, women without children, or both. Secondly, in country-level studies, indicators of institutional contexts are associated with average levels of female labour market participation and average levels of fertility. This approach does not allow the examination of whether and to what extent the outcomes of institutional determinants differ across women with different demographic backgrounds.

Studying the outcomes of family policies using person-level data has three advantages over studying country-level data. Firstly, using person-level data allows the examination of whether the family policies affect those women who are eligible for the benefits of those policies. Secondly, person-level studies differentiate between the effects of

family policies on all women, or specifically on mothers, thereby evaluating the degree to which the institutional context facilitates women in combining motherhood and employment. This cannot be done using country-level data. Finally, in all the studies using person-level data discussed above, the effects of family policies and other country-level determinants were controlled for various person-level demographics, such as women's age, marital status, and educational level. As countries differ not only in their institutional context but also in the demographic composition of their populations, these controls are relevant when studying whether differences in institutional context can explain cross-national variation in women's employment. Moreover, as the demographic composition of countries has changed over time, this should also be accounted for in trend analyses of how changes in family policies affected changes in women's employment.

Person-level studies on the outcomes of family policies have one key disadvantage compared to country-level studies: as a result of more appropriate data not yet being available, country-comparative studies using person-level data typically have only been able to compare cross-sections of countries, or pooled cross-sections of a single country at several moments in time. These studies could not identify either how changes in institutional context have changed the degree to which mothers are less likely to be employed than women without children, or how differences in institutional context between countries could explain inter-country variation in the employment gap between mothers and women without children. This limited capacity for studying differences in women's employment across countries or within countries over time is not a limitation inherent in using person-level data. Until recently, however, the collection of inter-country comparable person-level data had not been carried out long enough to be able to simultaneously study differences between countries and trends within countries using person-level data. Now data has become available that allows us to answer country-level questions (macro) on women's employment using person-level (micro) data, as will be detailed below. These questions are formulated in the next section.

1.6 MACRO–MICRO QUESTIONS ON WOMEN'S EMPLOYMENT

Explanations of cross-national variation and trends in women's employment based on the institutional context and explanations based on women's demographic background are not mutually exclusive. To a large extent these explanations have been tested separately, to a stage where the debate on the topic was referred to as polarised (Pettit & Hook, 2005, p. 780). This polarisation forgoes the possibility that women's decisions about employment are informed by an interplay of their demographic background and the institutional context in which they make these decisions, resulting in a tendency to assume that aspects of the institutional context affect women uniformly and overlooking socio-economic differences between women (Mandel, 2012). Not surprisingly, an integration of institutional (country-level) and demographic (person-level) explanations of women's employment has been called for (Van der Lippe & Van Dijk, 2002).

Testing the integration of institutional and demographic explanations of women's employment has now become possible. Comparable data, both person-level data on demographic determinants of women's employment and country-level data on institutional determinants, is increasingly available, in for instance the Luxembourg Income Study (LIS, 2013), the Mannheim Eurobarometer Trend File (Schmitt & Scholz, 2005), and the Comparative Family Benefits Database (Gauthier, 2010). These datasets provide data that is comparable across a number of countries and cover a period of decades; they are used throughout this dissertation.

As discussed earlier, mothers have been found to be less likely to be employed than women without children (Van der Lippe & Van Dijk, 2002), but the degree to which this '*motherhood-employment gap*' varies across countries and within countries over time has not been extensively studied. We discussed how the country-level correlation between total fertility rates and female labour force participation is not informative in this regard, and most person-level studies have been either based on a single country, or on data covering only a single point

in time. Thus, any comparison of these findings is limited, due to the use of different datasets, measures, and analytical techniques. A notable meta-analysis by Matysiak and Vignoli (2008) systematised many such findings, but could not overcome this limitation. There has therefore been not satisfactory study of how the motherhood-employment gap (based on the demographic explanation of women's employment stating that motherhood inhibits women's employment) has developed across countries and over time.

With the continued collection of comparable person-level data across OECD countries, it has now become possible to answer questions about trends in the size of the motherhood-employment gap. Hence, in this dissertation we develop the '*Comparative Motherhood-Employment Gap Trend File*', which consists of person-level observations and allows us to cover 18 OECD countries over 24 years (between 1975 and 1999). This dataset is introduced at the end of this Chapter and discussed in more detail in Chapter 2. In Chapter 2 we use this dataset to answer our first - descriptive - question on cross-national variation and trends in the size of the motherhood-employment gap in 18 OECD countries, and our second - explanatory - question about how this variability can be explained by reconciliation policies and financial support policies in these countries:

QUESTION 1 How has the size of the motherhood-employment gap changed between 1975 and 1999 in OECD countries?

QUESTION 2 To what extent can institutional developments in OECD countries between 1975 and 1999 explain cross-national variation and trends in the size of the motherhood-employment gap in these countries?

Various scholars have warned that although childcare leave allows women to combine motherhood and employment, very long durations of childcare leave actually negatively affect the employment of women (Bruning & Plantenga, 1999; Gornick & Meyers, 2003; Morgan & Zippel, 2003; Moss & Deven, 1999). Thus, short periods of childcare leave reduce the size of the motherhood-employment gap, whereas very long periods of leave increase the size of the motherhood-

employment gap. With our Comparative Motherhood-Employment Gap Trend File we can improve upon previous analyses by explicitly studying the impact of long periods of leave on the size of the motherhood-employment gap. Using these data in Chapter 3, we answer the question of whether or not childcare leave can be too long:

QUESTION 3 To what extent was the motherhood-employment gap larger between 1975 and 1999 in OECD countries providing long-term childcare leave than in countries providing short-term leave?

Next, we study how the association between family policies and the motherhood-employment gap varies between women with different levels of education. The reason we raise the question of education, is that, in addition to motherhood, it is a key demographic determinant of women's employment. In Chapters 2 and 3, we will answer questions on the extent to which family policies increased the employment of *all* mothers compared to women without children, but will not distinguish between women with different educational backgrounds. Education is often considered in country-comparative and longitudinal studies on the outcomes of family policies. However, typically the extent to which family policies affect the motherhood-employment gap differently among more and less educated women has not been considered. Hence, in Chapter 4, we will answer the question:

QUESTION 4 To what extent did the outcomes of reconciliation policies and financial support policies on the size of the motherhood-employment gap differ between more and less educated women in OECD countries between 1980 and 1999?

1.7 THE SECOND EMPIRICAL REGULARITY: RISING EARNINGS INEQUALITY

In addition to studying women's employment, we examine one aspect of earnings inequality: the contribution of women's earnings from employment to the inequality between households. The focus thus shifts

from the employment of individual women, to the inequality of earnings at the household level.

Women's earnings have been rising in OECD countries over recent decades, not only because of increased employment rates but also because employed women have gained a stronger position in the labour market with higher wages, more working hours, and higher status positions (Costa, 2000). With this stronger position of women in the labour market, women have contributed an increasingly large share of the total household earnings. Consequently, earnings inequality between men and women decreased (Blau & Kahn, 2000; Charles, 2011; Gregory, 2009), and as a result so did earnings inequality *within* households.

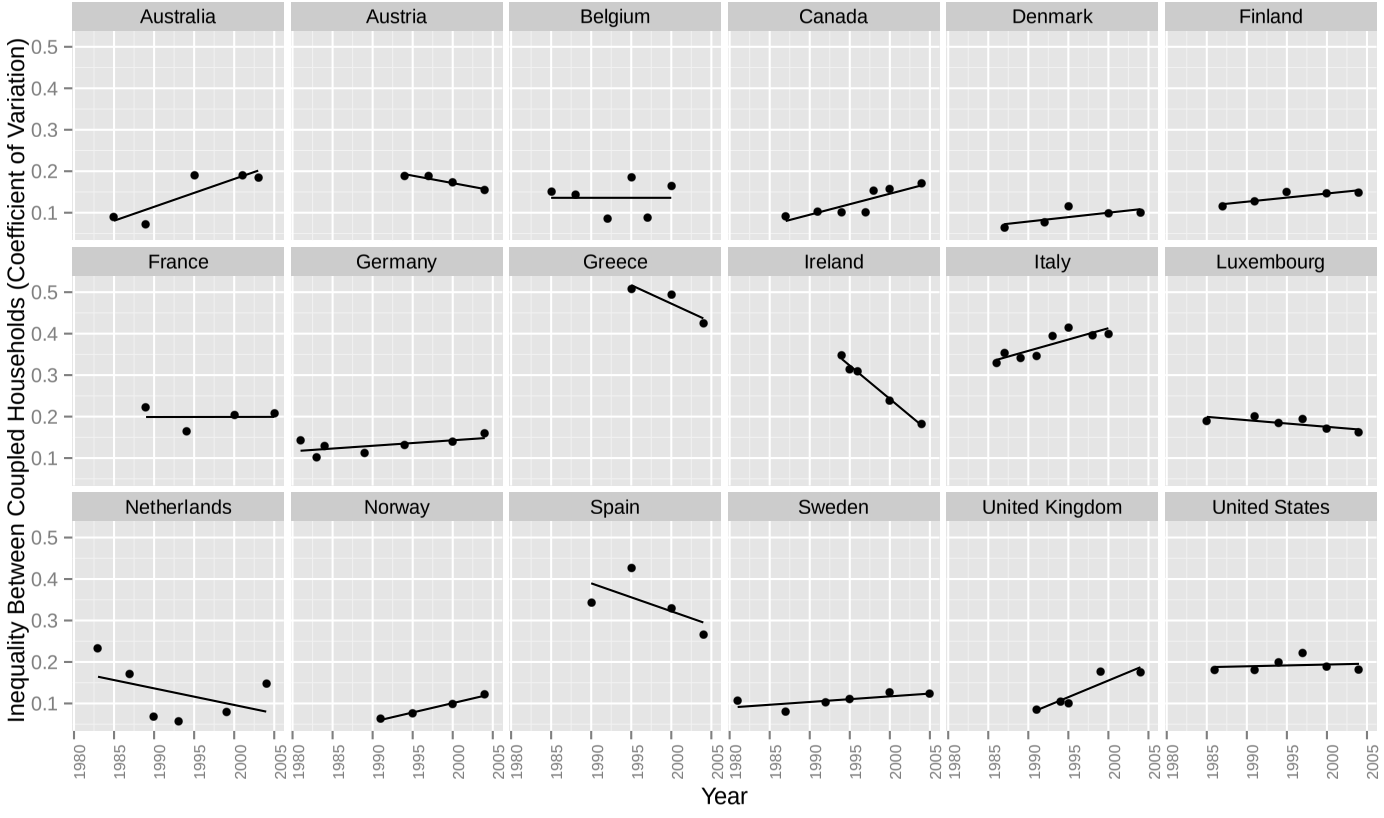
In the same period that women's earnings were rising and earnings inequality *within* households was declining, earnings inequality *between* households was rising in OECD countries. This is illustrated in Figure 1.6. In the majority of the countries in this Figure, earnings inequality between couples has been rising. The main exceptions to this overall pattern are Greece, Ireland and Spain where declines in inequality between coupled households have been observed in recent years. In addition, Figure 1.6 reveals variation in the general pattern of increasing inequality, as well as inter-country differences in the overall level of inequality.

1.8 MACRO–MICRO QUESTIONS ON RISING EARNINGS INEQUALITY

As the earnings of two spouses living in the same household tend to be positively correlated, scholars have often raised the question of how women's increased earnings have affected inequality *between* households, either by contributing to the levels of between-household inequalities, or by attenuating these levels of inequality (Burtless, 2009; Lam, 1997). Even though the consensus in the literature is that women's earnings have an attenuating effect on household inequality (Gregory, 2009), some scholars report that women's earnings in-

Figure 1.6: Earnings Inequality Between Coupled Households in 18 OECD countries, 1981-2005

Data represents earnings inequality between coupled households measured as the coefficient of variation. Source: Luxembourg Income Study. Details are discussed in Chapter 5. Lines represent linear trends. Note that Italy 2004, UK 1986 and Canada 1981 were deleted from the data, as they were overly influential outliers.



crease -rather than decrease- earnings inequalities between households in specific countries (Esping-Andersen, 2007, 2009; McCall & Percheski, 2010). Chapter 5 presents a mathematical argument, based on (a.) the correlation between spouses' earnings, (b.) earnings inequality among women, and (c.) women's share in total household earnings, that the increased earnings of women are indeed likely to attenuate between-household inequalities. This chapter also presents a visualisation of the mathematical argument of how spouses' earnings can be positively correlated while at the same time women's earnings attenuate (rather than exacerbate) earnings inequalities between households. We answer two questions:

QUESTION 5 To what extent have women's earnings attenuated earnings inequalities *between* households in 18 OECD countries from 1981 to 2005?

QUESTION 6 To what extent have changes in (a.) the correlation between spouses' earnings, (b.) earnings inequality among women, and (c.) women's share in total household earnings, affected the degree to which women's earnings attenuate household-level inequality in OECD countries between 1981 and 2005?

Comparative studies on how women's earnings affect inequalities between households, find considerable variation between countries - and within countries over time - in the degree to which women's earnings attenuate (and in a few exceptional cases: exacerbate) earnings inequalities between households (Harkness, 2013; Jäntti, 1997; Pasqua, 2002). Only a few studies have tested explanations for these differences, and those studies focused on household composition rather than institutional context. Here, we combine institutional and demographic explanations again. The demographic explanations here refer to the correlation between spouses' earnings, and how this correlation contributes to inequality between households. It will be shown that the earnings inequality among women decreased over time because fewer women had no earnings at all (a result of increase female labour force participation), which can be referred to as a demographic explanation as well. The institutional explanations here are the two types of family

policies: reconciliation policies and financial support policies to families with children. In Chapter 6 we therefore answer the question:

QUESTION 7 To what extent can cross-national variation in the degree to which women's earnings attenuate inequalities between households in 18 OECD countries from 1981 to 2005 be explained by (a.) reconciliation policies and (b.) financial support policies?

1.9 INNOVATIONS

1.9.1 Innovative Questions

The background of the research questions and the innovation that is achieved by answering these questions has been discussed above. Here, we briefly highlight the innovative character of combining institutional and demographic explanations of women's employment.

Combining institutional and demographic explanations of women's employment allows us to answer existing questions more effectively. This can be achieved by examining the effects of institutional contexts while controlling for women's demographic background characteristics. In addition, in Chapter 3 we revisit the question of whether childcare leave can be too long, and progress the existing research by explicitly examining how very long periods of leave affect the degree to which mothers are less likely to be employed than women without children.

The explicit interaction between institutional and demographic characteristics also allows us to answer new questions. In Chapter 2 this is done by explicitly testing how changes in the institutional context affected trends in the degree to which motherhood and employment are negatively associated. In Chapter 4 we go one step further, and differentiate this effect by educational level.

In Chapter 5, the demographic explanations are reflected in the argument that the rise in employment has particularly resulted in greater employment rates for more educated women, who also tend to have higher earnings. Moreover, these women are also more likely to have

more educated and higher earning spouses than less educated women. In Chapter 6 this is again combined with institutional explanations, based on the argument that women with different demographic backgrounds benefit to different extents from family policies. This allows us to answer the new research question of whether cross-national variation in the degree to which women's earnings affect between-household inequality can be explained by family policy arrangements in these countries.

1.9.2 Innovative Theories

In this dissertation, preliminary answers to the explanatory research questions will be derived from new home economics theory, which is a rational choice theory that is based on the human capital model. This theory was introduced in the previous section on combining institutional and demographic explanations of women's employment (Section 1.4). The application of rational choice theory on women's employment (Becker, 1991; Mincer, 1958) is not new, but will prove to be very useful in providing the person-level foundation for answering the new questions about the outcomes of family policies that were outlined above.

The theoretical innovation in this dissertation with respect to the two types of family policies, is that we hypothesise about both reconciliation policies and financial support policies based on a single theoretical argument that is based on the concept of opportunities. In addition, we test the contrasting effects of reconciliation policies and financial support policies on the employment of women simultaneously. This is particularly innovative with respect to financial support policies. Schwarz (2012, p. 18) described the question of how tax arrangements affect women's employment as "*under-researched*".

In Chapter 3 we challenge our assumption made in Chapter 2 about the opportunities provided by a specific type of reconciliation policies: childcare leave. Specifically, we challenge the assumption we make in Chapter 2 that when the duration of this childcare leave is increased, and thus more opportunities are provided to mothers, this will always

result in a smaller gap in employment between mothers and women without children. In other words, to answer our third research question whether leave can be too long, we formulate a hypothesis on how short and (overly) long periods of leave affect the employment of women differently.

To answer our fourth question, we have to improve our theory on family policy outcomes as formulated in Chapter 2. The opportunity-based explanation of the outcomes of family policies formulated in that Chapter cannot explain how these policies differently affect women with low and high levels of education. Hence, we introduce the concept of interests: the "*driving force*" of an individual person's actions (Coleman, 1990, p. 509, also see: Spillman and Strang, 2013). We argue that the mere opportunities provided by family policies do not have consequences for those without an interest, or conversely that mere interest in acting has no consequences without the opportunities to do to so (cf. De Graaf et al., 2000; Hedström, 2005; Ultee & Luijkx, 1998).

In Part III of this dissertation the unit of analysis changes from women to households, answering questions about how women's earnings have affected between-household earnings inequality. In Chapter 5 we derive (and test) hypotheses on the degree to which women's earnings increase or attenuate between-household earnings inequality based on three aspects of women's earnings: the correlation between spouses' earnings, earnings inequality among women, and the share of her earnings in the total earnings of the household. In Chapter 6 we present another test of how reconciliation policies and financial support policies have different outcomes, and test hypotheses on how these two types of family policies affect each of these aspects of women's earnings differently, and consequently how these policies have influenced the degree to which women's earnings increase or attenuate between-household earnings inequalities.

1.9.3 Innovative Methods

A Less Crazy Methodology: Answering Country-Level Questions Using Person-Level Data

Kittel referred to answering country-comparative questions using only (quantitative) aggregate data on the country-level as being a '*crazy methodology*' (2005a; 2006, also see: Cartwright (2002); Kittel (2005b); Pennings (2005); Ultee (2005)). Firstly, Kittel argued that when different theories predict the same country-level correlation, based on different arguments, the analysis of country-level data provides no means of testing which of these theories best explains the correlation. Secondly, Kittel argued against country-level analyses because they require the assumption of the 'modal mover' at the country-level: a representation of how countries typically 'behave'. However, the number of countries is too small and the heterogeneity between countries too big to realistically be able to assume such a country-level modal mover (Kittel, 2005a).

Kittel's critique (2005a) does not apply to the analyses presented in this dissertation, as we use person-level survey data to answer country-level questions. Kittel suggests two solutions for (quantitative) country comparative research. Firstly, he suggests not studying 'law-like regularities' (Kittel, 2006, p.667) at the country-level, but instead studying the person-level processes that bring about country-level regularities. This argument is not unlike the explanatory model proposed by Coleman (1990) in which country-level associations are explained by macro-to-micro, micro-to-micro, and micro-to-macro hypotheses. Secondly, Kittel (2006) acknowledges that the problems he identified do not apply to country-comparative analyses aggregative of behaviour (in contrast to emergent country-level phenomena) using person-level survey data.

Person-Level Data to Compare Countries

Throughout this dissertation, we analyse two comparative, person-level datasets to test our hypotheses. The first person-level dataset was developed by combining several pre-existing surveys. We call

this dataset the '*Comparative Motherhood-Employment Gap Trend File*'. The vast majority of person-level observations in the Comparative Motherhood-Employment Gap Trend File are from the Mannheim Eurobarometer Trend File, which provides pooled data from Eurobarometer surveys on selected trends in European countries (Schmitt & Scholz, 2005). To cover non-European countries, we added data on the United States and Canada that we obtained from the General Social Survey (Smith, Van Marsden, Hout, & Kim, 2010) and the Canadian Election Study (see http://ces-eec.org/). More details on the harmonisation are given in Chapter 2 and Appendix C. This dataset allowed us to analyse data on 192,484 individual women, and to cover 305 country-years from 1975 to 1999 in 18 OECD countries. The '*Comparative Motherhood-Employment Gap Trend File*' is used in Chapters 2, 3, and 4 and Appendix A of this dissertation.

The second person-level dataset we use in this dissertation is the Luxembourg Income Study (LIS). The Luxembourg Income Study harmonises pre-existing survey data on numerous aspects of income, taxes, social security contributions, transfers, expenditures, consumption, employment, and background information, covering nearly 40 countries with the first wave dating back to around 1980. All datasets in the LIS database are harmonised to a common template, allowing for comparisons across countries and over time. For our purposes, we have used data from 18 OECD countries from waves 1 through 6, covering the period from 1981 to 2005. In total, 99 country-years were covered, encompassing 1,114,444 person-level observations in 572,222 households. Data from the Luxembourg Income Study are used in Chapters 5 and 6, and in Appendix B.

Both person-level datasets were combined with country-level observations on family policies and labour market structure in those countries. These country-level observations were obtained from the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001) and the Comparative Family Benefits Database (Gauthier, 2003). During the process of writing this dissertation, a combined update of these databases became available under the combined name of the Comparative Family Policy Database (Gauthier, 2010).

Using these combined person-level and country-level data sets covering multiple countries over a long period of time has four advantages. Firstly, the repeated observation of countries over time allows us to study not only the effects of family policies on the differences in women's employment and earnings inequality *between* countries, but also how changes in family policies have resulted in changes *within* countries over time.

The second advantage of observing countries repeatedly over time using person-level data is that it allows for a multilevel model specification that separates variation between countries from variation within countries over time (Gelman & Hill, 2007; Hox, 2010). This model structure also controls for all unobserved, time-invariant factors at the country-level. In other words, by observing countries over time we are able to remove (unobserved) variation that is not of interest, thus providing a much stronger test of our hypotheses.

The third advantage of comparing countries using person-level data is that it allows us to measure, in Part II, the degree to which women combine motherhood and employment using the odds ratio. The advantage of using the odds ratio for studying the association between motherhood and women's employment is that it isolates the degree to which women combine motherhood and employment, from both the overall likelihood that women are employed as well as from the overall likelihood that women are a mother (Blanchet & Pennec, 1993). In technical terms: the odds ratio is insensitive to changes in the marginal distribution of employment and motherhood, so that the results of the association between these two concepts are not biased by increased female labour force participation and decreased fertility (Ganzeboom, Treiman, & Ultee, 1991; Lammers, Pelzer, Hendrickx, & Eisinga, 2007). The odds ratio is not informative regarding what the conflict between fertility and employment looks like or how it is experienced by individual women and families (cf. Stone, 2008; Williams & Boushey, 2010), but that is not the goal of this study. Instead, Blanchet and Pennec (1993, p. 121) characterise the odds ratio between motherhood and employment as a descriptive tool of "*the degree of incompatibility between work and child care*".

The final advantage of combining person-level and country-level data is testing two-way and three-way (cross-level) interaction hypotheses. These interaction hypotheses follow directly from our combining institutional and demographic explanations of women's employment and earnings inequality. Interaction terms are. of course, not new and have been applied many times in studies on women's employment. In our study we are using the interactions to not only answer questions about what explains women's employment, but also to answer questions about under what conditions such explanations apply. It has been well established that motherhood is negatively associated with employment but, as discussed above less is known about how this association varies across countries and within countries over time. We analyse this variation, applying interactions between measures of institutional context and the association between motherhood and employment. Moreover, and this has not been done often, in Chapter 4 we apply three-way interactions. These are highly complex, and will therefore be introduced in detail. Using these three-way interactions allows us differentiate between how family policies affect the association between motherhood and employment for more and less educated women.

Tools for Comparative Research

In this dissertation, we develop and apply two sets of tools for country-comparative research using survey data. The first is a set of tools for detecting influential data in multilevel regression models. These are presented in Appendix A, and all multilevel regression models used in this dissertation were evaluated for the presence of influential data. Secondly, in Chapters 5 and 6 data from the Luxembourg Income Study (LIS) are used. In some of these datasets income variables were measured net of income taxes and social security contributions, whereas in other datasets these variables were measured gross of taxes and social security contributions. In Appendix B we present and evaluate tools for netting down income data, improving the comparability of (net) income variables across countries. These tools were applied to all analyses of LIS data presented in Chapters 5 and 6.

1.10 OUTLINE OF THIS DISSERTATION

In this dissertation we combine institutional and demographic explanations of women's employment and earnings inequality to answer questions about two empirical regularities: women's rising labour force participation rates, and increasing earnings inequality between households. After this introduction (Part I), the questions pertaining to the size of the motherhood-employment gap are answered in Part II. This part of the dissertation is titled the 'The Motherhood-Employment Gap' and consists of three chapters in which Questions 1 to 4 are answered. Next, in Part III, questions regarding the contribution of women's earnings to between-household inequality are answered. This part is titled 'Earnings Inequality Within and Between Households' and consists of two chapters in which questions 5 to 7 are answered. Part IV provides a summary and discussion of the findings, as well as a summary in Dutch. In the course of this study, two sets of tools for country-comparative research were developed and these tools are presented in the Appendices. Appendix A provides tools for detecting influential data in multilevel regression models. Appendix B provides tools for comparing net and gross earnings data the LIS Database.

Part II

THE MOTHERHOOD-EMPLOYMENT GAP

This part answers questions pertaining to women's employment and specifically cross-national variation and trends in the degree to which women combine motherhood with employment. Chapter 2 concludes that reconciliation policies provide opportunities for women to combine motherhood with employment, while family allowances increase the gap in employment between women with children and women without. Chapter 3 finds that very long periods of childcare leave negatively affect the employment of mothers. Chapter 4 argues that employment decisions are based not only on the opportunities provided by family policies, but also on individual women's interest in employment. The Chapter concludes that more educated women, with a stronger interest in employment, respond more strongly to the opportunities to combine motherhood and employment provided by reconciliation policies.

2

INSTITUTIONAL AND DEMOGRAPHIC EXPLANATIONS OF WOMEN'S EMPLOYMENT

ABSTRACT

This Chapter combines institutional and demographic explanations of the employment decisions of women, describing and explaining the degree to which, in industrialised countries, mothers are less likely to be employed than women without children. We used data from the Comparative Motherhood-Employment Gap Trend File, covering 18 OECD countries, 192,484 observations, and 305 country-years between 1975 and 1999. These data were combined with measures of institutional context obtained from the Comparative Motherhood-Employment Gap Trend File, the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), the Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013). The combined data were analysed with multilevel logistic regression. The results indicate that, over time, women were increasingly likely to combine motherhood and employment in many, but not all, countries. Both mothers and women without children were more likely to be employed in societies with a large service sector and low unemployment. The employment of women without children was generally unaffected by family policies. Women were more likely to combine motherhood and employment in societies with extensive reconciliation policies and limited family allowances.

1 This Chapter is based on: Nieuwenhuis, R., Need, A., Van der Kolk, H. (2012a). Institutional and Demographic Explanations of Women's Employment in 18 OECD Countries, 1975-1999. *Journal of Marriage and Family*, 74(June), 614-630.

2.1 BACKGROUND AND RESEARCH QUESTIONS

Industrialised societies witnessed a marked rise in women's employment during the last decades of the Twentieth Century. This upward trend has been documented in a substantial literature, which puts forward two key explanations. One of these focuses on the person-level and emphasises women's shifting demographics, such as motherhood, marital status, and educational level; the other highlights the country-level and centres on the changing institutional characteristics of societies, such as family-related policies and labour market structure (Del Boca & Locatelli, 2006; Jaumotte, 2003; Van der Lippe & Van Dijk, 2002). Various authors have suggested integrating these two strands of explanation (Pettit & Hook, 2005; Van der Lippe & Van Dijk, 2002).

Studies of demographic explanations of women's employment have established the importance of determinants such as motherhood, marriage, education, and birth cohort on the returns, motives, interests, and costs associated with women's participation in the labour force (Becker, 1991; Bernhardt, 1993; Brewster & Rindfuss, 2000; Cramer, 1980; Stycos & Weller, 1967). At the same time, the rising employment of women in industrialised countries has been accompanied by demographic shifts such as declining fertility, declining marriage rates, and educational expansion (England & Farkas, 1986; Goldin, 1990; Pettit & Hook, 2005).

Of the demographic determinants, the presence of children in the household has been found to be the most important in explaining whether a woman is employed (Van der Lippe & Van Dijk, 2002). Nonetheless, the negative association between motherhood and employment of women cannot be interpreted simply as a causal effect of having children. Person-level studies have not only found that mothers are more likely to abstain from employment but also that employed women tend to limit their fertility (Bernhardt, 1993; Cramer, 1980; Waite & Stolzenberg, 1976). Notwithstanding this recursive relationship, the negative association between fertility and employment varies between countries and within countries over time. This variation, in

what we call the Motherhood-Employment Gap, is often attributed to the institutional context of the countries in which women live.

Institutional explanations of women's employment suggest that over time it has become easier for women to combine motherhood with employment because of the introduction of reconciliation policies such as maternity leave, pay during leave, and child care leave (Del Boca & Locatelli, 2006; Esping-Andersen, 1999; Rindfuss & Brewster, 1996). Conversely, it is argued that financial support policies to families, such as family allowances and tax benefits to families, decrease the employment of women (Gauthier, 1996; Thévenon, 2011). In addition, the evolving structure of the labour market in the late Twentieth Century has increasingly provided attractive options for women entering employment. Women are especially likely to be employed in countries with a large supply of jobs in the service sector, high female wages (Ahn & Mira, 2002; Brewster & Rindfuss, 2000; Del Boca & Locatelli, 2006), and low unemployment.

The literature on women's employment has mostly tested demographic and institutional explanations separately, to the point that the debate on the topic has been described as polarised (Pettit & Hook, 2005, p. 780). This polarisation, along with a number of other factors, limits previous studies in at least four ways.

Firstly, although they provide great detail on which demographic factors influence a woman's likelihood of participating in paid employment, person-level studies of women's employment have tended to cover only a small number of countries or a short period of time (Matysiak & Vignoli, 2008). The specificity of the various studies in regard to time and place makes it difficult to discern changes across multiple countries and over longer time periods in the negative association between motherhood and women's employment.

Second, country-level studies on the effects of institutional arrangements on women's employment have often used aggregate measures of women's employment, that is, measures that refer to all women. These data enable the study of extended time periods and multiple countries, but their coverage is insufficient for studying differences in employment between mothers and women without children.

In the current Chapter we argue that learning how country-level characteristics affect the person-level association between women's fertility and employment, requires data that are measured at the person-level and that cover a substantial number of countries over an extended period of time. Such data enables integration of the person-level and country-level explanations of women's labour force participation, enabling us to answer the question of whether institutional arrangements have a different influence on the employment of mothers than on the employment of women who do not have children (Cooke & Baxter, 2010). Before discussing the third and fourth factors that have limited previous research on women's employment, we raise our first, descriptive, research question on the motherhood-employment gap: the degree to which mothers were less likely to be employed than women without children:

QUESTION 1 How has the size of the motherhood-employment gap changed between 1975 and 1999 in OECD countries?

When seeking to explain differences in employment between mothers and women without children across multiple countries and over an extended period of time, a third limitation becomes clear: the use of solely a person-level or country-level approaches. Country-level studies that invoke institutional explanations of women's employment often compare countries without accounting for differences in their demographic compositions. Such analyses have tended to attribute all trends in women's employment, and all variation between countries, to differences in institutional arrangements or other contextual characteristics of the various countries. This overlooks the alternative explanation that women's demographic attributes may differ between countries and may have changed over time. For instance, in a number of countries women are now attaining higher educational levels than men. This points to the increasing importance of accounting for demographic attributes, to prevent biased estimates of the impact of institutional context.

The fourth and final factor that has limited previous studies, both country-level and person-level, is the tendency to describe the institutional context using the broad typology of social-democratic, liberal,

and conservative welfare state arrangements (Esping-Andersen, 1999, 2009; Matysiak & Vignoli, 2008). Studies based on welfare state typologies mask the exact aspects of institutional contexts that influence employment rates, both for all women and specifically for mothers. This limits their explanatory power, because it is unclear which features of the various welfare state arrangements affect women's employment. In addition, when a typology of welfare state arrangements is used with country-level data on female labour force participation, it ignores the possibility that some aspects of the institutional context may influence only the employment of mothers and not that of women without children, whereas other aspects may affect all women.

This leads to our second, explanatory, research question:

QUESTION 2 To what extent can institutional developments in OECD countries between 1975 and 1999 explain cross-national variation and trends in the size of the motherhood-employment gap in these countries?

In answering these two questions, the current study improves on existing research and addresses the four limitations just described. We describe the difference in employment between mothers and women without children - called the motherhood-employment gap - in 18 OECD countries over the 25 years from 1975 to 1999. We measured specific aspects of institutional context and tested how each had a different effect on the employment of mothers than on the employment of women without children. We also examined the extent to which differences in employment between mothers and women without children in OECD countries between 1975 and 1999 can be explained by indicators of institutional context. In addition to differentiating between mothers and women without children, our tests accounted for other demographic background characteristics.

2.2 THEORY AND HYPOTHESES

Demographic explanations of women's employment generally treat the decision to seek paid employment as the outcome of an evaluation of the costs and benefits involved. These costs and benefits are not exclusively monetary. They also include, but are not limited to, time spent at work, the practical difficulties of combining employment and motherhood, human capital development, contact with colleagues, and financial independence (Becker, 1991; Bernhardt, 1993; Brewster & Rindfuss, 2000).

In single-person households, the likelihood of a person seeking employment is expected to rise with increased investment in human capital and higher expected returns from employment. Shared households allow for task specialisation. This stimulates household members with more human capital and higher expected returns from employment to specialise in paid labour (Becker, 1991). Demographic explanations therefore hypothesise that women are more likely to be employed if they are childless, if they are single, and if they are highly educated. Because women from older cohorts generally have lower levels of education, married at an earlier age, and became mothers at an earlier age, and were socialised during a period of lower women's employment, the current study also controlled for women's birth cohort.

Institutional context influences the costs and benefits of employment as well, by affecting women's opportunities. To support families with children, governments have implemented two types of policy, each with a different goal: (a) reconciliation policies that facilitate the combination of employment and parenthood and (b) financial support to families, which aims to reduce poverty among those with children (Gauthier, 1996). Both types of policy are targeted exclusively at families with children. We therefore did not expect them to influence the employment decisions of women without children.

Reconciliation policies, such as maternity leave and child care leave, reduce the risk of a woman having to give up her job shortly before or after having a baby. Policies that mandate employers to continue to pay a substantial percentage of workers' wages while they are on ma-

ternity or child care leave reduce the likelihood of a woman refraining from having children while employed or leaving employment when she does have a baby (Joesch, 1997). This leads to:

RECONCILIATION POLICY HYPOTHESIS Reconciliation policies will increase the likelihood that mothers are employed but not affect the likelihood that women without children are employed. As a result, in societies with extensive reconciliation policies, the negative association between motherhood and employment will be weaker than in societies without reconciliation policies.

Family financial support policies reduce families' need to earn extra income. Extensive family allowances and tax benefits to families in effect reduce the cost of raising a child, diminishing the relative value of the monetary returns of mothers' employment (Apps & Rees, 2004, cf.). We therefore expected that financial support policies would widen the gap in employment rates between mothers and women without children. This leads to:

FINANCIAL SUPPORT POLICY HYPOTHESIS Policies that provide financial support for families will reduce the likelihood that mothers are employed but not affect the likelihood that women without children are employed. As a result, in societies that provide extensive financial support to families, the negative association between motherhood and employment will be stronger than in societies that do not provide financial support to families.

In addition to these two types of family policies, the labour market structure influences women's employment (Pettit & Hook, 2005). Unlike the two types of family policy, we expected that the labour market structure would have an equal effect on the employment of mothers and women without children.

Three aspects of the labour market are particularly important. Firstly, the growth of the service sector has contributed to the rise in women's employment (Del Boca & Locatelli, 2006). This holds for both mothers and women without children. Secondly, high average female wages in the manufacturing sector stimulate the employment of women, even

though this sector is regarded as providing less attractive options for women's employment. Finally, overall unemployment levels are important. We expect high unemployment rates to diminish employment options for both mothers and women without children. This leads to:

LABOUR MARKET STRUCTURE HYPOTHESIS Both mothers and women without children will be more likely to be employed in societies with a large service sector, high female wages in manufacturing, and low unemployment. As a result, the size of the service sector, the level of female wages in manufacturing, and unemployment rates will not affect the negative association between motherhood and employment.

2.3 DATA AND METHOD

2.3.1 Person-Level Data

We developed a person-level dataset called the '*Comparative Motherhood-Employment Gap Trend File*'. The vast majority of person-level observations in this dataset are from the Mannheim Eurobarometer Trend File, which provides pooled data from Eurobarometer surveys on selected trends in European countries (Schmitt & Scholz, 2005) as well as variables that were harmonised to ensure comparability. Extensive checks have been performed in the Mannheim Eurobarometer Trend File regarding both the coding of the variables and plausibility of trends over time. To cover non-European countries, we added data on the United States and Canada that we obtained from the General Social Survey (Smith et al., 2010) and the Canadian Election Study (see http://ces-eec.org/). All surveys provided samples that, with the use of sampling weights, were representative of the respective country populations.

The wording of the questions in the three surveys provided comparable measures for several demographic background characteristics of individual women. These enabled us to analyse women's employment in 18 countries from 1975 to 1999 and to differentiate between mothers

and women without children. Nevertheless, the number of comparable measurements was limited. For instance, we only had information on whether or not a woman was employed; no information was available on the number of hours worked per week. Our sample lacked data on the ages of children, and the U.S. data did not enable us to discern whether children were still living at home. By restricting our complete data to women aged 20 to 50, we limited the number of women in our sample whose children had already left home.

All variables at both person-level and country-level were available for 305 country-years. At least 4 years were covered per country. This initially yielded a total of 429,475 observations on individual women. After omitting women younger than 20 and those older than 50, the selection consisted of 244,221 observations. Because of split ballot designs, 9,275 observations had missing values. Of the remaining 235,946 observations, 43,462 (18.4%) contained at least one missing value; these were deleted listwise. Thus, 192,484 observations were left for analysis.[1]

In Table 2.1 we present descriptive statistics of all variables used. We had five person-level measurements:

EMPLOYMENT a binary dependent variable representing whether or not a woman was (coded 1) or was not (coded 0) involved in paid employment at the time of the survey.

MOTHERHOOD a binary variable representing whether or not a woman was a mother (coded 1) or not (coded 0).

PARTNERED HOUSEHOLD a binary variable indicating whether or not a woman was living in a single-person household (single, divorced, or widowed; coded 0) or in a partnered household (married or living as married; coded 1) at the time of the survey.

EDUCATION an interval-level variable indicating the age at which a woman completed or ended her education.

1 In Appendix C - also available on the *Journal of Marriage and Family* web site (http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1741-3737) - we provide detailed information on the data, including wording of the questions, missing data per country-year, and number of valid observations for each country-year.

COHORT an interval-level variable indicating the year a woman was born (range: 1925 – 1979).

2.3.2 Country-Level Data

We combined the pooled person-level survey data from the Comparative Motherhood-Employment Gap Trend File with several country-level indicators of the institutional context. These contextual data concerned both family policies and labour market structure. We obtained them from several databases designed for use in cross-country comparative research. Unfortunately, we had no access to indicators of child care availability. All country-level variables were measured at the country level and varied over time. We centred all of these variables at their respective mean values for all country-years, and divided their values by 10 or 100 for easier interpretation of the parameter estimates. In Table 1 we present both the original measurements and the centred measurements. Our country-level measurements were derived from a number of sources.

From the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), we obtained three measures of reconciliation policies:

MATERNITY LEAVE the duration of maternity and parental leave in weeks divided by 10.

PAY DURING LEAVE benefits during maternity and parental leave, expressed as a percentage of the average wage of women in manufacturing.

CHILD CARE LEAVE the duration of child care leave in weeks divided by 10.

From the Comparative Family Benefits Database (Gauthier, 2003) we obtained two measures of financial support policies:

FAMILY ALLOWANCE the monthly family allowance provided for the first child, expressed in 1970 USD (i.e., country-comparable units

Table 2.1: Descriptive Statistics on Demographic and Institutional Variables (N = 192,484 Individuals From 305 Country-Years From 18 Countries.
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

Variables	Original Measurements				Centred Measurements			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Employment status	0.54		0	1	0.54		0	1
Mother	0.61		0	1	0.61		0	1
Partnered household	0.75		0	1	0.75		0	1
Education	17.18	2.61	10	25	7.18	2.61	0	15
Cohort	53.25	8.48	25	79	28.25	8.48	0	54
Maternity leave (weeks ÷ 10)	15.88	7.69	0	64.00	0.00	0.77	−1.59	4.81
Pay during leave (÷ 100)	75.67	27.59	0	100.00	0.00	0.28	−0.76	0.24
Child care leave (weeks ÷ 10)	44.44	55.50	0	156.00	0.00	5.55	−4.44	11.20
Family allowance (÷ 100)	11.16	10.03	0	46.18	0.00	0.10	−0.11	0.35
Family tax benefits (÷ 10)	18.43	7.37	3.00	39.40	0.00	0.74	−1.54	2.10
Service sector (÷ 10)	61.04	8.00	39.48	75.04	0.00	0.80	−2.16	1.40
Female wages in manufacturing	1.76	1.09	0.08	4.42	0.00	1.09	−1.68	2.66
Unemployment (÷ 10)	8.52	3.69	0.20	23.80	0.00	0.37	−0.83	1.53

of purchasing power parity, corrected for inflation over time using consumer price indexes) divided by 100.

FAMILY TAX BENEFITS the added household-level income after taxes and cash transfers of a single-earner family with two children, relative to such a family without children. It is expressed as a percentage divided by 10.

OECD statistics (OECD, 2013) provided one of our measures of labour market structure:

SERVICE SECTOR the proportion of the total labour force employed in the service sector divided by 10.

From the Comparative Family Benefits Database (Gauthier, 2003), we obtained two additional measures of labour market structure:

FEMALE WAGES IN MANUFACTURING the average hourly wage earned by women working in manufacturing, again expressed in 1970 USD.

UNEMPLOYMENT the percentage of the civilian labour force that was unemployed divided by 10.

In Table 2.2 for each country in our dataset we present the earliest and latest year of observation, the number of country-years, and the number of individual observations.

2.3.3 Statistical Method

We analysed the data using logistic multilevel regression models, estimated using the lme4 package in R (Bates, Maechler, & Dai, 2010). We used a three-level model, with individual observations nested within country-years and countries. The contextual variables were measured at the level of country-years. This nesting structure accounted for time-invariant unobserved country heterogeneity at the country-level while explaining women's employment rates.

Table 2.2: Number of Observations, Countries, and Country-Years (N = 192,484 Individuals From 305 Country-Years From 18 Countries).
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

	Observed Years		No. Observations	
Country	Earliest	Latest	Country-Years	Individual
Austria	1995	1998	4	3309
Belgium	1975	1997	23	14,729
Canada	1984	1998	4	3699
Denmark	1975	1998	24	14,872
France	1975	1998	23	16,785
Germany East	1990	1998	9	6209
Germany West	1975	1998	24	15,060
Greece	1980	1997	18	10,966
Ireland	1975	1998	24	15,183
Italy	1975	1997	23	14,453
Luxembourg	1975	1996	22	4843
Netherlands	1975	1998	24	18,058
Norway	1990	1996	6	1801
Portugal	1985	1998	14	8514
Spain	1985	1999	15	8961
Sweden	1995	1998	4	2351
United Kingdom	1975	1999	25	21,733
United States	1975	1998	20	10,958
Total	1975	1999	305	192,484

2.4 RESULTS

To answer our first question, regarding the extent to which women in OECD countries combined motherhood and employment between 1975 and 1999, we calculated odds ratios for the association between motherhood and employment for each unique combination of country and year. These odds ratios represent the size of the motherhood-employment gap. In Figure 2.1 we present the log of these odds ratios (i.e., logit). A logit smaller than 0 means that mothers were less likely to be employed than women without children. Stronger negative values indicate that women were less likely to combine motherhood and employment in that specific year and country. The grey lines represent the linear trend. A rising slope indicates that women increasingly combined motherhood and employment within that country over time.

The findings depicted in Figure 2.1 answer our first research question: Substantial variation in the degree to which women combined motherhood and employment was observed between countries and within countries over time. Mothers were less likely to be employed than women without children in all but one country during the complete period from 1975 to 1999. The only exception is Denmark, where in several years mothers and women without children were equally likely to hold a paid job. Women were much more likely to combine motherhood and employment in some countries than in others during the entire observation period. For instance, in Denmark and Belgium, the negative association between motherhood and employment was much weaker than in Ireland, Luxembourg, and the United States. In addition to these absolute differences, countries also differed in the degree of change over time. The average trend was that the difference in employment between mothers and women without children grew smaller ($p < .05$, two-tailed). We also tested these trends for significance for each country separately and found no trend in Canada, Denmark, East Germany, France, Greece, Italy, and Spain. The same holds for Austria and Sweden, but these findings only apply to a limited number of years. In Belgium, Ireland, Luxembourg, The Netherlands, the United Kingdom, and the United States we found an upward trend,

indicating the employment gap between mothers and women without children narrowing over time. In most countries in which the employment gap between mothers and women without children grew smaller, the gap was relatively large at the start of our observation period. Finally, we noted that the employment gap increased in Portugal and particularly in West Germany.

Before we explained the variation in the degree to which motherhood and employment were combined between countries and within countries over time, and before we tested our hypotheses, we used regression models to estimate the extent to which demographic and institutional indicators accounted for variation in women's employment (see Table 2.3). Firstly, the null model contains no explanatory variables. It indicates the average proportion of women who were employed and variation between countries and within countries over time. The estimate of the intercept indicated that, on average, $exp(.29)/[1 + exp(.29)] = 57\%$ of the women were employed. Variation in the proportion of women holding a job was indicated by the variation of the random intercept. The variance in employment between countries (.36) was larger than the variation within countries (.13).

In Model 1, we estimated the effects of demographic variables on women's employment. The outcomes were in line with new home economics theory: Women were less likely to be employed if they were a mother, if they lived in a shared household, if they had a low educational level, and if they had been born in an older cohort. These demographic indicators explained 18% of the variation in women's employment between countries, but they were more effective in explaining the variation within countries: $(.13 - .05)/.13 = 62\%$.

In Model 2, we estimated the effects of the institutional indicators. The results show that women were more likely to be employed in societies with long periods of child care leave, low levels of tax benefits to families, a large service sector, and low unemployment. Variation between countries was explained much better by this model (now 50% explained variance). The institutional indicators also improved the explanation of variation within countries (now 75% explained variance). Adding the institutional determinants, however, reduced the deviance

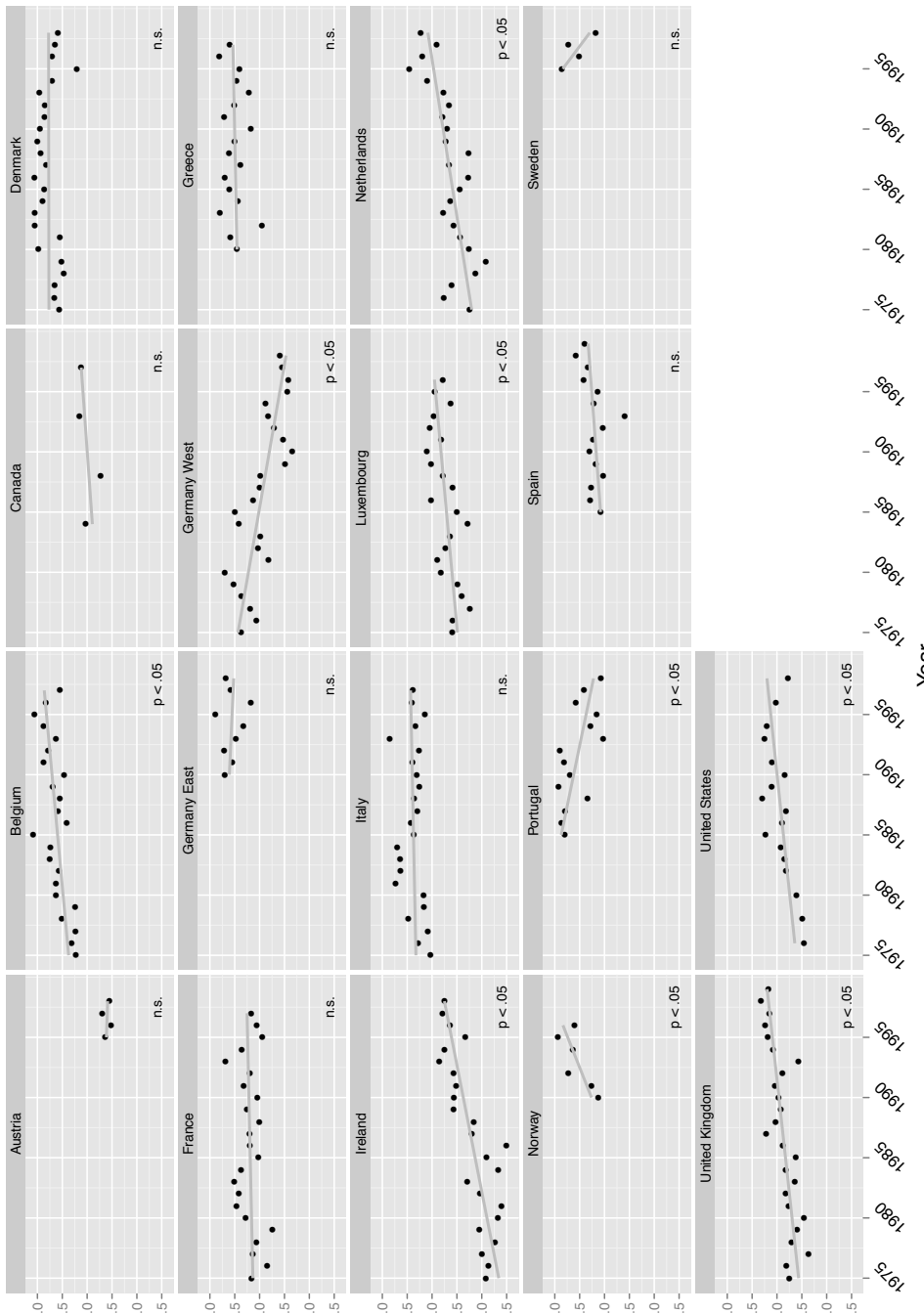


Figure 2.1: Trends in Combining Motherhood and Employment in 18 OECD Countries, 1975-1999.
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

Table 2.3: Multilevel Model Results Predicting Women's Employment From Demographic and Institutional Determinants (N = 192,484 Individuals From 305 Country-Years From 18 Countries).
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

	Null Model		Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE	B	SE
Intercept	0.30*	0.14	−0.33*	0.13	−0.33**	0.10	−0.29**	0.11
Individual level								
Mother			−0.70***	0.01	−0.69***	0.01	−0.72***	0.03
Partnered household			−0.61***	0.01	−0.61***	0.01	−0.61***	0.01
Education			0.16***	0.00	0.16***	0.00	0.16***	0.00
Cohort			0.01***	0.00	0.01***	0.00	0.01***	0.00
Country-year level								
Maternity leave					−0.02	0.05	−0.03	0.05
Pay during leave					0.08	0.16	0.14	0.16
Child care leave					0.02*	0.01	0.02**	0.01
Family allowance					0.08	0.30	0.08	0.29
Family tax benefits					−0.10**	0.04	−0.07*	0.04
Service sector					0.26***	0.03	0.26***	0.03
Female wages in manufacturing					0.06	0.05	0.05	0.04
Unemployment					−0.33***	0.07	−0.36***	0.07
Variances: Country-year								
σ^2 intercept	0.13		0.05		0.03		0.11	
σ^2 of mother							0.28	
R (intercept, mother)							−0.85	
Variances: Country								
σ^2 intercept	0.36		0.29		0.18		0.18	
Deviance	249,754		231,885		231,777		230,153	

* p<0.05, ** p<0.01, *** p<0.001

only very slightly compared with the reduction brought about by the demographic indicators. Thus, women's demographic background explained individual employment decisions better than the institutional arrangements.

Up to this point, our models were based on the assumption that the negative association between motherhood and employment was equally strong in all countries and in all years. However, we argued that this would not be the case, and in Figure 2.1 one can observe considerable variation between countries and within countries over time. We therefore allowed the strength of the association to vary by introducing a random slope at the country-year level in Model 3. The variance (.28) of this random effect represented the variation in the degree to which women combined having children with employment. The average estimate of the association between motherhood and employment was $-.72$. In 95% of the country-years this estimate was between -1.78 and 0.33 ($-.72 \pm 2 \times .28^{1/2}$). Finally, we observed a negative correlation (-.85) at the country-year level between the random intercept and the random slope of the motherhood-employment association. This indicates that the countries in which the negative association between a woman being a mother and being employed was weak were also those with low levels of women's employment. We return to this enigmatic finding in the Discussion section.

2.4.1 Explaining the Degree to Which Employment and Motherhood Are Combined

The previous tests were, of course, not informative about the extent to which institutional arrangements helped women to combine motherhood and employment; they merely provided information on the explanatory power of our regression models and the background for testing our hypotheses. In this section we aim to answer our second question, concerning the extent to which, over time, the institutional context in OECD countries has influenced the employment of mothers differently to the employment of women without children. We estimated cross-level interactions between country-level measurements of

institutional context and the person-level association between motherhood and employment. The results are presented in Tables 2.4, 2.5, and 2.6.

We used the models presented in Table 2.4 to test our reconciliation policy hypothesis. Contrary to our expectation, women without children were less likely to be employed in societies with long maternity leave periods compared to women without children in societies with short maternity leave periods (Model 1). In line with our expectations, we found that pay during leave (Model 2) and child care leave had no effect on the likelihood that women without children would be employed (Model 3). The significant and positive interaction terms indicate that all three reconciliation policies increased the likelihood of mothers being employed. This means that all three reconciliation policies narrowed the motherhood-employment gap.

We tested the financial support hypothesis using the models in Table 2.5. In line with our expectations, neither family allowances (Model 1) nor family tax benefits (Model 2) influenced the employment of women without children. Family tax benefits were also found to have no effect on the employment of mothers, contradicting our hypothesis. In support of our hypothesis, family allowances decreased the likelihood of a mother being employed. Together, these findings indicate that the difference in employment between mothers and women without children was larger in societies with high family allowances.

We tested the labour market structure hypothesis using the models in Table 2.6. As hypothesised, women without children were more likely to be employed in societies with a large service sector (Model 1), high female wages in manufacturing (Model 2), and low unemployment (Model 3). Non-significant estimates were found for the interaction terms between the motherhood – employment association on the one hand, and service sector size and unemployment levels on the other hand. These aspects of institutional context did not appear to have affected mothers' employment differently to the employment of women without children. Both mothers and women without children were more likely to be employed in societies with a large service sector and low unemployment. We calculated the effect of female wages

Table 2.4: Multilevel Model Results Predicting Women's Employment From Cross-Level Interactions Between Reconciliation Policies and a Woman's Motherhood (N = 192,484 Individuals From 305 Country-Years From 18 Countries)
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Intercept	−0.29**	0.11	−0.29**	0.11	−0.30**	0.11
Individual level						
Mother	−0.73***	0.03	−0.73***	0.03	−0.73***	0.03
Partnered household	−0.61***	0.01	−0.61***	0.01	−0.61***	0.01
Education	0.16***	0.00	0.16***	0.00	0.16***	0.00
Cohort	0.01***	0.00	0.01***	0.00	0.01***	0.00
Country-year level						
Maternity leave	−0.12*	0.05	−0.03	0.05	−0.03	0.05
Pay during leave	0.13	0.16	0.03	0.17	0.14	0.16
Child care leave	0.02**	0.01	0.02**	0.01	0.01	0.01
Family allowance	0.08	0.30	0.08	0.30	0.09	0.30
Family tax benefits	−0.07*	0.04	−0.07*	0.04	−0.07*	0.04
Service sector	0.26***	0.03	0.26***	0.03	0.26***	0.03
Female wages in manufacturing	0.05	0.04	0.05	0.04	0.05	0.04
Unemployment	−0.36***	0.07	−0.36***	0.07	−0.36***	0.07
Interaction: Mother ×						
Maternity leave	0.16***	0.04				
Pay during leave			0.20*	0.12		
Child care leave					0.02***	0.01
Variances						
Country-year						
σ^2 intercept	0.11		0.11		0.11	
σ^2 of mother	0.26		0.27		0.27	
R (σ^2 intercept, σ^2 of mother)	−0.84		−0.85		−0.85	
Country						
σ^2 intercept	0.18		0.18		0.18	
Deviance	230,136		230,150		230,141	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2.5: Multilevel Model Results Predicting Women's Employment From Cross-Level Interactions Between Financial Support Policies and a Woman's Motherhood (N = 192,484 Individuals From 305 Country-Years From 18 Countries)
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

	Model 1		Model 2	
	B	SE	B	SE
Intercept	−0.30**	0.11	−0.29**	0.11
Individual level				
Mother	−0.73***	0.03	−0.73***	0.03
Partnered household	−0.61***	0.01	−0.61***	0.01
Education	0.16***	0.00	0.16***	0.00
Cohort	0.01***	0.00	0.01***	0.00
Country-year level				
Maternity leave	−0.03	0.05	−0.03	0.05
Pay during leave	0.13	0.16	0.14	0.16
Child care leave	0.02**	0.01	0.02**	0.01
Family allowance	0.43	0.35	0.08	0.30
Family tax benefits	−0.07*	0.04	−0.05	0.04
Service sector	0.26***	0.03	0.26***	0.03
Female wages in manufacturing	0.05	0.04	0.05	0.04
Unemployment	−0.36***	0.07	−0.36***	0.07
Interaction: Mother ×				
Family allowance	−0.63**	0.33		
Family tax benefits			−0.03	0.04
Variances				
Country-year				
σ^2 intercept	0.11		0.11	
σ^2 of mother	0.27		0.28	
R (σ^2 intercept, σ^2 of mother)	−0.85		−0.85	
Country				
σ^2 intercept	0.18		0.18	
Deviance	230,149		230,152	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2.6: Multilevel Model Results Predicting Women's Employment From Cross-Level Interactions Between Labour Market Structure and a Woman's Motherhood (N = 192,484 Individuals From 305 Country-Years From 18 Countries)
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Intercept	−0.29**	0.11	−0.29**	0.11	−0.30**	0.11
Individual level						
Mother	−0.73***	0.03	−0.73***	0.03	−0.73***	0.03
Partnered household	−0.61***	0.01	−0.61***	0.01	−0.61***	0.01
Education	0.16***	0.00	0.16***	0.00	0.16***	0.00
Cohort	0.01***	0.00	0.01***	0.00	0.01***	0.00
Country-year level						
Maternity leave	−0.03	0.05	−0.03	0.05	−0.03	0.05
Pay during leave	0.13	0.16	0.14	0.16	0.14	0.16
Childcare leave	0.02**	0.01	0.02**	0.01	0.02**	0.01
Family allowance	0.08	0.30	0.08	0.30	0.09	0.30
Family tax benefits	−0.07*	0.04	−0.08*	0.04	−0.07*	0.04
Service sector	0.27***	0.04	0.26***	0.03	0.26***	0.03
Female wages in manufacturing	0.05	0.04	0.08*	0.05	0.05	0.04
Unemployment	−0.36***	0.07	−0.36***	0.07	−0.42***	0.08
Interaction: Mother ×						
Service sector	−0.02	0.04				
Female wage in manufacturing			−0.07**	0.03		
Unemployment					0.11	0.08
Variances						
Country-year						
σ^2 intercept	0.11		0.11		0.11	
σ^2 of mother	0.28		0.27		0.27	
R (σ^2 intercept, σ^2 of mother)	−0.85		−0.85		−0.85	
Country						
σ^2 intercept	0.18		0.18		0.18	
Deviance	230,153		230,147		230,151	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

for mothers by subtracting the interaction term (−.07) from the main effect of female wages (.08). The result, $.08 - .07 = .01$, is very close to 0, indicating that the employment of mothers was barely affected by high female wages in manufacturing. Therefore, in contradiction to our labour market structure hypothesis, we found that high female wages in manufacturing widened the employment gap between mothers and women without children.

To evaluate the robustness of our conclusions, we tested our models for the presence of overly influential data. The tools used for detecting influential data in multilevel models are discussed in Appendix A. Some countries in our sample were observed for only a limited number of years. This might have biased our results. After omitting the four countries with the fewest observations (Canada, Austria, Sweden, and Norway), our findings remained virtually unaltered. Tools for detecting influential data in multilevel models (Nieuwenhuis, Te Grotenhuis, & Pelzer, 2012c) also established that omission of singular country-years did not result in substantially different conclusions regarding the main effects of the institutional determinants. The combined outcomes of these tests suggested that our conclusions are robust to changes in the selection of the OECD countries in our sample.

Evaluation of the missing person-level data using techniques for multiple data imputation (Van Buuren & Groothuis-Oudshoorn, 2011) showed that neither the estimates of demographic variables nor the estimates of the cross-level interactions were affected by the presence of missing data in such a way that our general conclusions would change.

Finally, we calculated the predicted probability of employment of both mothers and women without children in different institutional contexts, using the estimates presented in Tables 2.4, 2.5, and 2.6. This provided a more informative interpretation of the outcomes in terms of how strongly various contexts influenced the employment of mothers and women without children. For each aspect of context, we calculated the predicted probability of employment of mothers and women without children, for both the lowest and the highest measured value of that specific aspect of institutional context. All person-level variables other than motherhood were kept constant at values indicating

a woman living in a shared household, born in 1950, and having completed her education at age 15. All institutional variables were controlled for at their respective average. The outcomes are presented in Figure 2.2.

The motherhood-employment gap was substantially smaller in countries with extended maternity leave, pay during leave, and child care leave. This is illustrated in the figures on reconciliation policies (see first row of Figure 2.2). Pay during leave has the weakest effect, but in countries with very long periods of maternity leave employment among mothers was estimated to be on a par with that of women without children.

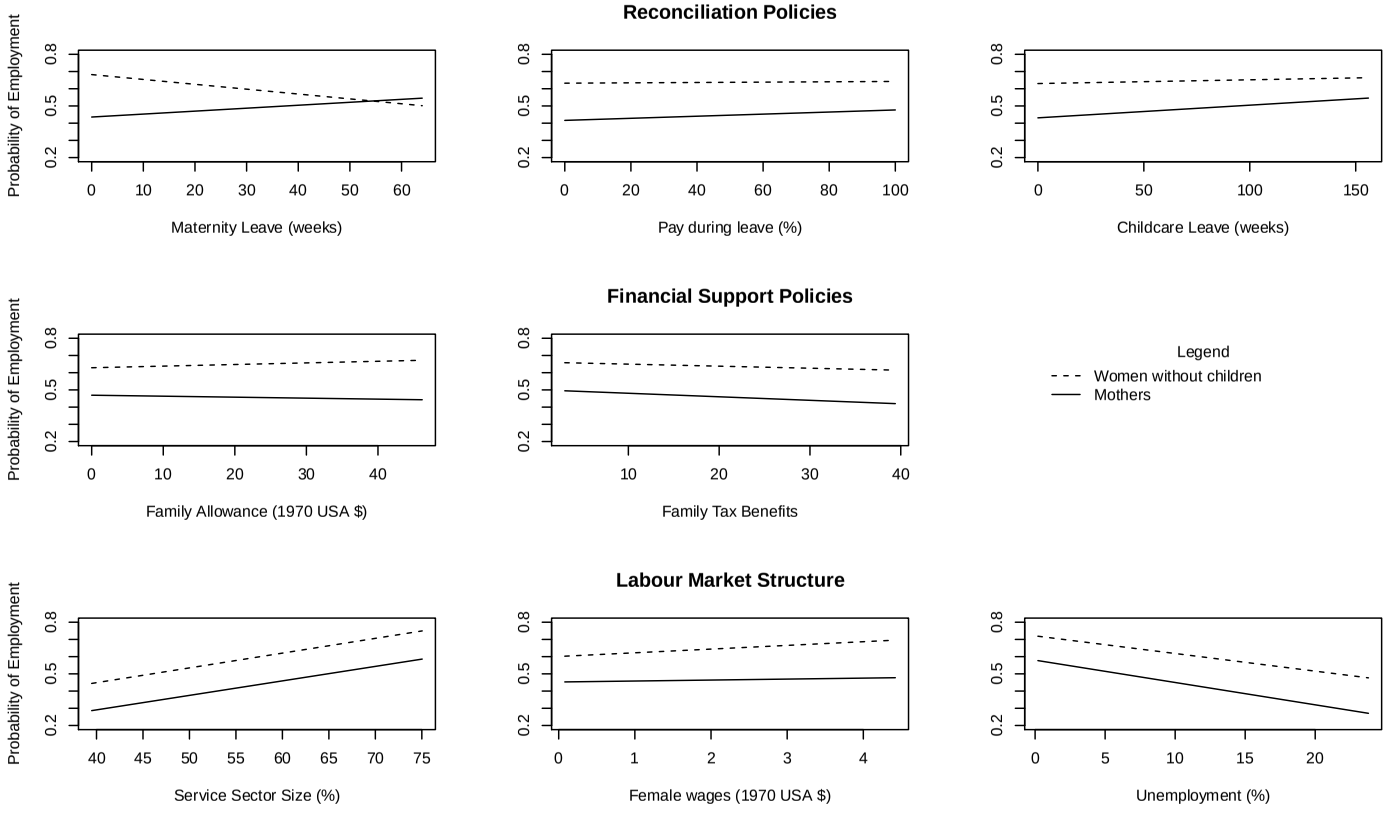
Financial support policies had less impact on women's employment than reconciliation policies. However, we found that extensive family allowances widened the employment gap between mothers and women without children.

Regarding labour market structure, the influence of service sector size and unemployment levels on the employment of both mothers and women without children was stronger than that of female wages in manufacturing. Moreover, the impact of several family policies was outweighed by service sector size and unemployment levels, again for both mothers and for women without children. Finally, mothers were barely affected by high female wages in manufacturing, although high wages did increase the employment of women without children.

2.5 CONCLUSION AND DISCUSSION

This Chapter combined institutional and demographic explanations enabling us to explore the extent to which their influence on the employment of mothers was different to their influence on the employment of women without children. Our data covered 18 OECD countries and the 25 years between 1975 and 1999. Our analyses showed that women's employment was associated with both demographic and institutional determinants. It also revealed an interplay between the effects of institutional context and motherhood.

Figure 2.2: Predicted Probability of Employment of Mothers and Women Without Children in Various Institutional Contexts. Source: Authors' calculations based on multilevel model results presented in Tables 2.4, 2.5, and 2.6.



We found substantial variation in the motherhood-employment gap across countries and within countries over time. This gap was smallest in Denmark and showed no significant trend there. This is consistent with Denmark's consistently high level of pay during leave and long duration of maternity leave. The increasing duration of maternity leave in Norway, higher levels of pay during leave in Ireland, and the implementation of child care leave in combination with high levels of pay during leave in The Netherlands are all consistent with the decreasing gap in employment between mothers and women without children that we observed in these countries. Mothers became less likely to be employed compared with women without children in West Germany. Germany has often been recognised as supporting the traditional breadwinner model (e.g. Cooke, 2011), which in our model was reflected in increasing levels of family allowances, particularly after 1990.

In the case of West Germany women became less likely to combine motherhood and employment, while during the same period the employment of all women in the country rose markedly. The same applies to Portugal. A demographic explanation of this finding is that fertility decreased in these countries. An institutional explanation refers to the increasing size of the service sector. As shown in Figure 2.2, a large service sector is associated with a high level of employment among women, and this influence is on a par with that of several family policies. However, a large service sector was not found to help women who combined motherhood and employment. Therefore, our findings suggest that as the growing size of the West German service sector stimulated the employment of mothers and women without children equally it did not counter the growing employment gap between mothers and women without children that resulted from increasing levels of family allowance. More generally, our findings thus suggest to policymakers that alleviating the difficulties involved in combining motherhood and employment is not the only thing that stimulates women's employment; labour market structure appears to be just as important.

Our findings reveal several advantages of integrating demographic and institutional explanations of women's employment. Firstly, using

the Comparative Motherhood-Employment Gap Trend File, we tested the effects of eight characteristics of institutional context on women's employment, controlling for demographic factors. The data enabled us to differentiate the institutional context's effects on mothers from the effects on women without children. This led to a stricter test and thus to a more detailed understanding of how institutional contexts influence women's decisions regarding employment. The cross-level interactions in Tables 2.4, 2.5, and 2.6 are more informative about the impact of family policies and labour market structure, compared with the findings presented in Table 2.3, which made no distinction between mothers and women without children. By making this distinction, we were able to show that most family policies consistently influence the employment of mothers but do not influence the employment of women without children. Future studies might take this approach a step further, using multi-actor models that include the characteristics of partners within the household. Such research could advance our understanding of the effects of institutional context on women's and men's employment.

Secondly, the literature on the association between women's employment and fertility often observes that the cross-country correlation between female labour force participation and total fertility rates turned positive after about 1985 (Ahn & Mira, 2002), as was discussed in Chapter 1. It was also suggested that the person-level association between women's employment and number of children became less negative over time (e.g. Kögel, 2004). Nevertheless, because most studies have used only country-level data, this had not actually been tested. To this literature we add the finding that in many countries, over time, women became more likely to combine motherhood and employment, but that this was not the case in all countries. Moreover, in our models motherhood and employment were combined more frequently in societies with low levels of women's employment. The statistical resolution of this enigmatic finding is that in our multivariate models the countries in which women were less likely to be employed, and in which women were more likely to combine having children with employment, were also those in which women were more likely to be a mother. Examples

of this finding include Greece, Portugal, and Italy, where we observed relatively small employment gaps between mothers and women without children and relatively low women's employment. During the first decades we covered in our study, these countries also had high fertility rates. We were not able to make causal inferences regarding the person-level association between motherhood and employment. This finding that the employment gap between mothers and women without children is smaller in countries with high fertility has three possible interpretations. The first interpretation is that if motherhood and employment are difficult to combine, women limit their motherhood. Indeed, a recent study showed that women postponed their fertility in the absence of supportive family policies (Mills, Rindfuss, McDonald, & Te Velde, 2011). A second possible interpretation is that countries with low fertility rates are less likely to enact reconciliation policies because the relatively small number of mothers has a limited political voice. A third interpretation is that policies aimed at increasing fertility rates, such as family allowances, do increase fertility but also reduce the likelihood of women combining motherhood and employment. We found evidence of this in the current chapter. All three interpretations are supported by our empirical findings, but they make different assumptions regarding the causal order underlying the motherhood – employment association. In addition, the second interpretation regards the institutional context as an outcome, rather than a cause, of women's behavioural decisions (cf. Linos, 2011). The use of country-comparative event-history analyses would allow future studies to disentangle the direction of causality of the person-level motherhood – employment association. This, however, is beyond the scope of this dissertation. This technique would also allow for a more detailed modelling of the timing of events such as childbirth and employment entry or exit and therefore provide a stricter test of whether changes in institutional context are indeed followed (rather than preceded) by changes in decisions regarding motherhood and employment.

Finally, as noted earlier, institutional contexts explained a considerable amount of the variation in women's employment between countries, although the demographic indicators were better than the institu-

tional ones in explaining the employment of individual women. Pettit and Hook (2005) reported a similar finding. Although the institutional context shapes the general pattern of how women organise their employment in a society, personal decisions associated with a woman's individual demographic background still play a key role. The current chapter emphasised the opportunities that institutional contexts provide for women to be employed and to combine motherhood and employment. Future studies could test the argument that opportunities do not lead to action on their own and nor do interests; rather it is the interaction of opportunities and interests that spurs behaviour. Just as demographic outcomes, such as those associated with having children, were seen to vary by institutional context (cf. Cooke & Baxter, 2010), institutional outcomes can vary by demographic strata. Similarly, Fuwa (2004) reported that, in regard to understanding household dynamics, women's individual assets were more important in countries that provide a gender egalitarian context than in countries without this context. To better understand the influence of institutional arrangements on women's employment, in Chapter 4 we focus on the interplay between opportunities and interests in determining women's employment and examine how this process can lead to family policies having socially stratified outcomes. We examine this by studying the extent to which the outcomes of family policies are stratified by women's level of education.

To conclude this Chapter, we found that mothers and women without children were more likely to be employed in societies with a large service sector and low unemployment. Mothers were more likely to be employed in societies with extensive reconciliation policies and limited family allowances. The employment of women without children was unaffected by family policies. Most importantly, by integrating person-level and country-level explanations of women's employment, this study constitutes an important step in examining the interplay between institutional and demographic factors in influencing the combination of work and family life. In this Chapter, we assumed that the effect of childcare leave on the employment of mothers was linear, whereas several authors have suggested that very long periods

of leave negatively affect women's employment (Bruning & Plantenga, 1999; Gornick & Meyers, 2003; Morgan & Zippel, 2003; Moss & Deven, 1999; Waldfogel, 2001). In the following chapter, we challenge the assumption of a linear effect of childcare leave, and answer the question whether there is such a thing as too long childcare leave.

3

IS THERE SUCH A THING AS TOO LONG CHILDCARE LEAVE?

ABSTRACT

In Chapter 2 we found that longer childcare leave facilitates women's employment by reducing the size of the motherhood-employment gap. In this Chapter we follow up on this finding and test whether women's employment is facilitated in societies with short-term childcare leave but negatively affected in societies with very long periods of childcare leave. We start by stating that this 'long-leave question' has not yet been satisfactorily answered. We argued that to correctly answer the long-leave question (1.) the relationship between duration of leave and employment of women should be explicitly hypothesised as being curvilinear and (2.) childcare leave should be expected to affect only mothers, not women without children. Based on this we formulated the long-leave hypothesis: In countries with short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, but in countries with long periods of childcare leave the motherhood-employment gap is larger than in countries with short periods of leave. In addition, we argued that to test the long-leave hypothesis one should use data in which countries are observed repeatedly over time, and one should evaluate for the presence of influential data. This can be done using the 'Comparative Motherhood-Employment Gap Trend File' on 192,484 individual women, 305 country-years, and 18 countries, combined with country-level data from the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001). We found that in countries with

short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, while in countries with long periods of childcare leave the motherhood-employment gap is bigger than in countries with short periods of leave.

3.1 BACKGROUND AND RESEARCH QUESTION

Childcare leave facilitates women in combining motherhood and employment in the period following childbirth. Countries vary substantially in the duration of childcare leave mothers are entitled to, from countries such as the United Kingdom and Ireland that have had very brief to no childcare leave entitlements in recent decades, to very long childcare leave arrangements in for instance France, Germany, and Sweden. Childcare leave policies either guarantee the possibility of returning to employment after a period of full-time leave, or provide the opportunity for part-time leave and part-time employment. Childcare leave thus provides women with the opportunity to continue their attachment to the labour market after childbirth (Pettit & Hook, 2005). Consequently, we found in Chapter 2 that women's employment is positively associated with the availability of childcare leave, because mothers' employment is facilitated by leave, a finding reported by several other scholars (Akgunduz & Plantenga, 2013; Gornick, Meyers, & Ross, 1999; Van der Lippe & Van Dijk, 2002).

However, despite findings that childcare leave facilitates the employment of mothers, various scholars have reported that very long periods of childcare leave actually reduce women's employment (Bruning & Plantenga, 1999; Gornick & Meyers, 2003; Morgan & Zippel, 2003; Moss & Deven, 1999; Pettit & Hook, 2009; Waldfogel, 2001). Extensive periods of leave have been argued to be associated with the traditional breadwinner model, with women combining motherhood and employment "*sequentially rather than simultaneously*" (Pettit & Hook, 2005, p.784) and thus leaving the labour market for extended periods of time after childbirth. In addition, very long periods of leave result in women becoming detached from the labour market by losing expe-

rience (human capital depreciation) (Gorlich & De Grip, 2008; Gornick & Meyers, 2003; OECD, 2001, 2011; Waldfogel, 2001). These findings on the effect of childcare leave contradict the findings we reported in Chapter 2. Hence, we reiterate the 'long-leave question':

QUESTION 3 To what extent was the motherhood-employment gap larger between 1975 and 1999 in OECD countries providing long-term childcare leave than in countries providing short-term leave?

In this Chapter, we resolve the contradiction between our own findings in our previous Chapter 2 regarding the question whether long-term childcare leave reduces women's employment and those reported in other studies (Pettit & Hook, 2009). Firstly, we argue that to correctly answer the long-leave question the relationship between duration of leave and employment of women should be explicitly hypothesised as being curvilinear. We build upon Chapter 2, where we assumed a linear effect of childcare leave, by testing the curvilinearity of this effect here. Secondly, we argue that childcare leave should be hypothesised to affect only mothers, not women without children. Here, as in Chapter 2, we do this by estimating the interaction between leave duration and the odds ratio between motherhood and employment.

3.2 THEORY AND HYPOTHESIS

3.2.1 The relationship between duration of leave and employment of women may be curvilinear

In Chapter 2 we argued that the effect of childcare leave on women's employment was based on the opportunities leave provides for women to combine motherhood and employment. Regarding childcare leave we hypothesised that, as they provide more opportunities, longer durations of leave would be associated with more employment among mothers and therefore with a smaller gap in employment between mothers and women without children. Pettit and Hook (2009), on the

other hand, argued that leave is a mechanism of exclusion for mothers on the labour market. Based on this argument they hypothesised that longer durations of leave decrease the employment of mothers.

It is also possible that short-term leave has a positive effect on mothers' employment while long-term has a negative effect (Akgunduz & Plantenga, 2013; Del Boca et al., 2009; Jaumotte, 2003). If that is the case, hypotheses which consider only the positive effect of leave on women's employment disregard possible negative effects, and hypotheses which consider only the negative effect of leave disregard possible positive effects. Thus, hypotheses should explicitly differentiate between the effects of short-term and long-term childcare leave.

3.2.2 Childcare leave is expected to affect only mothers, not women without children

In Chapter 2 we hypothesised that childcare leave policies only provide opportunities to mothers and not to women without children. We found that the employment of mothers was higher with longer periods of childcare leave, but not the employment of women without children. This differentiates leave policies from institutional factors - such as a large service sector size and low unemployment - which were found in Chapter 2 to stimulate the employment of both mothers and women without children. However, not all studies on the effect of childcare leave explicitly differentiate between the effect of leave on mothers and on women without children. This is the case in studies based on country-level data on women's employment using the aggregated female labour force participation rate of a country as the dependent variable, but also in several studies using person-level data.

A recent OECD (2011) report using country-level data showed that the employment of mothers with children under three was lower in countries with long leave, compared to countries with short leave. This finding supports the idea that long periods of leave detach mothers from the labour market, but it would have been more informative to hypothesise that long periods of leave increase the difference in employment between mothers and women without children. Akgunduz

and Plantenga (2013) found a curvilinear effect of leave duration on women's employment using country-level data. Although their study made a distinction between the effects of short-term and long-term leave, their analysis of country-level data could not differentiate between the employment of women with children and women without children. Jaumotte (2003) presented analyses of both country-level data and of person-level data. The analysis on country-level data showed a weakly positive (correlation of .05) linear association between the duration of paid leave and the (country-level) female labour force participation rate in a country. In multivariate regression models based on person-level data (Jaumotte, 2003), the curvilinear effect of leave was explicitly modelled and it was found that short periods of leave increased total female labour force participation while very long periods of leave were associated with lower levels of participation. Neither of these analyses made a distinction between the effect of leave on mothers and on women without children. Similarly, Del Boca et al. (2009) found a curvilinear effect of the duration of leave on the likelihood of a woman's employment, while controlling for the presence of children in the household. However, as it only applied a statistical control for the presence of children in the household, their analysis did not explicitly test whether the duration of leave only affected the employment of mothers.

Based on the arguments (1.) that the relationship between duration of leave and employment of women may be curvilinear and (2.) that childcare leave is expected to affect only mothers, not women without children, the long-leave hypothesis is:

LONG–LEAVE HYPOTHESIS In countries with short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, but in countries with long periods of childcare leave the motherhood-employment gap is larger than in countries with short periods of leave.

3.3 DATA AND METHOD

3.3.1 Testing the long-leave hypothesis requires a large number of cases

To test our long-leave hypothesis, both the curvilinear effect of leave and the interaction between this curvilinear effect of leave and motherhood need to be accounted for simultaneously, which increases the complexity of the models. Despite the complexity of estimating a curvilinear effect of leave simultaneously with an interaction between leave and motherhood, existing studies of the curvilinear effect of childcare leave on mothers' employment are based on a limited number of country-level observations. This is a natural data-limitation common to many country-comparative studies (Van der Meer, Te Grotenhuis, & Pelzer, 2010), as discussed in Chapter 1. However, the number of available degrees of freedom at the country-level is small, which increases the risk of over-fitting the model (Harrell Jr., 2001). Over-fitting a model increases the likelihood of influential cases and overly large standard errors.

Two examples show the limitations in the number of countries available for analyses of the effects of long childcare leave. Budig, Misra, & Boeckmann (2012) presented a country-comparative study on the curvilinear effect of leave on an aspect of women's employment: the motherhood-wage penalty. The analyses were based on 22 countries, but Budig et al. (2012) specified a complex model including three-way cross-level interactions (between motherhood, leave and leave squared, and a cultural variable), which accounted for 10 degrees of freedom, in addition to four country-level controls. Similarly, Pettit and Hook (2005) simultaneously estimated the interaction between seven country-level variables (including both the linear and the squared duration of parental leave) with the presence of young children in the household, based on observations from 19 countries. With so many country-level parameters to estimate compared to the number of country-level observations, the model is easily over-specified, which increases the risk of influential cases and overly large standard errors.

3.3.2 Person-Level Data

We test the long-leave hypothesis using person-level observations obtained from the '*Comparative Motherhood-Employment Gap Trend File*'. This dataset was introduced and used in Chapter 2, and merged from the Mannheim Eurobarometer Trend File (Schmitt & Scholz, 2005), the General Social Survey (Smith et al., 2010) and the Canadian Election Study (see http://ces-eec.org/). All surveys provided samples that, by using sampling weights, were representative of the respective country populations. The wording of the questions in the three surveys provided comparable measures for several demographic background characteristics of individual women. A detailed description of this dataset is given in Chapter 2 and in Appendix C. The Comparative Motherhood-Employment Gap Trend File dataset is based on 192,484 person-level observations, covering 305 country-years and 18 OECD countries.

Two person-level variables were used:

EMPLOYMENT a binary dependent variable representing whether or not a woman was (coded 1) or was not (coded 0) involved in paid employment at the time of the survey.

MOTHERHOOD a binary variable representing whether a woman was a mother (coded 1) or not (coded 0).

The size of the motherhood-employment gap was defined as the effect of motherhood on employment, and was indicated by estimating the odds-ratio (as was done in Chapter 2).

3.3.3 Country-Level Data

We combined the person-level data with one country-level variable on duration of leave. As we are interested here in the total effect of leave on the size of the motherhood-employment gap, we did not control for additional (country-level) variables. Our variable indicating leave is:

Table 3.1: Descriptive Statistics of Motherhood, Employment, and Childcare Leave
Source: Comparative Motherhood-Employment Gap Trend File, and the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001).

	M	SD	Min	Max
Employment status	0.54		0	1
Mother	0.61		0	1
Child care leave duration	0.00	5.55	−4.44	11.2
N.countries	18			
N.countryyears	305			
N.obs	192,484			

CHILDCARE LEAVE This country-level independent variable indicates the duration of childcare leave (in weeks) in a country. For easier interpretation, the duration of childcare leave in weeks was divided by 10 and centred in the full dataset on 0. This variable was measured at the country-level and was obtained from the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001). The measurement of childcare leave duration is time-varying.

Descriptive statistics on both the cross-sectional and the longitudinal datasets are presented in Table 3.1.

3.3.4 Statistical Method

The data were analysed using multilevel logistic regression, in which employment was regressed on motherhood, childcare leave duration, and the cross-level interaction between leave and motherhood. A curvilinear effect of childcare leave duration was estimated. Person-level observations nested in country-years and countries. The regression models were estimated using the lme4 package (Bates et al., 2010) in R (R Core Team, 2012).

We argued earlier that testing the long-leave hypothesis requires a large number of observations. This also reduces the risk of single influential cases biasing the conclusions drawn from the regression models. With a small number of country-level observations it is generally advised to test for influential cases at the country-level (Van der Meer et al., 2010). The importance of evaluating for the presence of influential cases is further emphasised by the complexity of the models discussed above, as the "*most common reason [for a case being influential] is having too few observations for the complexity of the model being fitted*" (Harrell Jr., 2001, p. 74). The fact that many multilevel models are based on large numbers of person-level observations does not help here, as the accuracy of country-level estimates and cross-level interactions in multilevel modelling are improved mainly by covering a large number of countries rather than by a large number of person-level observations (Hox, 2010).

In our analyses we used a large number of country-year level observations, and needed to evaluate whether this large number of country-year level observations resulted in stable estimates. We applied software for detecting influential data in multilevel models, for which statistical tools have become available (Nieuwenhuis et al., 2012c; Van der Meer et al., 2010). These tools are detailed in Appendix A.[1]

3.4 RESULTS

In Model I of Table 3.2, women's employment was regressed on the interaction between motherhood and both the linear and the curvilinear effect of leave. The results indicate that mothers were more likely to be employed in countries with longer periods of leave (compared to countries with no leave at all), while very long durations of leave tend to

1 For illustrative purposes, in Appendix A we also present a test of the long-leave hypothesis based on a cross-section of only 15 countries. This cross-section was obtained as a subset of the data used in this Chapter. Appendix A shows that the results of using cross-sectional data to test the long-leave hypothesis were indeed overly influenced by a single country.

Table 3.2: Women's Employment Regressed on the Interaction Between Motherhood and the (curvi-)linear effect of Childcare Leave, 1975-1999
Source: Comparative Motherhood-Employment Gap Trend File, and the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001).

	Model I		Model II	
	B	SE	B	SE
Fixed Effects				
Intercept	0.80***	0.157	0.77***	0.160
Mother	−0.62***	0.053	−0.63***	0.057
Childcare Leave	0.06***	0.008	0.06***	0.008
Leave Squared	0.00	0.001	0.00	0.001
Leave × Mother	0.07***	0.010	0.07***	0.010
Leave^2 × Mother	−0.01***	0.001	−0.01***	0.001
Random Effects (SD)				
Country				
Intercept	0.644		0.658	
Countryyear				
Intercept	0.304		0.297	
Mother	0.510		0.510	
N.countries	18		18	
N.countryyears	305		304	
N.obs	192,484		190,918	
-2×Loglikelihood	233,606		233,127	

* p<0.05, ** p<0.01, *** p<0.001 (one-tailed)

reduce the employment of mothers. This is in line with the long-leave hypothesis.

We evaluated the extent to which the conclusions drawn based on the longitudinal data were overly influenced by a single observation at the institutional level, in this case the country-year level. We tested whether the deletion of single country-years from the data would affect the fixed parameter estimates of model II in Table 3.2. To detect the level of influence each country-year has had on these parameter estimates, we calculated Cook's Distance for estimates of all fixed parameters. In Figure 3.1 the 25 country-years with the largest value of Cook's Distance are shown (sorted in descending order). The Cook's Distance of one country-year clearly stands out: East Germany in 1990.

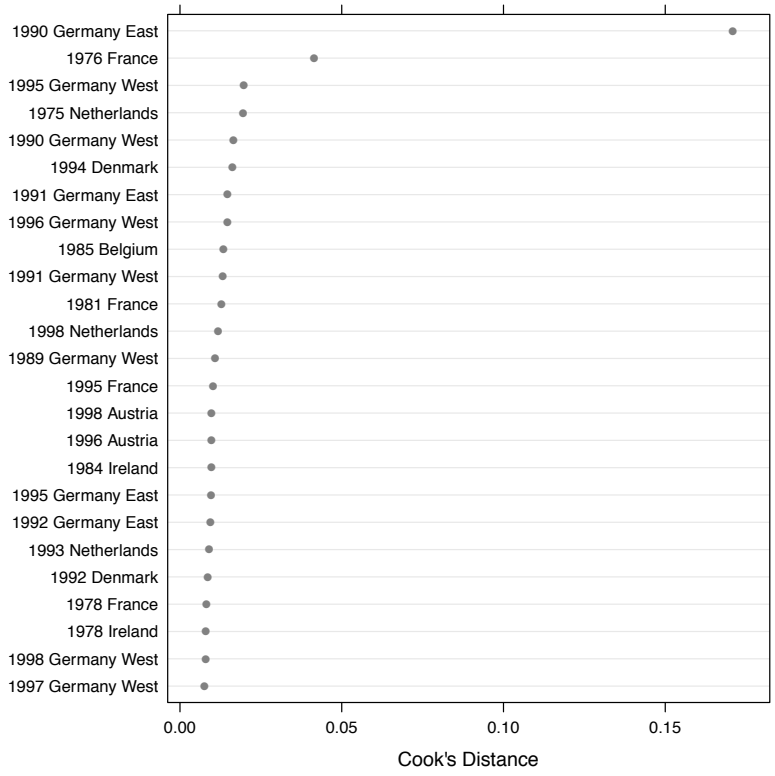
Model II of Table 3.2 shows the results of the same regression model as Model I of the same Table, after the observations of East Germany in 1990 were deleted from the data. All parameter estimates in Model II of Table 3.2 are similar to those originally reported in Model I. The conclusions regarding our hypothesis therefore remain the same, despite the deletion of a single country-year that was much more influential than all others.

3.5 CONCLUSION AND DISCUSSION

In this Chapter we confronted our Chapter 2 finding that long periods of childcare leave improve women's employment rates with Pettit and Hook's (2009) conclusion that long periods of leave exclude women from the labour market. Since these findings were contradictory, we considered the extent to which motherhood negatively affects employment more strongly in societies providing long-term childcare leave than in societies providing short-term leave.

We argued that to correctly answer the long-leave question (1.) the relationship between duration of leave and employment of women should be explicitly hypothesised as being curvilinear and (2.) childcare leave should be expected to affect only mothers, not women without children. Based on this we formulated the long-leave hypothesis

Figure 3.1: Influential Data Analysis Testing the Long-Leave Hypothesis (based on the Interaction Model in Table 3.2).
Cook's Distances of the 25 Country-Years With the Largest Influence on the Regression Parameters.
Source: Authors' calculations on interaction model in table 3.2



as: In countries with short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, but in countries with long periods of childcare leave the motherhood-employment gap is larger than in countries with short periods of leave.

The 'long-leave' hypothesis was tested against the data from the Comparative Motherhood-Employment Gap Trend File with 192,484 person-level observations, covering 305 country-years from 1975 to 1999 and 18 countries. The results corroborated both the hypothesised curvilinearity of the effect of childcare leave on women's employment, and the interaction between this curvilinear effect and motherhood. The model results were robust against the presence of influential cases: the deletion of the country-year with the largest statistical influence, East Germany in 1990, did not affect the conclusions regarding the long-leave hypothesis.

It is not difficult to present a post-hoc interpretation of the finding that East Germany in 1990 was an influential case: Rosenfeld, Trappe, and Gornick (2004) have documented how the regime change which began in East Germany in 1989 affected women's employment. Prior to reunification, East Germany had high rates of women's employment compared to other countries, particularly for mothers. After reunification, the duration of (childcare) leave was on a par with that of West Germany. The degree to which women combined motherhood and employment declined in East Germany (as we saw in Figure 2.1), but not as quickly as leave policy changed. This explains why we observed both very long leave and a high rate of women combining motherhood and employment in 1990.

The analyses we presented were not without limitations. Our measure of motherhood failed to distinguish between women with young children and women with older children. Additionally, our measurement of employment only distinguished between women currently employed, and those not in employment. Nevertheless, although using more detailed measures may detect more nuanced results the measures available were detailed enough to accurately test our long-leave hypothesis.

The finding that short-term leave reduces the gap in employment between mothers and women without children, while long-term leave increases it, has theoretical relevance. Pettit and Hook (2009) proposed a theory about how social contexts, including leave policies, affect various aspects of women's employment. In this theory, a distinction is made between mechanisms of inclusion (such as maternity leave, part-time employment, and child care) and mechanisms of exclusion (such as union density and long childcare leave). We contrasted their theorising with our theory on the opportunities provided by childcare leave, as formulated in Chapter 2. Our analysis suggested that leave policies cannot simply be classified as a mechanism of either inclusion or exclusion. Instead, the opportunities provided by short-term leave include keeping mothers in the labour market, while the opportunities of long-term leave have the (unintended) consequence of excluding them from the labour market.

To conclude, in this Chapter we answered the question of to what extent short-term childcare leave stimulates women's employment, while long-term childcare leave reduces their employment. We found support for our long-leave hypothesis that in countries with short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, while in countries with long periods of childcare leave the motherhood-employment gap is bigger than in countries with short periods of leave. Thus, in this Chapter we improved upon the analysis of the effect of childcare leave presented in Chapter 2. In Chapter 4, we challenge our assumption that family policies (both reconciliation policies and financial support policies for families with children) affect all mothers in the population equally. We will answer the question of to what extent these family policies reduce or increase the motherhood-employment gap differently among more and less educated women.

4

STRATIFIED OUTCOMES OF FAMILY POLICIES ON WOMEN'S EMPLOYMENT

ABSTRACT

In Chapter 2 we found that the size of the motherhood-employment gap was reduced by reconciliation policies and increased by financial support policies for families. In this Chapter, we answered the question of to what extent the outcomes of reconciliation policies and financial support policies differ between more and less educated mothers. Thus, we challenged the theoretical assumption held in Chapters 2 and 3 that all mothers are equally affected by family policies. Our data were obtained from the 'Comparative Motherhood-Employment Gap Trend File', combined with country-level data from the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001). The data covered 17 OECD countries, the period from 1980 to 1999, and 116,874 observations on individual women living in partnered households. The data were analysed using multilevel logistic regression. We found that paid leave - a reconciliation policy - reduces the motherhood-employment gap more strongly among more educated women than among the less educated.

1 This research was supported by the European Science Foundation (ESF), as part of the 'Quantitative Methods in the Social Sciences II' programme.

4.1 BACKGROUND AND RESEARCH QUESTION

In Chapter 2 we introduced a key explanation of the increased female labour force participation rates in industrialised societies in recent decades, reading that the introduction of family policies facilitated women in combining the responsibilities of motherhood and employment. These policies were important in explaining cross-national differences and trends in women's employment, as in country-comparative studies motherhood was found to be the most important demographic factor associated with women's employment (Van der Lippe & Van Dijk, 2002), with mothers less likely to be employed than women without children. The causal mechanism that underlies this motherhood-employment gap was found to be two-sided; mothers were more likely to opt out of employment, and employed women were less likely to have children (Bernhardt, 1993; Cramer, 1980; Waite & Stolzenberg, 1976).

The motherhood-employment gap was found to vary between countries and within countries over time, and to be affected by the policy context (Matysiak & Vignoli, 2008; Nieuwenhuis et al., 2012a; Pettit & Hook, 2005). Two types of family policies were be distinguished between in previous Chapters: reconciliation policies and financial support policies for families with children (Gauthier, 1996). The outcomes of these family policies on women's employment has received substantial amounts of attention, both in academic studies (Del Boca et al., 2009; Gornick et al., 1998; Hook, 2010; Jaumotte, 2003; Van der Lippe, De Ruijter, De Ruijter, & Raub, 2011) and from policy makers (OECD, 2011). The goal of reconciliation policies is to increase women's labour force participation by providing opportunities for combining motherhood and employment. Examples include maternity- and parental leave, pay during leave, and childcare arrangements. Policies financially supporting families, such as family allowances and tax benefits to families, were designed to reduce childhood poverty. These financial support policies also provide (financial) opportunities for mothers to stay at home, and have often been argued to be associated with a traditional male breadwinner model of the distribution of household

tasks. Financial support policies for families were found to increase the motherhood-employment gap by reducing the employment of mothers (Gauthier, 1996; Nieuwenhuis et al., 2012a; Schwarz, 2012; Thévenon & Luci, 2012). In Chapter 2 we found this motherhood-employment gap to be smaller in societies providing extensive reconciliation policies reduced, and to be bigger in societies with financial support policies for families with children.

The studies described above answered the question of to what extent family policies increased the employment of *all* mothers, without distinguishing between women with different educational backgrounds. Hence, in these studies it was assumed that all mothers would respond in the same way to the opportunities provided by reconciliation policies that assist them in combining motherhood with, and to the opportunities provided by financial support policies not to be employed. This was also the case in our Chapters 2 and 3. In the current Chapter we challenge this assumption, arguing that merely the opportunities to combine motherhood with employment provided by reconciliation policies do not have an impact on those mothers who do not have an interest in being employed. Conversely, the opportunities for a mother not to be employed provided by financial support policies will not have an impact on those mothers with a strong interest in employment.

More educated women tend to have a stronger interest in employment than less educated women (cf. Becker, 1991; Cunningham, 2008; Goldin, 2006; Pettit & Hook, 2005), for various reasons that are outlined in detail in the next section. However, many large-scale, cross-national comparative studies on women's employment have examined how the opportunities provided by family policies affect the employment of all women, or all mothers, alike (e.g. see Gornick et al., 1998; Jaumotte, 2003; Matysiak & Vignoli, 2008; Nieuwenhuis et al., 2012a). These studies failed to differentiate between the outcomes of reconciliation policies and financial support policies among more and less educated women.

The progress we made in Chapter 2 was to analyse the gap in employment between mothers and women without children and test how this motherhood employment gap was affected by reconciliation poli-

cies and financial support policies. In Chapter 3 we found that short durations of childcare leave reduced the degree to which mothers were less likely to be employed than women without children, but that very long durations of childcare leave actually increased the size of this motherhood-employment gap. In both Chapters, however, we assumed that family policies had the same effect on the motherhood-employment gap among all women. In this Chapter we examine to what extent the outcomes of family policies differ according to by women's educational level. We answer the following question:

QUESTION 4 To what extent did the outcomes of reconciliation policies and financial support policies on the size of the motherhood-employment gap differ between more and less educated women in OECD countries between 1980 and 1999?

In this Chapter we contribute to the literature that investigates the outcomes of family policy on women's employment, by responding to a call for furthering the integration of institutional and demographic explanations of women's employment (Van der Lippe & Van Dijk, 2002). We do so by differentiating the outcomes of family policies between more and less educated women. Pettit & Hook (2009) examined the extent to which more and less educated women responded differently to family policies, but did not differentiate between the outcomes of family policies on mothers, and on women without children. Analysing how family policies affect the size of the motherhood-employment gap, as we do in this study, results in a more informed understanding of how the outcomes of family policies are different among more and less educated women.

In the remainder of this Chapter we will first detail our theory on the outcomes of family policies for both more and less educated women. The hypotheses derived from this theory will be tested against data from the Comparative Motherhood-Employment Gap Trend File with observations on 116,874 women living in partnered households, from 17 OECD countries and covering the period 1980-1999.

4.2 THEORY AND HYPOTHESES

To answer our research question, we derived hypotheses from new home economics, a rational action theory based on that which we introduced in Chapter 1. In Chapter 2 we argued that family policies affect women's employment decisions by providing opportunities. Here, we introduce the concept of interests and then derive hypotheses on how the employment decisions of women are affected by the interplay between educational level, motherhood, and family policies. This theory is based on a distinction between women's interest in employment, and women's opportunity to pursue this interest. We formulate our theory in three steps.

In the first step, we define the two main theoretical concepts of interest and opportunity, and their interaction. Interests are the "*driving force*" of individual person's actions (Coleman, 1990, p. 509, also see: Spillman and Strand (2013). Women can have an interest in employment for various reasons. These are not limited to the (expected) earnings from employment, but also include fulfilment of a desire for having a career or beliefs about gender equality. Opportunities are most generally defined as the available options for action (Hedström, 2005). For a woman to have the opportunity to pursue employment, a job must be available, and she must have time available to spend on employment. In Chapter 2 we argued that motherhood also reduces women's opportunities for employment.

In Chapter 2 it was assumed that interests and opportunities independently determine the employment of women. Here, we introduce the notion that the outcomes of interests and opportunities depend on each other. A woman with an interest in employment can only be employed when she has the opportunity. A woman who has the opportunity to be employed will only take it up when she has an interest in being employed. The general hypothesis that mere opportunities to act do not have consequences without an interest in acting, or conversely that mere interests to act have no consequences without the opportunities to do to so, has been tested successfully in a wide variety of

situations (cf. De Graaf et al., 2000; Hedström, 2005; Ultee & Luijkx, 1998).

In the second step of formulating our theory, we use the concepts of interest and opportunity to derive a person-level hypothesis of how the interplay between educational level and motherhood determine women's employment. More educated women are assumed here to have a stronger interest in being employed than less educated women. New home economics theory (cf. Becker, 1991) regards a higher level of education as an investment in human capital, which on the labour market translates into higher earnings for the more educated. Next to the economic gains of a higher educational level, attitudes in favor of gender equality were also found to be positively associated with women's educational levels (Cunningham, 2008). More educated women marry later as a result of which their identities, including career decisions, were formed before marriage rather than during marriage. This results in a stronger interest in continuing to be employed during marriage (Goldin, 2006). Career aspirations have been found to be higher among women with higher levels of education (Coltrane, 2000; Del Boca et al., 2009, also see: Cooke and Baxter (2010)). Thus, in this study we assume that more educated women typically have a stronger interest in employment than less educated women, and that this assumption holds among women with children and among women without children.

Opportunities to be employed are limited for women with children, relative to the opportunities for employment of women without children. Reasons for this include the time-constraint that is associated with raising children, and the difficulties of reconciling the responsibilities of raising children with the responsibilities of employment. Of course, women with children are not completely without opportunity for employment, but they do have fewer opportunities, and are consequently expected to be less likely to be employed than women without children. Similarly, less educated women are not without interest in employment, but it can be assumed that less educated women have a weaker interest in employment than more educated women. This means that women with limited opportunities and limited interest are least likely to be employed, and women with extensive opportunities

and a strong interest in employment are most likely to be employed. Women with extensive opportunities but a weak interest in employment are less likely to be employed than women with both extensive opportunities and a strong interest in employment.

Interests will have a stronger effect when more opportunities are present, and correspondingly opportunities will affect those with more interest in employment more strongly. The reverse also holds: a lack of opportunities for employment will affect those with stronger interests most, compared to those who have neither the opportunities nor the interest in employment. The reason for this is that among the women with less interest in employment, employment was lower to begin with and thus relatively fewer women in that group will be affected by the reduction in opportunities. From this we derive the expectation that the effect of an increased level of education (increased interest) on the likelihood of a woman's employment is weaker for those who have a child (fewer opportunities). In other words, the motherhood-employment gap is expected to be *greater* among more educated women than among less educated women.

In the third, and final, step of formulating our theory we introduce how family policies affect less and more educated women differently based on the interplay between interests and opportunities. From the opportunity-based explanations it could be derived (e.g. in Chapter 2) why women with children are less likely to be employed than women without children, as well as why mothers are more likely to be employed when reconciliation policies are available and less likely when there is extensive financial support for families with children. These opportunity-based explanations do not explain why more and less educated mothers respond differently to family policies. Thus, the concepts of interests and opportunities need to be combined. Reconciliation policies, such as parental leave, increase mothers' opportunities to be employed, and do not affect the opportunities of women without children. Combining opportunity and interest based explanations, the opportunities provided by reconciliation policies are thus expected to be more effective in increasing the employment of more educated mothers, who have a stronger interest in employment than less edu-

cated mothers. Consequently, we expect reconciliation policies to reduce the motherhood-employment gap more strongly among higher educated women. We hypothesise:

RECONCILIATION POLICY HYPOTHESIS Reconciliation policies reduced the motherhood-employment gap more strongly among more educated women than among less educated women.

Financial support policies, such as family allowances, increase the opportunities for mothers not to be employed. We expect these opportunities to most strongly affect those mothers with a weaker interest in employment, and thus to increase the motherhood-employment gap more among less educated women. We hypothesise:

FINANCIAL SUPPORT POLICY HYPOTHESIS Financial support policies increased the motherhood-employment gap more strongly among less educated women than among more educated women.

Education, motherhood, and family policies are not the only factors determining women's interests in, and opportunities for, employment. Opportunities for employment are strongly dependent on the availability of jobs on the labour market, and with higher female wages women's interest in employment may rise. We therefore control for three aspects of the labour market structure that either affect interest in employment or opportunities for employment. Firstly, when female wages are high, women living in partnered households will have a stronger interest in employment because they can expect higher returns from that employment. Secondly, the growth of the service sector in all OECD countries is often argued to have increased the opportunities for employment of women (Del Boca & Locatelli, 2006). Thirdly, when unemployment rates in a country are high, the opportunities for individual women to find employment are reduced. Hence, we control for countries' unemployment rates.

4.3 DATA AND METHOD

4.3.1 Person-Level data

Our person-level observations were obtained from the '*Comparative Motherhood-Employment Gap Trend File*', that was also used in Chapters 2 and 3. This dataset was merged from the Mannheim Eurobarometer Trend File (Schmitt & Scholz, 2005), the General Social Survey (Smith et al., 2010) and the Canadian Election Study (see http://ces-eec.org/). All surveys provided samples that, with the use of sampling weights, were representative of the respective country populations. The wording of the questions in the three surveys provided comparable measures for several demographic background characteristics of individual women. A detailed description of this dataset is given in Chapter 2 and Appendix C.

Firstly, we selected all countries and years for which all variables, at both the person and country-level (detailed below), were available. The selection criterion for countries was that, per country, at least four years had to be covered. This resulted in a dataset of 361,729 individual-level observations. Secondly, to reduce the complexity of our analyses we limited ourselves to women living in a coupled household (married, living together), leaving us with 206,649 observations. Third, women younger than 20 years of age and those older than 50 were omitted from the data. Finally, from the remaining 138,329 observations, we deleted (list-wise) 21,455 (15,5%) observations with 1 or more missing values. Of the missing values, 3,669 originated from split ballot designs in the original surveys, and were therefore missing completely at random. 116,874 Observations were available for analysis, nested in 239 country-years and 17 countries. As we used different country-level variables in this Chapter, we cover 17 countries, rather than th 18 covered in Chapters 2 and 3, and cover a shorter period of time.

We used four person-level measurements:

EMPLOYMENT a binary dependent variable representing whether or not a woman was (coded 1) or was not (coded 0) involved in paid employment at the time of the survey.

MOTHERHOOD a binary variable representing whether or not a woman was a mother (coded 1) or not (coded 0).

EDUCATION an interval-level variable representing the age at which a woman completed or ended her education, ranging from 10 to 25 (indicating age 25 and up). We subtracted the overall mean value from this variable, for easier interpretation of the interaction models.

AGE an interval-level variable indicating the age of a woman at the time of interview, ranging from 20 to 50. It was centred so that the mean is 0, and divided by 10 for easier interpretation. We also account for age squared. Age and Age Squared are used as control variables to account for life-course effects.

4.3.2 Country-Level data

We combined the pooled person-level survey data with several country-level indicators of the institutional context. These contextual data indicate reconciliation policies, financial support policies for families, and controls for labour market structure. Descriptive statistics of both the person-level and country-level variables are presented in Table 4.1.

PAID LEAVE Our measure of reconciliation policies is an index of three leave policies: maternity leave, parental leave, and childcare leave. Each leave policy was measured as the number of weeks mothers were entitled to. The number of weeks of each of these leave policies were weighed by the percentage of wages that are paid during this leave. The final measure represents the total number of weeks of leave with full pay. These data were available from the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001).

Table 4.1: Descriptive Statistics of Demographic and Institutional Variables (N = 116,874 Individuals From 239 Country-Years From 17 Countries)
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

		(Centred) Measurements			
		M	SD	Min	Max
Individual level					
	Age	0	0.80	−1.58	1.42
	Employment status	0.50		0.00	1.00
Person-level Opportunities	Mother	0.69		0.00	1.00
Person-level Interests	Education (centred)	0	2.61	−7	8
Country-year level					
Country-level Opportunities	Parental Leave	20.49	17.81	0	78.40
	Family Allowance	1.27	0.67	0.13	2.90
	Service sector (÷10; centred)	0	0.80	−2.16	1.40
	Unemployment (÷10; centred)	0	0.35	−0.57	1.53
Person-level Interests	Male / Female wage ratio (centred)	0	0.06	−0.18	0.16

Table 4.2: Number of Observations, Countries, and Country-Years (N = 116,874 Individuals From 239 Country-Years From 17 Countries).
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

Countries	First Year	Last Year	N Country-years	N. Obs.
Austria	1995	1998	4	2127
Belgium	1980	1997	18	8812
Canada	1984	1998	4	2404
Denmark	1980	1998	19	8918
France	1980	1998	18	9938
Germany East	1990	1998	9	4487
Germany West	1980	1998	19	8221
Greece	1980	1997	18	8868
Ireland	1980	1998	19	8989
Italy	1980	1997	18	8040
Netherlands	1980	1998	19	11,978
Norway	1990	1996	6	1295
Portugal	1985	1998	14	6472
Spain	1985	1999	15	6289
Sweden	1995	1998	4	1742
United Kingdom	1980	1999	20	13,332
United States	1980	1998	15	4962
Total			239	116,874

FAMILY ALLOWANCE EXPENDITURE Our measure of financial support policies is the percentage of GDP a country spends on family allowances. These data were available from the Comparative Family Benefits Database (Gauthier, 2003)

Finally, we used three controls for the labour market structure.

SERVICE SECTOR SIZE Measured as the proportion of the total labour force employed in the service sector. Values were divided by 10 and centred at 0. These measurements were obtained from the OECD (OECD, 2013).

MALE / FEMALE WAGE RATIO Measured as the ratio between the average wages men and women could earn in manufacturing. These measurements were obtained from the Comparative Family Benefits Database (Gauthier & Bortnik, 2001), and centred at 0.

UNEMPLOYMENT Measured as the percentage of the civilian labour force that was unemployed. Values were divided by 10 and centred at 0. The measurements were obtained from the Comparative Family Benefits Database (Gauthier & Bortnik, 2001).

All country-level control variables were measured at the country-level and vary within countries over time - as do the family policy-variables.

4.3.3 Statistical Method: Three-way Interactions

We analysed the data using logistic multilevel regression models, estimated with the lme4-package in R (Bates et al., 2010). We used a three-level model, with individual observations nested within country-years and nested within countries. This nesting structure accounts for time-invariant unobserved country heterogeneity in women's employment rates at the country-level, as well as for possible bias resulting from the fact that some countries were observed for a long period of time, while some countries, such as Sweden and Austria, were observed for only a short period of time (cf. Table 4.2).

We tested our hypotheses with three-way, cross-level interactions between the effects of motherhood, educational level, and a policy variable. The complexity of this model is such that we specify the fully annotated model-formula in Page 103. This specification includes all constitutive terms of the three-way interaction (cf. Brambor, Clark, & Golder, 2005; Cox, 1984). Control variables were not specified in this equation, but will be included in the estimated models.

The equation, as well as the analyses in the next section, are restricted to modelling the interactions with a single type of policy at at time. The simultaneous estimation of two three-way interactions would be too complex. In total, the model (excluding the control terms) consists of 13 parameters: an intercept, three person-level variables, four variables measured at the level of the country-year (including the cross-level interactions), and five variance components. These are discussed below.

The intercept in the statistical model (indicated with (1.) in the equation) refers to the employment of women with an average level of education (since the education variable was centred), without children, and living in a country at a time with an average provision of (family) policy. The motherhood variable (2.) indicates the degree to which mothers are less likely to be employed compared to women without children, but only applies to women with an average level of education in an average-policy context. Similarly, the education variable (3.) indicates to what extent the employment of women without children and living in an average-policy context is associated with her educational level. The last person-level variable is the interaction between education and motherhood (4.), indicating the degree to which the effect of education on women's employment differs between women with children and women without children. The estimate of this interaction only applies to women living in a context with an average level of policy provision.

$$
\begin{aligned}
\text{logit}(P_{ijk}) = {} & \gamma_{000} && \text{Intercept referring to average educated childless women and average policy (1.)} \\
& + \gamma_{100}\text{Mother}_{ijk} && \text{Motherhood-employment gap, average educated, average policy (2.)} \\
& + \gamma_{200}\text{Education}_{ijk} && \text{Educational effect, childless women, average policy (3.)} \\
& + \gamma_{300}\text{Mother}_{ijk} * \text{Education}_{ijk} && \text{Degree to which the effect of education differs with motherhood, average policy (4.)} \\
& + \gamma_{010}\text{Policy}_{.jk} && \text{Policy effect, average educated, childless women (5.)} \\
& + \gamma_{020}\text{Policy}_{.jk} * \text{Mother}_{ijk} && \text{Policy effect, average educated mothers (6.)} \\
& + \gamma_{030}\text{Policy}_{.jk} * \text{Education}_{ijk} && \text{Increase in policy effect with higher education, for childless women (7.)} \\
& + \gamma_{040}\text{Policy}_{.jk} * \text{Mother}_{ijk} * \text{Education}_{ijk} && \text{Increase in policy effect with higher education, for mothers (8.)} \\
& + U_{0jk} && \text{Intercept variance, country-year level (9.)} \\
& + U_{1jk}\text{Mother}_{ijk} && \text{Slope variance of the effect of motherhood (10.)} \\
& + U_{2jk}\text{Education}_{ijk} && \text{Slope variance of the effect of education (11.)} \\
& + U_{3jk}\text{Mother}_{ijk} * \text{Education}_{ijk} && \text{Slope variance of the interaction between education and motherhood (12.)} \\
& + V_{00k}. && \text{Intercept variance, country level (13.)}
\end{aligned}
$$

The first parameter at the country-year level (5.) indicates the difference in employment associated with a one unit increase in the policy variable, among average educated women without children. The interaction between policy and motherhood (6.) represents to what extent the motherhood-employment gap is increased or decreased with each unit of change in the policy variable, for women with an average level of education. The interaction between policy and education (7.) indicates the degree to which the effect size of the education of women without children differs with each unit of change in the policy variable. This is an example of a parameter that has no substantive interest in this Chapter, as it is not used for testing our hypotheses, but is required for a proper estimation and interpretation of the other interaction terms (cf. Brambor et al., 2005). The parameter that is substantively the most important for testing our hypotheses is the three-way interaction between policy, motherhood, *and* education (8.). Parameter (6.) was defined as indicating the degree to which a policy increases or decreases the motherhood-employment gap for women with an average level of education, and parameter (8.) indicating the three-way indicates whether or not this policy-outcome is different for more or less educated women.

In addition to the fixed-effect parameters described above, the model also consists of five variance components (random effects). The intercept variance at the country-year level (9.) indicates the degree to which the employment of women varies between different countries and different years, controlled for the other parameters in the model. The slope variances (9.), (10.), and (11.) indicate the variation between country-year in the estimates of - respectively - motherhood, education, and the interaction between motherhood and education. The final variance component, at the country-level (12.), represents the degree to which countries differ in the likelihood that women are employed, controlled for the other parameters in the model. Incorporating this variation into the model accounts for time-invariant, unobserved heterogeneity at the country level.

Influential Data Analysis

We have evaluated our analyses for the presence of influential data, as we did in Chapters 2 and 3. We did so using the tools for the detecting of influential data in multilevel regression models that are described in Appendix A. These evaluations are not presented in this Chapter, since no influential cases were found that would change the reported findings in such a way that our conclusions did not hold.

4.4 RESULTS

4.4.1 Model Estimation

The multilevel models presented in Table 4.3 regress women's employment on demographic and institutional factors. The Null model is a random intercept model of the likelihood of employment for women living in a coupled household, which varies more between countries (intercept variance of .54) than within countries over time (intercept variance of .15).

In Model I, we estimated the effects of motherhood and education while controlling for age and age squared. The estimated effects are in line with common expectations. Mothers were found to be less likely to be employed than women without children. Older women were less likely to be employed than younger women, and the rate of decreasing employment with age accelerates at higher ages. More educated women were more likely to be employed than the less educated. These demographic factors account for some of the variation in the employment of women living in a coupled household between countries (variance dropped from .54 to .43) and within countries over time (variance dropped from .15 to .10). In Model II, institutional factors were added to Model I, including leave, family allowance, and controls for labour market structure. Long periods of paid leave were not found to be associated with the employment of women. However, in Model II this parameter applies to the employment of *all* women, rather than specifically the employment of mothers. High levels of expenditure on

family allowances were found to be associated with lower employment of women. The controls for labour market structure show that a large service sector and high female wages (relative to male wages) were associated with an increased likelihood that women were employed, and a high overall unemployment was found to be associated with a lower likelihood of women's employment. The inclusion of the contextual variables substantially decreased the variances of the random intercept within country (dropped from .43 to .22) and within country-years (dropped from .10 to .04).

Table 4.4 presents the outcomes of the complete three-way interactions between motherhood, education, and policy variables. The parameters are numbered corresponding with the model specification on page 103. The models presented in Table 4.4 were controlled for age, age squared, service sector size, unemployment rate and female/male wage ratio, but these control variables are not presented. The estimates of these control variables were virtually identical to those presented in Table 4.3.

In Model I (of Table 4.4), the individual-level interaction between motherhood and educational level was estimated. The negative estimate of -.04 indicates that the effect of educational level is weaker for mothers ($.19 - .04 = .15$) than for women without children (.19).

Model II presents the outcomes of paid leave policies. The intercept is 1.09, indicating that women with an average level of education, without children, and living in a country with an average provision of paid leave had a likelihood of $e^{-1.09}/(1 + e^{1.09}) = 75\%$ of being employed. Being a mother reduced this likelihood, while a higher level of education increased the likelihood that women without children were employed. In line with the findings presented in Model I, the interaction between motherhood and education is significant and negative. The difference is that this interaction in Model I referred to all women, whereas in Model II it only refers to women living in a country that has average provision of paid leave.

The interaction term between paid leave and the motherhood variable is not significant, indicating that the opportunities for employment provided by paid leave did not increase the employment of moth-

Table 4.3: Multilevel Model Results Predicting Women's Employment from the Interaction between Demographic and Institutional Factors, partnered households only (N = 116,874 Individuals From 239 Country-Years From 17 Countries).
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

		Model Null		Model I		Model II	
		B	SE	B	SE	B	SE
	Intercept	0.17	0.18	0.82***	0.16	1.06***	0.15
	Age (/10)			−0.10***	0.01	−0.10***	0.01
	Age (/10) ^{2}			−0.03***	0.00	−0.03***	0.00
Person-level Opportunities	Mother			−0.75***	0.02	−0.74***	0.02
Person-level Interests	Education			0.17***	0.00	0.17***	0.00
Country-level Opportunities	Leave (/10)					0.00	0.00
	Family Allowance					−0.26***	0.06
	Service Sector					0.41***	0.05
	Unemployment					−0.36***	0.08
Country-level Interests	Female / Male Wage ratio					2.37***	0.67
Variance Components:							
Countryyear	Intercept	0.15		0.10		0.04	
Country	Intercept	0.54		0.43		0.22	
	-2 * Loglikelihood	147,959		141,545		141,386	

* p<0.05, ** p<0.01, *** p<0.001

Table 4.4: Multilevel Model Results Predicting Women's Employment From the Interaction Between Motherhood, Education, and Family Policies, partnered households only (N = 116,874 Individuals From 239 Country-Years From 17 Countries).
Source: Comparative Motherhood-Employment Gap Trend File, Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001), Comparative Family Benefits Database (Gauthier, 2003), and OECD Statistics (OECD, 2013).

		Model I		Model II		Model III	
		B	SE	B	SE	B	SE
	Intercept (1.)	1.07***	0.16	1.09***	0.16	1.01***	0.14
Person-level Opportunities	Mother (2.)	−0.76***	0.03	−0.78***	0.05	−0.61***	0.06
Person-level Interests	Education (3.)	0.20***	0.01	0.22***	0.01	0.19***	0.01
Country-level Opportunities	Leave (/10) (5.)	0.02	0.02	0.01	0.02	0.04**	0.01
	Family Allowance (5.)	−0.26***	0.06	−0.26***	0.06	−0.24***	0.05
Opportunities and Interests	Mother * Education (4.)	−0.04***	0.01	−0.06***	0.01	−0.02	0.02
	Leave * Mother (6.)			0.01	0.02		
	Leave * Education (7.)			−0.01*	0.00		
	Leave * Mother * Education (8.)			0.01*	0.00		
	Family Allowance * Mother (6.)					−0.12**	0.04
	Family Allowance * Education (7.)					0.00	0.00
	Family Allowance * Mother * Education (8.)					−0.01	0.01
Variance Components:							
Countryyear	Intercept (9.)	0.13		0.13		0.00	
	Mother (10.)	0.21		0.21		0.17	
	Education (11.)	0.01		0.01		0.01	
	Mother * Education (12.)	0.01		0.00		0.00	
Country	Intercept (13.)	0.23		0.23		0.23	
	-2 * Loglikelihood	140,370		140,362		140,557	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

All models controlled for age, age squared, service sector size, unemployment, and female/male wage ratio

ers with an average level of education. The three-way-interaction term between motherhood, education, and policy is positive and statistically significant. The interpretation of this estimate is that among more educated women, the negative effect of motherhood on employment is reduced by longer periods of paid leave. Also, the higher the educational level of a mother, i.e. the higher level of interest in employment, the stronger the effect of the opportunities provided by paid leave. This is in line with our reconciliation policy hypothesis.

In Model III, higher levels of expenditure on family allowances were found to increase the motherhood-employment gap, as shown by the negative estimate of the interaction between family allowance and motherhood. This is in line with the financial support policy hypothesis formulated in Chapter 2. However, the three-way interaction term was not statistically significant, indicating that the effects of the opportunities provided by family allowances are not different among those with higher or lower levels of interest in employment, i.e. more or less educated women. The specific financial support policy hypothesis that was formulated here therefore needs to be rejected.

4.4.2 Predicted Employment

To allow for an easier interpretation of the complex model results presented above, we calculated the predicted likelihood of women's employment (expressed as a percentage) by motherhood, educational level, and duration of paid leave in a country. We only present these results for the leave policy, since in our multilevel models we found no support for our financial support policy hypothesis. The results are presented in Table 4.5. In addition we calculated the odds ratio between motherhood and employment, by education and leave. This odds ratio indicates the motherhood-employment gap. All odds ratios are smaller than 1, indicating than in all cases there is a motherhood-employment gap with mothers being less likely to be employed than women without children. The closer to 0 the odds ratio is, the bigger this motherhood-employment gap is. The predictions in Table 4.5 are based on Model IV in Table 4.4, with all variables except motherhood,

Table 4.5: Predicted Percentage of Women's Employment, by Motherhood, Education, and Leave Policy.
Source: Authors' calculations on Model IV in Table 4.4, with all variables except motherhood, education and leave policy on their average score. Odds Ratios (OR) calculated based on the percentages presented in this Table.

Completed Education at Age	Mother	20 Weeks Leave % Employed	OR	75 Weeks Leave % Employed	OR
17 Years	No	75		76	
	Yes	58	0.46	61	0.49
⩾23 Years	No	91		88	
	Yes	78	0.35	81	0.58

education, and leave policy on their average score (as presented in Table 4.1).

To give an example, the odds ratio that indicates the motherhood-employment gap for women with an average level of education in a country with an average duration of paid leave, was 0.46. This was calculated using the predicted percentages of women employed presented in Table 4.5: $\frac{(1-.75)*.58}{.75*(1-.58)} = .46$.

Although Table 4.5 does not provide a new test for our hypothesis, we draw six substantively interesting interpretations from the predicted percentages and odds ratios.

Firstly, more educated women were always more likely to be employed than women with an average level of education, irrespective of being a mother and irrespective of living in a country with an average or long duration of paid leave.

Secondly, within each category of education and leave policy, mothers were always less likely to be employed than women without children.

Thirdly, for each educational level, mothers were more likely to be employed in a country with 75 weeks of paid leave, than in a country with 20 weeks of paid leave.

Fourthly, the odds ratios indicate the magnitude of the motherhood-employment gap. For women with an average level of education this odds ratio is .46, and for more educated women .35. This suggests

that in countries with an average duration of leave, the motherhood-employment gap is *bigger* amongst more educated women. This finding is in line with the expectation that the consequences of having fewer opportunities for employment (in this case: motherhood) are most pronounced for those with the highest level of interest (in this case: education).

Fifthly, it should be noted that the *absolute* effect of an increase in leave duration from 20 to 75 weeks on the employment of mothers is small: 3 percentage points for both average and more educated women. On the other hand, for women with an average level of education this comes down to a *relative* increase of 3 percentage points out of the 42 percent (equals to 3/42=7%) of mothers not employed with 20 weeks of leave. For more educated mothers, this is an increase of 3 out of the 22 percent of mothers not employed (equals to 3/22=14%), which is a substantially bigger change.

Finally, the *relative* effect of longer leave is bigger for the more educated women than for women with an average level of education. For these women, the odds ratio indicating the motherhood-employment gap was .46 in countries with 20 weeks of paid leave, and .49 in countries with 75 weeks of paid leave (a 7% difference, indicating a smaller motherhood-employment gap). In the multilevel regression model in Table 4.4, however, this difference was found to be not statistically significant. For more educated women, the odds ratio indicating the motherhood-employment gap changed from .35 to .58 with long leave (a 66% difference), indicating that the motherhood-employment gap among highly educated women was smaller in countries with long periods of paid leave. Part of this substantial decrease in the motherhood-employment gap among more educated women is caused by the lower likelihood of employment for women without children in countries with 75 weeks of leave (88% versus 91% with 20 weeks leave). We have no substantive explanation for this finding, but we point out that even if the employment of women without children had been unaffected by longer leave at 91%, we would still have found that the odds ratio changed to .42 (a 20% increase from .35). Therefore, we conclude that we find support for our hypothesis that reconciliation policies

reduce the motherhood-employment gap more strongly among more educated women than among less educated women.

4.5 CONCLUSION AND DISCUSSION

In this Chapter we answered the question of to what extent the outcomes of reconciliation policies and financial support policies on the size of the motherhood-employment gap differ between more and less educated women. We found that the motherhood-employment gap is bigger among more educated women than among less educated women. Longer periods of paid leave (the combined paid duration of maternity, parental and childcare leave) reduce the motherhood-employment gap, and do so more strongly among more educated women. This is in line with our reconciliation policy hypothesis. We found that higher expenditure on family allowances increases the motherhood employment gap among all women, but does not have a greater effect on less educated women. While our general financial support policy hypothesis was corroborated by the analyses in Chapter 2, we had to reject the specific financial support policy hypothesis formulated here.

To test our hypotheses with three-way interactions, we used a subset from the Comparative Motherhood-Employment Gap Trend File, which was based on 239 pooled surveys covering 17 countries from 1980 to 1999, including 116,874 individual women living in a partnered household. Using these survey data, however, also came at a cost. We had no information on the timing of the births and therefore no indication of the age of the children. This means that the timing of becoming a mother is likely to differ between observed individuals, and that we could not precisely synchronise the policy measurements with the moment at which women make their decisions regarding both employment and motherhood. Since we used pooled cross-sections, our person-level analyses did not account for the time-dimension at the individual level. Our findings only apply to women aged 20 to 50, and living in coupled households. Despite these restrictions, we were able

to contribute to the emerging literature on how the outcomes of family policies interact with women's demographic background (Cooke, 2011; Korpi, Ferrarini, & Englund, 2013; Pettit & Hook, 2005, 2009; Yerkes, 2013), which in this study was indicated by women's educational level, in three ways.

Firstly, in studying the outcomes of family policies, our hypotheses on these outcomes were derived from the combination of opportunity-based explanations that only focus on the characteristics of policies (e.g. the duration of leave) and interest-based explanations that focus on the characteristics of individual women (in this case, their level of education). Indeed, our findings suggest an interest in seeking employment is not enough when no opportunities to do so are provided, and that mere opportunities without interests do not result in higher levels of employment. Instead, leave policies strongly affect the employment of mothers for whom interests and opportunities align.

Secondly, even though our questions did not focus on trends in the motherhood-employment gap, observing each country multiple times accounts for unobserved, time-invariant, factors at the country-level.

Finally, we derived two hypotheses on how family policies affect the motherhood-employment gap, and our empirical analyses allowed for tests of these hypotheses. Many country-comparative studies on the interplay between family policies and women's demographic background, however, do not take this interaction into account (Cooke, 2011; Korpi et al., 2013; Pettit & Hook, 2005, 2009). Our analyses were more informative as they estimated the three-way interaction between motherhood, educational level of the mother, and policy. For instance, in their country-comparative study Pettit and Hook (2009) hypothesised that the availability of childcare particularly benefits the employment of more educated women. In their empirical analyses this hypothesis was confirmed for *all* more educated women, which is in line with the findings reported in this study. However, we were also able to differentiate between how family policies were associated with the employment of women with children and women without children *differently* among more and less educated women. This is a stronger test of the outcomes of family policies, because these policies focus on

mothers and families with children and are therefore expected to affect the difference in employment between women with children and women without children.

In addition, our findings showed that more educated women are more likely to be employed than less educated women whether they are a mother or not, but also that the motherhood-employment gap was bigger among higher educated women than among the lower educated women. Due to our cross-sectional design, we cannot determine whether this association results from more educated employed women being more likely to postpone or refrain from having children compared to less educated women, or from motherhood having a stronger negative impact on the employment decisions of more educated women compared to less educated women, or whether both mechanisms play a role. We were, however, able to show that reconciliation policies narrow the motherhood-employment gap more strongly among more educated women. This could not have been found without explicitly interacting indicators of education, motherhood, and policy.

To conclude, in this Chapter we have contributed to the literature on the interaction between family policies and women's demographic background. We argued for explicitly testing how family policies have a different effect on the motherhood-employment gap among more educated women than among less educated women. It has shown that the difference in employment between women with children and women without children is bigger for more educated women. Moreover we have shown how more educated mothers respond more strongly than less educated mothers to the opportunities provided by reconciliation policies. For policy-makers seeking to reduce inequalities in employment between women with children and women without children - or between men and women - by implementing reconciliation policies, this suggests that the outcomes of these policies themselves can lead to stratification by educational level.

Part III

EARNINGS INEQUALITY WITHIN AND BETWEEN HOUSEHOLDS

This part answers questions pertaining to how women's increased earnings in recent decades have affected earnings inequalities *within* and *between* households. Chapter 5 addresses the enigmatic finding that despite a rising positive correlation between spouses' earnings *within* households, women's earnings increasingly attenuated the inequality *between* households in OECD countries from 1981 to 2005. Our explanation is that as women's employment increased, earnings inequality amongst women decreased due to fewer women having no earnings at all. Using a counter-factual analysis, the Chapter shows that if women's earnings had not increased during recent decades in OECD countries, ceteris paribus, inequality between households in 2005 would have been higher than it actually was. Chapter 6 answers the question of how family policies have affected the contribution of women's earnings to the inequality of earnings between households. In societies with extensive reconciliation policies women's earnings were found to have a stronger attenuating effect on household earnings inequalities, compared to societies without such policies. Financial support policies contribute to higher earnings inequality between households.

5

WOMEN'S EARNINGS: TRENDS IN EARNINGS INEQUALITY WITHIN AND BETWEEN COUPLED HOUSEHOLDS IN 18 OECD COUNTRIES, 1981-2005

ABSTRACT

In this Chapter we show that women's earnings attenuate inequality between coupled households, even though the earnings of spouses are positively correlated. We use data from the Luxembourg Income Study (LIS, 2013) on 572,222 coupled households, covering the period from 1981 to 2005 in 18 OECD countries. Three trends are described. Firstly, over time women's earnings increasingly contributed to total household earnings, thereby increasing equality within households. Secondly, the positive correlation between spouses' earnings increased over time. Thirdly, earnings inequality among women declined. With a counter-factual decomposition technique on earnings inequality, we show that the combined effect of these trends was that women's earnings increasingly attenuated earnings inequality between households. The trend towards women's earnings increasingly attenuating the inequality between households was mainly driven by decreasing inequality among women. If inequality among women had not declined as it did in recent decades, inequality between households would have been 25% higher than it actually was in 2005.

1 This Chapter is based on: Nieuwenhuis, R., Need, A., Van der Kolk, H. (2013c). Women's Earnings: Trends in earnings inequality within and between coupled households in 18 OECD countries, 1981-2005. *LIS Working Paper Series*, #598.

5.1 BACKGROUND AND RESEARCH QUESTIONS

Women's earnings have been rising in OECD countries during recent decades, because of increased employment rates, higher wages, longer working hours, and higher status positions (Costa, 2000). Because of this stronger position in the labour market, women contributed increasingly large shares to the total earnings of households. Consequently, earnings inequality between men and women in general decreased (Blau & Kahn, 2000; Charles, 2011; Gregory, 2009), and so did earnings inequality between men and women *within* coupled households.

In the same period in which women's earnings were rising, earnings inequalities *between* households were rising as well, particularly after the 1980's (Esping-Andersen, 2009). The trend towards higher inequality between households was represented in Figure 1.6 in Chapter 1 on Page 32. Scholars have often raised questions pertaining to how women's earnings affected inequality between households. Scholars are interested in the relationship between women's earnings and inequality between households, because spouses' earnings tend to be correlated (Lam, 1997; Burtless, 2009). It has been hypothesised that if the correlation between the earnings of spouses is positive, increased earnings by women would increase inequality between households (Esping-Andersen, 2007, 2009). Similarly, educational homogamy was hypothesised to contribute to inequality (Burtless, 1999; Blossfeld & Drobnič, 2001), by boosting the positive correlation between spouses' earnings (Breen & Salazar, 2009; Esping-Andersen, 2009). Furthermore, the rise of women's employment was '*stratified*' (Esping-Andersen, 2009, p.53), with more educated women being much more likely to be employed than less educated women. Earnings inequality among women has been hypothesised to further exacerbate inequality between households as long as the correlation between spouses' earnings was positive (ibid.).

Empirical results, however, suggest that women's earnings have an attenuating, rather than an exacerbating, effect on earnings inequality between households. Mincer (1962) expected that a positive cor-

relation between spouses' earnings would exacerbate inequality between households, while a negative correlation would attenuate inequality. After observing that wives' and husbands' incomes in the United States were negatively correlated, he thus hypothesised that wives' earnings would have an equalising effect on income inequalities between households. Later, Mincer found empirical support for this hypothesis in the 1960 census data, with inequalities between households observed to be lower than husbands' income inequalities (also see: Lam, 1997). This difference in inequality between men's earnings and inequality between households was to be attributed to wives' earnings, as Mincer argued that "*growth of the female labour force, while increasing the earnings inequality among all persons, has actually been a factor in the mild reduction of income inequality among families*" (1974, p. 125). Lam (1997) showed that while the negative correlation between spouses' incomes turned from negative to positive in the United States, women's earnings continued to have an attenuating effect on income inequalities between households.

Many other authors have reported the conclusion that women's earnings reduce inequality between households. This has been reported in studies on single countries, such as Sweden (Björklund, 1992), and the United Kingdom (Harkness, Machin, & Meguir, 1996; Machin & Waldfogel, 1994). Several studies on single countries evaluated trends. It was found that in the United States women's attenuating contribution to household inequality had become increasingly strong from 1968 to 1987 (Betson & Van der Gaag, 1984; Lam, 1997). Also in the United States, Cancian & Reed (1999) reported that even although the correlation between spouses' earnings became increasingly positive, rising levels of women's earnings could not explain the trend towards more inequality. This was also found in Ireland (Callan, Nolan, Neill, & Sweetman, 1998). Similarly, Breen & Salazar (2009) found that the increasing levels of educational homogamy could not explain rising inequality in the United Kingdom. Mastekaasa & Birkelund (2011) reported that in Norway in the 1970s women's earnings had a minor exacerbating effect, but with rising women's employment this effect

changed towards women's earnings equalising between-household inequality.

Other authors have compared the attenuating effect of women's earnings on inequality between households across countries. Cancian and Schoeni (1998) compared 10 developed countries to find that the earnings of women reduced inequality between coupled households in all of those countries, even though these countries differed markedly in the degree to which women contributed to total household earnings. Harkness (2013) found that this attenuating effect was stronger in countries with high female employment, such as the Nordic countries compared to the southern European countries. Pasqua (2002) found that inequality in 14 European countries was lower among households with two earners than among households in which only the man was employed.

Gregory (2009) summarised research findings and claimed that the consensus in the literature is that women's earnings have an attenuating effect on household inequality, rather than exacerbating these inequalities. In this Chapter we improve upon this literature on two ways.

Firstly, studies on the contribution of women's earnings to between-household inequality have focused either on a single country and analysed how inequality and women's attenuating effect on this inequality has developed over time, or compared this attenuating effect across countries at a single moment of time. As such, little is known about how trends in the degree to which women's earnings attenuated between-household inequality have differed across countries. Hence, we raise our descriptive question:

QUESTION 5 To what extent have women's earnings attenuated earnings inequalities *between* households in 18 OECD countries from 1981 to 2005?

Secondly, the explanation of trends and cross-country variation in the degree to which women's earnings attenuated inequality between households, is very limited. It has, however, been established that the contribution of women's earnings to between-household inequality not only depends on the correlation between spouses' earnings, but

on three factors: (a.) the correlation between spouses' earnings, (b.) women's share in total household earnings and the earnings inequality among women (relative to that of men), and (c.) women's share in total household earnings (Lam, 1997; Shorrocks, 1983). How these three aspects of women's earnings determine the degree to which women's earnings affect inequality between households will be specified in the next section of this Chapter. It has not been empirically examined how and to what extent trends in these aspects of women's earnings can explain trends in the attenuating contribution of women's earnings to between-household inequality. We therefore raise the explanatory question:

QUESTION 6 To what extent have changes in (a.) the correlation between spouses' earnings, (b.) earnings inequality among women, and (c.) women's share in total household earnings, affected the degree to which women's earnings attenuate household-level inequality in OECD countries between 1981 and 2005?

5.2 THEORY AND HYPOTHESES

The degree to which women's earnings attenuate (or exacerbate) the inequality of earnings between households, depends on three aspects of women's earnings:

A. CORRELATION BETWEEN SPOUSES' EARNINGS Women who have higher earnings tend to have higher earning spouses. The degree to which this is the case is reflected in the strength of the correlation between spouses' earnings. If the correlation between spouses' earnings is highly positive, women's earnings have a stronger tendency to increase inequality between households (Breen & Salazar, 2009).

B. EARNINGS INEQUALITY AMONG WOMEN (RELATIVE TO MEN) If the level of earnings inequality among all women in a country is higher than it is among men, women's earnings have a stronger tendency to increase inequality between households.

C. WOMEN'S SHARE IN TOTAL HOUSEHOLD EARNINGS The share of women's earnings in total household earnings, by itself, does not determine whether women's earnings attenuate or exacerbate inequality between households. That is determined by the combination of (A.) the correlation between spouses earnings, and (B.) inequality among women compared to that among men. However, given (A.) and (B.), a larger share of women's earnings in total household earnings increases the magnitude of the impact of women's earnings on between-household inequality.

In the next section the mathematical relationship between these components and earnings inequality between households will be detailed.

5.2.1 Decomposition of Inequality Between Households

To determine the extent to which women's earnings affect between-household inequality, the squared coefficient of variation is commonly used, as a measure of *relative* inequality.[1] The squared coefficient of variation indicating the inequality between households in a country (CV_h^2) is calculated as the variance of household earnings in that country (Y_h) divided by twice the average household earnings, as follows:

$$CV_h^2 = \frac{\sigma^2_{Y_h}}{2 \times \overline{Y_h}^2} \tag{5.1}$$

Next, this inequality is decomposed into aspects of women's earnings and men's earnings. To do this, Lam (1997) suggested re-writing the squared coefficient of variation at the household level as:

$$CV_h^2 = CV_m^2 \alpha_m^2 \times CV_w^2 \alpha_w^2 + 2\rho_{mw} CV_m CV_w \alpha_m \alpha_w \tag{5.2}$$

1 It has been shown that for all standard inequality indexes, including both the (squared) coefficient of variation and the GINI coefficient (Gini, 1912), the total between-household inequality can be decomposed using a single formula (Jenkins & Van Kerm, 2009; Shorrocks, 1983). Using different measures of relative inequality leads to similar conclusions (also see: Gronau, 1982; Harkness, 2013).

In this equation (5.2), the squared coefficient of variation at the household level (CV^2_h) is a function of:

- The correlation between men's and women's earnings within households (ρ_{mw}).
- The squared coefficient of variation in women's earnings (CV^2_w) and men's earnings (CV^2_m). Earnings inequality amongst women and amongst men are separate terms in this equation. The more unequal women's earnings are compared to men's, the more likely women's earnings are to increase inequality between households.
- The (squared) share of women's earnings (α^2_w) and men's earnings (α^2_m) in total household earnings. As total household earnings are assumed to be the sum of her and his earnings only, both shares sum to 1.

Once the inequality between households is broken down into different components, it can be determined to what extent the earnings of women attenuate or exacerbate inequality between households. The contribution of women's earnings to between-household inequality ($\mathtt{contrib}_w$) is expressed as the percentage by which the household inequality would change in a (counter-factual) scenario where women had no earning at all. This percentage is calculated based on the difference between men's inequalities and household inequalities (Lam, 1997):

$$\mathtt{contrib}_w = \frac{CV^2_h - CV^2_m}{CV^2_m} \times 100\% \tag{5.3}$$

It is a "*common misconception*" (Lam, 1997, p. 1026) that a positive correlation between spouses' earnings ρ_{mw} is a sufficient condition for women's earnings to increase the level of earnings inequality between households ($\mathtt{contrib}_w > 0$).

For women's earnings to increase inequality between households, it is *necessary* but not *sufficient* that the inequality amongst women's earnings is *greater* than that amongst men. Even when this condition is met,

women's earnings only increase inequality between households if the correlation between spouses' earnings is positive and strong. When the correlation is below 1, women's earnings are more likely to attenuate earnings inequality between households. This is because the weaker the correlation, the less likely extremely high (or low earnings) are to be matched by equally extreme earnings.[2]

5.2.2 Visualisation of Inequality Decomposition

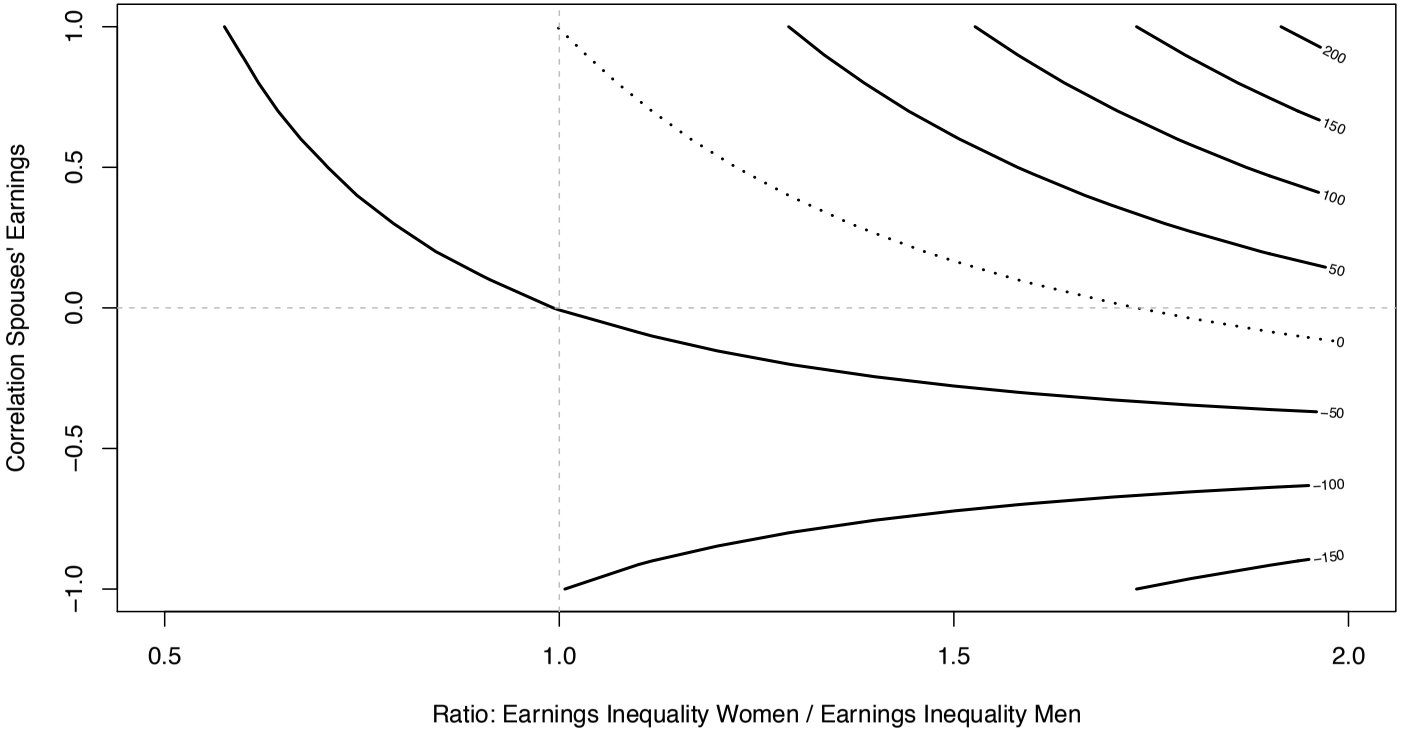
Figure 5.1 illustrates[3] that in order for women's earnings to exacerbate between-household earnings inequality, earnings inequality between women must exceed earnings inequality among men. The lines in the graphic represent the contribution of women's earnings to between-household inequality ($\mathtt{contrib}_w$) for different values of (a.) the correlation between spouses' earnings (ρ_{mw}) on the y-axis and (b.) the ratio between the earnings inequality of women compared to men (CV_w/CV_m) on the x-axis. The y-axis covers all possible values for ρ_{mw} and thus ranges from -1 to +1. The x-axis ranges from the earnings inequality among women being half that of the inequality among men ($CV_w/CV_m = .5$) to women's earnings being twice as unequal as men's ($CV_w/CV_m = 2$). These values were filled into Equation 5.2 to calculate the household inequality associated with these values, and then Equation 5.3 was used to calculate the contribution of women's earnings to between-household inequality ($\mathtt{contrib}_w$). It was also assumed that women and men have the same average earnings, or in other words that women's share of total household earnings (α_w^2) was .5.

The point in Figure 5.1 where the two straight (and dashed) lines cross, represents equal earnings inequality for men and women (x=1) and no correlation between spouses earnings (y=0). The curved line that crosses this point represents the value of women's earnings to between-household inequality ($\mathtt{contrib}_w$) of -50%, as indicated by the numbers on the right. This means that with identical earnings inequal-

2 This is similar to the principle of regression to the mean, a phenomenon originally observed by Galton (1886).

3 See Gronau (1982) for an alternative visualisation.

Figure 5.1: Contribution of Women's Earnings to Between-Household Inequality, by Correlation Between Spouses' Earnings and Earnings Inequality Among Women.
The lines represent the contribution of women's earnings to between-household inequality ($contrib_w$), percentages indicated by the numbers on the right.
Source: Data simulated by authors. Women's share to total household earnings (α_w^2) was .5.



ity among men and women and no correlation between spouses' earnings, women's earnings halve the inequality between households. Relative to this point (x=1, y=0) in Figure 5.1, we describe two scenarios.

Firstly, moving upwards along the vertical straight line represents a higher correlation between spouses' earnings (values of y exceeding 0). With a stronger positive correlation between spouses' earnings, women's attenuating contribution to between-household inequality becomes weaker and equals 0 with a perfect correlation between spouses' earnings. In other words, if men's and women's earnings have the same inequality and spouses' earnings are perfectly correlated, the earnings of women do not affect inequality between households.

Secondly, moving to the right along the horizontal straight line represents women's earnings being distributed more unequally than men's (values of x exceeding 1). With higher inequality among women than men, women's attenuating contribution to inequality between households becomes weaker. When there is no correlation between spouses' earnings (y=0) and women's earnings are twice as unequal as men's (x=2), women's contribution is positive, meaning that women's earnings exacerbate inequality between households.

The dotted curve in Figure 5.1 represents scenarios in which women's earnings do not affect between-household inequality. The area under this curved line shows that even when the correlation between spouses' earnings is positive (y>0), the inequality among women's earnings needs to be substantially higher than among men for women's earnings to exacerbate inequality between households.

5.2.3 Trends in Contribution Women's of Earnings to Household Inequality

Up to this point, we have explained how different aspects of women's earnings relate to inequality between households. Women's earnings tend to exacerbate earnings inequality between households when the correlation between spouses' earnings is strongly positive and when earnings inequality among women is high relative to inequality among men. When women earn a large share of total household earnings, this

increases the impact of women's earnings on the inequality between households, but whether this impact is attenuating or exacerbating depends on the correlation between spouses' earnings and the inequality of earnings among women. This theory needs to be combined with expectations about (A.) the correlation between spouses' earnings, (B.) the earnings inequality among women, and (C.) women's share in total household earnings to be able to derive hypotheses that answer our explanatory Question 6.

Regarding (A.), the correlation between spouses' earnings, it has been suggested by other studies (Breen & Salazar, 2009; Callan et al., 1998; Cancian & Reed, 1999; Oppenheimer, 1994, 1988; Sweeney, 2004) that this correlation has increased moderately over time in OECD countries. We thus expect that the correlation between spouses' earnings increased between 1981 and 2005. As a higher correlation between spouses' earnings results in women's earnings exacerbating inequality between households, we hypothesise that if the trend towards a stronger positive correlation had not taken place, the inequality between households would have risen *less* than it actually has between 1981 and 2005.

Regarding (B.), earnings inequality among women, various studies have shown that the increased labour force participation of women in OECD countries during recent decades has resulted in a reduction in the number of women with zero earnings, and consequently substantially lower earnings inequality among women (Cancian & Reed, 1999; Gregory, 2009). We argued above that higher inequality in earnings among women contributes to larger inequalities between households. The opposite also holds: with lower earnings inequality among women, women's earnings have a stronger attenuating contribution to between-household inequality. We thus expect a trend towards lower earnings inequality between women, and hypothesise that if this trend had not taken place, inequality between households would have risen *more* than it actually has between 1981 and 2005.

Regarding (C.), women's share in total household earnings, it has been found that women's share has increased over time (Charles, 2011; Costa, 2000; Gregory, 2009). It has been argued by itself that this factor

does not affect between-household inequality, but the size of the share women contributed affects the *degree to which* between-household inequalities are affected by women's earnings. The consensus is that women's earnings attenuate inequality between households (Gregory, 2009), even though women's earnings were distributed more unequally than men's and spouses' earnings were positively correlated. Given this, we hypothesise that if the trend towards women's earnings being a larger share of total household earnings had not taken place, inequality between households would have risen *more* than it actually has between 1981 and 2005.

To summarise this discussion, we hypothesise:

WOMEN'S EARNINGS HYPOTHESIS In OECD countries between 1981 and 2005, (A.) the positive correlation between spouses' earnings increased, (B.) earnings inequality among women's earnings decreased, and (C.) women's share of household earnings increased.

ATTENUATION HYPOTHESIS The trend towards higher inequality between households in OECD countries between 1981 and 2005 would have been less steep if (A.) the correlation between spouses' earnings had not risen, and more steep if (B.) inequality among women had not risen and if (C.) women's share in total household earnings had not risen.

5.3 DATA AND METHOD

5.3.1 Data From the Luxembourg Income Study

Our hypotheses were tested using data from the Luxembourg Income Study (LIS, 2013). LIS provides country-comparative household- and person-level surveys on income, organised in waves. We have used data from LIS waves 1 through 6 for 18 OECD countries (listed in Table 5.1), covering the period from 1981 to 2005. In total, 99 LIS datasets were used. By using sampling weights these LIS datasets provided

Table 5.1: Number of observations on coupled households, datasets, and time-span covered for 18 OECD countries
Source: Data from the Luxembourg Income Study (LIS, 2013).

Country	First Year	Last Year	N. Years	N. Obs.
Australia	1985	2003	5	13,619
Austria	1994	2004	4	5561
Belgium	1985	2000	6	11,067
Canada	1981	2004	8	74,215
Denmark	1987	2004	5	95,707
Finland	1987	2004	5	27,155
France	1989	2005	4	18,205
Germany	1981	2004	7	32,512
Greece	1995	2004	3	6518
Ireland	1994	2004	5	7375
Italy	1986	2004	9	36,043
Luxembourg	1985	2004	6	7585
Netherlands	1983	2004	6	14,144
Norway	1991	2004	4	19,132
Spain	1990	2004	4	23,321
Sweden	1981	2005	6	32,413
United Kingdom	1986	2004	6	28,242
United States	1986	2004	6	119,408
Total	1981	2005	99	572,222

representative samples of the respective countries' populations. Our sample was limited to coupled households, defined as two spouses living together who are married or in a consensual union. The sample was further limited to couples where both spouses were aged between 18 and 59 at the time of interview. Same-sex couples were removed from the data. These restrictions on the data were required to allow for the decomposition of earnings inequalities between households, and were necessary to determine the (influence of the changing) correlation between spouses' earnings. These decisions correspond to those made in similar studies (e.g. Harkness, 2013), ensuring comparability of the results. The total number of coupled households per country in our data are presented in Table 5.1.

The key variable observed in the data was earnings, defined as the monetary returns from paid employment. Negative earnings were recoded to 0, and earnings were trimmed at the level of the 99th percentile. We measured earnings for both of the spouses in the coupled households, and at the level of the household. Household earnings were defined as the sum of the earnings of two spouses, even when either or both spouses had no earnings. Based on these measurements, for each country-year we calculated the following measures:

INEQUALITY BETWEEN HOUSEHOLDS Calculated as the coefficient of variation (defined in Equation 5.1) of the total household earnings: CV_h.

INEQUALITY AMONG WOMEN Calculated as the coefficient of variation of women's earnings: CV_w.

INEQUALITY AMONG MEN Calculated as the coefficient of variation of men's earnings: CV_m.

WOMEN'S SHARE IN TOTAL HOUSEHOLD EARNINGS Women's earnings as a proportion of total household earnings: α_w.

CORRELATION OF SPOUSES' EARNINGS Pearson's correlation coefficient between spouses' earnings: ρ_{mw}.

5.3.2 Comparability of Net and Gross Datasets With LIS

LIS income variables were reported either net of taxes and social security contributions, or gross of taxes and social security contributions. These measures cannot be compared without accounting for the fact that net and gross earnings are different constructs. Where available, earnings net of taxes and social security contributions were used and when necessary net earnings were calculated by subtracting taxes and social security contribution from gross earnings. The procedures we developed for doing this are described in detail in Appendix B.

5.3.3 Statistical Method: The Role of Counter-Factuals

The contribution of women's earnings to inequality between households ($\texttt{contrib}_w$) cannot be directly observed, but can be inferred by comparing actual inequality between households with what the inequality would have been if women did not have any earnings (see Equation 5.3), or if women's earnings were different in another way. Thus, any assessment of women's contribution to between-household inequality is always based on a counter-factual scenario. In this Chapter we use two counter-factuals.

The first counter-factual is the scenario that all women had zero earnings, while the distribution of men's earnings remained unaltered. Thus, in Equation 5.3, a comparison is made between earnings inequality in the observed scenario in which household-earnings are the sum of both spouses' earnings and earnings inequality in the counter-factual scenario in which household-earnings are only comprised of men's earnings. If the latter level of inequality is higher, thus if the observed inequality between men is higher than the observed inequality between households, women's earnings are argued to have attenuated the inequality between households compared to the (counter-factual) scenario in which women had no earnings at all. This procedure is commonly applied (Folbre, Gornick, Connolly, & Munzi, 2013; Gronau, 1982; Harkness, 2013; Lam, 1997).

The second counter-factual is the scenario that aspects of women's earnings (A. correlation between spouses' earnings, B. earnings inequality among women, and C. women's share in total household earnings) had not changed during the period in which countries were observed in this study. Thus, for instance, Table 5.1 shows that we observed Australia for the first time in 1985 and the last in 2003. To assess the extent to which changes in women's earnings during that period affected the attenuating effect of women's earnings on between-household inequality, we calculated the between-household inequality in 2003 based on aspects of women's earnings (correlation, inequality, and share) as observed in 1985. We did this for each country separately and averaged the results which will be reported in Table 5.4.

Counter-factuals do not allow for an interpretation in causal terms (Cartwright, 2007), but are useful in determining the role of different aspects of women's earnings in the trend towards higher inequality of earnings between households. The results of this Chapter should thus be interpreted as a retrospective description of trends between 1981 and 2005, and not as model-based generalisable results.

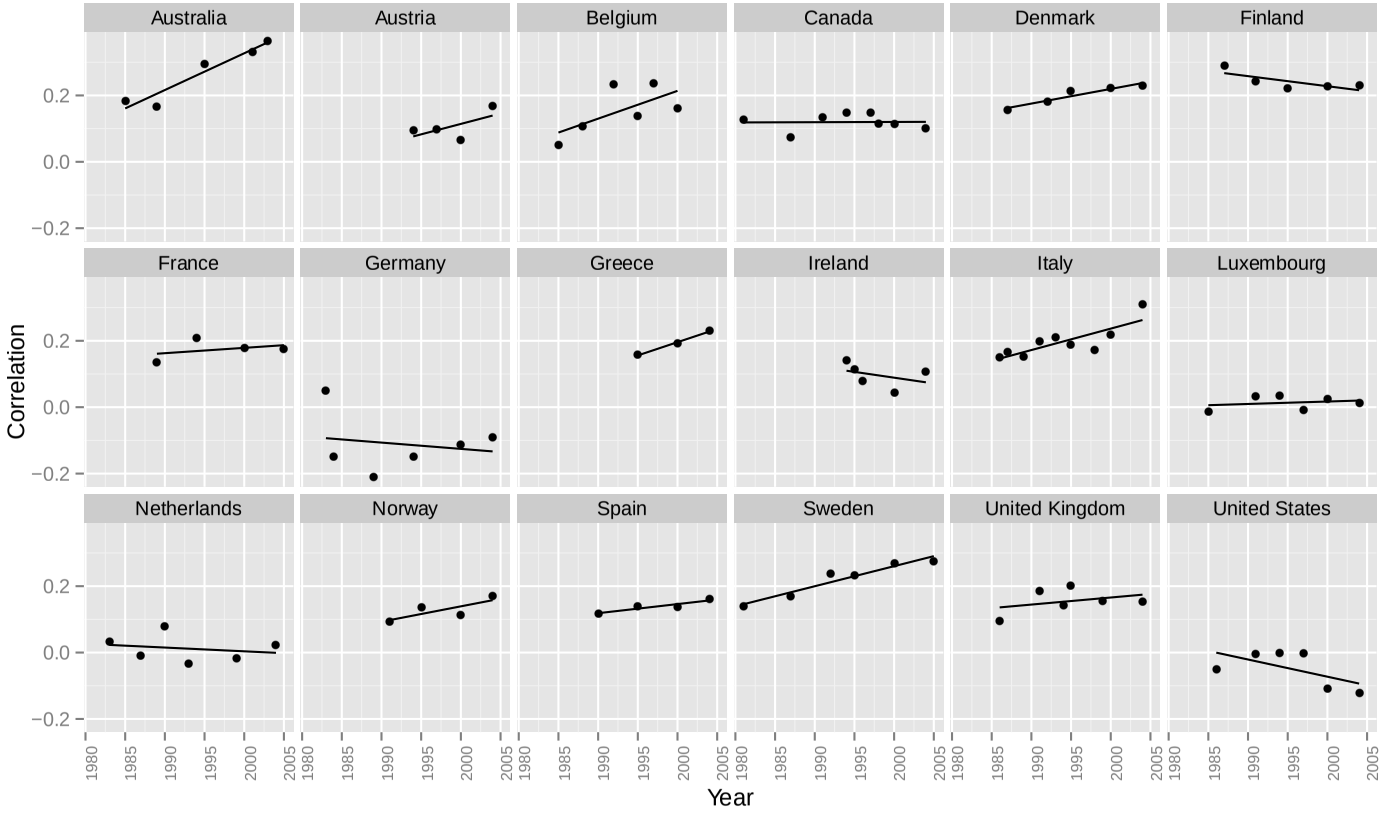
5.4 RESULTS

5.4.1 Factual Descriptions

The first step in our analysis is to present a (factual) graphical description of trends in (A.) the correlation between spouses' earnings (Figure 5.2), (B.) the earnings inequality among women (Figure 5.3) and (C.) women's share in total household earnings (Figure 5.4). Note that Figure 1.6 (on Page 32) showed the trend towards higher inequality between households in OECD countries. The black lines in the panels of Figures 5.2, 5.3, and 5.4 show the linear trends per country. These figures show that the correlation between spouses' earnings increased in about half of the observed countries, and confirm that earnings inequality among women decreased in several countries. The share of women's earnings as a percentage of total household earnings increased over time, with women's share rising towards 50%. These findings are in line with our women's earnings hypothesis.

Before continuing, it should be pointed out that in these descriptive statistics a few outliers are present. This is most clear in the case of inequality among women's earnings in the Netherlands. The first two observations show substantially higher inequality than later observations, most likely overly influencing the linear trend. Nevertheless, overall the linear trends in the Figures fit the data reasonably well.

Figure 5.2: Correlation Between Spouses' Earnings, 18 OECD countries 1981 - 2005
Source: Data from the Luxembourg Income Study (LIS, 2013).



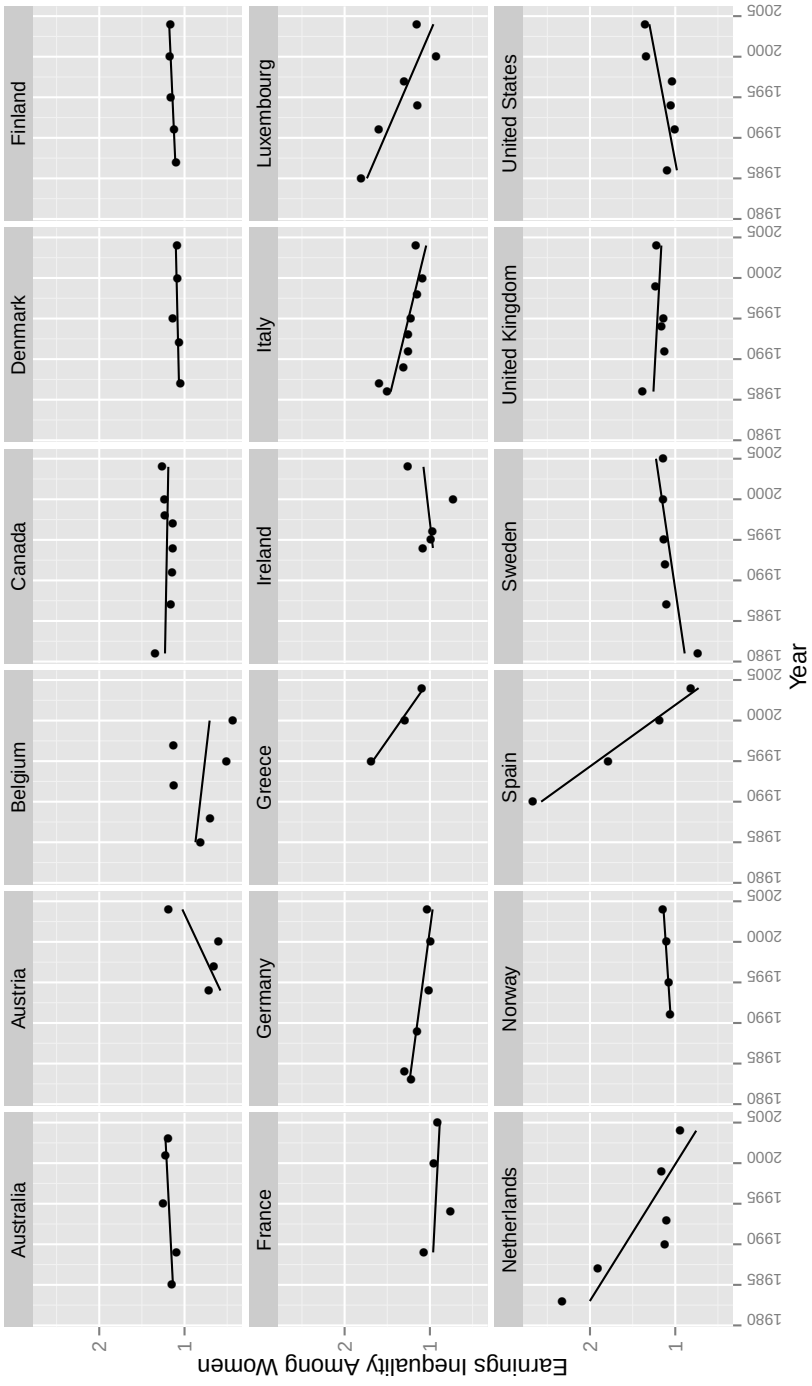
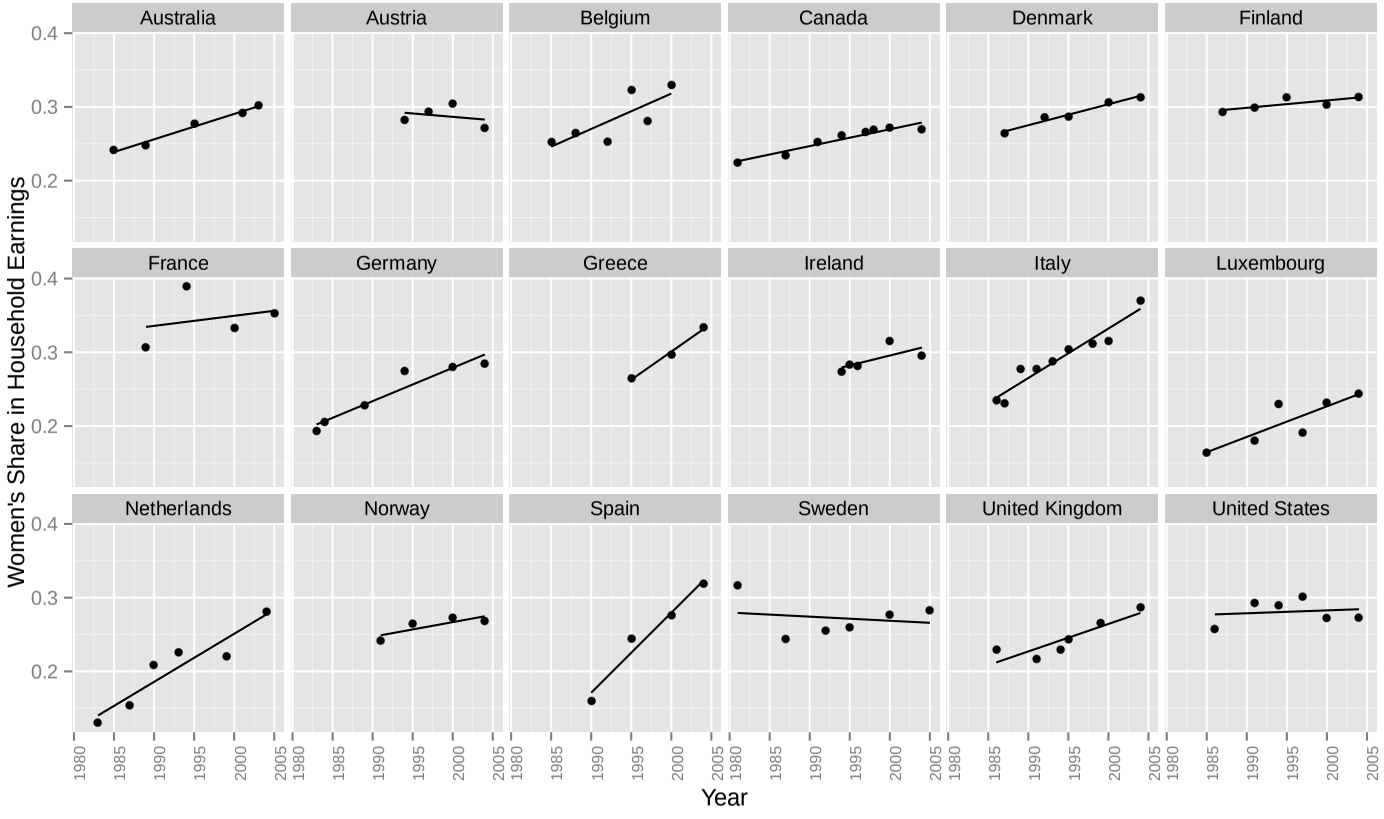


Figure 5.3: Earnings Inequality Among Women, 18 OECD countries 1981 - 2005
Source: Data from the Luxembourg Income Study (LIS, 2013).

Figure 5.4: Women's Share in Total Household Earnings, 18 OECD countries 1981 - 2005
Source: Data from the Luxembourg Income Study (LIS, 2013).



5.4.2 Counter-factual Decomposition

In this section, we answer our research questions. Question 5 pertained to describing the degree to which women's earnings attenuated earnings inequalities between households in 18 OECD countries from 1981 to 2005. In Table 5.2 we present the attenuating contribution of women's earnings for a cross-section of 18 OECD countries. For each country, Table 5.2 presents earnings inequality among men, among women, and between households. For instance, in Australia, the inequality among women exceeds that among men, but inequality between households is lower than inequality between men. In the counter-factual scenario where women had zero earnings, inequality between households would have been equal to inequality among men. Following Equation 5.3, in Australia in 2003 women's earnings attenuated between-household inequality by $(.18 - .26)/.26 * 100 = -29.66$ percent.

The remainder of the results presented in Table 5.2 lead to a similar answer to our descriptive question: inequality among women's earnings is larger than among men's (with the exception of Italy), and inequality between households is lower than inequality among either men's earnings or women's earnings. Again following Equation 5.3, women's earnings attenuated between-household inequality in the 18 OECD countries in our data. This is presented in the final column of Table 5.2. In the counter-factual scenario of women having zero earnings, inequality between households would have been substantially higher in OECD countries in around 2004 compared to the observed actual inequality.

Next, we determine trends in the degree to which women's earnings have attenuated inequality between households in OECD countries. The percentages presented in Table 5.3 are identical to those reported in the column labelled $\texttt{contrib}_w$ in Table 5.2, but cover a longer period of time. What stands out, as the answer to our descriptive Question 5, is that the degree to which women's earnings attenuate between-household inequality increased between 1981 and 2005 in the 18 OECD countries covered in this study. The rate of change, however, varied by country. A strong increase in the attenuating effect of women's earn-

Table 5.2: Attenuating Contribution of Women's Earnings to Between-Household Earnings Inequality in 18 OECD Countries
Source: Data from the Luxembourg Income Study (LIS, 2013).

		Earnings Inequality Between:			
Country	Year	Men	Women	Households	$\texttt{contrib}_w$ (%)
Australia	2003	0.26	0.28	0.18	−29.66
Austria	2004	0.25	0.27	0.15	−37.42
Belgium	2000	0.21	0.43	0.16	−22.73
Canada	2004	0.28	0.35	0.17	−38.32
Denmark	2004	0.15	0.17	0.10	−35.03
Finland	2004	0.23	0.25	0.15	−36.09
France	2005	0.29	0.45	0.21	−29.33
Germany	2004	0.25	0.57	0.16	−36.29
Greece	2004	0.51	1.10	0.42	−16.74
Ireland	2004	0.31	0.34	0.18	−41.07
Italy	2004	0.32	0.25	0.19	−41.57
Luxembourg	2004	0.21	0.69	0.16	−21.92
Netherlands	2004	0.21	0.48	0.15	−28.63
Norway	2004	0.19	0.23	0.12	−34.80
Spain	2004	0.32	0.82	0.27	−15.97
Sweden	2005	0.17	0.23	0.12	−27.28
United Kingdom	2004	0.29	0.30	0.17	−39.90
United States	2004	0.37	0.44	0.18	−51.03

ings was observed in, for instance, Canada (from -13 to -38), and Italy (from -7 to -42), while a smaller change was observed in the Nordic countries Denmark (from -33 to -35), Finland (from -33 to -36) and Sweden (from -26 to -27). These latter countries have frequently been observed to have had high levels of women's earnings throughout the period covered by this study (see Figure 5.4). The findings reported in Table 5.3 suggest that women's earnings have a stronger attenuating effect on household inequalities in those countries with strong female labour force participation, and that an increasing attenuating effect has been associated with women's increasing participation in other countries. This will be examined next.

Table 5.3: Trends in Attenuating Contribution of Women's Earnings to Between-Household Earnings Inequality in 18 OECD Countries
Values represent $\mathtt{contrib}_w$ (%)
Source: Data from the Luxembourg Income Study (LIS, 2013).

Year	Australia	Austria	Belgium	Canada	Denmark	Finland	France	Germany	Greece	Ireland	Italy	Luxembourg	Netherlands	Norway	Spain	Sweden	United Kingdom	United States
1981				-13												-26		
1982																		
1983								-29					-10					
1984								-28										
1985	-19		-8									-7						
1986											-7						-21	-32
1987				-30	-33	-33					-6		-10			-26		
1988			-4															
1989	-23						-24	-35			-10							
1990													-24		-1			
1991				-32		-36					-7	-6		-34			-23	-33
1992			-17		-34											-27		
1993											-12		-33					
1994		-16		-31			-34	-37		-20		-7					-29	-33
1995	-27		-23		-31	-38			-16	-22	-16			-34	-11	-28	-26	
1996										-25								
1997		-19	-14	-31								-14						-37
1998				-36							-18							
1999													-34				-38	
2000		-22	-23	-37	-35	-34	-25	-37	-21	-29	-16	-14		-40	-14	-29		-52
2001	-29																	
2002																		
2003	-30																	
2004		-37		-38	-35	-36		-36	-17	-41	-42	-22	-29	-35	-16		-40	-51
2005							-29									-27		

To analyse the impact of women's changing earnings over time on inequality between households, the final step in our analysis is based on the counter-factual scenario that women's earnings did not change within countries over time (since the first observation in our data of that country). These counter-factual analyses are presented in Table 5.4 and used to test our attenuation hypothesis. Because we did not observe each country in our data in exactly the same years or for the same period of time, we have clustered the presentation of these results by wave in the LIS data.

The top row of Table 5.4 presents the trend in earnings inequality between households, clustered by LIS wave of (approximately) every five years. These are calculated based on the measurements of inequality between households presented in Table 5.3, but averaged over the available countries per LIS wave. Over time, from Wave 1 (around 1980) to Wave 6 (around 2004), inequality between households has increased from .16 to .21. Our explanatory question (number 6) pertained to the extent to which this trend would have been different if women's earnings had not changed since around 1980. We calculated this counter-factual by keeping aspects of women's earnings constant and by using Equation 5.2.

Firstly, we calculated the counter-factual that the correlation between spouses' earnings had not changed over time, with all other aspects of women's earnings (inequality and share in total household earnings) as well as men's earnings remaining as observed in the data. The results in the row labelled "*A. Correlation Spouses' Earnings*" in Table 5.4 suggest that in this counter-factual scenario the inequality between households would have been virtually the same as observed. With respect to the counter-factual scenario in which the inequality of women's earnings had not decreased (presented in the row labelled "*B. Earnings Inequality Among Women*"), the results show that inequality between households would have been higher than actually observed, from around 1995 onwards. In this counter-factual scenario, between-household inequality would have been .26, which is about 25% higher than the observed inequality of .21. The same result was obtained for the counter-factual of no change in women's share in total household

earnings (row labelled "*C. Share Women's Earnings in Total Household Earnings*"), although in this case the impact was (.22-.21)/.21) = 5%. Finally, if all aspects of women's earnings had remained the same, i.e. the correlation between spouses' earnings, earnings inequality among women, and women's share in total household earnings, then between-household inequality would have been moderately higher than was actually observed in 2004 ([.23-.21]/.21=10%). Combined, these results corroborate our attenuation hypothesis that although the correlation between spouses' earnings, increased over time, the effect of that trend was offset by decreasing earnings inequalities among women, and as a result the rising share of women's earnings in the total earnings of households increasingly attenuated between-household inequalities in OECD countries between 1981 and 2005.

5.5 CONCLUSION AND DISCUSSION

Women's earnings attenuate earnings inequality between households. This had been observed before (Gregory, 2009; Lam, 1997), but studies typically observed only a single country over a longer period of time, or a number of countries cross-sectionally. We answered the descriptive question of to what extent women's earnings have attenuated earnings inequalities *within* households and *between* households in OECD countries from 1981 to 2005. Over time, women's earnings contributed a larger share of total household earnings, the correlation between spouses' earnings increased, and earnings inequality among women decreased. The combined effect of these trends was that the attenuating effect of women's earnings on between-household earnings increased over time in all 18 OECD countries observed in this study. This contribution of women's earnings was decomposed using the observed trends in (A.) the correlation between spouses' earnings, (B.) earnings inequality among women, and (C.) women's share in total household earnings have affected the degree to which women's earnings attenuated earnings inequality between households in OECD countries from 1981 to 2005. If the inequality between women's earn-

Table 5.4: Counter-factual Trends in Between-Household Inequalities
Values represent inequality between households, in 4 Counter-factual Scenarios
: Data from the Luxembourg Income Study (LIS, 2013).

Scenario	Wave 1 (1980)	Wave 2 (1985)	Wave 3 (1990)	Wave 4 (1995)	Wave 5 (2000)	Wave 6 (2004)
Observed Inequality Between Households	0.16	0.16	0.16	0.18	0.19	0.21
counter-factual. No change in:						
(A.) Correlation Spouses' Earnings	0.16	0.17	0.16	0.18	0.20	0.21
(B.) Earnings Inequality Among Women	0.16	0.15	0.16	0.19	0.21	0.26
(C.) Share Women' Earnings in Total Household Earnings	0.16	0.16	0.16	0.18	0.20	0.22
All of the Above	0.16	0.16	0.16	0.19	0.20	0.23

ings had not decreased as it did between 1981 and 2005, the inequality between households would have been approximately 25% higher in 2005. If all of the above-mentioned aspects of women's earnings had not changed between 1981 and 2005, inequality between households would also have been higher in 2005.

The results based on counter-factual scenarios in this Chapter cannot be interpreted in causal terms (Cartwright, 2007). Thus, the observed trends cannot be used to make predictions. These findings apply only to coupled households, and same-sex couples were excluded from the analysis. These restrictions were required for the decomposition analyses performed here, and are typical in the literature on the contribution of women's earnings to between-household inequality. The consequence of limiting our analyses to coupled households is that the increasing prevalence of single households was not taken into account. Thus, the findings reported in this Chapter cannot be generalised to the full populations of the OECD countries observed in this study. We were able to show that even while spouses' earnings became increasingly positively correlated, women's earnings attenuated between-household earnings inequality and have increasingly done so between 1981 and 2005 in 18 OECD countries.

Our findings contradict a commonly held intuition that as long as the correlation between spouses' earnings is positive, any increase in women's earnings contribute to inequality between households. Esping-Andersen argued that for women's earnings to attenuate inequality between households would require unrealistically low inequality among women and stated that "*the conditions required for an equalizing effect are quite steep*" (2007, p. 646). We showed that the correlation between spouses' earnings would need to be very high, much higher than it actually is, for women's earnings to exacerbate inequality between households. We thus conclude that the conditions for women's earnings to have a *de-equalising* effect are quite steep: high women's earnings have a strong tendency to reduce inequality between households.

6

FAMILY POLICIES, WOMEN'S EARNINGS, AND BETWEEN-HOUSEHOLD INEQUALITY: TRENDS IN 18 OECD COUNTRIES FROM 1981 TO 2005

ABSTRACT

This Chapter examines to what extent family policies have affected earnings inequality within and between coupled households. In Chapter 5 cross-country variation was found in the degree to which women's earnings attenuate earnings inequality between households. In this Chapter we explain this variation with reconciliation policies and financial support policies. We used person-level data from the Luxembourg Income Study (LIS, 2013) on 572,222 coupled households, covering the period from 1981 to 2005 in 18 OECD countries. These data were combined with country-level data from the Comparative Maternity, Parental, and Childcare Database (Gauthier, 2010). In countries with extensive reconciliation policies women contributed a larger share of total household earnings, and earnings inequality among women was relatively low. In societies with extensive financial support policies, women contributed a smaller share to total household earnings, and inequality among the earnings of women was relatively high. Women's earnings were found to attenuate inequality between households to a larger extent in countries with extensive reconciliation policies and limited financial support policies. Countries with family policy arrangements that facilitate women's employment and consequently smaller

1 This Chapter is based on: Nieuwenhuis, R., Need, A., Van der Kolk, H. (2013b). Family Policies, Women's Earnings, and Between-Household Inequality: Trends in 18 OECD countries from 1981 to 2005. *LIS Working Paper Series*, #599.

earnings inequalities *within* households also contribute to smaller inequalities *between* households.

6.1 BACKGROUND AND RESEARCH QUESTION

Reconciliation policies have been shown to have stimulated women's employment in OECD countries in recent decades. In Chapter 2 it was shown that extensive reconciliation policies reduce the gap in employment between mothers and women without children, although in Chapter 3 too long childcare leave was shown to negatively affect the employment of mothers. In Chapter 2 we showed that the rise in women's employment was explained not only by the implementation of family policies and other contextual factors (Charles, 2011; Gornick et al., 1998; Jaumotte, 2003; Matysiak & Vignoli, 2008; Pettit & Hook, 2005), but also by demographic determinants such as women's rising educational levels (Bradley, 2000) and decreasing fertility (Van der Lippe & Van Dijk, 2002).

Throughout this dissertation we have maintained that explanations of women's increased employment based on either changing institutional contexts or women's changing demographic background are not mutually exclusive. In Chapter 2 we showed that family policies affect mothers in a different way than women without children, and thus affect the motherhood-employment gap. In Chapter 4 we showed that more educated women benefit more from paid leave than less educated women. Other authors have challenged the idea that family policies have uniform effects on the earnings of women across more and less educated women (Mandel, 2012). For instance it was found that family policies selectively benefit those already in a strong position to have high earnings and consequently exacerbate earnings inequalities at the household level (Ghysels & Van Lancker, 2011; Lancker & Ghysels, 2012). The take-up of the benefits offered by reconciliation policies was found to be biased against low-income families (Ghysels & Van Lancker, 2011). Dual earnership was found to be less common

among couples with low earnings capacity (Cantillon, Ghysels, Mussche, & Van Dam, 2001).

The increased employment of women has been characterised as a polarisation between work-rich and work-poor households, because more educated women are more likely to be employed and have higher earnings and family policies have stratified outcomes. It was hypothesised that this resulted in an exacerbation of earnings inequalities between households (Esping-Andersen, 2007, 2009; McCall & Percheski, 2010).

In this Chapter we challenge the assertion that earnings inequality between households has increased as a result of the implementation of family policies. We do so on three accounts.

Firstly, not all family policies have been found to facilitate women's employment. A distinction between reconciliation policies and financial support policies has been identified (Gauthier, 1996; Thévenon, 2011; Thévenon & Luci, 2012). Whereas reconciliation policies such as leave and continued pay during leave were found to increase women's employment by facilitating women to combine motherhood and employment, financial support policies such as family allowances were found to provide women the financial opportunity not to be employed, as shown in Chapters 2 and 4

Secondly, it has been found that as women's participation in specific occupations was rising, wage levels in these occupations were declining (Mandel, 2013). The consequence of this development is that the earnings distribution in these occupations is compressed, contributing to lower earnings inequalities among women.

Finally, it was found in Chapter 5 that women's earnings attenuate earnings inequality between households. Women have been gaining a stronger position in the labour market with higher status positions and higher wages (Costa, 2000). As a result of women's stronger position for women in the labour market, earnings inequality between men and women, and *within households*, decreased (Blau & Kahn, 2000; Charles, 2011; Gregory, 2009). In Chapter 5 we concluded that earnings equality within households has a strong tendency to contribute to equality *between* households (also see: Cancian & Schoeni, 1998; Cancian & Reed,

1999; Gregory, 2009; Harkness, 2013; Jenkins & Van Kerm, 2009; Lam, 1997; Mastekaasa & Birkelund, 2011; Pasqua, 2002).

To summarise, on the one hand combining institutional and demographic explanations of women's employment has led to the suggestion that family policy outcomes are stratified by educational level, exacerbating earnings inequality across work-rich and work-poor households. On the other hand, it was found in Chapter 5 that women's rising employment and earnings had an attenuating contribution to between-household inequalities. This juxtaposition warrants further examination of how family policies have affected the degree to which women's earnings affect between-household inequality. In doing so, we contribute substantive explanations of cross-national variation in the degree to which women's earnings attenuate between-household inequalities of earnings. In this Chapter we improve upon both the literature on family policy outcomes and on the literature on the effect of women's earnings on between-household inequality by empirically testing to what extent the availability of reconciliation policies and financial support policies can explain differences between OECD countries from 1981 to 2005 in the degree to which women's earnings attenuate inequalities between households:

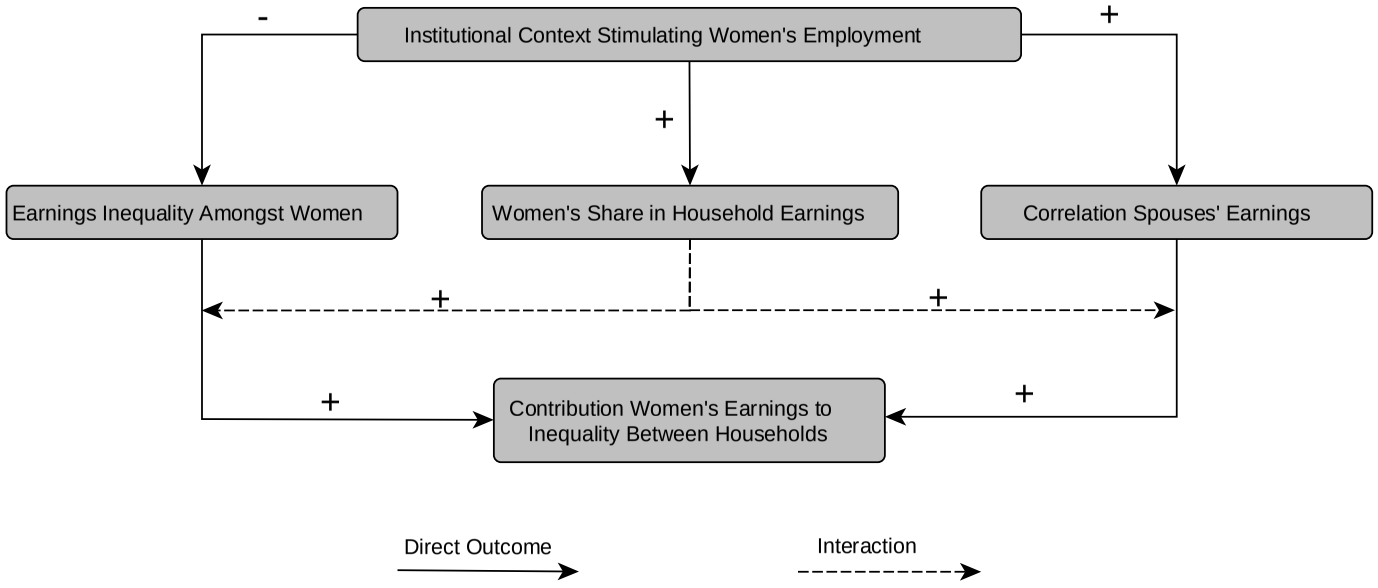
QUESTION 7 To what extent can cross-national variation in the degree to which women's earnings attenuate inequalities between households in 18 OECD countries from 1981 to 2005 be explained by (a.) reconciliation policies and (b.) financial support policies?

6.2 THEORY AND HYPOTHESES

In this section we hypothesise about how reconciliation policies and financial support policies affect the attenuating effect of women's earnings on household inequalities. Our theoretical framework is represented schematically in Figure 6.1.

The contribution of women's earnings to inequality between households is shown in the bottom row of Figure 6.1. In Chapter 5, we explained that the degree to which women's earnings attenuate (or ex-

Figure 6.1: Schematic Representation of Theoretical Framework on Family Policies and Attenuating Effect of Women's Earnings on Between-Household Inequality



acerbate) earnings inequality between households, is determined by (a.) the correlation between spouses' earnings, (b.) inequality among women's earnings, and (c.) women's share in total household earnings. These three aspects of women's earnings are shown in the middle row of Figure 6.1. In Chapter 5 it was shown, based on Equation 5.2 (Page 122) that if the correlation between spouses' earnings is positive, higher earnings inequality among women contributes to higher earnings inequality between households. This, however, will only hold when earnings inequality is higher among women than it is among men. Even if earnings inequality among women is greater than it is among men, the correlation between spouses' earnings needs to be much higher than it typically is for women's earnings to exacerbate earnings inequalities between households. As discussed in Chapter 5, it is a "*common misconception*" (Lam, 1997, p. 1026) that a positive correlation between spouses' earnings is a sufficient condition for women's earnings to increase inequalities between households.

The top row of Figure 6.1 represents how the aspects of women's earnings outlined above are hypothesised to be affected by institutional contexts. Here, we focus on family policies and how these affect women's employment. The relationship between family policies and women's employment will not be directly observed here (but was tested in Part II of this dissertation). If a family policy context facilitates many women having earnings (e.g. if female labour force participation is high), the share that women's earnings contribute to total household earnings is also expected to be high. Also, with high female labour force participation, earnings inequality among women will be low. The reason for this is that the number of women with zero earnings is reduced (Cancian & Reed, 1999; Gregory, 2009). Finally, when women's employment is high and women are likely to have earnings, it is to be expected that the positive correlation between spouses' earnings is stronger. The reason for this last expectation is given by Oppenheimer, who argues that with the stronger position of women in the labour market, the degree of educational homogamy increased because the marriage preferences of men and women converged (1988; 1994). Similarly, Sweeney (2004) found that with the increased participation of

women on the labour market, women's pre-marriage income became a more important determinant of partner selection. Our general assumption is therefore that if women's earnings are high in a country, this country has a high share of women's earnings in total household earnings, low inequality among women, and a stronger positive correlation between spouses' earnings.

Finally, we hypothesise about how family policies can affect women's employment, and in turn the attenuating effect of women's earnings on inequality between households. We again distinguish between two types of family policies: reconciliation policies and financial support policies to families. The different effects of these two types of family policies were studied in Chapters 2 and 4, and are tested here on different aspects of women's earnings.

Reconciliation policies provide opportunities to combine employment and motherhood (Gornick et al., 1998; Jaumotte, 2003; Matysiak & Vignoli, 2008; Nieuwenhuis et al., 2012a; Pettit & Hook, 2005). For maternity leave, this refers to the relatively short period before and after childbirth, and parental leave provides these opportunities when the child(ren) in the household are very young. Continued pay during leave further facilitates the opportunity to take up leave, without facing the consequences of reduced or no income. Hence, we expect that in a society with extensive reconciliation policies, women's employment and consequently women's earnings will be high:

RECONCILIATION POLICY HYPOTHESIS In countries with extensive reconciliation policies, (a.) the attenuating effect of women's earnings on between-household inequalities is *stronger* than in countries without extensive reconciliation policies, and (b.) the positive correlation between spouses' earnings is stronger, earnings inequalities among women are lower and women contribute a larger share of total household earnings.

In contrast, we expect that in countries with extensive financial support policies for families with children, women's employment will be lower. We found in Chapters 2 and 4 that financial support policies, also provide the opportunity to women not to be employed (also see: Gauthier, 1996; Thévenon, 2011). Consequently, we hypothesise:

FINANCIAL SUPPORT POLICY HYPOTHESIS In countries with extensive financial support policies to families, (a.) the attenuating effect of women's earnings on between-household inequalities is *weaker* than in countries without extensive financial support policies, and (b.) the positive correlation between spouses' earnings is weaker, inequality among women are greater and women contribute a smaller share of total household earnings.

6.3 DATA AND METHOD

6.3.1 Person-level Data

Our hypotheses were tested using data from the Luxembourg Income Study (LIS, 2013). We used data on 1,144,444 individuals in 572,222 households, covering 99 country-years from 1981 to 2005 in 18 OECD countries. With these data, we calculated four measures: (I.) the contribution of women's earnings to inequality between households, (II.) the correlation between spouses' earnings, (III.) earnings inequality among women, and (IV.) the share of women's earnings in total household earnings. These four measurements are the dependent variables in our analyses. To obtain these measurements, we have used the same datasets and procedures as in Chapter 5, where details about the sample, the computations, and the number of observations (Table 5.1 on Page 129) can be found. Descriptive statistics by country were presented in Table 6.1. Even though our measurements entail the aggregation of person-level data to the level of the country, we maintain that without the use of person-level data it would not have been possible to decompose between-household inequality.

6.3.2 Country-level data

We combined the data on four measurements of between-household inequality and aspects of women's earnings with indicators of fam-

Table 6.1: Descriptive Statistics on Earnings Inequality. Reported values apply to the Min / Max / Mean measurement across country-years, by country.
Source: Data from the Luxembourg Income Study (LIS, 2013).

	Share			Women's Inequality			Correlation			Attenuation		
Country	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Australia	35.85	41.93	38.92	1.81	3.36	2.67	16.56	36.37	26.72	−42.17	−24.02	−34.82
Austria	28.22	38.84	31.70	2.74	7.17	5.62	6.55	16.78	10.64	−59.79	−19.37	−32.61
Belgium	25.24	39.79	32.28	2.08	8.15	4.80	4.99	23.57	15.38	−30.50	−4.40	−18.26
Canada	23.82	38.91	36.03	2.22	8.82	3.48	7.29	14.76	11.97	−62.13	−15.16	−46.48
Denmark	38.15	43.03	40.84	1.31	2.21	1.68	15.58	22.89	20.02	−54.24	−45.14	−50.53
Finland	41.04	43.06	42.15	1.81	2.57	2.28	22.07	28.95	24.18	−60.07	−49.35	−54.56
France	32.06	40.32	35.93	2.95	6.10	4.61	13.47	20.77	17.39	−52.29	−31.29	−39.45
Germany	4.26	31.04	24.22	3.04	70.19	14.99	−34.96	4.91	−14.51	−59.88	−23.65	−47.06
Greece	26.47	33.36	29.83	10.95	16.90	13.60	15.76	22.98	19.32	−26.01	−19.18	−21.76
Ireland	27.37	41.26	31.31	3.44	10.85	8.24	4.35	14.10	9.66	−69.70	−24.27	−38.86
Italy	23.05	48.71	30.27	2.49	15.97	11.81	15.02	30.92	19.58	−71.14	−6.51	−19.75
Luxembourg	16.39	25.78	20.90	6.90	18.05	12.44	−1.36	3.45	1.35	−28.07	−6.09	−13.83
Netherlands	14.44	34.28	26.88	1.88	18.65	7.39	−3.35	7.88	1.21	−51.46	−10.52	−32.42
Norway	35.90	39.01	37.91	1.47	2.31	1.81	9.22	17.01	12.77	−65.48	−50.59	−55.10
Spain	15.98	31.86	24.97	8.19	26.72	16.17	11.67	16.09	13.85	−19.00	−0.99	−12.16
Sweden	33.03	39.99	37.24	1.89	2.72	2.23	13.90	27.45	22.00	−41.04	−35.15	−37.59
UK	24.33	40.38	34.51	2.10	9.22	3.71	9.44	20.10	15.48	−66.40	−25.80	−42.89
USA	27.11	39.01	32.93	4.25	6.35	5.35	−12.19	−0.17	−4.85	−106.42	−48.12	−69.61
Total	4.26	48.71	32.71	1.31	70.19	6.83	−34.96	36.37	12.34	−106.42	−0.99	−37.10

ily policies and institutional context. These contextual data were measured at the level of the country-year, and as such are time-varying.

PAID PARENTAL LEAVE Our measure of reconciliation policies is an index of three leave policies: maternity leave, parental leave, and childcare leave. Each leave policy was measured as the number of weeks mothers are entitled to. The number of weeks of each of these leave policies were weighted by the percentage of wages that are paid during this leave. The final measure represents the total number of weeks of leave with full pay. These data were available from the Comparative Family Policy Database (Gauthier, 2010).

FAMILY ALLOWANCES EXPENDITURE Our measure of financial support policies is the percentage of GDP a country spends on family allowances. These data were available from the Comparative Family Policy Database (Gauthier, 2010).

In addition, we control for two labour market variables. As the number of degrees of freedom in these analyses is smaller than it was in Chapters 2 and 4, we could only select two control variables. This is also the reason why we could not test for a curvilinear effect of leave (cf. Chapter 3). As the share of women's earnings in total household earnings is strongly dependent upon the female / male wage ratio, we control for this at the country-level. In addition, we control for the overall unemployment level as an indicator of the employment opportunities in an economy and because unemployment is an important determinant of inequality.

FEMALE / MALE WAGE RATIO Calculated as the hourly wages in manufacturing for women divided by the hourly wages in manufacturing for men. Our measure of financial support policies is the percentage of GDP a country spends on family allowances. These data were available from the Comparative Family Policy Database (Gauthier, 2010).

UNEMPLOYMENT Unemployment data available from the Comparative Family Policy Database, and defined as the "*number of unemployed persons as a percentage of the civilian labour force*" (Gauthier, 2010, p. 34).

Descriptive statistics of the country-level data in this Chapter are presented in Table 6.2.

Table 6.2: Descriptive Statistics on Family Policies and Labour Market Controls.
Source: Data From the Comparative Family Policy Database (Gauthier, 2010).

Variable	Min.	Mean	Max.	SD
Paid Leave	0.00	26.99	95.04	22.80
Family Allowance Expenditure	0.08	1.41	3.24	0.74
Unemployment	1.20	7.95	23.00	3.65
Female/Male Wage Ratio	0.54	0.77	0.91	0.08

6.3.3 Statistical Method: Multilevel Model for Change

The data described above will be analysed using the multilevel model for change (Singer & Willet, 2003). This model allows us to analyse the rate of change over time for each of the dependent variables described above for each country separately, and to test whether these are dependent on country-level independent variables. The multilevel model for change has several attractive features: first of all, it allows us to differentiate between the level of a dependent variable at the start of the observational period, and the actual rate of change in that variable. Secondly, the multilevel model for change does not require the length of the observed period to be equal in each country, nor that the observations took place in the same year. This flexibility is required for our data, since the nature of the LIS is such that not all countries have participated for an equally long period of time and the surveys were not held in all countries simultaneously.

We specify a multilevel model for change for each of our four dependent variables: (I.) the contribution of women's earnings to the inequal-

ity between households, (II.) the correlation between spouses' earnings, (III.) earnings inequality among women, and (IV.) the share of women's earnings as a percentage of in total household earnings. The models are specified to allow the trends in these variables to vary between countries, and to allow these differences in trends to be regressed on explanatory variables that are time-varying at the country-level.

The multilevel model for change differentiates between level-1 (here: trends within countries) and level-2 (here: the between-country differences in trends). The within-country part is specified as:

$$Y_{ij} = \pi_{0i} + \pi_{1i}\text{YEAR}_{ij} + \pi_{2i}\text{POLICY}_{ij} + \epsilon_{ij} \tag{6.1}$$

with Y_{ij} representing the value of one of the dependent variables for country i in year j, π_{oi} representing the onset of the trend in this variable for country i and π_{1i} representing the speed of the trend in country i expressed as the amount of change per YEAR. π_{2i} represents the effect of a policy variable in country i on the dependent variable Y_{ij}. Subscript $_{ij}$ to the POLICY variable indicates that this variable is allowed to vary within countries over time. The errors ϵ_{ij} are assumed to be distributed normally.

To relate both the onset of the trends and the speed of the trends to country-level variables, such as indicators of policy, the between-country part of the multilevel model for change is specified as:

$$\begin{aligned} \pi_{0i} &= \gamma_{00} + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \zeta_{1i} \\ \pi_{2i} &= \gamma_{20} \end{aligned} \tag{6.2}$$

The between-country part of the multilevel model for change that indicates the country-differences in the onset of the trends (π_{0i}) is specified here as an overall intercept (γ_{00}), and the variation of the per-country differences (ζ_{0i}). Similarly, the π_{1i} parameter of the within-country part is specified in the between-country part as an indicator of the overall trend (γ_{10}) and the per-country variation from this overall

trend (expressed as variance ζ_{1i}). Finally, the γ_{20} parameter indicates that the effect of policy in the within-country part of the model (π_{2i}) is assumed to be equally strong for each country.

Next, the within-country and between-country parts of the multilevel model for change can be integrated into:

$$Y_{ij} = \gamma_{00} + \gamma_{10}\text{YEAR}_{ij} + \gamma_{20}\text{POLICY}_{ij} + \zeta_{0i} + \zeta_{1i}\text{YEAR} + \epsilon_{ij} \quad (6.3)$$

Finally, the model can be extended to allow the trend to be subject to different levels of the POLICY_{ij} variable, by specifying:

$$Y_{ij} = \gamma_{00} + \gamma_{10}\text{YEAR}_{ij} + \gamma_{20}\text{POLICY}_{ij} + \gamma_{30}\text{POLICY}_{ij} \times \text{YEAR}_{ij} + \zeta_{0i} + \zeta_{1i}\text{YEAR} + \epsilon_{ij} \quad (6.4)$$

The multilevel model for change thus allows differentiation between trends within countries, and differences in these trends across countries. We centre our year-variable in such a way, that the value 0 represents the year 1995. As such, the reference category in the interaction models refers to 1995.

6.4 RESULTS

In this section, the four dependent variables are regressed on measurements of family policies. First we present the results regarding the attenuating contribution of women's earnings to between-household inequality. Then, the results regarding three aspects of women's employment are presented.

In the analyses presented in Table 6.3, the dependent variable is the degree to which women's earnings attenuate inequality between households (number I. in the above). In the first column, a baseline multilevel model for change is presented. The intercept represents the fact that in 1995, on average, women's earnings had an attenuating

effect of -36.6% on the earnings inequality between households. This means that in the counter-factual case that women had no earnings at all, earnings inequality between households would have been 36.6% higher than it actually was in 1995. Over time, as indicated by the year parameter, this attenuating effect increased in strength from (-36.64 - 14 * -1.49=) -15.78% in 1981 to (-36.64 + 10 * -1.49=) -51.45% in 2005. As indicated by the random effects, there is substantial variation in the average attenuating effect across countries.

In the second column of Table 6.3, the family policies (and controls for labour market) are introduced. In 1995, women's earnings had a stronger attenuating effect on between-household inequality in countries with long periods of paid leave than in countries with no paid leave. The strength of this effect of leave, as indicated by the interaction term, declined in the period from 1981 to 2005. In societies with high expenditure on family allowances, the attenuating effect of women's earnings on between-household inequality was weaker than in societies with low expenditure on family allowances. In contrast to the effect of paid leave, the effect of expenditure did not change over time. These findings are in line with our reconciliation policy hypothesis (part a.) and financial support policy hypothesis (part a.).

The inclusion of two family policies (and controls for the labour market) accounted for part of the between-country variation in the degree to which women's earnings attenuate earnings inequality between households. This is indicated by the reduction of the random effect for the intercept (ζ_{0i}) from 15.24 to 8.92, a 35% reduction. The variation in the rate of change of the attenuating effect of women's earnings (ζ_{1i}) was reduced from .91 to .28. This means that 69% of this variation was explained by the inclusion of paid leave and expenditure on family allowances (but not the by labour market controls as these were not interacted with year).

The attenuating effect of women's earnings on between-household inequality was thus found to be stronger in societies providing paid leave, and to be weaker in countries with high expenditure on family allowances. Based on Equation 5.2 (Chapter 5, on Page 122) it was already clear that the degree to which women's earnings atten-

Table 6.3: Multilevel Model for Change Regressing the Attenuating Contribution of Women's Earnings to Between-Household Inequality on Family Policies
The explanatory model (Model II) was controlled for female/male wage ratio and unemployment.
Source: Data from the Luxembourg Income Study (LIS, 2013) and the Comparative Family Policy Database (Gauthier, 2010).

	I. Baseline		II. Explanatory	
	B	S.E.	B	S.E.
Fixed Effects				
Intercept (γ_{00})	−36.64*	3.72	−26.27	17.17
Year (γ_{10})	−1.49*	0.27	−1.71*	0.33
Paid Leave			−0.13*	0.08
Family Allowance			5.04*	2.39
Year * Paid Leave			0.02*	0.01
Year * Family Allowance			0.23	0.21
Random Effects (SD)				
Residual (ϵ_{ij}), $N_{lvl1} = 99$	8.76		7.30	
Intercept (ζ_{0i}), $N_{lvl2} = 18$	15.24		8.92	
Random Slope Year (ζ_{1i}), $N_{lvl2} = 18$	0.91		0.28	

* $P < .05$

uate between-household inequality depends on three aspects: a high correlation between spouses' earnings exacerbates between-household earnings inequality, low inequality among women attenuates between-household earnings inequality, and when women's earnings contribute to a large share of total household earnings the potential for attenuating (or exacerbating) in increased. In the next set of analyses presented in Table 6.4, we test how paid leave and expenditure on family allowances affected each of these three aspects of women's earnings.

The first column of Table 6.4 reports the multilevel model for change, regressing the correlation between spouses' earnings on family policies. For each explanatory variable, the multilevel model for change differentiates between the effect of this variable on the onset of the change and on the rate of change. We did not find any statistically significant association between either type of family policy and the correlation between spouses' earnings in 1995. Over time, as indicated by the interaction terms between policy and year, the effects of paid leave and family allowances became slightly weaker, as the direction of the interaction term is opposite to that of the 'main' term (but note that they were statistically insignificant to begin with).

The second column in Table 6.4 presents the multilevel model for change with earnings inequality among women as the dependent variable. The results suggest that in 1995 in societies with extensive leave earnings inequalities among women were smaller than in countries without such leave policies. High expenditure on family allowances was associated with higher earnings inequalities among women. Over time, the effects of paid leave and expenditure on family allowances became weaker in explaining earnings inequalities among women.

In the third, and final, model presented in Table 6.4, the dependent variable is the share of women's earnings in total households earnings. The estimate for the leave scale (.06 and statistically significant) suggests that in 1995 in countries with extensive paid leave, women's share in household earnings was greater than in countries without extensive paid leave. The interaction between year and the effect of leave, however, indicates that the effect of leave declined over time. Over time differences in leave arrangements between countries became less im-

portant in explaining cross-national variation in the share of women's earnings in total household earnings. We interpret this as a ceiling effect, similar to that reported by Harkness (2013): countries in which women contributed a large share to total household earnings - as a result of the available paid leave - had relatively little room for higher female employment and earnings. This is exemplified by the Nordic countries, in which leave policies already were extensive the 1980s and women's earnings typically contributed between 40% and 50% of total household income. With respect to expenditure on family allowances, the results suggest that in 1995 in societies with high levels of expenditure on family allowances the share of women's earnings in total household earnings was lower than in countries without such expenditure on family allowances. The effect of the expenditure on allowances declined over time, as indicated by the interaction term between year and expenditure.

With these results we were not able to reject the reconciliation policy hypothesis (part b.) or the financial support policy hypothesis (part b.), except for the hypothesised effects of both types of family policies on the correlation between spouses' earnings.

Influential data

The analyses presented in this Chapter were based on observations from only 18 countries. Moreover, in Chapter 5 we identified several inconsistencies in the descriptive Figures (5.4, 5.2, and 5.3) on the dependent variables used in our analyses. We therefore evaluated the model that was central to our conclusion for the presence of influential data: the attenuating effect of women's earnings on between-household inequality (Model II in Table 6.3). The procedures for detecting influential data in multilevel models are detailed in Appendix A.

Two countries showed overly high levels of influence: Italy and Belgium. However, the deletion of these countries from our data did not result in a change in the conclusions, nor in the direction and significance of the regression parameters. We thus conclude that our findings are not biased by the presence of influential data.

Table 6.4: Multilevel Model for Change regressing (I. CORRELATION) correlation between spouses' earnings on Family Policies, (II. INEQUALITY) inequality among women, and (III. SHARE) women's share in household earnings on Family Policies.
Source: Data from the Luxembourg Income Study (LIS, 2013) and the Comparative Family Policy Database (Gauthier, 2010).

	I. Correlation		II. Inequality (×10)		III. Share (×100)	
	B	S.E.	B	S.E.	B	S.E.
Fixed Effects						
Intercept (γ_{00})	20.89	15.72	0.51	0.87	17.58*	2.94
Year (γ_{10})	0.21	0.23	0.02	0.03	0.24	0.15
Paid Leave	0.03	0.06	−0.01*	0.00	0.06*	0.03
Family Allowance	−0.56	1.85	0.30*	0.14	−2.45*	0.73
Year * Paid Leave	−0.01*	0.00	0.00*	0.00	−0.01*	0.00
Year * Family Allowance	0.27*	0.14	−0.06*	0.02	0.26*	0.09
Random Effects (SD)						
Residual (ϵ_{ij}), $N_{lvl1} = 99$	4.40		0.35		2.03	
Intercept (ζ_{0i}), $N_{lvl2} = 18$	10.18		0.05		3.05	
Random Slope Year (ζ_{1i}), $N_{lvl2} = 18$	0.27		0.55		0.31	

* $P < .05$

Results controlled for unemployment and female/male wage ratio.

6.5 CONCLUSION AND DISCUSSION

Women's earnings attenuate earnings inequality between households. This attenuating effect was stronger in societies with extensive reconciliation policies but weaker in societies with extensive financial support policies for families with children. Reconciliation policies stimulate women's employment, and as a result were found to be positively associated with women's earnings contributing a larger share of total household earnings, and with lower earnings inequality among women. The latter results from the fact that with higher female labour force participation, fewer women have zero earnings. Both these findings are in line with Stier et al. (2001), who found that family policies supporting the employment of mothers were associated with lower wage penalties for women who temporarily discontinued employment for childbirth.

Financial support policies were found to have outcomes opposite to those of reconciliation policies. In societies with extensive financial support policies women's earnings contributed a smaller share of total household earnings, and there was higher inequality among women.

The findings of this Chapter pertain to relative inequality. It should thus be pointed out that as long as a positive correlation between spouses' earnings exists, the absolute differences between the richest and poorest households are likely to increase in societies with an institutional context that facilitates women's employment. Also, our findings only apply to coupled households. Whereas this is common in the literature on inequality decomposition (Harkness, 2013; Lam, 1997), this also means that our findings do not apply to the increasing number of single-parent families (see: Casey & Maldonado, 2012). Finally, each of our two hypotheses were tested using only a single indicator. Future studies could repeat our analyses using different measures of reconciliation policies and financial support policies for families. We refrained from over-specifying our regression models: the number of available degrees of freedom was already low because the multilevel model for change (Singer & Willet, 2003) requires the specification

of both the independent variable and its interaction and because we sought to control for two labour market characteristics.

Our finding that between 1981 and 2005 family policies became less important in explaining cross-national variation in different aspects of women's employment deserves further discussion. Firstly, this finding shows the relevance of using the multilevel model for change, analysing trends by distinguishing between the starting point and the rate of change. Without this distinction, this ceiling effect could not have been detected. Secondly, this finding corresponds to Harkness (2013), who reported that the potential for increasing the attenuating effect of women's earnings was stronger in countries with relatively low female labour force participation. That finding was based on cross-sectional research, and is now corroborated by our trend analyses that showed that in countries traditionally providing an institutional context facilitating the employment of women women's earnings had a strong attenuating effect but a low increase in that effect over time.

Various authors have pointed towards the unintended consequences of the outcomes of family policies being biased against low-income families and exacerbating between-household inequalities (Cantillon et al., 2001; Lancker & Ghysels, 2012). In Chapter 4 we found that reconciliation policies are more effective for more educated women, and Ghysels & Van Lancker (2011) found low-income and low-educated families were less likely to take up leave. Others have shown, in contrast, that the increased participation of women in the labour market has compressed the earnings distribution among women and among households (Mandel, 2013). To this juxtaposition in the literature, we contribute the findings that despite selective uptake of the benefits of reconciliation policies, in the long run from 1981 to 2005 women's earnings in coupled households have increased and earnings inequality among women has decreased (also see: Cantillon, 2011). Thus, countries with family policy arrangements that facilitate women's employment and earnings, and consequently smaller earnings inequalities *within* households, also contribute to smaller inequalities *between* households.

Part IV

SUMMARY AND CONCLUSION

7

CONCLUSION

In this dissertation we combined institutional and demographic explanations of women's employment and earnings inequality. This led to the insights that reconciliation policies stimulate women's employment by closing the motherhood-employment gap, increase women's earnings, and reduce inequality among women and between households. Overly long childcare leave decreases women's employment, and higher educated women benefit more than lower educated women from (paid) leave. Financial support policies to families with children increase the motherhood-employment gap, reduce women's earnings, and increase inequality among women and between households.

7.1 ANSWERING THE RESEARCH QUESTIONS

Women's employment rates and earnings inequality between households have been rising in recent decades in OECD countries. Explanations of women's employment typically focused either on changing institutional contexts, or on women's changing demographic background. Few studies accounted for both, up to the point that the distinction between these two types of explanation has been referred the polarised (Pettit & Hook, 2005, p. 780). The starting point of this dissertation was the goal to combine institutional and demographic explanations, following a call for such integration (Van der Lippe & Van Dijk, 2002).

7.1.1 The Motherhood-Employment Gap

The first set of questions was answered in Part II and addressed describing and explaining country-variation and trends in the size of the motherhood-employment gap: the degree to which mothers were less likely to be employed than women without children. We discussed the empirical regularity that the country-level correlation between total fertility rates and female labour force participation in OECD countries turned from negative to positive in the period from 1970 to 2000. Several studies interpreted this as indicating that women's employment and motherhood had become easier to combine. In Chapter 1 we argued, however, that studies based on person-level data had yet to substantiate that claim. The first, descriptive, question of this dissertation therefore was:

QUESTION 1 How has the size of the motherhood-employment gap changed between 1975 and 1999 in OECD countries?

Covering 18 OECD countries from 1975 to 1999 in Chapter 2, we found no country in which mothers were more likely to be employed than women without children. We found substantial variation in the size of the motherhood-employment gap between countries, and within countries over time. We observed that the size of the motherhood-employment gap decreased in for instance Ireland and The Netherlands. Mothers became relatively less likely to be employed compared with women without children in West Germany and Portugal. The motherhood-employment gap was smallest in Denmark and showed no significant trend in that country. The next step was to explain, in Chapter 2, cross-national variation and trends within countries regarding the size of the motherhood-employment gap based on family policies and labour market structure. This was addressed in our second, explanatory, question:

QUESTION 2 To what extent can institutional developments in OECD countries between 1975 and 1999 explain cross-national variation and trends in the size of the motherhood-employment gap in these countries?

We hypothesised that the motherhood-employment gap would be smaller in societies that provide extensive reconciliation policies, and to be smaller in countries with extensive financial support policies to families with children. In general, in Chapter 2 we found these hypotheses corroborated. The motherhood-employment gap was smaller in countries providing longer maternity leave, childcare leave, and continued pay during leave. The motherhood-employment gap was larger in countries providing high levels of family allowance, but not in societies that provided tax benefits to families with children. In addition, we found that women's employment in a society was increased when the labour market was characterised by a large service sector and low unemployment, but that both mothers and women without children benefit equally from these labour market characteristics. In other words, the size of the service sector and the level of unemployment were found not to affect the size of the motherhood-employment gap.

The third research question was answered in Chapter 3:

QUESTION 3 To what extent was the motherhood-employment gap larger between 1975 and 1999 in OECD countries providing long-term childcare leave than in countries providing short-term leave?

The question whether there is such a thing as too long childcare leave has been addressed before, but the answers to this question were inconclusive. The employment of mothers in countries with moderate durations of childcare leave was found to be higher than in countries with no leave at all, but the motherhood-employment gap was found to be bigger in countries with very long periods of leave than with moderate periods of leave.

In the final chapter (4) of Part II we challenged the assumption that all mothers respond similarly to the two types of family policies discerned between in this dissertation:

QUESTION 4 To what extent did the outcomes of reconciliation policies and financial support policies on the size of the motherhood-employment gap differ between more and less educated women in OECD countries between 1980 and 1999?

We found that the size of the motherhood-employment gap is bigger for higher educated women than it is for the lower educated. We did not find that family allowances affected lower educated mothers differently from the higher educated. However, even though higher educated mothers are more likely to be employed than lower educated women without children, we found that higher educated mothers respond more strongly than lower educated mothers to the availability of reconciliation policies. We concluded that reconciliation policies thus further increase the inequality in employment between higher and lower educated women *within* societies.

7.1.2 Earnings Inequality Within and Between Households

In Part III of this dissertation we addressed questions pertaining to the earnings inequality within and between households. Studies had shown that women's earnings tend to attenuate the inequality between coupled households, even though the earnings of spouses are positively correlated. However, it was not known how trends in this attenuating contribution of women's earnings varied across OECD countries. Using harmonised data from the Luxembourg Income Study (LIS, 2013) were able to answer questions about trends:

QUESTION 5 To what extent have women's earnings attenuated earnings inequalities *between* households in 18 OECD countries from 1981 to 2005?

In Chapter 5 we first explained how it is logically possible that women's earnings have a strong tendency to attenuate inequality between coupled households, even though spouses' earnings are positively correlated. Next, we found that the attenuating contribution of women's earnings to inequality between coupled households increased in OECD countries between 1981 and 2005. The findings also suggested that women's earnings had a stronger attenuating effect on household inequalities in those countries with strong female labour force participation. This was further examined to answer:

QUESTION 6 To what extent have changes in (a.) the correlation between spouses' earnings, (b.) earnings inequality among women, and (c.) women's share in total household earnings, affected the degree to which women's earnings attenuate household-level inequality in OECD countries between 1981 and 2005?

Over time, in OECD countries, women's earnings constituted an increasingly large share in total household earnings, thereby decreasing the within-household earnings inequalities in coupled households. Furthermore, we showed that the correlation between spouses' earnings only moderately increased or remained stable over time, while earnings inequality among women decreased. The trend of women's earnings increasingly attenuating the inequality between households was found to be mainly driven by decreasing inequality among women. If inequality among women had not declined as it did in recent decades, inequality between households would have been 25% higher than it actually was in 2005.

Finally, in Chapter 6, we once more tested our reconciliation policy hypothesis and our financial support policy hypothesis, against data on earnings inequality. We did so at the country-level using person-level data, combined with measurements of family policies. We were able to explain cross-national variation and trends in the degree to which women's earnings attenuate inequality between coupled households, answering the question:

QUESTION 7 To what extent can cross-national variation in the degree to which women's earnings attenuate inequalities between households in 18 OECD countries from 1981 to 2005 be explained by (a.) reconciliation policies and (b.) financial support policies?

We found that in societies with extensive leave policies the share of women's earnings in total household earnings was higher, the correlation between spouses' earnings moderately higher, and earnings inequality among women's earnings to be lower. As a result, in countries with extensive paid leave, the resulting women's earnings had a stronger attenuating effect on between-household inequality. Financial support to families had the opposite outcomes, by reducing women's

earnings and increasing inequality among women. We found that the variation between countries in the availability of family policies was more important to explain cross-national variation in these aspects of women's employment in 1981 than in 2005. We argue that this results from a ceiling-effect in the outcomes of family policies in OECD countries on women's employment. Nevertheless, our conclusion is that family policy arrangements that facilitate women's employment not only contribute to smaller inequalities among women and *within* households, but also *between* households.

7.2 DIRECTIONS FOR FUTURE RESEARCH

7.2.1 Data for country-comparative research

This dissertation has been an exercise in combining explanations of women's employment at the levels of the country, the person and the household. Answering the country-comparative questions using person-level data has proven to have very clear advantages. Of course, increasing the scope of the study in terms of trend or number of countries inevitably reduced that availability of comparable measures (Van der Lippe & Van Dijk, 2002). When appropriate measures become available that are cross-nationally comparable and cover trends, it can be further examined how changes in factors other than those accounted for here, have affected women's employment. We mention four such factors.

At the country-level, the reconciliation policy hypothesis could also be tested using a measure on the availability of publicly supported childcare services. Such measure was not available in the database on family policies we used throughout this dissertation (Gauthier, 2010), and in other databases were only available cross-sectionally (Boeckmann, Budig, Misra, Evertsson, Gauthier, Gerstel, Gornick, Grunow, & Klüsener, 2012). Clearly, if trend data on childcare services are collected, future studies could evaluate how trends in childcare availability affected women's employment, and particularly replicate our

testing of the reconciliation policy hypothesis in Chapters 2, 4, and 6. In cross-sectional data (Boeckmann et al., 2012) the availability of parental leave and childcare services were found to be positively correlated across countries, suggesting that our reconciliation policy hypothesis would also hold when tested with childcare services, but this should be tested when data becomes available.

At the household level, our results either accounted for the presence of a partner in the household, or they were based on coupled households only. The focus of our research questions was, however, on the employment and earnings of women. Future studies could study country-variation in men's employment, and especially that of fathers, and attempt to explain this variation with family policy arrangements. It is, however, unlikely that more explicitly accounting for men would have changed the outcomes of our analyses, as the uptake of, for instance, childcare leave is very low among men (Gornick & Meyers, 2003; Kotsadam & Finseraas, 2013; Lappegard, 2012). Of course, this low uptake of leave among fathers suggests that future studies could seek explanations (based on institutional and demographic determinants) of why men's employment is not affected by family policies.

At the person-level, future work could provide a stronger test of our theory by directly measuring women's interest in employment. This is particularly important since women's interests, but also preferences and attitudes, have changed over time and hence may provide an alternative explanation for women's rising labour force participation. Budig et al. (2012) found that the effect of leave on women's employment was similar to that reported in this study, even while accounting for a variable representing societal norms about women's employment. Future studies could test whether changing interests in employment over time have affected the outcomes of family policies.

Finally, although we accounted for explanations at the levels of the country, the person, and the household, future research could address explanations at another important level: the employer. The importance of the employer in facilitating the combination of motherhood and employment is increasingly recognised (Abendroth, Van der Lippe, & Maas, 2012; Abendroth & Den Dulk, 2011; OECD, 2011). This provides

ample opportunity to study how cross-national trends in employer practices and the interplay between national policies and employer work-family arrangements have affected women's labour force participation.

7.2.2 Event-History Modeling

The data used in this dissertation were cross-sectional in nature. As a result, no causality at the person-level could be inferred from the analyses. We could not infer about the direction of causality between motherhood and employment, and only observe in what direction and to what extent family policies have affected the size of the motherhood-employment gap. The future use of event-history models could shed light on the question whether mothers were less likely to be employed, employed women were less likely to become a mother, or both. Using event-history models it can also be determined how family policies affects both women's employment and fertility decisions, by modeling these decision processes simultaneously (Steele, Goldstein, & Browne, 2004; Harrell Jr., 2001; Yamaguchi, 1991). As the data used for event-history models commonly are collected retrospectively, however, these do not allow for making inferences about population-level change over time and thus would not have allowed us to answer our research questions.

7.2.3 Changing Family Formation

In the final section of Chapter 6 it was already discussed how the findings on how family policies affected the inequality between households apply to *coupled* households only. This suited the goal of that Chapter 6, as well as that of Chapter 5. Future studies could study how family policies affect the increasing number of single-parent families (that by a majority are headed by women) (Casey & Maldonado, 2012; Kollmeyer, 2013). As women's employment and earnings increased, decline of marriage, and rising numbers of single-parent families are all related demographic developments, a next step would be to inves-

tigate how current family policy arrangements affect this specific type of households.

7.3 DISCUSSION

Combining institutional and demographic explanations of women's employment and earnigns inequality has brought forward the study of both inequality and policy outcomes. In addition to the advantages discussed before, we address three key innovations in terms of the questions we answered, the theory we tested, and the data used.

7.3.1 Macro-to-Micro Questions

We answered questions about women's employment and earnings, and thereby we answered questions about inequality. To address the combination of institutional and demographic explanations, we answered 7 macro-to-micro questions. These questions not only pertained to different inequalities, but also to the relationships between these inequalities.

First of all, we used person-level data to answer questions about differences in women's employment and earnings between countries, and trends within countries. By combining our person-level data with measurements of the policy context, we found in Chapter 2 that the institutional context could better explain differences in employment across countries, whereas women's demographic background was found to explain individual differences in the employment of women within countries.

Secondly, by simultaneously studying institutional and demographic factors, we could better address institutional determinants of inequalities within countries. We found that the motherhood-employment gap was smaller in countries providing reconciliation policies (although too long childcare leave increased the motherhood-employment gap) and smaller in countries with extensive financial support policies. Furthermore, we found that reconciliation policies

reduced the motherhood-employment gap more effectively among higher educated women than among lower educated, thereby increasing educational differences in women's employment within countries.

Finally, in our analyses of earnings inequality (Part III) we showed how different inequalities in earnings are related to each other. We showed how in coupled households women had fewer earnings then men. In addition, within countries, earnings were also distributed unequally among women. Between 1981 and 2005, as women's employment increased, women's share in total household earnings rose, and because over time fewer women had no earnings the inequality among women declined. It was shown that reconciliation policies increased women's share in total household earnings and decreased inequality among women, whereas financial support policies had the opposite effect of reducing women's earnings and boosting inequality among women. Furthermore, it was shown that as family policies facilitate women's employment and earnings, this also reduced both inequality among women and between households. This means that if family policy arrangements achieve equality among women and between men and women, this also reduces inequality between households.

The answers to our research questions are of clear relevance to policy makers, for three reasons. First, we showed that not all family policies are alike, and that reconciliation policies have opposite effects to those of financial support policies. That means that if both type of policies are implemented, these can (partially) cancel each others' effects. Secondly, it was shown that policy interventions affect inequalities within countries: family policies do not affect people in a homogeneous manner. In other words, as phrased in Chapter 4, family policy outcomes are socially stratified. Family policy intervention will only be effective if tailored to people's interests. Thirdly, as inequalities were shown to be related to each other, attempting to reduce an inequality with policy interventions can result in other inequalities. For instance (and this is related to the second point), implementing reconciliation policies to reduce the inequality in employment between mothers and women without children was shown to be increasing the inequality between the higher and lower educated. On the other hand, with respect to

earnings inequality among women and between households, it was shown that possible stratified outcomes of family policies were outweighed by how family policies facilitated the overall trend towards higher female labour force participation.

7.3.2 Theory

Our hypotheses were derived from our rational action theory that was based on new home economics theory. As was discussed in Chapter 1, applications of rational choice theory have often been predisposed to formulating explanations based on "social structural determinants" (Hechter & Kanazawa, 1997, p. 193), and have paid considerable less attention to personal motivational factors. To be able to combine institutional and demographic explanations of women's employment in a single theory, we thus extended new home economics based on the concepts of opportunities and interests. *Opportunities* referred to the social structural determinants of women's employment and *interests* to the whole of the reasons why women may be motivated to seek employment. By further assuming that higher educated women have a stronger interest in employment, we could formulate our hypothesis that opportunities provided by reconciliation policies will most strongly affect mothers with a strong interest in employment (i.e. the higher educated), and the opportunities of financial support policies only affect those with a weaker interest in employment (i.e. the lower educated).

Based on Lakatosian philisophy of science, the interaction between opportunities and interests means that we must reject explanations of women's employment solely based on the concept of opportunities. Lakatos argued that: "*For the sophisticated falsificationist a scientific theory T is falsified if and only if another theory T′ has been proposed with the following characteristics: (1) T′ has excess empirical content over T: that is, it predicts novel facts, that is, facts improbable in the light of, or even forbidden, by T; (2) T′ explains the previous success of T, that is, all the unrefuted content of T is included (within the limits of observational error) in the content of T′; and (3) some of the excess content of T′ is corroborated.*" (Lakatos, 1978, p.

30, also see: Levels & Nieuwenhuis, 2011). Based on the combination of opportunities and interests we (1.) could derive new hypotheses on how family policies affect the motherhood-employment gap differently among higher educated women than among lower educated women, we (2.) could still explain the opportunity-based results on family policy outcomes, and we (3) found empirical corroboration for our newly derived hypotheses. We emphasize that we only reject *applications* of rational choice theory that are solely based on opportunities, as in rational choice theory itself this interplay between structural and personal factors has been formulated before: "*Rational choice theorists regard both individual values and structural elements as equally important determinants of outcomes, but for methodological reasons their empirical applications typically place greater emphasis on social structural determinants.*" (Hechter & Kanazawa, 1997, p. 193)

7.3.3 Method

Answering questions pertaining to trends and comparisons between countries using person-level data would not have been possible without the continued collection and harmonisation of data, both at the person-level and at the country-level. We used two comparative person-level datasets. The first person-level dataset we used was the Comparative Motherhood-Employment Gap Trend File that was combined from several pre-existing datasets. This dataset allowed us to analyze data on 192,484 individual women, and to cover 305 country-years from 1975 to 1999 in 18 OECD countries. The second person-level dataset we used in this dissertation was the Luxembourg Income Study (LIS, 2013). The data covered 99 country-years in 18 OECD countries, and 1,114,444 person-level observations in 572,222 households. Both person-level datasets were combined with time-varying measures of family policy context and labour market structure, mainly obtained from Comparative Family Policy Database (Gauthier, 2010).

Using the person-level data to answer country-level questions required the development of two sets of statistical tools for country-comparative research. The first, a statistical package for the detection

of influential data in multilevel regression models, was used throughout this dissertation. The development and evaluation of a procedure for 'netting down' gross earnings data allowed us to make comparisons of earnings across the large number of datasets from the Luxembourg Income Study.

The availability of person-level, household-level, and country-level data that was comparable across countries and allowed us to analyze trends within countries, was invaluable in answering our research questions. Because of the prolonged collection of such data covering longer periods of time, the data allowed us to answer questions about how *changes* in family policies and labour market structure affected *changes* in women's employment. Secondly, the repeated observations of countries increased the amount of variability in both institutional context as well as in aspects of women's employment, as well as the number of observations at the country-year level. This allowed us to account for more institutional variables simultaneously.

The availability of multiple datasets allowed us to test our hypotheses using different measurements of both family policies and using different aspects of women's employment and earnings. This resulted in consistent findings, lending support for the idea that our conclusions are not dependent upon the selection of our measurements. For instance, reconciliation policies were represented by measurements of specific policies in Chapter 2, and using an index of the total of paid leave in Chapters 4 and 6. Our measurement of family allowance in Chapter 2 represented the absolute level of allowance, whereas in Chapters 4 and 6 it represented government expenditure on family allowances. The results regarding these different measures of family policies were very similar. Even though in Chapter 4 our financial support policy hypothesis had to be rejected, the results in Table 4.4 did show that high levels of government expenditure on family allowances increased the motherhood-employment gap among all women (but not particularly for lower educated women, as was hypothesised). Finally, our conclusion that reconciliation policies facilitate women's employment while financial support policies reduce women's employment

were found to hold when tested against the size of the motherhood employment gap and when tested against women's earnings.

7.4 CONCLUSION

To conclude, we have shown in this dissertation that family policies have been effective in shaping women's employment. The overall pattern in OECD countries was found to have been a decreased gap in employment between mothers and women without children, and of increased women's employment. The increased women's earnings that were associated with this increased employment attenuated inequalities of earnings between households. Reconciliation policies have the goal of facilitating the employment of mothers, and indeed result in a smaller gap in employment between mothers and women without children. Financial support policies to families have the goal to reduce poverty, but in this study were shown to also result in mothers being less likely to be employed. Two additional outcomes of reconciliation policies were found: overly long childcare leave negatively affects the employment of mothers and higher educated mothers benefit more from paid leave than the lower educated. Thus, a societal context that stimulates the employment of women thus also helps reducing inequality between households.

The central research question of this dissertation was:

CENTRAL RESEARCH QUESTION To what extent can (a.) women's employment and (b.) the contribution of women's earnings to inequality between households be explained based on the *combination* of institutional and demographic factors, in OECD countries between 1975 and 2005?

It has long been identified that women are less likely to be employed when they are a mother, lower educated, and living in a partnered household. Reconciliation policies were identified before to stimulate women's employment, and some studies found that financial support policies to families suppress women's employment. The findings of

this dissertation, however, could not have been reached by treating these two types of explanation of women's employment separately, as the outcomes of the demographic explanations were found to depend on the family policy context in a society, and the outcomes of reconciliation policies were found to depend on women's level of education. Without the distinction between the earnings inequality among women, and the earnings of the household she is part of, it could not have been identified how women's earnings attenuate the inequality between households. Nor could the outcomes of family policies on this attenuating contribution of women's earnings have been identified. Thus, our conclusions could not have been drawn without using the combination of institutional and demographic explanations of women's employment.

8

NEDERLANDSTALIGE SAMENVATTING

In dit proefschrift worden institutionele en demografische verklaringen voor de arbeidsparticipatie van vrouwen en voor inkomensongelijkheid met elkaar gecombineerd. Institutionele verklaringen voor het werken van vrouwen verwijzen naar verschillen tussen landen wat betreft gezinsbeleid en arbeidsmarkt. Demografische verklaringen verwijzen naar individuele eigenschappen van vrouwen zoals moederschap, opleidingsniveau en het hebben van een partner. Eerder onderzoek heeft slechts in zeer beperkte mate institutionele en demografische verklaringen gecombineerd, tot op het punt waarop het onderscheid tussen beide typen verklaringen werd beschreven als gepolariseerd (Pettit & Hook, 2005, p. 780). Het doel van dit proefschrift is het combineren van institutionele en demografische verklaringen, deels in antwoord op een oproep van Van der Lippe & Van Dijk (2002).

Zowel de arbeidsparticipatie van vrouwen als de inkomensongelijkheid tussen huishoudens zijn de afgelopen decennia in lidstaten van de OESO (Organisatie voor Economische Samenwerking en Ontwikkeling) gestegen. Dit was de aanleiding tot het beantwoorden van de volgende hoofdvraag:

HOOFDVRAAG In welke mate kunnen (a.) arbeidsparticipatie van vrouwen en (b.) de bijdrage van het inkomen van vrouwen aan de ongelijkheid tussen huishoudens in OESO-landen tussen 1975 en 2005 worden verklaard op basis van de *combinatie* van institutionele en demografische factoren?

De combinatie van institutionele en demografische verklaringen heeft nieuwe inzichten opgeleverd. Gezinsbeleid dat is gericht op het

combineren van werk en gezin verkleint de mate waarin moeders minder werken dan vrouwen zonder kinderen: door ons de moeder-werk discrepantie genoemd. Ook draagt dergelijk beleid bij aan hogere inkomsten van vrouwen, en minder inkomensongelijkheid tussen huishoudens. Te lang verlof reduceert overigens de arbeidsdeelname van vrouwen, en hoger opgeleide vrouwen bleken meer baat te hebben bij betaald verlof dan lager opgeleide vrouwen. Beleid gericht op financiële ondersteuning van gezinnen met kinderen vergroot de moeder-werk discrepantie, verlaagt de inkomens van vrouwen, en draagt zo bij aan een grotere inkomensongelijkheid tussen huishoudens.

8.1 DE MOEDER–WERK DISCREPANTIE

De eerste set deelvragen wordt beantwoord in deel II van dit proefschrift. Deze vragen hebben betrekking op het beschrijven en verklaren van de omvang van de moeder-werk discrepantie. Eerdere studies lieten zien dat de correlatie tussen de arbeidsparticipatie en het kindertal van landen positief werd tussen 1970 en 2000, maar wij beargumenteren in Hoofdstuk 1 dat dit niet betekent dat de moeder-werk discrepantie kleiner is geworden. Om dat laatste vast te kunnen stellen, is het nodig om gegevens over individuele personen te analyseren. Dat hebben we in dit proefschrift gedaan. De eerste, beschrijvende, vraag in dit proefschrift luidt daarom:

DEELVRAAG 1 Hoe groot was de moeder-werk discrepantie in OESO landen tussen 1975 en 1999?

In een analyse van gegevens over 192.484 individuele vrouwen, uit 18 OESO-landen en uit de periode van 1975 tot en met 1999, vinden we geen enkel land waarin moeders vaker werkten dan vrouwen zonder kinderen. Wel vinden we veel variatie in de omvang van de moeder-werk discrepantie, zowel tussen landen als binnen landen over de tijd. We zien dat de omvang van de moeder-werk discrepantie afnam in bijvoorbeeld Ierland en Nederland. In (voormalig) West-Duitsland en Portugal kwam het werken van moeders juist minder vaak voor in ver-

gelijking met vrouwen zonder kinderen. De omvang van de moeder-werk discrepantie was het kleinst in Denemarken en veranderde daar niet door de tijd heen.

De volgende stap in hoofdstuk 2 is het verklaren van deze variatie. We formuleren daartoe onze tweede, verklarende, deelvraag:

DEELVRAAG 2 In welke mate kunnen institutionele ontwikkelingen in OESO-landen tussen 1975 en 1999 trends in - en verschillen tussen - deze landen verklaren in de omvang van de moeder-werk discrepantie?

We formuleerden de hypothesen dat de moeder-werk discrepantie kleiner is in landen met uitgebreid gezinsbeleid gericht op het combineren van werk en gezin en dat deze groter is in landen met uitgebreid beleid ter financiële ondersteuning van gezinnen met kinderen. Op hoofdlijnen vinden we ondersteuning voor deze hypothesen. De moeder-werk discrepantie was kleiner in landen met langere perioden van zwangerschapsverlof, ouderschapsverlof en met een hogere mate van doorbetaling van loon tijdens verlof; de moeder-werk discrepantie is daarentegen groter in landen met hoge kinderbijslag. Belastingvoordelen voor gezinnen met kinderen bleken geen effect te hebben. Daarnaast vinden we dat de arbeidsparticipatie van vrouwen hoger was in landen met een grote dienstensector en lage werkloosheid; dit geldt in dezelfde mate geldt voor moeders en voor vrouwen zonder kinderen. Deze kenmerken van de arbeidsmarkt hadden dus geen invloed op de omvang van de moeder-werk discrepantie.

De derde deelvraag wordt beantwoord in hoofdstuk 3:

DEELVRAAG 3 In welke mate was de moeder-werk discrepantie tussen 1975 en 1999 groter in OESO-landen met erg lang ouderschapsverlof, vergeleken met landen met kort ouderschapsverlof?

De vraag of ouderschapsverlof ook te lang kan zijn, werd al in eerder onderzoek opgeworpen, maar de antwoorden op deze vraag waren tot op heden onbevredigend. Onze bevindingen laten zien dat moeders in landen met een middellange termijn van verlof vaker werken dan in landen zonder verlof. Echter, in landen met zeer lange perioden

van verlof bleken moeders minder vaak te werken dan in landen met middellange perioden van verlof.

In het laatste hoofdstuk (4) van deel II stellen we onze eigen aanname dat alle moeders in dezelfde mate baat hebben bij de gelegenheden die worden geboden door gezinsbeleid ter discussie:

DEELVRAAG 4 In welke mate verschilde de invloed van gezinsbeleid gericht op het combineren van werk en gezin en gericht op financiële ondersteuning van gezinnen met kinderen tussen hoger en lager opgeleide vrouwen?

We vinden dat de moeder-werk discrepantie groter is voor hoger opgeleide vrouwen dan voor lager opgeleide vrouwen. Gezinsbeleid gericht op financiële ondersteuning van gezinnen met kinderen blijkt wel de omvang van de moeder-werk discrepantie te verkleinen, maar dit effect blijkt even sterk voor zowel hoog als voor laag opgeleide vrouwen. Beleid gericht op het combineren van werk en gezin blijkt echter een sterker effect te hebben voor hoog opgeleide vrouwen dan voor laag opgeleide vrouwen. We concluderen dat dergelijk beleid dus weliswaar ongelijkheid tussen moeders en vrouwen zonder kinderen verkleint, maar ongelijkheid tussen laag en hoog opgeleide vrouwen vergroot.

8.2 INKOMENSONGELIJKHEID BINNEN EN TUSSEN HUISHOUDENS

In deel III van dit proefschrift beantwoorden we vragen met betrekking tot de inkomensongelijkheid *binnen* en *tussen* huishoudens. Eerder onderzoek liet zien dat de inkomens van vrouwen bijdragen aan minder ongelijkheid tussen huishoudens, ondanks het feit dat de inkomens van partners positief samenhangen. Er is echter minder bekend over de mate waarin deze bijdrage verschilde tussen landen van de OESO. Door gebruik te maken van gegevens over meer dan een miljoen personen in data van de Luxembourg Income Study (LIS) kunnen we de volgende deelvragen over trends beantwoorden:

DEELVRAAG 5 In welke mate hebben de inkomens van vrouwen bijgedragen aan een lagere ongelijkheid tussen huishoudens in 18 OESO-landen van 1981 tot en met 2005?

In hoofdstuk 5 leggen we eerst uit hoe het logisch mogelijk is dat de inkomsten van vrouwen een sterke neiging hebben om de inkomensongelijkheid tussen huishoudens te verlagen, zelfs terwijl de inkomens van partners positief gecorreleerd zijn. Vervolgens vinden we dat die egaliserende bijdrage van het inkomen van vrouwen in OESO-landen sterker is geworden tussen 1981 en 2005. Verder blijkt dat hoe hoger de arbeidsparticipatie van vrouwen in OESO landen is, hoe sterker het inkomen van vrouwen bijdraagt aan minder ongelijkheid tussen huishoudens. Om deze trend beter te begrijpen, stellen we de vraag:

DEELVRAAG 6 In welke mate hebben veranderingen in (a.) de correlatie tussen de inkomens van partners, (b.) de inkomensongelijkheid tussen vrouwen, en (c.) het aandeel van vrouwen in het totale huishoudinkomen, een invloed gehad op de mate waarin de inkomens van vrouwen hebben bijgedragen aan een lagere ongelijkheid tussen huishoudens in 18 OESO-landen van 1981 tot en met 2005?

Het aandeel dat vrouwen in OESO-landen tussen 1981 en 2005 bijdragen aan het totale huishoudinkomen blijkt gestegen te zijn, waarmee de ongelijkheid *binnen* huishoudens van stellen minder is geworden. Verder blijkt dat de positieve samenhang tussen de inkomens van partners gelijk bleef of iets toenam, terwijl de inkomensongelijkheid tussen vrouwen fors afnam. De trend dat de inkomens van vrouwen steeds sterker de ongelijkheid tussen huishoudens verminderden bleek voornamelijk voort te komen uit het feit dat de inkomstenongelijkheid tussen vrouwen af is genomen. Als de ongelijkheid tussen vrouwen van 1981 tot en met 2005 niet was afgenomen tussen, dan zou de inkomensongelijkheid tussen huishoudens in 2005 25% hoger zijn geweest.

Ten slotte toetsen we onze hypothesen over gezinsbeleid gericht op het combineren van werk en gezin en over gezinsbeleid gericht op financiële ondersteuning nog eenmaal op basis van gegevens over in-

komensongelijkheid. We doen dit in hoofdstuk 6. We beantwoorden de volgende deelvraag:

DEELVRAAG 7 In welke mate kunnen verschillen tussen landen in de mate waarin het inkomen van vrouwen ongelijkheid tussen huishoudens verlaagt in 18 OESO-landen tussen 1981 en 2005 worden verklaard met (a.) gezinsbeleid gericht op het combineren van werk en gezin en (b.) gezinsbeleid gericht op financiële ondersteuning van gezinnen met kinderen?

Onze analyses laten zien dat in landen met uitgebreid betaald verlof het aandeel van vrouwen in het totale huishoudinkomen groter was, de correlatie tussen de inkomens van vrouwen sterker positief was, en de inkomensongelijkheid tussen vrouwen lager was. Dit had als gevolg dat in landen met uitgebreid betaald verlof de inkomensongelijkheid tussen huishoudens sterker werd verlaagd door de inkomens van vrouwen. In landen met uitgebreid gezinsbeleid gericht op financiële ondersteuning bleken de inkomens van vrouwen een minder sterke egaliserende werking te hebben op de ongelijkheid tussen huishoudens, omdat de inkomens van vrouwen in die landen lager waren en de ongelijkheid tussen vrouwen groter. We concluderen dan ook dat landen, die met beleid de arbeidsparticipatie van vrouwen stimuleren, niet alleen bijdragen aan minder ongelijkheid *binnen* huishoudens, maar ook *tussen* huishoudens.

Part V

APPENDICES

A

INFLUENCE.ME: TOOLS FOR DETECTING INFLUENTIAL DATA IN MULTILEVEL REGRESSION MODELS

`influence.ME` provides tools for detecting influential data in mixed effects models. The application of these models has become common practice, but the development of diagnostic tools has lagged behind. `influence.ME` calculates standardized measures of influential data for the point estimates of generalized mixed effects models, such as DFBETAS, Cook's distance, as well as percentile change and a test for changing levels of significance. `influence.ME` calculates these measures of influence while accounting for the nesting structure of the data. The package and measures of influential data are introduced, a practical example is given, and strategies for dealing with influential data are suggested.

The application of mixed effects regression models has become common practice in the field of social sciences. As used in the social sciences, mixed effects regression models take into account that observations on individual respondents are nested within higher-level groups such as schools, classrooms, states, and countries (Snijders & Bosker, 1999), and are often referred to as multilevel regression models. Despite these models' increasing popularity, diagnostic tools to evaluate fitted models lag behind.

We introduce `influence.ME` (Nieuwenhuis, Pelzer, & Te Grotenhuis, 2012b), an R-package that provides tools for detecting influential cases in mixed effects regression models estimated with `lme4` (Bates et al.,

1 This chapter is based on: Nieuwenhuis, R., Te Grotenhuis, M., Pelzer, B. (2012c). Influence.ME: tools for detecting influential data in mixed effects models *R Journal*, 4(2), 38-47.

2010). It is commonly accepted that tests for influential data should be performed on regression models, especially when estimates are based on a relatively small number of cases. However, most existing procedures do not account for the nesting structure of the data. As a result, these existing procedures fail to detect that higher-level cases may be influential on estimates of variables measured at specifically that level.

In this paper, we outline the basic rationale on detecting influential data, describe standardized measures of influence, provide a practical example of the analysis of students in 23 schools, and discuss strategies for dealing with influential cases. Testing for influential cases in mixed effects regression models is important, because influential data negatively influence the statistical fit and generalizability of the model. In social science applications of mixed models the testing for influential data is especially important, since these models are frequently based on large numbers of observations at the individual level while the number of higher level groups is relatively small. For instance, Van der Meer et al. (2010) were unable to find any country-level comparative studies involving more than 54 countries. With such a relatively low number of countries, a single country can easily be overly influential on the parameter estimates of one or more of the country-level variables.

A.1 DETECTING INFLUENTIAL DATA

All cases used to estimate a regression model exert some level of influence on the regression parameters. However, if a single case has extremely high or low scores on the dependent variable relative to its expected value — given other variables in the model, one or more of the independent variables, or both — this case may overly influence the regression parameters by 'pulling' the estimated regression line towards itself. The simple inclusion or exclusion of such a single case may then lead to substantially different regression estimates. This runs against distributional assumptions associated with regression models,

and as a result limits the validity and generalizability of regression models in which influential cases are present.

The analysis of residuals cannot be used for the detection of influential cases (Crawley, 2007). Cases with high *residuals* (defined as the difference between the observed and the predicted scores on the dependent variable) or with high standardized residuals (defined as the residual divided by the standard deviation of the residuals) are indicated as outliers. However, an influential case is not always an outlier. On the contrary: a strongly influential case dominates the regression model in such a way, that the estimated regression line lies closely to this case. Although influential cases thus have extreme values on one or more of the variables, they can be *onliers* rather than *outliers*. To account for this, the *(standardized) deleted residual* is defined as the difference between the observed score of a case on the dependent variable, and the predicted score from the regression model that is based on data from which that case was removed.

Just as influential cases are not necessarily outliers, outliers are not necessarily influential cases. This also holds for deleted residuals. The reason for this is that the amount of influence a case exerts on the regression slope is not only determined by how well it's (observed) score is fitted by the specified regression model, but also by its score(s) on the independent variable(s). The degree to which the scores of a case on the independent variable(s) are extreme is indicated by the *leverage* of this case. A higher leverage means more extreme scores on the independent variable(s), and a greater potential of overly influencing the regression outcomes. However, if a case has very extreme scores on the independent variable(s) but is fitted very well by a regression model, and if this case has a low deleted (standardized) residual, this case is not necessarily overly influencing the outcomes of the regression model.

Since neither outliers, nor cases with a high leverage, are necessarily influential, a different procedure is required for detecting influential cases. The basic rationale behind measuring influential cases is based on the principle that when single cases are iteratively omitted from the data, models based on these data should not produce substantially

different estimates. If the model parameters change substantially after a single case is excluded, this case may be regarded as too influential. However, how much change in the model parameters is acceptable? To standardize the assessment of how influential a single case is, several measures of influence are commonly used. First, DFBETAS is a standardized measure of the absolute difference between the estimate with a particular case included and the estimate without that particular case (Belsley, Kuh, & Welsch, 1980). Second, Cook's distance provides an overall measurement of the change in all parameter estimates, or a selection thereof (Cook, 1977). In addition, we introduce the measure of percentile change and a test for changing levels of significance of the fixed parameters.

Up to this point, this discussion on influential data was limited to how single cases can overly influence the point estimates (or BETAS) of a regression model. Single cases, however, can also bias the confidence intervals of these estimates. As indicated above, cases with high leverage can be influential because of their extreme values on the independent variables, but not necessarily are. Cases with high leverage but a low deleted residual compress standard errors, while cases with low leverage and a high deleted residual inflate standard errors. Inferences made to the population from models in which such cases are present, may be incorrect.

A.1.1 Detecting Influential Data in Mixed Effects Models

Other options are available in R that help evaluating the fit of regression models, including the detection of influential data. The base R installation provides various plots for regression models, including but not limited to plots showing residuals versus the fitted scores, Cook's distances, and the leverage versus the deleted residuals. The latter plot can be used to detect cases that affect the inferential properties of the model, as discussed above. These plots, however, are not available for mixed effects models.

The `LMERConvenienceFunctions` package provides model criticism plots, including the density of the model residuals and the fitted values

versus the standardized residuals (Tremblay, 2012). However, while this package works with the `lme4`, it only is applicable for *linear* mixed effects models.

The `influence.ME` package introduced here contributes to these existing options, by providing several measures of influential data for *generalized* mixed effects models. The limitation is that, unfortunately, as far as we are aware, the measure of leverage was not developed for generalized mixed effects models. Consequently, the current instalment of `influence.ME` emphasizes detecting the influence of cases on the point estimates of generalized mixed effect models. It does, however, provide a basic test for detecting whether single cases change the level of significance of an estimate, and therefore the ability to make inferences from the estimated model.

To apply the logic of detecting influential data to generalized mixed effects models, one has to measure the influence of a particular higher level group on the estimates of a predictor measured at that level. The straightforward way is to delete all observations from the data that are nested within a single higher level group, then re-estimate the regression model, and finally evaluate the change in the estimated regression parameters. This procedure is then repeated for each higher-level group separately.

The `"influence"` function in the `influence.ME` package performs this procedure automatically, and returns an object containing information on the parameter estimates excluding the influence of each higher level group separately. The returned object of class `"estex"` (`EST`imates `EX`cluding the influence of a group) can then be passed on to one of the functions calculating standardized measures of influence (such as DFBETAS and Cook's Distance, discussed in more detail in the next section). Since the procedure of the `"influence"` function entails re-estimating mixed effects models several times, this can be computationally intensive. Unlike the standard approach in R, we separated the estimation procedure from calculating the measures of influence themselves. This allows the user to process a single model once using the `"influence"` function, and then to evaluate it using

various measures and plots of influence.

In detecting influential data in mixed effects models, the key focus is on changes in the estimates of variables measured at the group-level. However, most mixed effects regression models estimate the effects of both lower-level and higher-level variables simultaneously. Langford & Lewis (1998) developed a procedure in which the mixed effects model is modified to neutralize the group's influence on the higher-level estimate, while at the same time allowing the lower-level observations nested within that group to help estimate the effects of the lower-level predictors in the model. For each higher-level unit evaluated based on this method, the intercept-vector of the model is set to 0, and an (additional) dummy variable is added to the model, with score 1 for the respective higher level case. This way, the case under investigation does not contribute to the variance of the random intercept, nor to the effects of variables measured at the group-level. provides this functionality, which is accessed by specifying `delete=FALSE` as an option to the `"influence"` function. As a result of the specific modification of the model-specification, this specific procedure suggested by Langford and Lewis (1998) does not work when factor-variables are used in the regression model.

Finally, also allows for detecting the influence of lower-level cases in the mixed effects model. In social science applications of mixed effects models, with a great number of lower-level observations nested in a limited number of groups, this will not always be feasible. Detecting influence of lower-level observations is supported for applications in various disciplines where mixed effects models are typically applied to only a limited number of observations per group. This procedure is accessed by specifying `obs=TRUE` as an option to the `"influence"` function. The `"influence"` function can either determine the influence of higher-level cases, or of single observations, but not both at the same time.

A.2 THE OUTCOME MEASURES

The `"influence"` function described above returns an object with information on how much the parameter estimates in a mixed effects model change, after the (influence of) observations of higher-level groups and their individual-level observations were removed from it iteratively. This returned object can then be passed on to functions that calculate standardized measures of influence. offers four such measures, which are detailed in this section.

A.2.1 DFBETAS

DFBETAS is a standardized measure that indicates the level of influence observations have on single parameter estimates (Fox, 2002). Regarding mixed models, this relates to the influence a higher-level unit has on the parameter estimate. DFBETAS is calculated as the difference in the magnitude of the parameter estimate between the model including and the model excluding the higher level case. This absolute difference is divided by the standard error of the parameter estimate excluding the higher level unit under investigation:

$$\text{DFBETAS}_{ij} = \frac{\hat{\gamma_i} - \hat{\gamma_{i(-j)}}}{se\left(\hat{\gamma_{i(-j)}}\right)} \tag{A.1}$$

in which i refers to the parameter estimate, and j the higher-level group, so that $\hat{\gamma_i}$ represents the original estimate of parameter i, and $\hat{\gamma_{i(-j)}}$ represents the estimate of parameter i, after the higher-level group j has been excluded from the data.

In , values for DFBETAS in mixed effects models can be calculated using the function `"dfbetas"`, which takes the object returned from `"influence"` as input. Further options include `parameters` to provide a vector of index numbers or names of the selection of parameters for which DFBETAS is to be calculated. The default option of `"dfbetas"` is to calculate DFBETAS for estimates of all fixed effects in the model.

As a rule of thumb, a cut-off value is given for DFBETAS (Belsley et al., 1980):

$$2/\sqrt{n} \tag{A.2}$$

in which n, the number of observations, refers to the number of groups in the grouping factor under evaluation (and not to the number of observations nested within the group under investigation). Values exceeding this cut-off value are regarded as overly influencing the regression outcomes for that specific estimate.

A.2.2 Cook's Distance

Since DFBETAS provides a value for each parameter and for each higher-level unit that is evaluated, this often results in quite a large number of values to evaluate (Fox, 2002). An alternative is provided by Cook's distance, a commonly used measure of influence. Cook's distance provides a summary measure for the influence a higher level unit exerts on *all* parameter estimates simultaneously, or a selection thereof. A formula for Cook's Distance is provided (Snijders & Bosker, 1999; Snijders & Berkhof, 2008):

$$C_j^{0F} = \frac{1}{r+1}\left(\hat{\gamma} - \gamma_{\hat{(-j)}}\right)' \hat{\sum}_F^{-1} \left(\hat{\gamma} - \hat{\gamma}_{(-j)}\right) \tag{A.3}$$

in which $\hat{\gamma}$ represents the vector of original parameter estimates, $\hat{\gamma}_{(-j)}$ the parameter estimates of the model excluding higher-level unit j, and $\hat{\sum}_F$ represents the covariance matrix. In `influence.ME`, the covariance matrix of the model *excluding* the higher-level unit under investigation j is used. Finally, r is the number of parameters that are evaluated, excluding the intercept vector.

As a rule of thumb, cases are regarded as too influential if the associated value for Cook's Distance exceeds the cut-off value of (Van der Meer et al., 2010):

$$\frac{4}{n} \tag{A.4}$$

in which n to the number of groups in the grouping factor under evaluation.

In , values for Cook's distance in mixed effects models are calculated using the function `"cooks.distance"`, which takes the object returned from `"influence"` as input. Further options include `parameters` to provide a vector of index numbers or names of the parameters for which Cook's Distance is to be calculated. In addition, the user can specify `sort=TRUE` to have the values for Cook's distance returned in descending order.

As a final note, it is pointed out that if Cook's distance is calculated based on a single parameter, the Cook's distance equals the squared value of DFBETAS for that parameter. This is also reflected in their respective cut-off values:

$$\sqrt{\frac{4}{n}} = \frac{2}{\sqrt{n}} \tag{A.5}$$

A.2.3 Percentile Change

Depending upon the goal for which the mixed model is estimated (prediction vs. hypothesis testing), the use of formal measures of influence as DFBETAS and Cook's distance may be less desirable. The reason for this is that based on these measures it is not immediately clear to what extent parameter estimates change. For substantive interpretation of the model outcomes, the relative degree to which a parameter estimate changes may provide more meaningful information. A simple alternative is therefore offered by the function `"pchange"`, which takes the

same input-options as the "dfbetas" function. For each higher-level group, the percentage of change is calculated as the absolute difference between the parameter estimate both including and excluding the higher-level unit, divided by the parameter estimate of the complete model and multiplied by 100%. A percentage of change is returned for each parameter separately, for each of the higher-level units under investigation. In the form of a formula:

$$\left(\hat{\gamma} - \hat{\gamma_{(-j)}}\right) \frac{1}{\hat{\gamma}} * 100\% \tag{A.6}$$

No cut-off value is provided, for determining what percent change in parameter estimate is considered too large will primarily depend on the goal for which the model was estimated and, more specifically, the nature of the hypotheses that are tested.

A.2.4 Test for changes in significance

As discussed above, even when cases are not influential on the point estimates (BETAS) of the regression model, cases can still influence the standard errors of these estimates. Although cannot provide the leverage measure to detect this, it provides a test for changes in the statistical significance of the fixed parameters in the mixed effects model.

The "sigtest" function tests whether excluding the influence of a single case changes the statistical significance of any of the variables in the model. This test of significance is based on the test statistic provided by the lme4 package. The nature of this statistic varies between different distributional families in the generalized mixed effects models. For instance, the t-statistic is related to a normal distribution while the z-statistic is related to binomial distributions.

For each of the cases that are evaluated, the test statistic of each variable is compared to a test-value specified by the user. For the purpose of this test, the parameter is regarded as statistically significant if the test statistic of the model exceeds the specified value. The "sigtest" function reports for each variable the estimated test statistic after dele-

tion of each evaluated case, whether or not this updated test statistic results in statistical significance based on the user-specified value, and whether or not this new statistical significance differs from the significance in the original model. So, in other words, if a parameter was statistically significant in the original model, but is no longer significant after the deletion of a specific case from the model, this is indicated by the output of the `"sigtest"` function. It is also indicated when an estimate was not significant originally, but reached statistical significance after deletion of a specific case.

A.2.5 Plots

All four measures of influence discussed above, can also be plotted. The `"plot"` function takes the output of the `"influence"` function to create a dotplot of a selected measure of influence (cf. Sarkar, 2008). The user can specify which measure of influence is to be plotted using the `which=` option. The `which=` option defaults to `dfbetas`. Other options are to select `cook` to plot the cook's distances, `pchange` to plot the percentage change, and `sigtest` to plot the test statistic of a parameter estimate after deletion of specific cases.

All plots allow the output to be sorted (by specifying `sort=TRUE` and the variable to sort on using `to.sort=` (not required for plotting cook's distances). In addition, a cut-off value can be specified using (`cutoff=`). Values that exceed this cut-off value will be plotted visually differently, to facilitate the identification of influential cases. By default, the results for all cases and all variables are plotted, but a selection of these can be made by specifying `parameters=` and / or `groups=`. Finally, by specifying `abs=TRUE` the absolute values of the measure of influence are plotted.

A.3 EXAMPLE 1: STUDENTS IN 23 SCHOOLS

In our example, we are interested in the relationship between the degree of structure that schools attempt to enforce in their classrooms

and students' performance on a math test. Could it be that a highly structured class affects their performance?

The package contains the `school23` data.frame, that provides information on the performance of 519 students in 23 schools. Measurements include individual students' score on a math test, school-level measurements of class structure, and several additional independent variables. Student's class and school are equivalent in this data, since only one class per school is available. These data are a subset of the NELS-88 data (National Education Longitudinal Study of 1988). The data are publicly available on the internet,[1] and were reproduced with kind permission of Ita Kreft and Jan de Leeuw (1998).

First, using the `lme4` package, we estimate a multivariate mixed effects model with students nested in schools, a random intercept, a measurement of individual students' time spent on math homework, and a measurement of class structure at the school level. For the purpose of our example, we assume here that the math, homework, and structure variables were correctly measured at the interval level.

```
library(influence.ME)
data(school23)

school23 <- within(school23,
   homework <- unclass(homework))

m23 <- lmer(math ~ homework + structure
   + (1 | school.ID),
   data=school23)

print(m23, cor=FALSE)
```

This results in the summary of the model based on 23 schools (assigned to object `m23`), as shown on the next page.

1 http://www.ats.ucla.edu/stat/examples/imm/

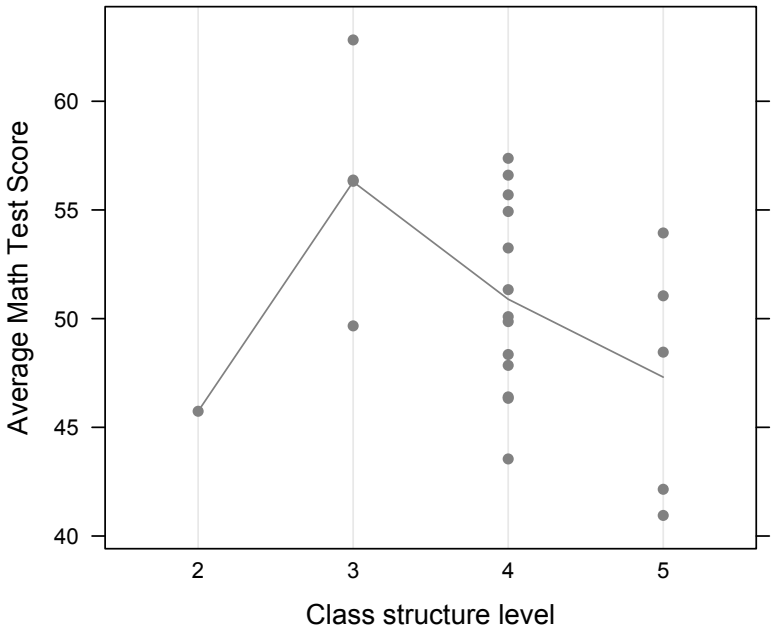
```
Linear mixed model fit by REML
Formula: math ~ homework +
   structure + (1 | school.ID)
   Data: school23
  AIC  BIC logLik deviance REMLdev
 3734 3756  -1862     3728    3724
Random effects:
 Groups    Name        Variance Std.Dev.
 school.ID (Intercept) 19.543   4.4208
 Residual              71.311   8.4446
Number of obs: 519, groups: school.ID, 23

Fixed effects:
            Estimate Std. Error t value
(Intercept)  52.2356     5.3940   9.684
homework      2.3938     0.2771   8.640
structure    -2.0950     1.3237  -1.583
```

Based on these estimates, we may conclude that students spending more time on their math homework score better on a math test. Regarding class structure, however, we do not find a statistically significant association with math test scores. But, can we now validly conclude that class structure does not influence students' math performance, based on the outcomes of this model?

A.3.1 Visual Examination

Since the analysis in the previous section has been based on the limited number of 23 schools, it is, of course, possible that observations on single schools have overly influenced these findings. Before using the tools provided in the package to formally evaluate this, a visual examination of the relationship between class structure and math test performance, aggregated to the school level, will be performed.

Figure A.1: Association Between Class Structure and Math Performance. Source: Authors' calculation on data described in this Appendix




```
struct <- unique(subset(school23,
    select=c(school.ID, structure)))

struct$mathAvg <- with(school23,
    tapply(math, school.ID, mean))

dotplot(mathAvg ~ factor(structure),
    struct,
    type=c("p","a"),
    xlab="Class structure level",
    ylab="Average Math Test Score")
```

In the syntax above, a bivariate plot of the aggregated math scores and class structure is created, which is shown in Figure A.1. In this plot,

it is clear that one single school represented in the lower-left corner of the graph seems to be an outlier, and - more importantly - the non-linear curve shown in this graph clearly indicates this single school with class structure level of 2 may overly influence a linear regression line estimated based on these data.

A.3.2 Calculating measures of influence

In the previous section, based on Figure A.1 we suspected that the combination in one specific school of the low average math test results of students, and the low level of class structure in that school, may have overly influenced the original analysis of our data. However, this only is a bivariate examination of the data, and therefore does not take into account other variables in the model. Hence, in our example, our preliminary conclusion that this may be an influential case is not controlled for possible effects of the homework variable. A better test is provided by standardized measures of influence, as calculated from the regression model rather than from the raw data.

The first step in detecting influential data is to determine the extent to which the parameter estimates in model `m23` change, when iteratively each of the schools is deleted from the data. This is done with the `"influence"` function:

```
estex.m23  <- influence(m23, "school.ID")
```

The `"influence"` function takes a mixed effects regression model as input (here: `m23`), and the grouping factor needs to be specified, which in our case is `school.ID`. We assign the output of the `"influence"` function to an object named `estex.m23`. Below, we use this object as input to the `"dfbetas"` function, to calculate DFBETAS.

```
dfbetas(estex.m23,
    parameters=c(2,3))
```

This results in a substantial amount of output, a portion of which is shown below. Only the DFBETAS for the `homework` and `structure`

variables were returned, since `parameters=c(2,3)` was specified.

```
         homework    structure
6053  -0.13353732 -0.168139487
6327  -0.44770666  0.020481057
6467   0.21090081  0.015320965
7194  -0.44641247  0.036756281
7472  -0.55836772  1.254990963
...
72292  0.62278508  0.003905031
72991  0.52021424  0.021630219
```

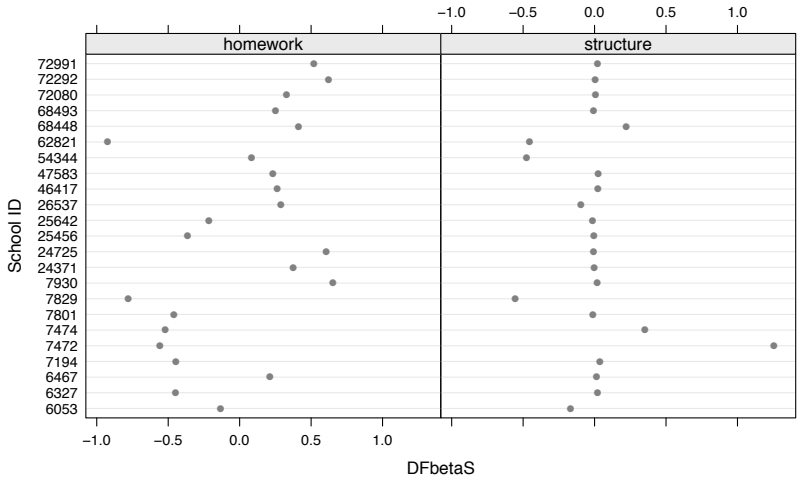
The numerical output given above by the `"dfbetas"` function provides a detailed report of the values of DFBETAS in the model. For each variable, as well as for each nesting group (in this example: each school), a value for DFBETAS is computed and reported upon. The cut-off value of DFBETAS equals $2/\sqrt{n}$ (Belsley et al., 1980), which in this case equals $2/\sqrt{23} = .41$. The estimate for class structure in this model seems to be influenced most strongly by observations in school number 7472. The DFBETAS for the `structure` variable clearly exceeds the cut-off value of .41. Also, the estimates of the `homework` variable changes substantially with the deletion of several schools, as indicated by the high values of DFBETAS.

A plot (shown in Figure A.2) of the DFBETAS is created using:

```
plot(estex.m23,
    which="dfbetas",
    parameters=c(2,3),
    xlab="DFbetaS",
    ylab="School ID")
```

Based on Figure A.2, it is clear that both the `structure` and the `homework` variables are highly susceptible to the influence of single schools. For the `structure` variable this is not all that surprising, since class structure was measured at the school level and shown in Figure A.1 to be very likely to be influenced by a single case: school number

Figure A.2: DFBETAS of Class Structure and Homework.
Source: Authors' calculation on data described in this Appendix



7472. The observation that high values of DFBETAS were found for the `homework` variable, suggests that substantial differences between these schools exist in terms of how much time students spend on average on their homework. Therefore, we suggest that in mixed effects regression models, both the estimates of individual-level and group-level variables are evaluated for influential data.

The measure of Cook's distance allows to determine the influence a single higher-level group has on the estimates of multiple variables simultaneously. So, since the `"cooks.distance"` function allows to specify a selection of variables on which the values for Cook's distance are to be calculated, this can be used to limit the evaluation to the measurements at the group-level exclusively. Note, that whereas DFBETAS always relates to single variables, Cook's distance is a summary measure of changes on all parameter estimates it is based on. Reports on Cook's distance thus should always specify on which variables these values are based.

To continue our example, we illustrate the `"cooks.distance"` function on a single variable, since class structure is the only variable measured at the school-level. In the example below, we use the same object

that was returned from the "influence" function. The specification of this function is similar to "dfbetas", and to create a plot of the cook's distances we again use the "plot" function with the specification which="cook". We specify two additional arguments to augment the figure. First, we specify sort=TRUE to have the resulting Cook's distances sorted in a descending order in the figure. The appropriate cut-off value for Cook's distance with 23 nesting groups equals to $4/23 = .17$. By specifying the cut-off value with cutoff=.17, Cook's distances exceeding the specified value are easily identified in the resulting figure. Thus, to receive both numeric output and a graphical representation (Figure A.3), the following specification of "cooks.distance" and "plot" is given:

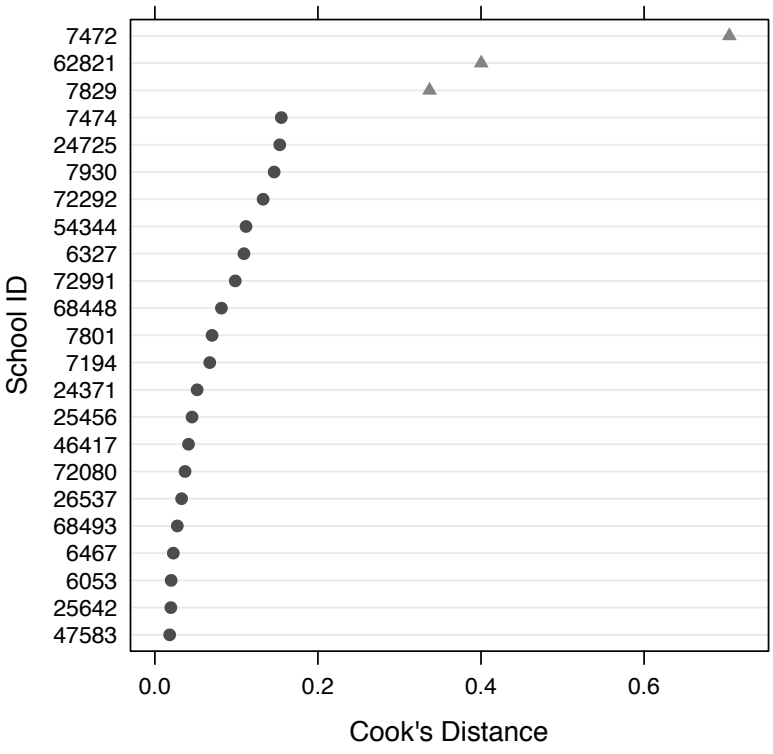
```
cooks.distance(estex.m23,
    parameter=3, sort=TRUE)

plot(estex.m23, which="cook",
    cutoff=.17, sort=TRUE,
    xlab="Cook's Distance",
    ylab="School ID")
```

The output below shows one value of Cook's distance for each nesting group, in this case for each school.

```
              [,1]
24371 6.825871e-06
72292 1.524927e-05
...
54344 2.256612e-01
7829  3.081222e-01
7472  1.575002e+00
```

Only a selection of the output is shown here. A few schools exceed the cut-off value (in Figure A.3 these are indicated with triangles), but one school stands out: 7472. Clearly, this school strongly influences the outcomes regarding the structure variable, as we already suspected based on our bivariate visual examination in Figure A.1.

Figure A.3: Cook's Distance based on Class Structure.
Source: Authors' calculation on data described in this Appendix



A.3.3 Testing for Changes in Statistical Significance (sigtest)

In the example below, the `"sigtest"` function is used to test for changing levels of significance after deletion of each of the 23 schools from our example model. We are specifically interested in the level of significance of the `structure` variable, for which it was already established above that school with number 7472 is very influential. Since we observed a negative effect in the original model, we specify `test=-1.96` to test for significance at a commonly used value (-1.96) of the test statistic. Note that since we estimated a normally distributed model, the test statistic here is the t-value.

```
sigtest(estex.m23, test=-1.96)$structure[1:10,]
```

In the example above, we only request the results for the `structure` variable and for the first 10 schools. In the results presented below, three columns are shown. The first column (`Altered.Teststat`) shows the value of the test statistic (here for the `structure` variable) after the deletion of the respective schools (indicated in the row labels). Especially school number 7472 stands out. In the original model, the test statistic for the structure variable was -1.583, which was not significant. When the influence of school number 7472 is excluded from the model, the test statistic now is -2.72, which exceeds the selected value of -1.96 selected by us. That the structure variable would be significant by deletion of school 7472 is indicated in the second column (`Altered.Sig`). The `Changed.Sig` column finally confirms whether the level of significance of the `structure` variable (which was not significant in the original model) changed to significant after each of the schools was deleted.

In the case of our example, the results for Cook's Distance and the results of this test for changing levels of significance both indicate that school number 7472 overly influences the regression outcomes regarding the school-level `structure` variable. Referring to the discussion on influential data above, however, we emphasize that this is not necessarily always the case. Cases can influence the point estimates without affecting their level of significance, or affect the level of significance

without overly affecting the point estimate itself. Therefore, both tests should be performed.

```
      Altered.Teststat Altered.Sig Changed.Sig
6053         -1.326409       FALSE       FALSE
6327         -1.688663       FALSE       FALSE
6467         -1.589960       FALSE       FALSE
7194         -1.512686       FALSE       FALSE
7472         -2.715805        TRUE        TRUE
7474         -1.895138       FALSE       FALSE
7801         -1.534023       FALSE       FALSE
7829         -1.045866       FALSE       FALSE
7930         -1.566117       FALSE       FALSE
24371        -1.546838       FALSE       FALSE
```

Before, using DFBETAS, we identified several several schools that overly influence the estimate of the `homework` variable. We have there performed `sigtest` test to evaluate whether deletion of any of the schools changes the level of significance of the `homework` variable. These results are not shown here, but indicated that the deletion of none of the schools changed the level of significance of the `homework` variable.

A.3.4 Measuring the influence of lower-level observations

Finally, it is possible that a single lower-level observation affects the results of the mixed effects model, especially for data with a limited number of lower-level observations per group. In our example, this would refer to a single student affecting the estimates of either the individual-level variables, the school-level variables, or both. Here, we test whether one or more individual students affect the estimate of the school-level `structure` variable.

To perform this test, the `"influence"` function is used, and `obs=TRUE` is specified to indicate that single observations (rather than groups) should be evaluated. The user is warned that this procedure often will

be computationally intensive when the number of lower-level observations is large.

Next, we request Cook's Distances specifically for the `structure` variable. Since the number of student-level observations in this model is 519, and cut-off value for Cook's Distance is defined as $4/n$, the cut-off value is $4/519 = .0077$. The resulting output is extensive, since a Cook's Distance is calculated for any of the 519 students. Therefore, in the example below, we directly test which of the resulting Cook's Distances exceeds the cut-off value.

```
estex.obs  <- influence(m23, obs=TRUE)
cks.d <- cooks.distance(estex.obs, parameter=3)
which(cks.d > 4/519)
```

The output is not shown here, but the reader can verify that students with numbers 88 and 89 exert too much influence on the estimate of the structure variable. Using the `sigtest` function, however, showed that the deletion of none of the students from the data affected the level of significance of the structure variable, nor of any of the other variables in the model.

A.4 EXAMPLE 2: TESTING THE LONG-LEAVE HYPOTHESIS ON A CROSS-SECTION OF 15 COUNTRIES

In this section, we present an alternative test of the long-leave hypothesis, that was detailed in Chapter 3. From the data used in that Chapter we took a subset of 15 countries: Austria, Belgium, Canada, Denmark, France, Germany (both the former Eastern Germany and West Germany), Greece, Ireland, Italy, The Netherlands, Portugal, Spain, Sweden, and the United Kingdom. We took the cross-section in the year in which most countries were observed: 1997. Descriptive statistics are presented in Table A.1.

Table A.1: Descriptive Statistics of Motherhood, Employment, and Childcare Leave in the Cross-sectional Dataset.
Source: Comparative Motherhood-Employment Gap Trend File, and the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001).

	M	SD	Min	Max
Employed	53		0.00	1.00
Mother	61		0.00	1.00
Childcare Leave	0	5.67	−4.47	11.13
N.countries	15			
N.countryyears	15			
N.obs	16,890			

A.4.1 Analyses based on the cross-sectional dataset

In this section we test the long-leave hypothesis that in countries with short periods of childcare leave the motherhood-employment gap is smaller than in countries with no childcare leave, but in countries with long periods of childcare leave the motherhood-employment gap is larger than in countries with short periods of leave. .

In Figure A.4 we plotted the bivariate relationship between the number of weeks of childcare leave per country and the degree to which mothers were less likely to be employed than women without children: the motherhood-employment gap. This gap is represented by the log of the odds ratio between motherhood and employment, that was calculated separately for each country. The continuous, curved line represents the LOESS (LOcal RegrESSion) estimate of this association (Fox, 2002). This line clearly is curved, with durations of childcare leave ranging from 0 to approximately 50 weeks being associated with smaller motherhood-employment gaps, but with further increases in this duration (up to 156 weeks in countries such as Germany, Spain and France) being associated with mothers being increasingly less likely to be employed than women without children.

Closer examination of the bivariate association in Figure A.4 shows that several of the observed countries are outliers to the estimated

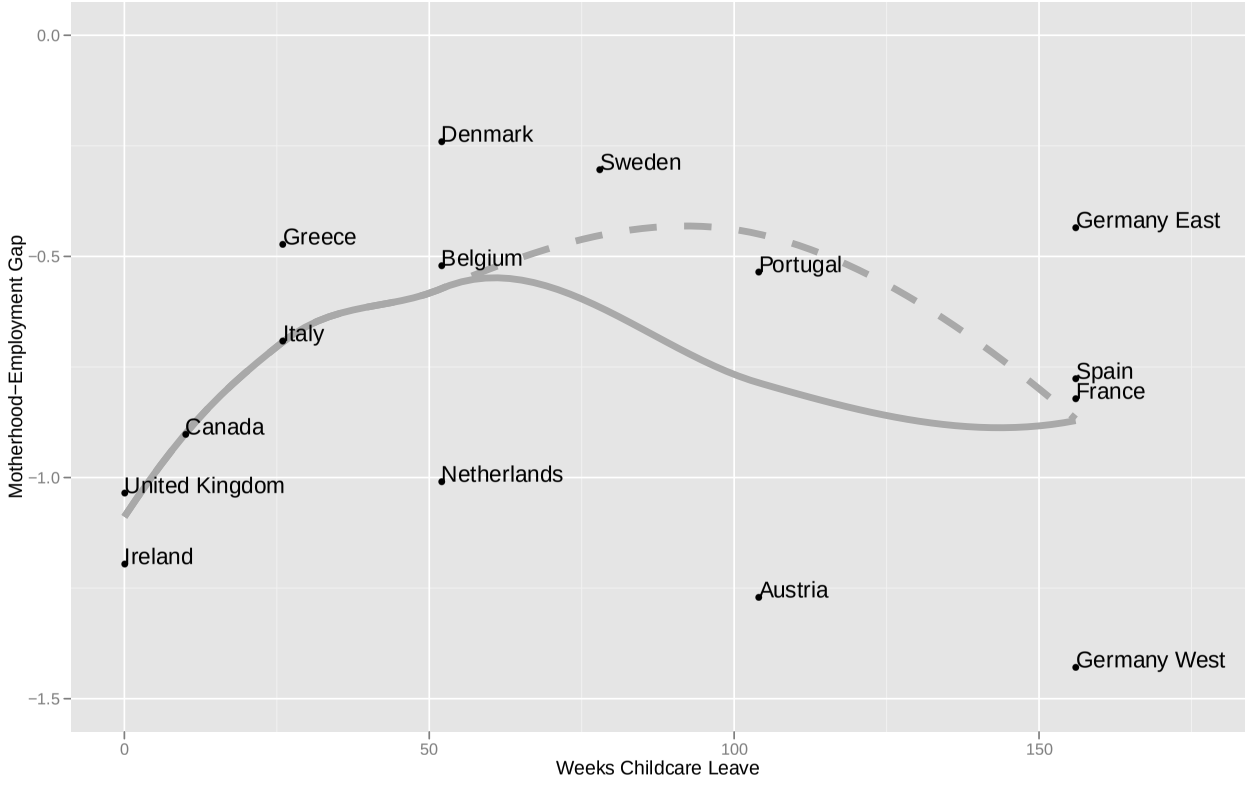


Figure A.4: Curvilinearity in the association between childcare leave and the motherhood-employment gap in 1997
Legend: Continuous line is based on all observations; Dashed line based on observations excluding Austria
Source: Authors' calculations on the cross-sectional dataset described in this chapter.

curve, and therefore may constitute influential cases. Deletion of a single country thus may result in a differently shaped curve, and thereby the observed association between the duration of leave and the size of the motherhood-employment gap. For instance, in 1997 The Netherlands and Austria show a relatively large motherhood-employment gap relative to other countries at the respective durations of childcare leave, such as Belgium and Portugal. Particularly Austria may have been an influential case, because the motherhood-employment gap was very large in this country (indicated by the strongly negative logit). The evaluate the degree to which Austria has affected the shape of the curve, the curve drawn with the dashed line represents the association between leave duration and the motherhood-employment gap after Austria was deleted from the data.

The graphical analysis shown in Figure A.4 has two disadvantages for detecting influential cases. First of all, it only provides an informal test of the influence of Austria on the (shape of the) association between duration of leave and the motherhood employment gap. Secondly, the selection of Austria was based on this country being an outlier. As influential cases are not necessarily outliers (Crawley, 2007), the procedure should be performed for all countries in the data. Hence, we next estimate a regression model to test our long-leave hypothesis, and then evaluate this model for the presence of influential data. This allows for a formal test, and is more efficient.

In Table A.2 two multilevel logistic regression models are presented based on the cross-sectional data covering 1997. In Model I, women's employment is regressed on motherhood, the duration of childcare leave and the squared duration, and the interaction between motherhood and the (squared) duration of leave. We thereby account for the first two limitations we identified in studies answering the long-leave question. In addition to the intercept, only the estimate for the effect of motherhood on employment is significant, and indicates that mothers on average are less likely to be employed than women without children. This motherhood-employment gap varies substantially between countries, as is represented by the random effect of motherhood (a variance of .325). The results thus also suggest that there is neither a

Table A.2: Multilevel Model Results Predicting Women's Employment From the Interaction Between Motherhood and the (curvi-)linear effect of Childcare Leave, 1997
Source: Comparative Motherhood-Employment Gap Trend File, and the Comparative Maternity, Parental, and Childcare Database (Gauthier & Bortnik, 2001).

	Model I		Model II	
	B	SE	B	SE
Fixed Effects				
Intercept	1.00***	0.149	0.96***	0.158
Mother	−0.65***	0.125	−0.56***	0.112
Childcare Leave	0.02	0.039	0.01	0.041
Leave Squared	0.00	0.004	0.00	0.005
Leave × Mother	0.02	0.032	0.05*	0.029
Leave^2 × Mother	0.00	0.000	−0.01*	0.003
Random Effects (SD) Country				
Intercept	0.407		0.408	
Mother	0.325		0.268	
N.countries	15		14	
N.obs	16,890		15,534	
-2×Loglikelihood	20,560		19,198	

* p<0.05, ** p<0.01, *** p<0.001 (one-tailed)

linear nor a curvilinear effect of the duration of childcare leave on the size of the motherhood-employment gap. Based on this analysis, the long-leave hypothesis should be rejected.

In the graphical analyses in Figure A.4 we observed how the shape of the association between the duration of childcare leave and the size of the motherhood-employment gap could substantially change after the deletion of a single country from the data. To formally test to what extent each of these countries have overly influenced the estimate of the hypothesized curvilinear association between the duration of childcare leave and the degree to which mothers are less likely to be employed than women without children, we have used the influence.ME software presented in this Appendix (Nieuwenhuis et al., 2012c). First,

we calculated the Cook's Distance for each country in the data, to evaluate the level of influence these countries had. Since the goal of the analyses is to test our hypothesis, we have determined for the country with the largest Cook's Distance whether the deletion of the observations in this country from the data would lead to different conclusions regarding the level of statistical significance on each of the parameters in Model I of Table A.2. This procedure showed Austria had the largest value for Cook's Distance, and we decided to regard Austria as an influential case and to delete it from the data. The updated model is presented in Model II of Table A.2.

After the deletion of Austria from the data, the model results have changed. The results presented in Model II of Table A.2 now indicate that indeed there was a curvilinear association between the duration of childcare leave and the degree to which mothers were less likely to be employed than women without children. Please note that this model is not nested in any of the previous models, as it is based on 14 (instead of 15) countries and 15,534 (instead of 16,890) individual level observations. Therefore, the variances and the -2×Loglikelihood of these models cannot be compared directly. Nevertheless, the models outcomes suggest that while short durations of childcare leave facilitate women to combine motherhood and employment, very long periods of childcare leave reduce the attachment of mothers to the labour market. Women without children are not affected by either the linear or the curvilinear effects of leave duration. Thus, these results corroborate the long-leave hypothesis.

Further deletions of countries from the data, such as for instance either (formerly) Eastern Germany or Western Germany, did not affect the presented results in such a way that the conclusion regarding the long-leave hypothesis would no longer hold. The detection and deletion of Austria as an influential case thus has improved the stability of the analysis.

A.5 DEALING WITH INFLUENTIAL DATA

Now that overly influential cases have been identified in our model, we have to decide how to deal with them. Generally, there are several strategies, including getting more data, checking data consistency, adapting model specification, deleting the influential cases from the model, and obtaining additional measurements on existing cases to account for the overly influential cases (Van der Meer et al., 2010; Harrell Jr., 2001).

Since overly influential data are a problem especially encountered in models based on a limited number of cases, a straightforward remedy would be to observe more cases in the population of interest. In our example, if we would be able to sample more schools, it may very well turn out that we observe several additional schools with a low score on the `structure` variable, so that school number 7472 is no longer influential. Secondly, there may have been measurement, coding, or transcription errors in the data, that have lead to extreme scores on one or more of the variables (i.e. it may be worthwhile, if possible, to check whether class structure and / or students' math performance in school 7472 really is that low). Thirdly, the model specification may be improved. If the data are used to estimate too complex models, or if parameterization is incorrect, influential cases are more likely to occur. Perhaps the `structure` variable should have been treated as categorical.

These are all general strategies, but cannot always be applied. Depending on the research setting, it is not always feasible to obtain more observations, to return to the raw data to check consistency, or to reduce model complexity or change parameterization.

The fourth strategy, deleting influential cases from the model, can often be applied. In general, we suggest deleting influential cases one at the time and then to re-evaluating the model. Deleting one or more influential cases from a mixed effects model is done with the function `"exclude.influence"`. The input of this function is a mixed effects model object, and it returns an updated mixed effects model from which a specified group was deleted. To illustrate, we return our

school-example, and we delete school number 7472 (which was identified as being overly influential) and its individual-level observations, using the example code below:

```
m22 <- exclude.influence(m23,
   "school.ID", "7472")

print(m22, cor=FALSE)
```

The `"exclude.influence"` function takes a mixed effects model as input, and requires the specification of the grouping factor (`school.ID`) and the group to be deleted (`7472`). It returns a re-estimated mixed effects model, that we assign to the object `m22`. The summary of that model is shown below:

```
Linear mixed model fit by REML
Formula: math ~ homework + structure
   + (1 | school.ID)
   Data: ..1
  AIC  BIC logLik deviance REMLdev
 3560 3581  -1775     3554    3550
Random effects:
 Groups    Name        Variance Std.Dev.
 school.ID (Intercept) 15.333   3.9157
 Residual              70.672   8.4067
Number of obs: 496, groups: school.ID, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept)  59.4146     5.9547   9.978
homework      2.5499     0.2796   9.121
structure    -3.8949     1.4342  -2.716
```

Two things stand out when this model summary is compared to our original analysis. First, the number of observations is lower (496 versus 519), as well as the number of groups (22 versus 23). More importantly, though, the negative effect of the `influence.MEstructure`

variable now is statistically significant, whereas it was not in the original model. So, now these model outcomes indicate that higher levels of class structure indeed are associated with lower math test scores, even when controlled for the students' homework efforts.

Further analyses should repeat the analysis for influential data, for other schools may turn out to be overly influential as well. These repetitive steps are not presented here, but as it turned out, three other schools were overly influential. However, the substantive conclusions drawn based on model `m22` did not change after their deletion.

Finally, we suggest an approach for dealing with influential data, based on Lieberman (2005). He argues that the presence of outliers may indicate that one or more important variables were omitted from the model. Adding additional variables to the model may then account for the outliers, and improve the model fit. We discussed above that an influential case is not necessarily an outlier in a regression model. Nevertheless, if additional variables in the model can account for the fact that an observation has extreme scores on one or more variables, the case may no longer be an influential one. For instance, pertaining to our long-leave example, a post-hoc interpretation of the finding that Austria was an influential case in the cross-sectional analysis reads that in Austria the period of time during which a woman is guaranteed she can return to her original employer after taking leave is shorter than the period during which income is supplemented, reducing the likelihood of women's return to employment after leave (OECD, 2011). This could be measured in future analyses and then it could be tested whether this explanation accounted for Austria being an influential case.

Thus, adding important variables to the model may solve the problem of influential data. When the observations in a regression model are, for instance, randomly sampled respondents in a large-scale survey, it often is impossible to return to these respondents for additional measurements. However, in social science applications of mixed effects

models, the higher-level groups are often readily accessible cases such as schools and countries. It may very well be possible to obtain additional measurements on these schools or countries, and use these to remedy the presence of influential data.

A.6 CONCLUSION

`influence.ME` provides tools for detecting influential data in mixed effects models. The application of these models has become common practice, but the development of diagnostic tools lag behind. The R package `influence.ME` calculates standardized measures of influential data such as DFBETAS and Cook's distance, as well as percentile change and a test for changing in statistical significance of fixed parameter estimates. The package and measures of influential data were introduced, a practical example was given, and strategies for dealing with influential data were suggested.

B

NETTING DOWN GROSS EARNINGS DATA IN THE LIS DATABASE: AN EVALUATION OF TWO PROCEDURES

ABSTRACT

LIS researchers who seek to perform country-comparative and / or trend analyses have to account for the fact that in some LIS datasets income variables were reported net of taxes and social security contributions, while in other datasets income variables were reported gross of taxes and social security contributions. In this technical paper we discuss, develop, and evaluate two 'netting down procedures' that help reduce the bias that would be introduced by directly comparing net and gross datasets. Evaluations of the performance of these netting down procedures indicate that the validity of the comparison of net and gross datasets can be greatly improved when netting down procedures are applied. In several cases, however, substantial amounts of bias remain.

B.1 INTRODUCTION

LIS researchers who seek to perform country-comparative or trend analyses have to account for the fact that in some LIS datasets earn-

1 This chapter is based on: Nieuwenhuis, R., Munzi, T, and Gornick, J. (2013a). Netting Down Gross Earnings Data in the LIS Database: An Evaluation of Two Procedures. *LIS Technical Paper Series, number 6*. The LIS Technical Papers Series are mostly written by, or commissioned by, the LIS staff. The Technical Paper version also provides the programme code.

ings variables were reported net of income taxes and social security contributions (paid by the employee), while in other datasets earnings variables were reported gross of income taxes and social security contributions. Directly comparing net and gross earnings variables while assuming that both measure the same earnings concept introduces bias to the analysis. Net earnings variables represent earnings after income taxes and contributions are subtracted, while gross earnings variables represent earnings before income taxes and contributions are subtracted. For users of LIS who seek to perform country-comparative or trend analyses, this results in the challenge that their selected earnings variable(s) refer(s) to different earnings concepts in the different datasets used, and therefore in most applications should not be compared directly. In the remainder of this technical paper, wherever the word 'taxes' is used, we refer to the combination of income taxes and social security contributions. In addition, when we use the term 'contributions', we refer to 'social security contributions' (paid by the employee); in other words, we use those terms interchangeably

LIS researchers have available (and applied) four different strategies for comparing net and gross datasets. The first is to include both types of datasets in the same (comparative) analysis, explicitly stating that the comparison might be biased. The second strategy is to limit the analysis either to only net datasets, or to only gross datasets. This results in correct analyses, but clearly limits the scope of the research. Thirdly, LIS users can present separate analyses for the net datasets, and for the gross datasets. This also results in correct analyses, and users can discuss the differences in results between the two groups of countries. The limitation of this strategy is that differences in the results between the net and gross analyses can originate both from the different earnings concepts that were used and from the analyses being based on different countries. In addition, because separate analyses were performed, no statistical tests can be performed to compare differences in outcomes between (groups of) countries. The fourth strategy for comparing net and gross datasets is to modify the gross earnings data to approximate net earnings data. This process is referred to as '*netting down*' gross data, and entails subtracting taxes from the gross

data. Research on such netting down procedures, however, is limited. Although some netting down procedures for LIS are available informally, they are undocumented and there has been no empirical evaluation of whether such netting down procedures result in measurements of earnings that are equivalent across datasets. Hence, the goals of this technical paper are, in order:

- Provide background information on the comparison of earnings in net and gross LIS datasets.
- Provide practical guidelines on using two netting down procedures.
- Quantify the degree of bias introduced by directly comparing net and gross earnings variables, answering the question "*To what extent is bias introduced by directly comparing gross to net earnings data?*"
- Evaluate the degree to which the two netting down procedures improve the comparability of net and gross earnings variables, answering the question: "*To what extent do 'netting down' procedures result in measures of net earnings that are equivalent (unbiased) across datasets?*"

B.2 COMPARING NET AND GROSS EARNINGS DATA

The Luxembourg Income Study Database (LIS) provides harmonised survey data on numerous aspects of income, taxes, social security contributions, transfers, expenditures, consumption, employment, as well as background information, covering nearly 40 countries, with the first wave dating back to around 1980. All datasets in the LIS database are harmonised to a common template, allowing for comparisons between countries and over time. This makes LIS an invaluable source of data for country-comparative and/or trend studies on various aspects of income. From the 205 LIS datasets available at the time of writing,

55 (27%) were classified as providing earnings variables net of taxes, 139 (68%) provide earnings variables gross of taxes, and 11 (5%) were classified as 'mixed'.[1] Newly added datasets are increasingly likely to provide earnings gross of taxes.

The difference between datasets with net or gross earnings variables affects country-comparative research, as the earnings variables for some countries were always classified as net (e.g. Hungary, Mexico, and Slovenia) and others always as gross (e.g. Australia, Canada, and the Netherlands). As the measurement of earnings variables over time changed from net to gross in some countries(e.g. Greece, Italy), the (lack of) comparability between net and gross earnings variables can also affect trend studies on a single country.

B.2.1 Why is the difference between net and gross datasets important?

Comparing net and gross data on income can either be a challenge in comparability, or of substantive interest. This applies to income in general, and to earnings as a specific form of income on which this technical paper is focused.

Firstly, the differences between net and gross earnings becomes a challenge in comparability, when comparisons are made across datasets of which some are gross and others are net. This is the case in country-comparative analyses and / or in trend analyses. It has been shown that country-comparative studies based on different earnings concepts across countries can be "*seriously misleading*" (p. 777), for instance when net and gross earnings variables are mixed (Atkinson & Brandolini, 2001). Also, in measuring inequality, the earnings concept used was found to affect not only the level of inequality, but also the trend in inequality (ibid.).[2] The comparison between net and gross

1 For an updated overview, see: http://www.lisdatacenter.org/our-data/lis-database/datasets-information/

2 The literature on comparability of measurements across countries and / or time is well developed (Davidov, Schmidt, & Biliet, 2011; Verhagen, 2012; Kline, 2005). However, the methodology in this literature is mostly based on a generalised-latent variable approach. This approach assumes a latent construct that is measured by multiple manifest indicators. Since earnings are not a latent construct, this approach does not apply here.

datasets can then be improved by netting down the gross earnings, so that the comparison is based on a common earnings concept: net of taxes.

Secondly, the difference between gross and net income is of substantive interest when a single dataset contains information on both gross and net income or earnings, when information on taxes is available, and when additional information on social transfers is available. This allows LIS users to, for instance, compare pre-tax-pre-transfer income to post-tax-post-transfer income, and thereby to answer a set of research questions on how taxes, but also social transfers, affect the income distribution.[3] This approach has been applied to study the effects of social-welfare policies on poverty rates in the total population of countries (see, e.g., Kenworthy, 1999). Other studies have evaluated how taxes and transfers affected poverty of specific subgroups, such as children (see, e.g., Gornick & Jäntti, 2012) and migrant households (Sainsbury & Morissens, 2012; Morissens & Sainsbury, 2005). With respect to earnings, questions on redistribution include whether the gender gap in earnings differs between net and gross earnings (cf. England, Gornick, & Shafer, 2012; Blau & Kahn, 2000), and the whether women's net or gross earnings provide more bargaining power in household economic models (cf. Becker, 1991). For such 'redistribution studies', the actual differences between net and gross income are of substantive interest, and both are compared *within* a single dataset.

The issue of comparability between net and gross datasets in LIS also applies to redistribution studies. This is clarified using a detailed example of a typical redistribution study on the comparison between *pre-tax-and-transfer* income (referred to as 'household market income') to *post-tax-and-transfer* income (referred to as 'disposable household income') (Gornick & Jäntti, 2012). However, market income was reported gross of income taxes in some LIS datasets, and net of income taxes in other LIS datasets. As for these 'redistribution' studies, parallel to the issues raised with earnings, bias is introduced when gross and net datasets are mixed and the poverty/inequality reductions associated

3 Note that with comparing pre-tax-pre-transfer income to post-tax-post-transfer income, solely using the concept of 'earnings' is not meaningful, as total income can be derived from labour (including earnings), capital and transfers.

with taxes and transfers are compared across datasets. That clearly understates poverty/inequality reduction in the net datasets, as the comparison between market income and disposable household income in these datasets only captures the effects of transfers, whereas in gross datasets this comparison would capture the combined effect of taxes and transfers. Thus, although the quantity of interest (the difference between market household income and disposable household income) is derived by comparison *within* a dataset, here again the challenge of comparisons arises when this quantity of interest is compared between net and gross datasets. Thus, here too the comparison between net and gross datasets can be improved by netting down the gross datasets, so that the comparison is based on a common market household income concept.

B.2.2 Netting Down, or Grossing up?

It should be noted that in principle there is no difference between netting down gross earnings data, or grossing up net earnings data. Both options can be attractive, depending upon the substantive questions to be answered or hypotheses to be tested with the data.

With LIS, however, grossing up is not possible as the datasets that are labelled as 'net' do not contain information on taxes. To then estimate the gross earnings would require country-specific details on the tax system, which is beyond the scope of this technical paper. Detailed simulations to this end are available for many countries through the Euromod project.[4]

The scope of this technical paper is limited to developing and evaluating netting down procedures.

4 On the internet: https://www.iser.essex.ac.uk/euromod

B.3 GUIDELINES ON NETTING DOWN PERSON-LEVEL EARNINGS

In this section, practical guidelines are provided on netting down earnings measured at the person-level. The steps to be taken in netting down person-level earnings in gross datasets are summarised in Figure B.1.

B.3.1 Netting Down Specific Income Sources

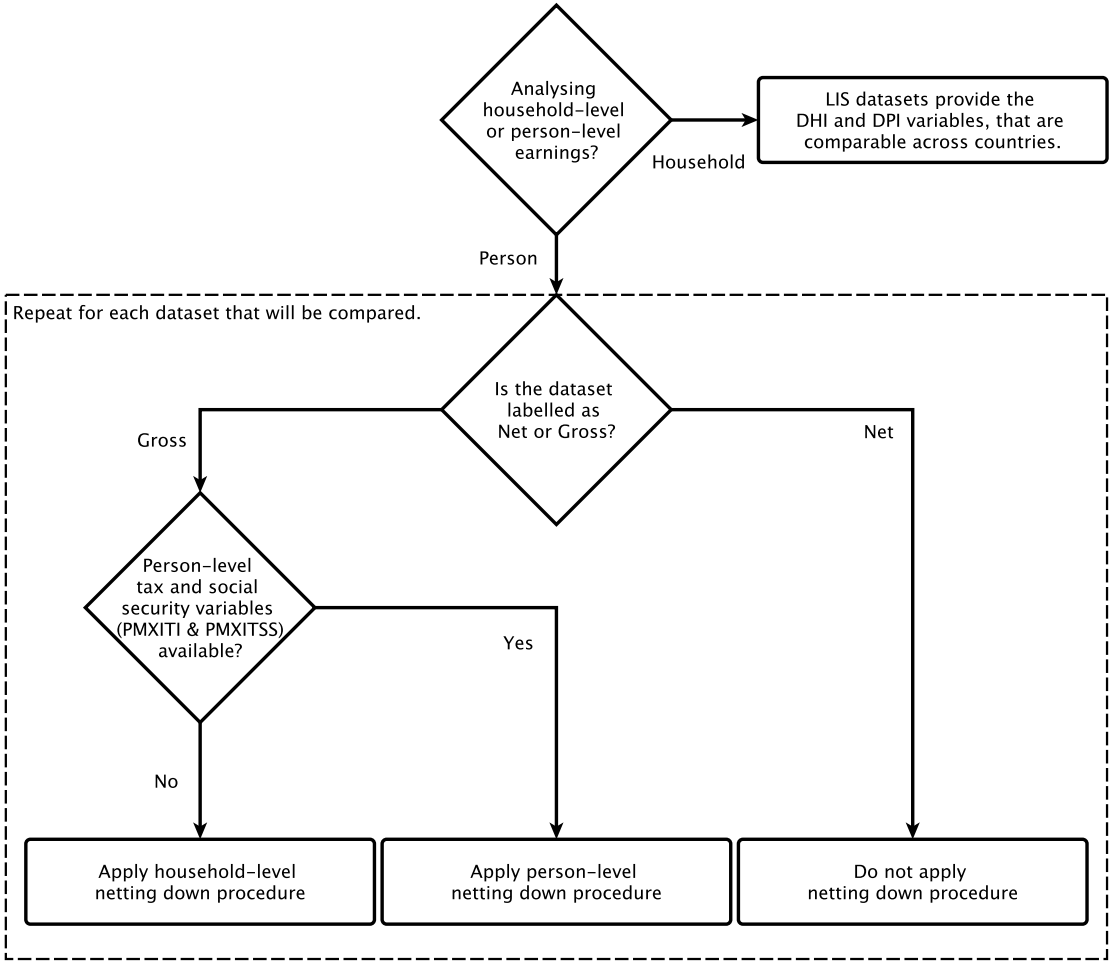
Total income is typically recognised as being obtained from three different sources: labour, capital, and transfers. A researcher may be interested in comparing the total income derived from these sources combined, or be interested in the income derived from one of these sources, such as labour. These three main categories can be divided further, for instance by differentiating between earnings derived from dependent employment or from self-employment (both part of labour).

Netting down the income from either labour, capital, or transfers is challenging both conceptually and practically. The conceptual problem with net income from separate sources lies in the fact that earners typically pay their taxes based on all (taxable) income. So, if a household or person has income from multiple sources the amount of income tax paid is based on the amount of total income, rather than the amount of earnings from each of these sources separately. Nevertheless, studying income from a single source, such as the earnings obtained from employment, may be interesting for many researchers.

The practical problem with net income from separate sources is that, as a result of the above, only information on total taxes is available. To calculate the net income from a separate source, the assumption must be made that income from each source was taxed at the same rate.

In the empirical part of this technical paper we will evaluate the netting down of just those *earnings* that were obtained from dependent employment (not including self-employment).

Figure B.1: Schematic Representation of Selecting the Correct Netting Down Procedure for Person-Level Gross Earnings in LIS



B.3.2 Netting Down Person Level Earnings

All LIS datasets provide disposable household income (LIS variable name: DHI) and monetary disposable household income (LIS variable name: DPI). These household-level variables are comparable across countries.

However, it may also be desirable to calculate net earnings at the person level, rather than at the household level. To do so, person-level taxes (thus both income taxes and social security contributions) are subtracted from person-level gross earnings. Doing this is challenging in countries with joint taxation of members of the same household. So, if no person-level tax variables are available, netting down person-level earnings requires the assumption that the taxes paid at the household level were paid by each household member proportionally to the share of the total household income received by that member.

The decision on whether to compare household-level or person-level earnings depends, of course, on the substantive research interest. Netting down procedures can be developed for both person-level and household-level income concepts. As LIS provides comparable data on several specific income concepts on the household-level, but not on the person-level, the focus of this technical paper is on developing and evaluating netting down procedures for person-level earnings. This is represented in the top row of Figure B.1.

B.3.3 Programmes for Netting Down

We developed two programs that perform netting down procedures, available for STATA, SPSS, R, and SAS. These procedures either use information on taxes on the person-level or, if these are not available, or household level tax information. The LIS website has a table providing information on whether datasets are gross or net (URL was given above). Datasets classified as mixed should be treated with more caution, as the earnings reported in these datasets can be gross of income taxes but net of contributions, or vice versa. This is reported in detail in the LIS data documentation per country. All LIS datasets also contain a

variable named 'GROSSNET', providing information on how earnings (and other income variables) were reported.

Users should carefully apply the correct netting down programme for each dataset in their analysis. The second row from the top in Figure B.1 indicates that netting down procedures should not be applied to datasets that were classified as net.

If person-level variables on taxes (LIS variable PMXITI) and (self-paid) social security contributions (LIS variable PMXITSS) are available, the person-level netting down procedure should be applied. Otherwise, the household-level netting down procedure should be applied. This is represented by the third row from the top in Figure B.1. The person-level and household-level netting down procedures are described below. These descriptions also state the assumptions that are required for the procedure to result in information on person-level net earnings.

PERSON–LEVEL NETTING DOWN PROCEDURE: When person level tax information is available, the netting down procedure first calculates the total taxable income (earnings from dependent employment, self-employment, unemployment compensation benefits, short-term sickness and work injury benefits, family leave benefits, and pensions). Next, it calculates the proportion of that total income that was obtained by dependent employment. Next, it is assumed that the total amount of taxes was distributed proportionally over all sources of income. As taxable income is made up of different components across countries, this procedure is based merely on an approximation of taxable income. The person level net earnings are then calculated by subtracting the paid income taxes from the gross earnings, proportional to the amount of total income obtained from earnings.

HOUSEHOLD–LEVEL NETTING DOWN PROCEDURE: When tax information is only available at the household level, the netting down procedure first calculates the percentage of the total household income that was paid as taxes. Next, it is assumed that this percentage is equal for all members of the household, and applies equally to all sources of income. The person level net earnings

are then calculated by reducing the person level gross earnings by the percentage taxes.

B.4 METHOD AND DATA

B.4.1 Method

In a select number of LIS datasets both gross and net earnings variables were reported at the person level, along with information on taxes and social security contributions on both the person-level and the household-level. These datasets therefore provide a unique opportunity for quantifying the amount of bias that arises from directly comparing net and gross datasets, as well as for evaluating netting down procedures.

To quantify the degree of bias that is introduced by directly comparing datasets with either net or gross earnings, we calculated different measures of the earnings distribution, such as the average earnings. We did so twice for each dataset: once using the gross earnings variable, and once using the net earnings variable. The two resulting averages thus refer to the same country and year, and they were calculated based on exactly the same respondents. The difference between gross and net does not always indicate bias, as they represent different earnings concepts. However, this difference does represent the amount of bias that would have been introduced to analyses directly comparing average earnings from gross and net datasets while assuming both are indeed directly comparable. This difference is expressed as a percentage of the value of the measure based on net earnings:

$$\mathrm{Difference}(\%) = \frac{\overline{X}_g - \overline{X}_n}{\overline{X}_g} \times 100\% \tag{B.1}$$

in which $\overline{X}_g$ represents the average earnings calculated on the gross data as reported by respondents, and $\overline{X}_n$ represents the average earnings calculated based on the net data as reported by respondents.

We have calculated not only the average earnings, but have also four commonly used measures of inequality. In total, we evaluate bias using five measures at the person-level:

- Average earnings;
- Ratio between 25th and 75 percentile of earnings;
- Gini of earnings;
- Low earnings rate (Defined as percentage population with earnings below 2/3 of median earnings);
- Gender gap in earnings (Defined as: (male earnings - female earnings) / male earnings).

In addition to calculating the difference between gross and net gross measures of the earnings distribution, we evaluated to what extent the netting down procedures described result in an unbiased approximation of net earnings. To evaluate a netting down procedure, we applied this procedure to a gross earnings variable, calculated the average earnings (or one of the other measures of the earnings distribution) based on the netted down earnings variable, and compared the results to those based on the reported net earnings. The reported net earnings thus serve as a benchmark against which the netted down net earnings are evaluated. We calculated the degree to which the netted down results are biased (compared to the reported net results) using the following equation:

$$Bias(\%) = \frac{\overline{X}_{nd} - \overline{X}_{n}}{\overline{X}_{n}} \times 100\% \tag{B.2}$$

in which $\overline{X}_{nd}$ represents the (for example) average earnings calculated on the netted down net data, and $\overline{X}_{n}$ represents the average earnings calculated based on the net data as reported by respondents. The resulting bias is expressed as a percentage of the reported net earnings. So, a bias of 0% means that the results based on the approximated net earnings ($\overline{X}_{nd}$, obtained using the netting down procedure) are identical to those based on the net earnings as reported by respondents ($\overline{X}_{n}$).

In that case the netting down procedure results in an unbiased measure of net earnings. If the bias % is larger than 0, this means that the results based on the approximated net earnings are higher than those based on the reported net earnings, a percentage below 0 indicates that the approximated results are lower.

To evaluate whether the netting down procedure improves the quality of a comparison of earnings across net and gross datasets, the bias of the netting down procedure (Bias %, defined in B.2) is to be compared to the difference between reported net and gross earnings (Difference %, defined in B.1). If the (absolute) percentage of bias of the netting down procedure is smaller than the difference between the reported net and gross earnings, this indicates that the netted down earnings are closer to the benchmark of the reported net earnings than the reported gross earnings are.

B.4.2 Data

The netting down procedures described in this technical paper can be applied to LIS datasets of the 'new' (post-2011) template that are classified as gross. The evaluation of these netting down procedures, however, required the availability of both gross and net earnings as reported by the respondent. This could only be the case in the 'old' template (pre-2011; these are still available to users).[5] The required earnings variables, as well as person- and household level variables on taxes and social contribution were available in 7 datasets: Austria 2004, Belgium 1992, Belgium 1997, Estonia 2004, Ireland 2004, United Kingdom 1999, and United Kingdom 2004.

We have restricted our analyses to those observations with valid information on both the gross and net earnings variables. Although this introduced some missing observations (either by persons not having reported their gross earnings, their net earnings, or both), the goal of this technical paper is not to obtain valid inferences regarding the measures of earnings, but to understand the bias that was introduced by di-

5 *Luxembourg Income Study Database (LIS)*, www.lisdatacenter.org (multiple countries; november 2012-june 2013). Luxembourg: LIS

rectly comparing net and gross earnings data. By deleting observations in which any or both of the earnings measures were missing, we ensured that our measurement of bias was not affected by gross and net earnings variables which were based on different observations. Person-level sampling weights were applied. The total number of person-level observations is presented in Table B.1.

Table B.1: Number of Observations and Countries
Source: Data from the Luxembourg Income Study (LIS, 2013)

Country	Year	Number Observations
Austria	2004	5,563
Belgium	1992	4,138
Belgium	1997	4,001
Estonia	2004	5,155
Ireland	2004	3,297
United Kingdom	1999	21,791
United Kingdom	2004	24,161

B.5 RESULTS

B.5.1 Results on Average Earnings

Firstly, we present in detail our results regarding netting down average earnings. In the next sub-section, we present the outcomes of netting down the measures of inequality. In Figure B.2 the results of our analyses are presented graphically. In panel A, the average earnings are shown, standardised by setting the average gross earnings to 100 to account for widely different averages between countries. The average gross earnings were directly observed in the data and are represented by the white bars. The observed net earnings are shown as black bars, and unsurprisingly the average net earnings are much lower than the average gross earnings. This is, of course, the result of paying taxes, but

if directly compared between datasets this would result in the amount of bias as calculated.

Next, the dark-grey bars represent an approximation of the average net earnings, which was based on the netting down procedure using person-level taxes. This average of the approximated net earnings is very close to the average of the net earnings that were reported by respondents, suggesting that the netting down procedure performed well. The same holds for the netting down procedure based on the household-level tax information household, represented by the light-grey bars.

Next, we quantify the difference between the measures of net and gross earnings, as an indicator of the amount of bias that would be introduced by directly comparing measures of the net and gross earnings distribution. In addition, we calculate the bias associated with the two netting down procedures described above. In Table B.2, the results are presented for the average person-level earnings. The first two columns present results that were directly observed from the data: the average gross earnings and the average net earnings. The third column presents the difference between the first two columns, expressed as a percentage of the net average earnings.

For instance, in Austria 2004 the gross earnings averaged at 24,555.88 and the average net earnings 17,268.49. The absolute difference equals to 7287.39, which is reported in the third column as 42.20 percent of the net average earnings (following equation B.1). Based on the third column in Table B.2 we conclude that the amount of bias introduced by directly comparing net and gross average earnings ranges from about 27 percent in Estonia 2004 to 85 percent in Belgium 1997.[6] These percentages are based on the net earnings, suggesting that in Belgium 1997 the gross earnings were close to twice as high as the net earnings.

Next, under the header 'Netted Down: Person', an approximation is presented of the average net earnings based on the netting down procedure using the person-level taxes. We observe that using the per-

6 It should be noted that the average earnings (both net and gross) in Belgium are much higher in 1997 than in 1992. This, however, is simply due to the fact that in 1997 the currency was expressed in Belgian Francs and in 1992 the currency was expressed in 1000s of Belgian Francs.

Figure B.2: Bias Comparing Net and Gross Earnings Data and the Performance of Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

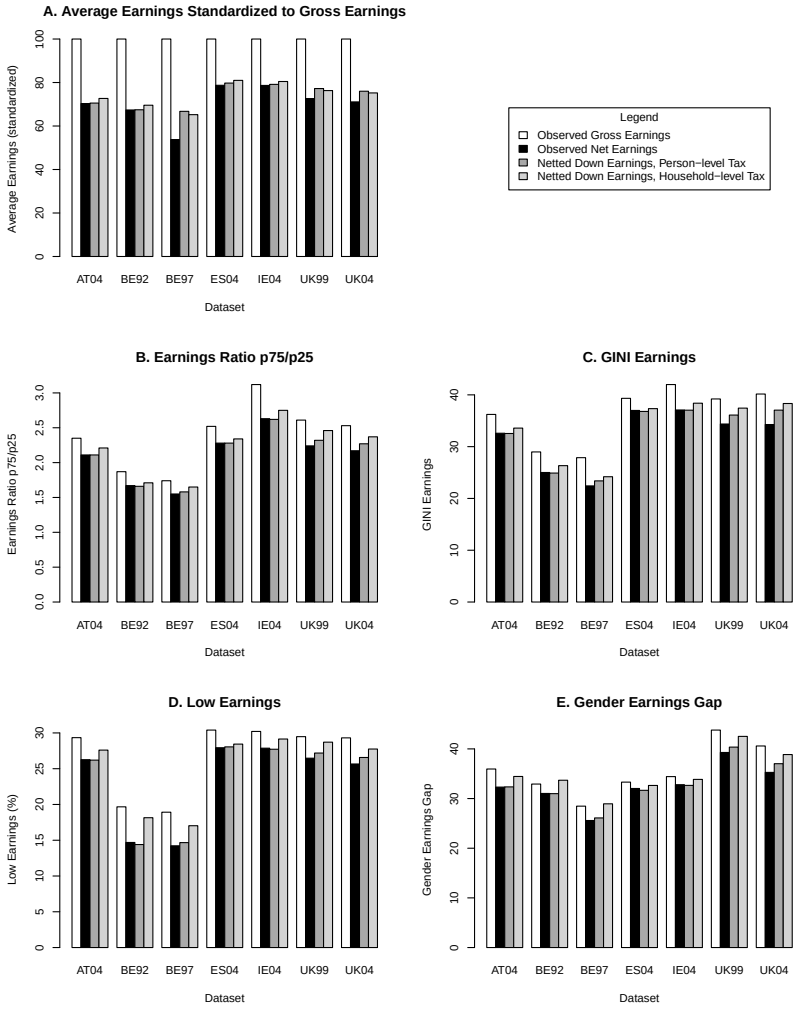


Table B.2: Average Earnings: Quantifying Bias and Evaluating Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

		Observed from Data			Person		Household	
Country	Year	Gross	Net	Difference (%)	abs.	Bias (%)	abs.	Bias (%)
Austria	2004	24555.88	17268.49	42.20	17323.61	0.32	17846.89	3.35
Belgium	1992	7863.32	5297.93	48.42	5305.25	0.14	5469.41	3.24
Belgium	1997	1105850.00	594700.10	85.95	738104.20	24.11	721286.80	21.29
Estonia	2004	76243.25	60022.50	27.02	60796.69	1.29	61745.79	2.87
Ireland	2004	26523.79	20864.45	27.12	20992.20	0.61	21343.59	2.30
UK	1999	16893.31	12266.87	37.71	13047.11	6.36	12887.84	5.06
UK	2004	21335.35	15164.91	40.69	16212.84	6.91	16047.22	5.82

son level tax data, much of the difference between average gross and net earnings is accounted for by the netting down procedure. In other words, the netting down procedure results in approximations of net earnings that have very little bias compared to the benchmark of reported net earnings. For instance, applying the netting down procedures on the Austrian gross earnings data, results in an average netted down wage of 17,323.61 Euros. The reported net earnings averaged at 17,268.49. This means that the netting down procedure performed very well using the person level tax information in the data, and resulted in an approximation of average net earnings that was very close to the reported net earnings. Following equation B.2, the bias in this case was only 0.32%. This is, of course, much lower than the 42.20% bias that would have been introduced by directly comparing average earnings based on gross and net earnings data. When only household level tax data was available (presented in the final two columns), the netting down procedure performed slightly less well with a calculated bias of 3.35%. Based on the performance of the two netting down procedures in other countries, it is clear that the netting down procedures substantially reduce the bias for calculating average earnings, but no conclusions can be drawn on whether the netting down procedure based on person level taxes or the netting down procedure based on household level taxes performed better.

B.5.2 Results on Measures of Inequality

Looking back at Figure B.2, panels B through E represent the results of the ratio of the earnings of the 75th and 25th percentile of the earnings distribution, the GINI, the percentage of individuals with low earnings, and the gender gap in earnings. The overall patterns in each of these four measures of inequality are similar. The results based on the reported gross earnings are different from those based on the reported net earnings, with inequalities being smaller in the net data. In all cases, this suggests that bias would be introduced by directly comparing net and gross earnings. In most cases, the results for the netted down earnings are closer to the reported net than the reported gross.

This holds both for the person-level and household-level netting down procedures, suggesting that both procedures perform well in reducing bias.

These findings are generally supported by the quantified results presented in Tables B.3 (earnings ratio of the 75th and 25th percentile), B.4 (GINI), B.5 (low earnings rate), and B.6 (gender gap in earnings). Closer examination of these quantified results reveals several further findings, including some exceptions to the general pattern. We discuss five.

Firstly, the bias that is introduced when directly comparing gross and net earnings (indicated in the columns labelled *Difference (%)*) is bigger for average earnings than it is for the four measures of inequality.

Secondly, the performance of the netting down procedures is typically poorer for measures of inequality than it is for average earnings. This means that the remaining bias is smallest where the differences between net and gross earnings were largest to begin with.

Thirdly, for all measures of inequality and in all countries that were evaluated, the person-level netting down procedure outperformed the household-level netting-down procedures. The bias associated with the person-level netting down procedure typically reduced the bias by 0-8%. When household-level tax variables were used, however, performance degraded: the bias associated with this netting down procedure typically ranged from 1% to 24%. So, for comparing measures of inequality, using the person-level netting down procedure is preferable when person-level tax information is available.

Fourthly, in some cases the netting down procedures corrected more than 100% of the existing difference between net and gross earnings. This is indicated by a negative estimate of bias. An example is the 75p/25p earnings ratio in Belgium 1992: the person-level netting down procedure is associated with a bias of -1.02%.

Finally, and finally, in the specific case of the gender gap in earnings in Belgium (1992 and 1997), applying the household-level netting down procedure actually introduced bias. This is indicated by the fact that the bias associated with this netting down procedure was larger

than the actual differences between the measurements of the gender gap in earnings based on the (reported) net and gross earnings.

B.5.3 Two Notes on the Benchmark

As indicated, the benchmark we used to evaluate the netting down procedures was the difference between the reported gross and reported net earnings from dependent employment. Closer examination of the original micro-level surveys that were used to create the LIS data, however, showed that this benchmark was not correct for all countries. The reason for this is that the reported gross and net earnings were calculated in different ways (other than the exclusion of income taxes and social security contributions in the net earnings). Two such exceptions are described in this section.

Firstly, the LIS dataset on Belgium in 1997 was based on the Socio-Economic Panel. In this original dataset, the information on holiday- and end-of-year bonuses was available net of taxes, but unavailable gross of taxes. Hence, in calculating the yearly gross earnings in the preparation of the LIS dataset, the monthly earnings were multiplied by 13,85 (approximating the average bonuses). For the net yearly earnings the information on bonuses was available in the original data. Hence, whereas in the LIS dataset the net yearly earnings account for person-level variation in bonuses independent of other earnings, in the gross yearly earnings such person-level variation was not accounted for.

Secondly, the LIS datasets on the United Kingdom (both in 1999 and 2004) were based on the Family Resources Survey. During the recoding of these datasets to the LIS templates, the gross earnings were specified to include income from odd jobs, while net earnings could not be specified to include this source of income. Hence, the difference between gross and net yearly earnings is an overestimation of the 'real' difference. Therefore, the netted down results presented in the Tables of this technical paper may actually be a better representation of true net earnings than the net earnings reported in the data. It should be

Table B.3: 75p/25p Earnings Ratios: Quantifying Bias and Evaluating Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

		Observed from Data			Person		Household	
Country	Year	Gross	Net	Difference (%)	abs.	Bias (%)	abs.	Bias (%)
Austria	2004	2.35	2.11	11.22	2.11	0.05	2.21	4.35
Belgium	1992	1.87	1.67	12.08	1.66	-1.02	1.71	2.27
Belgium	1997	1.74	1.55	12.48	1.58	2.39	1.65	6.60
Estonia	2004	2.52	2.28	10.52	2.28	0.00	2.34	2.72
Ireland	2004	3.12	2.63	18.62	2.62	-0.61	2.75	4.71
United Kingdom	1999	2.61	2.24	16.39	2.32	3.71	2.46	10.09
United Kingdom	2004	2.53	2.17	16.79	2.27	4.66	2.37	9.50

Table B.4: GINI of Earnings: Quantifying Bias and Evaluating Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

		Observed from Data			Person		Household	
Country	Year	Gross	Net	Difference (%)	abs.	Bias (%)	abs.	Bias (%)
Austria	2004	36.22	32.59	11.13	32.54	-0.16	33.58	3.05
Belgium	1992	28.98	25.01	15.88	24.90	-0.44	26.33	5.28
Belgium	1997	27.87	22.42	24.32	23.39	4.34	24.19	7.92
Estonia	2004	39.33	36.99	6.35	36.80	-0.49	37.34	0.95
Ireland	2004	41.99	37.08	13.26	37.04	-0.09	38.39	3.54
United Kingdom	1999	39.21	34.37	14.09	36.10	5.05	37.44	8.93
United Kingdom	2004	40.16	34.27	17.19	37.05	8.14	38.33	11.85

Table B.5: Low Earnings Rate: Quantifying Bias and Evaluating Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

		Observed from Data			Person		Household	
Country	Year	Gross	Net	Difference (%)	abs.	Bias (%)	abs.	Bias (%)
Austria	2004	29.34	26.26	11.73	26.20	-0.23	27.60	5.10
Belgium	1992	19.67	14.69	33.90	14.40	-1.97	18.15	23.55
Belgium	1997	18.92	14.22	33.05	14.67	3.16	17.03	19.76
Estonia	2004	30.40	27.93	8.84	28.05	0.43	28.44	1.83
Ireland	2004	30.21	27.87	8.40	27.72	-0.54	29.15	4.59
United Kingdom	1999	29.48	26.47	11.37	27.19	2.72	28.71	8.46
United Kingdom	2004	29.31	25.66	14.22	26.57	3.55	27.75	8.14

Table B.6: Gender Gap in Earnings: Quantifying Bias and Evaluating Two Netting Down Procedures
Source: Data from the Luxembourg Income Study (LIS, 2013)

		Observed from Data			Person		Household	
Country	Year	Gross	Net	Difference (%)	abs.	Bias (%)	abs.	Bias (%)
Austria	2004	34.95	32.31	8.17	32.35	0.12	34.45	6.62
Belgium	1992	32.92	31.03	6.09	30.99	-0.13	33.69	8.57
Belgium	1997	28.47	25.58	11.30	26.10	2.03	28.94	13.14
Estonia	2004	33.31	32.03	4.00	31.66	-1.16	32.64	1.90
Ireland	2004	34.42	32.79	4.97	32.65	-0.43	33.86	3.26
United Kingdom	1999	43.79	39.28	11.48	40.36	2.75	42.52	8.25
United Kingdom	2004	40.59	35.27	15.08	37.02	4.96	38.86	10.18

noted, that within the scope of this paper it was not possible to empirically test this statement.

B.5.4 Alternative Netting Down Procedures

The two netting down procedures evaluated in this technical paper were designed to use as few variables as possible, thereby allowing them to be used on as large a number of LIS datasets as possible. Despite their simplicity, the netting down procedures presented outperformed several alternatives that were based on more variables, but also required more assumptions. Several of these alternative netting down procedures were evaluated, but not presented. These procedures performed less well and were associated with more bias. These procedures are described below, so that users do not need to evaluate them themselves.

The following alternatives were evaluated to perform less well than those presented, or not to apply to the new LIS template:

PERSON-LEVEL GROSS/NET WAGES In some LIS datasets, the person level files provide both gross and net hourly wages (earned in a set period of time). Although this period of time (hourly, weekly, monthly, etc.) varies between datasets, it is always identical in the gross and net variables of a single dataset. Based on these variables, it was expected that a good approximation of person level taxes for earnings from paid employment could be established. This procedure could only be evaluated in Austria and Ireland, and generally performed worse compared to the netting down procedure based on both the household- and person level taxes.

HOUSEHOLD-LEVEL GROSS/NET WAGES In the pre-revised datasets of LIS, household level variables representing both gross and net 'wages and salaries' (V1 / V1NET) were available. These were used to calculate the percentage of gross wages that were paid as taxes by the household, and this percentage was used to net down the person-level gross earnings. This netting down procedure actually performed better than the one using household

level tax variables presented above (but worse than the one based on person level tax data). These variables (or equivalents) are not available in the revised LIS-template. Therefore, this procedure cannot be used with the most recent LIS datasets, and therefore these results are not presented here.

B.6 CONCLUSION

In this technical paper guidelines were presented for comparing earnings using both net and gross LIS datasets. Two netting down procedures were developed that approximate net earnings based on information regarding gross earnings, taxes and social security contributions. One netting down procedure uses tax variables measured on the person-level, and one netting down procedure uses household-level tax variables.

Descriptive analyses quantified the difference between measures of gross and net earnings, as an indication of the bias that would be introduced if net and gross datasets are directly compared. The difference between net and gross earnings was (unsurprisingly) substantial, but varied with the measure of the earnings distribution used. The difference was smallest, about 5-15%, in the gender gap in earnings, and largest, about 30-85%, in the average earnings. This suggests that, depending on the measure of earnings distribution that is used, country-comparative and / or trend analyses should treat comparisons between net and gross LIS datasets with caution.

The netting down procedures that were developed in this technical paper were typically associated with lower amounts of bias than the original difference between net and gross earnings. Generally, this suggests that applying a netting down procedure is preferable over not netting down. The exception to this, we found in the Belgium 1992 and 1997 datasets, was with the gender earnings ratio when netted down using only household-level tax variables. In these specific cases the netting down procedure actually resulted in more bias than a direct comparison between net and gross earnings.

The results of our analyses also strongly suggest that when seeking to net down earnings to estimate measures of inequality, using person level taxes is desirable over using the netting down procedure based on only household-level tax variables. However, person-level tax variables are not available in all LIS datasets.

In many cases, data availability will dictate which of the two netting down procedures users can apply. It should be noted, however, that the household-level netting down procedure is expected to perform better in a country with joint-taxation than in countries in which spouses pay taxes separately. Furthermore, in both netting down procedures it is assumed that all sources of income are taxed at the same rate. From this, the expectation follows that the procedures will perform better in countries with a single, rather than a dual tax system in which separate tax rates exist for capital income and other income. Finally, we expect the procedures to perform better in countries with a tax system that is close to proportional (=flat rate).

A user seeking to compare a large number of both net and gross datasets, may want to statistically control for the different netting down procedures used. In regression-based analyses, for instance, this could be done by adding dummy-variables indicating the observations derived from datasets netted down with the person-level procedure, and another dummy for the observations from datasets that were netted down using the household-level procedure (the observations from those datasets that were reported as net then function as the reference category). This would capture the average bias associated with the different netting down procedures, thereby further improving the analyses.

To conclude, country-comparative and trend analyses of earnings based on both net and gross LIS datasets should be executed with caution. The netting down procedures presented here typically improve comparability. However, depending on the outcome measure of interest, and especially when no person-level tax variables were available, netting down procedures result in approximations of net earnings that are substantially biased.

C

THE COMPARATIVE MOTHERHOOD-EMPLOYMENT GAP TREND FILE

C.1 INTRODUCTION

In this Data Appendix we provide information on the Comparative Motherhood-Employment Gap Trend File, used in Chapters 2, 3 and 4 and in Appendix A of this dissertation.

The vast majority of our micro-level observations were obtained from the Mannheim Eurobarometer Trend File, which provides pooled data from Eurobarometer surveys on selected trends in European countries. The Mannheim Eurobarometer Trend File provides data and variables harmonised to ensure comparability, and extensive checks have been performed regarding both the coding of the variables and plausibility of trends over time. To cover non-European countries, we added data on the United States and Canada. These data were obtained from the General Social Survey and the Canadian Election Survey. All surveys provide samples that, by the use of sampling weights, are representative for the country's population. The number of valid observations that were derived from each of these surveys is shown in Table C.1.

C.2 QUESTION WORDING

The wording of the questions in the three different surveys allowed us to have comparable measures for several demographic background characteristics of individual women. It also enabled us to provide anal-

Table C.1: Total Number of Valid Observations

Survey	Number of Observations	%
Eurobarometer	177,827	92.4
General Social Survey	10,958	5.7
Canadian Election Survey	3,699	1.9
Total	3,699	100

Table C.2: Question Wording: Employment

Survey	Question
Eurobarometer	What is your occupation?
General Social Survey	Last week were you working full time, part time, going to school, keeping house, or what?
Canadian Election Survey	(Please indicate your) present employment status: Working now / Laid off / Unemployed / Retired / Disabled / Student / Homemaker / Self-employed

yses of women's employment in a large number of countries and years, while differentiating between mothers and women without children. Nevertheless, the number of comparable and relevant measures was limited to a woman's year of birth, employment status, motherhood status, education, and marital status. The question wordings of the original surveys are presented in Tables C.2 through C.5.

Questions about the employment variables were asked slightly differently in the three surveys used, but the question wording was such that the responses could easily be recoded to whether a woman was employed at the time of the interview. Nevertheless, we had information only on whether a woman was employed, with no information on the number of hours of work per week.

The original questions on motherhood asked the actual number of children women had. Note that for almost all of our observations we actually knew whether a child was living at home, the exception being the General Social Survey. By restricting our sample to women ages 20 to 50, we limited the number of women in our sample whose children

Table C.3: Question Wording: Motherhood

Survey	Question
Eurobarometer	How many children under 15 are there living at home?
General Social Survey	How many children have you ever had? Please count all that were born alive at any time (including any you had from a previous marriage).
Canadian Election Survey	(Please indicate the) number of children under 18 living in home.

Table C.4: Question Wording: Education

Survey	Question
Eurobarometer	How old were you when you finished your full-time education?
General Social Survey	No question wording was available in the codebook. Educational level was measured in grades, which we recalculated to the (approximate) age at which education was finished.
Canadian Election Survey	(What is the) Highest level of education (you have) completed? Educational level was measured in grades, which we recalculated to the (approximate) age at which education was finished.

had already left home. Our sample does not contain data on the age of children. We have made these data comparable, and the resulting analyses more concise, by recoding these measurements to a straightforward binary variable indicating motherhood.

Questions about marital status were asked in all three surveys in a similar way, and all answer categories could be recoded to a categorical variable indicating whether a woman was living in a single household (singe, divorced, widowed; coded 0) or in a partnered household (married, living as married) at the time of interview (coded 1).

Table C.5: Question Wording: Marital Status

Survey	Question
Eurobarometer	Are you: Single / Married / Living as Married / Divorced / Separated?
General Social Survey	Are you currently: Married, Widowed, Divorced, Separated, or have you Never been married?
Canadian Election Survey	Respondent's marital status: Married / Partner / Divorced / Separated / Widowed / Never Married

C.3 NUMBER OF OBSERVATIONS PER SURVEY

In this appendix, Table C.6, contains the total number of valid observations for each country by year. In Table C.7, we present the proportion of data that were lost because of missing values on one or more of the micro-level observations.

Table C.6: Total Number of Valid Observations, by Country and Year (N = 192,484 Individuals From 305 Country-Years From 18 Countries).
Note: AT = Austria; BE =Belgium; CA = Canada; DK = Denmark; FR = France; GR = Greece; IE = Ireland; IT = Italy; LU = Luxembourg; NL = The Netherlands; NO = Norway; PT = Portugal; ES = Spain; SE = Sweden; UK = United Kingdom; US = United States; WG = West Germany; EG = East Germany.)

Year	AT	BE	CA	DK	ES	FR	WG	EG	GR	IE	IT	LU	NL	NO	PO	SE	UK	US
1975		643		608		781	528			545	572	169	662				744	473
1976		539		509		776	518			518	508	143	596				726	456
1977		527		515		630	527			507	613	173	606				732	472
1978		790		911		1075	768			717	793	251	1025				1120	505
1979		500		636		614	526			500	611	134	644				753	
1980		511		518		642	517		275	540	642	134	714				824	460
1981		521		587		648	512		526	529	582	214	635				758	
1982		570		584		600	580		535	505	554	158	709				758	594
1983		543		537		624	557		516	538	566	140	607				737	535
1984		552	1134	549		621	573		520	539	561	149	626				757	523
1985		597		573	257	616	502		530	500	559	142	640		262		775	478
1986		596		557	520	594	491		497	516	572	161	646		528		748	484
1987		608		552	497	681	477		502	526	557	143	651		549		773	593
1988		568	1053	507	536	635	540		510	529	520	166	677		526		760	465
1989		979		1017	951	1161	1193		1028	1043	1047	314	1336		991		1417	490
1990		837		728	748	923	757	621	794	851	850	154	922	258	724		1012	441
1991		869		760	751	876	783	975	804	880	805	385	894	240	704		1083	512
1992		1088		1003	1036	1111	1099	1155	1038	1136	973	496	1214	481	989		1446	
1993		256	285	274	246	255	276	297	258	282	247	128	322		230		360	532
1994		535		528	494	541	591	552	503	582	482	325	646	446	533		719	1016
1995	311	253		242	246	286	288	269	270	290	233	231	327	184	245	242	356	
1996	984	786		738	740	888	836	775	814	889	717	533	963	192	760	756	1141	1002
1997	1356	1061	1227	989	963	1207	1078	1070	1046	1158	889		1338		984	914	1610	
1998	658			450	486		543	495		563			658		489	439	822	927
1999					490												802	

Table C.7: Proportion of Missing Values, by Country and Year (N = 192,484 Individuals From 305 Country-Years From 18 Countries) Note: AT = Austria; BE =Belgium; CA = Canada; DK = Denmark; FR = France; GR = Greece; IE = Ireland; IT = Italy; LU = Luxembourg; NL = The Netherlands; NO = Norway; PT = Portugal; ES = Spain; SE = Sweden; UK = United Kingdom; US = United States; WG = West Germany; EG = East Germany.)

Year	AT	BE	CA	DK	ES	FR	WG	EG	GR	IE	IT	LU	NL	NO	PO	SE	UK	US
1975		0.04		0.03		0.05	0.04			0.01	0.07	0.05	0.04				0.02	0.04
1976		0.05		0.07		0.07	0.05			0.02	0.10	0.01	0.04				0.03	0.00
1977		0.03		0.09		0.14	0.05			0.04	0.05	0.02	0.13				0.04	0.05
1978		0.02		0.04		0.04	0.06			0.06	0.07	0.02	0.08				0.03	0.00
1979		0.08		0.05		0.09	0.07			0.04	0.05	0.02	0.12				0.03	
1980		0.03		0.06		0.04	0.08		0.04	0.03	0.05	0.02	0.08				0.04	0.00
1981		0.03		0.06		0.03	0.04		0.07	0.03	0.05	0.01	0.22				0.04	
1982		0.03		0.04		0.01	0.04		0.02	0.02	0.03	0.01	0.04				0.01	0.00
1983		0.03		0.07		0.04	0.15		0.07	0.03	0.05	0.06	0.13				0.06	0.00
1984		0.06	0.04	0.09		0.04	0.03		0.07	0.05	0.07	0.04	0.09				0.05	0.00
1985		0.03		0.05	0.11	0.05	0.09		0.08	0.05	0.07	0.06	0.12		0.06		0.04	0.00
1986		0.04		0.08	0.11	0.04	0.08		0.07	0.04	0.07	0.04	0.10		0.07		0.06	0.00
1987		0.09		0.10	0.08	0.04	0.10		0.08	0.06	0.07	0.07	0.10		0.04		0.07	0.00
1988		0.11	0.08	0.14	0.12	0.07	0.12		0.10	0.04	0.11	0.01	0.10		0.06		0.08	0.00
1989		0.12		0.11	0.10	0.08	0.07		0.08	0.04	0.09	0.07	0.10		0.07		0.05	0.00
1990		0.08		0.16	0.12	0.07	0.09	0.02	0.09	0.02	0.11	0.37	0.11	0.22	0.06		0.12	0.00
1991		0.09		0.15	0.12	0.10	0.08	0.04	0.09	0.04	0.10	0.08	0.13	0.26	0.13		0.05	0.00
1992		0.09		0.18	0.08	0.08	0.05	0.04	0.09	0.03	0.11	0.06	0.11	0.22	0.10		0.06	
1993		0.70	0.39	0.71	0.71	0.71	0.70	0.69	0.70	0.68	0.70	0.70	0.69		0.73		0.69	0.00
1994		0.34		0.37	0.36	0.35	0.34	0.34	0.35	0.34	0.38	0.32	0.35	0.00	0.32		0.36	0.00
1995	0.76	0.78		0.79	0.77	0.77	0.77	0.76	0.76	0.76	0.78	0.71	0.76	0.00	0.77	0.78	0.77	
1996	0.06	0.09		0.19	0.11	0.11	0.06	0.04	0.09	0.03	0.17	0.08	0.10	0.32	0.11	0.14	0.07	0.00
1997	0.07	0.09	0.08	0.17	0.11	0.11	0.07	0.04	0.08	0.06	0.17		0.07		0.06	0.14	0.06	
1998	0.38			0.48	0.41		0.39	0.38		0.37			0.38		0.38	0.46	0.36	0.00
1999					0.10												0.05	

BIBLIOGRAPHY

Aassve, A., & Lappegard, T. (2008). Childcare Cash Benefits and Fertility Timing in Norway. *European Journal of Population / Revue européenne de Démographie*, *25*(1), 67–88.

Abendroth, A., & Den Dulk, L. (2011). Support for the work-life balance in Europe: the impact of state, workplace and family support on work-life balance satisfaction. *Work, Employment & Society*, *25*(2), 234–256.

Abendroth, A., Van der Lippe, T., & Maas, I. (2012). Social support and the working hours of employed mothers in Europe: The relevance of the state, the workplace, and the family. *Social science research*, *41*(3), 581–97.

Adserà, A. (2004). Changing fertility rates in developed countries. The impact of labor market institutions. *Journal of Population Economics*, *17*(1), 17–43.

Ahn, N., & Mira, P. (2002). A note on the changing relationship between fertility and female employment rates in developed countries. *Journal of Population Economics*, *15*(4), 667–682.

Akgunduz, Y. E., & Plantenga, J. (2013). Labour market effects of parental leave in Europe. *Cambridge Journal of Economics*, *37*(4), 845–862.

Albrecht, J., & Edin, P. (1999). Career interruptions and subsequent earnings: a reexamination using Swedish data. *Journal of Human Resources*, *34*(2), 294–311.

Apps, P., & Rees, R. (2004). Fertility, Taxation and Family Policy. *Scandinavian Journal of Economics*, *106*(4), 745–763.

Atkinson, A., & Brandolini, A. (2001). Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries as a case study. *Journal of Economic Literature*, *39*(3), 771–799.

Bates, D., Maechler, M., & Dai, B. (2010). *lme4: Linear mixed-effects models using S4 classes*.

Becker, G. (1965). A Theory of the Allocation of Time. *The economic journal*, *75*(299), 493–517.

Becker, G. (1985). Human capital, effort, and the sexual division of labor. *Journal of labor economics*, *3*(1), S33–S58.

Becker, G. S. (1991). *A Treatise on the Family*. Cambridge: Harvard University Press.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics. Identifying Influential Data and Sources of Collinearity*. Wiley.

Bernhardt, E. M. (1993). Fertility and Employment. *European Sociological Review*, *9*(1), 25–42.

Betson, D., & Van der Gaag, J. (1984). Working Married Women and the Distribution of Income. *The Journal of Human Resources*, *19*(4), 532–543.

Bickel, P., Hammel, E., & O'Connell, J. (1975). Sex Bias in Graduate Admissions: Data From Berkeley. *Science*, *187*(4175), 398–404.

Björklund, A. (1992). Rising Female Labour Force Participation and the Distribution of Family Income - the Swedish Experience. *Acta Sociologica*, *35*(4), 299–309.

Blake, J. (1965). Demographic Science and the Redirection of Population Policy. In M. Sheps, & J. Ridley (Eds.) *Public Health and Population Change*. Pittsburgh.

Blanchet, D., & Pennec, S. (1993). A simple model for interpreting cross-tabulations of family size and women's labour force participation. *European Journal of Population*, *9*(2), 121–142.

Blau, D., & Robins, P. (1991). Child care demand and labor supply of young mothers over time. *Demography*, *28*(3), 333–351.

Blau, D. M., & Robins, P. K. (1988). Child-Care Costs and Family Labor Supply. *The Review of Economics and Statistics*, *70*(3), 374.

Blau, D. M., & Robins, P. K. (1989). Fertility, employment, and child-care costs. *Demography*, *26*(2), 287–99.

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *The Journal of Economic Perspectives*, *14*(4), 75–99.

Blossfeld, H.-P., & Drobnič, S. (2001). *Careers of Couples in Contemporary Society*. Oxford: Oxford University Press.

Blyth, C. R. (1972). On Simpson's Paradox and the Sure-Thing Principle. *Journal of the American Statistical Association*, *67*(338), 364.

Boeckmann, I., Budig, M., Misra, J., Evertsson, M., Gauthier, A. H., Gerstel, N., Gornick, J., Grunow, D., & Klüsener, S. (2012). Work-Family Policy Indicators. Tech. rep.

Bradley, K. (2000). The incorporation of women into higher education: Paradoxical outcomes? *Sociology of education*, *73*(1), 1–18.

Brambor, T., Clark, W. R., & Golder, M. (2005). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, *14*(1), 63–82.

Breen, R., & Salazar, L. (2009). Has Increased Women's Educational Attainment Led to Greater Earnings Inequality in the United Kingdom? A Multivariate Decomposition Analysis. *European Sociological Review*, *26*(2), 143–157.

Brewster, K., & Rindfuss, R. (2000). Fertility and women's employment in industrialized countries. *Annual Review of Sociology*, *26*(1), 271–296.

Bruning, G., & Plantenga, J. (1999). Parental Leave and Equal Opportunities: Experiences in Eight European Countries. *Journal of European Social Policy*, *9*(3), 195–209.

Budig, M. J., Misra, J., & Boeckmann, I. (2012). The Motherhood Penalty in Cross-National Perspective: The Importance of Work-Family Policies and Cultural Attitudes. *Social Politics: International Studies in Gender, State & Society*, *19*(2), 163–193.

Burtless, G. (1999). Effects of growing wage disparities and changing family composition on the US income distribution. *European Economic Review*, *43*(4-6), 853–865.

Burtless, G. (2009). Demographic Transformation and Economic Inequality. In W. Salverda, B. Nolan, & T. Smeeding (Eds.) *The Oxford Handbook of Economic Inequality*, chap. 18, (pp. 435–454). Oxford: Oxford University Press.

Callan, T., Nolan, B., Neill, D. O., & Sweetman, O. (1998). Female Labour Supply and Income Inequality in Ireland.

Cancian, M., & Reed, D. (1999). The impact of wives' earnings on income inequality: issues and estimates. *Demography*, *36*(2), 173–84.

Cancian, M., & Schoeni, R. F. (1998). Wives' earnings and the level and distribution of married couples' earnings in developed countries. *Journal of Income Distribution*, *8*(1), 45–61.

Cantillon, B. (2011). The paradox of the social investment state: growth, employment and poverty in the Lisbon era. *Journal of European Social Policy*, *21*(5), 432–449.

Cantillon, B., Ghysels, J., Mussche, N., & Van Dam, R. (2001). Female employment differences, poverty and care provisions. *European Societies*, *3*(4), 447–469.

Cartwright, N. (2002). The Limits of Causal Order, From Economics to Physics. In U. Mäki (Ed.) *Fact and Fiction in Economics: Models, Realism, and Social Construction*, (pp. 137–151). Cambridge: Cambridge University Press.

Cartwright, N. (2007). Counterfactuals in Economics: A Commentary. In J. K. Campbell, M. O'Rourke, & H. Silverstein (Eds.) *Causation and Explanation*, chap. 10, (p. 191). Cambridge, Massachusetts: MIT PRess.

Casey, T., & Maldonado, L. (2012). Worst Off: Single-Parent Families inn the United States. A Cross-National Comparison of Single Parenthood in the U.S. and Sixteen Other High-Income Countries. Tech. Rep. December, Legal Momentum, New York.

Charles, M. (2011). A World of Difference: International Trends in Women's Economic Status. *Annual Review of Sociology*, *37*(1), 355–371.

Charles, M., Buchmann, M., Halebsky, S., Powers, J. M., & Smith, M. M. (2001). The Context of Women's Market Careers: A Cross-National Study. *Work and Occupations*, *28*(3), 371–396.

Cleveland, G., Gunderson, M., & Hyatt, D. (1996). Child Care Costs and the Employment Decision of Women: Canadian Evidence. *The Canadian Journal of Economics*, *29*(1), 132.

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, Massachusetts: Harvard University Press.

Coltrane, S. (2000). Research on Household Labor: Modeling and Measuring the Social Embeddedness of Routine Family Work. *Journal of Marriage and Family*, *62*(November), 1208–1233.

Connelly, R. (1992). The Effect of Child Care Costs on Married Women's Labor Force Participation. *The Review of Economics and Statistics*, *74*(1), 83.

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, *19*(1), 15–18.

Cooke, L. (2011). *Gender-class equality in political economies*. New York: Routledge.

Cooke, L., & Baxter, J. (2010). Families in International Context: Comparing Institutional Effects Across Western Societies. *Journal of Marriage and Family*, *72*(3), 516–536.

Costa, D. (2000). From mill town to board room: the rise of women's paid labor. *Journal of Economic Perspectives*, *14*(4), 101–122.

Cox, D. (1984). Interaction. *International Statistical Review*, *52*(1), 1–31.

Cramer, J. C. (1980). Fertility and female employment: Problems of causal direction. *American Sociological Review*, *45*(2), 167–190.

Crawley, M. J. (2007). *The R Book*. Wiley.

Cunningham, M. (2008). Changing Attitudes toward the Male Breadwinner, Female Homemaker Family Model: Influences of Women's Employment and Education over the Lifecourse. *Social Forces*, *87*(1), 299–323.

Daly, M. (2000). A fine balance: Women's labour market participation in International Comparison. In W. Scharpf, Fritz, & V. Schmidt (Eds.) *Welfare and work in open economy, Vol. 2: Diverse responses to common challenges*. Oxford: Oxford University Press.

Davidov, E., Schmidt, P., & Biliet, J. (Eds.) (2011). *Cross-Cultural analysis. Methods and Applications*. Routledge.

De Graaf, N. D., Need, A., & Ultee, W. (2000). Levensloop en kerkverlating. Een nieuwe en overkoepelende verklaring voor enkele empirische regelmatigheden. *Mens & Maatschappij*, *75*(3), 229–257.

Del Boca, D., & Locatelli, M. (2006). *The Determinants of Motherhood and Work Status: A Survey*. 2414. IZA Discussion Paper Series.

Del Boca, D., Pasqua, S., & Pronzato, C. (2009). Motherhood and market work decisions in institutional context: a European perspective. *Oxford Economic Papers*, *61*(Supplement 1), i147–i171.

Dingeldey, I. (2001). European Tax Systems and their Impact on European Tax Systems and their Impact on Family. *Journal of Social Policy*, *30*(04), 653–672.

Diprete, T. A., Morgan, S. P., Engelhardt, H., & Pacalova, H. (2003). Do Cross-National Differences in the Costs of Children Generate Cross-National Differences in Fertility Rates? *Population Research and Policy Review*, *22*(5/6), 439–477.

Drobnič, S. (2000). The effects of children on married and lone mothers' employment in the United States and (West) Germany. *European Sociological Review*, *16*(2), 137–157.

Drobnič, S., Blossfeld, H., & Rohwer, G. (1999). Dynamics of women's employment patterns over the family life course: A comparison of the United States and Germany. *Journal of Marriage and the Family*, *61*(1), 133–146.

Engelhardt, H., Kögel, T., & Prskawetz, A. (2004). Fertility and women's employment reconsidered: a macro-level time-series analysis for developed countries, 1960-2000. *Population studies*, *58*(1), 109–20.

Engelhardt, H., & Prskawetz, A. (2005). *A pooled time-series analysis on the relation between fertility and female employment*. European Demog aphic Research Papers.

England, P., & Farkas, G. (1986). *Households, employment and gender: A social, economic, and demographic view*. New York: Aldine.

England, P., Gornick, J., & Shafer, E. (2012). Women's employment, education, and the gender gap in 17 countries. *Monthly Labor Review*, (April), 20–29.

Esping-Andersen, G. (1999). *Social Foundations of Postindustrial Economics*. Oxford: Oxford University Press.

Esping-Andersen, G. (2003). *Why we need a new welfare state*. Oxford: Oxford University Press.

Esping-Andersen, G. (2007). Sociological Explanations of Changing Income Distributions. *American Behavioral Scientist*, *50*(5), 639–658.

Esping-Andersen, G. (2009). *The Incomplete Revolution: Adapting to Women's New Roles*. Polity Press.

Evertsson, M., & Grunow, D. (2012). Women's work interruptions and career prospects in Germany and Sweden. *International Journal of Sociology and Social Policy*, *32*(9), 561–575.

Fawcett, C. D., & Lee, A. (1902). A second study of the variation and correlation of the human skull, with special reference to the Nagada crania. *Biometrika*, *1*(4), 408–467.

Folbre, N., Gornick, J. C., Connolly, H., & Munzi, T. (2013). Women's Employment, Unpaid Work and Economic Inequality. In J. C. Gornick, & M. Jäntti (Eds.) *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*, (pp. 234–260). Stanford, California.

Fox, J. (2002). *An R and S-Plus Companion to Applied Regression*. Sage.

Fuwa, M. (2004). Macro-level gender inequality and the division of household labor in 22 countries. *American Sociological Review*, *69*(6), 751–767.

Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *15*, 246–263.

Galton, F. (1896). Note to the Memoir by Professor Karl Pearson, F.R.S., on Spurious Correlation. *Proceedings of the Royal Society of London (1854-1905)*, *60*(1), 498–502.

Ganzeboom, H. B., Treiman, D. J., & Ultee, W. (1991). Comparative intergenerational stratification research: Three generations and beyond. *Annual Review of Sociology*, *17*, 277–302.

Gauthier, A. H. (1996). *The State and the Family. A comparative Analysis of Family Policies in Industrialized Countries*. Oxford: Oxford University Press.

Gauthier, A. H. (2003). Comparative Family Benefits Database.

Gauthier, A. H. (2007). The impact of family policies on fertility in industrialized countries: a review of the literature. *Population Research and Policy Review*, *26*(3), 323–346.

Gauthier, A. H. (2010). Comparative Family Policy Database, Version 3.

Gauthier, A. H., & Bortnik, A. (2001). Comparative Maternity, Parental, and Childcare Database, Version 2 (Online).

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical methods for social research. Cambridge University Press.

Gheng, Z. (1992). Collapsibility of Relative Risk in Contingency Tables with a Response Variable. *Journal of the Royal Statistical Society. Series B (Methodological)*, *54*(2), 585–593.

Ghysels, J., & Van Lancker, W. (2011). The unequal benefits of activation: an analysis of the social distribution of family policy among families with young children. *Journal of European Social Policy*, *21*(5), 472–485.

Gini, C. (1912). Variabilità e mutabilità. In E. Pizetti, & T. Salvemini (Eds.) *Reprinted in Memorie di metodologica statistica*. Rome: Libreria Eredi Virgilio Veschi.

Goldin, C. (1990). *Understanding the Gender Gap*. Oxford, UK: Oxford University Press.

Goldin, C. (2006). The Quiet Revolution That Transformed Women's Employment, Education, and Family. *American Economic Review*, *96*(2), 1–21.

Good, I. J., & Mittal, Y. (1987). The Amalgamation and Geometry of Two-by-Two Contingency Tables. *The Annals of Statistics*, *15*(2), 694–711.

Gorlich, D., & De Grip, A. (2008). Human capital depreciation during hometime. *Oxford Economic Papers*, *61*(Supplement 1), i98–i121.

Gornick, J., & Jacobs, J. (1998). Gender, the welfare state, and public employment: a comparative study of seven industrialized countries. *American Sociological Review*, *63*(5), 688–710.

Gornick, J., Meyers, M., & Ross, K. (1998). Public policies and the employment of mothers: A cross-national study. *Social Science Quarterly*, *79*(1), 35–54.

Gornick, J. C., & Jäntti, M. (2012). Child poverty in cross-national perspective: Lessons from the Luxembourg Income Study. *Children and Youth Services Review*, *34*(3), 558–568.

Gornick, J. C., & Meyers, M. K. (2003). *Families that work. Policies for Reconciling Parenthood and Employment*. New York: Russel Sage Foundation.

Gornick, J. C., Meyers, M. K., & Ross, K. E. (1997). Supporting the Employment of Mothers: Policy Variation Across Fourteen Welfare States. *Journal of European Social Policy*, *7*(1), 45–70.

Gornick, J. C., Meyers, M. K., & Ross, K. E. (1999). Public Childcare, Parental Leave, and Employment. In D. Sainsbury (Ed.) *Gender and Welfare State Regimes*, (pp. 117–145).

Gregory, M. (2009). Gender and Economic Inequality. In W. Salverda, B. Nolan, & T. M. Smeeding (Eds.) *The Oxford Handbook of Economic Inequality*, chap. 12, (pp. 284–312). Oxford University Press.

Grindstaff, C. (1988). Adolescent marriage and childbearing: the long-term economic outcome, Canada in the 1980s. *Adolescence*, *23*(89), 45–58.

Gronau, R. (1982). Inequality of Family Income: Do Wives' Earnings Matter? *Population and Development Review*, *8*, 119–136.

Grunow, D., Hofmeister, H., & Buchholz, S. (2006). Late 20th-Century Persistence and Decline of the Female Homemaker in Germany and the United States. *International Sociology*, *21*(1), 101–131.

Gustafsson, S., & Wetzels, C. (1996). Women's labor force transitions in connection with childbirth: A panel data comparison between Germany, Sweden and Great Britain. *Journal of Population Economics*, *9*(3), 223–246.

Harkness, S. (2013). Women's Employment and Household Income Inequality. In J. C. Gornick, & M. Jäntti (Eds.) *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*, (pp. 207–233). Stanford, California: Stanford University Press.

Harkness, S., Machin, S., & Meguir, C. (1996). Women's Pay and Household Incomes in Britain, 1979-1991. In J. Hills (Ed.) *New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom*. New York: Cambridge University Press.

Harrell Jr., F. E. (2001). *Regression Modeling Strategies. With Applications to Linear Models, Logistics Regression, and Survival Analysis*. New York: Springer.

Hechter, M., & Kanazawa, S. (1997). Sociological Rational Choice Theory. *Annual Review of Sociology*, *23*(1), 191–214.

Heckman, J. J. (1974). Effects of Child-Care Programs on Women's Work Effort. *Journal of Political Economy*, *82*(S2), S136.

Hedström, P. (2005). *Dissecting The Social. On Principles of Analytical Sociology*. New York: Cambridge University Press.

Hook, J. L. (2010). Gender Inequality in the Welfare State: Sex Segregation in Housework, 1965 to 2003. *American Journal of Sociology*, *115*(5), 1480–1523.

Hox, J. J. (2010). *Multilevel Analysis. Techniques and Applications*. New york: Routledge, second ed.

ILO (1924). Family allowances. The Remuneration of Labour According to Need. Tech. Rep. 13, International Labour Office, Geneva.

Jacobsen, J., & Levin, L. (1995). Effects of intermittent labor force attachment on women's earnings. *Monthly Labor Review*, (September), 14–19.

Jäntti, M. (1997). Inequality in Five Countries in the 1980s: The Role of Demographic Shifts, Markets and Government Policies. *Economica*, *64*(255), 415–440.

Jaumotte, F. (2003). Labour Force Participation of Women: Empirical Evidence on the role of policy and other determinants in OECD countries. *OECD Economic Studies*, *2003*(2), 51–108.

Jenkins, S. P., & Van Kerm, P. (2009). *The Measurement of Economic Inequality*. Oxford University Press.

Joesch, J. M. (1997). Paid Leave and the Timing of Women's Employment Before and After Birth. *Journal of Marriage and the Family*, *59*(4), 1008.

Kenworthy, L. (1999). Do social-welfare policies reduce poverty? A cross-national assessment. *Social Forces*, *77*(3), 1119–1139.

Kittel, B. (2005a). De Paradox Van de Modal Mover en de Discussie of Een Methodologie Gek Kan Zijn. *Sociologie*, *1*(3).

Kittel, B. (2005b). Enkele verduidelijkingen in het het debat over de gekheid van een methodologie. *Sociologie*, *1*(2).

Kittel, B. (2006). A Crazy Methodology?: On the Limits of Macro-Quantitative Social Science Research. *International Sociology*, *21*(5), 647–677.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press, second ed.

Kögel, T. (2004). Did the association between fertility and femaly employment within OECD countries really change its sign? *Journal of Population Economics*, *17*(1), 45:65.

Kollmeyer, C. (2013). Family Structure, Female Employment, and National Income Inequality: A Cross-National Study of 16 Western Countries. *European Sociological Review*, *29*(4), 816–827.

Korpi, W., Ferrarini, T., & Englund, S. (2013). Women's Opportunities under Different Family Policy Constellations: Gender, Class, and Inequality Tradeoffs in Western Countries Re-examined. *Social Politics: International Studies in Gender, State & Society*, *20*(1), 1–40.

Kotsadam, A., & Finseraas, H. (2013). Causal Effects of Parental Leave on Adolescents' Household Work. *Social Forces*, *92*(1), 329–351.

Kreft, I., & Leeuw, J. D. (1998). Introducing Multilevel Modeling. *London: Sage. Lancaster, T*, *84*, 881–896.

Lakatos, I. (1978). Falsification and the methodology of scientific research programmes. In J. Worall, & G. Currie (Eds.) *The Methodology of Scientific Research Programmes*, chap. Falsificat, (pp. 8–101). Cambridge: Cambridge University Press.

Lam, D. (1997). Demographic variables and income inequality. In M. Rosenzweig, & O. Stark (Eds.) *Handbook of population and family economics*, vol. 1, chap. Chapter 18, (pp. 1015–1059). Elsevier.

Lammers, J., Pelzer, B., Hendrickx, J., & Eisinga, R. (2007). *Categorische Data Analyse met SPSS. Inleiding in loglineaire analysetechnieken*. Uitgeverij van Gorcum.

Lancker, W. V., & Ghysels, J. (2012). Who benefits? The social distribution of subsidized childcare in Sweden and Flanders. *Acta Sociologica*, *55*(2), 125–142.

Langford, I. H., & Lewis, T. (1998). Outliers in multilevel data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *161*(2), 121–160.

Lappegard, T. (2012). Couples' Parental Leave Practices: The Role of the Workplace Situation. *Journal of Family and Economic Issues*, *33*(3), 298–305.

Levels, M. (2011). *Abortion laws in Europe between 1960 and 2010. Legislative developments and their consequences for women's reproductive decision-making*. Phd dissertation, Radboud University.

Levels, M., Need, A., Nieuwenhuis, R., Sluiter, R., & Ultee, W. (2012). Unintended Pregnancy and Induced Abortion in the Netherlands 1954-2002. *European Sociological Review*, *28*(3), 301–318.

Levels, M., & Nieuwenhuis, R. (2011). Groei van kennis en stagnatie in de empirisch- theoretische sociologie - een onkiese beschouwing. In G. Kraaykamp, M. Levels, & A. Need (Eds.) *Problemen en Theorieen in Onderzoek. Een staalkaart van de hedendaagseempirisch-theoretische sociologie*, (pp. 1–20). Van Gorcum.

Lieberman, E. S. (2005). Nested Analysis as a Mixed-Method Strategy for Comparative Research. *American Political Science Review*, *99*(3), 435–452.

Linos, K. (2011). Diffusion through Democracy. *American Journal of Political Science*, *55*(3), 678–695.

LIS (2013). Luxembourg Income Study. Multiple Countries, Database Accessed Between May 2012 and June 2013.

Machin, S., & Waldfogel, J. (1994). The Decline of the Male Breadwinner. Changing Shares of Husbands' and Wives' Earnings in Family Income.

Mandel, H. (2012). Winners and Losers: The Consequences of Welfare State Policies for Gender Wage Inequality. *European Sociological Review*, *28*(2), 241–262.

Mandel, H. (2013). Up the Down Staircase: Women's Upward Mobility and the Wage Penalty for Occupational Feminization, 1970-2007. *Social Forces*, *91*(4), 1183–1207.

Mandel, H., & Semyonov, M. (2006). A Welfare State Paradox: State Interventions and Women's Employment Opportunities in 22 Countries. *American Journal of Sociology*, *111*(6), 1910–1949.

Mastekaasa, A., & Birkelund, G. (2011). The equalizing effect of wives' earnings on inequalities in earnings among households. Norway 1974-2004. *European Societies*, *13*(2), 219–237.

Matysiak, A., & Vignoli, D. (2008). Fertility and Women's Employment: A Meta-analysis. *European Journal of Population*, *24*, 363–384.

McCall, L., & Percheski, C. (2010). Income Inequality: New Trends and Research Directions. *Annual Review of Sociology*, *36*(1), 329–347.

Messick, D., & Van de Geer, J. (1981). A reversal paradox. *Psychological Bulletin*, *90*(3), 582–593.

Mills, M., Rindfuss, R. R., McDonald, P., & Te Velde, E. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human reproduction update*, *17*(6), 848–860.

Mincer, J. (1958). Investment in human capital and personal income distribution. *The Journal of Political Economy*, *66*(4), 281–302.

Mincer, J. (1962). Labor Force Participation of Married Women: A Study of Labor Supply. In U.-N. Bureau (Ed.) *Aspects of Labor Economics*, vol. I, (pp. 63–106). Prince.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Columbia University Press.

Mincer, J., & Ofek, H. (1982). Interrupted work careers: Depreciation and restoration of human capital. *Journal of Human Resources*, *17*(1), 3–24.

Mittal, Y. (1991). Homogeneity of Subpopulations and Simpson's Paradox. *Journal of the American Statistical Association*, *86*(413), 167.

Morgan, K. J., & Zippel, K. (2003). Paid to Care: The Origins and Effects of Care Leave Policies in Western Europe. *Social Politics: International Studies in Gender, State & Society*, *10*(1), 49–85.

Morissens, A., & Sainsbury, D. (2005). Migrants' Social Rights, Ethnicity and Welfare Regimes. *Journal of Social Policy*, *34*(04), 637.

Moss, P., & Deven, F. (1999). *Parental Leave: Progress or Pitfall?*. Brussels: NIDI CBGS Publications.

Myrdal, A. (1941). *Nation and the Family: The Swedish Experiment in Democratic Family and Population Policy*. Cambridge: MIT Press.

Myrdal, A., & Klein, V. (1956). *Women's Two Roles. Home and Work*. London: Routledge Paperback, reprinted ed.

Nieuwenhuis, R., Munzi, T., & Gornick, J. (2013a). Netting Down Gross Earnings Data in the LIS Database: An Evaluation of Two Procedures. *LIS Technical Paper Series*, (#6), 46.

Nieuwenhuis, R., Need, A., & Van der Kolk, H. (2012a). Institutional and Demographic Explanations of Women's Employment in 18 OECD Countries, 1975-1999. *Journal of Marriage and Family*, *74*(June), 614–630.

Nieuwenhuis, R., Need, A., & Van der Kolk, H. (2013b). Family Policies, Women's Earnings, and Between-Household Inequality: Trends in 18 OECD countries from 1981 to 2005. *LIS Working Paper Series*, (#599), 21.

Nieuwenhuis, R., Need, A., & Van der Kolk, H. (2013c). Women's Earnings: Trends in earnings inequality within and between coupled households in 18 OECD countries, 1981-2005. *LIS Working Paper Series*, (#598), 26.

Nieuwenhuis, R., Pelzer, B., & Te Grotenhuis, M. (2012b). *influence.ME: Tools for detecting influential data in mixed effects models. Version 0.9.3*.

Nieuwenhuis, R., Te Grotenhuis, M., & Pelzer, B. (2012c). Influence.ME: tools for detecting influential data in mixed effects models. *R Journal*, *4*(2), 38–47.

OECD (2001). Balancing work and Family Life: Helping Parents into Paid Employment. In *Employment Outlook*, (pp. 129–166).

OECD (2011). *Doing Better for Families*. OECD Publishing.

OECD (2013). OECD.StatExtracts.

Ondrich, J., Spiess, C., & Yang, Q. (1996). Barefoot and in the German kitchen: Federal parental leave and benefit policy and the return to work after childbirth in Germany. *Journal of Political Economy*, *9*(3), 247–266.

Oppenheimer, V. (1988). A theory of marriage timing. *American Journal of Sociology*, *94*(3), 563–591.

Oppenheimer, V. (1994). Women's Rising Employment and the Future of the Family in Industrial Societies. *Population and Development Review*, *20*(2), 293–342.

Pasqua, S. (2002). Wives' work and income distribution in European countries. *The European Journal of Comparative Economics*, *5*(2), 197–226.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Pearson, K. (1896). Mathematical Contributions to the Theory of Evolution.–On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. *Proceedings of the Royal Society of London (1854-1905)*, *60*(1), 489–498.

Pearson, K., Lee, A., & Bramley-Moore, L. (1899). Mathematical Contributions to the Theory of Evolution. VI. Genetic (Reproductive) Selection: Inheritance of Fertility in Man, and of Fecundity in Thoroughbred Racehorses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *192*(1899), 257.

Pennings, P. (2005). Wat mogen we verwachten van macrokwantitatief sociaal-wetenschappelijk onderzoek? *Sociologie*, *1*(2).

Pettit, B., & Hook, J. (2005). The Structure of Women's Employment in Comparative Perspective. *Social Forces*, *84*(2), 779–801.

Pettit, B., & Hook, J. L. (2009). *Gendered Tradeoffs. Family, Social Policy, and Economic Inequality in Twenty-One Countries*. New York: Russel Sage Foundation.

Presser, H., & Baldwin, W. (1980). Child care as a constraint on employment: Prevalence, correlates, and bearing on the work and fertility nexus. *American Journal of Sociology*, *85*(5), 1202–1213.

R Core Team (2012). R: A language and environment for statistical computing.

Rahmqvist, M. (2006). The close relation between birth, abortion and employment rates in Sweden from 1980 to 2004. *Social Science & Medicine*, *63*(5), 1262–1266.

Ribar, D. (1992). Child care and the labor supply of married women: Reduced form evidence. *Journal of Human Resources*, *27*(1), 134–165.

Ribar, D. (1995). A structural model of child care and the labor supply of married women. *Journal of Labor Economics*, *13*(3), 558–597.

Rindfuss, R. R., & Brewster, K. L. (1996). Childrearing and Fertility. *Population and Development Review*, *22*, 258.

Rindfuss, R. R., Guzzo, K. B., & Philip, S. (2003). The changing institutional context of low fertility. *Population Research and Policy Review*, *22*(5-6), 411–438.

Rocha, J. D., & Fuster, L. (2006). Why are fertility rates and female employment ratios positively correlated across OECD countries? *International Economic Review*, *47*(4), 1187–1222.

Rönsen, M., & Sundström, M. (1996). Maternal employment in Scandinavia: A comparison of the after-birth employment activity of Norwegian and Swedish women. *Journal of population economics*, *9*, 267–285.

Rosenfeld, R. A., Trappe, H., & Gornick, J. C. (2004). Gender and Work in Germany: Before and After Reunification. *Annual Review of Sociology*, *30*(1), 103–124.

Saari, D. G. (1995). A chaotic Exploration of Aggregation Paradoxes. *SIAM Review*, *37*(1), 37.

Sainsbury, D., & Morissens, A. (2012). Immigrants' Social Rights across Welfare States. In *Welfare States and Immigrant Rights. The Politics of Inclusion and Exclusion*, chap. 6. Oxford: Oxford University Press.

Samuels, M. L. (1993). Simpson's Paradox and Related Phenomena. *Journal of the American Statistical Association*, *88*(421), 81.

Sarkar, D. (2008). *Lattice. Multivariate Data Visualization with R*. Springer.

Schmitt, H., & Scholz, E. (2005). *The Mannheim Eurobarometer Trend File, 1970 - 2002*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (ICPSR).

Schober, P. S. (2013). Maternal labor market return and domestic work after childbirth in Britain and Germany. *Community, Work & Family*, (August), 1–20.

Schwarz, P. (2012). Tax disincentives and female employment in Organisation for Economic Co-operation and Development (OECD) countries. *Journal of European Social Policy*, *22*(1), 17–29.

Semyonov, M. (1980). The social context of women's labor force participation: A comparative analysis. *American Journal of Sociology*, *86*(3), 534–550.

Shorrocks, A. (1983). The impact of income components on the distribution of family incomes. *The Quarterly Journal of Economics*, *98*(2), 311–326.

Simpson, E. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *13*(2), 238–241.

Singer, J. D., & Willet, J. B. (2003). *Applied Longitudinal Data Analysis*. Oxford University Press.

Sleebos, J. (2003). Low fertility rates in OECD countries: facts and policy responses.

Smith, T., Van Marsden, P., Hout, M., & Kim, J. (2010). *General Social Surveys, 1972 - 2010*. Chicago: National Opinion Research Center.

Snijders, T., & Berkhof, J. (2008). Diagnostic Checks for Multilevel Models. In J. De Leeuw, & E. Meijer (Eds.) *Handbook of multilevel analysis*, vol. 2, chap. Chapter 3, (pp. 141–175). New York: Springer.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis, an introduction to basic and advanced multilevel modelling*. Sage.

Spillman, L., & Strand, M. (2013). Interest-Oriented Action. *Annual Review of Sociology*, *39*(1), 85–104.

Steele, F., Goldstein, H., & Browne, W. (2004). A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modelling*, *4*(2), 145–159.

Stier, H., Lewin-Epstein, N., & Braun, M. (2001). Welfare Regimes, Family Supportive Policies, and Women's Employment along the Life Course. *American Journal of Sociology*, *106*(6), 1731–1760.

Stone, P. (2008). *Opting Out? Why Women Really Quit Careers and Head Home*. University of California Press.

Stycos, J., & Weller, R. (1967). Female Working Roles and Fertility. *Demography*, *4*(1), 210–217.

Sundström, M., & Stafford, F. P. (1992). Female labour force participation, fertility and public policy in Sweden. *European journal of population*, *8*(3), 199–215.

Sweeney, M. (2004). The Changing Importance of White Women's Economic Prospects for Assortative Mating. *Journal of Marriage and Family*, *66*(November), 1015–1028.

Sweet, J. (1981). Work and Fertility. In G. Fox (Ed.) *Decision-Making, Family Planning and Child-Bearing*, chap. 9, (pp. 197–218). New York: Sage.

Thévenon, O. (2011). Family policies in OECD countries: a comparative analysis. *Population and Development Review*, *37*(1), 57–87.

Thévenon, O., & Luci, A. (2012). Reconciling Work, Family and Child Outcomes: What Implications for Family Support Policies? *Population Research and Policy Review*, *31*(6), 855–882.

Tremblay, A. (2012). LMERConvenienceFunctoins: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions.

Ultee, W. (2005). Niks Raar! *Sociologie*, *1*(3).

Ultee, W., & Luijkx, R. (1998). Jewish-Gentile intermarriage in six european cities 1900-1940. Explaining differences and trends. *The Netherlands journal of social sciences*, *34*(2), 165–195.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal Of Statistical Software*, *45*(3), 1–67.

Van der Lippe, T., De Ruijter, J., De Ruijter, E., & Raub, W. (2011). Persistent Inequalities in Time Use between Men and Women: A Detailed Look at the Influence of Economic Circumstances, Policies, and Culture. *European Sociological Review*, *27*(2), 164–179.

Van der Lippe, T., & Van Dijk, L. (2002). Comparative Research on Women's Employment. *Annual Review of Sociology*, *28*(1), 221–241.

Van der Meer, T., Te Grotenhuis, M., & Pelzer, B. (2010). Influential Cases in Multilevel Modeling: A Methodological Comment. *American Sociological Review*, *75*(1), 173–178.

Vandenbroucke, F., & Vleminckx, K. (2011). Disappointing poverty trends: is the social investment state to blame? *Journal of European Social Policy*, *21*(5), 450–471.

Verhagen, J. (2012). *Bayasian Item Response Theory Models for Measurement Variance*. Ph.D. thesis, University of Twente.

Wagner, C. H. (1982). Simpson's Paradox in Real Life. *The American Statistician*, *36*(1), 46.

Waite, L., & Stolzenberg, R. (1976). Intended childbearing and labor force participation of young women: Insights from nonrecursive models. *American Sociological Review*, *41*(2), 235–252.

Waldfogel, J. (1997). The effect of children on women's wages. *American sociological review*, *62*(2), 209–217.

Waldfogel, J. (2001). International policies toward parental leave and child care. *The Future of Children*, *11*(1), 99–111.

Waldfogel, J., Higuchi, Y., & Abe, M. (1999). Family leave policies and women's retention after childbirth: Evidence from the United States, Britain, and Japan. *Journal of Population Economics*, *12*(4), 523–545.

Wellington, A. (1993). Changes in the male/female wage gap, 1976-85. *Journal of Human Resources*, *28*(2), 383–411.

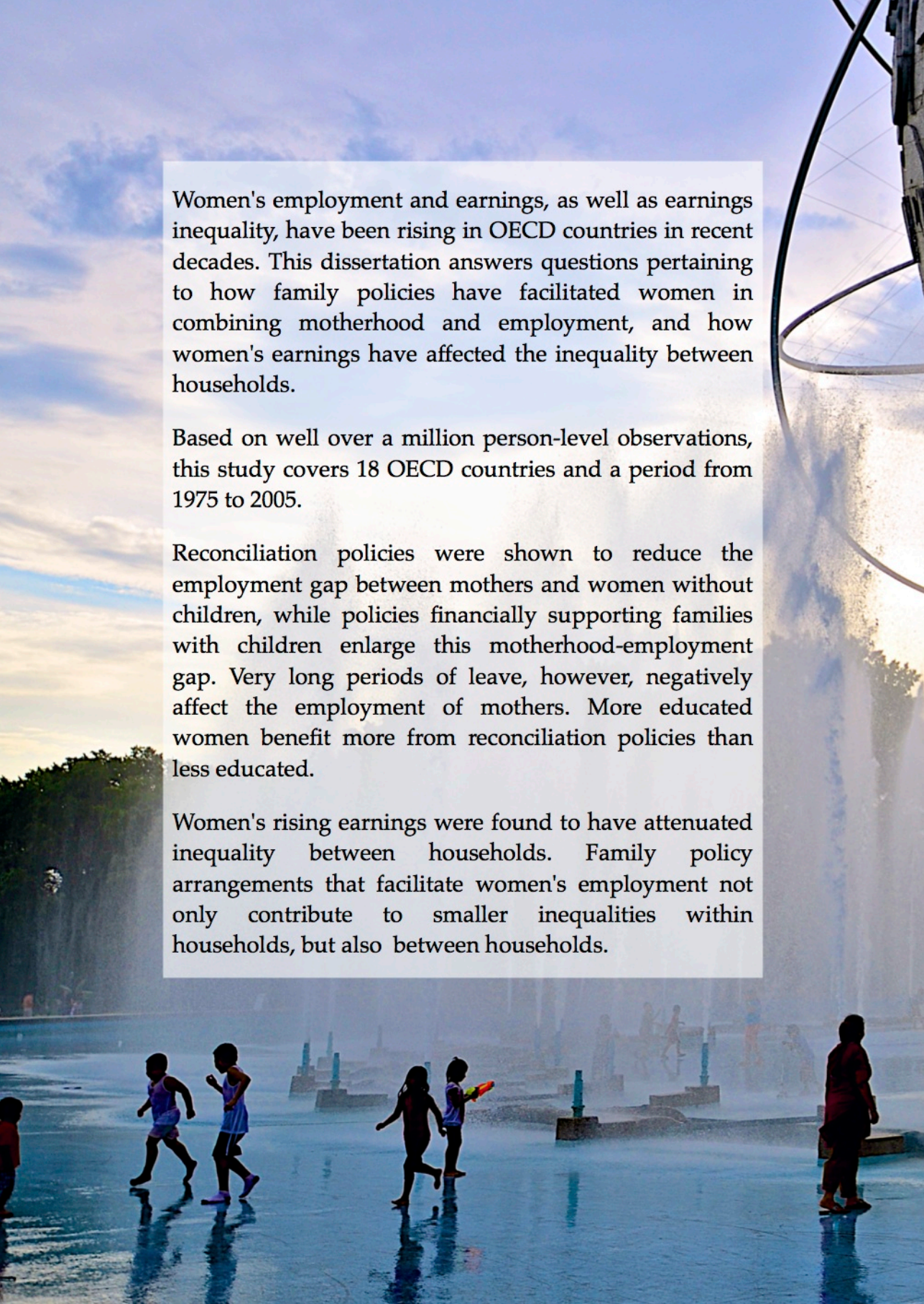
Williams, J. C., & Boushey, H. (2010). The Three Faces of Work-Family Conflict. The Poor, the Professionals, and the Missing Middle. Tech. Rep. January, Center for American Progress.

Yamaguchi, K. (1991). *Event History Analysis*. Applied Social Research Methods Series. Sage.

Yasuoka, M. (2012). Effectiveness of Child Care Policies in an Economy with Child Care Services. *Modern Economy*, *03*(05), 658–670.

Yerkes, M. (2013). Choice or Constraint? Women's weekly working hours in comparative perspective. *Sociologia, Problemas E Práticas*, (72), 9–30.

Yule, G. U. (1903). Notes on the Theory of Association of Attributes in Statistics. *Biometrika*, *2*(2), 121–134.



Women's employment and earnings, as well as earnings inequality, have been rising in OECD countries in recent decades. This dissertation answers questions pertaining to how family policies have facilitated women in combining motherhood and employment, and how women's earnings have affected the inequality between households.

Based on well over a million person-level observations, this study covers 18 OECD countries and a period from 1975 to 2005.

Reconciliation policies were shown to reduce the employment gap between mothers and women without children, while policies financially supporting families with children enlarge this motherhood-employment gap. Very long periods of leave, however, negatively affect the employment of mothers. More educated women benefit more from reconciliation policies than less educated.

Women's rising earnings were found to have attenuated inequality between households. Family policy arrangements that facilitate women's employment not only contribute to smaller inequalities within households, but also between households.