Women, work, and wages:

How family policies continue to shape the gender wage gap

Sandra Elise Bakkevoll

Master's thesis

Abstract

Despite a dramatic convergence in the gender wage gap at the end of the 20th century, there has since been a stability which, considering the social, demographic, and economic changes that should have attenuated the gender inequality, seems particularly puzzling. The thesis addresses some of this puzzling stability by asking the research question: How do family policies intended to help balance work and family obligations affect the gender wage gap, and are the effects different across the wage distribution? In order to answer the research question, I use a large-N quantitative approach with time-series cross-sectional data from 23 member countries of the Organisation for Economic Co-operation and Development (OECD) from 1990 to 2018.

Building on the conceptual debate discussing the disaggregation of the concept of *family policy*, I first separate the effects of policies enabling defamilisation, i.e., facilitating work, from the effects of policies enabling familisation, i.e., reducing work. I expected the former to reduce the gender wage gap, while the latter to increase it and find support for the first expectation and thereby also for the disaggregation of the concept. Furthermore, I build on the literature on the *glass ceiling effect* and expected the effects of the family policy dimensions to affect women differently depending on their placement in the wage distribution. I find that familisation positively and defamilisation negatively affects the gender wage gap for top earners, but I do not find any effect on the gender wage gap for bottom earners.

I conclude that by focusing only on the average trends, may it be family policies or wage gaps, we lose out on important information on what drives the gender wage gap. This information is important both for future research and for policymakers by shedding light on what kinds of policies to pursue in order to decrease economic inequality between genders.

Acknowledgements

Jeg vil først og fremst takke min veileder Georg Picot for god hjelp det siste året. Takk for alle grundige tilbakemeldinger, gode innspill og for å ha holdt meg strukturert selv i min rotete skriveprosess.

Takk til alle fine mennesker på Sofie Lindstrøms hus. Masterstudiene hadde ikke vært det samme uten alle vinlotteri, kakefredager, spontane veggavstemninger, lunsjquizer og lesesal/terapihunder. Takk særlig til Eline, Marte, Lucas, Rebekka, Trygve og Maria på rom 101 (eller «Trygvekontoret») for god stemning både på og utenfor lesesalen, for mental støtte og ikke minst for utrolig god hjelp med masteroppgaven i form av besvarte spørsmål, tittel-workshop og delte koder i R.

En ekstra spesiell takk til Sara for alle sene kvelder fylt av tårer av både latter og frustrasjon over masteroppgaven som ikke ville skrive seg selv. Takk for at du ikke lot meg sitte igjen alene med masterinnspurten de siste ukene.

Takk også til Synne og Ylva for språkvask av oppgaven, og generelt for å ha vært verdens beste venner til tross for min mye manglende tilstedeværelse det siste halvåret.

Sist, men overhodet ikke minst, takk til Simen for tålmodigheten, for gjennomlesing av oppgaven, for å være den aller beste heiagjengen og for å ha passet på at jeg har fått i meg ordentlig mat og hatt rene klær under siste innspurten. Jeg er så utrolig heldig som har deg.

Table of contents

List of tables and figures

List of tables

List of figures

1. INTRODUCTION

1.1. Setting the stage

After a number of societal transformations in the 1970s challenged the traditional gender division of labour, women increasingly joined paid employment and started to catch up with men in terms of both education and work experience (Kleider 2015, 505). This, in turn, led to an immense convergence in the gender wage gap across countries, broadly defined as the difference in wages between men and women (Blau and Kahn 2017, 790; Kunze 2018, 382). Today, men and women are more equal in the labour market than they were just sixty years ago. In addition to women's increasing human capital, the declining wage gap was also attributed to the introduction of family policies specifically intended to help balance work and family commitments for women, thus strengthening their labour force attachment and, in turn, increasing their wage levels. Therefore, because women's labour market outcome was highly associated with potential motherhood, by the early 21st century, most high-income countries had effectively put in a host of family policies with, among others, the goal of achieving economic equality between genders (Kolberg and Esping-Andersen 1991, 4; Olivetti and Petrongolo 2017, 205).

Yet, this dramatic narrowing in the gender wage gap during the late 20th century has been followed by a stability which, in light of the many social, demographic, and economic changes that should have attenuated gender inequality in labour market outcomes, seems particularly puzzling (Boye, Halldén and Magnusson 2017, 595; Cha and Weeden 2014, 458). The persisting differences in earnings between genders therefore remains an important challenge facing democracies. Furthermore, when it comes to the gender wage gap, new puzzles keep emerging. Although women have caught up to the human capital, they still receive lower wages; the gender wage gap is especially large in highly paid and prestigious occupations; and the wage differences are surprisingly large in Scandinavian countries despite their highly developed welfare states (Grönlund and Magnusson 2016, 92). So, why is it that equally educated and productive women are not advancing in the labour market, and how come that even in countries with the highest rate of female employment and most generous welfare systems, a persistent wage gap persists?

In accounting for these puzzling findings, researchers have explored the possibility of a potential paradox in the quest for gender equality in which there is a trade-off between women's economic and occupational achievements and the "family-friendliness of welfare states" (Grönlund and Magnusson 2016, 92). While the family policies are formally committed to advancing gender equality, have they achieved what they set out to do, or are there unintended consequences which further impede women's labour market outcomes? Early research on this potentially detrimental effect of family policy on women's relative earnings have furthermore been criticised for seeing such policies as unidimensional and based mostly on regime typologies (Korpi, Ferrarini and Englund 2013, 4). I will therefore look further into the argument of disaggregating this concept, asking whether we can realistically expect that all family policies will have the same effect on women's relative wages. Furthermore, in light of research finding a larger gender wage gap at the top of the wage distribution, typically interpreted as a "glass ceiling effect", can we expect that the potentially different effects of family policies will affect all women similarly irrespective of their place in the wage distribution?

Based on these observations, the research question for the thesis is therefore:

How do family policies intended to help balance work and family obligations affect the gender wage gap, and are the effects different across the wage distribution?

1.2. Why study the gender wage gap?

This is an important topic to study because if there are unintended consequences of policies attempting to relieve the work-family balance instead increasing gender inequalities, this information is important especially for policymakers. Also, if there are systematic barriers to women's opportunities and achievements in the labour market created and upheld by national policies, this implies that these barriers can be alleviated by the same measures, thus looking away from deterministic explanations of gender essentialism. It is therefore clear that both political scientists and policymakers need to pay attention to what drives these inequalities and how these policies may have unintended consequences and affect women differently depending on their wage levels.

Furthermore, understanding the mechanisms connecting family policies to women's labour market outcomes is especially important because the pressure on the welfare state from major

societal and demographic changes, most notably the ageing population, declining fertility and increasing economic inequality, forces policymakers to think differently about several policies. On top of these challenges, family policy – a policy area which has not undergone drastic changes for decades – faces the additional challenge of accommodating and responding to the changing structures of modern-day families, such as the growing number of single-parent households, same-sex parents and increasing divorce rates (Adema, Clarke and Thévenon 2020, 195; Estévez-Abe 2006, 175; White 2020, 79). Expanding cross-national research in this area is therefore important not only for understanding the heterogeneity in policy effects and inequality more broadly, but also in order to better prepare for the changes to come and work out sustainable solutions decreasing economic inequality both at large and between genders (Hook and Li 2020, 261).

1.3. Research approach and contribution

In order to study how family policies affect the gender wage gap and whether the effects are different for various points in the wage distribution, I use a large-N quantitative approach with time-series cross-sectional data from 23 member countries of the Organisation for Economic Co-operation and Development (OECD) from 1990 to 2018. In addressing the first part of the research question, I deconstruct the concept of family policy into dimensions of familisation and defamilisation and estimate their effects on the gender wage gap. Then, addressing the second part of the research question, I alternate between three different dependent variables from three points of the wage distribution and compare their results.

My contributions are twofold. First, the thesis contributes to the empirical evidence by using a longer time frame and thereby looking both at longer and more recent processes. Second, I contribute to the theoretical framework by detangling the concept of family policies and then applying it to different parts of the wage distribution.

First, by utilising macro-level panel data, I am able to incorporate a time dimension into a research field which have been mostly focused on cross-sections. There have been major advances during the last decades in both data availability and data quality, but most studies rely on microdata from the early 1990s to mid-2000s for industrialised Western European countries (Schäfer and Gottschall 2015, 473). I therefore take a step back and trade off some precision by being able to look at more recent and longer processes. In the light of recent

changes, such as the spread of a separate parental leave exclusively for fathers, and campaigns such as *#metoo* putting gendered labour market issues back on the agenda, it is important to capture these recent changes in order to understand this complex socio-economic phenomenon.

Second, one of the main critiques of research on the gender wage gap has been the conceptualisation of family policies into one single encompassing indicator and the failure to explore how the effects on the gender wage gap might differ for different parts of the wage distribution. I therefore detangle this concept of family policies into two separate dimensions of familisation and defamilisation and test whether there are different effects of policies with the same aim of relieving the work-family balance. Also, using three dependent variables across the wage distribution sheds some light on how these policies affect the gender wage gap differently depending on relative earnings. There has not been much theoretical or empirical research connecting the dimensions of family policy to the differences in wage gaps across the wage distribution. I therefore further contribute to the literature by exploring this connection. This also contributes to explain why there has been a puzzling stagnation in the convergence of the gender wage gap.

1.4. Chapter outline

This thesis is structured as follows:

In the second chapter, I look at previous research and theory explaining what affects the gender wage gap. I first present an overview of early research on the issue, then look closer at the conceptual debate regarding the concept of *family policy* and at theory expecting effects to be different depending on the placement in the wage distribution.

In the third chapter, I present the data and methodological approach. This first section of this chapter presents the three dependent variables, the construction of the independent variables, and gives a descriptive overview of the main variables. In the second section of this chapter, I present the choice of method and estimation techniques, and discuss some important advantages as well as limitations to the approach.

In the fourth section, I present the findings from the analysis. Here, I find that looking at overall family policy generosity gives an inaccurate account of the policies' different effects. I

also find that what affects the gender wage gap for bottom earners is not the same as for the gender wage gap for top earners.

Lastly, in the fifth and final section I discuss the findings in relation to the hypotheses and their implications for both theory and practice. I also give suggestions for future research. I finish with some concluding remarks where I conclude that paying attention to the different effects of family policies for the gender wage gaps across the wage distribution is important in order to better understand what shapes the gender wage gap.

2. PREVIOUS RESEARCH AND THEORY

There is a rich interdisciplinary literature on what drives the gender wage gap with important theoretical framework dating back as far as the 1960s. In the first section of this chapter, I start by giving an overview of previous research and present the most influential theories explaining the gender wage gap. In the second section, I look further into theory arguing in favour of disaggregating the family policy concept into more precise dimensions in terms of the policies' aims and effects and then how these effects are expected to affect the gender wage gap differently across the wage distribution. I argue that doing this provides a better understanding of what drives the wage gap by avoiding the puzzling empirical results that arise when variables are linked to aggregate and average measures. I formulate hypotheses throughout this second section and finish by summarising the hypotheses at the end of the chapter.

2.1. What explains the gender wage gap?

What are some of the main theoretical explanations for the gender wage gap cross-nationally? In this section, I present three traditional and central approaches to understanding the gender wage gap. First, I look at the human capital theory arguing that the wage gap is determined by gender differences in individual endowments. Second, I present the demand-side explanation of gender discrimination in the labour market as hindering women's work-related achievements. Both of these first approaches were developed mainly in the economics field and later incorporated as important explanatory mechanisms in the third approach looking at the constraints and opportunities embedded in family policy. Although there is a larger consensus in the literature today that the different policies should be differentiated, I start here by giving a general review of the early literature on the effects of family policy on the gender wage gap and return afterwards to the argument in favour of disaggregating the concept.

2.1.1. Human capital

The first classic approach to explaining the gender wage gap has been that of human capital, indicating that differences in individual characteristics are the most important drivers of gender differences in labour market outcomes such as wages. The human capital theory puts

forward the idea that people increase their productivity in the labour market, and therefore their earnings, through education and training, i.e., through activities that produce skill, knowledge, and personal growth, and was mainly developed in the economics field by pioneers such as Gary Becker, Jacob Mincer and Solomon Polachek. In line with the concept of *capital*, these activities produce "human, not physical or financial capital because you cannot separate a person from his or her knowledge" (Becker 1993, 16).

By relating the application of economic theory to the human condition, Gary Becker's (1993, 3rd edition) analysis on how an individual's education and training affect their earnings, first published in 1964, became a major contribution to the field. First, education is assumed to raise individual human capital through the mechanisms of providing knowledge, skills, and a way of analysing problems or – in an alternative view – through credentialism, i.e., the way in which "education convey information about the underlying abilities, persistence, and other valuable traits of people" (Becker 1993, 19). This means that a person, simply from having been through education and regardless of results, will be expected to hold a series of redeemable qualities that justifies higher wage levels. Another way of learning is through training and work experience, which Becker further distinguished into the two categories of general and specific on-the-job training. The former relates to skills that can be "useful in many firms besides those providing it" (Becker 1993, 33). Since they are transferable, the general skills acquired increases one's marginal product in various employment. Firm-specific training, on the other hand, leads to unique skills that cannot easily be transferred to other firms, for example proficiency with a specific tool used for a specific job (Becker 1993, 30). Investing in this latter type of training often places greater costs on the specific firm and is seen as an investment in long-term employment.

Therefore, because wages are expected to increase with increases in human capital as this, in turn, increases productivity, this approach provides a convincing explanation for the gender wage gaps that existed before the 70's and for the following convergence in the wage gap as women started attaining both education and work experience (Boye, Halldén and Magnusson 2017, 597; Padavic, Ely and Reid 2020, 62).

Further relating human capital to gender differences in labour market outcomes, Jacob Mincer and Solomon Polachek (1974) extended the human capital theory to include expected work continuity in the explanation. They argued that the difference in productivity between men and women, which in turn leads to a difference in wages, could further be attributed to women

having shorter actual working time because of expected periods of workforce interruptions related to motherhood (Mincer and Polachek 1974, 78). An important aspect is therefore the continuity and discontinuity of work:

> "To the extent that earnings in the labour market are a function of the human-capital stock accumulated by individuals, a sequence of positive investments give rise to growing earning power over the life cycle. When net investment is negative, that is, when market skills are eroded by deprecation, earning power declines" (Mincer and Polachek 1974, 78).

Simply stated, there is a high penalty for workforce interruption on wages because of skill depreciation and forgone experience. Therefore, because women are more likely to spend time away from the labour market due to potential childbearing and care responsibilities, their wages decline as their built-up human capital erodes (Grönlund and Magnusson 2016, 94; Hipp 2020, 252). Mincer and Polachek (1974) further argued that because of this, families were expected to invest more in skill acquisition for men than women, assuming that allocating time and care responsibilities this way would be the most rewarding (89).

This theory was first tested in an early study by Mincer and Ofek (1982) using panel data on wages of married women in the United States. With the data available, they were able to look at the wages for the same women before and after workforce interruptions and test whether these interruptions had led to a statistically significant decline in their wages, as expected by the theory. In their work, they found that real wages at re-entry to the labour market were significantly lower than at the point of labour force withdrawal, and that the decline was greater the longer the interruption (Mincer and Ofek 1982, 16-17). This was interpreted as evidence for the human capital theory in which time away from the workplace erodes human capital. They also found a rapid growth in wages after the return to work, as a restoration or repair of human capital, further supporting the human capital theory as "the reconstruction of occupational skills is more efficient – that is, less costly – than new construction of human capital" and would therefore entitle a faster growth in wages than starting over in a new job (Mincer and Ofek 1982, 4).

Since this study, the importance of individual characteristics – and thus productivity – in explaining the gender wage gap has been explored by numerous studies. This theoretical early work has also been important for the development of research on the gender wage gap in terms of methodology. Building on these theoretical expectations, the economics field developed statistical decomposition methods for testing how much of the gender wage gap is

explained by these observable characteristics and pinpointing which characteristic is the most important (Djurdjevic and Radyakin 2007, 366; Kunze 2018, 373; Nopo 2008, 290). Decomposition methods – the most widely known and used being the Blinder-Oaxaca decomposition method developed in 1973 – generates the counterfactual: "what would men earn if the compensation scheme for his individual characteristics aligned with that of a female" (Nopo 2008, 290). Based on this counterfactual, the raw gender wage gap is decomposed into an explained part, i.e., the individual endowments, and an unexplained residual part which is usually interpreted as labour market discrimination (Kunze 2018, 374; Maczulskij and Nyblom 2020, 2239).

Summing up the main argument then, as women historically have not had a large place in either education or the workforce, women have also had lower wages – if any – than their male counterparts because of their lower productivity. However, women are catching up to men in both education and work experience and are today more similar than men in terms of productivity determined by these observable characteristics (Kunze 2018, 392; Pettit and Hook 2005, 781). This has been observed across a variety of contexts, for example by Djurdjevic and Radyakin (2007) in Switzerland, García, Hernández and López-Nicolás (2001) in Spain, and McDonald and Thornton (2016) in Canada. If the observable characteristics explain the gender wage gap, then, the gender wage gap following these developments should either be eliminated entirely or even turned around in some cases. Yet, there remains a persistent wage gap, indicating that the explanatory power of the human capital theory has declined.

This indication that human capital theory explains less today than before further receives support by Blau and Kahn (2017) in their encompassing overview of the gender wage gap in developed countries. Using longitudinal data from the United States, for example, they find that changes in human capital such as education and work experience explained around one third of the gender wage gap in the United States in 1980, but only eight percent of the gender wage gap thirty years later in 2010 (Blau and Kahn 2017, 801). Although the wage gap declined drastically during the same time period, they identify a persisting gender wage gap. This further supports the account of a declining importance of traditional human capital characteristics for explaining the wage differences between men and women, and points towards the importance of some other explanatory factor for the persisting wage gap today.

2.1.2. Discrimination

Another traditional approach to understanding the gender wage gap has been the possibility of discrimination in the labour market. Gender discrimination in the labour market is defined as "a situation in which equally productive men and women are rewarded differently" either through wages, employment, promotions, and so on (Azmat and Petrongolo 2014, 33). This has been one of the main explanations of the gender wage gap and puts forward the idea that women, although equally productive as men, meet discrimination in the labour market that hinders their advancement and achievement. In contrast to theories focusing on the supply-side of the labour market, such as human capital, theories of discrimination provide a demand-side explanation for the gender wage gap by looking at how employers group employees and potential employees based on beliefs about productivity, the cost of potential training and expected rewards (Azmat and Petrongolo 2014, 32; Hipp 2020, 251).

There are two main theories of gender discrimination: taste-based discrimination and statistical discrimination. First, taste-based discrimination is defined as when "employers derive disutility from employing women relative to men" simply from preferring members of this demographic group over others (Azmat and Petrongolo 2014, 35). By suggesting that employers discriminate some groups based on their own taste, this explanation is argued to be based on more vague and irrational grounds for discrimination (Hipp 2020, 253).

Statistical discrimination, on the other hand, has received more attention and has gotten a larger hold in the literature. This is defined as when "employers use gender to extrapolate a signal of unobserved components of productivity" (Azmat and Petrongolo 2014, 35). This approach therefore assumes that employers are rational actors seeking to maximise their revenues by employing and/or promoting the most productive employees. Because the required pre-hire information for correctly assessing one's productivity is difficult or costly to obtain – expected family formation and fertility are private questions for example – employers extrapolate the needed information based on gender and systematically base their decisions on the assumed mean performance of a given demographic in order to reduce the risk of hiring low-performing workers (Booth 2007, 141; Hipp 2020, 252; Mandel and Semyonov 2006, 1915). In other words, employers make their decisions on the basis of the higher statistical probability of women quitting (Estévez-Abe 2005, 184). Thus, according to these arguments, because women are likely to have more frequent and lengthier workforce interruptions related to childbearing, employers rationally discriminate against women

because of the lower expected productivity (Hipp 2020, 253). Consequently, women face the risk of not getting hired, get pay increases or promotions that they otherwise could have gotten.

Furthermore, in the theory of gendered careers developed by Lazear and Rosen in 1990, the statistical discrimination is expected to increase when the job at question requires high-skilled workers and needs to invest in specific on-the-job training (Lazear and Rosen 1990, 107). This is because in jobs where specific training is needed, the employer wants their employees to stay with the employer long enough to receive the required returns from the investment in such training (Ehrenberg and Smith 2018, 188). Simply stated, the longer the employee stays with in the same employer, the better the investment. Therefore, when women are expected to provide lower returns to the investment, which is based on the assumption that women work shorter and more inflexible hours and thus have lower productivity, the higher the investment, the more likely the statistical discrimination (Ekberg, Eriksson and Friebel 2013, 134; Hipp 2020, 253).

Additionally, while the theory of taste-based discrimination assumes subjective – and arguably irrational – bases for discrimination, statistical discrimination takes into account the general context. In this way, the extent of statistical discrimination is heavily influenced by the contextually prevalent ideas and attitudes about roles, relationships, and practice (Padavic, Ely and Rider 2020, 103). If it is widely accepted that women are the primary caregivers, men the breadwinners, and the institutional structures and policies accommodate this idea by making it easier for women to stay at home, then statistical discrimination will be more prominent (Budig, Misra and Boeckmann 2012, 164; Datta Gupta, Smith and Verner 2008, 78; Kamerman and Moss 2009, 266).

This also indicates that cross-country differences can determine the extensiveness of statistical discrimination. A prominent example of theories expecting context to influence the extent of discrimination is provided by Estévez-Abe (2005, 2006). By building on Hall and Soskice's (2001) *Varieties of Capitalism* and Becker's (1993) definition of general and specific training and skills, Estévez-Abe (2005, 2006) provides a skill-based theory on gender differences in labour market outcomes. The general claim is that firm-specific skills discriminate against women, whereas general skills are more gender-neutral. This is because specific skills are less transferable, consequently making workers and employers more hesitant to invest in specific skills in cases of job uncertainty, which is more likely for women

(Estévez-Abe 2006, 153). Therefore, institutions that promote specific skills exacerbate gender gaps in skills and lead to further differences in labour market outcomes for men and women (Estévez-Abe 2006, 181-182). Thus, the systems that are argued to be more discriminatory are coordinated market economies (CMEs) – in contrast to liberal market economies (LMEs) – because of the presence of institutions that lock key economic actors into long-term relationships, such as strict employment legislation, for example for the use of temporary contracts, and cooperation with strong unions, making it less costly for employers to invest in specific on-the-job training (Estévez-Abe 2005, 189-191; Hall and Soskice 2001, 8). Therefore, Estévez-Abe expects that "employers who invest in their workers' skill formation will be more discriminatory against women", especially in CMEs (2006, 153).

Moreover, statistical discrimination is argued to affect all women, independently of whether a particular woman has children or not or plans to have a family (Estévez-Abe 2005, 190). On one hand, research has found that part of the gender wage gap can be explained by inequalities among women due to what is known as "the motherhood penalty" defined as "the amount each additional child lowers women's earnings" (Budig, Misra and Boeckmann 2016, 120). Blau and Kahn (2017), for example, show that there is considerable evidence indicating a negative relationship between number of children and women's wages (823). Budig, Misra and Boeckmann (2012) similarly find that mothers earn less than women without children in most European countries (186). On the other hand, research has also found that there are no differences in the gender wage gap or the mechanisms that are theorised to shape it between women with and without children. For example, Datta Gupta, Smith and Verner (2008) when looking at the Nordic countries, find that there is a tendency for all women to have a have a flatter wage profile during their childbearing ages, indicating that "there might be negative effects on *all* women's wages, not just on mothers' wages" (78). This can be explained by the heterogeneity of mothers' preferences and employers' beliefs about how these preferences are distributed. If employers cannot screen the types of women that chose to have children from those choosing not to, and those choosing to have children that do not want to take prolonged time off work from those that do want this time off, "future mothers with a strong preference for parental leave impose an externality on more career-oriented women" (Ekberg, Eriksson and Friebel 2013, 143).

Nevertheless, although employer discrimination is argued to be both a persistent disadvantage that women face in the labour market and a consistent mechanism driving the gender wage gap, a fundamental challenge for this line of research has been to accurately identify incidents

of discrimination (Hipp 2020, 250). This is because identifying the extent and nature of discrimination relies on non-statistical evidence in the sense that it is not observable as is often a subconscious process that people rarely admit to (Booth 2007, 135; García, Hernández and López-Nicolás 2001, 152).

In order to study the extent of discrimination in determining the gender wage gap, early work have extensively used decomposition techniques on observational data, as already discussed, to compare male and female earnings based on their observable characteristics and have usually attributed the unexplained differences to labour market discrimination (Azmat and Petrongolo 2014, 33; García, Hernández and López-Nicolás 2001, 150). However, this method has been criticised for doing this because "accrediting all unexplained variation to discrimination assumes that all possible confounding factors are accounted for in the analysis, leaving discrimination as the only possible answer" (Azmat and Petrongolo 2014, 33). This is a hard assumption not to break for two reasons. First, because early studies using decomposition methods are related to the human capital theory, this research has been criticised for not taking into account the context in which the individuals exist and thus leaving out important confounding factors. Second, some of the important determinants of individual observable characteristics, such as work experience or willingness to work overtime, are inherently hard to measure and are therefore not sufficiently accounted for in these analyses. This can potentially create an upward bias of the true extent of discrimination (Azmat and Petrongolo 2014, 33).

Because of an increased awareness of these limitations, recent research has turned towards experimental methods in order to detect and research gender discrimination in the labour market. Experimental methods, by randomly assigning subjects different treatments, are able to isolate and study causal mechanisms, which seems promising for the studies of discrimination (Gerring 2012, 258; Kellstedt and Whitten 2018, 85). In a methodological review of experimental research on gender and the labour market, Azmat and Petrongolo (2014) discuss the contribution of the experimental literature to understanding mechanisms leading to gender differences in labour market outcomes, such as discrimination and women's preferences. For discrimination, experimental studies have for example found that women appear to gain less from negotiation and that they are less likely to be invited to interviews, especially in typically male-dominated occupations (Azmat and Petrongolo 2014, 34; Hipp 2020, 251). Azmat and Petrongolo (2014) thus conclude that "the experimental approach

offers a way to answer questions previously believed to be unanswerable because of data limitations" (39).

Nevertheless, although measuring and unravelling potential gender discrimination in the labour market is outside the scope of this thesis, the presence of discrimination remains one of the main mechanisms in explaining the gender wage gap from the demand-side of the labour market.

2.1.3. Family policy

Finally, a third traditional approach to explaining the gender wage gap has been looking at the degree of which family policy either facilitates or constrains women's choices and opportunities in the labour market. This approach thus addresses one of the human capital theory's main critiques, namely that it does not take into account the context (Maczulskij and Nyblom 2020, 2239). The labour market consists of actors demanding and supplying labour, but these actors' decisions are constricted by the context in which they operate, for example by responsibilities outside work, cultural expectations, institutional structure, or policy configurations (Ehrenberg and Smith 2018, 60). Therefore, because women's labour market outcomes are highly associated with potential motherhood, as already discussed, countries' family policies attempting to regulate the work-family balance are expected to powerfully influence gender wage gaps (Kolberg and Esping-Andersen 1991, 4).

First, we need to define what is meant by *family policy*. This is usually regarded as an umbrella term covering a broad range of legislation and policies. Thus, the scope of family policy is not clear and there is consequently no consensus about a definition in the literature. In its broadest definition, the concept of *family policy* includes all policies associated with families and intergenerational dependencies, covering everything from parental leave to care for the elderly (Daly 2020, 26). A more common approach, which I will adhere to, is narrower and defines family policy as policies "centred around the well-being, functioning, and responsibilities of families with children" (Daly 2020, 28). This definition therefore only looks at intergenerational dependencies for children and includes policies such as childcare arrangements and parental leave schemes. Although having this commonality, the policies falling under this definition are diverse and compromise different aims, ranging from supporting fertility and educating children to relieving barriers to female and employment and reducing economic inequality between genders (Fleckenstein 2010, 796; Thévenon 2008, 165;

Zagel and Lohmann 2020, 122). Motives, of course, overlap, but in addressing the gender wage gap, there is a strong belief today that family policy should not only support families but "foster gender equality and promote the chances of women" (Ekberg, Eriksson and Friebel 2013, 132).

Having defined family policy, early research built on the assumption that women's involvement in the labour market is largely a response to the cost and benefits of paid labour relative to domestic work and argued that regimes with the most generous family policies relieved the costs of entering and staying in the labour force for women, while increasing the cost of leaving (Kleider 2015, 506). One central contribution to the literature was made by Gornick, Meyers and Ross (1998) demonstrating a strong positive association between women's labour force participation rates and family policy generosity. They argued that by the state taking some of the care responsibilities traditionally entrusted to women, family policies would reduce the work-family conflict and, in turn, "decrease the probability that childcaring responsibilities would force mothers to withdraw from paid work" (Gornick, Meyers and Ross 1998, 37).

In other words, in contexts with generous family policies then, potentially having children would no longer entail a choice between either work or family for women because they had care arrangements available and institutionalised time off, thus keeping the connection to the paid work. This, in turn, would strengthen women's labour force attachment and consequently increase their labour force participation, which is further expected to decrease the gender wage gap by first enabling women to have a wage at all, and – as discussed earlier – by decreasing the labour market discrimination by providing employers with a guarantee that women would return after having children and at the same time increasing their human capital (Estévez-Abe 2006, 146; Gornick, Meyers and Ross 1998, 37; Hook and Li 2020, 2650). The greatest penalties for women in terms of relative wages is therefore in countries with the least developed family policies supporting female employment (Gornick, Meyers and Ross 1998, 38).

Furthermore, because family policies did not emerge in a political vacuum but in the context of broader welfare configurations, early research approached this from a welfare regime perspective building on Esping-Andersen's (1990) influential typology of the three worlds of capitalism (Nieuwenhuis and Van Lancker 2020, 9; Thévenon 2008, 165). Esping-Andersen originally formulated the typology based on the concept of *decommodification*, i.e., the degree

to which individuals are able to uphold an acceptable standard of living independent of employment. However, he introduced another concept of *defamilisation* relating to policies relieving parents from their care responsibilities in response to feminist critiques of the typology being "gender-blind" not taking into account the additional risk of motherhood for women's employment (Kolberg and Andersen 1991, 25; Nieuwenhuis and Van Lancker 2020, 10). Nevertheless, in relation to both concepts, the regimes differ in the belief of who is the legitimate provider of welfare and cultural expectations regarding maternal employment, thus emphasising the role of the state in social-democratic regimes, the market in liberal regimes and the family in conservative regimes (Budig, Misra and Boeckmann 2012, 167; Korpi, Ferrarini and Englund 2013, 13).

Keeping in line with Esping-Andersen's three worlds, then: First, social-democratic regimes, typically the Nordic states, are characterised by the state to a large extent providing services so that parents can combine work and family, usually in the form of universal access to publicly subsidised childcare and highly developed parental leave schemes (Boye, Halldén and Magnusson 2017, 600; Pettit and Hook 2005, 782). These countries are thus expected to have high female labour force participation and, in turn, smaller gender wage gaps. Second, liberal regimes, typically the English-speaking countries, take a "hands-off" approach by leaving the market to be the main provider of services. These countries therefore have the least developed family policies and are expected to have the lowest female labour force participation and thus higher gender wage gaps (Mandel and Shalev 2009, 1876; Pettit and Hook 2005, 782). Finally, conservative states emphasise the role of the family as the provider of services and occupy a middle ground in terms of family policies, female employment rates and gender wage gaps (Mandel and Shalev 2009, 1876; Kamerman and Moss 2009, 7).

In sum, family policies are argued to be capable of creating "women-friendly" states reducing gender inequalities in the labour market (Schäfer, Tucci and Gottschall 2012, 5). However, although researchers have found that family policies have increased women's labour force participation, there is limited evidence that this is the case for gender wage gaps (Ekberg, Eriksson and Friebel 2013, 131). This is particularly evident in the Nordic states which were considered forerunners in designing family policies and achieving high female labour force participation rates, but despite this have a persistent gender wage gap that is larger than expected from this theory (Datta Gupta, Smith and Verner 2008, 66).

Trying to account for this puzzling outcome in the Nordic countries, another strand of research suggests that there is a trade-off between family friendly policies and women's relative wages resulting in a welfare state paradox due to unintended consequences of these policies. These theories suggest that while family policies increase women's labour force attachment, this does not automatically translate into smaller gender wage gaps. Therefore, these policies may have contributed to undermining women's earnings by institutionalising female workforce interruptions in a way that has negative consequences for women's relative wages (Budig, Misra and Boeckmann 2016, 12; Grönlund and Magnusson 2016, 92).

A central contribution to this theory has been that of Mandel and Semyonov (2005, 2006). In their work, they identify this welfare state paradox and argue that these policies, while facilitating women's access into the labour force, do not facilitate women's access into powerful and desirable position and thus end up "reproducing the gendered division of labour" (Mandel and Semyonov 2006, 1913). They argue that the welfare state, in its role as both legislator of family policies and employer of women in the jobs created by the family policies, most notably care work in the public sector, create "sheltered labour markets for women" (Mandel and Semyonov 2006, 1911).

This is because adjusting the demands of employment to women's home duties and allowing women to work reduced hours and take long leaves leads to occupationally segregated labour markets. Occupational segregation, defined as the different representation of men and women in different kinds of work, is therefore argued to be a main driver of the persisting gender wage gap. It is not that women do not work or that they are not paid equally in the same jobs, women simply work in different jobs, mostly in the lower paid public sector (Boye, Halldén and Magnusson 2017, 600; Djurdjevic and Radyakin 2007, 373; Maczulskij and Nyblom 2020, 2242). These jobs, typically in health, education, and social services, are argued to be more attractive to women – and especially mothers – because of their protective nature, flexible working hours and greater tolerance for absenteeism, but are also characterised by restricted career opportunities and lower wages (Mandel and Semyonov 2005, 952). Another mechanism through which this occupational sorting happens is argued to be statistical discrimination. By institutionalising women's workforce interruptions, instead of the family policies providing employers with a guarantee that women return to work after potential childbearing – as argued above – these policies instead provide employers with a guarantee that there will be a potentially costly workforce interruption when employing women (Datta Gupta, Smith and Verner 2008, 78; Nielsen, Simonsen and Verner 2005, 7229). Furthermore,

if women are expecting to meet discrimination in some occupations, it is less likely that women would be motivated to compete with men and therefore self-select into more safe public sector positions, further increasing the aggregate wage gap (Datta Gupta, Smith and Verner 2008, 79; Mandel and Semyonov 2005, 952).

In sum, by concentrating women in these kinds of jobs, generous family policies do not challenge the traditional distribution of market-family responsibilities between men and women and thus limit women's occupational opportunities and earnings capacity (Mandel and Semyonov 2005, 950).

Relating this to the welfare regimes again, this approach provides an explanation to the puzzling pattern observed in the Nordic countries. In these countries, women's integration into the labour market is accompanied by their crowding in female-dominated occupations, thus creating a larger wage gap (Grönlund and Magnusson 2016, 94; Mandel and Semyonov 2006, 1940). This has also been shown empirically by the Nordic countries where more than half of the female workforce work in the public sector compared to the 20 percent of the male workforce working in the public sector, making the labour markets in the Nordic countries "among the most gender segmented labour markets in the world" (Datta Gupta, Smith and Verner 2008, 79). Liberal states, on the other hand, are less effective in mobilising women into employment but do not impede their chances of reaching highly paid positions outside the public sector and lastly, the conservative model typically disadvantages women in both respects (Mandel and Semyonov 2006, 1940).

Therefore, summarising the early theories on the effects of family policy on the gender wage gap, Arulampalam, Booth and Bryan (2007) note that:

"Family-friendly policies can be a double-edged sword. On the one hand, they might raise women's relative earnings by preserving their ties with the firm, thereby increasing incentives to invest in specific human capital and leading to higher female pay. But on the other hand, family-friendly policies could increase women's time out of work force for childbearing, thus widening the average gap" (Arulampalam, Booth and Bryan 2007, 177)

2.2. Breaking down family policy and the wage distribution

Having elaborated on the traditional approaches introducing important mechanisms affecting the gender wage gap, I will now turn towards the research question and argue that breaking down both family policy and the wage distribution allows for a better understanding of how family policies affect the gender wage gap and how the effects differ across the wage distribution. I formulate the hypotheses to be tested throughout this section. I start by elaborating on the large conceptual debate regarding the concept of *family policy* and argue in favour of disaggregating it into the dimensions of *familisation* and *defamilisation*. I then look closer at how the effects of the family policy dimensions are expected to affect gender wage gaps across the wage distribution building on literature identifying a glass ceiling phenomenon and argue that what affects the gender wage gap for bottom earners is not necessarily the same as what affects the gender wage gap for top earners.

2.2.1. Separating familisation from defamilisation

Summarising the argument above, it has been argued that overall family policy generosity will decrease the gender wage gap by enabling women to balance work and family and therefore have a stronger labour force attachment to paid work (Gornick, Meyers and Ross 1998). On the other hand, there is a strong theoretical argument suggesting the opposite, i.e., that family policy generosity institutionalises a gendered division of labour and thus increases the gender wage gap (Mandel and Semyonov 2005, 2006). Thus, the theoretical expectations are ambiguous, and the resulting empirical findings have been equally inconclusive. I nevertheless expect that:

H1: More family policy generosity should affect the gender wage gap.

However, there is an important limitation to this approach. This line of research has often subsumed a multitude of policies – such as parental leave, childcare arrangements, and public sector employment – into an overall index to assess their impact on female labour market outcomes. While these policies are often viewed as a package deal, there is – as already mentioned – a growing consensus in the literature that these policies should be differentiated by their objectives as some of these may have markedly different implications for women in the labour market (Budig, Misra and Boeckmann 2016, 125; Pettit and Hook 2005, 797). Although family policies have in common that they centre around the well-being, functioning

and responsibilities of families with children (Daly 2020, 8), the policies falling under this definition are diverse and cover different aims (Fleckenstein 2010, 796; Thévenon 2008, 165; Zagel and Lohmann 2020, 122). Therefore, Pettit and Hook (2005) for example argue that we cannot expect extended paid parental leave to have the same effect on the gender wage gap as high-quality subsidised childcare. This, they argue, is because while the former institutionalises prolonged absence from work and decreases the costs of choosing to stay outside paid work, the latter provides possibilities for childcare outside the family, making it easier for women to choose paid work (797).

This relates to a well-known conceptual debate in social sciences about maximalist versus minimalist concepts, questioning what to include and what not to include in a theoretical as well as an empirically measured concept (Gerring 2012, 157). While minimalist concepts are criticised for missing important indicators, maximalist concepts are criticised for including too many, thereby making the concept less precise and in the worst case stretching it beyond recognition (Gerring 2012, 135). This is especially the case when indicators included in a concept is argued to have different effects. Therefore, although studies such as by Gornick, Meyers and Ross (1998) and Mandel and Semyonov (2005, 2006), among others, have shown that there is a relationship between family policy generosity and women's labour market outcomes, using a generalised index that combines policies that may have contradictory effects on women's labour market outcomes – such as either promoting or deterring paid employment or being assumed to either increase or decrease statistical discrimination – will obscure the policies' actual effects and lead to unreliable results (Budig, Misra and Boeckmann 2016, 126; Korpi, Ferrarini and Englund 2013, 5; Pettit and Hook 2005, 783). This may explain some of the theoretical ambiguity and contradictory empirical findings in the early literature.

Therefore, it has been argued that "rather than […] analysing a single index of family policy generosity", the policies should be disaggregated and explored separately according to their aims and expected effects (Budig, Misra and Boeckmann 2016, 143). Along these lines, a prominent approach reconceptualising *family policy generosity* is presented by Lohmann and Zagel (2016). In their article, they discuss the shortcomings in the conceptualisation and measurement of *family policy* and the lack of consensus on the effects of these policies. They argue that instead of looking at overall family policy generosity, the concepts should be defined as the degree of support of individuals' independence from or dependence on family relationships (Lohmann and Zagel 2016, 49). Further developing on concepts already

suggested in the literature, they propose differentiating the dimensions of *defamilisation* and *familisation* from the overall concept of family policy, defined instead in terms of their aims and expected outcomes. First, defamilisation is defined as policies that "reduce care and financial responsibilities and dependencies between family members" (Lohmann and Zagel 2016, 52). Second, familisation is defined as policies that instead "foster dependencies among family members by actively lowering their negative social and economic outcome" (Lohmann and Zagel 2016, 53).

By separating the effects of the different policies, they argue that these concepts are helpful analytical tools for comparing family policy across countries and is especially suitable when "analysing a policy field with opposing policy aims" (Lohmann and Zagel 2016, 49). The main advantage is the concepts being gradual instead of categorical, thereby considering the level of which variables are observed in each country instead of considering countries as static cases, such as in welfare regime approaches (Zagel and Lohmann 2020, 124). This can therefore give a more precise description of country differences in family policies. In addition, gradual concepts are able to take into account hybrid cases that fail to fit specific regime types, such as the female-friendly policies in France, which is typically considered a conservative welfare regime, or the slightly more traditional family policies in Ireland, typically considered a liberal welfare regime (Kleider 2015, 511; Korpi, Ferrarini and Englund 2013, 8).

First, defamilising policies aim at reducing the dependencies between family members, which in this case entails reducing children' potential dependencies on their mothers, by providing a care alternative to women having, or planning to have, children (Lohmann and Zagel 2016, 52). Defamilising policies are therefore argued to be "instrumental in supporting women's labour force participation" by the state alleviating some of the care responsibilities, thus enabling women to have continuity in paid employment, minimising lost human capital and potentially missed skill acquisition opportunities (Budig, Misra and Boeckmann 2016, 126). The policies also challenge the traditional expectation of women's and men's roles regarding paid work and care responsibilities by normalising the working mother. They are therefore argued to further decrease the statistical discrimination often facing women when enhanced labour market attachment is incorporated into employers' beliefs (Ekberg, Eriksson and Friebel 2013, 134; Olivetti and Petrongolo 2017, 212). This, in turn, increases women's work opportunities and relative earnings. I therefore expect that:

H2: More policies enabling defamilisation should decrease the gender wage gap.

The policies argued to enable defamilisation are subsidised childcare arrangements, paternity leave and moderate length paid parental leave (Kleider 2015, 510; Korpi, Ferrarini and Englund 2013, 5; Podestà 2017, 340)

First, publicly available childcare is argued to be "a prime example of a policy with a clear defamilising effect" (Lohmann and Zagel 2016, 53). This is because when childcare arrangements are generally available and subsidised, the state assumes some of the parents' caregiver responsibilities and provides an accessible substitute to dependence on parental care (Olivetti and Petrongolo 2017, 221). This, in turn, enables women to pursue paid employment when the children are very young and reduces the workforce interruptions related to childbirth (Kleider 2016, 510; Olivetti and Petrongolo 2017, 221; Pettit and Hook 2005, 796). While childcare programs have been adopted with both the goal of educating children and of supporting parents' employment, programs for children under three years old are explicitly recognised as helping families balance care and employment (Budig, Misra and Boeckmann 2015, 126; Vandenbroeck 2020, 170).

A second policy argued to reduce children's potential dependencies on their mothers is paid paternity leave, enabling fathers to take time off around the time of childbirth instead of mothers (Kamerman and Moss 2009, 2). Although leave schemes are typically sharable between mothers and fathers by their own choice, leaves have largely been taken by women (Brandt and Kvande 2009, 191; Daly 2020, 36; Kotsadam and Finseraas 2011, 1611). The motivation behind introducing leaves reserved for the exclusive use of fathers was therefore to give incentives to fathers to partake in care responsibilities traditionally entrusted to women, thus encouraging greater gender equality both within the household and in the labour market (Cools, Fiva and Kirkebøen 2015, 801; Kamerman and Moss 2009, 2; Rege and Solli 2013, 2256). As this is a relatively new policy for most countries, it has not been included in studies of family policies and the effects on gendered labour market outcomes. Nevertheless, by redistributing the care responsibilities, paternity leave is argued to be defamilising in enabling women to have a stronger workforce attachment (Kleider 2015, 510).

Third, paid parental leave can be defamilising as it "reduces family dependencies by allowing employed parents of young children to maintain high labour market attachment, guaranteeing the return to their jobs after a given period" (Lohmann and Zagel 2016, 53). Women are therefore able to take care of children for a period of time while being able to return to the

same job or an equivalent one (Budig, Misra and Boeckmann 2016, 127). Thus, while institutionalising women's workforce interruptions, having paid parental leave has a defamilising effect at large by reducing the chances of women having children leaving paid work entirely. Yet, the case of parental leave is somewhat particular in terms of the expected effects. This is because of the different compositions of parental leave schemes. On the other hand, parental leave is also argued to be familising because these policies effectively institutionalise women's workforce interruptions. This is especially the case if leaves are prolonged – usually over several years – or unpaid, further incentivising parents to stay at home to take care of their child(ren) (Kleider 2015, 510; Lohmann and Zagel 2015, 53). Researchers have therefore found that paid parental leave has a curvilinear U-shaped effect on women's employment and earnings, thereby being defamilising when moderate and well paid – enabling women to have a strong labour force attachment, but familising when there either is no parental leave provisions – thus forcing women to leave the labour market when having children – or when the leave is either prolonged or unpaid, further incentivising women to stay outside the workforce by reducing the costs of choosing so and impacting employers' perceptions of women's employability (Datta Gupta, Smith and Verner 2008, 66; Olivetti and Petrongolo 2017, 214; Ruhm 1998, 293).

Second, familising policies foster dependencies among family members, which in this case focusing only on intergenerational dependencies for children either means children's dependence on their parent's care or women's financial dependence on a breadwinner (Lohmann and Zagel 2016, 53). Familising policies thus have work-reducing properties institutionalising women's workforce interruptions, thereby also reducing the cost of choosing to stay outside paid work and – contrary to defamilising policies – support the cultural expectation of women as mothers first, workers second (Korpi, Ferrarini and Englund 2013, 5; Padavic, Ely and Rider 2020, 103; Pettit and Hook 2005, 335). This, in turn, can lead to increasing statistical discrimination, lost human capital acquisition and further cluster women into low-paying jobs that better tolerate absenteeism (Mandel and Semyonov 2005, 951). I therefore expect that:

H3: More policies enabling familisation should increase the gender wage gap.

Familising policies are therefore those supporting family care at home, which – as argued above – can either be unpaid parental leaves or prolonged leaves (Podestà 2017, 340). Furthermore, in the construction of a familising concept, Kleider (2015) also includes certain

tax rules as a familising policy, arguing that such policies can increase the cost for women to work and to maintain a workforce attachment. This is because the tax systems in question, i.e., that of joint taxation of spouses, essentially add up the income of both parents and divide it by two to determine a common tax rate, which quickly adds up to increased household taxes when both parents work full-time (Kleider 2015, 510). Such taxation systems are therefore most favourable for households where the main earner receives a high income and the second earner – who is typically a woman – has no or a small income (Budig, Misra and Boeckmann 2016, 124; Kleider 2015, 510). However, as these policies are not typically included as the main family policies and are complexly related to poverty-relieving policies, they cannot necessarily be categorised simply as a family policy trying to aid the work-family balance (Budig, Misra and Boeckmann 2016, 128; Daly 2020, 36; Ekberg, Eriksson and Friebel 2013, 133).

Finally, as the duration of paid parental leave is argued to be both defamilising and familising, I also expect that:

H4: The duration of paid parental leave should have a curvilinear U-shaped effect on the gender wage gap

In sum, the concept of family policy can be disaggregated as presented in *Figure 2.1.*, showing the overreaching concept of family policy and its division into the dimensions of familisation and defamilisation based on their main ideas and expected effects. The individual indicators are presented under each dimension.

FIGURE 2.1. *The concept of family policies disaggregated.*
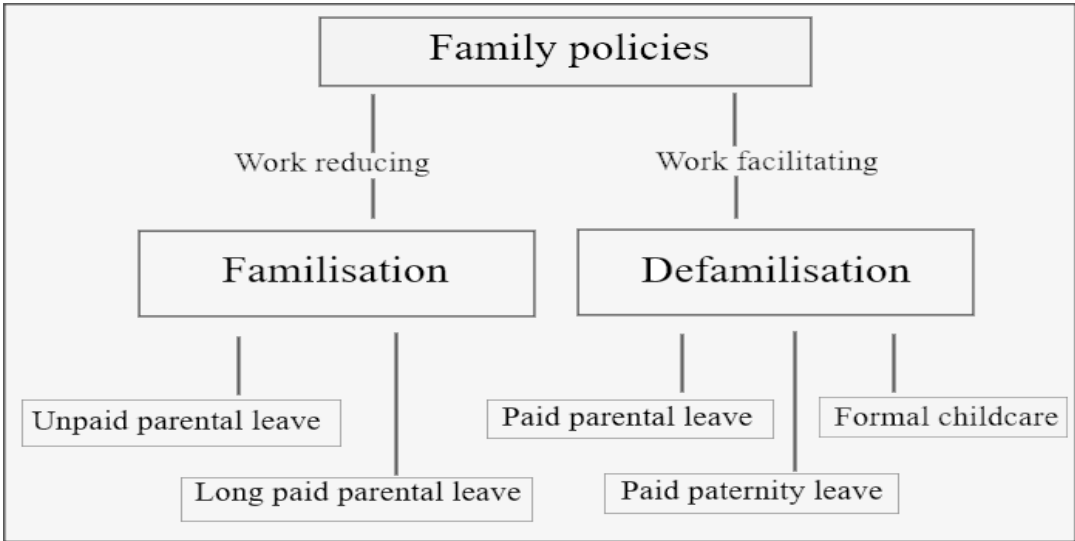
2.2.3. Differences across the wage distribution

Finally, previous research on the gender wage gap has often focused on the average gender wage gap between men and women, thus assuming that the wage gap is constant across the wage distribution and that, in turn, the factors shaping the gender wage gap will be the same for all irrespective of their wage levels. However, resent research has questioned to what extent these assumptions are reasonable. A main critique of such approaches has been that the focus on mean wage differences neglects "important variations of the wage difference across the entire wage distribution" (Bonjour and Gerfin 2001, 408). Therefore, just as looking at family policy generosity is misleading when the policies are argued to have different effects, looking at the gender wage gap at median can be equally misleading if the wage gaps and what drives them are not uniform across the wage distribution. In response to this shortcoming, there is a growing body of literature looking into the gender wage gap not only at median and addressing the question of whether the mean gender wage gap hides larger or smaller differences between high paid men or women, or between low paid men or women (Albrecht, Björklund and Vroman 2003, 146; Arulampalam, Booth and Bryan 2007, 167). The essential question is therefore whether women meet larger inequalities at certain points of the wage distribution (Albrecht, Björklund and Vroman 2003, 146).

First, research has addressed the question of whether there are differences across the wage distribution and has identified that this is indeed the case by showing differences in the magnitude of the gender wage gap for top and bottom earners. Two central contributions to this line of argument have been that of Albrecht, Björklund and Vroman (2003) and that of Arulampalam, Booth and Bryan (2007).

First, Albrecht, Björklund and Vroman (2003) identified differences across the wage distribution in their seminal study. By using a quantile regression approach with data from Sweden, they both found and showed empirically that the gender wage gap in Sweden increases throughout the wage distribution and accelerates rapidly at the upper tail (Albrecht, Björklund and Vroman 2003, 147). The finding that the gender wage gap was greater at the top of the wage distribution held even after adjusting for a set of basic individual and aggregate controls, which they interpret as a glass ceiling effect, defined as "the phenomenon whereby women do quite well in the labour market up to a point after which there is an effective limit on their prospects" (Albrecht, Björklund and Vroman 2003, 146). This implies

that women fall behind men more at the top of the wage distribution than at the middle or bottom.

Second, building on the work by Albrecht, Björklund and Vroman (2003), Arulampalam, Booth and Bryan (2007) extended the sample and investigated whether the gender wage gap differed both between countries and across the wage distribution in 11 European countries using a similar quantile regression approach. They found that the magnitude of the gender wage gaps varied substantially both across countries and across the wage distribution. They also observed a widened gender wage gap at the top of the wage distribution in most countries, which they also interpreted as a glass ceiling phenomenon (Arulampalam, Booth and Bryan 2007, 164). In addition, they observed that in two cases – in Spain and Italy respectively – the gender wage gap not only widened at the top of the distribution, but at the bottom of the distribution as well. They interpreted and described this effect as a *sticky floor* phenomenon in which there are barriers preventing women from moving up the ladder, i.e., reaching top positions, when included in the workforce (Arulampalam, Booth and Bryan 2007, 164).

This glass ceiling effect has since been a consistent finding in the literature studying the gender wage gap in similarly industrialised countries. Blau and Kahn (2017) for example find that the gender wage gap is larger at the top of the wage distribution in their longitudinal overview of the gender wage gap in the United States (807); Schäfer and Gottschall (2015) similarly find a disproportionately large gender wage gap at the top of the wage distribution in the Nordic countries (479); and Christofides, Polycarpou and Vrachimis (2013) also find evidence of substantial glass ceilings, in which the gender wage gap is larger at the top of the wage distribution, in most of the 26 European countries in their sample. On the other hand, there has been less evidence of a sticky bottom-effect in these countries. One potential explanation for this could be that the sticky-bottom effect is more likely to be found in developing countries in which the female labour force is still increasing, while the glass ceiling phenomenon is mostly found in developed countries (Duraisamy and Duraisamy 2016, 4099; Hara 2018, 217). For example, the sticky floor effect has been observed by Duraisamy and Duraisamy (2016) in India, by Hara (2018) in Japan, by Xiu and Gunderson (2012) in China and by Carillo, Gandelman and Robano (2014) in Latin America. This, however, does not explain its presence in Spain and Italy.

Therefore, these findings taken together showing that the magnitude of the gender wage gap does change across the wage distribution further motivates and justifies a focus on class-gender differences, here defined by wage levels (Grönlund and Magnusson 2017, 96). This further provides support for the notion that measuring the gender pay gap at the mean of each distribution, i.e., comparing the 'average' man with the 'average' woman, can produce "a misleading simple picture of how men's and women's wages differ" (Arulampalam, Booth and Bryan 2007, 167). This, in turn, further gives reason to expect that there either are different factors shaping the gender wage gaps at bottom, median and top levels of the wage distribution, or that the same factors have different effects for different points at the wage distribution (Korpi, Ferrarini and Englund 2013, 6).

However, early research – such as the research already already presented – only briefly hypothesise how family policies are expected to have affected this difference in magnitude of wage gaps across the wage distribution. Arulampalam, Booth and Bryan (2007), for example, finish their article by checking correlations between a family policy generosity index and the gender wage gap at either end of the wage distribution. They find that "countries with more generous work-family policies have a lower wage gap at the bottom of the wage distribution and a wider gender wage gap at the top", which they argue could be because women at the bottom of the wage distribution can be assumed to have a lower workforce attachment to begin with and therefore would be more incentivised to join and stay in the workforce by such policies (Arulampalam, Booth and Bryan 2007, 177). The positive correlation, on the other hand, is explained by women at the top of the wage distribution being incentivised to take time off and therefore lose out on wages they would otherwise have gotten (Arulampalam, Booth and Bryan 2007, 177).

Similarly, Albrecht, Björklund and Vroman (2003) argue that differences in wage gaps are partially due to different family policy constellations after observing that the gender wage gap is larger at the top of the wage distribution in Sweden than in the United States (172). One potential explanation, they argue, is that family policy generosity discourages career commitment which means that "women may have strong incentives to participate in the labour force but not do so very intensively" (Albrecht, Björklund and Vroman 2003, 172). This, in turn, could increase statistical discrimination as employers expect less commitment, and therefore also less productivity, from women having these alternatives or women self-selecting away from high-demanding, high-paid jobs to begin with.

Furthermore, Mandel and Shalev (2009) find that family policy generosity does not affect the overall gender wage gap, but instead increases the wage gap in higher socio-economic groups and decrease it in lower socio-economic groups, which they define by educational attainment (Mandel and Shalev 2009, 1877). This, they argue, is because family policy generosity, on one hand allows women in lower socio-economic groups a better labour force attachment, but on the other hand, by institutionalising workforce interruptions and the traditional division of labour, decreases women's chances of joining desirable positions and therefore further increases the wage gap among top earners (Mandel and Shalev 2009, 1877).

Yet, none of the above-mentioned studies differentiated between the dimensions of familisation and defamilisation, thus making the inferences somewhat unreliable due to the fact that family policy generosity could be argued to have contradictory effects, as discussed in the previous section. Therefore, in contrast to the relative consensus on how policies enabling defamilisation and familisation either facilitate or reduce women's labour force attachment, the theoretical expectation of how these dimensions affect women differently depending on their place in the wage distribution has not been properly accounted for in the literature and the mechanisms are argued to be somewhat conflicting (Grönlund and Magnusson 2016, 93; Hook and Li 2020, 257). The comparative research on this is therefore extremely limited. It is nevertheless possible to make general inferences based on the nature of these policies and the mechanisms discussed throughout this chapter.

First, where research finds common ground, as seen above, is in arguing that the mechanism of discrimination is expected to be more prevalent for high-earning women in qualified-positions. This is argued to contribute to occupational segregation either horizontally – by providing women with incentives to self-select into low-paying public jobs that tolerate absenteeism – or vertically – by hindering women's labour force achievement and advancements in the labour market (Estévez-Abe 2005, 183; Grönlund and Magnusson 2016, 96; Nielsen, Simonsen and Verner 2004, 743). Of course, this mechanism also applies for women at the other end of the wage distribution, but as already discussed, discrimination is expected to increase in occupations valuing workforce continuity either because they require high-skilled workers that are hard to replace, need a continuous investment in on-the-job-training, require more interactions with each other or clients, or have more time pressure, all of which typically characterise occupations with higher wage levels (Boye, Halldén and Magnusson 2017, 596-598; Cha and Weeden 2014, 460; Goldin 2014, 1108; Lazear and Rosen 1990, 107). Simply stated then, discrimination is more likely to "weigh heavily in

present-day high-skilled, high-intensity workplaces, where pay and promotions are tied to [expected] performance and productivity rather than to qualifications and tenure" (Datta Gupta, Oaxaca and Smith 2006, 262). This is further supported empirically by several studies finding that women's experience of discrimination does increase as they rise in the workplace hierarchy or are employed in high-status high-paid occupations (Azmat and Petrongolo 2014, 35; Boye, Halldén and Magnusson 2017, 598; Hipp 2020, 253; Newman 2016, 1010).

In a related argument, research by Goldin (2014) and Cha and Weeden (2014) shows that it is not only workforce interruptions but the assumed inflexibility for women that further extend the gender wage gap at the top of the wage distribution. This is argued to be the case because of how firms disproportionately reward individuals who work long, flexible, and particular hours, which is typically more concentrated among highly educated, professional, and managerial workers (Cha and Weeden 2014, 459-460; Goldin 2014, 1110). Therefore, because "certain occupations impose heavy penalties on employees who want fewer hours and more flexible employment", the lower renumeration can result in shifts to an entirely different occupation, to a lower position within the same occupation to being out of the labour force altogether (Goldin 2014, 1106). In other words, because a lower proportion of women than men overwork, "women are less likely to enter jobs that require extremely long [and flexible] hours" (Cha and Weeden 2014, 459).

Furthermore, this means that when policies institutionalise long workforce interruptions and foster dependencies between family members, i.e., when policies enable familisation, these policies signal extra risk to employers and further emphasises the traditional roles of women as caregivers. Statistical discrimination is thus expected to increase and be particularly detrimental for career-minded women's workforce achievements (Mandel and Semyonov 2006, 1914; Padavic, Ely and Reid 2020, 64). Therefore, while familising policies in the same vein are argued to extend the gender wage gap for women at the lower end of the wage distribution, it could be argued that the effect is more prominent for top earners. I therefore expect that:

H5: The positive effect of policies enabling familisation on the gender wage gap should be stronger for the top earners

On the other hand, defamilising policies enabling women to quickly return to work and have a stronger workforce attachment can therefore be expected to decrease the statistical discrimination and in turn decrease the gender wage gap for top earners. It could, however, be

argued that policies enabling defamilisation will have a stronger negative effect on the gender wage gap at the bottom end of the wage distribution. For women earning less, the main mechanism through which defamilising policies are expected to decrease the wage differences is by facilitating and reinforcing their workforce attachment. As it could be argued that they would be less attached to the workforce to begin with, defamilising policies would therefore bring women into the labour market on a large scale (Arulampalam, Booth and Bryan 2007, 176; Grönlund and Magnusson 2016, 95).

This is because if, for example, childcare arrangements are not subsidised and private childcare options prevails, these arrangements can still be defamilising for women who can afford to pay for such arrangements which typically entails that they are already strongly attached to the workforce (Arulampalam, Booth and Bryan 2007, 176; Podestà 2017, 323). However, in the absence of publicly subsidised childcare, women who are already disadvantaged would not be able to employ costly childcare options (Grönlund and Magnusson 2016,95; Keck and Saraceno 2013, 304; Podestà 2017, 323). Therefore, since low earners are more dependent on defamilising policies to have a strong workforce attachment – thus decreasing the cost of having children and further incentivising human capital accumulation and further participation in the labour market – it could be expected that defamilisation policies would have a stronger negative effect on the gender wage gap for bottom earners than for the top earners (Arulampalam, Booth and Bryan 2007, 176; Blau and Kahn 2003, 111). It could therefore be expected that:

> **H6:** The negative effect of policies enabling defamilisation on the gender wage gap should be stronger for the bottom earners.

3. DATA AND METHDOLOGICAL APPROACH

In this chapter, I will present the data and methodological approach. I begin by introducing the data before looking closer at the dependent and independent variables and giving a short account of the control variables included. In the second part, I account for the strategies employed to estimate the effects of family policies on the gender wage gap. I begin broadly by introducing the general approach and then elaborate on specific choices related to estimation techniques and dealing with regression assumptions.

3.1. Data

For the analysis, I will use time-series cross-sectional (TSCS) data which is characterised by having repeated observations on fixed units. The sample consists of 23 member countries of the Organisation of Economic Co-operation and Development (OECD). Because some countries – especially the ones that have been members of the OECD the longest – have a better observational coverage, the sample is unbalanced. The longest time period is from 1990 to 2018. See *Table 3.1.* for an overview the countries and yearly observations in the sample for the dependent variable – the gender wage gap at median.

The sample is chosen in consideration of first, data availability and second, including as many countries as possible while avoiding conceptual stretching. First, because of the availability of comparable data over a longer time period, the main source of data is from the OECD. The OECD, in cooperation with both national and international statistical agencies, provide comparable and credible macro-panel data with good coverage both in time and space. Their data are widely used, highly accessible and cover a wide range of subjects, including detailed data on social policies. This includes my dependent and main independent variables – the gender wage gap and family policies respectively. The sample is therefore restricted to OECD countries with good coverage on the dependent variables. This further excludes OECD countries with less than 10 observations on the dependent variable in order to get a more precise picture of the development over time[1].

---

[1] These are: Estonia, Latvia, Lithuania, Luxembourg, Malta, Slovenia, Spain, the Netherlands, and Turkey.

TABLE 3.1. *Complete sample overview.*

Country, Year					
Australia	1990-2018	Germany	1992-2018	Poland	1991-2018
Austria	2000-2018	Greece	2004-2018	Portugal	2004-2018
Belgium	1999-2017	Hungary	1992-2018	Slovak Republic	2002-2018
Canada	1997-2018	Iceland	2004-2018	Sweden	1995-2018
Czech Republic	1996-2018	Ireland	1994-2018	Switzerland	1996-2018
Denmark	2002-2018	Italy	1991-2018	United Kingdom	1990-2018
Finland	1990-2018	New Zealand	1990-2017	United States	1990-2018
France	1995-2018	Norway	1997-2018		

Notes: Sample overview of countries and years for the dependent variable, the gender wage gap at median from the OECD (2022a). Most countries do not have complete coverage of yearly observations for all years between the first and last observation.

Second, in order to avoid conceptual stretching, a problem identified by Sartori (1970) as one of the main problems of comparative research, the analysis is restricted to advanced democracies with an established female employment rate – which has not been under 40% after year 2000 for the countries included - and similar family policy systems. The challenge of conceptual stretching concerns applying and comparing concepts across diverse context without stretching the concept beyond recognition, resulting in the loss of meaning (Gerring 2012, 123). Sartori's solution is climbing the "ladder of abstraction" in which making the concept more general is possible through losing some of its defining attributes (Sartori 1970, 1052). For this analysis however, there are several attributes to the concept making it less generally applicable. Therefore, for conceptual and comparative purposes of looking at as similar policies as possible across different contexts, this excludes the Latin-American, Asian and Middle Eastern OECD members[2]. These countries also have less observations on all indicators included which further justifies their exclusion from the analysis.

In order to achieve best possible coverage, data have been collected from several sources. This includes data from Eurostat, the World Bank, and the Comparative Family Policy Database. A summary of the main variables' operationalisation and descriptive statistics is presented in *Table 3.2*. For a full overview of all variables, their operationalisations, and original sources, see *Table A.1.* in Appendix A.

---

[2] These are: Chile, Colombia, Costa Rica, Mexico, Korea, Japan, and Israel.

TABLE 3.2. *Descriptive statistics.*

Variable	Operationalisation	N	Mean	St. dev.	Min	Max
Dependent variables						
Gender wage gap	Gender wage gap at median, in percentages.	487	15.03	5.81	0.38	32.74
Gender wage gap, 1st decile	Gender wage gap at the first earnings decile.	328	10.49	6.69	-9.50	32.00
Gender wage gap, 9th decile	Gender wage gap at ninth earnings decile.	343	21.36	5.52	-5.60	36.70
Independent variables *1. Parental leave, in weeks*						
Paid parental leave	Paid maternity and parental leave available to women after childbirth.	463	59.79	53.35	0	164
Paid paternity leave	Paid paternity leave reserved for the exclusive use of fathers.	463	4.01	6.43	0	28
Unpaid leave	Estimate of unpaid parental leave available to women after childbirth.	436	30.34	42.68	-10.7	156
Long leave	Long paid parental leave. 0 = less than 60 weeks, 1 = 60 weeks or more.	463	0.33	0.47	0	1
2. Childcare, proportion of enrolled children						
Childcare 0-2 years	Proportion of children aged 0-2 enrolled in formal childcare or preschool, in percentages.	274	31.02	18.92	1	78
Childcare 3-5 years	Proportion of children aged 3-5 enrolled in formal childcare or preschool, in percentages.	285	84.27	13.70	45	100
3. Added indices, from 0 (least) to 10 (most)						
Family policy generosity	Added index of *paid parental leave, paid paternity leave, childcare 0-2 years* and *childcare 3-5 years*.	467	4.700	2.707	0.00	10.00
Defamilisation	Added index of *paid paternity leave* and *childcare 0-2 years*.	463	2.124	2.380	0.00	10.00
Familisation	Added index of *unpaid leave* and *long leave*.	463	2.941	2.564	0.00	10.00

Notes: Summary of descriptive statistics. Data from various sources. See Appendix for full overview of data sources for each variable.

It is also worth mentioning that the gender definitions in the databases are based on legal definitions within countries, distinguishing exclusively between men and women. There are no comprehensive databases yet consisting of a third gender option. Ideally, I would therefore

have liked to consider wage differences in the labour market using a more flexible gender definition. This is also the case for the family policies, especially the leave policies, distinguishing exclusively between a mother and a father and do not take into account neither family constellation nor same sex couples.

3.1.1. Dependent variables: The gender wage gap

The median gender wage gap

The main dependent variable is the gender wage gap from the OECD Employment Database's indicators on earnings and wages (2022a). The gender wage gap is measured as the difference between male and female median wages divided by the male median wages for all full-time employees, defined as those working more than 30 hours per week (OECD 2022a). This is a standard procedure for measuring wage gaps across countries. As raw wages and the resulting wage gap will depend on the countries' economies and wage dispersion within the country, they cannot be compared cross-nationally. However, since this variable is measured as a percentage away from the male median wages within each country, the variable is standardised and can be compared across countries. A drawback from using this measure which might impact the results is that this estimate is based on the countries' reported average earnings, which for 17 of the 23 countries are based on monthly averages. Canada, Australia, and the United States, on the other hand, are based on weekly estimates; Denmark and New Zealand on hourly estimates; and Finland and Austria on yearly estimates (OECD 2022a). *Figure 3.1.* gives an overview of the median gender wage gaps for the whole sample and shows considerable variation between the countries.

The variation between the countries' wage gaps is presented by the dispersion of the individual lines in *Figure 3.1*. The largest difference in gender wage gap is found between the United Kingdom in 1990 with a gender wage gap of 32.74 compared to Hungary in 2006 with 0.38. This large difference can, of course, also be due to the difference in time between the measured values. For the latest observations in 2018, the difference is smaller but still notable. For 2018, the highest measured wage gap was in the United States with 18.91, while the lowest is found in Denmark with 4.86.

FIGURE 3.1. *Overview of all countries' median wage gaps, 1990-2018.*
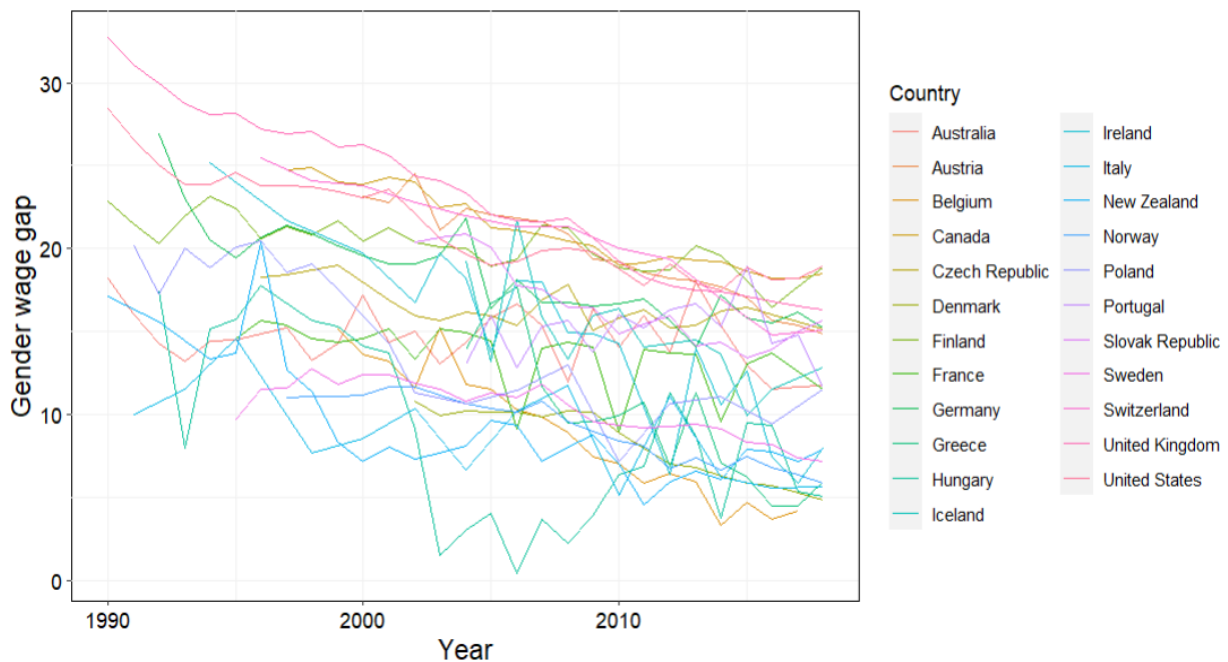
Notes: Overview of the differences in the gender wage gap. Visualisation based on R-package *ggplot2* (Wickham 2016). Data from OECD (2022a).

The gender wage gap at first and ninth earnings decile

There is an important limitation to looking at the median gender wage gap. As discussed in the theoretical background, the median wage gap in a country can hide large differences in wage gaps across the wage dispersion within countries (Albrecht, Björklund and Vroman 2003, 146; Bonjour and Gerfin 2001, 408). In order to account for potential differences in effects of family policies on the wage gap across the wage dispersion, I use two additional dependent variables: the gender wage gap at the first and ninth earnings decile. The variables are based on gross earnings of full-time employees reported in the OECD Distribution of Earnings Database (OECD 2021a). The first decile refers to the wage gap for the ten percent lowest earners, while the ninth decile refers to the ten percent highest earners within each country. The observations for the deciles start in year 2000 and thus cover a shorter time period. The sample consequently declines by around 150 observations, which can have an impact on the results. Nevertheless, this decline leaves a viable number of observations and the ability to check for diverse effects of the same policies.

From the descriptive statistics presented in *Table 3.2.* we see that there are variations between the three dependent variables, giving an indication that there is important variation lost in

looking only at the median. For example, we can see that the mean value is higher for the ninth decile than the first and at median, which indicates a glass ceiling effect where the wage gap is larger for the top earners. This is presented graphically in *Figure 3.2.* where each line indicates the average wage gap for OECD-countries from 2000-2018. The figure shows that there has been a notable decline in the average wage gaps for all measures – from 13.3 to 9.3 for the first decile and from 23.7 to 17.3 for the ninth decile – but the wage gap at the ninth decile is notably larger than the median and first decile for the whole time period.

FIGURE 3.2. *Average gender wage gaps, 2000-2018.*
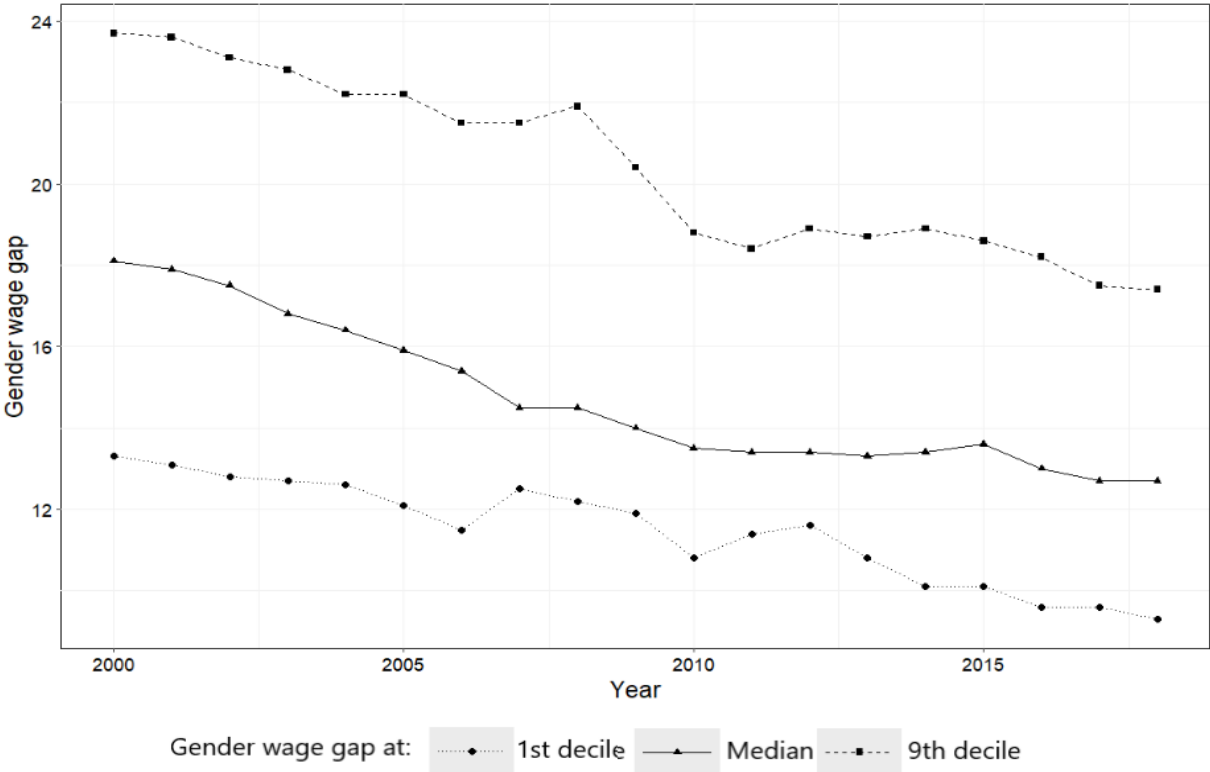
Notes: Average gender wage gaps for OECD-countries at median, first and ninth earnings decile from 2000-2018. Visualisation based on R-package *ggplot2* (Wickham 2016). Data from OECD (2021a).

3.1.2. Independent variables: Family policies

The independent variables consist of family policies, specifically indicators on parental leave and childcare. The main variables are constructed indices on familisation and defamilisation, for which I follow Lohmann and Zagel (2016, 61) in their empirical construction of the concepts. As I focus on intergenerational dependencies for children, I do not include indicators on care for elderly. I also include an index on family policy generosity – consisting

of the available parental leave and childcare policies – against which I test the dimensions of familisation and defamilisation. I therefore start by estimating the effects of overall family policy generosity and then compare the results of the disaggregated dimensions to empirically test the argument of disaggregating the concept. Because the indices consist of different observations and thus vary in sample sizes, they are not completely comparable. However, the indices are constructed in the same way and rescaled to a scale from 0 (least) to 10 (most) in order to be more comparable across model specifications.

Family policy generosity

The family policy generosity is created to resemble the family policy generosity indices discussed in the literature review of early research on gender wage gap and welfare state constellations. It is an average unweighted index consisting of four individual standardised indicators: (1) paid parental leave, (2) paternity leave, (3) childcare for children aged two or younger and (4) childcare from ages three and up to compulsory school age. Creating an average indicator like this is preferable to an additive index when it is theoretically uncertain whether "all social policies contribute equally to the underlying dimension" (Kleider 2015, 511). It is important to mention, however, that while an average index provides better coverage, some observations may be more unreliable than others. This is the same for all indices as they are constructed in the same way. All parental leave indicators are from the OECD Family Database (OECD 2022b) and have been supplemented with data from the Comparative Family Policy Database (Gauthier 2011). The Comparative Family Policy Database is based on OECD data, but estimates have been added where there are missing observations. I therefore supplement the family policy indicators where observations are missing with Gauthier's estimates (Gauthier 2011). Indicators on childcare are also from the OECD Family Database but supplemented with the same indicators from Eurostat (Eurostat 2022). Eurostat also provides high-quality comparable indicators for European countries in cooperation with national statistical authorities. They overlap with the OECD countries and have therefore been used to supplement the data where there are missing observations.

First, paid parental leave is measured as the total number of weeks of which women can be on paid leave after the birth of a child (OECD 2022b). I focus on a very parsimonious indicator as it is the most comparable across countries and across time. It therefore does not take into account the multitude of rules on issues like length, replacement rate or potential

transferability between parents (Kleider 2015, 511). In addition, the measure only reflects entitlements at national or federal level and does not reflect potential regional or firm-specific variations. Furthermore, a distinction is made between maternity and parental leave, and as the former is considered a health-related policy for both mother and child, I focus on paid parental leave with the added economic goal of balancing work and family (Kamerman and Moss 2009, 3). As shown in *Table 3.2.,* there is considerable variation in available weeks of paid leave between the countries. While the Slovak Republic, the Czech Republic and Finland all offer over 160 weeks paid parental leave, Canada, Australia, and the United States offer none nationally. This, of course, does not mean that they do not have any kind of parental leave, but that there is no national legislation on the right to paid parental leave.

Second, paternity leave refers to weeks of paid leave reserved for the exclusive use of fathers, including entitlements to paid paternity leave, "father quotas" or weeks of paid parental leave that must be taken by the father to qualify as "bonus" weeks (OECD 2022b). Paid paternity leave is a relatively new policy for several countries and has a smaller overall scope than the total offered parental leave. The countries offering the longest paternity leave is France and Austria with 28 and 26 weeks respectively. Canada and Ireland, on the other hand, do not offer any weeks specifically for the father.

For the indicators on childcare, the first is childcare for children up to two years of age. This is measured as a share of children in that age group enrolled in formal childcare and pre-school (OECD 2022b). Formal childcare here refers to all arrangements other than the family and therefore also includes private childcare. The indicator is supplemented by Eurostat's indicator "Percentage of children under 3 cared for by formal arrangements" (Eurostat 2022). This indicator is an aggregated variable based on the EU Statistics on Income, Social Inclusion and Living Conditions (EU-SILC). Similar to the parental leave policies, there is a lot of variation between the countries. Finally, children aged three and up until compulsory school age enrolled in early education and formal childcare are measured as the share of all children in that age group. This is the indicator in which the OECD and Eurostat indicators differ the most, but the values are marginally different. The OECD includes children aged from three to five years, while Eurostat has operationalised it as "to the starting age of compulsory primary school" which differs between five and six years (OECD 2022b; Eurostat 2017). It is also worth mentioning that the childcare variables are those with the most missing observations, especially for the English-speaking countries. This can affect the number of observations, sample composition and thus the comparability of the models.

Defamilisation

The defamilisation indicator consists of two indicators: (1) paternity leaves and (2) childcare for children aged two or younger. I only include childcare for the youngest children in this index for two reasons. First, childcare for children under three years old has been explicitly designed in order to help families balance care and employment, while childcare programs for children aged three years and until compulsory school age are more often designed as educational programs (Budig, Misra and Boeckmann 2016, 159; Vandenbroeck 2020, 169). Second, in most European and OECD countries there is some kind of care and education provision for all children over the age of three years old, meaning that there is hardly any variation between or within countries (Keck and Saraceno 2013, 304). For this dimension, the countries with the least amount of defamilisation are the Slovak Republic, the Czech Republic and Hungary – all with under 0.10 out of 10 for several years. The countries with the most defamilisation, on the other hand, are France, Portugal, and Belgium – all with over 7.00 for a longer period and closely followed by Austria and the Scandinavian countries Norway, Sweden, and Denmark.

Familisation

Two indicators make up the dimension of familisation. These are (1) unpaid parental leave and (2) long paid parental leave. First, the measure of unpaid leave is created as an estimate by the total number of weeks of maternity and parental leave which women can take with her job protected but disregarding payment conditions, minus total maternity and parental leave available for mothers with pay. Both paid and job protected leave are measured in weeks and are gathered from the OECD Family Database, supplemented with data from the Comparative Family Policy Database (OECD 2022b; Gauthier 2011). Second, I include a dichotomous measure indicating long paid parental leave over 60 weeks where 1 indicates paid parental leave over 60 weeks and 0 indicates less than 60 weeks. Long leave here is defined as over average, the average of all countries in the sample being 59.79 which coincides perfectly with what is generally considered a cut-off point for long paid parental leaves (Olivetti and Petrongolo 2017, 209).

The lowest scorers for this index are Belgium and Canada – both under 0.40 out of 10 for a considerable amount of time – while the countries with the most familisation are the Czech

Republic, Germany, and Finland – all with over 8.00 of 10 for several years. These values somewhat correspond with the defamilisation index, only inverted. For example, while Belgium has the least amount of familisation they are among the countries scoring the highest on the defamilisation index. This indicates that these policies – because they do coexist together – present somewhat of a trade-off, where more familisation (defamilisation) also indicates less defamilisation (familisation). This is further supported by their statistically significant negative correlation[3].

Paid parental leave

Finally, as the duration of paid parental leave is argued to be both defamilising and familising, I do not incorporate this indicator into the concepts above but instead test this effect separately. I therefore add a quadratic term of paid parental leave in order to account for nonlinearities in the effect. I expect a U-shaped effect where the shortest and longest paid parental leave increases the gender wage gap in general, as presented graphically in *Figure 3.3.* on the next page. As its effects on the first and ninth earnings decile are ambiguous, I expect a similar U-shaped effect for these as well. As mentioned earlier, the indicator of paid parental leave is a general one and does not take into account the various conditions embedded in the policies. Ideally, I would therefore also have controlled for the replacement rate as it gives information on the specific financial amount covered by the parental leave, and not just on whether or not the leave is paid. However, data on this are not available for the countries over a satisfactory time period.

3.1.3. Control variables

Finally, in order to account for additional factors that are known, both from empirical work and the theoretical literature, to affect the gender wage gap, I include control variables for these factors in the analysis. There is of course the always-present problem of omitted variable bias, and I would ideally also have controlled for whether these policies are regulated at a supranational or regional level. However, because of both data availability and the scope of the thesis – looking at the wage gap and family policies nationally – such controls are not included.

---

[3] Pearson's R = -0.312 with a p-value of > 0.000.

FIGURE 3.3. *Expected curvilinear effect of paid parental leave.*
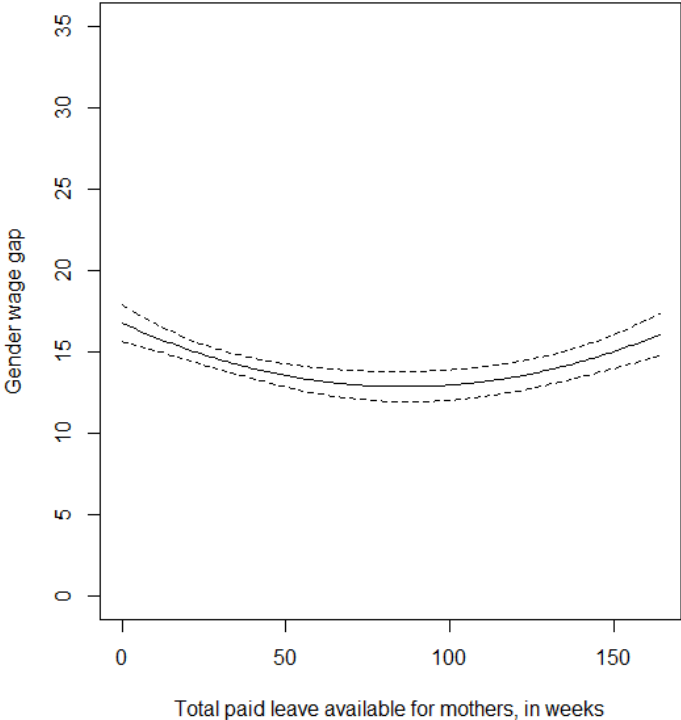
Notes: Graphical illustration of the expected U-shaped effect of paid parental leave.

The control variables included are: (1) female employment rate, (2) educational difference, (3) female public sector employment, (4) economic inequality and gross domestic product (GDP) per capita and (5) central bargaining coverage and strictness of employment protection legislation.

First, I control for the female employment rate because the effect of policies on wage gaps depends, in part, on women being in the labour force, i.e., if women work less, they also earn less (Budig, Misra and Boeckmann 2012, 169). The variable is from the OECD and measured as the female labour force divided by the total working-age female population, defined as people aged 15 to 64 (OECD 2022c). I do not control for the incidence of part-time employment. While higher incidence of part-time employment among women is one important explanation of the gender wage gap (Budig, Misra and Boeckmann 2016, 140; Kunze 2018, 382), empirically the sample is not biased towards part-time employment as the dependent variables exclude part-time completely.

Second, in order to control for aggregated differences in human capital I include a measure of the difference in tertiary enrolment ratio between men and women, constructed by taking the

female enrolment ratio minus the male enrolment ratio. I include tertiary education enrolment as there are virtually no differences between secondary enrolment for men and women within the countries included, nor between the countries[4], but some variation for tertiary enrolment. The gross enrolment ratio for both genders is from the World Bank's World Development Indicators and is defined as "the ratio of enrolment regardless of age, to the population of that age group that officially corresponds to the level of education", here defined as five years after finished secondary education (World Bank 2022a). As a measure it does not take into account completion of education or the gender differences for each field of study but indicates the capacity of the educational system as a whole and the general differences in enrolment. Ideally, I would also have included a control for work experience – the other important human capital indicator (Becker 1993, 33; Mincer and Polachek 1974, 78), but because of data availability and the inherent difficulties measuring such a diverse concept at a macro-level and cross-nationally, such a control is not included.

Third, previous research also indicates that occupational segregation drives the gendered wage gap (for example Estévez-Abe 2005, 2006; Mandel and Semyonov 2005, 2006; Hook and Pettit 2016). Therefore, to control for occupational segregation, I include a measure of how many females are employed in the public sector. This is measured as the share of women employed in public sector as a share of female employment and is from the World Bank's Worldwide Bureaucracy Indicators (World Bank 2022b).

Fourth, since it is expected that a more compressed wage structure, i.e., less overall economic inequality, raises the wage floors in occupations that are female dominated and overall makes less of a difference between the wages of people in general (Blau and Kahn 2003; Hook and Li 2020, 259), I add a control variable of economic inequality. As a measure, I use the wage dispersion between the first and ninth earnings decile. This is measured as the ratio of the ninth to first earnings decile of gross earnings of full-time employees and is from the OECD Employment database of decile ratios (OECD 2022d). I also include Gross Domestic Product (GDP) per capita as a control, mostly used to measure macroeconomic changes both within and between countries. The indicator is from the World Bank's World Development Indicators and measured in US dollars (World Bank 2021).

---

[4] The secondary enrolment rate is around 100 percent for all countries. The average total enrolment rate in secondary education for all countries is 109 percent, 110 for girls and 108 for boys. Values are over 100 percent as the ratio also takes into consideration those repeating years.

Finally, I control for the institutional features based on the Varieties of Capitalism-literature suggested to have an effect on the gender wage gap by Estévez-Abe (2005, 2006), thereby separating features of coordinated market economies (CMEs) from liberal market economies (LMEs) in terms of central bargaining coverage and employment protection legislation. Therefore, I add a measure of both as control variables. First, central bargaining coverage is an added unweighted measure consisting of union density – measured as the ratio of wage earners that are trade union members to the total number of wage earners – and collective bargaining coverage – measured as the share of employees covered by a collective agreement in force as a proportion of the total number of eligible employees (OECD 2022e, 2022f). Second, strictness of employment protection legislation is an added unweighted measure of the strictness of regulation of both individual and collective dismissal and the strictness of the use of fixed and temporary contracts, both indicators from the OECD Indicators of Employment Protection (OECD 2021b, 2021c).

3.2. Methodological approach

In this part, I will first account for the choice of using a time-series cross-sectional approach and its advantages. I then go further into the estimation techniques used, more specifically the three dependent variables, the choice of random two-ways effects over fixed effects modelling and how to deal with the regression assumptions. Lastly, I look at some important limitations to the approach.

3.2.1. Why use time-series cross-sectional modelling?

In order to assess the effects of the different family policies on the gender wage gap, I will use a large-N quantitative method and carry out a multivariate regression analysis using time-series cross-sectional data. In general, large-N quantitative approaches have the ability to achieve both breadth and comparability, both of which I aim for in this thesis when looking for general patterns across contexts. In addition, as TSCS data combine cross-section and time-series analyses, I am able to look at how the policies affect the wage gap within as well as between the countries (Beck and Katz 1995, 643).

There are three main advantages of this approach, as listed by Baltagi (2021, 6-9). These are (1) the possibility to model individual heterogeneity, (2) that the data provided is more informative and (3) it is suitable for looking at dynamics of change.

First, one advantage when combining time-series and cross-sections in a macro-panel structure is that it allows for more advanced regression techniques that can model the specific context both within and between countries (Baltagi 2021, 6). While regular Ordinary Least Squares (OLS) assumes homogeneity and does not consider heterogeneity across groups, not controlling for potential dependencies across both space and time run the risk of obtaining biased results (Baltagi 2021, 6). For example, there might be country-invariant variables, such as state religion or educational system, or time-invariant variables – such as large external shocks – affecting the dependent variable that regular time-series or cross-section studies cannot account for (Baltagi 2021, 6). There are two dominant approaches to account for this grouped nature of the data, fixed and random effects modelling, which I will discuss further down (Clark and Linzer 2015, 339).

Another, but also related, advantage is that since both dimensions increase the sample size, the data structure gives more informative data by providing more variability and degrees of freedom (Baltagi 2021, 7). One of the main advantages is therefore simply the increase in observations. When it comes to statistical methods, size does matter. This is because as the sample size increases, the confidence that the observed relationship is accurately representing the underlying population also increases (Kellstedt and Whitten 2018, 164). This means more evidence for testing the hypotheses and providing insurance against errors caused by the presence of random noise in the data (Gerring 2012, 365).

Third, and perhaps most important, adding a time dimension makes the approach suitable to study dynamics of change. This is particularly important when the variables of interest are slow-moving variables with little within variation, such as in this case where I look at specific policies and their institutional setting, because a longer time frame can shed light on longer processes (Baltagi 2021, 8). In contrast to much of the previous research on the gender wage gap that have been based on cross-sections, either through the mentioned decomposition methods or multilevel modelling including both an individual- and a country-level dimension (Budig, Misra and Boeckmann 2016; Hook and Li 2020, 250; Kunze 2018, 367), the approach is able to specifically model slow changes or address reverse causality by forcing the theorised cause to proceed the outcome. The difficulty of including a longer time frame at the individual level is partially because they rely heavily on comparable micro-data across both time and space for characteristics that are often hard to measure at an individual level (Ehrenberg and Smith 2018, 195). Therefore, while these methods achieve more precision in looking at both individual characteristics and their institutional contexts, this is an essential trade-off in which modelling longer processes at a higher level trades off some of this precision.

3.2.2. Estimation techniques

Estimating the effect across the wage distribution

As already discussed, there is a strong argument made in the literature for looking at different effects on the wage gap at different points of the wage distribution (see for example Albrecht, Björklund and Vroman 2003; Arulampalam, Booth and Bryan 2007). In order to achieve this and estimate whether the effects differ across the wage distribution, I use the three dependent

variables as presented above. This approach is similar to quantile regression which allows for heterogenous effects across the outcome's distribution in each model and thus "identify more subtle effects which would be missed by the application of mean regression" (Alexander, Harding and Lamarche 2011, 48). By using these three dependent variables already defined in deciles of the wage distribution and comparing the results from the different models, it is therefore possible to see if the effects are similar for the wage gaps at first and ninth earnings decile, and thus if the median wage gap hides important differences.

Choosing between fixed and random effects modelling

As already mentioned, as regular OLS treats all observations as conditionally independent, it cannot take into account potential group-level variation, either within or between the units (Baltagi 2021, 6). Therefore, in cases where the dependent variable(s) exhibits – or is expected to exhibit – grouped variation beyond what can be explained by the independent variables, fitting a standard OLS can lead to poorly fitting models (Bell and Jones 2015, 133; Clark and Linzer 2015, 399). This, in turn, makes the results unreliable (Bell, Fairbrother and Jones 2019, 1058). Using TSCS data, there are two dominant approaches to remedy this: fixed effects (FE) and random effects (RE) modelling. The choice between these two has led to some debate in the methodological literature because both approaches have convincing advantages and disadvantages, and the choice is as much a technical fix to the modelling itself as it is related to the substantive interest of the research.

In social sciences, the FE has been the "golden standard" as it is credited with theoretically controlling for all unseen factors and therefore deal with the always-present problem of omitted variable bias (Bell and Jones 2015, 133). This bias results from the failure to either include a variable that belongs in the regression model or to not include a noncausal variable in the model and will cause the expected value of the parameter estimate that we obtain from the sample to "not be equal to the true population parameter" (Kellstedt and Whitten 2018, 229). One of the greatest merits of the FE approach is therefore to relieve this problem completely, which it does by controlling for all variation between the units by using the units themselves included as dummy variables (Bell and Jones 2015, 138). In this way, all variation between the units and stable characteristics within the units are effectively controlled for, keeping only the within variation for each unit. The estimated coefficients therefore indicate

how much the dependent variable changes over time, on average per country when x increases by one unit.

However, there are some important limitations to this approach. First, using the units' characteristics themselves as the dummy variables assumes homogeneity across units, i.e., that the units are as similar as possible at a higher between-level (Bell and Jones 2015, 137). Although the countries do have similar policies, assuming homogeneity is debatable. Another limitation, Bell and Jones (2015) argue, is that by controlling instead of modelling context, FE models lose out on important information: "in controlling out context, FE models effectively cut out much of what is going on – goings-on that are usually of interest to the researcher, the reader and the policy maker" (Bell and Jones 2015, 134). FE models are therefore a good choice when the within effect is of greatest interest and the units are somewhat similar. Furthermore, FE models are especially problematic when estimating time-invariant or slow-moving variables with minimal variation within each unit (Allison 2009, 2; Bell, Fairbrother and Jones 2019, 1053), which is the case here with the independent variables being relatively stable policy configurations. This is because all degrees of freedom at higher-level are consumed and in the case of time-invariant variables, the presence of multicollinearity – when there is a perfect linear relationship between two or more independent variables (Kellstedt and Whitten 2018, 232) – makes it impossible to estimate the unique effect of the variable (Bell and Jones 2015, 139; Clark and Linzer 2015, 403). So, while controlling out all potential factors, it cannot be used to estimate how and which of these do have an effect. Consequently, the FE models introduce more variance, are sample-dependent and make out-of-sample predictions and generalisability unsuitable (Clark and Linzer 2015, 404).

Therefore, because I aim to model the context of relatively stable characteristics between countries and believe these differences will have important effects on the dependent variables, RE models assuming and modelling heterogeneity between units are more suitable. RE models do this by partitioning the added unexplained residual variance into two: higher-level variance between units and lower-level variance within units, and instead of these being drawn from an estimation of the units themselves, these are random estimates drawn from a normal distribution (Bell and Jones 2015, 135). The regression coefficient therefore represents the average effect on the dependent variable when x changes both over time *and* between countries by one unit. In addition, because the RE models say something about the underlying context, generalisations and out-of-sample predictions are feasible (Clark and Linzer 2015, 404).

On the other hand, the most common objection to the use of RE is the violation of the modelling assumption that there is no correlation between the covariates of interest and the random effects, i.e., no confounding due to omitted variables (Clark and Linzer 2015, 407). Of course, there is always the possibility of omitting important but unknown factors that are relevant when attempting to simplify complex societal structures, but when this assumption is broken the RE models run the risk of introducing too much bias and underfitting the model. One solution, in contrast to FE models, is to add all important stable variables, which I do in my models. This entails another trade-off in which the model is better fitted but the sample size decreases because of missing observations. Also, Clark and Linzer (2015) running simulations to estimate the trade-offs between FE and RE for various specifications find that in the case of smaller datasets, which they define as fewer than 200 total observations, and slow-moving independent variables – both of which is the case here – the root mean square error (RMSE) of the estimated coefficient is smaller with the RE than FE even when the estimates are biased (404). Furthermore, this lends support to the statement of Bell and Jones (2015) that the assumption of zero correlation completely ruling out RE as an option is misguided (144).

Another widely used method of choosing between FE and RE is the Hausman test. The Hausman test tests whether the unique errors are correlated with the regressors – the null hypothesis being that they are not – and if significant is usually interpreted as FE is the preferred model (Allison 2009, 89). The results from the Hausman test for the models used here are somewhat disorienting, as for some models the Hausman test clearly indicates either RE or FE, but for others the significance is just around the usually accepted five-percent cut-off-point. However, there has been some critique to basing this decision solely on this test. Bell and Jones (2015), for example, argue that the test only tells if the RE is consistent, not if the FE is best according to other assumptions and the research objectives (138). Clark and Linzer (2015) also argue that the test does not differentiate between a small correlation and no correlation at all, and that, as shown above, this should be regarded in relation to other trade-offs (403). The choice should therefore ultimately be based on the study's objective, which here is to model the context between countries as well as within countries, justifying the choice of RE over FE. This is further justified by the data itself showing more variation

between than within the countries, indicating the importance of modelling these differences[5].

In addition, random effects modelling allows the choice between within, between or two-ways effects averaging the within- and between effects. I follow Bell, Fairbrother and Jones (2019) in their argument that the two-ways effects model is the most general because it allows for "distinct within and between effects […] with heterogeneity modelled at both the cluster (level 2) and observation (level 1) level", for which level 2 refers to the between variation and level 1 to the within variation (Bell, Fairbrother and Jones 2019, 1052). This can be argued to be a hybrid solution adding effects within and between. It is therefore a technical fix if the assumption of no correlation between covariates and residuals is broken, as well as if when the normality assumption of the random effects is violated because "the small biases that can arise in such models will often be a price worth paying for the added flexibility" (Bell, Fairbrother and Jones 2019, 1069).

In sum and based on these arguments, random effects models with two-ways effects have been chosen as the best suited estimation technique. I use the R-package *plm* to estimate all models (Croissant and Millo 2018). I nevertheless estimate a fixed effects alternative as a robustness test for the analysis both to check the robustness of the results and whether these change when only looking at the within-effects.

Dealing with regression assumptions

In addition to the regression assumptions already mentioned of omitted variable bias and normality, there are other assumptions that have to be met for both the cross-sectional and time-series dimension of the data. These are (1) homoskedasticity, (2) no serial correlation, (3) stationarity and (4) no multicollinearity.

First, one of the key regression assumptions is that of homoskedasticity, i.e., that residuals are distributed with equal variance (Kellstedt and Whitten 2018, 208). When this assumption is broken, there is heteroskedasticity, which means that the effect of a variable differs across the distribution of a dependent variable (Beck and Katz 1995, 635). This, in turn, can have the consequence of making the results unreliable by fitting some cases better than others and

---

[5] The mean variance for each dependent variable within countries are: median = 10.378, first decile = 10.784, ninth decile = 14.901 compared to the mean variance between countries: median = 30.022,
first decile = 27,207, ninth decile = 27.428.

effecting the estimation of the variables' confidence intervals. The results can therefore still be consistent estimates of the regression coefficients, but the standard errors will be biased and unreliable (Baltagi 2021, 109; Kellstedt and Whitten 2018, 209). When testing for heteroskedasticity using the Breusch-Pagan test for the null hypothesis of homoskedasticity, the results show that there is some degree of heteroskedasticity present in all models[6]. This means that the parameters can still be estimated consistently but for valid inferences, applying a covariance matrix with robust standard errors is essential (Zeileis 2004, 2). I therefore add heteroskedasticity consistent standard errors, allowing for the correct computation of the confidence intervals and statistical significance in all models[7].

A second assumption which applies when using time-series is that of no serial correlation, or autocorrelation. This happens when "the stochastic term for any two or more cases are systematically related to each other" (Kellstedt and Whitten 2018, 209), or more simply stated: when there is a high degree of dependence between values around the same time period. This can cause the regression results to be unreliable as the standard errors decrease and the adjusted R-squared increase, making the model appear more significant than it really is. To check for autocorrelation, I use the Breusch-Godfrey test of serial correlation in panel models where the null hypothesis is that there is no serial correlation[8]. For most models, some degree of autocorrelation is present in the models, with the highest p-value being 0.622 and the lowest being 0.254. This is not particularly alarming, but because of the presence of heteroskedasticity I nevertheless estimate robust standard errors for all models which also help with serial correlation.

Third, the data is assumed to be stationary. This means that the means, variance, and covariance are constant over time for the variables (Baltagi 2021, 339). This is especially a concern when the variables are known to follow a pattern for the years, which initially seems to be the case with the dependent variables decreasing across the years. However, running an Augmented Dickey-Fuller test where the null hypothesis is non-stationarity[9], the p-value does not exceed 0.030 for any variables and does not increase drastically when changing the time

---

[6] The Breusch-Pagan test for heteroskedasticity is from R-package *lmtest* (Zeileis and Hothorn 2002).

[7] I use the R-package *sandwich* for adding robust standard errors. For the FE models, I use the suggested Arellano-method, and the White method for the RE models. All models are estimated with the *HC3* type of bias, which is recommended when the number of observations is relatively small (Zeileis, Köll and Graham 2020).

[8] The Breusch-Godfrey test used is from the R-package *plm* (Croissant and Millo 2018).

[9] The Augmented Dickey-Fuller test is from the R-package *tseries* (Trapletti and Hornik 2022).

lag between zero, five and ten years, thus indicating that the data does not violate this assumption.

Lastly, the final regression assumption, as already mentioned above, is that of no perfect multicollinearity. There are two main ways of checking this assumption, either by checking the independent variables' bivariate correlation or using a Variance Inflation Factor (VIF) test for each model. When testing both the bivariate correlations and the VIF test, none of the independent variables show signs of multicollinearity. First, the highest correlation is between *wage dispersion* and *central bargaining coverage*, with a Pearson's R of 0.667, followed by a Pearson's R of 0.648 between *GDP per capita* and *female employment,* both with a p-value smaller than 0.000. See *Figure A.1.* in *Appendix A* for a correlation matrix of all control variables. The VIF does not show high values for the models here either, with the highest being a value of 2.775 for *GDP per capita* in model 1C estimating the effect of familisation on the median gender wage gap. For the models adding a quadratic term to paid parental leave, the VIF is naturally somewhat higher, but this is expected in second degree estimations[10].

3.2.3. Limitations to the approach

Following a growing dominance of positivist philosophy of science, technological advancement and the development of advanced statistical tools, quantitative methods have become both more popular and accessible for social scientists (Laitin 2002, 632; Moses and Knutsen 2012, 93), but while quantitative methods have their advantages there are some important limitations. As Baltagi writes: "Panel data is not a panacea and will not solve all the problems that a time-series or a cross-section study could not handle", or in a larger setting, the problems a qualitative approach could not handle (Baltagi 2021, 11). For this thesis, there are two important limitations to the approach: the first one regarding data quality, and the second regarding the inability to address causal complexities.

First, regarding data quality, one of the main criticisms of quantitative approaches is that they do not pay enough attention to the quality of the data being analysed, herein the conceptualisation, measurement, reliability, and validity. As Sartori famously stated:

---

[10] Highest VIF for the quadratic models is 13.161 for the quadratic term of paid parental leave in the model estimating the effect of paid parental leave on the median wage gap.

"Concept formation stands prior to quantification" (1970, 1038) and without proper concepts and measurements, there is no meaning in the statistical inferences. There are three main concerns, in which the first two consider reliability and the third the validity.

A first concern is unbalanced panels and missing coverage, which decreases the sample drastically and impedes the overall comparability as the sample will be somewhat different under different specifications (Baltagi 2021, 9). A second concern is the potential distortions of measurement errors, thereby lacking reliability in the measurement themselves so that they are not repeatable and consistent across observations (Baltagi 2021, 9; Kellstedt and Whitten 2018, 112). This is especially the case with large-N-large-T data based on yearly observations in different countries. Here, I rely on the variables being measured the same way across the whole time period and in all countries, which is not always the case and can further decrease comparability. Some measurements for example, such as the wage gaps, are based on national estimates and thus differ somewhat between countries. Fortunately, all data sources aim at comparability and thus provide standard definitions for most indicators, for example full-time employment being defined as more than 30 hours per week.

A third concern regarding data quality is the validity, both internal and external; Does the measure accurately represent the concept it is supposed to measure, and can it be applied elsewhere? (Kellstedt and Whitten 2018, 114). Here, the institutional indicators represent a very rough measure and do not fully capture national specificities, a problem resulting from aggregation and the conflicting goal of applying the same knowledge to diverse settings while saying something substantial about the individual processes (Gerring 2012, 112; Slater and Ziablatt 2013, 1305). While using such general measures decreases the problem of conceptual stretching and increases comparability, a lot of information is lost in the process. For example, the family policy indicators, in order to be comparable, do not take into account the various criteria for which the policies can be used. The comparison can also be biased by the fact that regional and local variations are not taken into account (Thévenon 2008, 174) Also, the analysis does not account for whether the policies are actually being used, by whom, and the flexibilities that exist for these policies, it only looks at the overall policy setting, (Rubery 2011, 1106). In addition, by looking only at full time employment, I am not able to account for part-time employment, as already mentioned, and potential self-selectivity, i.e., why some do not work full time or only work part time (Baltagi 2021, 10). Because their wages are missing from the wage gap calculations, these are dropped from the sample completely.

The second important limitation to the approach is addressing causal complexities and looking closer at causal processes and mechanisms. As Gerring (2004) argues: "causal arguments depend not only on measuring causal effect [,] they also presuppose the identification of a causal mechanism" (348), and one of the limitations of statistical approaches is the inability to examine such a causal mechanism (Moses and Knutsen 2012, 93). Rather than looking at specific cases and the particular contexts in which these causal connections exist, the logic of the statistical method is probabilistic in that it assesses the probability that the relationship we find is due to random chance (Kellstedt and Whitten 2018, 15). In short, it presents a simplification of the processes at study. Therefore, quantitative methods can only achieve a probabilistic estimate of how much a variable affects the phenomenon that is being studied but cannot give a substantial explanation of the causal pathways. Of course, this is not always the goal and what I attempt to do here is indeed drawing large, more general inferences about the connections between the gender wage gap and family policy dimensions.

4. RESULTS AND ANALYSIS

In this chapter, I present the results from the analysis and show both how different family policies affect the gender wage gap and how these differ for various points of the wage distribution. In the first models, I address the questions of how family policy affects the gender wage gap and whether the effects are different for the dimensions of familisation and defamilisation when disaggregating the encompassing family policy generosity index. I also include a separate model estimating the potential curvilinear effect of the duration of paid parental leave. For the following models, I address the second part of the research question asking whether these effects differ across the wage distribution and estimate the effect of family policies on the gender wage gap at first and ninth earnings decile. I finish this chapter by checking the sensitivity and robustness of the results.

Throughout the analyses, the models are numerated based on the dependent variable and differentiated by letters based on the independent variable. Therefore, models *(1)* have the median wage gap as the dependent variable, models *(2)* the gender wage gap at first earnings decile and models *(3)* the wage gap at ninth earnings decile. Similarly, all models *(A)* have the family policy generosity index as the independent variable, models *(B)* the familisation index, models *(C)* the defamilisation index and models *(D)* have paid parental leave and its quadratic term as the independent variables. All models are estimated using a random two-ways effects model with robust standard errors included as a corrective for heteroskedasticity.

4.1. The family policy dimensions' effects on the median wage gap

The first regression models (1A-D) estimate the effect of the family policy dimensions on the median wage gap, addressing the first part of the research question: How do family policies affect the gender wage gap?

The first model (1A) starts by estimating the effect of family policy generosity on the median gender wage gap. The main purpose of starting with the general effect of family policy on the gender wage gap is constructing a base against which I test the dimensions of familisation and defamilisation, thereby empirically testing the support for this disaggregation as proposed in the literature.

TABLE 4.1. *Regression results, family policies' effects on the median wage gap.*

Dependent variable: Median wage gap	Generosity (1A)	Familisation (1B)	Defamilisation (1C)	Paid leave (1D)
Family policy generosity	**-0.753 *** (0.000, 0.221)			
Familisation		0.197 (0.223, 0.161)		
Defamilisation			**-0.348 *** (0.088, 0.203)	
Paid parental leave				**-0.102 *** (0.058, 0.053)
Paid parental leave^2				0.000 (0.511, 0.00)
Controls				
Female employment	0.127 (0.940, 6.006)	0.024 (0.774, 0.085)	0.114 (0.140, 0.077)	0.017 (0.868, 0.099)
Female public sector empl.	-0.450 (0.874, 4.014)	-1.919 (0.790, 7.197)	-0.770 (0.901, 6.161)	-0.493 (0.954, 8.460)
Educational difference	**-0.100 **** (0.009, 0.038)	**-0.121 **** (0.004, 0.041)	**-0.099 **** (0.012, 0.039)	**-0.138 **** (0.005, 0.049)
Wage dispersion	**1.132 *** (0.083, 0.642)	1.177 (0.155 0.824)	**1.190 *** (0.063, 0.637)	0.706 (0.512, 1.075)
GDP per capita	**-0.203 *** (0.000, 0.053)	**-0.152 **** (0.027, 0.068)	**-0.161 **** (0.007, 0.059)	**-0.268 *** (0.000, 0.074)
Central bargaining	**6.426 *** (0.000, 0.989)	**6.161 *** (0.000, 1.011)	**6.003 *** (0.000, 1.030)	**5.632 *** (0.000, 1.473)
Employment protection	**-5.163 *** (0.000, 1.184)	**-3.047 **** (0.021, 1.310)	**-4.070 *** (0.001, 1.260)	**-3.134 *** (0.087, 1.824)
Intercept (p-value)	14.35 (0.012)	13.84 (0.027)	9.721 (0.091)	25.79 (0.000)
Number of obs.	216	216	212	216
Adjusted R-squared	0.069	0.039	0.019	0.039
F-statistic, p-value	0.000	0.000	0.000	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Random two-ways effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Data: OECD (2021b, 2022 a-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b).

The model (1A) confirms the first hypothesis and shows that the overall generosity of family policies does have a significant effect on the gender wage gap at median. This effect is here estimated to be negative. Being a random effects model, a literal interpretation of this is that the gender wage gap at median on average between units and over time decreases by 0.753 when the family policy generosity increases by one unit. Knowing that the gender wage gap at median ranges from 0.38 to 32.74, the impact – while being statistically significant at the one percent level with a p-value less than 0.000, meaning that there is less than one percent chance that this pattern due to chance – is only around two percent. Furthermore, as the generosity index goes from zero to ten, the total difference from the least to most family policy generosity decreases the wage gap by 7.53 points. Relating it to the intercept means that even in the cases with the most amount of family policy generosity, a gender wage gap remains.

While the effect of the generosity index is statistically significant, I do expect this negative effect to be spurious and hide important differences resulting from the policies that make up the generosity index having different effects. For this reason, in models 1B, C and D, I disaggregate the generosity index into the dimensions of familisation and defamilisation and estimate a separate model for the effect of paid parental leave. For the index to be consistent then, its dimensions must also be significant and/or at least similar in effect. This, it turns out, is only partially the case here.

In the second model (1B), familisation indeed has a positive effect on the median gender wage gap, as expected from the hypothesis (H3) and showing inconsistencies with the family policy generosity index, but the effect is not statistically significant at any accepted level of confidence with a p-value of 0.222. In contrast, the third model (1C) estimates the effect of policies enabling defamilisation on the median wage gap and shows that more policies enabling defamilisation is likely to decrease the median wage gap on average by 0.348. The effect is significant at the one percent level with a p-value of 0.088, which is right at the cut-off point of what is usually accepted but might be related to the small number of observations, and somewhat smaller than the family policy generosity indicator. This model therefore does lend some support to the hypothesis (H2) that more defamilisation negatively affects the gender wage gap.

In the last model (1D), I do not find support for the fourth hypothesis (H4) that there should be a curvilinear effect of the duration of paid parental leave on the median wage gap. The

model shows that the curved effect is as expected with a negative first-degree term and positive second-degree term indicating a U-shaped/convex effect in which the general effect is negative, but the shortest and longest paid parental leaves in contrast increases the wage gap at median. Yet, for this effect to be consistent both terms have to be significant and/or at least show some kind of effect, which they are not. Here, the second-degree term has a high p-value and a virtually non-existent effect. However, this does not mean that paid parental leave does not have any effect on the median wage gap. In fact, the first-degree linear term has a statistically significant negative, though small, effect on the median wage gap indicating that as paid parental leave increases by one unit – here measured in weeks – the wage gap decreases by 0.102.

In addition, there are some interesting effects of the control variables that we are able to estimate because of the random effects modelling. Most prominently, both central bargaining coverage and strictness of employment protection legislation have highly significant and consistent effects on the gender wage gap at median. First, more central bargaining coverage is expected to increase the gender wage gap with around six points. This initially looks like a large effect but taking into account that the scale for central bargaining coverage – going from 0.55 to 1.89 – a one-unit increase will entail much of the same effect as the other variables on the gender wage gap. Second, more strictness of employment protection legislation is likely to decrease the gender wage gap at median. Moreover, educational differences consistently have a negative and significant effect right around the one percent level for all models 1A to D. This means that more educational difference – which here indicates a larger ratio of women than men enrolled in tertiary education – is associated with a decrease in the gender wage gap. Lastly, GDP per capita has a negative effect which is interpreted as more GDP per capita will likely decrease the wage gap.

In sum, when taking all these models into consideration it is clear that disaggregating the family policy generosity indicator has its merits, showing some interesting variations for the family policies' effects on the median wage gap. Overall, however, the family policies do not explain much of the variation in the gender wage gap at median, considering both within and between-variation. This is illustrated by their small effects and further by the small adjusted R-squared that also decreases when disaggregating the generosity indicator.

4.2. Are the effects the same for top and bottom earners?

As seen from the previous models, family policies do have some significant effects on the gender wage gap, but are these effects consistent across the wage distribution? Addressing the second part of the research question, I alternate the dependent variable and estimate the same regression models as above. I start by looking at what affects the wage gap at first earnings decile, then the ninth earnings decile.

4.2.1. The effect of family policies on the wage gap for bottom earners

Table 4.2. presents the regression results from estimating the effect of family policy dimensions on the wage gap at first earnings decile. In the first model (2A), the results are similar to that of the median wage gap in model (1A). The effect is negative and statistically significant with a p-value of 0.002, indicating that more family policy generosity on average also decreases the wage gap at first earnings decile by 0.793. For models 2B and 2C, both familisation and defamilisation are estimated to have a decreasing effect on the gender wage gap at first earnings decile. This, in contrast to models 1B and 1C, does to some degree support the consistency of the generosity index. These effects, however, are not statistically significant and cannot surely be drawn conclusions from. For the fourth model (2D) and similar to the effects on the median wage gap, the effects estimated to show a U-shaped curve in which the effect of paid parental leave first decreases the wage gap, then increases the wage gap again. However, only the former is statistically significant at five percent level and indicates that one more week of paid parental leave decreases the wage gap at first earnings decile by 0.125. I thus do not find support for the estimated curvilinear effect and the linear term's effect is marginal.

What is notable for these models is that the intercepts are much lower than in the models 1 A-D. Interpreting the effect of the family policy generosity index then, for example, indicates that when going from the least to most generosity of family policies, the wage gap at first earnings decile is actually reversed. The intercepts for these models, on the other hand, are not statistically significant. This means that there is not sufficient statistical evidence that the intercept is different from zero – which in terms of values is not far away from the estimated intercepts. This can either be because there are less observations for these models or because the gender wage gap is indeed smaller at first earnings decile.

TABLE 4.2. *Regression results, family policies' effects on the wage gap at first earnings decile.*

Dependent variable: Wage gap at 1st decile	Generosity (2A)	Familisation (2B)	Defamilisation (2C)	Paid leave (2D)
Family policy generosity	**-0.793 ** ** (0.002, 0.258)			
Familisation		-0.058 (0.748, 0.180)		
Defamilisation			-0.382 (0.104, 0.234)	
Paid parental leave				**-0.125 ** ** (0.035, 0.058)
Paid parental leave^2				0.000 (0.125, 0.00)
Controls				
Female employment	**0.182 ** ** (0.013, 0.072)	0.074 (0.398, 0.087)	**0.137 ** ** (0.048, 0.069)	0.084 (0.433, 0.107)
Female public sector empl.	1.459 (0.783, 5.282)	0.419 (0.947, 6.295)	1.682 (0.761, 5.530)	1.758 (0.788, 6.518)
Educational difference	**-0.343 *** ** (0.000, 0.047)	**-0.351 *** ** (0.000, 0.052)	**-0.351 *** ** (0.000, 0.051)	**-0.349 *** ** (0.000, 0.050)
Wage dispersion	**2.972 ** ** (0.004, 1.022)	**2.805 ** ** (0.018, 1.179)	**2.968 ** ** (0.004, 1.023)	**2.341 * ** (0.078, 1.319)
GDP per capita	-0.047 (0.356, 0.051)	-0.027 (0.673, 0.063)	0.000 (0.994, 0.060)	-0.088 (0.186, 0.067)
Central bargaining	**9.680 *** ** (0.000, 1.338)	**9.145 *** ** (0.000, 1.423)	**9.285 *** ** (0.000, 1.421)	**8.630 *** ** (0.000, 1.524)
Employment protection	**-4.572 * ** (0.050, 1.745)	**-3.828 * ** (0.053, 1.961)	**-4.485 * ** (0.055, 1.575)	-3.532 (0.140, 2.385)
Intercept (p-value)	4.421 (0.162)	1.358 (0.189)	6.489 (0.141)	5.846 (0.157)
Number of obs.	187	187	183	187
Adjusted R-squared	0.312	0.138	0.244	0.140
F-statistic, p-value	0.000	0.000	0.000	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Random two-ways effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Data: OECD (2021 a-b, 2022 b-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b)

What is also notable is that the adjusted R-squared increases drastically for these models, which can be interpreted as these variables being better at explaining the variation in the wage gap at first earnings decile than at median. Still, as the effects of the independent variables are both small and insignificant, this gives some indication that the other variables contribute more to this explained variation, such as the control variables.

Especially three controls are highly significant and consistent for all four models: educational difference, wage dispersion and central bargaining. First, educational difference has a statistically significant negative effect at one percent level. This is similar to the effect on the median wage gap and means that the more women are enrolled in tertiary education, the smaller the wage gap is for the first earnings decile, with an average decrease of around 0.350. Second, the effect of economic inequality – here measured by the wage dispersion between first and ninth earnings decile – is estimated to have a positive effect on the gender wage gap. This means that when there is a larger wage dispersion, i.e., a larger distance between first and ninth earnings decile, the wage gap also increases for the first earnings decile between men and women. Finally, central bargaining has a highly significant positive effect on the gender wage gap at first earnings decile. It is therefore likely that more central bargaining coverage, in terms of people covered by a collective agreement and union density, increases the wage gap for the first earnings decile.

4.2.2. The effect of family policies on the wage gap for top earners

On the other side of the wage distribution, how does the family policies affect the wage gap? *Table 4.3.* presents the regression results from estimating the effect of family policy dimensions on the wage gap at ninth earnings decile.

First, in model 3A, family policy generosity has a similar effect on the wage gap at ninth earnings decile as it did at median and first earnings decile. The highly significant effect of family policy generosity is somewhat larger for the ninth earnings decile, but marginally so, indicating an average decrease in the wage gap of 0.729 when family policy generosity increases. However, this finding is questioned by the estimated effects in models 3B and 3C. What is interesting for these models is that in models 3B and 3C the effects on the gender wage gap of the independent variables are statistically significant but contrasting. This further puts the effects and accuracy of the family policy generosity index into question as it consists of indicators with divergent effects when decomposed.

TABLE 4.3. *Regression results, family policies' effects on the wage gap at ninth earnings decile.*

Dependent variable: Wage gap at 9th decile	Generosity (3A)	Familisation (3B)	Defamilisation (3C)	Paid leave (3D)
Family policy generosity	**-0.729 **** (0.048, 0.366)			
Familisation		**0.485 **** (0.002, 0.155)		
Defamilisation			**-0.547 *** (0.071, 0.302)	
Paid parental leave				0.057 (0.270, 0.052)
Paid parental leave^2				-0.000 (0.258, 0.000)
Controls				
Female employment	0.075 (0.516, 0.115)	-0.047 (0.630, 0.099)	0.080 (0.475, 0.111)	-0.019 (0.855, 0.106)
Female public sector empl.	-0.245 (0.763, 8.120)	-2.606 (0.697, 6.673)	-1.221 (0.876, 7.812)	-3.596 (0.657, 8.089)
Educational difference	-0.040 (0.350, 0.042)	**-0.083 *** (0.094, 0.037)	-0.041 (0.325, 0.041)	**-0.070 *** (0.098, 0.042)
Wage dispersion	**-5.924 ***** (0.000, 1.058)	**-5.636 ***** (0.000, 0.971	**-5.884 ***** (0.000, 1.017)	**-5.915 ***** (0.000, 1.131)
GDP per capita	**-0.408 ***** (0.000, 0.081)	**-0.244 ***** (0.001, 0.069)	**-0.376 ***** (0.000, 0.073)	**-0.337 ***** (0.000, 0.086)
Central bargaining	0.563 (0.653, 1.252)	-1.101 (0.262, 0.979)	0.458 (0.702, 1.195)	-0.507 (0.628, 1.045)
Employment protection	**-10.034 ***** (0.000, 1.952)	**-10.408 ***** (0.000, 1.514)	**-10.110 ***** (0.000, 1.858)	**-9.784 ***** (0.000, 1.777)
Intercept (p-value)	65.14 (0.000)	62.70 (0.000)	60.19 (0.000)	64.68 (0.000)
Number of observations	198	198	194	198
Adjusted R-squared	0.057	0.185	0.122	0.032
F-statistic, p-value	0.000	0.000	0.000	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Random two-ways effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Data: OECD (2021 a-b, 2022 b-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b).

In model 3B, the estimated effect of familisation is positive and highly significant. This means that on average, taking into consideration variations both within and between the countries, the family policies enabling familisation do increase the gender wage gap at the ninth earnings decile by 0.458. In contrast, the effect of defamilisation, although being slightly less significant, is negative, which indicates that more of policies enabling defamilisation decreases the wage gap at ninth earnings decile on average by 0.547. These effects are therefore somewhat contradictory. Lastly, I do not find support for a curvilinear effect of paid parental leave on the gender wage gap at ninth earnings decile in model 3D. What is interesting, however, is that the estimated curvilinear effect shows an inverted U-shape. Looking away from the significance then, these estimates show that paid parental leave will first increase the wage gap at ninth earnings decile, then decrease it. This is in contrast to both the theoretical expectations and previous research. Nevertheless, neither the first-degree nor the second-degree term are significant, the estimated effects are minimal and further inference would just be speculating.

All models also show an increase in the adjusted R-squared, indicating that more of the variation in the wage gap at ninth earnings decile can be explained by these variables than the wage gap at median. In addition, it is worth noting that the intercepts of these models drastically increase compared to the previous models. One explanation for this is that there is simply a higher wage gap at the ninth earnings decile, lending some support to the existence of a glass ceiling phenomenon.

Finally, briefly looking to the control variables, there are three control variables that have a statistically significant and consistent effect on the gender wage gap at ninth earnings decile. These are: wage dispersion, GDP per capita and strictness of employment protection legislation. First, the effect of wage dispersion is here negative and can therefore be interpreted as on average actually decreasing the gender wage gap at ninth earnings decile when there is a larger distance between first and ninth earnings decile. This is quite interesting as it indicates that more overall wage inequality actually benefits the relative wage levels of high-earning women. Second, GDP per capita also has a negative effect on the wage gap, but this effect is much smaller. Finally, an increase in the strictness of employment protection legislation, including both strictness of work dismissal and use of temporary contracts, on average between countries and over time decreases the wage gap at ninth earnings decile.

4.3. Comparing the effects across the wage distribution

Based on these results, it is clear that there are different effects of the family policy dimensions on the gender wage gap across the wage distribution. As has been shown, what has an impact on the wage gap at median is not necessarily true for the top or bottom earners. *Figure 4.1.* shows this more clearly in a coefficient plot. The figure shows the coefficient estimates for all independent variables' effects on the three dependent variables. Further addressing the second part of the research question and the last hypotheses then, this section discusses how the effects of the family policy dimensions differ across the wage distribution.

FIGURE 4.1. *Coefficient plot, independent variables*
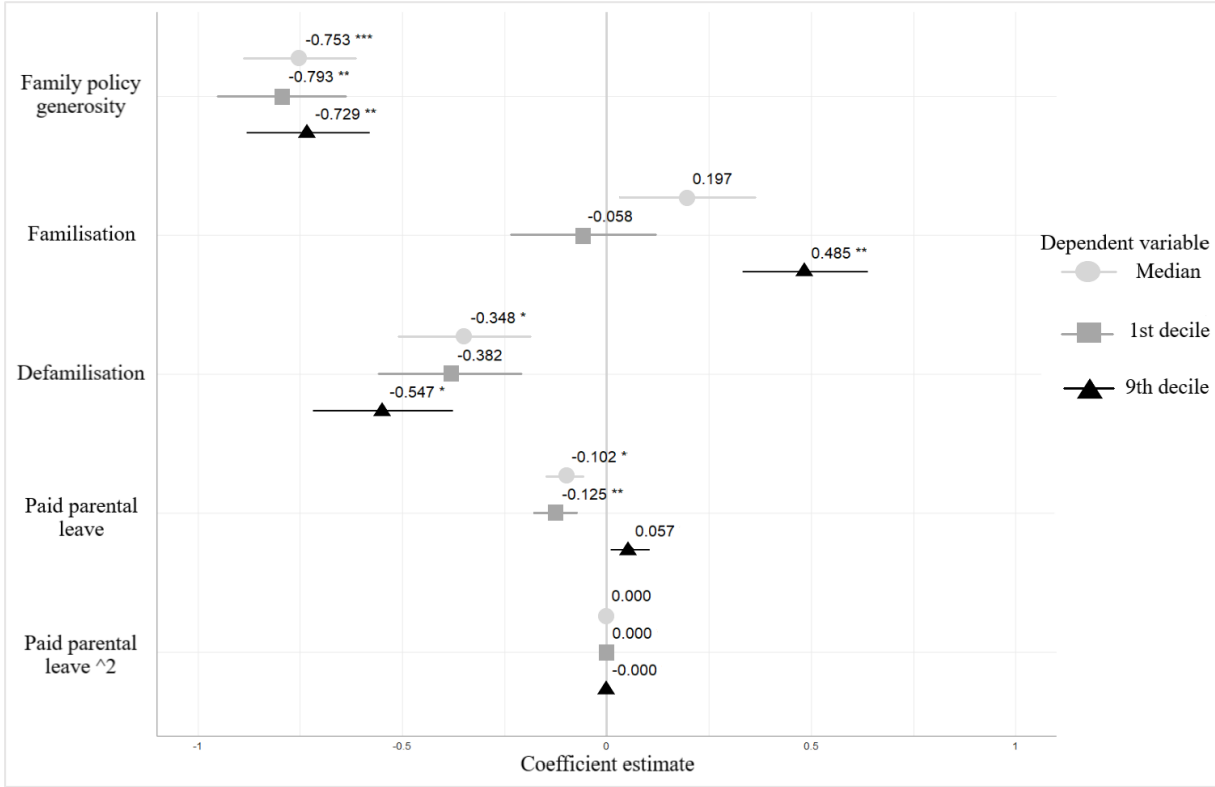
Notes: Coefficient plot showing the coefficient estimates and their 95 percent confidence interval. Visualisation based on R-packages *sjPlot* (Lüdecke 2021) and *ggplot2* (Wickham 2016). Data: OECD (2021a, 2022 a-b), Gauthier (2011), Eurostat (2017, 2022).

First, the plot shows that overall family policy generosity does have a significant effect on the gender wage gap at median, first and ninth earnings decile. The effect is somewhat stronger for the gender wage gap at first earnings decile, but marginally so. Nonetheless, although the overall family policy generosity has a similar and significant effect on all dependent variables,

Figure 4.1. graphically shows that this does not give a good picture of what is actually going on. When disaggregating this indicator into the dimensions of familisation and defamilisation – all building on the same variables – the differences as to what drives the wage gap is notable. This illustrates that the family policy generosity index hides important differences that result from the policies having different effects, thus emphasising the value of disaggregating the overall family policy generosity into its constituent dimensions.

Furthermore, *Figure 4.1.* also shows the importance of looking at effects across the wage distribution in order to understand what really drives the wage gap, here illustrated by the differences of the coefficient estimates for the same variables between the first and ninth earnings decile. This difference is most striking for the dimension of familisation. For the first earnings decile, the effect of familisation is estimated to be negative but with both a high p-value and standard error. Familisation, on the other hand, does have a highly significant and relatively strong effect on the wage gap for top earners. This shows that for one side of the wage distribution, these policies do not have any effect on the gender wage gap, but for the other side this dimension is likely to increase the gender wage gap. Thus, the fifth hypothesis (H5) stating that the positive effects of policies enabling familisation should be stronger for the top earners is only partially supported. There is a positive and significant effect of familisation on the gender wage gap at ninth earnings decile, but I also expected this effect to be positive for the gender wage gap at first earnings decile, which it is not.

Moreover, although the figure shows that for the defamilisation indicator, the effects differ for the different points of the wage distribution, I do not find support for the sixth hypothesis (H6) that the negative effect of policies enabling defamilisation should be stronger for the gender wage gap at first earnings decile than at ninth earnings decile. On the contrary, defamilisation only has a statistically significant negative effect on the wage gap at ninth earnings decile, not the first. This indicates that with more defamilising policies, the wage gap decreases for the top earners, while the wage gap for the bottom earners remains unaffected.

I similarly do not find support for the expected curvilinear U-shaped effect of the duration of paid parental leave on either measure of the gender wage gap. This is consistent for the whole wage distribution, and in contrast to previous research, as the second-degree terms are not significant or have any noticeable effect in either model. There is nevertheless some statistically significant negative effect of paid parental leave on the gender wage gap at first earnings decile, indicating that a longer paid parental leave decrease the wage gap for those

earning the least. But what is interesting, as already mentioned, is that for the ninth earnings decile – although not statistically significant – the curvilinear effect is estimated to be an inverted U-shape, meaning that only the least and most amount of paid parental leave decreases the wage gap for the top earners. It is, of course, uncertain whether this is due to chance or is consistent with the underlying population, but it is nonetheless an interesting potential effect.

Besides the independent variables, there are also some interesting patterns in the control variables when comparing the models to each other, as presented in *Figure 4.2.* showing a coefficient plot of the control variables for models 2C and 3C[11].

FIGURE 4.2. *Coefficient plot, control variables*
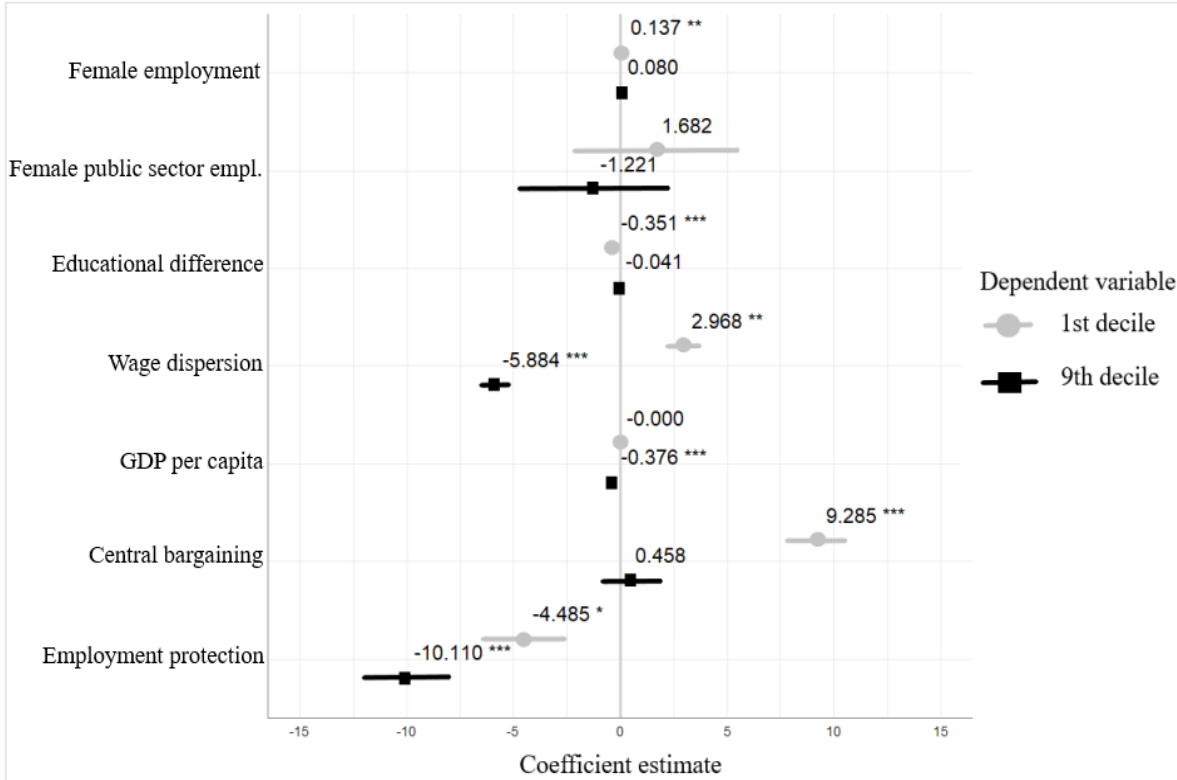
Notes: Coefficient plot showing the coefficient estimates and their 95 percent confidence interval from models estimating the effect of defamilisation on the wage gap at first and ninth earnings decile. Visualisation based on R-packages *sjPlot* (Lüdecke 2021) and *ggplot2* (Wickham 2016). Data: OECD (2021 b-c, 2022 c-f), World Bank (2021, 2022 a-b).

---

[11] I only include these models in the coefficient plot as a graphic illustration of the variations in the control variables across the wage distribution for two reasons: (1) these are the best models in terms or adjusted R-squared (0.244 and 0.122 respectively) and (2) the control variables are stable and consistent for all models 2 and models 3, so these are representative for the other models as well.

This figure clearly shows that what affects the wage gap at one end of the distribution is not necessarily what will affect the wage gap for the other end, and this is also the case with the control variables.

For educational difference, for example, there is a consistent negative effect for all estimated models, but when decomposing the wage distribution this effect is only significant for the gender wage gap at first earnings decile and not at the ninth earnings decile. This means that more educational equality is only relevant and has a negative effect on the wage gap for the bottom earners but does not make a difference for those earning the most. In addition, wage dispersion measuring economic inequality shows similar patterns as that of education. The effects of this variable are completely different for the gender wage gap at first and ninth earnings decile, where for the first earnings decile more wage inequality is likely to increase the wage gap, as opposed to the ninth earnings decile where more wage inequality instead decreases the gender wage gap. GDP per capita as well is only significant and negative for the ninth earnings decile, indicating that the top earners' gender wage gap is more sensitive to the overall economic situation of the country. Similar patterns are also found for the last two variables. First, central bargaining coverage only has a significant effect for the first earnings decile, indicating that more central bargaining coverage positively affects the gender wage gap for bottom earners but does not have any effect on the gender wage gap for top earners. Contrary to this variable, strictness of employment legislation is revealed to have a strong and significant effect only on the gender wage gap at ninth earnings decile.

In sum, disaggregating the family policy generosity indicator and looking at different places at the wage distribution reveal some interesting differences and is an important distinction to make in order to better understand what drives the gender wage gap. This is further supported by the adjusted R-squared increasing drastically and the residuals being closer to normally distributed for models 2 A-B and 3 A-B estimating the effect on the wage gap at first and ninth earnings decile compared to models 1 A-D estimating the effect on the median wage gap.

4.4. Assessing the robustness of the results

In order to assess the robustness of the results, I test whether the results are sensitive to changes in the parameters and to changes in the modelling techniques. First, I test whether the results are sensitive to the conceptual specification of the independent variables, then whether extreme outliers substantially influence the results. Finally, I test the robustness of the overall model design by estimating fixed effects models.

4.4.1. Alternative independent variable specifications

First, I test whether the results are sensitive to changes in the concept formation of the independent variables. As Kleider (2015) argued, one of the main obstacles to cross-national empirical work has been the operationalisation of family policy dimensions, and changes in the operationalisation can have major consequences for the results of the empirical analysis (508). Therefore, further following Kleider's (2015) argument, I test the sensitivity of the operationalisation by adding a measure of the total public spending on tax breaks and alterations to the family policy generosity index as well as to the familisation index. As discussed, certain tax rules – especially co-taxation of spouses – can be argued to be important family policy indicators but were not included in the main model as they are closely related to anti-poverty measures and therefore do not exclusively intend to aid the work-family balance for parents and potential parents (Daly 2020, 36; Kleider 2015, 510; Lohmann and Zagel 2016, 61). In order to test the sensitivity of the defamilisation index as well, I add the measure for childcare for children aged 3 and up to school age, which were initially excluded as this is more related to children's education than to women's employment (Budig, Misra and Boeckmann 2016, 159; Vandenbroeck 2020, 169).

Re-estimating the models with these variables is therefore a good sensitivity test to see whether the extended conceptualisation of all indices gives the same results. Looking at the overall results, these changes do not alter the overall results from the main analysis but does further decrease the number of observations and by consequence increase the estimated standard errors.

4.4.2. Are there any influential outliers?

Second, I test whether the results are influenced by extreme outliers, i.e., values of variables that are extremely high or low relative to the result of the values of that variable (Kellstedt and Whitten 2018, 258). Although all dependent variables are mostly normally distributed, there are some outliers which I remove and then re-estimate the models as above. This does not change the overall results. Removing the outliers does increase the adjusted R-squared somewhat, but only marginally. It does not change the overall results either, but the coefficient estimates decline somewhat. Here, the most dramatic decrease is for central bargaining coverage in the re-estimated model 2B estimating the effect of familisation on the wage gap at first earnings decile, for which the coefficient decreases from 9.145 to 8.444. Most drastically, for the models 3 A-D removing the outliers decreases the intercepts by 10 to around 55 for all models. They are still statistically significant, but this indicates that the large intercepts are somewhat influenced by the outlying values in the gender wage gap at ninth earnings decile, most notably Greece in 2006 with a gender wage gap at ninth earnings decile of 36.7. Overall, however, removing the outliers does not change the results, which implies that while these are outliers in the univariate sense, they are not influential cases in the regression models. This is not surprising considering the mostly normally distributed observations for each variable.

4.4.3. Estimating fixed effects models

Finally, I check the robustness of the model design by estimating the fixed effects model alternative. As already discussed, FE only model within-variation but can be used as a robustness to the RE models and show how effects are when looking at changes over time. If the models are significantly better, it can be an indication of processes within the countries. Also, as mentioned in the discussion on random and fixed effects, the Hausman test was somewhat undecidable. For all models estimating the effects on the gender wage gap at first earnings decile, however, the Hausman test was consistently significant indicating that the random effects model is inconsistent and thus preferring the FE model. The results from estimating the family policy dimensions' effects on the first earnings decile using fixed effects modelling are therefore presented in *Table 4.4.*, for which I expect the most differences between the RE and FE models.

TABLE 4.4. *Regression results using fixed effects, dependent variable: first earnings decile.*

Dependent variable: 1st earnings decile	Generosity (4A)	Familisation (4B)	Defamilisation (4C)	Paid leave (4D)
Family policy generosity	**-0.619 ***** (0.000, 0.165)			
Familisation		-0.288 (0.330, 0.295)		
Defamilisation			-0.082 (0.760, 0.268)	
Paid parental leave				-0.161 (0.360, 0.176)
Paid parental leave^2				0.000 (0.579, 0.001)
Controls				
Female employment	0.308 (0.104, 0.189)	0.210 (0.334, 0.217)	0.262 (0.210, 0.208)	0.133 (0.564, 0.230)
Female public sector empl.	4.246 (0.314, 4.206)	2.574 (0.579, 4.628)	3.786 (0.442, 4.915)	4.973 (0.307, 4.852)
Educational difference	**-0.240 **** (0.012, 0.094)	**-0.257 **** (0.041, 0.125)	**-0.254 **** (0.017, 0.106)	**-0.267 **** (0.019, 0.113)
Wage dispersion	**3.075 **** (0.036, 1.453)	**2.679 *** (0.085, 1.547)	**2.874 **** (0.050, 1.452)	2.467 (0.140, 1.664)
GDP per capita	**-0.584 **** (0.002, 0.188)	**-0.617 **** (0.001, 0.185)	**-0.591 **** (0.003, 0.198)	**-0.584 **** (0.003, 0.191)
Central bargaining	**11.156 ***** (0.000, 2.553)	**10.616 ***** (0.000, 2.767)	**10.696 ***** (0.000, 2.733)	**9.116 **** (0.001, 2.783)
Employment protection	1.520 (0.802, 6.062)	4.845 (0.487, 6.950)	3.997 (0.588, 7.366)	3.129 (0.656, 7.017)
Hausman test, p-value	0.000 (RE < FE)	0.000 (RE < FE)	0.000 (RE < FE)	0.000 (RE <FE)
Number of obs.	187	187	183	187
Adjusted R-squared	0.444	0.407	0.383	0.421
F-statistic, p-value	0.000	0.000	0.000	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Fixed effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Includes results from the Hausman test indicating which model is preferred of RE and FE. Data: OECD (2021 a-b, 2022 b-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b).

For these models (4 A-D), *l*ooking only at within-variation, the overall effect of the institutional variables does not change considerably, but the standard error increases and the effects loses all significance for all but family policy generosity. In model 6A, the family policy generosity is still negative with a regression coefficient of -0.619 and a significance at one percent level with p-value less than 0.000. A literal interpretation of this result is that with a one unit increase in family policy generosity, the gender wage gap at first earnings decile decreases by 0.619 on average per country. In models 4B, 4C and 4D, however, none of the independent variables have a significant within-effect. Most of the other control variables do not change their effect drastically in these models either. What does change, however, is the effects of GDP per capita. This effect was not significant for the random effects models but is now both negative consistently and significant for the corresponding fixed effects models. This indicates that while GDP per capita does not have an effect taking into account both within- and between variation, it does negatively affect the gender wage gap at first earning decile within countries. Another interpretation is simply that GDP per capita is one of the only variables left with considerable within-variation and that its effect is therefore overestimated. The opposite happens to the strictness of employment protection legislation indicator, for which the effects are no longer significant within countries.

These results are somewhat similar when estimating the same models with fixed effects for the median gender wage gap and the gender wage gap at ninth earnings decile, presented in *Table B.1. and Table B.2.* in Appendix B. For these models, the effect of the independent variables does not change notably, but all effects besides the family policy generosity index lose all significance. This again can either be because the independent variables do not have any effect within the countries over time, or because these stable effects have been effectively controlled out. The only significant variable for the models estimating the effect on the median wage gap using fixed effects (*Table C.1.*) is central bargaining coverage, which has a positive effect. This is consistent with the corresponding models estimated using random effects. For the wage gap at ninth earnings decile (*Table C.2.*) only GDP per capita is consistent in effect and significance.

These results are not surprising because, as FE models control out the between variation, there is not much variation left in the independent variables (Clark and Linzer 2015, 402). Consequently, fixed effects estimates often have higher standard errors and less significant results, which can either be because there are actually no effects of these policies within the countries or that all variation has been controlled out (Allison 2009, 3). On the other hand, the

adjusted R-squared for models 4A to D drastically increases, which is unexpected for fixed effects models as the country dummies decrease the degrees of freedom. This, in addition to the results provided by the Hausman test, gives some indication that what might be important in determining the wage gap at first earnings decile are processes within the countries, while what is important for the gender wage gap at median and ninth earnings decile are the processes that vary between countries.

5. DISCUSSION AND CONCLUSION

In this last chapter, I begin by giving a short summary of what we know so far in terms of theoretical expectations and the findings. I then move on to discussing the findings and their implications for both theory and practice and give some suggestions for future research. I finish with some concluding remarks.

5.1. What do we know so far?

So, how do family policies intended to help balance work and family obligations affect the gender wage gap, and are the effects different throughout the wage distribution? From the theory, I expected first that while family policy generosity should affect the gender wage gap, I expected this to be misleading in terms of the policies' actual effects. Thus, building on theory arguing in favour of separating the effects of policies enabling defamilisation, i.e., facilitating work, and the policies enabling familisation, i.e., reducing work, I expected the former to decrease the gender wage gap and the latter to increase it. Since paid parental leave can be both defamilising and familising, I expected it to have a U-shaped effect on the gender wage gap where the least and the most leave increase the wage gap. Furthermore, I expected these effects to differ across the wage distribution. First, because statistical discrimination is arguably more prominent in high-wage jobs, I expected familisation to have a stronger positive effect on the wage gap for top earners. Next, because defamilising policies to a larger degree bring disadvantaged women into paid work, I expect the negative effects of these policies to be stronger for the wage gap for bottom earners.

In the results, I find empirical support for the disaggregation of family policies into the dimensions of familisation and defamilisation. Defamilisation is found to have a negative effect on the gender wage gap, supporting the second hypothesis, but familisation does not have a significant effect, thus not supporting the third hypothesis. In addition, I do not find support the fourth hypothesis expecting a curvilinear U-shaped effect of the duration of paid parental leave on the gender wage gap, neither at median, first nor ninth earnings decile. Furthermore, addressing the potential different effects across the wage distribution, I find first that familisation has a positive and significant effect only for the gender wage gap at the ninth earnings decile, lending some support to the fifth hypothesis. Second, contrary to what was expected in the last hypothesis, I find that the effect of defamilisation is only negative and

significant for the gender wage gap at the ninth earnings decile, not the gender wage gap at the first earnings decile.

5.2. Implications for theory and practice

The findings have implications going beyond this thesis, providing support for some theoretical explanations, and giving some indication to what kinds of policies to pursue in order to decrease differences in wages between men and women.

5.2.1. Implications for theory

Starting with the first hypothesis, I find a negative effect of family policy generosity on the median gender wage gap. The expected effect was ambiguous *a priori*, but the estimated negative effect lends some support to the early arguments made by Gornick, Meyers and Ross (1998) and questions the trade-off argument presented by Mandel and Semyonov (2005, 2006). However, as already mentioned, this does not give a good picture of what is actually going on. Because the estimated effects of familisation and defamilisation on the median gender wage gap are different, I find support for the disaggregation of the concept, as proposed in the literature. This further shows that disentangling these policies is crucial for understanding the relationship between women's labour force outcomes and specific policy conditions. Or, as summarised by Budig, Misra and Boeckmann (2016): "It is not that generous welfare states create trade-offs regarding gender equality but particular policies […] that do so" (161).

Looking at the different policy dimensions, I first find that policies enabling familisation has the expected positive effect on the gender wage gap at ninth earnings decile. This supports the presumption that institutionalised workforce interruptions are worse for women's relative wages at the top of the wage distribution. One potential explanation is that gender discrimination, or expected discrimination, is more prevalent in top positions, thereby making highly qualified women fall further behind men at the top of the wage distribution or giving incentives to self-select out of the top earning jobs. Familisation, on the other hand, does not have any effect on the gender wage gap at first earnings decile. This might indicate that there are no adverse effects of policies enabling women to stay at home with care responsibilities on women's relative wages at the bottom of the wage distribution.

73

Furthermore, and in contrast to what I expected, defamilisation only has a negative effect on the gender wage gap at ninth earnings decile and no significant effect on the gender wage gap at first earnings decile. First, the negative effect on the gender wage gap at the top of the wage distribution further supports the interpretation that when women's workforce attachment and continuity is supported by the policy setting, they further catch up with men in terms of relative wages. This shows that, as already mentioned in the theory section, it is not the qualifications in high-paying jobs that matter *per se*, but the assumed productivity and relative performance (Datta Gupta, Oaxaca and Smith 2006, 262). This is further supported by the effect of educational equality not being a significant factor in explaining the gender wage gap for top earners. In addition, the importance of workforce attachment and continuity also lends support to Goldin's (2014) and Cha and Weeden's (2014) more recent theories about how the structure of high-paying jobs themselves prefer flexibility and the possibility to work long hours, and how those who do are disproportionately rewarded.

Second, policies enabling defamilisation do not have an effect on the gender wage gap at the bottom of the distribution, thus questioning the presumption that by integrating women into paid labour and increasing their labour force attachment at the lower end of the wage distribution, their wage levels would increase. One potential explanation could be that while defamilising policies get women into the workforce, this does not necessarily mean into full-time employment. Women, and especially mothers, are indeed more likely to work part-time, especially in lower paying public sector jobs that tolerate workforce interruptions better, either because of restricted opportunities of by self-selection (Budig, Misra and Boeckmann 2016, 140). Therefore, building on the argument by Mandel and Semyonov (2005, 2006), inclusion into the labour force not only leads to occupational segregation in which women cluster into these kinds of jobs, but also a kind of occupational segregation in which women more often work part-time. As the dependent variables used here are only based on wages of full-time employees, I am not able to take into account this potential effect on part-time employees which could explain the lack of significant effect of defamilisation on the gender wage gap at the first earnings decile.

Furthermore, I do not find support for a curvilinear effect of the duration of paid parental leave on the gender wage gap at any point in the wage distribution, which stands in contrast to previous research. In fact, I do not find much effect of this policy at all, only a small negative effect on the gender wage gap at first earnings decile. This is somewhat puzzling as this effect has been a consistent finding in the literature since the early study by Ruhm (1998) first

identifying this effect. The lack of effect is also puzzling because of the centrality of leave schemes both in theory and in practice. Maternity leave was indeed the first family policy, designed first to promote the health of women and children close to childbirth but later became fundamental in terms of balancing paid work and family responsibilities (Booth 2007, 132; Ruhm 1998, 291). The lack of finding could therefore either be because paid parental leave simply is not important in shaping the gender wage gap anymore, or because it was not properly assessed by the measure used here, looking only at the duration in weeks. As already discussed, while most countries have some kind of entitlement to parental leaves, there is substantial variation between the countries not only in terms of length and whether or not it is paid, but also in terms of *how* much it is paid, the flexibility of who receives it and other criteria regarding who can take leaves from paid work (Kamerman and Moss 2009, 3; Pettit and Hook 2005, 783). Therefore, it might be that this measure simply does not capture the different conditions under which we might expect this U-shaped effect, or – as argued above – that the curvilinear effect is more prominent among part-time employees, who have not been accounted for here.

Finally, these findings may contribute to explain some of the contrasting and puzzling findings regarding the gender wage gap, such as the stagnation in the wage gap conversion and the persistent gender wage gap often identified in the Nordic countries (Hook and Li 2020, 263). First, as family policies – both familising and defamilising – are estimated to have a consistent negative effect on the gender wage gap for bottom earners, the gender wage gap at first earnings decile could be argued to have experienced a decline together with the development of such policies. However, the contrasting effects of family policies – with familisation policies having a strong positive effect and defamilising policies an equally strong negative effect – on the gender wage gap at ninth earnings decile may have contributed to the gender wage gap staying more stable for the top earners or even increased in some cases. Therefore, the difference in the development of the gender wage gap at different ends of the wage distribution may partially explain, as also observed by Boye, Halldén and Magnusson (2017), why the gender wage gap "at the aggregate level has not decreased in the past 30 years even though the proportion of women in [employment] has increased" (164).

In sum, all these findings taken together emphasise the importance of looking at contextual explanations in order to understand what shapes the gender wage gap and show that family policies do contribute to shaping the gender wage gap in important ways. The greater awareness of the gendered consequences of family policies could also benefit comparative

research looking into the consequences of social policies more generally (Estévez-Abe 2006, 174). In addition, these findings further support arguments in favour of looking at various points of the wage distribution and not only at the average.

5.2.2. Implications for practice

As the results have shown, there are complex and sometimes contradictory effects of the family policies on the gender wage gap at different points in the wage distribution. Therefore, as Booth (2007) somewhat pessimistically concluded: "There is no unique policy solution to the problems of the gender pay gap" (141). Nevertheless, it is clear from both theory and the results that the policies that help strengthen women's labour force attachment by reducing the cost of entering and staying in paid work, i.e., policies enabling defamilisation, could be most effective at reducing wage inequalities (Budig, Misra and Boeckmann 2016, 120; Estévez-Abe 2005, 192). Considering the indicators used here, this entails either strengthening the availability of childcare arrangements for the youngest children and/or further supporting fathers' involvement in care responsibilities through offering parental leave reserved for the exclusive use of fathers.

First, when it comes to childcare arrangements, what seems most important is that childcare is publicly subsidised, high-quality, and highly accessible. This is argued to not only have positive effects on women's relative earnings, but also on children's development of social skills and early education (Vandenbroeck 2020, 170). Second, since countries are still developing paternity leave schemes, what seems most important is simply offering paid paternity leave. This has been argued to be beneficial for women's relative earnings by enabling women to return to work earlier, but also in being a long-term investment in changing attitudes towards both parents working and sharing care responsibilities (Cools, Fiva and Kirkebøen 2015, 825). This, in turn, is argued to level out some of the statistical discrimination facing women when normalising – and in turn, not heavily penalising – workforce interruptions for all genders (Cools, Fiva and Kirkebøen 2015, 825; Padavic, Ely and Rider 2020, 103). Having a reserved paternity leave has also been associated with other positive outcomes, such as improving children's school performance – especially in cases where the fathers are highly educated (Cools, Fiva and Kirkebøen 2015, 802), and facilitating bonding with children which strengthens their cognitive development from early ages (Rege and Solli 2013, 2256).

Second, because policies enabling familisation, here measured with indicators indicating long paid parental leave over 60 weeks or the duration of unpaid leave, are likely to decrease women's chances of catching up to male wages at the top of the wage distribution, limiting these policies would likely further increase gender equality among top earners. However, this raises the normative question of whether these policies *should* be limited as a means of securing gender equality in earnings amongst top earners. This is because taking away some of the policy options for women could in turn negatively affect those in need of longer leaves, which often coincides with already disadvantaged women not having other options. Or, in another line of argument, this could limit the options available for women who for some reason either want to or need to stay at home with their children longer than a few months. This could for example be if the mother is still breastfeeding, experience health related issues post-birth or that the children have health related difficulties making them unable to attend publicly funded childcare from the earliest years. Another related normative question further asks how much the state should be able to regulate one's personal life in the private sphere. Therefore, as noted by Korpi, Ferrarini and Englund (2013), the tension between class inequality and gender inequality:

"Bring up key political and normative questions: is gender inequality to be counteracted over the whole range of social classes? Should the concerns of well [paid] women be secondary to those of the less-advantaged women? Should gender inequality at the top be decreased even at the expense of increasing socio-economic inequalities among other women?" (30).

The observation that the effect of overall economic inequality affects the gender wage gap differently at first and ninth earnings decile further leads me to pose the same questions. As shown in the results, the effect of economic inequality indicates that for the bottom earners, more wage inequality is likely to increase the gender wage gap. This could be explained as by Mandel and Semyonov (2005) arguing that "any reduction in the dispersion of the earnings distribution benefits workers in its lower tail, most of whom are women" (953). In contrast, the results show that more economic inequality is likely to decrease the gender wage gap at the ninth earnings decile, which is more puzzling and potentially something to explore for future research. Yet, in terms of practical policy implications, compressing the wage structure and creating more overall wage equality leads to more gender inequality in earnings for the top earners. This again poses the normative question of whether economic equality at large should be sacrificed in order to further decrease the gender wage gap for the highest earners.

In sum, the findings reflect a diversity of outcomes for family policies, suggesting that they are not either entirely good or bad (Budig, Misra and Boeckmann 2016, 120). While this does not offer a clear solution to gender inequalities in earnings and instead poses several normative questions, it is important to draw attention to the unintended consequences of family policies so that we can better understand how the policies might affect women's labour market outcomes and, in turn, continue to develop institutional arrangements that not only support women's employment, but also assist them in successfully competing with men for top positions.

5.3. Suggestions for future research

Finally, although the thesis does contribute to explain some of the variation in the gender wage gap, there remains variation in the gender wage gap for which these results alone cannot account for. Based on the results and the drawbacks of the thesis, there are some suggestions for future research that should be undertaken in order to further understand how family policies shape the gender wage gap and how this differs across the wage distribution.

First and foremost, future research should keep in mind the disaggregation of the concept of *family policy* and look at differences across the wage distribution. As seen here, there are important patterns to be discovered when not looking at average trends. Also, as mentioned above, future research should consider how part-time employment fits into this explanation. For example, are the factors shaping the gender wage gap for full-time employees the same as what shapes the gender wage gap for part-time employees? As Budig, Misra and Boeckmann (2012) argue: Employed women are not a random subgroup of all women and estimating the determinants of earnings from the subpopulation that are employed may introduce bias and not fully account for how policies actually affect wage inequalities (176).

There is, of course, always a question of data availability, but these considerations should be made whenever possible. Data availability and quality is also a general limitation to this thesis, and this line of research in general, and most comprehensive databases taking into account the specificities of the various family policies are old and somewhat outdated. Therefore, future research would also need more and better data with better coverage and in which different criteria and characteristics of these policies are taken into account. For

example, are childcare arrangements universal, or can parental leave be shared freely between the parents? And what effect would these characteristics entail?

Furthermore, future research should also consider how these policies interact with other policies. As Keck and Saraceno (2013) writes: "Policies are not a single, straightforward indicator [and] have to be understood as a part of a national setting" (318). Thus, since policies are not shaped in a political vacuum, they should not be analysed in isolation (Kamerman and Moss 2009, 10). While I have here analysed several policies together, I have not studied how they might interact with each other. It would therefore be interesting for future research to look into how these policies interact and shape each other, instead of viewing the policies as packages.

It would also be interesting for future research to look at potential interactions between family policies and the wage-setting structures, such as central bargaining coverage and employment protection legislation, which were here only applied as control variables. Most literature explaining the gender wage gap by contextual factors focuses either on family policies or the institutional setting, but rarely combine the two dimensions. Yet, some researchers have hypothesised some interesting interactions. White (2020), for example, observes that "coordinated market economies tend to provide more extensively subsidised childcare for families, in contrast to liberal market economies" (79). Grönlund and Magnusson (2016), in another example, argue that "CME institutions produce occupational and vertical gender segregation and exacerbate gender differences in skill and wages, and by institutionalising work interruptions, generous family policies further aggravate the situation" (95). These are patterns that would be interesting for future research to look further into.

Moreover, another drawback of this thesis has been only looking at national policies. With the historical dominance of the nation-state as main policymaker, this has led most research to focus on the national level (Betcher and Meardi 2012, 187; Nieuwenhuis and Van Lancker 2020, 4). However, several societal developments point towards an increased importance in the multidimensionality of family policies. For example, several organisations and firms now offer family policies such as childcare or leave schemes, regionalisation increasingly transfers policy responsibility towards regions, and supranational and international organisations increasingly issue recommendations or directives in the area of social policy (Nieuwenhuis and Van Lancker 2020, 5). It would therefore be interesting to study this issue from a different level, especially the organisational firm-level because we already know that women

are more likely to work in the public sector and be occupationally segregated (Albrecht, Björklund and Vroman 2003, 163). Thus, if there are any pattern in the types of policies the firms offer in different sectors, what kind of further inequalities would that entail? For example, if only highly paid private occupations provide childcare arrangements to their workers and there are no nationally provided alternatives, how would this affect the gender wage gap across the wage distribution? These are interesting questions for future research.

Finally, as mentioned in the data section, the gender definitions as well as definitions for mothers and fathers are based on strict – and increasingly challenged – legal definitions within countries. It would therefore be interesting to approach the research on the gender wage gap with a more fluid gender definition and see if the same patterns hold when the definition is extended. Because questions of gender fluidity are relatively recent, databases – especially the older ones – do not yet reflect this non-binarity of how we increasingly view gender. Those not fitting the binary definition of gender is therefore a highly understudied group (Davidson 2016, 2). In addition, the family policies also need to be seen in terms of changes in the family structures. One of the most important challenges for family policies today are indeed the changing structures of modern-day families with a growing number of single-parent households and same-sex parents, increasing divorce rates and declining fertility (Adema, Clarke and Thévenon 2020, 195), and how family policies accommodate these changes – and potentially change themselves – is an interesting topic for future research.

5.4. Concluding remarks

Summing it all up, based on the theory and results presented, it is clear that family policies that are intended to help balance work and family obligations play an important part in shaping the gender wage gap. The findings also show that these policies have different effects on the wage gaps across the wage distribution. Overall, these findings are consistent with a large and still growing body of research documenting the importance of family policies for women's relative wages, as well as how these vary by women's position in the wage distribution. Thus, there is no unique solution to deal with the gender wage gap, and in the most extreme cases, if what decreases the wage gap for the top earners is what increases the wage gap for bottom earners, this raises a set of normative questions regarding which inequalities we accept and which ones we do not.

Finally, what is most important is that by focusing only on the average trends, may it be family policies or wage gaps, we lose out on important information on what drives the gender wage gap. As a conclusion, it is important to pay attention to the different effects of family policies for the wage gap across the wage distribution in order to better understand what shapes the gender wage gap.

REFERENCES

Adema, Willem, Chris Clarke and Olivier Thévenon. 2020. "Family Policies and Family Outcomes in OECD Countries." *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker, 193-217. Chum: Palgrave Macmillan.

Albrecht, James, Anders Björklund and Susan Vroman. 2003. "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics* 21 (1): 145-177.

Alexander, Marcus, Matthew Harding and Carlos Lamarche. 2011. "Quantile Regression for Time-Series-Cross-Section Data." *International Journal of Statistics and Management System* 6 (1-2): 47-72.

Allison, Paul D. 2009. *Fixed Effects Regression Models.* Los Angeles: SAGE.

Arulampalam, Wiji, Alison L. Booth and Mark L. Bryan. 2007. "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution." *Industrial & labor relations review* 60 (2): 163-186.

Azmat, Ghazala and Barbara Petrongolo. 2014. "Gender and the labor market: What have we learned from field and lab experiments?" *Labour Economics* 30: 32-40.

Baltagi, Badi H. 2021. *Econometric Analysis of Panel Data*. 6th edition. Switzerland: Springer.

Beck, Nathaniel and Jonathan N. Katz. 1995. "What to do (and not to do) with Time-Series Cross-Section Data." *The American Political Science Review* 89 (3): 643-647.

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* 3rd edition. Chicago: The University of Chicago Press.

Bell, Andrew and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3 (1): 133-153.

Bell, Andrew, Malcolm Fairbrother and Kelvyn Jones. 2019. "Fixed and random effects models: making an informed decision." *Quality & Quantity* 53: 1051-1074.

Betcher, Barbara and Gulielmo Meardi. 2012. "Sectors or countries? Typologies and levels of analysis in comparative industrial relations." *European Journal of Industrial Relations* 18 (3): 185-202.

Blau, Francine D. and Lawrence M. Kahn. 2003. "Understanding International Differences in the Gender Pay Gap." *Journal of Labor Economics* 21 (1): 107-144.

Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789-865.

Bonjour, Dorothe and Michael Gerfin. 2001. "The unequal distribution of unequal pay – An empirical analysis of the gender wage gap in Switzerland." *Empirical Economics* 26: 407-427.

Booth, Alison L. 2007. "The Glass Ceiling in Europe: Why Are Women Doing Badly in the Labour Market?" *Swedish Economic Policy Review* 14: 121-144.

Boye, Katarina, Karin Halldén and Charlotta Magnusson. 2017. "Stagnation only on the surface? The implications of skill and family responsibilities for the gender wage gap in Sweden, 1974-2010." *The British Journal of Sociology* 68 (4): 595-619.

Brandt, Berit and Elin Kvande. 2009. "Norway: the making of the father's quota." in *The politics of parental leave policies. Children, parenting, gender and the labour market.* Edited by Sheila B. Kamerman and Peter Moss, 191-206. Great Britain: The Policy Press.

Budig, Michelle J. Joya Misra and Irene Boeckmann. 2016. *"*Work-family policy trade-offs for mothers? Unpacking the cross-national variation in motherhood earnings penalties.*" Work and Occupations* 43 (2): 119-177.

Budig, Michelle, J. Joya Misra and Irene Boeckmann. 2012. "The motherhood penalty in cross-national perspective: The importance of work-family policies and cultural attitudes." *Social Politics* 19 (2): 163-193.

Carrillo, Paul, Néstor Gandelman and Virginia Robano. 2014. "Sticky floors and glass ceilings in Latin America." *The Journal of Economic Inequality* 12: 339-361.

Cha, Youngjoo and Kim A. Weeden. 2014. "Overwork and the Slow Convergence in the Gender Gap in Wages." *American Sociological Review* 79 (3): 457-484.

Charles, Maria and David B. Grusky. 2004. *Occupational Ghettos. The Worldwide Segregation of Women and Men*. Standford: Standford University Press.

Christofides, Louis N., Alexandros Polycarpou and Konstantinos Vrachimis. 2013. "Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe." *Labour Economics* 21: 86-102.

Clark, Tom S. and Drew A. Linzer. 2015. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods* 3 (2): 399-408.

Cools, Sara, Jon H. Fiva and Lars J. Kirkebøen. 2015. "Causal Effects of Paternity Leave on Children and Parents." *Journal of Economics* 117 (3): 801-828.

Croissant, Y. and Millo, G. 2018. *plm: Panel Data Econometrics with R.* R-package, version 2.6-1. URL: https://cran.r-project.org/web/packages/plm/plm.pdf .

Daly, Mary. 2020. "Conceptualizing and Analyzing Family Policy and How It Is Changing." *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker, 25-41. Chum: Palgrave Macmillan.

Datta Gupta, Nabanita, Nina Smith and Mette Verner. 2008. "The impact of Nordic countries' family friendly policies on employment, wages, and children." *Review of Economics of the Household* 6: 65-89.

Datta Gupta, Nabanita, Ronald L. Oaxaca and Nina Smith. 2006. "Swimming Upstream, Floating Downstream: Comparing Women's Relative Wage Progress in the United States and Denmark." *Industrial and Labor Relations Review* 59 (2): 243-266.

Davidson, Skylar. 2016. "Gender inequality: Nonbinary transgender people in the workplace." *Coagent Social Sciences* 2 (1), 1-12. DOI: 10.1080/23311886.2016.1236511.

Djurdjevic, Dragana and Sergiy Radyakin. 2007. "Decomposition of the Gender Wage Gap Using Matching: An Application for Switzerland." *Swiss Journal of Economics and Statistics* 143 (4): 365-396.

Duraisamy, Malathy and P. Duraisamy. 2016. "Gender wage gap across the wage distribution in different segments of the Indian labour market, 1983-2012: exploring the glass ceiling or sticky floor phenomenon." *Applied Economics* 48 (43): 4098-4111.

Ehrenberg, Ronald G. and Robert S. Smith. 2018. *Modern Labor Economics. Theory and Public Policy.* New York: Routledge.

Ekberg, John, Rickard Eriksson and Guido Friebel. 2013. "Parental leave – A policy evaluation of the Swedish 'Daddy-Month' reform." *Journal of Public Economics* 97: 131-143.

Estévez-Abe, Margarita. 2005. "Gender Bias in Skills and Social Politics: The Varieties of Capitalism Perspective on Sex Segregation." *Social Politics* 12 (2): 180-215.

Estévez-Abe, Margarita. 2006. "Gendering the Varieties of Capitalism. A Study of Occupational Segregation by Sex in Advanced Industrial Societies." *World Politics* 59 (1): 142-175.

Eurostat. 2017. "Participation in early childhood education by sex (children aged 3 and over)." Indicator from *Eurostat's Sustainable Development Goals indicator set*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/sdg_04_31_esmsip2.htm (accessed on March 3rd, 2022).

Eurostat. 2022. "Children aged less than 3 years in formal childcare". Indicator from *Eurostat*'s indicators on *Social protection and inclusion*. URL: https://ec.europa.eu/eurostat/databrowser/view/tepsr_sp210/default/table?lang=en (accessed on March 3rd, 2022).

Fleckenstein, Timo. 2010. "Party Politics and Childcare: Comparing the Expansion of Service Provision in England and Germany." *Social Policy & Administration* 44 (7): 789-807.

García, Jaume, Pedro J. Hernández and Angel López-Nicolás. 2001. "How wide is the gap? An investigation of gender wage differences using quantile regression." *Empirical Economics* 26: 149-167. 104 (4): 1091-1119.

Gauthier, Anne H. (2011). Comparative Family Policy Database, Version 3. Netherlands Interdisciplinary Demographic Institute and Max Planck Institute for Demographic Research (distributors). Retrieved from: www.demogr.mgp.de .

Gerring, John. 2004. "What is a Case Study and What is it Good for?" *American Political Science Review* 98 (2): 341-254.

Gerring, John. 2012. *Social Science Methodology: A Unified Framework*. Cambridge: Cambridge University Press.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091-1119.

Gornick, Janet C., Maria K. Meyers and Katherin E. Ross. 1998. "Public Policies and the Employment of Mothers: A Cross-National Study." *Social Science Quarterly* 79 (1): 35-54.

Grönlund, Anne and Charlotta Magnusson. 2016. "Family-friendly policies and women's wages – is there a trade-off? Skill investment, occupational segregation and the gender pay gap in Germany, Sweden and the UK." *European Societies* 18 (1): 91-113.

Hall, Peter A., and David Soskice. 2001. *Varieties of Capitalism. The Institutional Foundations of Comparative Advantage.* Oxford: Oxford University Press.

Hara, Hiromi. 2018. "The gender wage gap across the wage distribution in Japan: Within- and between-establishment effects." *Labour Economics* 53: 213-229.

Hipp, Lena. 2020. "Do Hiring Practices Penalize Women and Benefit Men for Having Children? Experimental Evidence from Germany." *European Sociological Review* 36 (2): 250-264.

Hook, Jennifer L., and Becky Pettit. 2016. "Reproducing Occupational Inequality: Motherhood and Occupational Segregation." *Social Politics* 23 (3): 329-362.

Hook, Jennifer L., and Meiying Li. 2020. "Gendered Tradeoffs." *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker, 249-266. Chum: Palgrave Macmillan.

Kamerman, Sheila B. and Peter Moss. 2009. *The politics of parental leave policies. Children, parenting, gender and the labour market.* Edited by Sheila B. Kamerman and Peter Moss. Great Britain: The Policy Press.

Keck, Wolfgang and Chiara Saraceno. 2013. "The Impact of Different Social-Policy Frameworks on Social Inequalities among Women in the European Union: The Labour-Market Participation of Mothers." *Social Politics* 20 (3): 297-328.

Kellstedt, Paul M. and Guy D. Whitten. 2018. *The Fundamentals of Political Science.* Cambridge: Cambridge University Press.

Kleider, Hanna. 2015. "Paid and unpaid work: The impact of social policies on the gender division of labour." *Journal of European Social Policy* 25 (5): 505-520.

Kolberg, Jon Eivind and Gøsta Esping-Andersen. 1991. "Welfare States and Employment Regimes." *International Journal of Sociology* 21 (1): 3-35.

Korpi, Walter, Tommy Ferrarini and Stefan Englund. 2013. "Women's Opportunities under Different Family Policy Constellations: Gender, Class, and Inequality Tradeoffs in Western Countries Re-Examined." *Social Politics* 20 (1): 1-40.

Kotsadam, Andreas and Henning Finseraas. 2011. "The state intervenes in the battle of the sexes: Causal effects of paternity leave." *Social Science Research* 40: 1611-1622.

Kunze, Astrid. 2018. "The Gender Wage Gap in Developed Countries." In *The Oxford Handbook of Women and the Economy*, edited by Susan Averett, Laura M. Argys and Saul D. Hoffman, 369-394. New York: Oxford University Press.

Laitin, David. 2002. "Comparative Politics: The State of the Subdiscipline." In *Political Science: State of the Discipline* edited by Ira Katznelson and Helen V. Milner, 630-661. New York: Norton.

Lazear, Edward P. and Sherwin Rosen. 1990. "Male-Female Wage Differentials in Job Ladders." *Journal of Labour Economics* 8 (1): 106-123.

Lohmann, Henning and Hannah Zagel. 2016. "Family policy in comparative perspective: The concepts and measurement of familization and defamilization." *Journal of European Social Policy* 26 (1): 48-65.

Lüdecke, D. 2021. *sjPlot: Data Visualization for Statistics in Social Science.* R-package version 2.8.10, URL: https://CRAN.R-project.org/package=sjPlot .

Maczulskij, Terhi and Jukka Nyblom. 2020." Measuring the gender wage gap – a methodological note." *Applied Economics* 52 (21): 2239-2249.

Mandel, Hadas and Michael Shalev. 2009. "How welfare states shape the gender pay gap: A theoretical and comparative analysis." *Social Forces 87* (4): 1873-1911.

Mandel, Hadas and Moshe Semyonov. 2005. "Family Policies, Wage Structures, and Gender Gaps: Sources of Earnings Inequality in 20 Countries." *American Sociological Review* 70 (6): 949-967.

Mandel, Hadas and Moshe Semyonov. 2006. "A Welfare State Paradox: State Interventions and Women's Employment Opportunities in 22 Countries." *American Journal of Sociology* 111 (6): 1910-1949.

McDonald, Judith Ann and Robert Thornton. 2016. "Have Pay Equity Laws in Canada Helped Women? A Synthetic-Control Approach." *American Review of Canadian Studies* 46 (4): 452-473.

Mincer, Jacob and Haim Ofek. 1982. "Interrupted work careers: Depreciation and restoration of human capital." *Journal of human resources*: 3-24.

Mincer, Jacob and Solomon Polachek. 1974. "Family investments in human capital: Earnings of women." *Journal of political economy* 82 (2): 76-108.

Moses, Jonathon and Torbjørn L. Knutsen. 2012. *Ways of Knowing: Competing Methodologies in Social and Political Research*. 2nd edition. Palgrave Macmillan.

Newman, Benjamin J. 2016. "Breaking the Glass Ceiling: Local Gender-Based Earnings Inequality and Women's Belief in the American Dream." *American Journal of Political Science* 60 (4): 1006-1025.

Nielsen, Helena S., Marianne Simonsen and Mette Verner. 2004. "Does the Gap in Family-Friendly Policies Drive the Family Gap?" *The Scandinavian Journal of Economics* 106 (4): 721-744.

Nieuwenhuis, Rense and Wim Van Lancker. 2020. *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker. Chum: Palgrave Macmillan.

Nopo, Hugo. 2008. "Matching as a tool to decompose wage gaps." *The Review of Economics and Statistics* 90 (2): 290-299.

OECD. 2021a. "Earnings: Gross earnings: decile ratios (Edition 2020)". Indicator from the *OECD Employment and Labour Market Statistics Database.* DOI: 10.1787/5177e25c-en (Accessed on February 19th, 2022).

OECD. 2021b. "Strictness of employment protection – temporary contracts." Indicator from the *OECD Indicators of Employment Protection*. URL: https://www.oecd.org/employment/emp/oecdindicatorsofemploymentprotection.htm (Accessed on April 26th, 2022).

OECD. 2021c. "Strictness of employment protection – individual and collective dismissal." Indicator from the *OECD Indicators of Employment Protection*. URL: https://www.oecd.org/employment/emp/oecdindicatorsofemploymentprotection.htm (Accessed on April 26th, 2022).

OECD. 2022a. "Gender wage gap." Indicator from the *OECD Employment Database – Earnings and Wages.* DOI:10.1787/7cee77aa-en (accessed on February 19th, 2022).

OECD. 2022b. "Public policies for families and children." Indicators from the *OECD Family Database.* URL: https://www.oecd.org/els/family/database.htm#public_policy (accessed on February 27th, 2022).

OECD. 2022c. "Labour force participation rate." Indicator from the *OECD Employment Database*. DOI: 10.1787/8a801325-en (accessed on April 20th, 2022).

OECD. 2022d. "Distribution of gross earnings of full-time employees." Indicators from the *OECD Employment Database.* URL: https://stats.oecd.org/Index.aspx?DatasetCode=DEC_I# (accessed on April 27th, 2022).

OECD. 2022e. "Trade Unions: Trade union density." Indicator from the *OECD Employment and Labour Market Statistics database*. DOI: https://doi.org/10.1787/data-00371-en (accessed on April 27th, 2022).

OECD. 2022f. "Collective bargaining coverage." Indicator from the *OECD/AIAS ICTWSS database*. URL. https://www.oecd.org/employment/ictwss-database.htm (Accessed on April 26th, 2022).

Olivetti, Claudia and Barbara Petrongolo. 2017. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives* 31 (1): 205-230.

Padavic, Irene, Robin J. Ely and Erin M. Reid. 2020. "Explaining the Persistence of Gender Inequality: The Work-family Narrative as a Social Defence against the 24/7 Work Culture." *Administrative Science Quarterly* 65 (1): 61-111.

Pettit, Becky and Jennifer L. Hook. 2005. "The Structure of Women's Employment in Comparative Perspective." *Social Forces* 84 (2): 779-801.

Podestà, Frederico. 2017. "The impact of 'free choice': Family reforms of France and Belgium, a synthetic control analysis." *International Journal of Social Welfare* 26: 340-352.

Rege, Mari and Ingeborg F. Solli. 2013. "The impact of paternity leaves on fathers' future earnings." *Demography* 50 (6): 2255-2277.

Rubery, Jill. 2011. "Towards a gendering of the labour market regulation debate." *Cambridge Journal of Economics* 35 (6): 1103-1126.

Ruhm, Christopher J.. 1998. "The Economic Consequences of Parental Leave Mandates: Lessons from Europe." *Quarterly Journal of Economics* 113 (1): 285-317.

Sartori, Giovanni. 1970. "Conceptual Misformation in Comparative Politics." *American Political Science Review* 64 (4): 1033-1053.

Schäfer, Andrea and Karin Gottschall. 2015. "From wage regulation to wage gap: how wage-setting institutions and structures shape the gender wage gap across three industries in 24 European countries and Germany." *Cambridge Journal of Economics* 39 (2): 467-496.

Schäfer, Andrea, Ingrid Tucci and Karin Gottschall. 2012. "Top Down or Bottom Up? A Cross-National Study of Vertical Occupational Sex Segregation in 12 European Countries." in *Firms, Boards and Gender Quotas: Comparative Perspectives* edited by Englestad, F. and Teigen, M., 3-43. Emerald Group Publishing: Bingley.

Slater, Dan and Daniel Ziablatt. 2013. "The Enduring Indispensability of the Controlled Comparison." *Comparative Political Studies* 46 (10): 1301-1327.

Thévenon, Olivier. 2008. "Family policies in Europe: available datasets and initial comparisons." *Vienna Yearbook of Population Research* 6: 165-177.

Trapletti, Adrian and Kurt Hornik. 2022. *tseries. Time Series Analysis and Computational Finance* (Version 0.10-51). Available from CRAN - Package tseries (r-project.org).

Vandenbroeck, Michel. 2020. "Early Childhood Care and Education Policies that Make a Difference." *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker, 169-191. Chum: Palgrave Macmillan.

Wei, Taiyun and Viliam Simko. 2021. *R package 'corrplot': Visualization of a Correlation Matrix* (Version 0.92). Available from https://github.com/taiyun/corrplot .

White, Linda A. 2020. "Do International Organizations Influence Domestic Policy Outcomes in OECD Countries?" *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker. Chum: Palgrave Macmillan.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag: New York. URL: https://ggplot2.tidyverse.org .

World Bank. 2021. "GDP per capita (current US$)." Indicator from *World Development Indicators*. URL: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD (accessed on April 27th, 2022).

World Bank. 2022a. "School enrolment, tertiary (% gross)." Indicator from *World development indicators*. URL: data.worldbank.org/indicator/SE.TER.ENRR (accessed on April 27th, 2022).

World Bank. 2022b. "Public sector employment as a share of total employment by gender (Female)." Indicator from *The Worldwide Bureaucracy Indicators*. URL: https://govdata360.worldbank.org/indicators/he75ced3a?country=BRA&indicator=42 306&viz=line_chart&years=2000,2017 (accessed on April 27th, 2022).

Xiu, Lin and Morley Gunderson. 2012. "Glass ceiling or sticky floor? Quantile regression decomposition of the gender pay gap in China." *International Journal of Manpower* 35 (3): 306-326.

Zagel, Hannah and Henning Lohmann. 2020. "Conceptual Approaches in Comparative Family Policy." in *The Palgrave Handbook of Family Policy*. 1st edition, edited by Rense Nieuwenhuis and Win Van Lancker, 119-140. Chum: Palgrave Macmillan.

Zeileis, A., Köll, S. and Graham, N. 2020. "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R." *Journal of Statistical Software* 95 (1): 1-36. DOI: 10.18637/jss.v095.i01.

Zeileis, Achim and Torsten Hothorn. 2002. "Diagnostic Checking in Regression Relationships." *R News* 2 (3): 7-10.

Zeileis, Achim. 2004. "Econometric Computing with HC and HAC Covariance Matrix Estimators." *Journal of Statistical Software* 11 (10): 1-17.

APPENDIX A: Variable descriptions

TABLE A.1. *Full overview of descriptive statistics and data sources.*

Variable	Operationalisation and original data source	N	Mean	St.dev	Min	Max
Dependent variables						
Gender wage gap	Difference between male and female median wages divided by the male median wage for full time employees, in percentages. From the OECD Employment Database.	487	15.03	5.81	0.38	32.74
Gender wage gap, 1st decile	Gender wage gap at the first full-time earnings decile. From the OECD Employment Database.	328	10.49	6.69	-9.50	32.00
Gender wage gap, 9th decile	Gender wage gap at ninth full-time earnings decile. From the OECD Employment Database.	343	21.36	5.52	-5.60	36.70
Independent variable indicators *1. Parental leave*						
Paid parental leave	Total duration of paid maternity and parental leave available to women after childbirth in weeks. From the OECD Family Database, supplemented with the same indicator from the Comparative Family Policy Database when missing observations.	463	59.79	53.35	0	164
Paid paternity leave	Total duration of paid paternity leave reserved for the exclusive use of fathers, in weeks. From the OECD Family Database.	463	4.01	6.43	0	28
Unpaid leave	Estimate of weeks of unpaid parental leave available to women after childbirth, in weeks. Constructed by subtracting *total parental leave* from *maternal and parental leave with job protection, disregarding payment*, both from the OECD Family database.	436	30.34	42.68	-10.70	156
Long leave	Dichotomous measure indicating long paid parental leave. 0 = less than 60 weeks, 1 = 60 weeks or more.	463	0.33	0.47	0	1
2. Childcare						
Childcare 0-2 years	Proportion of children aged 0-2 enrolled in formal childcare or preschool, in percentages. From the OECD Family Database, supplemented with the same indicator from Eurostat aggregated EU-SILC indicators.	274	31.02	18.92	1	78
Childcare 3-5 years	Proportion of children aged 3 to compulsory school age (usually 5 or 6) enrolled in formal childcare or preschool, in percentages. From the OECD Family Database, supplemented with similar indicator from Eurostat's Sustainable development indicators.	285	84.27	13.70	45	100
3. Tax breaks						
Public expenditure on tax breaks	Public social expenditure on tax breaks for families, as a percentage of GDP. From the OECD Family Database.	375	0.26	0.29	0.00	1.05

4. Added indices						
Family policy generosity	Average index of the variables *paid parental leave, paid paternity leave, childcare 0-2 years* and *childcare 3-5 year* standardised. Rescaled from 0 (lest) to 10 (most).	467	4.700	2.707	0.00	10.00
Familisation	Average index of the variables *unpaid leave* and *long leave* standardised. Rescaled from 0 (lest) to 10 (most).	463	2.941	2.564	0.00	10.00
Defamilisation	Average index of the variables *paid paternity leave* and *childcare 0-2 years* standardised. Rescaled from 0 (least) to 10 (most).	463	2.142	2.38	0.00	10.00
Control variables						
Female employment rate	Female employment rate for full-time employees. From OECD Employment Outlook indicators, supplemented with the same indicator from Eurostat.	453	62.49	8.55	34.60	83.57
Share of women in public employment	Share of women in public employment as a percentage of total female employment. From the World Bank's Worldwide Bureaucracy Indicators database.	321	0.65	0.07	0.45	0.78
GDP per capita	Gross domestic product per capita, in US dollars in thousands. From the World Bank's World Development Indicators.	487	36.39	11.04	8.48	72.39
Wage dispersion	Earnings dispersion as the ratio of goss earnings of 9th and 1st earnings decile. From the OECD Distribution of Earnings database.	455	3.25	0.79	1.79	5.22
Educational difference	Difference between male and female enrolment in tertiary education (ISCED 5 to 8). Calculated by the female enrolment rate minus the male enrolment rate. Enrolment rates are from the World Bank's World Development Indicators.	415	16.59	11.41	-15.87	45.98
Central bargaining	Consisting of trade union density and collective bargaining coverage standardised then added. From least to most covered by centralised bargaining (0-2). Both from the OECD Indicators of Trade Unions and Collective Bargaining.	323	0.88	0.55	0.02	1.89
Strictness of employment protection legislation	Consisting of strictness and hiring restrictions for using fixed-term and temporary contracts and of individual and collective dismissal standardised then added. From least to most strictness of employment protection legislation (0-2). Both are from the OECD Indicators of Employment Protection.	407	0.72	0.35	0.00	1.53

FIGURE A.1. *Correlation matrix, control variables.*
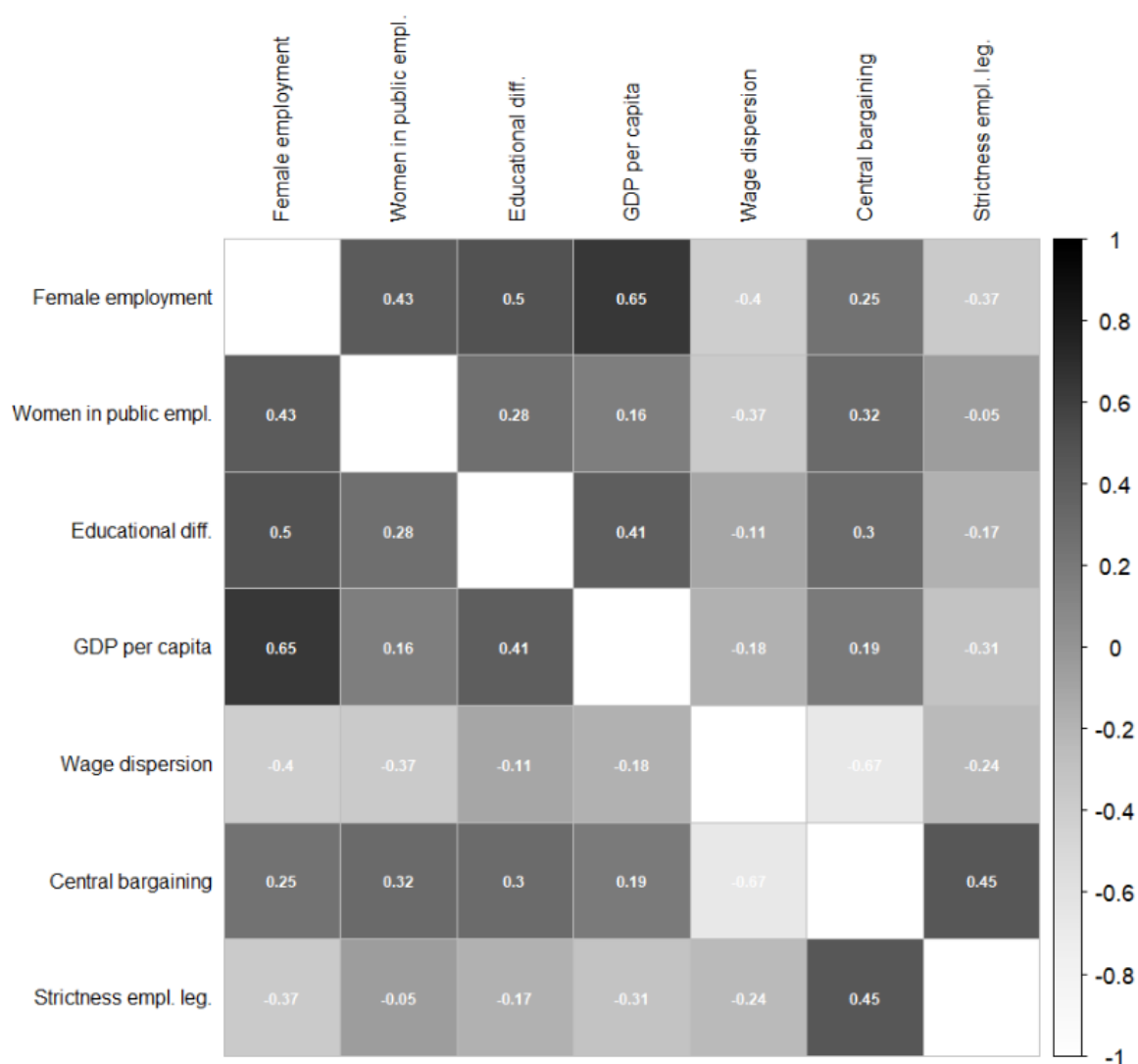
Notes: Correlation matrix for all control variables. Stronger colours indicate strong positive correlation, weaker colours indicate strong negative correlation. Visualisation based on R-package *corrplot* (Wei and Simko 2021). Data: OECD (2021 b-c, 2022 c-f), World Bank (2021, 2022 a-b).

APPENDIX B: Regression results using FE modelling

TABLE B.1. *Regression results using fixed effects, dependent variable: median wage gap.*

Dependent variable: Median wage gap	Generosity (B1.A)	Familisation (B1.B)	Defamilisation (B1.C)	Paid leave (B1.D)
Family policy generosity	**-0.659 **** (0.006, 0.239)			
Familisation		0.133 (0.604, 0.256)		
Defamilisation			-0.135 (0.573, 0.238)	
Paid parental leave				-0.122 (0.175, 0.089)
Paid parental leave^2				0.000 (0.654, 0.000)
Controls				
Female employment	0.112 (0.535, 0.180)	0.019 (0.933, 0.227)	0.111 (0.567, 0.193)	-0.054 (0.797, 0.208)
Female public sector empl.	0.487 (0.852, 2.595)	-572 (0.875, 3.620)	0.218 (0.946, 3.248)	1.297 (0.666, 3.00)
Educational difference	-0.068 (0.517, 0.104)	-0.096 (0.459, 0.130)	-0.057 (0.623, 0.116)	-0.164 (0.193, 0.126)
Wage dispersion	0.868 (0.151, 0.602)	0.819 (0.181, 0.610)	0.856 (0.184, 0.641)	0.569 (0.309, 0.558)
GDP per capita	-0.260 (0.158, 0.183)	-0.241 (0.249, 0.208)	0.281 (0.133, 0.186)	-0.305 (0.113, 0.192)
Central bargaining	**8.354 ***** (0.000, 1.610)	**8.064 ***** (0.000, 1.732)	**8.308 ***** (0.000, 1.537)	**6.466 **** (0.003, 2.117)
Employment protection	-4.048 (0.484, 5.769)	0.103 (0.983, 4.884)	-1.190 (0.830, 5.529)	-0.742 (0.893, 5.524)
Hausman test, p-value	0.726 (RE > FE)	0.356 (RE > FE)	0.219 (RE > FE)	0.477 (RE > FE)
Number of observations	216	216	212	216
Adjusted R-squared	0.279	0.196	0.185	0.290
F-statistic, p-value	0.000	0.000	0.000	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Fixed effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Includes results from the Hausman test indicating which model is preferred of RE and FE. Data: OECD (2021b, 2022 a-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b).

TABLE B.2. *Regression results using fixed effects, dependent variable: ninth earnings decile.*

Dependent variable: 9th earnings decile	Generosity (B2.A)	Familisation (B2.B)	Defamilisation (B2.C)	Paid leave (B2.D)
Family policy generosity	**-0.712 ** ** (0.003, 0.239)			
Familisation		0.079 (0.769, 0.260)		
Defamilisation			-0.361 (0.281, 0.334)	
Paid parental leave				0.272 (0.429, 0.351)
Paid parental leave^2				-0.001 (0.350, 0.002)
Controls				
Female employment	0.228 (0.206, 0.180)	0.109 (0.695, 0.227)	0.229 (0.282, 0.212)	0.111 (0.731, 0.321)
Female public sector empl.	-1.126 (0.665, 2.595)	-1.779 (0.547 2.950)	-0.354 (0.920, 3.436)	-2.704 (0.402, 3.216)
Educational difference	0.055 (0.600, 0.104)	0.004 (0.978, 0.152)	0.064 (0.597, 0.121)	-0.060 (0.697, 0.154)
Wage dispersion	**-6.416 ** *** (0.000, 0.602)	-6.624 (0.100, 4.006)	**-6.505 *** (0.098, 3.910)	**-6.236 *** (0.076, 3.494)
GDP per capita	**-0.808 *** ** (0.000, 0.181)	**-0.791 *** ** (0.001, 0.236)	**-0.808 *** ** (0.000, 0.225)	**-0.770 ** ** (0.001, 0.238)
Central bargaining	2.260 (0.162, 1.610)	1.563 (0.490, 2.258)	2.158 (0.187, 1.628)	3.339 (0.475, 4.663)
Employment protection	-8.802 (0.129, 5.769)	-5.251 (0.301, 5.058)	-7.948 (0.164, 5.691)	-1.593 (0.693, 4.028)
Hausman test, p-value	0.183 (RE > FE)	0.000 (RE < FE)	0.105 (RE > FE)	0.090 (RE > FE)
Number of observations	198	198	194	198
Adjusted R-squared	0.378	0.338	0.345	0.368
F-statistic, p-value	0.000	0.000	0.002	0.000

Notes: Significant p-values: * = p-value < 0.1, ** = p-value < 0.05, *** = p-value < 0.001. Fixed effects models with robust standard errors correcting heteroskedasticity. Shows regression coefficient with (p-value, robust standard error). Includes results from the Hausman test indicating which model is preferred of RE and FE. Data: OECD (2021b, 2022 a-f), Eurostat (2017, 2022), Gauthier (2011), World Bank (2021, 2022 a-b).