Essays on the Gender- and the Part-Time Wage Gap

A Distributional Approach

Inaugural-Dissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Patricia Gallego Granados
geboren in Barcelona, Spanien

Januar 2018

Gedruckt mit der Genehmigung des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin.

Dekan:
Prof. Dr. Dr. Andreas Löffler

Erstgutachter:
Prof. Dr. Peter Haan, Freie Universität Berlin

Zweitgutachter:
Prof. Dr. Christoph Breunig, Humboldt-Universität zu Berlin

Datum der Disputation:
16. Februar 2018

Erklärung über Zusammenarbeit mit Koautoren und Vorveröffentlichungen

Kapitel 1

In Zusammenarbeit mit Johannes Geyer (Eigenanteil: 50%).

Veröffentlicht als: Gallego Granados, Patricia and Geyer, Johannes, "Distributional and Behavioral Effects of the Gender Wage Gap", DIW Discussion Paper 1469, 2015.

Kapitel 2

In Zusammenarbeit mit Katharina Wrohlich (Eigenanteil: 50%). Bislang unveröffentlicht.

Kapitel 3

Bislang unveröffentlicht.

Acknowledgments

This doctoral thesis was prepared while I was a member of the Graduate Center of the German Institute for Economic Research (DIW Berlin) and a research associate at DIW's Public Economics department. I am grateful for the very stimulating research environment as well as for the financial support for my research project.

I want to express my deepest gratitude to Peter Haan for the excellent supervision, continued support and encouragement during all these years. My work has immensely benefited from his invaluable advice, constructive discussions and lasting patience over many previous drafts. I am also grateful to Christoph Breunig for co-supervising my dissertation and providing very valuable comments.

A felt thank you goes also to my co-authors Johannes Geyer and Katharina Wrohlich, from whom I have learned a lot and who have always kept an eye on how my dissertation and myself were doing over the last years.

I am equally indebted to the three generations of colleagues and friends with whom I have shared an office: the "original" ones, Michael Neumann and Andreas Thiemann, for sharing the ups and downs of the beginnings; the "young" ones, Renke Schmacker and Aline Zucco, for a warm welcome after my parental leave; and the "experienced" one, Anna Hammerschmid, for always having a nice word for me during the end phase of this dissertation. And Songül Tolan, with whom I never shared an office but it often felt as if I would.

I am also grateful to the colleagues at the DIW's Department of Public Economics for helpful comments in our internal seminars, all kinds of good ideas as well as much-appreciated chatting during coffee breaks and Christmas parties: Stefan Bach, Hermann Buslei, Sascha Drahs, Stefan Etgeton, Ronny Freier, Nicole Haase, Luke Haywood, Daniel Kemptner, Annekatrin Lüdecke, Kai-Uwe Müller, Richard Ochmann, Julia Schmieder, Ulrich Schneider and Clara Welteke.

My parents, Jordi and Isabel, and my sister, Irene, have my profound gratitude for a long list of things but, first and foremost, for always believing that this endeavour would eventually take an end, a happy one, especially in those moments I was no longer sure.

Most important of all, I would like to thank Patrick, who has always been there for me and whose unconditional support and love have made this dissertation possible. I am deeply grateful. And, finally, many thanks to Alma, who -without knowing it- has always done her bit in reminding me of the bigger picture.

Table of Contents

List of Tables

List of Figures

General Introduction

Motivation

The gender wage gap is a persistent and pervasive phenomenon observable in virtually every country's labour market. One of the most prominent questions in the literature has been whether unexplained gender wage gaps reflect gender discrimination in the labour market or, alternatively, unobserved differences in productivity between men and women (Becker 1971, Altonji and Blank 1999, Hellerstein et al. 1999, Altonji and Pierret 2001, Black and Strahan 2001, Hellerstein et al. 2002, among others). Most studies find evidence for both the existence of discrimination and of unobserved productivity differences. Without taking sides in the debate about discrimination versus productivity differences, Chapter 1 of this dissertation quantifies the effect of the unexplained gender wage gap in terms of net household income and female labour supply. Calculating the consequences of the gender wage gap for net household income and female labour supply requires considering intra-household bargaining, division of household chores as well as implicit gender-specific features of tax-benefit systems, all issues also widely dealt with in the literature (e.g. Baldwin and Johnson 1992 on the labour supply effects of wage discrimination, Hersch and Stratton 2002 on the effect of housework on wages, Alesina et al. 2011 on optimal gender-based taxation and Bertrand et al. 2015 on the effect of gender identity rules on intra-household income distribution and labour supply of women, just to mention a few).

Although the unexplained component of the gender wage gap has been constantly decreasing over time, the overall gender wage gap has decreased much less since the beginning of the 2000s. Consequently, the literature has also explored other factors contributing to the gender wage gap. Whereas human capital levels (such as educational attainment) accounted for a large share of the gap in the 1980s and 1990s (Blau and Kahn 2017), women have caught up with men's educational achievement in virtually all high-income countries (see Goldin et al. 2006, Becker et al. 2010). However, the gender wage gap widens with age over the life cycle - both for older and younger cohorts - which researchers attribute to marriage and motherhood (see Anderson et al. 2002, Angelov et al. 2016 or Juhn and McCue 2017, among others). Other factors examined in the literature are occupational segregation (Groshen 1991, Fitzenberger and Kunze 2005, Ludsteck 2014), intra-firm gender wage gaps (Heinze and Wolf 2010, Card et al. 2016) as well as labour market institutions such as wage-setting mechanisms (Blau and Kahn 2003), unions (Blau and Kahn 1992, 1996, Booth and Francesconi 2003) and family policies (Christofides et al. 2013, Olivetti and Petrongolo 2017).

Over time, in many countries the narrowing of the gender wage gap has coincided with the rise of female labour market participation rates and, also, of wage inequality. This has attracted the interest of several studies on the effect of changing selection into employment on the gender wage gap. Olivetti and Petrongolo (2008) show that a large share of the cross-country differences in gender wage gaps can be explained by differences in female employment rates. Mulligan and Rubinstein (2008) present a theoretical framework that links increasing positive selection into employment to rising wage inequality. Importantly, they show that convergence of male and female wages in a context of rising wage inequality can be overestimated if increasing selection into employment is not taken into account.

One last strand of literature relevant for this dissertation focuses on examining the large variation of the gender wage gap across the distribution as well as the different roots behind gender wage gaps among low- and high-earnings individuals (e.g. Albrecht et al. 2003, Arulampalam et al. 2007, Kassenböhmer and Sinning 2014).

Interweaving these different threads of research, Chapter 2 of this dissertation draws on both the literature about the role of changing selection into employment as well as the importance of considering the gender wage gap across the distribution. In particular, Chapter 2 makes use of state-of-the-art econometrics literature to examine the changing

effect of selection into employment on the gender wage gap across the distribution and over time. The goals of Chapter 2 are fourfold: First, I provide a descriptive analysis of the gender wage gap across the distribution and over time, including its heterogeneity by human capital levels and family status. Next, I characterise the magnitude and evolution of male and female selection into employment across the distribution and over time. Third, I present estimates of selection-corrected gender wage gaps. Finally, I carry out a decomposition exercise to better understand the trends in the overall gap.
In Chapter 2, I restrict the analysis to the gender wage gap that arises among full-time employees. The reasoning is that part-time employment may act as a confounding factor if the wage structures of full- and part-time employment differ, especially given that most part-time employees are women. However, the literature increasingly hints at part-time employment as a factor behind the persistence of the gender wage gap (see Manning and Petrongolo 2008, Bowlus and Grogan 2009, Mumford and Smith 2009, Goldin 2014 and Blau and Kahn 2017). Against this backdrop, Chapter 3 of this dissertation specifically takes on this issue and examines the female part-time wage gap across the distribution and over time.

The part-time wage gap, besides its effect on the gender wage gap, is worth studying for several reasons. First, the female part-time wage gap is large and has been steadily increasing over time (Manning and Petrongolo 2008, OECD 2010). Second, in several countries part-time employment has virtually captured all the increase in female labour supply over the last decades. Furthermore, from a life cycle perspective, a big share of women report having worked part-time at some point in their lives (see Connolly and Gregory 2010 for the UK and Paul 2016 for Germany). Therefore, any potential part-time pay penalty is affecting a large and increasing share of the female population. Third, a good understanding of this penalty is crucial because it amplifies the adverse economic consequences of part-time employees' inherently lower earnings, including weaker insurance in social security schemes as well as lower old-age pension incomes (see Bardasi and Gornick 2008). Moreover, in the presence of elastic labour supply curves, the part-time penalty is likely to induce negative labour supply effects.

In Chapter 3, I argue that the observed part-time wage gap can be the result of lower returns to skill in part-time employment or the consequence of changing selection patterns into full- and part-time employment, especially in a context of rising wage inequality (Blundell et al. 2007, Mulligan and Rubinstein 2008, Arellano and Bonhomme 2017, Biewen et al. 2017) and a general increase in female labour market participation (Jacobsen et al. 2015, Blau and Kahn 2017). The aim of the third chapter is to analyse to which extent the raw part-time wage gap can be explained by part-time employment being paid worse than full-time employment or, alternatively, by composition effects resulting from different selection patterns into full- and part-time work.

Methodology

The uniting methodological approach of this dissertation is the use of conditional quantile regressions as a tool for the distributional analysis of the gender- and part-time wage gap. Otherwise, the results presented here rely on decomposition exercises (Chapters 1 and 2), correction for selection into employment across the wage distribution (Chapters 2 and 3), and microsimulation combined with a simple discrete choice labour supply model (Chapter 1).

Decomposition techniques are widely used in economics in order to examine phenomena like the gender wage gap. In the case of wages, decomposition techniques aim at isolating differences in the characteristics of the two groups under comparison (often referred to as the explained component of the gap) from the returns to those characteristics (also called unexplained component). Since the seminal work of Oaxaca (1973) and Blinder (1973), who pioneered decompositions in wage gaps, a vast literature has suggested many extensions to their original decompositions (see Fortin et al. 2011 for an overview). Chapters 1 and 2 make use of one such extension and perform a decomposition across the wage distribution as proposed by Machado and Mata (2005) and Melly (2005). They rely on conditional quantile regressions to generate the counterfactual distribution needed for the decomposition and yield equivalent results, but differ in how they derive the unconditional distribution from the conditional quantile regression results. As an alternative to regression-based decompositions, reweighing methods

have also been proposed to reach the same goal (e.g. DiNardo et al. 1996). Moreover, several studies issue a word of caution on the causal interpretation of decomposition results (e.g. Black et al. 2008, Huber 2015), not least because of their failure to control for general equilibrium effects which likely arise from the counterfactual distributions.
Chapters 2 and 3 use an imputation-based method to correct for sample selection over the distribution developed by Melly and Santangelo (2014, 2015). Imputation-based approaches used to correct for selection into employment have been widely used in the literature (Neal 2004, Blau and Kahn 2006, Olivetti and Petrongolo 2008, among others). Mostly, those rely on informed guesses of whether non-realized wages fall above or below the observed median wages - conditional on the education level of the individual. The imputation method proposed by Melly and Santangelo (2015) goes one step further and suggests imputing a point-identified wage for each individual out-of-work which takes into account both observable and unobservable characteristics of the individual. To this end, the authors suggest an extension of Athey and Imbens (2006)'s changes-in-changes model. Once each individual in the sample gets assigned a wage - which can be observed or imputed - the resulting wage distribution is by definition selection-corrected.
More concretely, Melly and Santangelo (2015)'s imputation method recovers information on an individual's unobservables from an individual's realised wage obtained from longitudinal data and assumes the time-invariance of the unobservables conditional on the observables to impute this individual's wage whenever she or he is out of work. Intuitively, this implies that the distribution of unobservable characteristics which have an effect on wages (such as innate ability, professional ambition) stay constant over time conditional on observable characteristics (such as education level and previous working experience).
For the analysis of the female part-time wage gap in Chapter 3, I exploit the fact that many women in Western economies work both full- and part-time at some point in their lives (see, e.g. Connolly and Gregory 2010, for the UK or Paul 2016, for Germany) to extend Melly and Santangelo (2015)'s method to the imputation of non-realized part-time wages in addition to the imputation of full-time wages. Deriving

these two selection-corrected wage distributions requires imputing both a full- and a part-time wage for each woman each year whenever one (or both) of these wages is not realised. For those women only observed working either full- or part-time, or never observed working at all, an additional assumption regarding their unobservables is needed. I discuss different options and perform several robustness checks that confirm the stability of the results presented.

Contribution and Main Findings

This dissertation contributes to the literature in three important ways. First, it shows that the observed convergence in male and female (full-time) wages in Germany from mid 1980s until nowadays is mostly explained by increasing positive selection of women into full-time employment. Once selection into employment is controlled for, the median gender wage gap by the end of the 2000s turns out to be almost as high as twenty-five years before. Second, my results suggest that the raw female part-time wage gap - which has been widening over the last decades - is best explained by opposite patterns of selection into employment in the full- and part-time sectors. However, the wage structures of full- and part-time employment have become more equal over time. Third, from the methodological perspective, I extend an existing econometric procedure to examine the part-time wage gap across the distribution and over time, while making it possible to control for selection into full- and part-time employment.

The main findings of the dissertation can be summarized as follows. Chapter 1 quantifies the effect of the unexplained gender wage gap in terms of household disposable income and labour supply effects. The results show that, on average, the unexplained gender wage gap amounts to approximately 5% of households' disposable income. Furthermore, in the aggregate the unexplained component of the gender wage gap is associated with the labour market non-participation of around 105,500 women as well as with a reduction of working hours of around 247,600 full-time equivalents.

Chapter 2 analyses the effect of changing selection into employment on the evolution of the gender wage gap across the distribution and over time. The results indicate that the sample of women working full-time is strongly positively selected at all points of the wage distribution - and the magnitude of this selection increases over time. The sample of men working full-time is also slightly positively selected, but less so compared to women. Nonetheless, the magnitude of selection at the lower bound of the male distribution has increased steeply from the beginning of the 1990s onwards. Second, as a result of the revealed trends regarding selection into full-time employment, the selection-corrected gender wage gap is much higher than the observed gap. Moreover, while by the end of the 2000s the observed gender wage gap was smaller than 25 years before, the selection-corrected gender wage gap was not. Third, the decomposition analysis shows that the unexplained gender wage gap is decreasing over time. This is particularly true for the bottom of the wage distribution. In other parts of the wage distribution, unexplained gaps are still present and increase with wages. For women earning high wages, differences in coefficients still contribute more to the total gender wage gap than differences in their characteristics.

Chapter 3 examines the female part-time wage gap across the distribution and over time, while accounting for changing selection into full- and part-time employment. My findings suggest that the increasing female part-time wage gap over time is driven by opposite patterns of selection into full- and part-time employment. More precisely, I find that selection into full-time employment increases over time, whereas the sample working part-time becomes less selected over time. By the end of the 2000s, selection into part-time employment is non-existent for median wages and becomes negative at the lower end of the distribution. My findings also suggest that the wage structures of full- and part-time employment have converged over time and, by the end of the 2000s, the results point at a small but statistically significant corrected part-time wage premium.

This dissertation points to two major policy recommendations. If the goal is to reduce the gender- and part-time wage gap, then policies need to first reduce restrictions that hinder female full-time employment and, secondly, further facilitate the switch between full- and part-time employment in both directions and for all individuals, regardless of skill level and gender.

Chapter 1

Distributional and Behavioural Effects of the Gender Wage Gap*

1.1 Introduction

In this chapter we examine the distributional (available income) and behavioural (labour supply) consequences of the well-known gender wage gap. The gender gap in gross wages is a persistent and pervasive phenomenon observable in virtually every country's labour market. The OECD (2012, p.166) reported an average gender wage gap of 16% for full-time employees in its member countries in 2010. Numerous studies have analysed its magnitude, determinants and development over time (see, for instance, Bertrand 2011 or Blau and Kahn 2003). The influential decomposition literature, starting with the seminal work of Oaxaca (1973) and Blinder (1973), has sought to divide the observed differences between men's and women's wages into the effects of different human capital endowments and their returns. These studies find that observed wage differences can be only partly explained by observable productivity differences, and that a large share remains unexplained.

* This Chapter is based on joint work with Johannes Geyer.

Given the systematic and sizeable gross wage differences between men and women, the wage gap obviously affects disposable income and probably behaviour as well. It is likely to impact female labour supply negatively, and women generally tend to react more elastically than men to changes of the wage rate. Depending on the distribution of the gender wage gap, it also affects the income distribution between and within households in a non-trivial way. To the extent that the gender wage gap affects the relative incomes of household members, it is likely to play an important role in most intra-household bargaining models (see, for instance, Browning et al., 1994). However, there is very little research on distributional or behavioural effects of the wage gap. We contribute to the literature by proposing a flexible approach to its effect on income distribution and labour supply. Our main contribution is to link the gender wage gap in gross hourly wages to net disposable income. Using a labour supply model, we estimate the effect of a reduction of the gender wage gap on female labour supply. Our proposed framework builds on the decomposition literature, which we combine with simulation of the tax-benefit system as well as structural estimation of labour supply. We provide a systematic discussion of each necessary step, underlying assumptions as well as data requirements[1].

The distribution of tangible financial consequences deriving from gender-based hourly wage differentials are, *a priori*, unknown – mainly for three reasons: First, the distribution of working hours needs to be factored in. Especially in countries where part-time work is widespread, the magnitude of the effect of a given (gross) wage differential in terms of gross earnings varies along the hours and the wage distribution. Second, the tax-benefit system is crucial for disposable income. Tax-benefit systems have an implicit gender dimension, which is best exemplified by schemes of joint taxation. But also implicit equivalence scales which underlie all tax-benefit systems often differ by household type and can thus produce different incentives for different household types. Third, to the degree that households share resources, the distributional impact depends on the income sharing rule within the household.

[1] In one of the few papers that focuses on behavioural effects of the gender wage gap, Aizer (2010) shows that a reduction of the gender wage gap had a negative effect on domestic violence in the United States. However, it is often difficult to find exogenous variation of the gender wage gap in order to identify a causal effect.

The presented framework for deriving the net gender pay gap involves four steps: (1) The first step is a quantile decomposition of hourly wage differences between men and women along the wage distribution. To this end, we estimate conditional quantile wage regressions and apply the decomposition method proposed by Machado and Mata (2005). We derive the explained and unexplained portion of the wage gap. (2) In the second step of our analysis, we simulate counterfactual gross hourly wages for women in the absence of gender wage gap (or a component thereof). The semi-parametric decomposition method provides a characterization of a counterfactual distribution for female wages but not the desired counterfactual wages for each individual. Therefore, we derive individual wages by assuming rank invariance of female individuals in the counterfactual distribution with respect to the baseline distribution. This allows us to increase individual female wage rates by a mark-up. On this basis, we can then derive the counterfactual gross monthly earnings distribution by accounting for hours worked. (3) The third step is the calculation of disposable income. At this stage, in addition to the counterfactual gross earnings obtained from step two, we require information on the household level (other own income, income from other household members, household characteristics relevant to the tax and transfer system such as number of children, etc.) which is then fed into a simulation model of the applying tax-benefit system (for details of the model, see Steiner 2012). The difference in household disposable income between the observed and the counterfactual scenario can be interpreted as the representation of the gender wage gap – originally estimated in hourly gross wages – in terms of household disposable income. However, the level of analysis is now the household and no longer the individual. In order to recover an individual measure of the net pay gap, information about an income sharing rule within the household is needed. Given that we do not observe this information, we use two common approaches in the literature, intra-household income pooling and proportional income sharing, which respectively yield a lower and upper bound of the net pay gap at individual level. (4) In a final step, we go beyond a static analysis and allow for labour supply adjustments. Using a discrete choice labour supply model, we derive net pay gaps with labour supply adjustments as well as aggregated labour supply effects resulting from the pay gap.

The empirical analysis is carried out using data from the German Socio-Economic Panel (GSOEP) for West Germany. Analysing the distributional impact of the gender pay gap in Germany is interesting for several reasons. First of all, the gender pay gap is relatively large and persistent in Germany: in 2010, the median wage gap of full-time working women was about 22 percent and has only slightly narrowed since 2000. Comparing female employees working full-time, several studies also suggest that the wage gap varies across the wage distribution in Germany (OECD 2012; Christofides et al. 2013; Arulampalam et al. 2007). Secondly, the share of women working part-time is particularly high in Germany: in 2010, nearly 40 percent of all employed women worked part-time (OECD, 2012, p.161). Thirdly, Germany is one of the few remaining countries, together with France and the United States, retaining a system of joint taxation of married couples. This leads to high marginal tax rates for the second earner which is typically the female spouse. On the other hand, income redistribution might mitigate disadvantageous labour market treatment of women with low household income.

We find that, on average, the unexplained household net gap is around 5.2 percent of household's equivalised net income and is higher for women living in single households (with or without children) than in couple households. Furthermore, within each of these two groups, the net financial consequences of the unexplained gap are higher for households without children. This pattern can be explained by larger earnings inequality of households with children. In particular, the share and level of female earnings in couple households with children are smaller than in households without children. At individual level, we can only identify bounds for the net pay gap. These range, on average, from 5.2 to 9.7 percent. The more (less) the actual sharing rule approaches perfect income pooling, the lower (higher) the net unexplained gap at individual level. When we allow for labour supply effects, we find a higher impact of closing the gender wage gap for women living in couple households than in single households. This s consistent with the higher labour supply elasticities of the first group. In the aggregate, we find the unexplained gap to be associated with the labour market non-participation of around 105,500 women and with a reduction of working hours of around 247,600 full-time equivalents.

The paper proceeds as follows: The next section provides an overview of the related literature. Our methodological approach is detailed in Section 3. Section 4 presents our application and findings. Section 5 concludes.

1.2 Previous findings

Empirical research on gender wage gaps is mainly focused on the identification of factors that explain systematic differences in gross wages of men and women. In this sense, Blau and Kahn (1996, 2003)'s cross-country studies find that wage-setting institutions (such as collective bargaining agreements) are associated with lower gender wage gaps. Olivetti and Petrongolo (2008) analyse the effect of women's non-random selection into employment on the gender wage gap of various OECD countries and find sizeable negative effects for southern EU countries. Goldin (2014), based on US data, points at wage penalties arising from part-time work as a major source of gender differences in wage rates. Most studies that analyse differences in wage rates between women and men use a standardized wage measure (e.g. gross hourly wage) to identify the factors that explain pay differences.

Studies on the gender wage gap show that observed wage differences can be only partly explained by observable productivity differences. A part of this literature uses decomposition approaches to explain differences between female and male wages (see Fortin et al. 2011 for an overview of the state of decomposition techniques in economics). Having its origins in the seminal work of Oaxaca (1973) and Blinder (1973), the decomposition literature has additionally developed semi- and non-parametric techniques that go beyond the mean decomposition; this enables analyses of the gender wage gap across the whole wage distribution (see, e.g., DiNardo et al., 1996; Fortin and Lemieux, 1998; Machado and Mata, 2005; Melly, 2005). The idea behind these techniques is that the wage gap might vary across the wage distribution. Albrecht et al. (2003), for example, argue that a small average wage gap might conceal "glass ceiling effects" at the top of the wage distribution and find evidence for such an effect in Sweden. Rica et al. (2008) find both a glass ceiling and a "sticky floor" effect for Spain. Arulampalam et al. (2007), in a study on eleven European countries, find a glass ceiling effect for most analysed countries. One key question in this literature is whether gender differences in the returns to labor market endowments which cannot be explained by differences in productivity are due to gender discrimination (e.g., Hellerstein et al., 1999; Black and Strahan, 2001).

In our application we focus on Germany. Several studies find a relatively large gender pay gap for Germany that varies with estimation sample, dataset and year. The average raw gap in gross hourly wages lies between 23 and more than 30 percent (e.g., Busch and Holst (2008); Anger and Schmidt (2010); OECD (2012)). Decomposition exercises à la Oaxaca-Blinder show that a large part of this gap cannot be explained by differences in the labour market skills of women and men (e.g., Anger and Schmidt (2010)). In addition, most studies find some variation of the gender wage gap across the wage distribution although its estimated shape across the distribution differs depending on estimation sample, year and dataset used (see, e.g.,Fitzenberger and Kunze 2005, Heinze and Wolf 2010, Antonczyk et al. 2010, Arulampalam et al. 2007).

Parallel to the body of gender wage gap research and the decomposition of income inequality, a seemingly unrelated literature examines the gender aspects of tax-benefit systems. Even though nowadays tax-benefit systems in Western countries do not distinguish between men and women, the application of the same rules can have implicit gender specific effects (Stotsky, 1996). This branch of scholarship examines the role of tax-benefit systems on several objects of interest such as the intra-couple distribution of earnings (see Figari et al., 2011), working (dis)incentives (Immervoll et al., 2011), as well as optimal (gender-based) taxation (Alesina et al., 2011; Bastani, 2013). Very relevant to the present paper is the issue of shifting the level of analysis from the individual to the household (and vice versa) – which becomes even more complex if the tax-benefit system is taken into account (Bargain, 2008; Sutherland, 1997).

Lastly, there are some studies on income inequality which relate to our analysis. These focus on whether the secular increase in female labour force participation had an influence on the development of income inequality. The increase in female labour force participation was mainly driven by a higher participation rate of married women in many countries. Cancian and Reed (1998, 1999) discuss different methods to decompose inequality measures by sources of income. For the US, they find that female earnings reduced inequality compared to an income distribution without female earnings. Other studies found similar effects, e.g., Del Boca and Pasqua (2003) for Italy, and Campos-Vázquez et al. (2012) for Mexico. Pasqua (2002) uses the ECHP to analyse to what extent country differences in inequality can be attributed to differences in female employment rates. The author shows that the effect on inequality depends mainly on the employment rate. In order to decompose the contribution of female earnings to income inequality, the common approach in these studies is a counterfactual distribution in which women have zero earnings.

1.3 Methodological approach

In this section, we describe all necessary steps to analyse the consequences of the gender wage gap for disposable income and labour market behaviour. The first step is a quantile decomposition of hourly wage differences between men and women along the wage distribution. In the second step of our analysis, we simulate counterfactual gross hourly wages for women in which the unexplained gap is closed. The third step is the simulation of disposable income. In a final step, we go beyond a static analysis and allow for labour supply adjustments to take place.

1.3.1 Wage model and decomposition method

Our starting point is a standard decomposition of wage differences between women and men along the wage distribution. Departing from the work of Oaxaca (1973) and Blinder (1973) – hereinafter OB –, these decompositions aim at separating observed differences in wages (in our case between women and men) into a component that can be explained by differences in labour market endowments and another component that can be explained by differences in the returns to those labour market endowments.[2]

[2]See Fortin et al. (2011) for an in-depth discussion of the assumptions underlying decompositions of wage differentials.

The literature refers to these two components as explained and unexplained factors. We follow Machado and Mata (2005) and apply the principle of the OB decomposition to differences in wage quantiles between women and men. Observed differences between male wages w_m and female wages w_f at the θ-percentile can be expressed as:

$$\hat{Q}_\theta(w_m) - \hat{Q}_\theta(w_f) = \left[\hat{Q}_\theta(w_m) - \hat{Q}_\theta(w_C)\right] + \left[\hat{Q}_\theta(w_C) - \hat{Q}_\theta(w_f)\right] \tag{1.1}$$

where w_C stands for a counterfactual wage distribution. The counterfactual distribution is not observed and we have to make an assumption about its shape. For the case in which the choice of counterfactual wage distribution consists of women's labour market characteristics being paid as if they were men, the first term on the right hand side is the component explained by differences in covariates and the second term is the unexplained component – at each quantile of the wage distribution.[3]

We follow the algorithm proposed by Melly (2006) in order to obtain estimates for unconditional quantiles that are consistent with our conditional quantile regression model.[4] First, we estimate the conditional quantile regression coefficients β_θ separately for men and women for a grid of 1000 percentiles, $\theta \in (0, 1)$, from the model:[5]

$$Q_\theta(w_g|X) = X_g\beta_{g,\theta} + u_g,\ g = f, m\ , \theta \in (0, 1) \tag{1.2}$$

where X stands for relevant labour market characteristics entering the wage model, $g = f$ stands for females and $g = m$ for males.

Unconditional quantiles of $\widehat{w}_m$, $\widehat{w}_f$ and $\widehat{w}_C = X_f\widehat{\beta}_m$ are then computed as:

$$\hat{Q}_\theta(w_g) = \inf\left\{q : n^{-1}\sum \hat{F}_w\left(q|X_i\right) \geq \theta\right\}, \theta \in (0, 1) \tag{1.3}$$

[3]Alternatively, the counterfactual wage distribution could consist of men's labour market characteristics being paid as if they would be women, in which case the correct interpretation of the two components would be the other way around.

[4]As explained in Melly (2006), this procedure is numerically identical to the procedure proposed by Machado and Mata (2005) when the number of simulations used approaches infinity.

[5]We would ideally like to also account for selection issues in the quantile regression set-up. Unfortunately, the main estimator available for this purpose, developed by Buchinsky (1998), is only consistent when the slope coefficients are equal for all quantiles or when selection is randomly determined (see, for instance, Huber and Melly, 2012).

where inf$\{\ldots\}$ picks the smallest value for which the condition in curly brackets is true. w_g stands for any of the three wage distributions $\widehat{w}_m$, $\widehat{w}_f$, $\widehat{w}_C$ and $\hat{F}_w(q|X_i)$ is the cumulative density of wages conditional on X and can be computed as $\hat{F}_w(q|X_i) = \sum 1\left(X_i\hat{\beta}(\tau_j) \leq q\right)$. This way, we obtain the characterization of the three wage distributions required in Equation (1.1) and can decompose into the gender wage gap along the wage distribution. Point-wise and uniform standard errors of the overall gap as well as its two components can be computed via bootstrap as suggested in Chernozhukov et al. (2013).

1.3.2 Constructing individual counterfactual wages

In order to examine the distributional effect of the gender wage gap (and its components), a counterfactual wage for each woman in our sample is needed. However, the method described above results in the characteristics of three wage distributions (female, male and counterfactual) but not in counterfactual wages for each female observation. The mapping of these results to individual wages necessarily requires an additional assumption imposing rank invariance. At this stage, we have in principle two options: preserving the rank in the conditional or the unconditional wage distribution.[6]

Assuming rank invariance in the unconditional wage distribution implies that the gap to which each woman is exposed only depends on her rank in the observed distribution of gross hourly wages. Furthermore, under this assumption the unexplained gap is only job-specific to the point that different jobs pay wage rates that are in different segments of the distribution. This procedure can be understood as adding a mark-up to observed (predicted) female wages that "closes up" the gap found in the decomposition exercise. Concretely, each observation is assigned the counterfactual wage rate corresponding to its quantile in the counterfactual distribution. The assumption of rank invariance in the unconditional wage distribution allows to separately close the unexplained, explained

[6]Note that this would also be the case if working with a mean decomposition. In this case, preserving the conditional rank would translate into computing female counterfactual wages as a regression-based prediction (with the beta coefficients of the male wage regression) while preserving the unconditional rank would boil down to adding the estimated gap to each status quo (observed or predicted) female wage.

and overall wage gap, thereby preserving a great deal of flexibility.

Alternatively, assuming rank invariance in the conditional wage distribution implies that the unexplained gap to which women are exposed will depend on their labour market relevant characteristics. In this case, counterfactual female wages will be computed as a prediction, using the β_θ-coefficients of the male (female) wage estimation applied to the values that the model covariates take on for each female (male) observation. This alternative yields counterfactual wages that respond to a hypothetical exercise: female (male) wages if women (men) would be paid as if they were men (women).

Whereas the assumption of rank invariance in the unconditional wage distribution allows to separately/sequentially close the unexplained, explained and overall wage gap, it also suggests that the used reference wage distribution is the "true" distribution (the wage distribution that would prevail in absence of labour market discrimination, which cannot be determined within the decomposition exercise). On the contrary, assuming rank invariance in the conditional wage distribution allows to abstract from a true reference wage distribution but restricts the possible counterfactuals to "women paid as men are currently paid" (or, vice versa, "men paid as women are currently paid"). This renders it impossible to trace the net financial consequences from the explained and the overall gross gender wage gap. Although our measure of the net gender pay gap (derived in section 1.3.4) is technically compatible with both approaches, we believe that maintaining the rank in the unconditional wage distribution suits our research question and motivation for this paper best. Under both assumptions, counterfactual wages of men (the reference group) equal their status quo wages.

1.3.3 Simulation of net household income

In order to simplify the notation, from now on it will be assumed that the chosen reference wage distribution is that of men, and consequently counterfactual wages are being computed for women. In what follows, index i denotes individuals whereas index j denotes households. The superscript C stands for counterfactual.

On the basis of the counterfactual wage rates calculated above, we can derive gross monthly earnings for each individual under the status quo and the counterfactual scenario, which we denote $E_{i,f}$ and $E^C_{i,f}$, respectively. To this end, we multiply the (exponentiated) wage rate times observed working hours. The difference between the two can be expressed as $\Delta E^f_i = E^C_{i,f} - E^f_i = \left(exp\left(w^C_i\right) - exp\left(w_i\right)\right) h_i$.

The next step is computing households' net income under status quo and counterfactual earnings. Note that this is a necessary step, since under most tax-benefit systems in the world it is not possible to compute an individual's disposable income without applying a sharing rule to the previously computed household net income (except in the case of childless single-adult households). This is a data intensive step, because it requires very detailed information on the household (other household members, their demographic characteristics, their labour-market status, other income sources, among others).

We denote household net income as $N(E^C_j,\, X_j)$ (counterfactual) and $N(E_j,\, X_j)$ (observed), where E_j are gross labor earnings for household j and X_j is a vector of household characteristics (e.g., marital status, number of children) relevant for the tax-benefit system, which does not change between the status quo and the counterfactual scenario. Since $N(\cdot)$ is a highly complex non-linear function capturing the rules of the tax-benefit system, household net income is obtained via microsimulation.[7]

1.3.4 The net gender pay gap

This section presents our measures of the net gender wage gap. In the gender wage gap literature, it is often the case that gender differences are expressed in terms of women's relative wages (to men). In the present exercise, however, we are interested in the net financial consequences of women' relative income to a reference income that does not include the unexplained gross wage gap. As discussed above, in order to do so we need to shift the analysis level to the household. As a result, it is no longer meaningful (nor possible) to frame the resulting income gaps in terms of "relative incomes". Instead, our measures of the net pay gap will be based on the difference in terms of disposable

[7]We use the tax-transfer-simulation-model STSM for Germany. Based on variables drawn from the GSOEP, gross earnings, the taxable income, income taxes, all important transfers and finally the disposable net income can be derived at the household level. See Appendix A.3 and Steiner et al. (2012) for a detailed description of the model.

income that each woman would experience should her hourly wage gap be closed. Table 1.1 provides an overview of the measures presented below.

Net gender pay gap – in absolute values

We can express the difference in disposable income, ΔN_j, as:

$$\Delta N_j = N(E_j^C,\, X_j) - N(E_j,\, X_j) \tag{1.4}$$

From the static perspective, i.e., without labour supply adjustment, the only change between the status quo and counterfactual scenario is the increase in women's wages. Therefore, any change in net household income, ΔN_j, can be fully attributed to the unexplained component of the gender wage gap.[8]

Net gender pay gap – relative to household income

Starting with the absolute difference ΔN_j, it is straightforward to obtain a measure of the (standardized) net gender gap relative to household net income:

$$\Omega_j = \frac{N(E_j^C,\, X_j)}{N(E_j,\, X_j)} - 1 \tag{1.5}$$

Ω_j describes the relative change of household net income following an increase of female gross wages for each household. The denominator of Equation (1.5) includes all income components of the household. Hence, the relative gender income gap is smaller the higher other income components are. There are also interactions with the tax-benefit system. For example, conditional on the position in the gross hourly earnings distribution, the progressivity of the tax-benefit system leads to net gender wage gaps that decrease the more hours women work. This is true for women both in an individual as well as in a joint taxation regime. Under joint taxation the net gender wage gap will increase slower with hours for married than for non-married women. In addition, fol-

[8] In section 1.3.5 we allow for adjustments in labour supply resulting from the existence of the gap. In this case, the difference in disposable income can capture – in addition to increased female hourly wage rates – changes in labour supply of both women and men. However, due to the fact that men's cross-elasticities of labour supply are very small, most of the change in disposable income is due to changes in female wages and women changing working hours or labour market participation.

lowing this definition of the net gender gap, a woman exposed to a very high marginal tax will have a very small net gender gap independently of how big the unexplained gap allocated to her is.

Net gender pay gap – relative to female income

Beyond the household measure, we quantify the relative change with respect to the female component of household income. This measure is better comparable to the gross wage gap which is also derived at the individual level. However, for non-single households determining the denominator in Equation 1.5 is not straightforward, since it requires further information on how net income is distributed within the household. This is because many elements of the tax-benefit system are evaluated and determined at the household level. Moreover, some income components can only be attributed to the household and not to its individual members. As a consequence, any measure of individual net income necessarily implies a sharing rule within the household. A measure of the net gender gap in terms of female income would be constructed as:

$$\omega_i = \frac{\delta_i^C}{\delta_i} \frac{N(E_j^C,\, X_j)}{N(E_j,\, X_j)} - 1 \tag{1.6}$$

δ and δ^C represent the value of the sharing rule for the female household member under the status quo and the counterfactual female wage distribution, respectively. δ is bounded between zero and one, as it represents the proportion of household net income available to the female spouse. For single households, δ equals one in both cases, and ω_i equals Ω_j.

Given that we do not observe the true sharing rule in the households of our sample, in our application we use two opposite concepts of sharing rules: (1) perfect income pooling within the household and (2) proportional sharing of household net income. Under income pooling, female net income in couple households amounts to half of the overall household net income (regardless of the (in)existence or level of labour earnings of any other household member). In this case, $\delta = \delta^C = 1/2$ and as a result $\omega_i = \Omega_j$. Under proportional income sharing, we let the share of female labour earnings (over total household labour earnings) determine the share of net household income available to the female spouse. Thus:

$$\delta_i = \frac{E_{i,f}}{E_{i,f} + E_{i,m}} \tag{1.7}$$

$$\delta_i^C = \frac{E_{i,f}^C}{E_{i,f}^C + E_{i,m}^C} \tag{1.8}$$

where $E_{i,k}$ denotes labour earnings earned by women ($k = f$) and men ($k = m$) within a household j.

While none of these two extreme sharing rules are (expected to be) realistic, we have chosen them so as to obtain a lower and upper bound of the net gender pay gap at individual level[9]. The assumption of perfect income pooling, although it has been questioned in the literature (see, e.g., Browning et al., 1994; Lundberg et al., 1997), yields a lower bound for the net gap because it fully "socializes" the gender gap within the household. Contrarily, the assumption of proportional sharing of household net income implies that women alone bear the whole financial consequences of the gap – providing thus an upper bound for our estimate of interest.

Given that, by construction, $E_{i,f}^C > E_{i,f}$ and $E_{i,m}^C = E_{i,m}$, the quotient δ_i^C/δ_i will equal one for single women and greater than one for partnered women. Thus, the net gap under income pooling will be by construction equal or smaller than the net gap under proportional sharing.

1.3.5 Labor supply effects of the net gender pay gap

We now go beyond a static wage gap decomposition and show how we can use a structural labour supply model to estimate net pay gaps that allow for behavioural changes. These estimates take into account labour supply adjustments derived from the existence of the gender wage gap itself. This relaxes the usual assumption of empirical studies on gender wage gaps that working hours and participation are not modified by the existence of a wage differential. Furthermore, such a model allows us to estimate aggregate labour supply responses from the gap in the intensive as well as in

[9]The estimation of the actual sharing rule is beyond the scope of this paper.

the extensive margin. The adjustment in terms of working hours is computed for each household by comparing the predicted number of working hours under the status quo and the counterfactual female wages. To calculate participation effects, we compare the probabilities of choosing working hours combinations.

We use a structural labour supply model as proposed by van Soest (1995) to estimate labour supply elasticities. In this model, households choose labour supply by maximizing a utility function that includes net income, leisure and other household characteristics (such as number of children in the household, age, health, etc.). The household utility level depends on the combination of working hours chosen. For couple households, men and women are assumed to choose among k and l alternatives of working hours (including non-employment in both cases), so that the household choice set includes $k \times l$ combinations of working hours. For single households, the choice set reduces to the own working hours categories. Under the assumption that the error term in the utility function is type I extreme-value distributed, the probability of choosing each given combination of working hours can be estimated via a conditional logit model (McFadden, 1974). For more details, see Appendix A.3 and Steiner et al. (2012).

Once we have estimated the discrete choice model explained above, – both under status quo and counterfactual wage distributions – we can proceed to calculate household net incomes that allow for labour supply adjustments. To this end, we multiply the estimated probabilities of choosing the diverse working hours categories times the corresponding net income for each category – both under the status quo and the counterfactual wage distribution. The net pay gaps with behavioural reactions can now be expressed as:

$$\widetilde{\Omega}_j = \frac{N(\widetilde{E}_j^C,\, X_j)}{N(\widetilde{E}_j,\, X_j)} - 1 \tag{$\widetilde{5}$}$$

$$\widetilde{\omega}_i = \frac{\widetilde{\delta}_i^C}{\widetilde{\delta}_i^S} \frac{N(\widetilde{E}_j^C,\, X_j)}{N(\widetilde{E}_j,\, X_j)} - 1 \tag{$\widetilde{6}$}$$

In this formulation, adjustments in labour supply of men deriving from changes in female wages would theoretically show up in $\widetilde{\Omega}_i$ and $\widetilde{\omega}_i$. This means it is no longer possible to claim that all changes in $N(\cdot)$ can be exclusively traced back to the change in female wages (as it was the case in the static calculation). However, cross-elasticities of men are empirically very low and therefore not likely to play a major role (Steiner and Wrohlich, 2004).

Table 1.1: Summary of different representations of the gender wage gap in terms of disposable household income

Measure	Description
ΔN_j	*Absolute difference* in disposable household income between observed and counterfactual wages without labour supply adjustments
Ω_j	*Relative difference* in disposable household income between observed and counterfactual wages without labour supply adjustments (range between 0 and 1). Measured at the *household level.*
ω_i	*Relative difference* in disposable household income between observed and counterfactual wages without labour supply adjustments (range between 0 and 1). Measured at the *individual level.*
$\widetilde{\Omega}_j$	*Relative difference* in disposable household income between observed and counterfactual wages with labour supply adjustments (range between 0 and 1). Measured at the *household level.*
$\widetilde{\omega}_i$	*Relative difference* in disposable household income between observed and counterfactual wages with labour supply adjustments (range between 0 and 1). Measured at the *individual level.*

1.4 The unexplained net gap in West Germany

In this section we estimate the net wage gap for Germany applying the framework detailed in the previous section. Germany is a good example to study the distributional and behavioural consequences of the gender wage gap for several reasons: the gender pay gap in Germany is relatively large and persistent. In 2010, the median wage gap of full-time working women was about 22 percent and has only slightly narrowed since 2000. Comparing female employees working full-time, several studies also suggest that the wage gap varies across the wage distribution in Germany (OECD, 2012; Christofides et al., 2013; Arulampalam et al., 2007). Secondly, the share of women working part-time is particularly high in Germany: in 2010, nearly 40 percent of all employed women worked part-time (OECD, 2012, p.161). Thirdly, the German tax-benefit system has an implicit gender dimension given its progressive tariff and the joint taxation regime for married couples. In what follows, we show the consequences of the unexplained component of the gross hourly wage gap in terms of disposable income.[10]

We distinguish four different types of households to study the effect of the gross wage gap. We group women along household characteristics (partner in household and existence of children) that were identified in the literature to be important for either the level and distribution of gross wages, the level and distribution of working hours and participation behaviour, and net household income: First of all, we differentiate between couples and women living alone because of different labour supply behaviour. Estimated labour supply elasticities for women living in couples are usually higher than for single women (see, e.g., Bargain et al., 2014). Furthermore, most couples are married and therefore taxed jointly which – compared to individual taxation – leads to lower average tax rates but higher marginal tax rates for the second earner, which is most often the female spouse. The high marginal tax rate for women entails negative labour supply incentives (e.g., Crossley and Jeon, 2007). Furthermore, it also increases incentives to take up lowly paid marginal employment which is exempted from social security contributions and tax free. Secondly, we differentiate women with and without children. Children have a negative impact on labor supply and are one important factor that leads to disadvantageous wage growth for women (e.g., Anderson et al., 2002, 2003; Meurs et al., 2010).

[10] We have chosen the unexplained component of the gender wage gap as the core of our application because of its relatively large magnitude in West Germany and since it is the main focus of many studies in this field and can be cautiously interpreted as a proxy for discrimination. It is difficult to identify and quantify discrimination based on the estimated residuals. If the unobservables are correlated with characteristics, the estimated betas are biased (Altonji and Blank, 1999). Even if there is no endogeneity, the residuals might include unobserved group differences related to productivity which would bias the estimated discrimination (Blau and Kahn, 2007). Conversely, observed characteristics could also be influenced by discrimination, e.g., vertical and horizontal occupational segregation.

1.4.1 Data and Descriptive Statistics

Our study is based on data from the GSOEP which is a representative longitudinal micro database that provides a wide range of socio-economic information on private households in Germany. In 2012, the sample included about 21,000 respondents living in more than 12,000 households.[11] The GSOEP provides information about employment status, earnings and working hours of individuals. Moreover, it includes detailed income information on the individual and household level and other demographic characteristics. One important feature of the data is the availability of information at the household level that is relevant for the tax-benefit system (e.g., earnings of the spouse, dependent children, other household income), which is not available in administrative data.

Sample for the Wage Decomposition

Our estimation sample for the wage regression comprises all working individuals aged between 15 and 64 years with residence in West Germany, except for the self-employed, people in vocational training or military service, students, and pensioners (in line with the literature). We focus on West Germany because the wage gap and female labour supply differ markedly between regions and the gap is particularly large in West Germany. We excluded observations earning less than 1.5 euro and more than 100 euro per hour, as well as observations with missing data in the variables entering the wage model.[12] We have not imposed any exclusion based on working hours, so that our estimation sample includes part-time and full-time workers. We pool data for the years 2008 to 2012 in order to have enough observations for our richly specified wage model. This yields a total number of 29,975 observations, of which 14,949 are females and 15,026 are males.

As Table 1.2 shows, women's hourly wages are on average about 15 euro – amounting to only 77 percent of male wages. It also shows that women work, on average, 11

[11]A detailed description of the GSOEP is provided by Wagner et al. (2007); more information is available at `http://www.diw.de/en/soep`.

[12]Variables included are polynomials in age, experience, and tenure, and sets of dummies that control for public sector, education, occupation, industry, firm size, and year effects.

hours per week less than men. Lower wages and lower working hours result in an average gross earnings difference of 1,355 euro per month. This represents 42 percent of male earnings. The unconditional difference of hourly wages increases in absolute and relative terms along the distribution of monthly earnings, reaching about 8 euro or 25 percent of the male average at the top quintile. In terms of working hours, the pattern we observe is the opposite. The difference is the highest at the bottom quintile of the distribution, where women work on average 22 hours per week less than men. Whereas working hours of men stay mostly steady around 40 hours per week along the gross earnings distribution, female work hours increase steeply, reducing the average difference in the top quintile to 5 hours per week.

Table 1.2: Average hourly wages and working hours by quintiles of gross monthly earnings

Quintiles of gross monthly earnings	Women		Men	
	Wage	Hours	Wage	Hours
1	8.36	13.55	9.91	35.97
2	11.72	25.87	14.53	40.37
3	13.60	33.23	17.50	40.76
4	16.37	36.80	21.66	41.04
5	23.95	38.98	31.84	44.20
Average	14.80	29.68	19.03	40.43
Gross monthly earnings	1,896		3,252	

Notes: Quintiles were calculated separately for the distribution of monthly gross earnings for men and women. Wages in euro. Hours worked per week. Weighted observations. Years 2008 through 2012.

Source: SOEP.v29.1, own calculations.

The difference in the distribution of working hours between men and women is also illustrated in Figure 1.1 below. The upper panel illustrates the differences in the distribution of hourly wages between men and women. The distribution of hourly wages of men is shifted to the right compared to the distribution of women's wages. The lower panel shows the differences in the distribution of gross monthly earnings. Since men generally work longer hours and earn higher wages, the average difference in monthly earnings is larger than the difference in hourly wages. But we also observe that monthly earnings of women follow a bimodal distribution. The bunching of women around 400

euro per month can be explained by the existence of marginal employment. Up to this income threshold, employees do not pay taxes or social security contributions. Women find themselves in thiy type of employment more often than men.
Descriptive statistics of the variables entering the wage model can be found in Tables A.1 and A.2 in the Appendix. In our sample, men are on average slightly older than women, have significantly more work experience and longer tenure. The share of women working in the public sector is higher than for men. There are no large differences in the level of education.

Sample for the estimation of net household income and labor supply effects

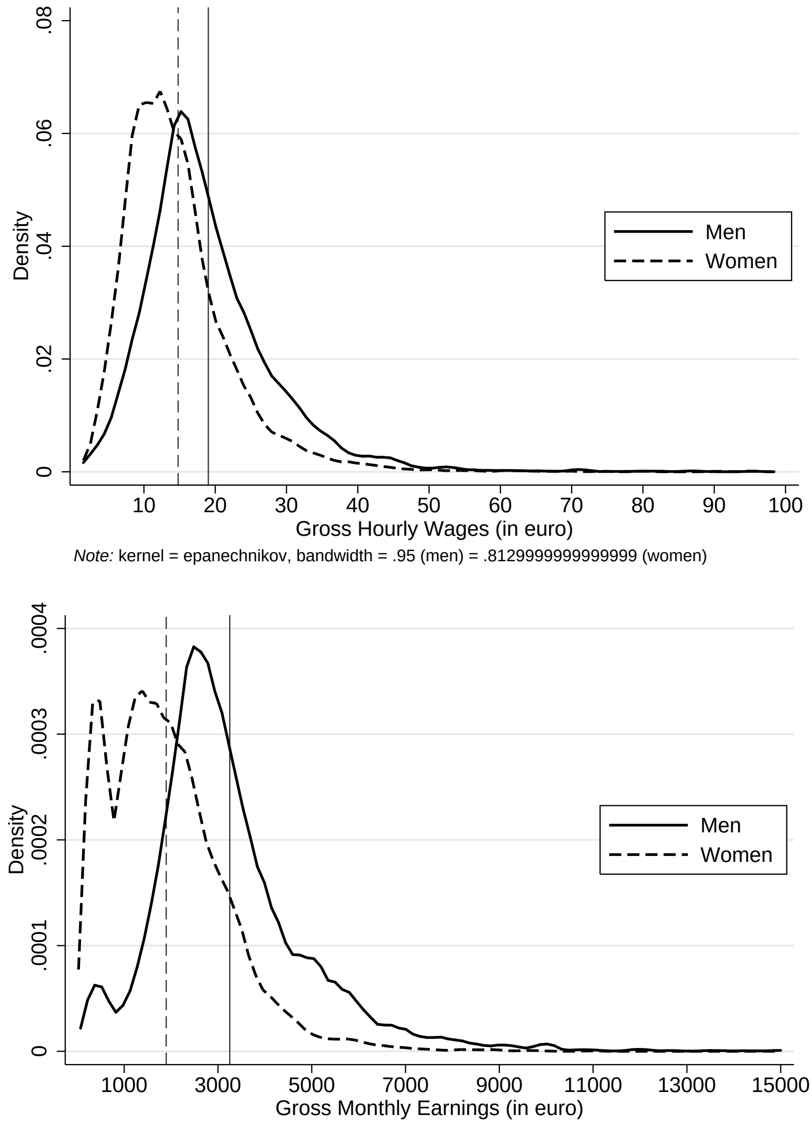
The simulation of net income and the estimation of labor supply is based on the year 2010. We choose only one year in order to keep the rules of the tax and benefit system constant. That is, all variation in the income distribution between the observed and the counterfactual scenarios are related to our simulation and not induced by changes in tax rates, social security contributions or transfers. For this part of the analysis, we restrict our sample to households with a woman of working age (working or not working) and thereby those households that may be directly affected by the existence of a gender wage gap. A detailed description of the imputation procedure of wages for non-working women can be found in Appendix A.2. This leaves us with a total of 3,063 households, of which 2,174 are couple households and 889 are single households.

1.4.2 Gross gender wage gap: decomposition results

In this section, we report the results of our decomposition exercise. It is based on counterfactual female wages that would result from rewarding female labour market skills with male skill prices. In other words, we take male wages as the reference return to labour market characteristics.[13]

[13]Besides the straightforward alternative of using the betas estimated for female wages as reference returns to labour market skills, further alternatives explored by the literature – though only in the context of mean decompositions – consist of reference returns to skills constructed with the variation in the wage regressors of the two groups (e.g. Fortin, 2008; Neumark, 1988) and using the shares of women and men in the population to construct a weighted distribution (Oaxaca and Ransom, 1994).

Figure 1.1: Kernel densities of gross hourly wages and gross monthly earnings (2008-2012)

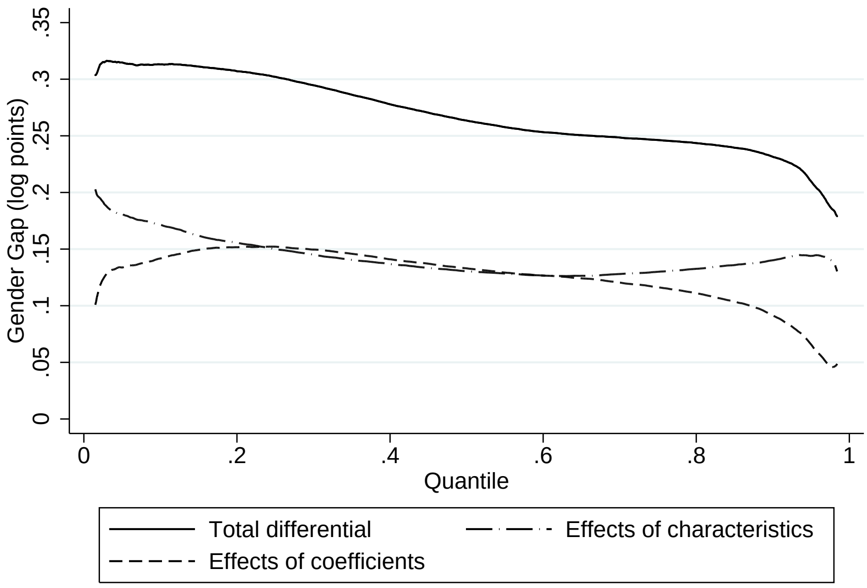


Note: kernel = epanechnikov, bandwidth = .95 (men) = .8129999999999999 (women)

Note: kernel = epanechnikov, bandwidth = 160.432 (men) = 157.085 (women)

Source: SOEP.v29.1, own calculations.

Figure 1.2: Decomposition results



Source: SOEP.v29.1, own calculations. Weighted observations.

Figure 1.2 presents the results of the decomposition exercise (see Table A.5 in the Appendix for the exact point estimates and standard errors). The y-axis depicts the absolute gap between the quantiles of two logarithmic gross hourly wage distributions and is approximately the relative gap between the quantiles of the exponentiated distributions.

The overall gap shows a decreasing shape from nearly 32 log points in the lower tail of the distribution down to approximately 18 log points at the upper end. The unexplained component displays a slightly inverted u-shape. It starts at about 10 log points and increases up to 15 log points at the twentieth percentile. From there on, it decreases steadily and reaches the 5 log points at the top of the distribution. The explained component has a weak u-shape. It starts from about 20 log points, decreases until the third quintile to 12 and increases again to nearly 15 log points in the 95^{th} percentile.

A comparison to previous findings from other studies is difficult because samples, data, and year of the analysis are in general different. In terms of the raw gender wage gap, some studies find a similar decreasing trend (e.g. Fitzenberger and Kunze 2005; Hübler 2005; Heinze and Wolf 2010). Nevertheless, other studies like Arulampalam et al. (2007), Antonczyk et al. (2010), and Christofides et al. (2013) show either a u-shaped overall gap or a gap that increases over the distribution.

As explained in Section 1.3.3, the next step consists of assigning a specific value for the unexplained gap to each woman's observed wage in the sample. In order to do so, we assume rank invariance in the unconditional hourly wage distribution. Allocating an unexplained gap to each individual woman enables us to conduct a distributional analysis of the unexplained gap by household types. However we do not find large differences by household type with respect to the unexplained gap in hourly wages. This is consistent with our observation that households are distributed evenly across observed hourly wages. Note that the height of the gap is solely determined by the rank in the observed wage distribution.

We can now turn to analysing how the unexplained gap is distributed across gross monthly earnings of women. Empirically, we do observe women with high hourly wage rates who have low monthly wages because they work few hours. Alternatively, we also observe women with very low wage rates who still reach female median gross monthly earnings by working many hours. To the extent that women are ranked differently in the hourly wage distribution than in the monthly earnings distribution, the shape of the unexplained gap across the monthly earnings distribution is an open empirical question.

We assume the number of working hours to be constant between the status quo and the counterfactual scenario and multiply the counterfactual hourly wage rates by the observed number of working hours for each woman in our sample. Table 1.3 depicts our results. The unexplained gap as percentage of gross monthly earnings is decreasing by earnings quintiles. It is lowest at the upper end of the earnings distribution for each type of household. This corresponds to the results of the decomposition which showed a decreasing gap at the top of the wage distribution. However, in absolute values, the gap is increasing. The unexplained gap amounts to 60 euro on average in the lowest quintile to 372 euro in the top quintile. We find the wage gap in most quintiles to be smaller for households with children.

Table 1.3: Unexplained gross monthly gaps (in % and absolute values) by quintiles of gross earnings and household type

Quintiles	Couples		Singles		All
	without children	with children	without children	with children	
% of female monthly gross earnings					
1	15.1	14.6	14.6	15.1	14.8
2	15.0	13.9	15.5	15.1	14.6
3	14.2	12.9	15.3	14.2	14.2
4	13.3	11.4	13.7	12.9	13.2
5	10.6	8.8	10.6	10.6	10.4
Average	13.6	13.3	13.3	14.0	13.4
Absolute difference (in Euro)					
1	61	60	62	51	60
2	160	143	171	161	154
3	244	219	262	251	244
4	322	270	332	302	318
5	379	346	373	356	372
Average	234	157	297	210	229

Notes: Unexplained gaps represented as % of gross monthly earnings. Quintiles of the overall predicted female monthly gross earnings distribution in 2010.

Source: SOEP.v29.1, own calculations. Weighted observations.

Table 1.4 shows that for a given quantile of monthly earnings, households with children reach an average hourly wage that is above the wage of households without children. In other words, they work fewer hours and reach the same income as their childless counterparts in the same quantile. Given the results of the decomposition, higher wages correspond to lower unexplained wage gaps (see Table 1.3).

Table 1.4: Hourly wages and working hours by earnings quintiles and household type

Quintiles	Couples		Single		All
	without children	with children	without children	with children	
Median Hourly Wages					
1	7.8	8.2	7.5	7.8	7.8
2	11.4	13.2	9.2	10.8	11.5
3	13.8	15.1	11.0	14.4	12.9
4	14.9	19.9	14.2	16.6	14.9
5	19.9	25.5	19.5	19.2	20.1
Average	13.6	13.5	14.2	13.4	13.8
Median Working Hours per Week					
1	9.1	8.6	8.6	8.6	8.6
2	25.7	15.6	27.7	20.2	25.7
3	26.7	25.7	38.8	27.7	27.7
4	38.3	25.7	38.8	34.1	38.8
5	38.8	38.3	38.8	38.8	38.8
Average	26.7	15.6	38.8	27.7	27.7

Notes: Median hourly wages in Euro. Quintiles of the overall predicted female monthly gross earnings distribution in 2010.

Source: SOEP.v29.1, own calculations.Weighted observations.

1.4.3 Static net gender pay gap: the role of the tax-benefit system

Due to the tax-benefit system, the change in gross monthly earnings will have very different consequences in terms of net household income depending on the characteristics of each household.

Table 1.5 presents results for the net gender pay gap, Ω_j, at the household level. Ob-

Table 1.5: Static net gender pay gap by income quintiles and household type

Quintiles	Couples		Singles		All
	without children	with children	without children	with children	
% of household disposable income, Ω_j					
1	6.6	3.1	7.5	4.5	5.6
2	4.4	2.6	9.9	6.7	5.8
3	3.4	2.1	9.7	6.8	5.0
4	4.8	2.6	9.3	6.4	5.8
5	3.9	1.9	7.1	2.7	4.0
Average	4.4	2.4	8.9	5.5	5.2
Absolute difference (in Euro), ΔN_j^{hh}					
1	96	86	92	75	88
2	107	66	123	125	100
3	107	71	151	139	110
4	142	99	172	162	141
5	146	78	186	152	138
Average	126	79	141	102	116

Notes: Gaps presented as % of net household income. Quintiles of overall equivalized net household income. Weighted observations.

Source: SOEP.v29.1, own calculations.

servations are sorted by quintiles of household equivalised[14] disposable income. On average, the unexplained gap accounts for 5.2 percent of households' net income. This figure stays fairly constant along the income distribution.

If we look at the results disaggregated by household types, differences in net gender gaps along the income distribution become more pronounced. For couple households without children the net gender gap is highest at the lowest income quintile – more than 6 percent – and decreases at higher income quintiles, where it amounts nearly 4 percent. For single households without children, the net gender gap is highest at the middle of the income distribution, with gaps ranging between 9.3 and 9.9 percent, and decreases at both distribution ends, though not as strongly as for couple households. Households with children show lower average gaps in each quintile.

On average, household equivalised income increases by about 116 euro per month in our counterfactual scenario. We find that the absolute difference increases monotonically across income quintiles for households without children and singles but shows a heterogeneous pattern for couples with children. This pattern can be explained by a larger intra-household earnings inequality of households with children. In particular, the share and level of female earnings in couple households with children are smaller than in households without children.

As explained in Section 1.3, a point-identification of the net unexplained gap in terms of female net income for couple households is not possible without information on the actual sharing rule within the household. However, by choosing opposite sharing rules, we can produce bounds for the estimate of interest. These are reported in Panel A of Table 1.6. Our calculations suggest that the average net pay gap for women in couple households lies between 5.2 and 9.7 percent depending on the actual sharing rule of the household. The closer the household is to income pooling (proportional sharing), the more (less) the financial consequences of the gap are internalized within the household and the smaller (bigger) the burden faced by the women. In other words,

[14] We use the new OECD-equivalence scale that assigns a value of 1 to the first adult in the household, a value of 0.5 to each further member of the household above the age of 14, and a value of 0.3 to children in the household below the age of 14.

Table 1.6: Different measures of the average net gender pay gap by household type

	Couples		Singles		All
	without children	with children	without children	with children	
PANEL A: Measures without labor supply adjustment					
Household level					
Ω_j	4.4	2.4	8.9	5.5	5.2
Individual level					
$\omega_i^{inc\,pool}$	4.4	2.4	8.9	5.5	5.2
$\omega_i^{ind\,inc}$	10.7	10.0	9.2	6.0	9.7
PANEL B: Measures with labor supply adjustment					
Household level					
$\widetilde{\Omega}_j$	5.5	3.3	9.9	6.3	6.2
Individual level					
$\widetilde{\omega}_i^{inc\,pool}$	5.5	3.3	9.9	6.3	6.2
$\widetilde{\omega}_i^{ind\,inc}$	14.32	14.20	10.32	6.9	12.6

Notes: Gaps presented as % of net household income in the case of Ω_j and $\widetilde{\Omega}_j$, and as % of female individual income in the case of ω_i and $\widetilde{\omega}_i$. Weighted observations.

Source: SOEP.v29.1, own calculations.

to the degree that households pool resources, the risk of wage discrimination is also shared and mitigated within the household.

1.4.4 Net gender pay gap with labor supply adjustments

The estimation of a structural model of labour supply enables us to take the analysis a step further and to present results on net gender pay gaps that take into account changes in labour supply. These are results that allow for adjustments in terms of labour supply of households deriving from the existence of the unexplained gap. Table A.6 in the Appendix reports the estimates from the conditional logit model underlying our results, and Table A.7 reports the associated labour supply elasticities.
Panel B of Table A.8 replicates the results of the previous section allowing for labour supply adjustments.[15] In terms of household net income, the gaps with adjustments are slightly higher than without – suggesting that behavioural reactions on the working hours margin may raise the net financial consequences of the unexplained gap. In terms of female net income, the more individualistic the sharing rule is, the higher the difference when we allow for labour supply effects. Furthermore, the effects are considerably bigger for women in couple households than in single households (consistent with the smaller elasticities of the latter, see Table A.7 in the Appendix). Thus, the upper bound for females in couple households goes up to 14 percent when allowing for labour supply effects to take place, a four percentage points difference with respect to the static scenario.

Table 1.7 reports the aggregate labour supply effects deriving from our counterfactual female wage distribution. According to our results, the existence of the unexplained component of the gender wage gap corresponds to an increase of labour market participation of about 109,000 individuals (most of them women). In terms of working hours supplied, the existence of the unexplained gap accounts for a reduced labour supply of ca 251,100 full-time equivalents (taken to be 40 hours per week). However, most of the change would affect individuals who have already taken up employment.

[15] For the sake of presentation, results on net pay gaps with behavioral adjustments by income quintiles can be found in Table A.8 in the Appendix.

Table 1.7: Labor supply effects

	People	Working hours (Full-time equivalents, in 1000)		
	(in 1000)	Total change	Additional hours	Newly employed
Couple households				
w/o children				
Women	36.5	101.5	77.4	24.1
Men	1.7	2.6	3.9	-1.3
with children				
Women	42.7	63.5	42.8	20.7
Men	1.9	2.9	1.2	1.7
Single households				
w/o children	18.8	62.0	47.6	14.4
with children	7.5	18.6	13.9	4.7
Total	109.1	251.1	186.6	64.4

Source: SOEP.v29.1, own calculations.

1.5 Conclusion

The aim of this chapter was to provide a net measure of the raw gender wage gap as well as its explained and unexplained components. To the best of our knowledge, our study is the first that attempts to relate the outcome of a gross wage decomposition to net household income. It is important to quantify the distributional impact of the gender wage gap in terms of disposable income because of its nature as an inequality measure as well as potential behavioural responses deriving from it. From a policy perspective, this is highly relevant for female labour supply, gender-specific distribution of income within couple households as well as the economic independence of women.

The point of departure for our framework was a quantile regression-based decomposition with which we obtain an estimate of the raw gender wage gap as well as its explained and unexplained components across the whole wage distribution. The construction of individual counterfactual wages enables us to derive the gender gap both in terms of gross monthly earnings and in terms of net household income. This step of the analysis rests on two important assumptions: We assume the prices of labour market skills paid to male individuals to be the reference prices of labour market skills. Whereas the literature has explored some alternatives in the context of mean decompositions, such a technique is missing for quantile-based decompositions. The main consequence for our analysis was that it takes a partial equilibrium perspective. The second assumption is that we need to impose rank invariance in either the conditional or the unconditional wage distribution. In our application we choose the latter, although our methodological framework would be compatible with both of them.

Once we assign a counterfactual hourly wage to each individual, we proceed to compute the resulting household net income via microsimulation. By comparing household net income with counterfactual and observed wages, we can derive the net financial consequences of the gender pay gap at household level. Whereas for single households the net financial consequences at household level equal those at individual level, for couple households we require further information on how the income is distributed within the household. Given that we do not observe the true sharing rule in the our sample, we suggest to use two opposite concepts of sharing rules: (1) perfect income pooling within the household and (2) proportional sharing of household net income. While none of these two extreme sharing rules are (expected to be) realistic, we have chosen them so as to obtain a lower and upper bound of the net gender pay gap at individual level.

In a last step, we estimate a structural labour supply model to estimate net pay gaps with labour supply adjustments. This relaxes the assumption that working hours are not modified by the existence of the gender wage gap. Additionally, we can also estimate aggregate labour supply effects of the gender wage gap.

We exemplify our procedure with an application to the unexplained gap in West Germany. In particular, we derive the effects in terms of disposable income of the unexplained component of the gender wage gap in West Germany. The quantile-based decomposition exercise shows that, in terms of hourly wages, the overall gender wage gap is highest at the lower end of the distribution and decreases with wages. The unexplained component of the gender wage gap follows an inverted u-shape and is similarly distributed for different household types. As expected, both female gross monthly earnings and households' equivalised net income rise in the absence of the unexplained gap. We find that on average the household unexplained net gap lies around 5.2 percent of household's equivalised net income. This is higher for women living in single households (with or without children) than in couple households. Furthermore, within each of these two groups, the net financial consequences of the unexplained gap are higher for households without children. This pattern can be explained by larger earnings inequality of households with children. In particular, the share and level of female earnings in couple households with children are smaller than in households without children. At the individual level, we can only identify bounds for the net pay gap. These range, on average, from 5.2 to 9.7 percent. The more (less) the actual sharing rule approaches perfect income pooling, the lower (higher) the net unexplained gap at individual level. The gender wage gap also has negative labour supply effects. When we allow for labour supply adjustments, we find a higher impact for women living in couple households than in single households, which is consistent with the higher labour supply elasticities of the first group. In the aggregate, we find the unexplained gap to be associated with the labour market non-participation of around 105,500 women and with a reduction of working hours of about 247,600 full-time equivalents.

Chapter 2

The Gender Wage Gap across the Distribution and over Time*

2.1 Introduction

In all countries of the world, men earn - on average - higher wages than women. This differential, known as the gender wage gap, has been decreasing over the last decades, although recently at a lesser pace (e.g. Blau and Kahn 2017, OECD 2017b). Many factors - such as rising female educational attainment and labour market attachment - have contributed to the decrease of the gender wage gap over the last decades, not least changing selection patterns into employment (e.g. Blau and Kahn 2006, Blundell et al. 2007, Mulligan and Rubinstein 2008). In addition, distributional studies about the gender wage gap have revealed a large variation over the distribution, particularly substantial glass ceilings and sticky floors in selected countries (see Chapter 1.2, Albrecht et al. 2003, Arulampalam et al. 2007, Rica et al. 2008). However, in the prevailing context of rising wage inequality as well as a general increase in female labour market participation, selection into employment is likely to differ along the distribution and

* This Chapter is based on joint work with Katharina Wrohlich.

also over time (e.g. Arellano and Bonhomme 2017). As a consequence, the effect that selection into employment has on the gender wage gap may also change across the distribution and over time. It is the aim of this chapter to make use of recent developments in the econometrics literature to analyse selection-corrected gender wage gaps across the distribution and over time.

In particular, we follow Melly and Santangelo (2015)[16] and provide a distributional analysis of the gender wage gap for West Germany in the years 1985 to 2009. Importantly, the method controls for changing patterns of selection into employment across the distribution and over time. The goals of this paper are fourfold: First, we provide a descriptive analysis of the gender wage gap across the distribution and over time, including its heterogeneity by human capital levels and family status. Then we characterise the magnitude and evolution of male and female selection into employment across the distribution and over time. Third, we present estimates of selection-corrected gender wage gaps and, finally, we carry out a decomposition exercise to better understand the trends in the overall gap.

The empirical strategy used to take potential selection effects into account consists of imputing non-realised wages for those individuals who are not full-time employed in a given time period (see Neal 2004, Blau and Kahn 2006, Olivetti and Petrongolo 2008 for similar approaches). Concretely, we use a new econometric method proposed by Melly and Santangelo (2014, 2015) that extends Athey and Imbens (2006)'s changes-in-changes model by accounting for covariates and adapts it to impute non-realised wages which account for both individuals' observable and unobservable characteristics. This method uses information from an individual's realized wage obtained from longitudinal data and assumes the time-invariance of the unobservables conditional on the observables to impute this individual's wage whenever he or she is out of work. This imputation-based approach allows us to characterize female and male selection-corrected wage distributions without having to rely on specific variables as exclusion restrictions.

Our analysis is based on the German Socio-Economic Panel (GSOEP), a rich longi-

[16] Melly and Santangelo (2015) provide an analysis of the gender wage gap for the United States from 1968 to 2008.

tudinal dataset with detailed data on earnings, working hours and individual- and household-level characteristics. We restrict our analysis to full-time employees residing in West Germany over the time period 1985 to 2009. Our results show that in contrast to many other countries, the observed gender wage gap follows a slight u-shaped curve over the wage distribution in Germany, with lower values in the middle and higher values at the bounds. Moreover, we find that Germany's observed gender wage gap has constantly decreased between 1985 and 2009, although the decrease in the 1980s and 1990s was larger than since the beginning of the 2000s. This is in accordance with results for other countries such as the United States (see Blau and Kahn 2017). Conditioning on age, we find the observed gap to slightly increase with age and, for the younger age segments, to decrease over time. Conditioning on education, we find the gap to be highest among those with a basic degree. Furthermore, we find sticky floors for individuals with low educational attainment and glass ceilings for those with medium and high educational attainment. This is in line with the findings of Rica et al. (2008) for Spain.

Our aim is to also quantify the effect of selection into employment for the male and female wage distribution. We find the sample of women working full-time to be strongly positively selected at all points of the wage distribution, and the magnitude of the selection to increase over time. This is consistent with the findings of Biewen et al. (2017) using administrative data for West Germany. The sample of men working full-time is also slightly positively selected, but less so than it is the case for women. However, the magnitude of selection at the lower bound of the distribution has increased steeply from the beginning of the 1990s onwards. These findings occur in a context of constant (full-time) employment rates as well as rising wage inequality (in line with Mulligan and Rubinstein 2008). An analysis of the selection by age and education reveals the different nature of selection in both genders. Whereas selection has gained strength among the youngest male segment over time, the female distribution was much more selected among its oldest segment during the 1980s and 1990s, but does not differ any longer by age segments during the 2000s.

Taking into account selection into full-time employment, this chapter shows that the selection- corrected gender wage gap turns out to be much larger than the observed one. This means that studies failing to control for positive selection into employment are likely to overestimate the convergence of male and female wages over time. After a strong decline in the 1980s and 1990, when the selection-corrected gender wage gap of median wages decreased from 38 to 31 log points, it has been increasing since the early 2000s (amounting to 35 log points in the period 2005 to 2009).

Finally, the decomposition of the gender wage gap according to differences in men's and women's characteristics as well as differences in coefficients shows that the relative importance of coefficients diminishes over time. In other words, characteristics such as education and work experience are rewarded more equally for men and women in the late 2000s than in the beginning of our observation period. Ultimately, our results also suggest that only the closing of gaps in characteristics (particularly full-time working experience) will contribute to a substantial lowering of the gender wage gap in the years to come.

This chapter contributes to the large and growing literature examining differences in male and female wages. Whereas human capital levels (such as educational attainment) accounted for a large share of the gap in the 1980s and 1990s (Blau and Kahn 2017), women have caught up with men's educational achievement in virtually all high-income countries (see Goldin et al. 2006, Becker et al. 2010). However, the gender wage gap widens with age over the life cycle - both for older and younger cohorts - which researchers attribute to marriage and motherhood (see Anderson et al. 2002, Angelov et al. 2016 or Juhn and McCue 2017, among others). Other factors examined in the literature are occupational segregation (Groshen 1991, Fitzenberger and Kunze 2005, Ludsteck 2014), intra-firm gender wage gaps (Heinze and Wolf 2010, Card et al. 2016) as well as labour market institutions such as wage-setting mechanisms (Blau and Kahn 2003), unions (Blau and Kahn 1992, 1996, Booth and Francesconi 2003) and family policies (Christofides et al. 2013, Olivetti and Petrongolo 2017).

Furthermore, the literature has also focused on the effect changing selection into employment has on the gender wage gap. Olivetti and Petrongolo (2008) show that a large share of the cross-country differences in gender wage gaps can be explained by differences in female employment rates. Mulligan and Rubinstein (2008) present a theoretical framework that links increasing positive selection into employment to rising wage inequality and show that convergence of male and female wages in a context of rising wage inequality can be overestimated if increasing selection into employment is not taken into account. One last strand of literature relevant for this paper focuses on examining the large variation of the gender wage gap across the distribution as well as the different roots behind gender wage gaps among low- and high-earning individuals (e.g. López-Nicolás et al. 2001, Albrecht et al. 2003, Gupta et al. 2006, Arulampalam et al. 2007, Kassenböhmer and Sinning 2014).

In the next two sections, we provide a detailed description of our empirical strategy and the data. In Section 2.4.1, we provide descriptive evidence on the shape of the gender wage gap across the distribution and its evolution over time. Section 2.4.2 shows the results of the imputation of non-realised wages and presents estimates on the magnitude and evolution of selection into employment of men and women. Next, in Section 2.4.3 results on the selection-corrected gender wage gap as well as its heterogeneity by age and level of education are discussed. The results of the decomposition exercise are presented in Section 2.4.4. Finally, Section 2.5 summarizes our findings and draws policy conclusions.

2.2 Empirical strategy

The aim of this paper is to present a selection-corrected gender wage gap across the wage distribution and to decompose it into the effect of different returns to skill and the effect of different distribution of skills.

To this end, we need a method to correct for selection into employment that is compatible with our distributional approach. We use an imputation-based method proposed in Melly and Santangelo (2015) that imposes the time-invariance of the distribution of unobservables (conditional on observables) to recover non-realized wages from individuals out of work.

Imputation-based approaches correcting for selection into employment have been widely used in the literature (Neal 2004, Blau and Kahn 2006, Olivetti and Petrongolo 2008 among others). Mostly, however, they rely on informed guesses of whether non-realized wages fall above or below the observed median wages - conditional on the education level of the individual. The imputation method proposed by Melly and Santangelo (2015) goes a step further and suggests imputing a point-identified wage for each out-of-work individual that takes into account both observable and unobservable characteristics of the individual. Once each individual in the sample gets assigned a wage, which can be observed or imputed, the resulting wage distribution is by definition selection-corrected. Throughout the paper, we denote the observed cumulative distributions of men's and women's full-time log wages as $F_{W^M,t}$ and $F_{W^F,t}$, respectively, which are only defined for individuals pursuing full-time employment at time t. We refer to the selection-corrected counterparts as $\widehat{F}_{W^M,t}$ and $\widehat{F}_{W^F,t}$, which are defined for all individuals in the sample.

Thus, the observed gender wage gap at point $\tau \in (0,1)$ in the unconditional distribution is given by:

$$G_{obs}(\tau,t) = F^{-1}_{W^M,t}(\tau) - F^{-1}_{W^F,t}(\tau) \tag{2.1}$$

whereas the corrected wage gap can be expressed as:

$$\widehat{G}_{corr}(\tau,t) = \widehat{F}^{-1}_{W^M,t}(\tau) - \widehat{F}^{-1}_{W^F,t}(\tau) \tag{2.2}$$

As a summarizing measure of selection into employment for each gender, we compare the observed and selection-corrected log wage distributions. Formally:

$$\widehat{Sel}^{M}(\tau,t) = F^{-1}_{W^M,t}(\tau) - \widehat{F}^{-1}_{W^M,t}(\tau) \tag{2.3}$$

$$\widehat{Sel}^{F}(\tau,t) = F^{-1}_{W^F,t}(\tau) - \widehat{F}^{-1}_{W^F,t}(\tau) \tag{2.4}$$

The rest of this section is structured as follows. Next, we present Melly and Santangelo (2015)'s imputation method of non-realised wages[17]. Then, we discuss the specification

[17] See Melly and Santangelo (2015) for details. The aim of this subsection is merely to sketch the imputation method proposed by the authors so as to make it understandable - for a thorough explanation of the model it is advised to refer to the original sources.

of the conditional wage model. Last, we present the decomposition method used for disentangling the effects of the coefficients and of the characteristics between men and women.

2.2.1 Imputation of non-realised wages

Intuitively, Melly and Santangelo (2015) suggest to build subsamples with individuals that work in two given periods (group 0) and compare them to subsamples of individuals that only work in one of these two periods (group 1). Group 1 individuals reveal information on their unobservables in the one period when they work, which is captured by their conditional rank in the wage distribution. The evolution of wages of group 0 allows imputing a conditional wage for group 1, which responds to the wage structure of the time when the imputation is required and accounts for both individuals' observable and unobservable characteristics.

This exercise is carried out for all possible combinations of two survey years in the data, which we refer to as $t \in \{k, l\}$. The assignment to a group according to the description above is captured by variables G_{kl}, each with realisation $g_{kl} \in \{0, 1\}$. The covariates entering the wage equation are captured in vector of variables X and the unobservable component of wages in random variable U.

The identifying assumption behind this imputation method is the time-invariance of the unobservables conditional on the observables, $f_{U|T,G,X} = f_{U|G,X}$. In words, this implies that the distribution of unobservable characteristics which have an effect on wages (such as innate ability, professional ambition) stay constant over time conditional on observable characteristics (such as education level and previous working experience). The specification of the covariate vector needs to capture as much variation in unobservable components as possible. It is therefore discussed separately in the next subsection.

Formally, Melly and Santangelo (2015) show that the conditional wage distribution of those individuals not working in time period t=k but working in time period t=l can be derived as:

$$F^{-1}_{W|g=1,t=k,x}\left(\theta\right) = F^{-1}_{W|g=0,t=k,x}\left(F_{W|g=0,t=l,x}\left(F^{-1}_{W|g=1,t=l,x}\left(\theta\right)\right)\right) \tag{2.5}$$

and individual wages included in $F^{-1}_{W|g=1,t=k,x}(\theta)$ can be imputed as:

$$\widetilde{w}_{ikl} = x_i \hat{\beta}_{g=0,t=k} \left(\int_0^1 \mathbb{1} \left(x_i \hat{\beta}_{g=0,t=l}(u) \leq x_i \hat{\beta}_{g=1,t=l}(\theta) \right) du \right) \tag{2.6}$$

where $\hat{\beta}_{g,t}(\theta)$ are the wage equation coefficients for quantile θ coming from the estimated conditional quantile regression processes. As each realized full- or part-time wage provides enough information to impute a non-realized wage, for individuals observed working several years, the imputation rule of expression 2.6 produces multiple available imputations for a single non-realized wage. For these cases, Melly and Santangelo (2015) suggest to weigh all available imputations so as to obtain a final imputed wage for each individual i in year k, $\widetilde{w}_{ik}$:

$$\widetilde{w}_{ik} = \sum_{m=1}^{M_i} d_{ikm} \cdot \widetilde{w}_{ikl} \tag{2.7}$$

where M_i is the number of available imputations for a given $\widetilde{w}_{ik}$, different for each individual i, and d_{ikm} is the weighting factor for each $\widetilde{w}_{ikl}$, so that $\sum_{m=1}^{M_i} d_{ikm} = 1$. Equation 2.6 only recovers non-realised wages from individuals observed working full-time for at least at one point in time. For the rest, we predict wages by means of the median conditional quantile coefficients estimated on the imputed wages according to Equation 2.6[18].

2.2.2 Conditional wage model

The wage equation is estimated separately for men and women using a linear conditional quantile regression model (Koenker and Bassett, 1978):

$$Q_\theta (w_{it}|x_{it}) = x'_{it} \beta_t (\theta) \tag{2.8}$$

The dependent variable, w_{it}, is the natural logarithm of the hourly wage. The independent variables, x_{it}, consist of an intercept, age (polynomially, up to the power of three), years of education, an indicator variable for an intermediate degree, an indica-

[18]In Chapter 3, Section 3.2.2 discusses the implications of the median imputation rule.

tor variable for an advanced degree and actual full- and part-time working experience (both polynomially up to the power of three). Age is included as a regressor in addition to years of education and actual working experience in order to account for unemployment spells. In Section 2.4.2, we examine the goodness of fit of the conditional wage model with this specification. The fulfilment of the identifying assumption with this specification of the wage model is discussed in the Appendix to Chapter 3.

2.2.3 Decomposition of the gender wage gap

Having estimated the selection-corrected gender wage gap, we analyse to which extent this differential is driven by differences in labour market endowments or by the returns to those endowments. To this end, we carry out a decomposition in the spirit of Oaxaca (1973) and Blinder (1973) and its quantile equivalent as suggested by Machado and Mata (2005)[19].

Thus, the selection-corrected gender wage gap at the τ-percentile can be expressed as:

$$\widehat{G}_{corr}(\tau, t) = \left[\widehat{F}^{-1}_{W^M,t}(\tau) - \widehat{F}^{-1}_{W^C,t}(\tau)\right] + \left[\widehat{F}^{-1}_{W^C,t}(\tau) - \widehat{F}^{-1}_{W^F,t}(\tau)\right] \tag{2.9}$$

where $\widehat{F}_{W^C,t}$ stands for a (selection-corrected) counterfactual wage distribution. Naturally, the counterfactual distribution is not observed and we have to make an assumption about its shape. In our case, we define the counterfactual wage distribution to be the wage distribution that would result if women, with their given characteristics, would be paid as if they were men. Under this choice of counterfactual, the first term on the right hand side in Equation 2.9 is the component explained by differences in characteristics (the covariates in the wage model) and the second term captures differences in the coefficients attached to those characteristics (often referred to as the unexplained component of the gender wage gap in the literature)[20].

Constructing $\widehat{F}_{W^C,t}$ requires a two-step procedure. First, we run a tight grid of condi-

[19] See Fortin et al. (2011) for an in-depth discussion of the assumptions underlying decompositions of wage differentials.

[20] Alternatively, the counterfactual wage distribution could consist of men's labour market characteristics being paid as if they would be women. As explained in Chapter 1, in this case the correct interpretation of the two components would be the other way around.

tional quantile regressions on the entire male sample (i.e. those in full-time employment and those out of it) at each point in time, as in Equation 2.8. This allows us to compute counterfactual conditional wages for all women, $\left\{\hat{w}_i^C(\theta) = x_i^F \hat{\beta}^M(\theta)\right\}_{\theta=0.01}^{\theta=0.99}$.
In order to recover the unconditional log wage distribution corresponding to the estimated counterfactual wages, we use the algorithm proposed by Melly (2005)[21], concretely:

$$\widehat{F}_{W^C}^{-1}(\tau) = \inf\left\{w^C : N^{-1}\sum_{i=1}^{N}\mathbb{1}\left(\hat{w}_i^C(\theta) \leq w^C\right) \geq \tau\right\}, \tau \in (0,1) \qquad (2.10)$$

Equation 2.10 means that the unconditional quantile function can be recovered by the sample quantiles computed on $\left\{\hat{w}_i^C(\theta) = x_i^F \hat{\beta}^M(\theta)\right\}_{\theta=0.01}^{\theta=0.99}$, which are counterfactual conditional wages calculated for all (female) individuals in the data for a grid of θ at 99 points (each percentile) of the conditional distribution.

2.3 Data and descriptive statistics

2.3.1 Data

Four our empirical analysis we use the German Socio-Economic Panel (GSOEP) for years 1985 to 2009. The GSOEP is a rich dataset for Germany that brings together a longitudinal dimension and detailed information on the number of working hours, which are both essential for the imputation of non-realised wages. Official statistics about the gender wage gap for Germany are computed on the basis of the German Structure of Earnings Survey (SES), which is the most reliable data source for estimating the gender wage gap at one point in time. Unfortunately, the SES is not a longitudinal dataset and therefore cannot be used to answer our research question. The administrative datasets from the Institute for Employment Research (IAB), which are based on social security reports, could be an alternative as they are longitudinal datasets with high-quality earnings information. Unfortunately, they lack detailed information on individuals' working hours. This poses a problem given the systematic differences betweeen working

[21] As explained in Chernozhukov et al. (2013), this procedure is numerically identical to the procedure proposed by Machado and Mata (2005) when the number of simulations used approaches infinity.

hours of men and women, which prevail even if the analysis is restricted to full-time employment (see Table B.3 in the Appendix) and are likely to introduce a bias in the estimates of the gender wage gap. However, the GSOEP consists of survey data which may suffer from measurement error issues. Therefore, in Section 2.4 we validate our estimates of the observed gender wage gap based on the GSOEP with studies based on the SES working with a similar sample. We find that point estimates of the gender wage gap are similar for both datasets.

We restrict our estimation sample to individuals aged 20 to 55 with residence in West Germany[22]. Furthermore, we exclude individuals in retirement, the military, and disabled people. Individuals who only appear once in the data are dropped, as the imputation procedure requires at least two observations per individual. In addition, individuals with missing information in one or more covariates entering the wage model are also dropped. This leaves us with 77,963 male and 86,553 female person-year observations in our sample (see Tables B.1 and B.2 in the Appendix).

We focus only on the gender wage gap among full-time employees. We define full-time employment based on whether individuals work at least 30 hours per week. Working hours are defined as actual working hours. The dependent variable is the natural logarithm of the hourly wage. The hourly wage variable has been constructed by dividing gross monthly earnings over actual working hours. Only earnings and working hours from individuals' main job is taken into account.

Following the literature, for individuals in part-time employment or not employed, a wage is imputed. The same is applies to individuals who report being on full-time education (irrespectively of whether they have a side job) as well as for individuals with missing information on earnings but complete information on the covariates entering the wage equation.

The number of individuals for whom we never observe a full-time wage is important for the imputation procedure. This ranges between 2% and 5% of all male person-year observations and between 25% and 33% of all female person-year observations (detailed

[22] The labour market of East Germany during these years had very distinctive features that would require a separate estimation. Unfortunately, the number of observations for East Germany also happens to be too small for our data-intensive imputation method.

year-by-year figures are provided in Tables B.1 and B.2 in the Appendix).

2.3.2 Descriptive statistics

Consistent with the chapter's focus on the changing effect of selection into employment on the gender wage gap, Figure 2.1 displays male and female full-time employment rates over time (conditional on the threshold defining full-time employment at 30 hours per week). The male employment rate has stayed relatively stable over time at around 80%, and the female employment rate at around 35% (although by the end of the 2000s one observes a light increasing trend approaching a rate of 40%).

Figure 2.1: Full-time employment rates, by gender

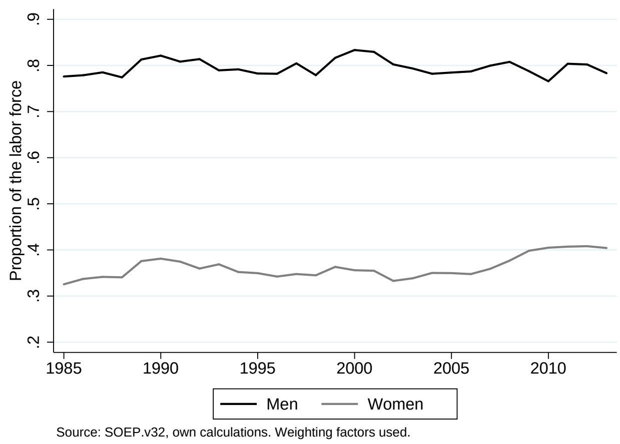


Table 2.1 show sample means of the covariates entering the wage equation. The two upper panels refer to the sample of men and women working full-time for whom we estimate the gender wage gap. The male sample was older than the female sample throughout the 25 years under study, but the difference has become smaller over time. The average years of education between the two groups are similar in the studied time period, slightly higher for men by the mid 1980s and higher for women during the 2000s. The share of women with an intermediate degree was higher than the share of men throughout the entire period, and the gap in terms of advanced degrees has been closing too over these 25 years, a phenomenon observed in most high-income countries (Becker et al. 2010).

The two lower panels show sample means of the wage covariates for individuals - men and women - who do not work full-time. The idea here is to obtain a first impression on the sign and evolution of the selection on observables into full-time employment.
Men not working full-time are younger than men in full-time employment, probably because a large share of them is still in education. At the same time, the average years of education of male individuals are very similar for the two subgroups. However, the share of men with basic and advanced degrees is higher among those in full-time employment, whereas the share of men with intermediate degrees is higher among those working part-time or out of work. Previous working experience spells are also shorter for those who are not working full-time, in line with the younger age of the sample. From these descriptive statistics one cannot conclude whether or not men working full-time are a selected sample on observables.
In the case of women, positive selection on observables is much more evident (see Table 2.1). The share of women with both intermediate and advanced degrees is higher in the group working full-time and previous full-time working spells are also longer for that group, even though the group working full-time is on average younger than the one not working full-time. This suggests that work interruptions are more frequent and longer among the subsample of women not working full-time. That, in turn, is a sign of positive selection into full-time employment on education and previous working experience.

Table 2.1: Sample means (standard deviations) and sample size

	Age		Years of education		Intermediate degree (0/1)		Advanced degree (0/1)		Experience full-time		Experience part-time		Sample size
(A) Men working full-time													
1985-1989	37.8	(9.8)	11.6	(2.5)	0.24	(0.43)	0.14	(0.34)	16.8	(10.5)	0.2	(1.2)	12104
1990-1994	38.2	(9.7)	11.7	(2.6)	0.27	(0.45)	0.16	(0.36)	16.7	(10.4)	0.2	(1.2)	10967
1995-1999	38.1	(9.0)	12.0	(2.7)	0.33	(0.47)	0.18	(0.39)	16.0	(9.7)	0.3	(1.2)	10566
2000-2004	38.9	(8.7)	12.2	(2.6)	0.36	(0.48)	0.20	(0.40)	16.6	(9.5)	0.3	(1.3)	16372
2005-2009	39.8	(8.7)	12.4	(2.7)	0.38	(0.49)	0.22	(0.41)	16.9	(9.4)	0.5	(1.6)	13229
(B) Women working full-time													
1985-1989	34.0	(10.4)	11.4	(2.3)	0.42	(0.49)	0.08	(0.28)	10.8	(8.9)	1.4	(4.0)	5495
1990-1994	35.4	(10.2)	11.5	(2.4)	0.45	(0.50)	0.09	(0.29)	11.1	(8.9)	1.9	(4.1)	5356
1995-1999	36.8	(9.8)	11.9	(2.6)	0.45	(0.50)	0.14	(0.35)	11.9	(9.0)	2.0	(4.0)	5276
2000-2004	37.8	(9.6)	12.3	(2.5)	0.49	(0.50)	0.18	(0.38)	12.3	(9.1)	2.4	(4.5)	8494
2005-2009	38.6	(9.8)	12.6	(2.7)	0.51	(0.50)	0.21	(0.41)	12.3	(9.4)	3.0	(4.9)	7626
(C) Men not working full-time													
1985-1989	30.9	(10.4)	11.5	(2.3)	0.47	(0.50)	0.07	(0.25)	8.2	(11.0)	0.4	(1.4)	3125
1990-1994	31.0	(10.1)	11.8	(2.7)	0.48	(0.50)	0.11	(0.31)	7.3	(10.2)	0.5	(1.3)	2589
1995-1999	32.4	(10.3)	11.7	(2.6)	0.47	(0.50)	0.11	(0.31)	8.0	(10.1)	0.8	(1.8)	2630
2000-2004	32.1	(10.3)	11.8	(2.5)	0.50	(0.50)	0.11	(0.32)	7.6	(9.6)	0.8	(1.8)	3483
2005-2009	32.2	(10.4)	11.6	(2.4)	0.50	(0.50)	0.09	(0.29)	7.2	(9.5)	1.2	(2.4)	2898
(D) Women not working full-time													
1985-1989	37.5	(10.3)	10.8	(2.1)	0.32	(0.47)	0.05	(0.21)	6.0	(6.2)	3.0	(5.3)	10014
1990-1994	37.7	(10.1)	11.1	(2.3)	0.37	(0.48)	0.07	(0.26)	6.0	(5.8)	3.5	(5.6)	8778
1995-1999	37.4	(9.7)	11.4	(2.4)	0.42	(0.49)	0.09	(0.29)	6.4	(6.2)	3.7	(5.3)	9132
2000-2004	37.9	(9.5)	11.6	(2.3)	0.47	(0.50)	0.10	(0.31)	6.2	(6.3)	4.2	(5.9)	14509
2005-2009	38.4	(9.6)	11.8	(2.4)	0.51	(0.50)	0.12	(0.32)	6.1	(6.3)	4.9	(5.9)	11873

Source: SOEP.v32, weighting factors used, own calculations.

2.4 Results

In this section, we start by offering a descriptive analysis on the observed gender wage gap across the distribution and over time, as well as the heterogeneity that arises when the gap is analysed conditional on selected characteristics such as age, education or family status. Next, we describe the imputation of non-realised wages that enables us to derive selection-corrected wage distributions. We present detailed findings on both the magnitude and evolution of male and female selection into employment by selected characteristics. Subsection 2.4.3 presents the selection-corrected gender wage gap. Finally, subsection 2.4.4 shows the results of our decomposition of the gap.

2.4.1 Raw gender wage gap

Our analysis shows that the observed gender wage gap among full-time workers displays a slight u-shape over the distribution for the period 1985-2005, with somewhat higher numbers at the very bottom (10^{th} percentile) and the very top (90^{th} percentile). In these early years, the gap at the very bottom and at the very top is mostly 6 to 7 log points higher than at the median. In the more recent years 2005-2009, the profile of the gender wage gap across the distribution becomes flatter and amounts to 18 log points at the median as well as at the 75^{th} percentile; and it is only one log-point higher at the 90^{th} percentile. At the bottom of the distribution, the gender wage gap is slightly higher than at the median, amounting to 22 log points at the 10^{th} and to 21 log points at the 25^{th} percentile.

Our estimate of a median gender wage gap of 18 log points for the years 2005-2009 compares fairly well to estimates obtained from administrative datasets. In particular, the German Federal Statistical Office calculates a gender wage gap based of 23 percent based on the SES for the year 2006 (Finke 2010). However, the underlying samples of both studies differ in two aspects: first, the Federal Statistical Office includes all workers, part-time and full-time, while we only include full-time workers. Given that most part-time employees are women and observed part-time wages are substantially lower than full-time wages (Paul 2016), one would expect that their estimate is larger

Table 2.2: Observed gender wage gap

	$\tau = .10$	$\tau = .25$	$\tau = .50$	$\tau = .75$	$\tau = .90$
1985-1989	.32	.29	.26	.25	.29
1990-1994	.32	.26	.23	.26	.28
1995-1999	.26	.21	.20	.21	.26
2000-2004	.26	.22	.19	.20	.25
2005-2009	.22	.21	.18	.18	.19

Comments: Log-point differences between the male and female inverse cumulative wage distributions evaluated at $\tau = .10, .25, .50, .75, .90$.
Source: SOEP.v32, own calculations.

than ours. This is indeed confirmed by Antonczyk et al. (2010), who also use the SES but restrict their sample to full-time employees. They find a median gender wage gap of 20 log points for the year 2006[23]. Second, our estimate is an average over a five-year period in a general context of a slowly decreasing gender wage gap. We argue that the similarity of the figures between studies validate the use of the GSOEP for the analysis of the gender wage gap in Germany.

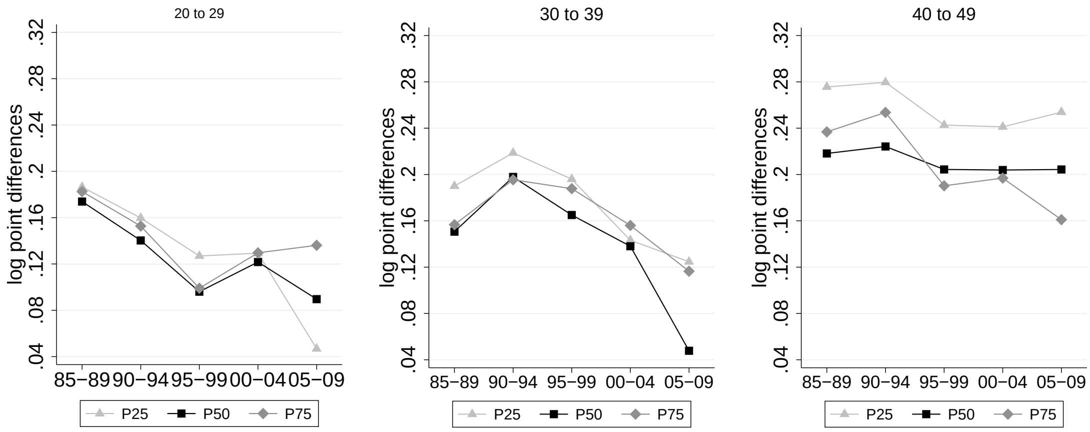
Comparing the observed gender wage gap at different points in time, we find that it has been decreasing within our observation period (1985 to 2009). This decrease was more or less parallel over the wage distribution and was much more pronounced in the 1980s and 1990s than in the 2000s. As Table 2.2 shows, from 1985 to 1999, the gender wage gap decreased from 32 to 26 log points at the 10^{th} percentile, from 29 to 21 log points at the 25^{th} percentile, from 26 to 20 log points at the median and from 25 to 21 log points at the 75^{th} percentile. In the following period, from 2000 to 2009, the decrease was much lower from the 10^{th} to the 75^{th} percentile. The development of the gender wage gap at the 90^{th} percentile, however, followed a different trend: it has also been decreasing over the whole observation period; but here, the decrease was more pronounced in the 2000s than in the late 1980s and 1990s (see Figure B.1 in the Appendix for a graphical presentation of Table 2.2)[24].

[23] Our results at other points of the distribution are also similar to those of Antonczyk et al. (2010)

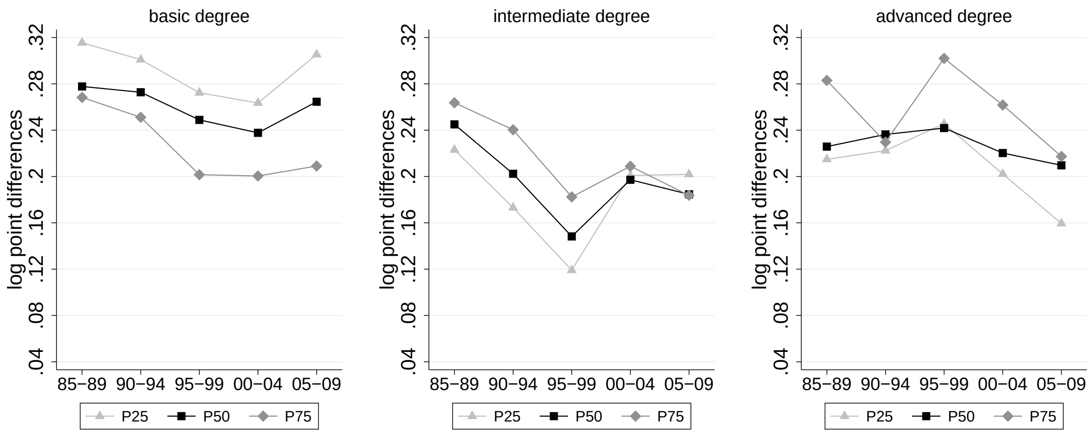
[24] Our finding that the decrease of the gender wage gap over time was more or less parallel over the whole wage distribution - except at the very top - is in line with the fact that the increase of wage inequality followed a similar pattern for men and women, at least in the 1990s (see Dustmann et al. 2009).

Figure 2.2: Observed gender wage gap, by age and education

(a) Age



(b) Education



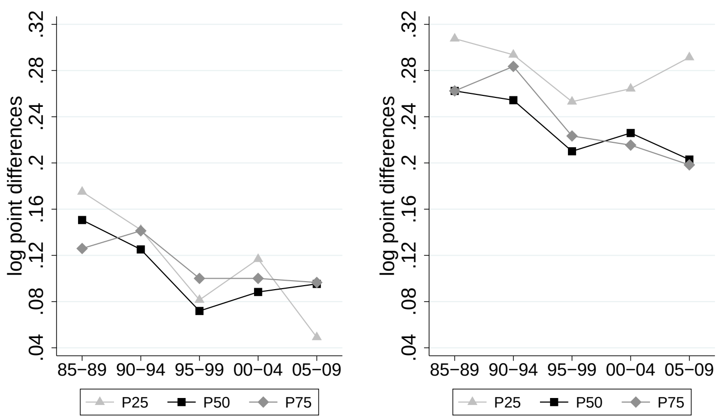
Source: SOEP.v32, own calculations.

A closer look at the observed gender wage gap conditional on age and education reveals more interesting insights (see Figure 2.2). Consistent with the literature, the observed gender wage gap increases with age (e.g. Goldin 2014, Juhn and McCue 2017). Moreover, the gap has been decreasing over time at all points of the distribution among individuals in their 20s and 30s. This has not been the case for individuals in their 40s, for whom the median observed gap has stayed constant at around 20 log points during the entire period under study. Furthermore, the gap displays an almost flat profile over the distribution for the two younger groups up to the mid 2000s, but becomes more unequal in the latest period; this means a glass ceiling for both groups and a sticky floor for the group in their 30s. The dispersion of the gap across the distribution is more pronounced for the group in their 40s, especially in the form of a persistent sticky floor.

The lower panel of Figure 2.2 depicts observed gender wage gaps by groups with different educational attainment. From this perspective, the observed gender wage gap is highest among those with a basic (or no) degree. For this group, the gap is higher at the lower end of the distribution, pointing to the existence of a sticky floor. However, among those with an intermediate and advanced degrees, one observes glass ceilings in the 1980s and 1990s. This mirros the findings of Rica et al. (2008), according to which sticky floors arise among low-educated individuals and glass ceilings occur among highly-educated individuals. Over time, the median observed wage gap by educational levels stays fairly constant for those with a basic and an advanced degree. It decreases for those with an intermediate degree.

Figure 2.3: Observed gender wage gap, by family status

(a) Single **(b)** Married



Source: SOEP.v32, own calculations.

Finally, Figure 2.3 shows that the incidence of the gender wage gap strongly correlates with family status. Thus, by 2009 the median gap among single individuals was approximately 10 log points, whereas it was around 20 log points at the same point in time.

2.4.2 Imputation of non-realised wages and selection into employment

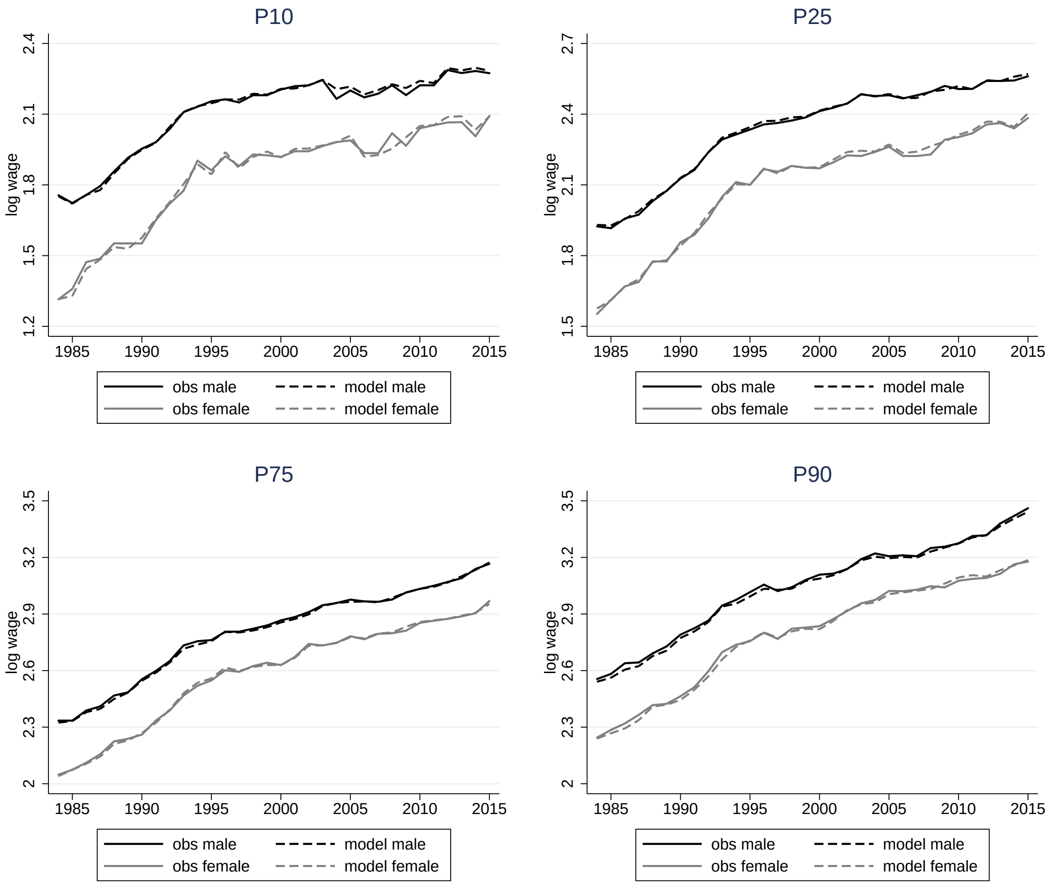
The imputation of non-realized wages is based on the conditional quantile regression wage model presented in Section 2.2. In order to evaluate the goodness of fit of the specification of the wage quantile regression, Figure 2.4 below contrasts observed log hourly wages (solid lines) with those produced by the conditional wage model (dashed lines) at different points of the unconditional distribution. In order to obtain unconditional wage distributions from the grid of conditional quantile regressions, we use the algorithm provided by Chernozhukov et al. (2013)[25]. Figure 2.4 shows that the conditional wage model fits the observed wage data very well at most points of the distribution, both for male and female wages.

According to the imputation method described in Section 2.2, the imputation of all non-realised wages requires several hundred thousand conditional quantile regressions run on all possible pairs of subsamples (defined by $G_{k,l}$). Those are not reported for the sake of brevity. Instead, Figure 2.5 shows the results of the imputation by gender and over time - at three selected points in the distribution (the 25^{th}, 50^{th} and 75^{th} percentiles).

In all cases, imputed wages (dashed lines in Figure 2.5) are lower than observed wages (solid lines in Figure 2.5) at all points of the distribution - which points to positive selection into employment for both men and women. Furthermore, during the end of the 1980s and beginning of the 1990s both observed and imputed wages follow a parallel trend. From the mid 1990s onwards, the difference between observed and imputed wages broadens dramatically for both genders at all points of the distribution. Dotted lines in Figure 2.5 represent the selection-corrected wage distribution, which is made up of both observed and imputed wages, and therefore bound to lie between the solid and dashed lines depicting observed and imputed wages, respectively.

[25] This algorithm yields results equivalent to those suggested by Machado and Mata (2005) for a large enough number of replications.

Figure 2.4: Goodness of fit of the conditional wage model

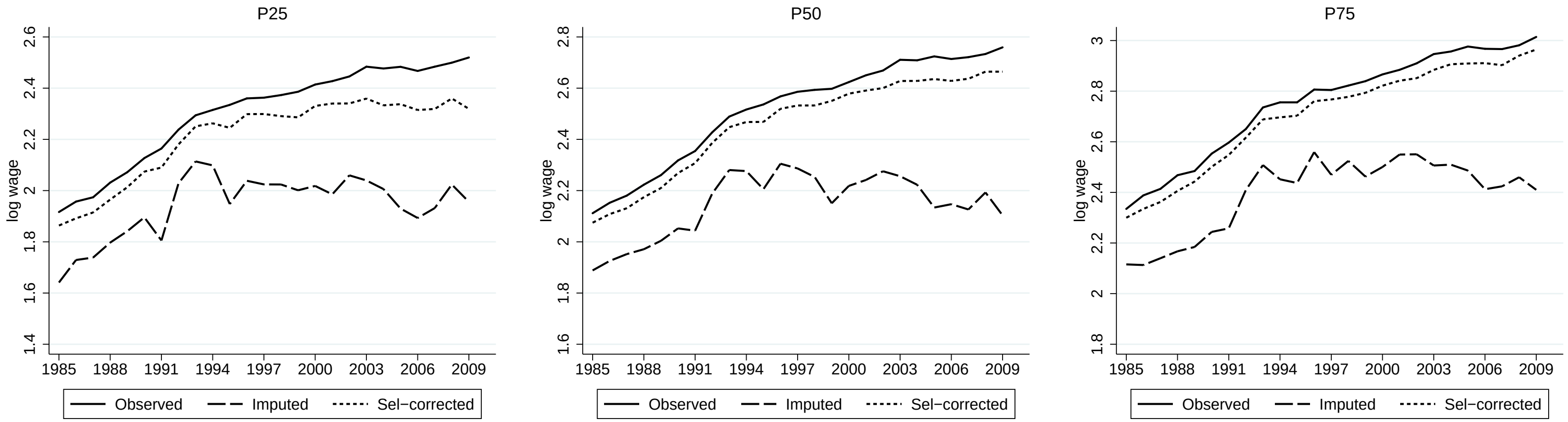


Source: SOEP.v32, own calculations.

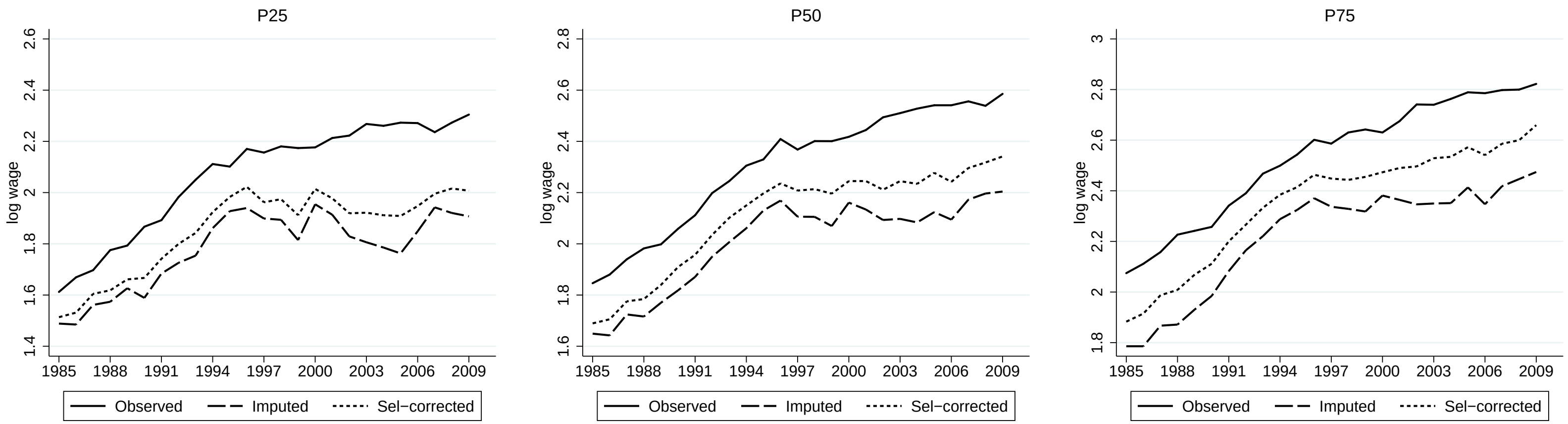
How close or far away the selection-corrected distribution is from the observed one depends on two factors: the share of individuals requiring imputations, and the difference in levels between imputed and observed wages. For men, for whom the full-time employment rate stays at around 80% (see Figure 2.1), the selection-corrected distribution is much closer to the observed distribution than for women, for whom the full-time employment rate stays at around 35% in the time period under study. However, the level difference between the observed and imputed wages is more sizeable for men than for women at all points of the distribution. It is precisely the distance between the selection-corrected distribution and the observed distribution that summarises the effect of selection into full-time employment on men's and women's wages.

Figure 2.5: Observed, imputed and selection-corrected wage distributions, by gender

(a) Men



(b) Women



Source: SOEP.v32, own calculations.

For each five-year period, Table 2.3 below shows averages of the effect of selection into full-time employment on the wage distributions for men and women. Given that the difference is specified as the observed distribution minus the selection-corrected one, positive values hint at positive sample selection into full-time employment with respect to the entire population.

Panel A shows that the effect of selection on male wages is positive but rather small, except for at the bottom of the distribution, where the selection effect has strongly increased over the observation period. For women, as seen in Panel B, the selection effect is much stronger and has also been increasing over time, particularly in the lower half of the distribution. While the selection effect has tripled at the 10^{th} percentile, more than doubled at the 25^{th} and increased by a third at the median, it has remained fairly constant over time at the 75^{th} and 90^{th} percentile.

Table 2.3: Effect of selection on the wage distributions

	$\tau = .10$	$\tau = .25$	$\tau = .50$	$\tau = .75$	$\tau = .90$
(A) Male Wages: $\widehat{Sel}^{M}(\tau,t) = F^{-1}_{W^M,t}(\tau) - \widehat{F}^{-1}_{W^M,t}(\tau)$					
1985-1989	.08*	.06*	.05*	.05*	.05*
1990-1994	.07*	.05*	.04*	.04*	.04*
1995-1999	.13*	.08*	.06*	.05*	.04*
2000-2004	.15*	.11*	.07*	.05*	.04*
2005-2009	.22*	.16*	.08*	.06*	.05*
(B) Female Wages: $\widehat{Sel}^{F}(\tau,t) = F^{-1}_{W^F,t}(\tau) - \widehat{F}^{-1}_{W^F,t}(\tau)$					
1985-1989	.10*	.12*	.17*	.19*	.15*
1990-1994	.15*	.17*	.16*	.13*	.13*
1995-1999	.21*	.19*	.17*	.16*	.13*
2000-2004	.31*	.27*	.24*	.21*	.16*
2005-2009	.31*	.29*	.26*	.21*	.18*

Comments: Differences in log-points between the observed and the selection-corrected distributions. Unweighted average of multiple imputations. * Statistical significance at the 5% level (computed by bootstrap with 200 replications).
Source: SOEP.v32, own calculations.

These findings are in line with the framework presented by Mulligan and Rubinstein

(2008), according to which positive selection into employment is expected to increase in a context of rising wage inequality, even if full-time employment rates stay constant. Their reasoning is that higher returns to skill over-proportionally induce high-skilled individuals to participate in the labour market, as their market wages increasingly exceed their reservation wages. This seems to be the case here, since the effect of selection on wages is strongest at the lower end of the distribution, which indicates that individuals with low potential wages increasingly stay out of full-time employment.

We can now go a step further and analyse the potential heterogeneity of the selection effect by relevant characteristics such as age and education. Table 2.4 summarizes the effect of selection on median wages for different age categories. The numbers in Table 2.4 can be interpreted as the log-point difference between the observed median wage for a given age category minus its selection-corrected counterpart.

Table 2.4: Effect of selection on median wages, by age

Ages	20 to 29	30 to 39	40 to 49
(A) Male Wages $\widehat{Sel}^{M}_{t\|age}(\tau = .5) = F^{-1}_{W^M,t\|age}(\tau = .5) - \widehat{F}^{-1}_{W^M,t\|age}(\tau = .5)$			
1985-1989	.06*	.02	.03
1990-1994	.06*	.03	.01
1995-1999	.11*	.03	.02
2000-2004	.12*	.03*	.03
2005-2009	.22*	.04*	.02
(B) Female Wages $\widehat{Sel}^{F}_{t\|age}(\tau = .5) = F^{-1}_{W^F,t\|age}(\tau = .5) - \widehat{F}^{-1}_{W^F,t\|age}(\tau = .5)$			
1985-1989	.10*	.23*	.34*
1990-1994	.10*	.12*	.24*
1995-1999	.17*	.17*	.21*
2000-2004	.23*	.23*	.23*
2005-2009	.25*	.27*	.26*

Comments: Differences in log-points between the observed and the selection-corrected distributions. *Statistical significance at the 5% level (computed by bootstrap with 200 replications).

Source: SOEP.v32, own calculations.

Interestingly, selection in the male distribution appears to come from the youngest age category, where the selection effect at median wages has risen from 6 log-points in the mid 1980s to 22 log-points by 2009. However, median wages of those aged 30 and older are only minimally or not selected at all over the whole time period under study. For women, we see a completely different picture. First, selection is the smallest among the youngest age group, although it has been increasing uninterruptedly over time. For the two older groups, selection at the median has been decreasing up to mid 1990s and then rising again during the 2000s. Thus, whereas by mid 1980s selection at the median increased steeply with age, by the end of the 2000s this is no longer the case.

Table 2.5: Effect of selection on median wages, by education

Type of Degree:	Basic	Intermediate	Advanced
(A) Male Wages $\widehat{Sel}^{M}_{t\|edu}(\tau=.5)=F^{-1}_{W^M,t\|edu}(\tau=.5)-\widehat{F}^{-1}_{W^M,t\|edu}(\tau=.5)$			
1985-1989	.02	.11*	.03
1990-1994	.02	.10*	.03
1995-1999	.03	.10*	.02
2000-2004	.05*	.09*	.04
2005-2009	.08*	.10*	.02
(B) Female Wages $\widehat{Sel}^{F}_{t\|edu}(\tau=.5)=F^{-1}_{W^F,t\|edu}(\tau=.5)-\widehat{F}^{-1}_{W^F,t\|edu}(\tau=.5)$			
1985-1989	.16*	.12*	.09*
1990-1994	.19*	.10*	.08*
1995-1999	.17*	.13*	.12*
2000-2004	.28*	.17*	.11*
2005-2009	.28*	.20*	.13*

Comments: Differences in log-points between the observed and the selection-corrected distributions. *Statistical significance at the 5% level (computed by bootstrap with 200 replications).
Source: SOEP.v32, own calculations.

The separate analysis of the effect of selection on wages by educational attainment also reveals substantial differences between men and women. For men, median wages among those with an intermediate degree are the most selected. For the group holding an advanced degree, median wages are practically selection-free. The same applies to the male group holding a basic or no degree, whose median wages have been selection-free up to the end of the 1990s and have increased only during the 2000s. However, the magnitude of selection at male median wages for the two other groups has been very stable over time.

For women, the selection effect on median wages clearly decreases with educational attainment at all points in time. At the same time, the magnitude of selection has increased within each group over time but at different speeds. The increase has been most pronounced for the group with lowest educational attainment (a 12 log-point increase over 25 years) followed by the group with an intermediate degree (8 log-point difference).

2.4.3 Selection-corrected gender wage gap

Having derived the selection-corrected male and female wage distributions, we can calculate the selection-corrected gender wage gap. Table 2.6 below summarises our findings.

Table 2.6: Selection-corrected gender wage gap

	$\tau = .10$	$\tau = .25$	$\tau = .50$	$\tau = .75$	$\tau = .90$
1985-1989	.33*	.35*	.38*	.39*	.40*
1990-1994	.40*	.38*	.35*	.35*	.37*
1995-1999	.33*	.32*	.31*	.32*	.35*
2000-2004	.42*	.39*	.37*	.36*	.37*
2005-2009	.31*	.35*	.35*	.33*	.32*

Comments: Log-point differences between the relevant inverse cumulative distributions. Unweighted average of multiple imputations. *Statistical significance at the 5% level (computed by bootstrap with 200 replications).
Source: SOEP.v32,own calculations.

Our main finding is that the selection-corrected gender wage gap is much higher than the observed gap. As Section 2.4.2 has shown, this phenomenon is driven by the fact that the female wage distribution is much more selected than the male distribution. The observed gender wage gap has been constantly decreasing over time - but the selection-corrected gap has not. More precisely, the selection-corrected gap decreased from 1985 to 1999, and then increased almost to the original levels again in the period from 2000 to 2004. Since then, in the years 2005 to 2009 it has again been slightly decreasing. This pattern is found for the whole wage distribution except for the very top, where the increase in the selection-corrected gap between 2000 and 2004 was much less pronounced than at the other parts of the wage distribution.

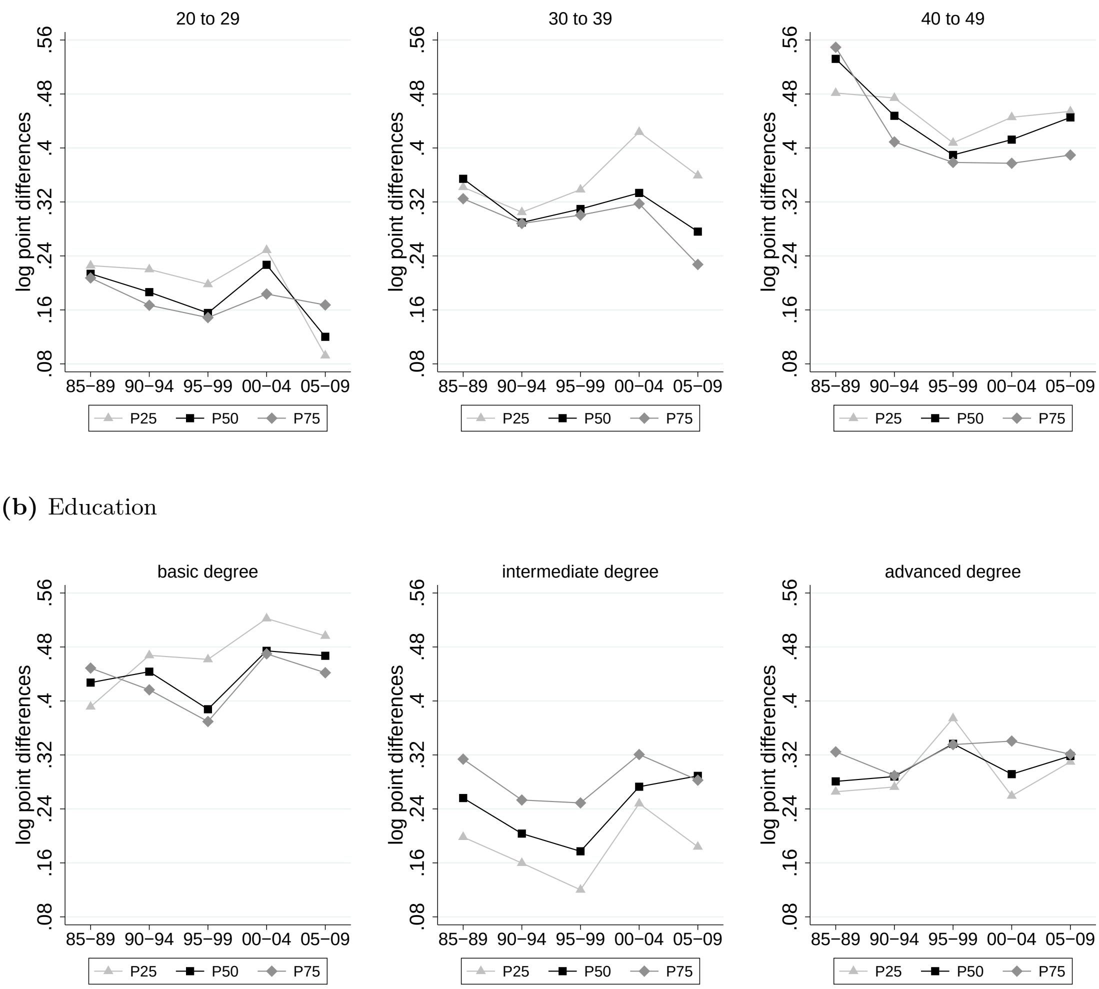
In addition, the shape of the selection-corrected gap over the distribution also differs to that of the observed gap (see Figure B.3 in the Appendix). By the mid 1980s, the selection-corrected gender wage gap displayed an increasing profile across the wage distribution, unlike the u-shaped profile of the observed gap. However, during the 1990s the selection-corrected gap mimicked the observed gap and displayed a slight u-shape. By the beginning of the 2000s, the selection-corrected gap was decreasing along the distribution. Finally, for the period from 2005 to 2009, the selection-corrected gap took on an inverted u-shape.

Given the different developments of selection into employment by age and education between men and women, Figure 2.6 shows the selection-corrected gender wage gap by age and education. Selection-corrected gender wage gaps by age make much more apparent to which extent the gender wage gap increases with age. Thus, among the youngest group in their 20s, the gap amounts to 20 log points. For the group in their 30s, the gender wage gap climbs up to 32 log points. Lastly, among those in their 40s, the gender wage gap exceeds the 40 log point mark.

The lower panel of Figure 2.6 shows the selected-gender wage gap by type of educational degree. After correcting for selection, the gap becomes much higher for those holding a basic degree and exceeds 40 log points over the whole period under study (unlike the observed gap for this group, which stayed fairly constant at around 25 log points). The results for the group holding an intermediate degree are not altered by correcting for selection. The gap for those with an advanced degree increases 6 log-points at the median and the observed glass ceiling disappears with respect to the observed figures in the data. All in all, our selection-corrected results confirm that a slight sticky floor

Figure 2.6: Selection-corrected gender wage gap over time, by age and education

(a) Age



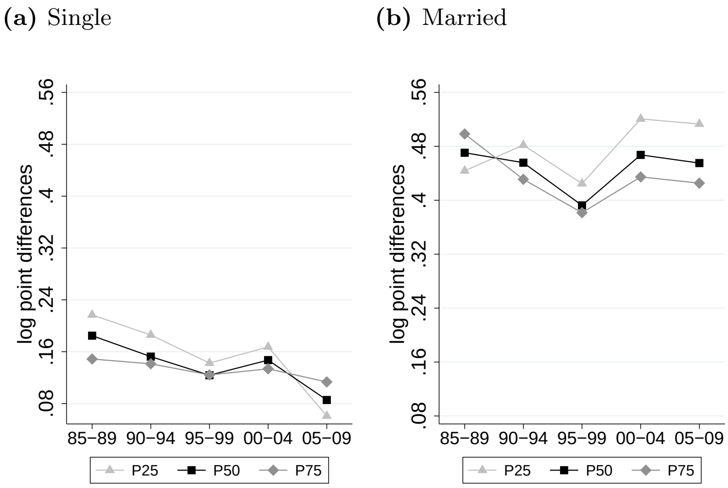
(b) Education

Source: SOEP.v32, own calculations.

arises for women with a basic degree, and there is a moderate glass ceiling for women with intermediate degrees.

Correcting for selection into employment has a strong impact on the gender wage gap by marital status, especially for those married (see Figure 2.7). The level and evolution of the selection-corrected gender wage gap among single individuals mimics the one observed in the data and confirms the shrinking of the gap for this group. However,

Figure 2.7: Selection-corrected gender wage gap, by family status



Source: SOEP.v32, own calculations.

among married individuals the selection-corrected wage gap is much higher than the observed one and no such declining trend over time can be discerned.

2.4.4 Decomposition results

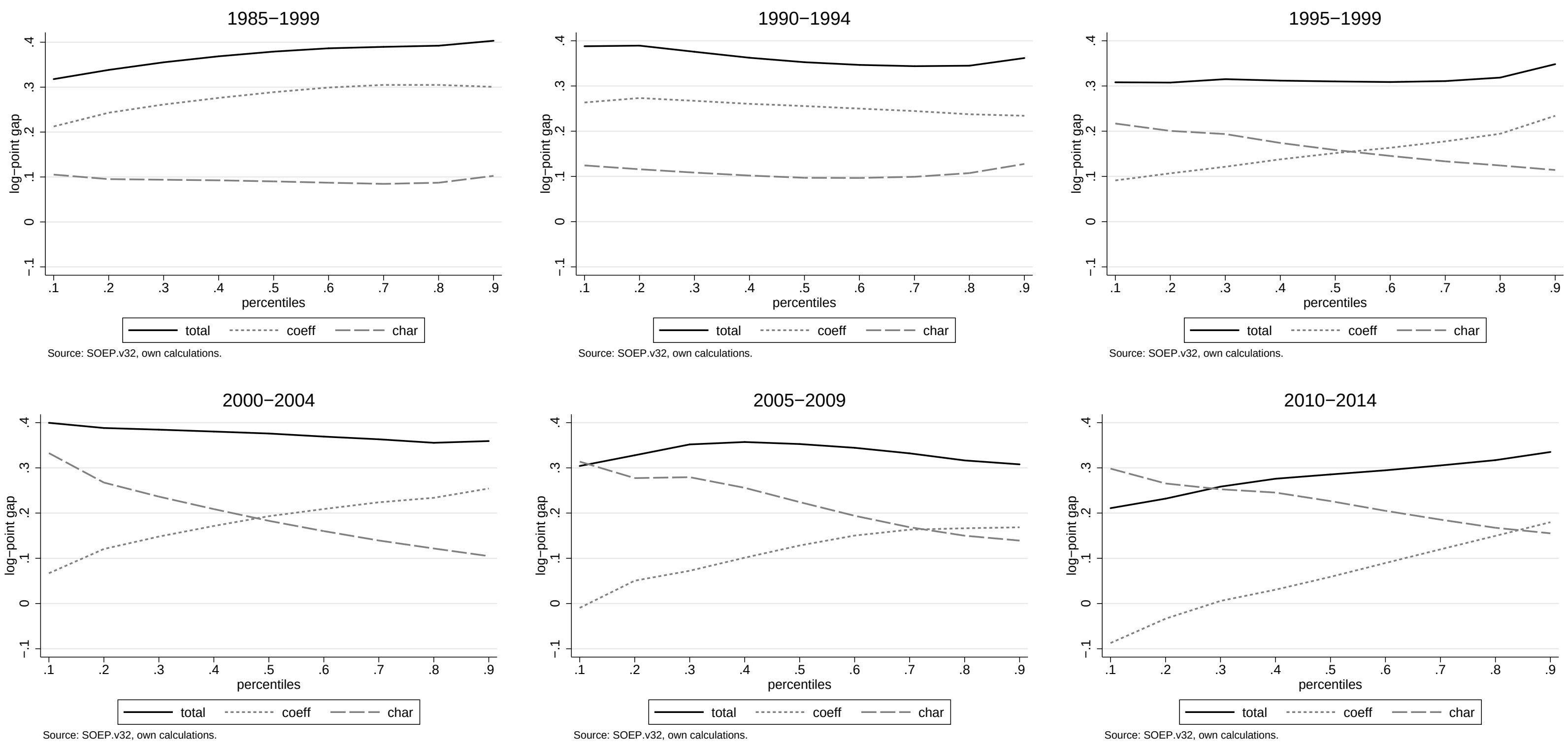
Until here, we have presented figures on the gender wage gap over the distribution and over time. Next, we present the results of decomposing the selection-corrected gender wage gap into the effect of labour-market relevant characteristics and the returns to those characteristics in order to better understand the trends in the gender wage gap described so far.

The decomposition of the selection-corrected gender wage gap into a part explained by differences in characteristics and a part that is explained by differences in coefficients reveals that the relative importance of these two parts has been changing over time. At the beginning of our observation period, the differences in men's and women's characteristics such as age, education and actual full-time and part-time work experience, explained about 10 log-points of the gap over the whole wage distribution. The remaining part of the gap can be explained by differences in coefficients. The relatively high importance of differences in coefficients in explaining the gender wage gap is partly due to the fact that several important variables such as occupation, industry or company size are not accounted for in the wage equation. Hence, the effect of coefficients partly captures the existence of strong occupational segregation between men and women in Germany (see, e.g. Wrohlich and Zucco 2017) as well as segregation into industries and companies of different firm size (Hinz and Gartner 2005).

Beginning in the mid-1990s, the differences in coefficients become less important in explaining the gap at the lower bound of the wage distribution, while differences in characteristics gain importance. The latter is in line with the increasingly positive selection of women into full-time employment (see Section 2.4.2). In the years 2005 to 2009, differences in coefficients have essentially disappeared at the bottom of the wage distribution. In other words, the selection-corrected gender wage gap of those (potentially) earning very low wages can be entirely explained by differences in characteristics such as education and actual work experience. That we still find a large selection-corrected gender wage gap in this part of the wage distribution stems from the fact that characteristics of women in this group, in particular actual work experience, strongly differ from those of men in this group.

For the median, the relative importance of differences in characteristics and differences in coefficients has also changed: in the 1990s, differences in coefficients explained almost two thirds of the gap at the 50^{th} percentile. In the following years, their relative importance has been decreasing (and is lower than the importance of characteristics in the years 2005 to 2009). The same holds true for the top of the wage distribution: in the late 1980s, differences in characteristics explained two thirds of the gap at the 90^{th} percentile. In other words, if women at the top of the wage distribution had exactly the same characteristics in terms of education and actual work experience, there would have still been a (selection-corrected) gap of 30 log points. In the years 2005 to 2009, this number has dropped to below 20 log points. This means that the decline of the selection-corrected gender wage gap at the top of the wage distribution can be mostly explained by the fact that personal characteristics such as age, education and actual work experience are rewarded more similarly for men and women in the late 2000s than 20 years before.

Figure 2.8: Decomposition of the gender wage gap, across the distribution and over time



2.5 Conclusion

The analysis of the observed and the selection-corrected gender wage gap over the past decades in West Germany has revealed five major insights. First of all, the observed gender wage gap displays a slight u-shape over the wage distribution, with slightly higher values at the bottom and at the top of the distribution than at the median. However, these differences are not as pronounced as in other countries, where the gender wage gap at the top of the distribution is much higher than at the median.

Second, we find that the observed gender wage gap in Germany has been declining over time. This is particularly true for the 1980s and 1990s, when the observed gender wage gap at the median dropped from 26 to 20 percent. Since the beginning of the 2000s, the decline of the observed gender wage gap has slowed down markedly, but it still amounts to 18 percent at the end of our observation period in 2009. The analysis also shows that the decline of the gender wage gap over time has occurred more or less in parallel over the entire wage distribution.

Third, we find the sample of women working full-time to be strongly positively selected at all points of the wage distribution, and the magnitude of the selection to increase over time. The sample of men working full-time is also slightly positively selected, but less so than for women. However, the magnitude of selection at the lower bound of the male distribution has also increased steeply from the beginning of the 1990s onwards. These findings are noteworthy because they happen in a context of constant (full-time) employment rates as well as of rising wage inequality.

Fourth, as a result of the revealed trend of selection into full-time employment, the selection-corrected gender wage gap is much higher than the observed gap. This suggests that studies failing to correct for positive selection into employment are likely to overestimate the convergence of male and female wages over time. In fact, we show that the selection-corrected gender wage gap has been narrowing during the 1980s and 1990s but it increased again in the 2000s.

Finally, the decomposition analysis shows that the effect of differences in coefficients decreases over time. This means that characteristics such as education and work experience are rewarded more equally for men and women by the late 2000s than in the beginning of our observation period. This is particularly true for the bottom of the wage distribution. At the tenth percentile, the total selection-corrected gender wage gap can be explained by differences in characteristics such as education and work experience - there are no differences in coefficients any more. In other parts of the wage distribution, differences in coefficients are still present and increase with wages. For women earning high wages, differences in coefficients still contribute more to the total gender wage gap than differences in characteristics.

To sum up, by the end of the 2000s, the observed gender wage gap is smaller than 25 years before - but the selection-corrected gender wage gap is not. As this chapter has shown, this can be mainly attributed to stronger selection into full-time employment by women over this time period. These decades have been marked by a dramatic expansion of female labour market participation - particularly in part-time employment. However, this has not led to a more representative sample of women working full-time.

From a policy perspective, we would argue that the key question is: to which extent does this development reflect women's choices or rather constraints they face in terms of participating in full-time employment? Policies aiming at reducing barriers - such as measures taken in order to improve the availability and affordability of childcare, avoid tax disincentives for second-earners or facilitate the transition between full- and part-time employment - should have a positive effect on lowering the gender wage gap. Furthermore, as the decomposition analysis in this chapter has shown, differences in characteristics between women and men are still prevalent and explain a large part of the gender wage gap. Differences in terms of educational attainment have already narrowed. In fact, among younger generations, women now even surpass men regarding years of schooling. However, there are still large differences in actual work experience, both due to differences in employment interruptions and periods of part-time employment. As long as these large differences in work experience between men and women remain, we will continue to observe a gender wage gap even if returns to skills are similar for men and women. Policies aimed at reducing inequalities in the employment careers between men and women will thus also contribute to reducing the gender wage gap.

Recent reforms of the parental leave benefits that have shortened employment interruptions of mothers and created incentives to take parental leave for fathers are one step in this direction. Additionally, child care reforms have also facilitated the reconciliation of family and work life and helped shorten family-related employment interruptions. Therefore, these two types of policy reforms can be expected to also contribute to a decrease in the gender wage gap in the future.

Chapter 3

The Part-Time Wage Gap across the Wage Distribution*

3.1 Introduction

Part-time work has undergone a major expansion in many labour markets during the last decades (OECD 2016). At the same time, full- and part-time wages have grown increasingly apart, leading to an ever-increasing raw part-time wage penalty (Manning and Petrongolo 2008, OECD 2010). A good understanding of this penalty is crucial, as it amplifies the adverse economic consequences of part-time employees' inherently lower earnings[26] and also potentially yields negative labour supply effects. The observed part-time wage gap could be the result of lower returns to skill in part-time employment as well as the consequence of changing selection patterns into full- and part-time employment, especially in a context of rising wage inequality (Blundell et al. 2007, Mulligan and Rubinstein 2008, Arellano and Bonhomme 2017, Biewen et al. 2017) and a general increase in female labour market participation (Jacobsen et al. 2015, Blau and Kahn 2017). It is the aim of this chapter to analyse to which extent the raw part-time wage gap can be explained by part-time employment being paid worse

[26] Bardasi and Gornick (2008) identify weaker insurance in social security schemes and lower old-age pension incomes as economic costs associated to part-time employment.

than full-time employment or by composition effects resulting from different selection patterns into full- and part-time work.

Until now, the literature has focused on estimating the average effect of part-time work on wages controlling for several factors such as individuals' labour market characteristics or occupational choices. At the same time, the literature has shown that the effect of composition effects, returns to skill and institutional factors on wages can strongly vary over the wage distribution, and even more so in the prevailing context of rising wage inequality (see DiNardo et al. 1996 and Dustmann et al. 2009, among others). Therefore, this paper goes beyond the analysis of average wage effects of part-time employment and analyses the part-time wage gap along the wage distribution. The goal of this chapter is threefold: First, I present a measure of *corrected* part-time wage gap that isolates differences in the wage structures of full- and part-time employment from composition effects of the two groups. Secondly, I present results on the magnitude of sample selection into those two kinds of work arrangements. Last, I analyse the evolution of these factors over time for the twenty years from 1990 to 2009.

The empirical strategy is straightforward: Using quantile regression methods, I characterize selection-corrected cumulative log hourly wage distributions for full- and part-time employment. Deriving these two selection-corrected wage distributions requires imputing both a full- and a part-time wage gap for each woman each year whenever one or both of them is not realized. To this end, I use an imputation-based method to correct for sample selection issues developed by Melly and Santangelo (2015) which is based on an extension of Athey and Imbens (2006)'s changes-in-changes model. Once I obtain selection-corrected full- and part-time wage distributions, I evaluate how distant these resulting distributions are from each other at different quantiles. I refer to this difference as the corrected part-time wage gap. This measure presents estimates on the wage effects of part-time employment that correct for different observed and largely also unobserved characteristics of individuals working in each of the two work arrangements, as it relies on two wage distributions with the same underlying joint distribution of labour market characteristics.

Melly and Santangelo (2014)'s imputation method - originally conceived to recover non-realized full-time wages - uses information from an individual's realized wage obtained from longitudinal data and assumes the time-invariance of the unobservables conditional on the observables to impute this individual's wage whenever she is out of work. For the analysis of the female part-time wage gap[27], I exploit the fact that many women in Western economies work both full- and part-time at some point in their lives (see e.g. Connolly and Gregory 2010 for the UK or Paul 2016 for Germany) to extend Melly and Santangelo (2015)'s method to the imputation of non-realized part-time wages in addition to the imputation of full-time wages. For women only observed working either full- or part-time, or never at all, I need an additional assumption regarding their unobservables. I discuss different options and perform several robustness checks that confirm the stability of the results presented.

For the empirical analysis in this paper, I use the German Socio-Economic Panel (GSOEP), a rich longitudinal dataset with detailed information on individuals' earnings and working hours. I restrict the sample to prime-age women residing in West Germany and focus on the period 1990 to 2009, two decades in which part-time employment has dramatically expanded and the raw part-time wage gap has widened across the whole distribution. In a context of rising (female) wage inequality, I show that the widening of the raw part-time wage gap over time is the joint result of part-time employment gaining presence in the low-wage sector and full-time employment increasing its presence in the high-wage sector. My findings reveal that differences between the wage structures of the two kinds of employment, as captured by the corrected part-time wage gap, have also undergone substantial changes over this time period and, from the 2000s onwards, have operated in the opposite direction than the raw gap. In fact, at the beginning of the 1990s a sizeable corrected wage penalty for part-time workers is found for the lower end of the distribution, a zero effect at the median, and a substantial premium at the top of the distribution. During the 2000s, the returns penalty at the lower half of the distribution disappears and I find a part-time returns premium for the entire wage distribution.

[27]It would be interesting to carry out the empirical analysis for men too, as the data also reveals a part-time wage gap for male employees. Unfortunately, this is not feasible because of too few observations in the data of men working part-time in Germany during this time period.

The broadening difference between the raw and corrected part-time wage gap hints at important dynamics of sample selection into the two kinds of employment. Concretely, I find strong positive selection into full-time work - indicating that women with high potential wages tend to select into full-time work much more often than women with lower potential wages. In addition, the magnitude of (positive) selection is highest at the lower end of the distribution and has been steeply increasing over time. On the contrary, sample selection into part-time employment changes from being positive by the beginning of the 1990s to almost disappearing by the end of the 2000s, and even turning slightly negative at the lower end of the wage distribution. Both findings fit and can be explained by the theoretical framework linking wage inequality and the employment rate to the selection bias as proposed by Mulligan and Rubinstein (2008). These findings emphasize the substantive role of sample selection in the analysis of female wages, even in the presence of stable employment rates. This is particularly important for studies on the evolution of the gender wage gap over time that are based on full-time workers, which are likely to underestimate the gender wage gap and to overestimate the convergence of women's relative wages to men[28] if not controlling for sample selection.

This chapter contributes to the existing literature on the effect of part-time employment on wages, which can be classified into three big strands according to its methodology: studies which use standard least square wage models with or without selection correction - often accompanied by decomposition exercises; analyses which are based on panel estimators of wage equations; and joint estimations of hours and wage equations.

Within the first strand, Blank (1990) was the first study to produce estimates on the part-time pay penalty controlling for individuals' skills and sample selection, and she found substantial penalties for the United States in the late 1980s. Manning and Petrongolo (2008) estimate a variety of models and find that differing workers characteristics in the full- and part-time group as well as growing occupational segregation and rising wage inequality practically explain the totality of the observed rising part-time wage gap in the UK for the period 1975 to 2001. Bardasi and Gornick (2008) and

[28]This is the case if small and constant sample selection into full-time employment for men is assumed.

Matteazzi et al. (2014) carry out cross-country studies on the average wage effects of part-time work and find occupational segregation to be a major determinant of the observed part-time pay penalty in most countries. In all these studies, sample selection is taken care of through the Heckman Two-Stage Model Heckman (1979) and the chosen exclusion restrictions are based on household characteristics such as - but not limited to - marital status or number of children.

The next strand of the literature analyses the wage effects of part-time work by estimating a fixed-effects wage model. In such models, the average effect of part-time work on wages is identified by the wage changes of individuals transitioning between full- and part-time status relative to the wage changes of individuals who stay in each type of employment. Hirsch (2005) finds a very small penalty for women in the United States once he controls for individuals' and job skills, whereas Booth and Wood (2008) find a female part-time premium for Australia. Connolly and Gregory (2009) and Fouarge and Muffels (2009) also look at the long-term consequences on part-time work on wages and find long-term earnings losses for all countries under study.

The third group of papers takes into account the potential endogeneity between working hours and wages by conducting joint estimations of hours and wage equations. Aaronson and French (2004), making use of an old-age part-time social security rule in the United States as exclusion restriction, find a part-time penalty for men but not for women. Paul (2016), using exclusion restrictions from the German institutional context, distinguishes between short- and long-hours part-time work and finds a female part-time penalty on current wages for the first group but not for the latter.

This chapter is structured as follows. Sections 3.2 and 3.3 cover the empirical strategy and describe the data. Next, Section 3.4 presents descriptive evidence on the raw part-time wage gap across the distribution and over time and discusses it in the context of dramatic changes in the distribution of female wages. Section 3.5 presents the results of the imputation of non-realized full- and part-time wages. Section 3.6 discusses the main findings of the paper in terms of the corrected part-time wage gap. The magnitude and evolution of selection into full- and part-time employment is then examined in Section 3.7. Last, Section 3.8 concludes and draws policy recommendations.

3.2 Empirical strategy

In this section I discuss the empirical strategy for deriving a measure of the part-time pay gap that isolates the wage effect of part-time employment while controlling for different characteristics of the individuals working in full- and part-time employment. Section 3.2.1 presents formally the concepts used throughout the paper. Section 3.2.2 describes the imputation method used to correct for sample selection into full- and part-time employment. Last, Section 3.2.3 discusses the specification of the conditional wage model used for the imputations.

3.2.1 Corrected part-time wage gap and selection-corrected wage distributions

In what follows, I denote X_t^{FT}, X_t^{PT} and X_t^{OW} to be the joint distribution of human capital variables at time t entering the wage model for three different subsamples: individuals working full-time (X_t^{FT}), part-time (X_t^{PT}) and out-of-work (X_t^{OW}). The joint distribution of human capital variables for the whole sample is denoted $X_t = \left\{X_t^{FT}, X_t^{PT}, X_t^{OW}\right\}$. The distributions of log hourly wages for full- and part-time employment are referred to as W^{FT} and W^{PT}, respectively. Throughout this section, $F(\cdot)$ denotes unconditional cumulative density functions and τ stands for the unconditional rank in a given inverse cumulative function. The observed full- and part-time log wage distributions, $F_{W_t^{FT}}$ and $F_{W_t^{PT}}$, are only defined for X_t^{FT} and X_t^{PT}, respectively.

However, $\widehat{F}_{W_t^{FT}}$ and $\widehat{F}_{W_t^{PT}}$ are both defined for all $X_t = \left\{X_t^{FT}, X_t^{PT}, X_t^{OW}\right\}$ and consist of observed (realised) as well as imputed (non-realised) log wages. $\widehat{F}_{W_t^{FT}}$ and $\widehat{F}_{W_t^{PT}}$ are selection-free by definition, as each woman in the sample has been inputed both a full- and a part-time wage, whenever one (or both) of them are not realised. This builds on the idea that sample selection issues can be controlled for once a wage rate for every individual in the sample - especially those individuals for which a wage is not observed - is imputed (see Neal 2004, Blau and Kahn 2006 and Olivetti and Petrongolo 2008 for earlier imputation-based approaches to control for sample selection issues).

Following this notation, the raw part-time wage gap at percentile τ and time t can be expressed as the difference between the observed full- and part-time log wage distributions:

$$G_{raw}(\tau,t) = F^{-1}_{W^{FT},t}(\tau) - F^{-1}_{W^{PT},t}(\tau) \tag{3.1}$$

The corrected part-time pay gap is defined by the difference between the selection-corrected log wage distributions for full- and part-time work in a given year t, $\widehat{F}_{W^{FT},t}$ and $\widehat{F}_{W^{PT},t}$. Because these two distributions apply to the same underlying population, any difference between the two is due to unequal wage structures of full- and part-time employment:

$$\widehat{G}_{corr}(\tau,t) = \widehat{F}^{-1}_{W^{FT},t}(\tau) - \widehat{F}^{-1}_{W^{PT},t}(\tau) \tag{3.2}$$

Note that by adding and subtracting equation 3.2 to equation 3.1, and rearranging the terms, the raw part-time wage gap can be also expressed as[29]:

$$G_{raw}(\tau,t) = \widehat{G}_{corr}(\tau,t) + \widehat{Sel}^{FT}(\tau,t) - \widehat{Sel}^{PT}(\tau,t) \tag{3.3}$$

where $\widehat{Sel}^{FT}(\tau,t) = F^{-1}_{W^{FT},t}(\tau) - \widehat{F}^{-1}_{W^{FT},t}(\tau)$ quantifies the magnitude of selection into full-time employment and $\widehat{Sel}^{PT}(\tau,t) = F^{-1}_{W^{PT},t}(\tau) - \widehat{F}^{-1}_{W^{PT},t}(\tau)$ the selection into part-time employment.

Equation 3.3 implies that the difference between the raw and the corrected gap can be explained in terms of sample selection into both types of employment. In terms of interpretation, this implies that in case of identical returns to skill of full- and part-time employment $\widehat{G}_{corr}$ would be zero and any observed difference between full- and part-time wages would be only due to composition effects of individuals selecting into each kind of work arrangement.

[29] This can be understood as a usual decomposition exercise in the spirit of Oaxaca (1973) and Blinder (1973), or its quantile equivalent proposed by Machado and Mata (2005), with a particular interpretation of the composition effect.

3.2.2 Imputation of non-realised wages

This section explains the imputation procedure used for recovering the sets of non-realized full- and part-time wages needed to obtain the selection-corrected wage distribution, $\widehat{F}^{-1}_{W^{FT},t}(\tau)$ and $\widehat{F}^{-1}_{W^{PT},t}(\tau)$.

First, I briefly outline the method proposed by Melly and Santangelo (2015)[30] and then proceed to present its application to the study of full- and part-time wages. Intuitively, Melly and Santangelo (2015) suggest to build subsamples with individuals who work in two given periods (group 0) and compare them to subsamples of individuals that only work in one of these two periods (group 1). Group 1 individuals reveal information on their unobservables in the period when they work, which is captured by their conditional rank in the wage distribution. The evolution of wages of group 0 allows imputing a conditional wage for group 1 which responds to the wage structure of the time when the imputation is required and accounts for both individuals' observable and unobservable characteristics.

This exercise is carried out for all possible combinations of two survey years in the data, which I refer to as $t \in \{k, l\}$. Group assignment according to the description above is captured by variables G_{kl}, each with realisation $g_{kl} \in \{0, 1\}$. The covariates entering the wage equation are captured in vector of variables X and the unobservable component of wages in random variable U.

The identifying assumption behind this imputation method is the time-invariance of the unobservables conditional on the observables, $f_{U|T,G,X} = f_{U|G,X}$. In words, the distribution of unobservable characteristics which have an effect on wages - such as innate ability or professional ambition - stay constant over time conditional on observable characteristics such as education level and previous work experience. The specification of vector X needs to capture as much variation in unobservable components as possible. It is therefore discussed separately in the next subsection.

Formally, Melly and Santangelo (2015) show that the conditional wage distribution of those individuals not working in time period $t = k$ but working in time period $t = l$

[30] See Melly and Santangelo (2015) for details. The aim of this section is merely to sketch the imputation method proposed by the authors so as to make it understandable - for a thorough explanation of the model it is advised to refer to the original sources.

can be derived as:

$$F^{-1}_{W|g=1,t=k,x}\left(\theta\right) = F^{-1}_{W|g=0,t=k,x}\left(F_{W|g=0,t=l,x}\left(F^{-1}_{W|g=1,t=l,x}\left(\theta\right)\right)\right) \tag{3.4}$$

and individual wages conforming $F^{-1}_{W|g=1,t=k,x}\left(\theta\right)$ can be imputed as:

$$\widetilde{w}_{ikl} = x_i\hat{\beta}_{g=0,t=k}\left(\int_0^1 \mathbb{1}\left(x_i\hat{\beta}_{g=0,t=l}\left(u\right) \leq x_i\hat{\beta}_{g=1,t=l}\left(\theta\right)\right) du\right) \tag{3.5}$$

where $\hat{\beta}_{g,t}\left(\theta\right)$ are the wage equation coefficients for quantile θ coming from the estimated conditional quantile regression processes. Because each realised full- or part-time wage provides enough information to impute a non-realized wage, for individuals observed working several years the imputation rule of expression 3.5 produces multiple available imputations for a single non-realized wage. For these cases, Melly and Santangelo (2015) suggest to weigh all available imputations so as to obtain a final imputed wage for each individual i in year k, $\widetilde{w}_{ik}$:

$$\widetilde{w}_{ik} = \sum_{m=1}^{M_i} d_{ikm} \cdot \widetilde{w}_{ikl} \tag{3.6}$$

where M_i is the number of available imputations for a given $\widetilde{w}_{ik}$, different for each individual i, and d_{ikm} is the weighting factor for each $\widetilde{w}_{ikl}$, so that $\sum_{m=1}^{M_i} d_{ikm} = 1$.
In my application, I allow the wage structure of full- and part-time employment to differ from each other, and carry out the imputation procedure separately for the two kinds of employment[31]. The identification of non-realised full-time wages requires the expression $f_{U^{FT}|T,G^{FT},X^{FT}} = f_{U^{FT}|G^{FT},X^{FT}}$ to hold, whereas the identification of non-realised part-time wages requires $f_{U^{PT}|T,G^{PT},X^{PT}} = f_{U^{PT}|G^{PT},X^{PT}}$. Thus, not only the returns to observable characteristics are allowed to differ between full- and part-time employment but the structure of unobservables is also allowed to differ in both kinds of work arrangements. Hence, individuals can have a different unobservable in the full- than in the part-time wage distribution. This feature accommodates the cases in which the transition from one kind of work to the other is associated with wage losses (a

[31]See Ermisch and Wright (1993) for a discussion of reasons why full- and part-time employment are likely to present different wage structures.

downward move in the conditional rank) or gains (an upward move in the conditional rank).

Consequently, the assignment to a two-year subsample is done separately for full- and part-time employment and is captured by variables G_{kl}^{FT} and G_{kl}^{PT}, each with realisation $g_{kl}^{FT} \in \{0,1\}$ and $g_{kl}^{PT} \in \{0,1\}$ for all $\{k,l\}$, respectively. The algorithm in equation 3.5 is applied separately for full- and part-time wages. Given the high computational burden and the relatively small size of the estimation subsamples in the empirical application, I use a slightly modified imputation algorithm:

$$\widetilde{w}_{ikl}^{FT} = x_i \hat{\beta}_{g^{FT}=0,t=k} \left(\int_0^1 \mathbb{1} \left(x_i \hat{\beta}_{g^{FT}=0,t=l} \left(u\right) \leq \overline{w}_{i,t=l}^{FT} \right) du \right) \tag{3.7}$$

$$\widetilde{w}_{ikl}^{PT} = x_i \hat{\beta}_{g^{PT}=0,t=k} \left(\int_0^1 \mathbb{1} \left(x_i \hat{\beta}_{g^{PT}=0,t=l} \left(u\right) \leq \overline{w}_{i,t=l}^{PT} \right) du \right) \tag{3.8}$$

where $\overline{w}_{i,t=l}$ is the observed wage for person i in $t = l$ and replaces its estimated equivalent $x_i \hat{\beta}_{g=1,t=l} \left(\theta\right)$ in Equation 3.5 above. Practically, this enables me to recover non-realised wages using only two sets of quantile-regression coefficients, $\left\{ \hat{\beta}_{g=0,t=l} \left(\theta\right) \right\}_{\theta=0.01}^{\theta=0.99}$ and $\left\{ \hat{\beta}_{g=0,t=k} \left(\theta\right) \right\}_{\theta=0.01}^{\theta=0.99}$, and do not need to estimate a third grid of conditional quantile regressions, $\left\{ \hat{\beta}_{g=1,t=l} \left(\theta\right) \right\}_{\theta=0.01}^{\theta=0.99}$.

The main results presented in Section 3.6 use an unweighed average scheme, $d_{ikm} = \frac{1}{M_i}$, as suggested by Melly and Santangelo (2015). Other possibilities include a closest neighbour approach or a weighted average that penalizes wider distances between year k and year l. The sensitivity of the results to different weighting schemes is examined in the Appendix (Section C.4.3).

With the imputation algorithm of Equation 3.7, I can only recover non-realised wages for those individuals observed working both full- and part-time for at least one time period in the data. There are three groups for whom this is not possible: those never observed working full-time, those never observed working part-time, and those never observed working at all. For these individuals, I apply a different imputation rule than the one described above, which I refer to as alternative imputation rule.

In particular, I predict wages for them by means of the median conditional quantile coefficients estimated on the most suitable distribution[32]. Missing full-time wages for women who have never worked full-time are filled with predictions based on conditional quantile regression coefficients from either $Q_{0.5}\left(\widetilde{w}_{it}^{FT}|X_t^{PT}\right) = x_{it}'\beta_t\,(0.5)$ if at period t the individual was working part-time or from $Q_{0.5}\left(\widetilde{w}_{it}^{FT}|X_t^{OW}\right) = x_{it}'\beta_t\,(0.5)$ if she was out of work. In both cases, $\widetilde{w}_{it}^{FT}$ refers to full-time wages imputed according to the main model as captured by Equation 3.3. Non-realised part-time wages for women who have never worked part-time are filled equivalently. Whereas this alternative rule fully controls for observables, its drawback is that - contrary to the main imputation rule - it controls only partially for unobservables. Furthermore, this median imputation is arbitrary but does not *a priori* impose a sign on the potential sample selection into each of the two work arrangements under study. The effect of this additional assumption on the imputation results is discussed in Section 3.5 and several robustness checks are available in the Appendix.

3.2.3 Specification of the conditional wage model

The wage equation is estimated separately for the full- and the part-time wage distribution as a linear conditional quantile regression model (Koenker and Bassett, 1978):

$$Q_\theta\left(w_{it}|x_{it}\right) = x_{it}'\beta_t\,(\theta) \tag{3.9}$$

The dependent variable w_{it} is the natural logarithm of the hourly wage and the independent variables x_{it} consist of an intercept, age (polynomially, up to the power of three), years of education, an indicator variable for an intermediate degree, an indicator variable for an advanced degree and actual full- and part-time working experience (both polynomially up to the power of three, following Blau and Kahn (2013) on the importance of controlling for actual experience as well as the literature on the different long-run returns of full- and part-time working experience). Age is included as a regressor in addition to years of education and experience because many women experience spells of non-market work. This category is excluded to avoid an overspecified wage

[32]This is a different imputation rule than the one suggested by Melly and Santangelo (2015) for such cases. Section C.4.4 elaborates on the reasons in favour of this imputation rule.

model. The goodness of in-sample fit of the conditional wage model is examined in Appendix C.3.

Besides the goodness of fit, the specification of the conditional wage model is crucial for the fulfilment of the identifying assumption. A flexible, rich specification clearly contributes to this end, although the econometric approach used in this paper restrains the choice of covariates in the wage model to characteristics observed for all women (including those out of work) in all survey years. As a result, variables such as current occupation and industry branch cannot be included in the wage model without further assumptions. In order to determine whether this poses a problem, Appendix C.5 explores how well a specification in terms of human capital variables does with regard to the fulfilment of the identifying assumption and the results are predominantly positive.

3.3 Data

The empirical analysis is carried out on the basis of the German Socio-Economic Panel (GSOEP), version 32, for the years 1990 to 2009. The GSOEP is a representative household longitudinal study with data, among other information, on gross labour earnings and precise working hours (see Wagner et al. 2007 for details on the dataset), which renders it a suitable dataset to answer the present research question.

The GSOEP data has two shortcomings. First, information on earnings in survey data is known to contain measurement error[33]. However, Fuchs-Schündeln et al. (2010) find that wage dispersion measures based on the GSOEP and an administrative dataset from the Institute for Employment Research (IAB) follow similar trends during the 1990s. Second, the GSOEP includes a relatively small number of observations, which has a potential impact on the quality of the imputations. Unfortunately, available administrative datasets for Germany have either no longitudinal dimension or they lack detailed information on working hours, which are two necessary pre-requisites for the analysis at hand.

[33] At the same time, the literature has questioned the extent to which measurement error in survey data is more severe than in administrative datasets (see Kapteyn and Ypma 2007 and Abowd and Stinson 2013).

I restrict the estimation sample to women aged 20 to 55 with residence in West Germany. Furthermore, individuals in retirement, self-employment, the military, and disabled people are excluded. Women pursuing education or on vocational training are kept on the estimation sample and are coded as non-working. Women who only participate one year in the GSOEP are dropped, since I require at least two observations per individual in order to be able to use this information in the imputation procedure. These sample restrictions leave us with a total of 6,448 female individuals in our sample. In terms of person-year observations, these represent a total of 71,044 observations which oscillates between 2,700 to 4,700 observations per year (see column (1) in Tables C.1 and C.2 in the Appendix for the exact figures).

The dependent variable, the natural logarithm of the hourly wage, has been constructed by dividing gross labour earnings over working hours in the main job. Earnings and working hours from a second job are not taken into account. Observations with a resulting hourly wage under one euro have been dropped.

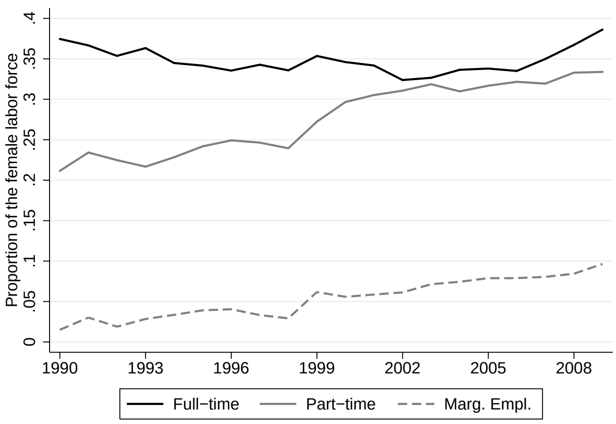
Working hours are defined as actual weekly working hours. In the main specification, women working more than 32 hours per week have been coded as working full-time, while those working up to 32 hours have been coded as working part-time[34]. This threshold has been chosen in order to obtain balanced full- and part-time samples (i.e. the highest possible number of women working in both full- and part-time employment). In the Appendix, I replicate the results with alternative thresholds defining full- and part-time employment (Section C.4.1) as well as using contractual hours plus paid overtime instead of actual working hours (Section C.4.2). In all cases, the results differ only minimally from the main ones and remain qualitatively the same. Women working less than 4 hours per week have been coded as not working[35]. Marginal employment, the term used for employment that pays below a given threshold established by law and for which alternative taxation and social security rules apply, is treated as part-time employment throughout the paper. This is because the low limit of earnings allowed under this kind of work arrangement generally imply a low number of working hours.

[34]The literature has also explored statistically at which hours threshold it arises a part-time penalty (see Averett and Hotchkiss 1996), but this is outside the scope of this paper.

[35]Hourly wages corresponding to very low working hours appear to be unreasonably high in the GSOEP, hinting at measurement error issues.

The evolution of full- and part-time employment rates according to the 32 weekly hours threshold are shown in Figure 3.1. The full-time employment rate has been quite stable around the 35 percent level whereas the part-time employment rate has dramatically increased during the twenty years under study and by the 2010s has approached the level of full-time employment[36]. Figure 3.1 also depicts the evolution of marginal employment over time (the dashed line in the graph)[37]. It has gained a lot of relevance in terms of the number of individuals engaging in this kind of work, which has gone from being every tenth part-time worker at the beginning of the 1990s to slightly more than every fourth twenty years later.

Figure 3.1: Full- versus part-time employment rates



Source: SOEP.v32, own calculations. Weighting factors used.

Relevant for the imputation procedure are the number of individuals who I never observe working either full- or part-time, which needs to be sufficiently low. However, the share of person-year observations for whom I never observe a part-time wage is very large (higher than 50 percent) during the mid 1980s and the share of individuals for whom I never observe a full-time wage is very large from 2010 onwards[38]. For this

[36]See also Figure C.1 in the Appendix for the evolution of the distribution of weekly working hours over time.

[37]Figure 3.1 only depicts individuals in marginal employment who exclusively work in such an arrangement. Individuals with a regular part- or full-time job additionally working in marginal employment are picked up by the part- or full-time lines. This is consistent with the definition of gross earnings used, which only takes into account the main job of each individual, and explains why the depicted figures are lower than official statistics on the overall number of people working in so-called Minijobs.

[38]See column (5) in Tables C.1 and C.2 in the Appendix for exact year-by-year figures.

reason, the main analysis is restricted to years 1990 to 2009[39].

During this time period and according to the 32 hours threshold, I observe both a full- and a part-time hourly wage for 24 percent of all women in the sample. For 31 percent of women I only observe a full-time wage, whereas for 26 percent I only observe a part-time wage. For the rest 19 percent I do not observe any wage. This contrasts with the figures according to actual full- and part-time working experience, where 60 percent of women report having worked both full- and part-time, 26 percent say having worked only full-time, 9 percent part-time and only 5 percent report not having ever worked at all. This discrepancy is mainly due to women joining the GSOEP - as part of refreshment samples - at an age where they have already accumulated working experience, as well as to panel attrition issues. For the period 1990 to 2009, in terms of person-year observations, the percentage of women having never worked in any of the two kinds of arrangement oscillates between 29 percent and 41 percent, depending on the year (see column (5) in Tables C.1 and C.2 in the Appendix for year-by-year figures).

The second data issue regarding the imputation procedure are the subsample sizes on which the many quantile regressions are run. Subsamples are defined separately for individuals working full- and part-time. Within each of these two groups, subsamples are built with individuals working in two given years for all possible two-year combinations in the data. In terms of the previous notation, each two-year combination stands for $\{k, l\}$, where $k \in \{1990, ..., 2009\}$ and $l \in \{1984, ..., 2015\}$, which yields a total of 430 two-year combinations. Tables C.9 and C.10 in the Appendix show the number of observations contained each two-year combination in the data for full- and part-time wages, respectively. Table 3.1 shows descriptive statistics about the subsample sizes.

Quantile regression samples range from 47 to 1,189 observations in the case of full-time work and from 12 to 1,092 observations in the case of part-time work. Around half of all tight grids of quantile regression are estimated on subsamples that have at least 252 observations in the case of the imputations of full-time wages and 164 observations in

[39] However, the imputation procedure uses information on realised wages from all available survey years in the GSOEP at the time of writing (1984 to 2015).

Table 3.1: Subsample sizes of all possible two-year combinations

	min	P50	max	N
Full-time wages	47	252	1189	430
Part-time wages	12	164	1092	430

Source: SOEP.v32, own calculations.

the case of part-time wages.

For the imputation of part-time wages, I need to establish a cut-off regarding minimum subsample size. I have set this to 30 observations. As a consequence, I cannot use the imputation algorithm for 32 out of the 430 possible two-year combinations. This does not affect the results, because for most observations on which I would otherwise use the information from the dropped two-year combinations, there are multiple available imputations.

3.4 The raw part-time wage gap

This section presents descriptive evidence about the raw part-time wage gap across the distribution and over time and discusses it in a context of dramatic changes in the distribution of female wages.

Table 3.2 displays figures on the raw part-time wage gap, G_{raw}, summarized for four time periods covering the overall time span 1990 to 2009 and for five selected quantiles ($\tau = .1, .25, .5, .75, .9$). Given that differentials are computed as the difference between full-time log wages minus part-time log wages, positive figures imply part-time wage penalties whereas negative figures hint at a part-time wage premium.

Table 3.2 shows that the raw part-time wage gap differs greatly across the distribution and has undergone substantial changes over time. By the beginning of the 1990s, the raw part-time penalty amounted to 22 log points at the bottom of the distribution and 9 log points at the median. At the upper end of the distribution, one observes a 17 log point part-time premium. Over time, the raw part-time gap has increased at all points of the distribution, augmenting the wage penalties at the lower half of the distribution

Table 3.2: Raw part-time wage gap

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
G_{raw} ($\tau = .10$)	.22	.26	.35	.33
G_{raw} ($\tau = .25$)	.19	.21	.22	.28
G_{raw} ($\tau = .50$)	.09	.08	.16	.18
G_{raw} ($\tau = .75$)	-.02	.01	.08	.13
G_{raw} ($\tau = .90$)	-.17	-.21	-.00	.08

Comments: Units are log-point differences between the full- and part-time inverse cumulative wage distributions evaluated at $\tau = .10, .25, .50, .75, .90$.

Source: SOEP.v32,own calculations.

and turning the wage premiums at the upper end of the distribution into moderate wage penalties. By the end of the 2000s, the raw wage penalty at the lower end of the distribution had reached the 33 log points; at the median it amounted to 18 log points and at the 90^{th} percentile it rose to 8 log points.

In order to better understand these figures on the raw part-time wage gap, the two wage distributions underlying the figures in Table 3.2 are now examined separately. Whereas changes in inequality have been well-documented for the female full-time wage distribution in the literature (see Dustmann et al. 2009, Antonczyk et al. 2010 and Card et al. 2013) little is known about the evolution of the part-time wage distribution.

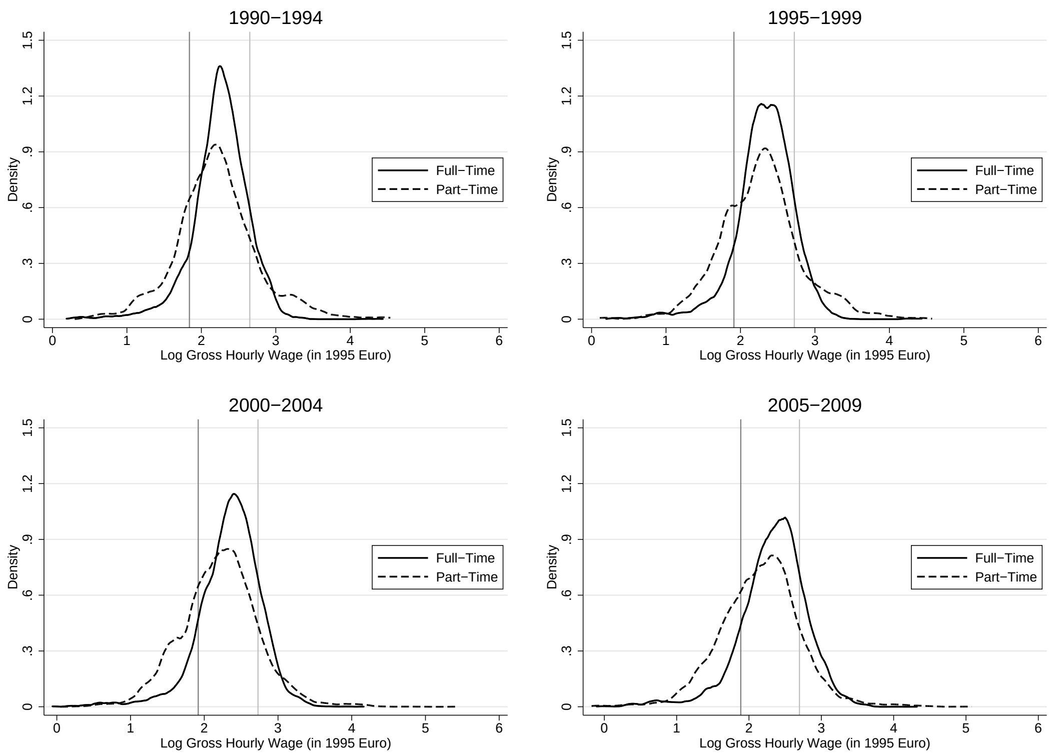
To this end, Figure 3.2 shows full- and part-time wage densities for the four time periods under study, deflated to 1995 euro[40]. Additionally, each graph displays two vertical lines depicting 67 and 150 percent of the overall female median wage, respectively. I refer to wages falling below 67 percent of the female median as the low-wage sector and to wages above 150 percent of the median as the high-wage sector.

Figure 3.2 clearly shows that the part-time wage distribution is more dispersed than the full-time distribution, as indicated by the crossing of the two densities at the lower- and upper-end of the distributions from the beginning of the 1990s to the mid 2000s[41]. Full-time wages are less dispersed than part-time wages throughout the whole period, but inequality within full-time wages has increased substantially from year 2000 onwards in

[40]Nominal prices deflated according to CPI indexes provided by the German Federal Statistical Office.

[41]Wage dispersion measures of the full- and part-time wage distributions are quantified in Table C.13 in the Appendix.

Figure 3.2: Observed full- and part-time wage densities



Source: SOEP.v32, own calculations.

a rather symmetric pattern between lower- and upper- tail inequality. On the contrary, dispersion measures of part-time wage gaps, which are higher throughout, display a moderate increase in the second half of the 1990s but are otherwise stable for the complete time span. Interestingly, the stability of the $P90 - P10$ measure for the part-time distribution over time results from the compression of the upper-half of the distribution and the dispersion of the lower-half of the distribution[42].

Consistent with the unequal evolution of full- and part-time wage dispersion measures, Figure 3.2 also shows that the distance between the full- and part-time marginal densities at the low-wage sector increases to the detriment of part-time wages over time. Over the years, the percentage of part-time workers with a low wage increases whereas the percentage of full-time workers receiving a low wage stays constant. Thus, whereas by the beginning of the 1990s slightly more than every second woman earning a low

[42]See Table C.13 in the Appendix for concrete figures on wage dispersion in my working sample.

wage was part-time employed, by the end of the 2000s, two thirds of women receiving a low wage were working part-time. At the upper tail of the distribution, Figure 3.2 reveals that the part-time distribution dominates during the 1990s but the trend is also reversed during the 2000s. In fact, 17 percent of part-time workers had a high wage at the beginning of the 1990s as compared to only 13 percent of full-time workers. By the end of the period under study, the picture had turned the other way around and only 13 percent of part-time workers and up to 20 percent of full-time workers had a high wage. As a consequence, part-time employees have become less frequent in the high-wage sector over time.

These developments in terms of inequality within full- and part-time distributions explain the widening of the raw part-time wage gap at the lower end of the distribution and the disappearance of the premium and emergence of a penalty at the higher end of the distribution. The next sections show to which extent these developments over time are due to the wage structures of full- and part-time employment, and to which extent they are the result of changing composition effects. In order to disentangle these factors, I need to characterize selection-corrected wage distributions which rely on the results of the imputation of non-realized full- and part-time wages.

3.5 Results of the imputation of non-realised wages

The imputation of non-realized full- and part-time wages requires estimating 170,280 conditional quantile regressions; the resulting wage equation coefficients are not reported for the sake of brevity[43].

Instead, Figure 3.3 summarizes the result of the imputation procedure[44] and depicts the evolution over time of observed (solid lines), imputed (dashed lines) and selection-corrected (dotted lines) hourly wages. The three upper row graphs plot observed,

[43]The number of regressions is calculated as the number of all two-year combinations in the data, 430, times the 99 percentiles at which the wage equations are estimated for each two-year combination, times 2 (once for each of the two years), and again times 2 - once for the full- and once for the part-time wage distribution.

[44]These imputation results correspond to using an unweighed average of multiple imputations (when available) and the alternative imputation rule presented in Section 3.2.2. A sensitivity analysis of the imputation results to these choices is provided in Section C.4.

imputed and selection-corrected wages for full-time employment at three selected points of the unconditional distribution ($\tau = .25; .5; .75$). The lower row plots the equivalent for part-time wages.

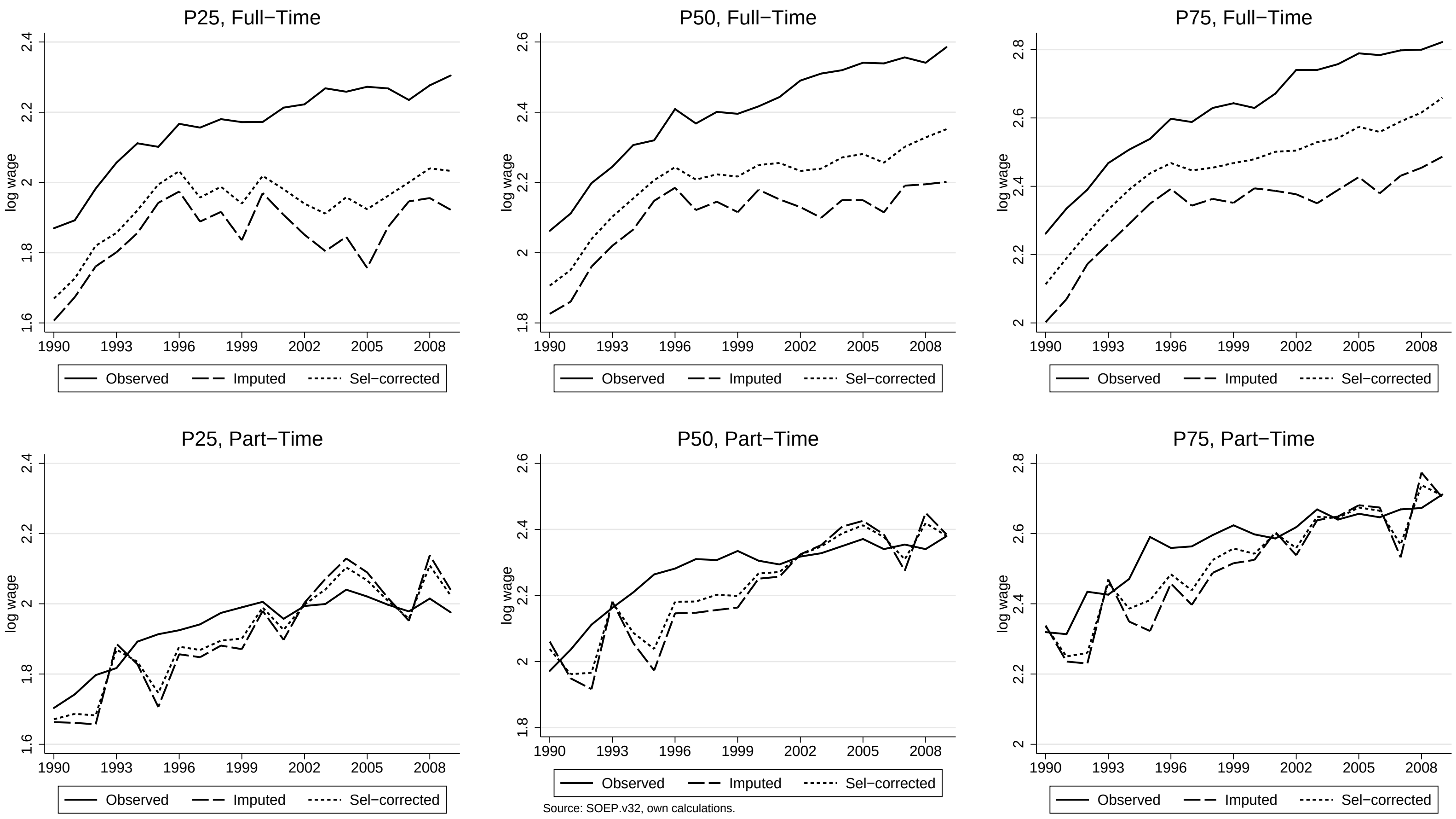
Dotted lines in Figure 2.5 represent the selection-corrected wage distributions, which are made up of both observed and imputed wages, and bound to lie between the two. The distance between the selection-corrected distribution and the observed one depends on the share of individuals requiring imputations as well as the difference in levels between imputed and observed wages. In the case of full-time employment, imputed wages are much lower than observed ones - hinting at strong positive selection into employment. For part-time employment, the selection-corrected distribution is much closer to the observed one and by the mid 2000s both are very similar.

A closer look at the imputation results reveals interesting insights. Full-time imputed wages are lower than observed values at all points of the distribution, which is consistent with large differences in terms of education and working experience between the subsamples working full-, part-time and out of work (see Table C.4 in the Appendix). Despite this level difference, the evolution of imputed wages in terms of (nominal) growth is parallel to observed wages for the period 1990 to 1996. From 1997 onwards, imputed wages display significantly lower growth than realised ones, causing the level difference between the two to rise over time. While the concrete causes behind this development are not identified by this approach, the flat profile of realized part-time wages suggests that one reason could be the slower accumulation of human capital stocks of the group working part-time relative to the group working full-time (a phenomenon which has been also documented for the UK in Blundell et al. 2016). The context of increasing full-time wage inequality (and corresponding higher returns to skill) can explain the growing divergence between realized and imputed wages, especially given the less favourable labour market characteristics of imputation subsample part-time employees and non-working individuals relative to the actual full-time sample.

Imputed part-time wages are closer to their observed counterparts than in the case of full-time employment. The reason for the smaller distance is that part-time imputations are carried out for two groups (those working full-time and those out-of-work) with large differences in terms of human capital, whose differences partially cancel out when regarded as one group. During the 1990s, part-time imputed wages are lower than realized ones but increase more steeply, so that during the 2000s they even surpass the level of observed part-time wages. In order to better understand this, Figure 3.4 examines imputed wages by full-, part-time or out-of-work status. One can explain the dynamics of imputed part-time wages as the result of three factors. First, imputed part-time wages for the sample working full-time display relative high (nominal) growth over the entire period. This is consistent with the evolution of observed full-time wages, which also exhibits such a trend, and hints at the fact that favourable labour market skills are highly rewarded in the part-time wage structure as well. Second, imputed part-time wages for the out-of-work group display a certain convergence to observed part-time wages, especially at the lower half of the distribution. Third, over time the weight in the overall part-time imputation between the group working full-time and the out-of-work group has evolved to the advantage of the full-time working group, as the out-of-work group represents a decreasing share of the overall female sample over time.

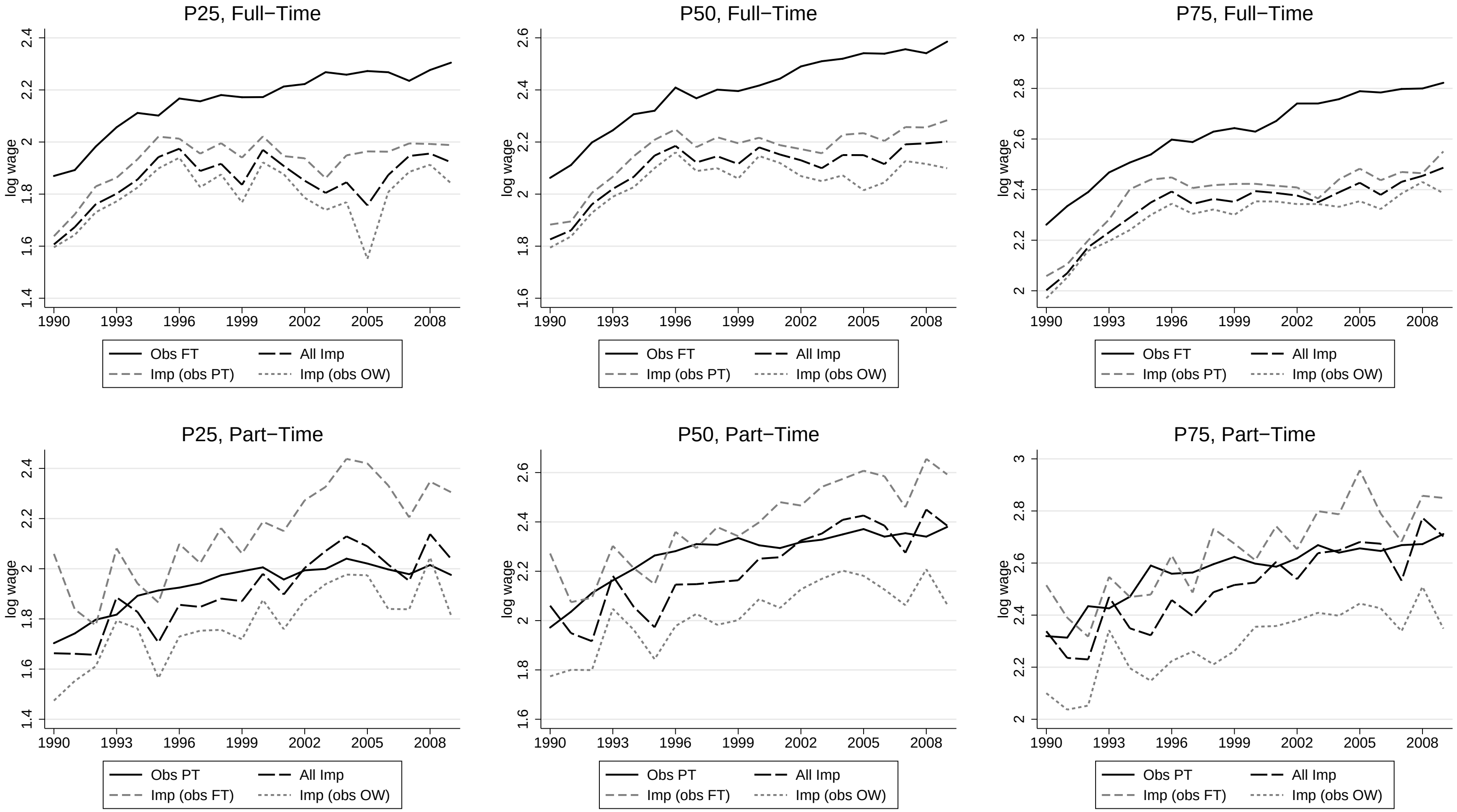
The imputations depicted in Figures 3.3 and 3.4 entail imputed wages according to both the main imputation method as well as the median imputation used for those women never observed working in one of the two kinds of employment. In fact, the distribution of wages imputed with median coefficients differs from that of wages imputed according to the main model (see Figure C.3 in the Appendix). However, this is consistent with the fact that individuals only working in one of the two kinds of employment have different characteristics (see Table C.5). Nonetheless, this does not yet clarify to which extent my results are driven by the additional assumption made on the unobservables of this group of individuals, which depends on the dispersion of conditional wages. In order to explore this issue, in Section C.4.4 I impute wages for these two groups with randomly assigned conditional ranks. This check suggests that the median imputation is not driving the results for the central part of the distribution, but it does have an effect on the magnitude of the results at the two tails of the distribution - the sign of the results, however, remains unchanged across the whole distribution.

Figure 3.3: Observed, imputed and selection-corrected wage distributions



Source: SOEP.v32, own calculations.

Figure 3.4: Imputations by full-time, part-time and out-of-work status



Source: SOEP.v32, own calculations.

The median imputation approach entails a second caveat. Given the large number of women who are not observed in both types of employment, the question is whether the counterfactual full- or part-time employment for this group exists in reality. Given that a majority of women do work both full- and part-time over the life cycle - I just do not observe their wages because they are part of refreshment samples in the GSOEP - it seems that transitioning between states is feasible for a large share of women. However, it is possible, at least theoretically, that individuals with particular skill combinations are not able at all to engage in one or the other type of employment. My approach does not control for this. Nevertheless, the median imputation can control - at least partially - for the much likelier scenario that transitioning between full- and part-time employment may be possible but associated with large wage losses. Precisely, it could be argued that individuals who only work in one kind of employment are those for whom the transition is more likely to imply a wage loss. If this is the case, these individuals surely have a high (above-median) unobservable in their current employment - and the median (counterfactual) wage that will be imputed for them will reflect the potential wage loss associated to their transition.

3.6 Corrected part-time wage gap

Once non-realised full- and part-time wages have been imputed, one has all the ingredients required to disentangle differences in the wage distributions from differences in the composition of the groups taking up full- and part-time employment to explain the increasing raw part-time wage gap.

Table 3.3 displays results on the corrected part-time wage gap, $\widehat{G}_{corr}$. It condenses the information contained in Figure 3.3 by reporting the log point differences between selection-corrected full- and part-time wage distributions (the black dotted lines in Figure 2.5). Again, analogously to the analysis of the raw part-time wage gap, results are summarized for four time periods covering the overall time span 1990 to 2009 and include - additionally to those quantiles contained in Figure 3.3 - figures on the 10^{th} and the 90^{th} percentiles. Given that differentials are computed as the difference between full-time log wages minus part-time log wages, positive figures in Table 3.3 can be interpreted as part-time wage penalties whereas negative figures hint at a part-time wage premium[45].

Table 3.3: Corrected part-time wage gap

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
$\widehat{G}_{corr}$ ($\tau = .10$)	.10*	.16*	-.05*	-.06*
$\widehat{G}_{corr}$ ($\tau = .25$)	.05*	.13*	-.04*	-.05*
$\widehat{G}_{corr}$ ($\tau = .50$)	-.02	.06*	-.07*	-.08*
$\widehat{G}_{corr}$ ($\tau = .75$)	-.08*	-.02*	-.09*	-.08*
$\widehat{G}_{corr}$ ($\tau = .90$)	-.19*	-.17*	-.14*	-.13*

Comments: Log-point differences between the relevant inverse cumulative distributions. Unweighted average of multiple imputations. *Statistical significance at the 5% level (computed by bootstrap with 200 replications).

Source: SOEP.v32,own calculations.

The corrected part-time wage gap displays great variation across the distribution and over time. At the beginning of the 1990s, I find a sizeable wage penalty for part-time workers at the lower end of the distribution (around 15 log points at the 10^{th} percentile), zero effect at the median, and a substantial premium at the top of the distribution (around 18 log points). During the 2000s, the returns penalty at the lower half of the distribution disappears and I find a part-time returns premium for the whole of the wage distribution, reaching 8 log points at the median and 13 log points at the upper end of the distribution[46].

[45] Figures in Table 3.3 can only be interpreted as differences between distributions, i.e. they do not capture the gap to which particular individuals at one point of the distribution are exposed. One reason thereto is that the methodological framework used allows individuals to have different ranks in the full- and part-time wage distributions.

[46] The results at the median are compatible with Paul (2016), who finds - on average for the period 1984 to 2011 - a negative wage effect for short-hours part-time and a slight premium for long-hours part-time, and Fouarge and Muffels (2009), who find no effect of part-time work on current wages for the period 1994 to 2006. On the contrary, Wolf (2002) finds a part-time penalty even after controlling for human capital characteristics. My findings for the median are also in line with the more broader, international context, in which most studies find very small or no penalties for female part-time work once they control for individuals' labor characteristics (see Aaronson and French 2004 and Hirsch 2005). Booth and Wood (2008) find a part-time wage premium for Australia and so do Mocan and Tekin (2003) for the non-profit sector in the United States. Furthermore, the increasing part-time premium found during the 2000s could be due to increases in productivity of part-time workers. This would be in line with the results of Künn-Nelen et al. (2013), who find a productivity plus of firms in the Dutch pharmacy sector with large shares of part-time employees, and partly of Garnero et al. (2014), who

The results on the corrected part-time wage gap can be summarized in three main messages. First, during the 1990s the corrected part-time wage gap mimics the pattern of the raw part-time wage gap in that it presents part-time wage penalties at the lower half of the distribution and premiums at the upper half. This finding suggests that the higher dispersion of part-time wages is not driven by the composition of the individuals actually working part-time but rather as a result of the larger heterogeneity of part-time employment in terms of economic activities, working hours range and resulting wages.

Second, I observe a drastic change in the pattern of the corrected part-time wage gap from the late 1990s until the 2000s, signalled by the emergence of a corrected wage premium across the whole distribution. This is due to the lower/no growth of the lower half of the selection-corrected full-time wage distribution combined with some growth for the 10^{th} and 25^{th} percentiles of the part-time distribution. While the economic causes behind this development are not revealed by my approach, one possible explanation consists of increased returns to high skills in both distributions together with stagnated or decreased returns to lower skills (more frequent in the part-time distribution). This trend can be interpreted as a decrease in occupational segregation by full- and part-time status. Part-time employment has become more popular and a broader spectrum of jobs may have been made available under a part-time scheme. This makes it increasingly possible to transition between full- and part-time status without changing one's occupation.

Third, the differences between the corrected and the raw part-time wage 1gaps increase dramatically over time at all points of the distribution, implying that different wage structures of both kinds of employment are no longer explaining the raw penalty. This is most evident by the end of the 2000s, when the raw part-time penalty ranges from 8 to 33 log-points (see Table 3.2) and the corrected part-time premium amounts from 5 to 13 log points, depending on the percentile. As differences between the two measures can be explained by different selection patterns into full- and part-time employment[47],

find productivity premiums for male workers but not for female workers.

[47] Recall expression 3.3, which implies $G_{raw}(\tau,t) - \widehat{G}_{corr}(\tau,t) = \widehat{Sel}^{FT}(\tau,t) - \widehat{Sel}^{PT}(\tau,t)$.

I now turn to discuss the estimates on $\widehat{Sel}^{FT}(\tau,t)$ and $\widehat{Sel}^{PT}(\tau,t)$ over time and across the distribution.

3.7 Selection into full- and part-time employment

The effect of sample selection on each of the two wage distributions can be summarized as the log-point difference between the observed and the selection-corrected full- and part-time log wage distributions at selected percentiles (see Table 3.4). Given that the difference is specified as the observed distribution minus the selection-corrected one, positive (negative) values hint at positive (negative) sample selection into each type of work arrangement with respect to the entire population[48].

Table 3.4: Effect of selection on the wage distributions

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) Full-Time Distribution: $\widehat{Sel}^{FT}(\tau,t) = F^{-1}_{W^{FT},t}(\tau) - \widehat{F}^{-1}_{W^{FT},t}(\tau)$				
$\tau = .10$	.14*	.19*	.28*	.30*
$\tau = .25$	.17*	.17*	.26*	.28*
$\tau = .50$	.16*	.16*	.22*	.24*
$\tau = .75$	.13*	.15*	.20*	.20*
$\tau = .90$	.12*	.12*	.16*	.18*
(B) Part-Time Distribution: $\widehat{Sel}^{PT}(\tau,t) = F^{-1}_{W^{PT},t}(\tau) - \widehat{F}^{-1}_{W^{PT},t}(\tau)$				
$\tau = .10$	.02*	.09*	-.11*	-.10*
$\tau = .25$	.03*	.10*	-.01	-.04*
$\tau = .50$	.05*	.13*	-.00	-.02
$\tau = .75$	.07*	.11*	.03*	-.01
$\tau = .90$	.11*	.15*	.02	-.04

Comments: Differences in log-points between the observed and the selection-corrected distributions. Unweighted average of multiple imputations. * Statistical significance at the 5% level (computed by bootstrap with 200 replications).
Source: SOEP.v32, own calculations.

The first finding is that the female full-time wage distribution is strongly positively

[48]Figures on the magnitude of selection into full- and part-time employment do not add to zero because each of them is computed with respect to the entire population instead of only the non-working population.

selected at all points in the distribution (see the upper panel of Table 3.4). The selection effect at the median hourly wage takes on values around 16 to 24 log points depending on the year. For the complete time span under study, positive selection in full-time employment is found to be strongest in the lower end and weakest in the upper end of the distribution. Moreover, I find that positive selection starts increasing at all points of the distribution over time from 2000 onwards. This is particularly so for the lower half of the distribution where the selection gap at the 10^{th} percentile rises from roughly 14 log points by 1990 to 30 log-points by 2009.

The pattern is very different for selection into part-time employment. During the 1990s, I find the part-time wage distribution to be positively selected throughout the distribution but less so than the full-time distribution. In the 2000s, sample selection almost disappears for most of the distribution and it turns negative for the lower end of the distribution (at around 10 log points).

These patterns coincide with a substantial rise in part-time employment, particularly in marginal employment, as well as a steady full-time employment rate (see Figure 2.1). In order to better understand the mechanisms behind these patterns, Figures 3.5 and 3.6 explore the role of selected observable characteristics (such as education, working experience and age) as well as unobservable characteristics.

To this end, Figure 3.5 shows full- and part-time employment rates over time by human capital variables included in the wage model. The level of the full-time employment rate rises with educational achievement and previous working experience[49], pointing at positive selection into full-time employment on these two observable factors[50]. Over time, Figure 3.5 reveals substantial changes of the full-time employment rate by educational achievement and age, but not working experience. Hence, the share of women working full-time conditional on having an advanced degree has increased around 10 percentage points while decreasing 5 percentage points for those women with an intermediate degree over the twenty years under study. This points at education as a main

[49]Total working experience in Figure 3.5 is constructed as the sum of previous years of working experience in full-time, part-time and marginal employment without any further weighting (i.e. one year of full-time employment counts the same as one year of part-time employment).

[50]Figure 3.5 also shows that the level of the full-time employment rate decreases with age. However, due to the u-shaped returns to age, no conclusion in terms of selection can be derived from this.

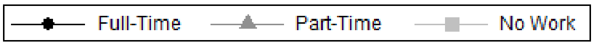
observable factor behind the increase in positive selection into full-time employment over time. In terms of age, the full-time employment rate is highest for the youngest category of women. However, from the beginning of the 1990s onwards it has decreased roughly 8 percentage points, amounting to 40% by the end of the 2000s. For all other age categories, the full-time employment rate displays a flat profile over time at 30%. Contrary to the full-time employment rate, over time the part-time employment rate has risen relatively homogeneously across most categories of human capital variables. This is consistent with the finding that part-time employment is less selected than full-time employment and also decreasingly so over time. Interestingly, the only category in which part-time employment displays a flat profile is among women with advanced educational degrees, which is the only category for which full-time employment has a steep increasing profile. The second insight of Figure 3.5 is the rise of part-time employment among the youngest group of women. This could be a consequence of the observed longer years of education for women, as this is the only age group for which non-employment has increased during the late 1990s and early 2000s.

The methodology used in this chapter enables examining the role of unobservable characteristics in the selection patterns. The conditional rank captures individuals' unobservable component. Therefore, Figure 3.6 shows how the probability of working full- and part-time (conditional on working) has evolved over time, by rank in the overall female conditional wage distribution[51]. The dashed lines in Figure 3.6 represent full- and part-time employment rates (conditional on employment) and serve as benchmark to evaluate the role of unobservables regarding the selection into full- and part-time employment.

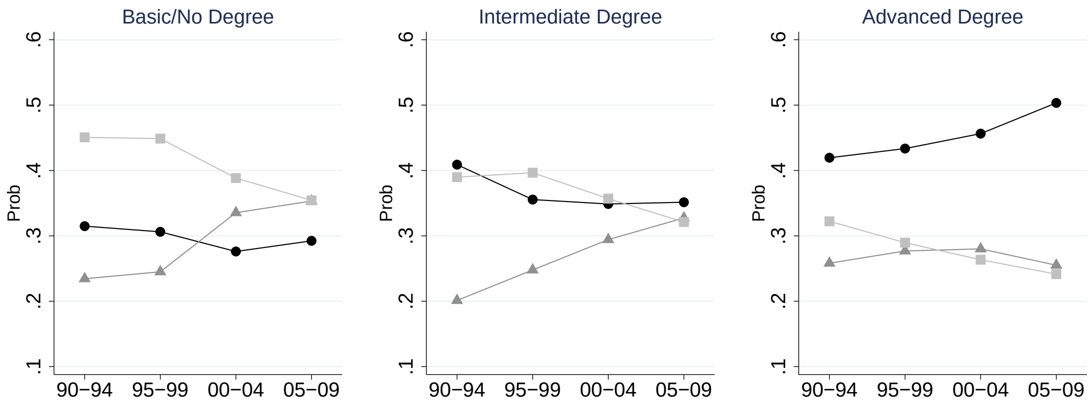
In the 1990s, working women were more likely to have a full-time job than a part-time job regardless of their unobservables. This is shown by circle-marked lines lying above triangle-marked lines in Figure 3.6. However, women with both low- (below the 25^{th} percentile) and high conditional ranks (above the 75^{th} percentile) were over-proportionately likely to work part-time, while women in the middle half of the distribu-

[51]See Chernozhukov and Hansen (2005) for a discussion of the interpretation of the conditional rank of an individual as a summarizing variable of her wage-relevant unobserved characteristics.

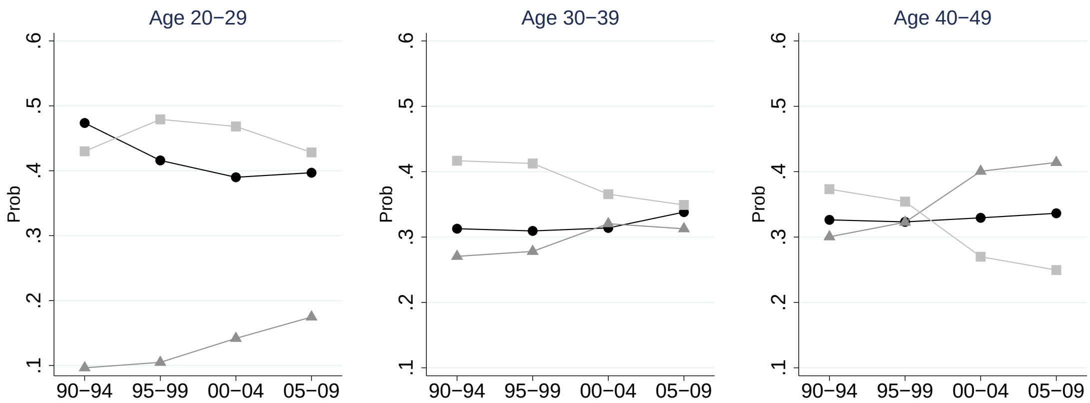
Figure 3.5: Share of women working full-time, part-time and out of work, by observable characteristics and over time



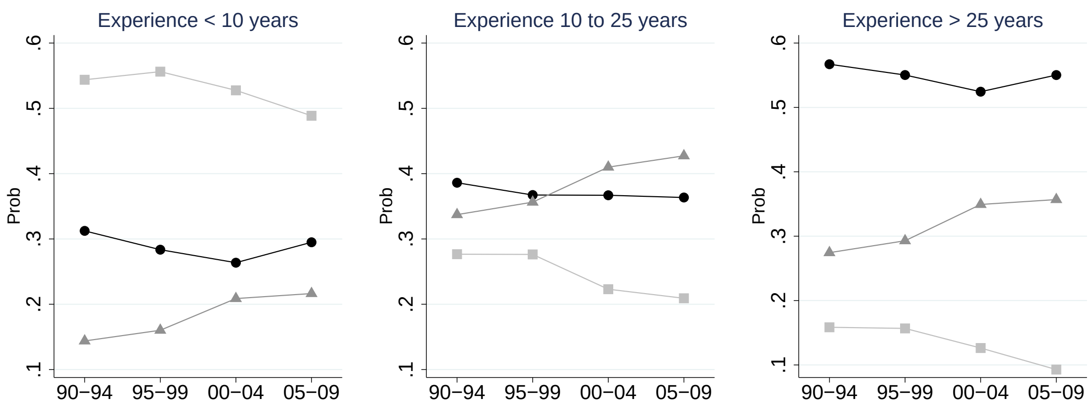
(a) Education



(b) Age



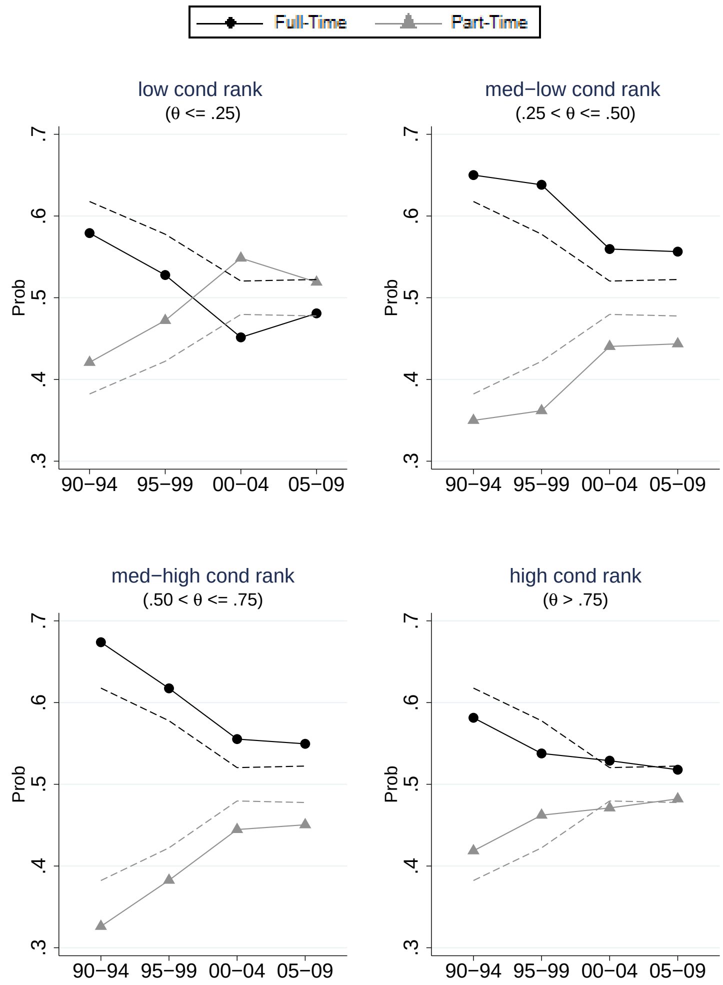
(c) Total Working Experience



Source: SOEP.v32, own calculations.

tion were over-proportionately likely to work full-time. Remarkably, in the first half of the 2000s, women with low ranks in the conditional distribution become 10 percentage points more likely to have a part-time job than a full-time job (although, conditional on employment, full-time work was more common). Furthermore, selection bias for unobservables at the middle of the distribution has remained constant over time, and selection on unobservables for women with high conditional ranks has disappeared over time (as can be seen by the marked lines coinciding with the dashed, benchmark lines).

Figure 3.6: Probability of working full- or part-time, by conditional rank



Source: SOEP.v32, own calculations.

To sum up, Figure 3.6 indicates that there is selection in factors other than human capital and that those have been evolving over time. Concretely, my findings suggest that individuals with low unobservables are over-represented in the part-time distribution, whereas individuals with middle unobservables are over-represented in the full-time distribution. Moreover, individuals with high unobservables were over-represented during the 1990s in the part-time distribution. In the 2000s, this is no longer the case and individuals with high unobservables are proportionally represented in each kind of employment.

Next, I discuss my findings in a broader economic context. From a supply-side perspective, the pattern of selection into full-time employment presented in Table 3.4 fits well the theoretical framework proposed by Mulligan and Rubinstein (2008), in which the authors relate the magnitude of the selection bias to inequality of realised wages and the employment rate[52]. According to their framework, rising wage inequality - resulting from higher market returns to skill - favours the decision of highly skilled individuals to participate in the labour market because their wage offers increasingly exceed their reservation wages. This mechanism seems to be at work for female full-time employment in Germany. In a time period in which full-time wage inequality has increased, Figure 3.5 shows how women with an advanced degree increasingly select into full-time employment while women with an intermediated degree (and to a lesser extent, those with a basic degree) participate less in full-time employment. Consistent with Mulligan and Rubinstein (2008), this development occurs in a context of a stable full-time employment rate.

The conditions surrounding the decrease in part-time selectivity are very different from those applicable to full-time employment. Wage dispersion in the part-time distribution has remained constant, though this has been the result of wage compression at the upper-half of the distribution combined with large dispersion in the lower-half (see Table C.13 in the Appendix). At the same time, the part-time employment rate has risen substantially during this time period. The rising part-time employment rate is likely contributing in a mechanical manner to the decrease in selection: the closer the share of women participating in the labour market is to the entire female population,

[52]The authors make use of the seminal work by Roy (1951), Heckman (1979) and Gronau (1974) to formalize their point.

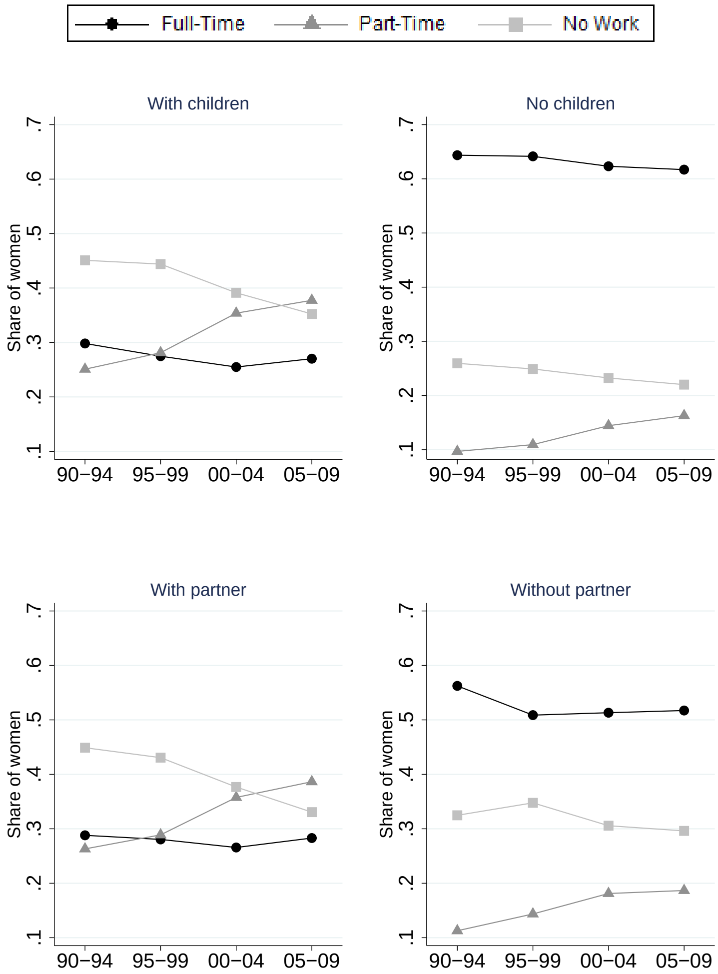
the smaller the selection bias.

However, the increasing representativeness of women working part-time surely reflects much more than only a mechanical effect of higher female labour market participation. In fact, these opposite selection patterns into full- and part-time employment are likely to reflect both women's choices and constraints. While the econometric approach used in this chapter cannot disentangle the two, I now discuss them in the context of the existing literature.

On the side of choices, according to the OECD (2017a), part-time employment among prime-age women in Germany is mostly voluntary (87% and 81% of female part-time dependent employment was voluntary in the time periods 2000 to 2004 and 2005 to 2009, respectively). Furthermore, some studies point at high rates of job satisfaction of female part-time employees (Booth and van Ours 2008, 2013). Furthermore, while the shape of the full-time employment rate is flat with respect to family status, the rise in part-time employment has taken place predominantly for women with children and women in couple households (see Figure 3.7). Among mothers, part-time employment over the twenty years under study has gone up by 13 percentage points. In comparison, for women without children, it has risen by only 6 percentage points. I can only speculate about the mechanisms behind these developments. Certainly, the evolution on social norms regarding female labour supply of mothers and wives may have contributed to this trend. In addition, these two decades - the 1990s and 2000s - have been marked by periods of low or no economic growth and - by German standards - also high unemployment rates (see Dustmann et al. 2014). This has led to increasing job insecurity (Bergmann and Mertens 2011) as well as decreasing real wages for individuals in the lower part of the distribution (e.g. Dustmann et al. 2009). Taking into account the high degree of assortative mating in Germany (see Ermisch et al. 2006), the gradual incorporation of women with relatively low wages into part-time employment could also partially be a reaction to counterbalance these developments at the household level.

However, there is also strong evidence on the existence of constraints behind these selection patterns. Data on desired versus actual working hours reveal substantial dif-

Figure 3.7: Share of women working full-time, part-time and out of work, by family status



Source: SOEP.v32, own calculations.

ferences between the two (see Müller et al. 2017 for Germany). Moreover, short-time work schemes have been widely used as employment adjustment mechanisms in crisis years, resulting in higher rates of involuntary unemployment (e.g. Hijzen and Martin 2013 and Borowczyk-Martins and Lalé 2016). These have had a large incidence in Germany (Hijzen and Venn 2011). Furthermore, there can be demand-side hours restrictions on certain occupations (Euwals and van Soest 1999). For instance, if the range of offered working hours correlates with the required skill level for the job, this can result in highly-paid occupations being offered only in full-time and low-paid occupations being offered only in part-time. While there is evidence for this phenomenon in some countries (e.g. Manning and Petrongolo 2008 for the UK), my results on the convergence of full- and part-time wage structures suggest that this is not the case for Germany. In addition, lack of affordable (full-time) childcare can pose a constraint too for women to take up full-time employment (e.g. Del Boca and Vuri 2005 and Wrohlich 2011). Lastly, tax- and social security disincentives for second earners may hinder female full-time employment.

Finally, there have been numerous reforms of Germany's labour market over this time period. For instance, legislation on marginal employment has been modified several times over these twenty years, which may have contributed to the rise in part-time work (e.g. Haywood and Neumann 2017 estimate that one million of these jobs were created as a result of the reform in 1999)[53]. In addition, the second half of the 2000s has also seen the implementation of two further major reforms, the social assistance reform and the parental leave reform, which may have further altered the incentives for joining full- and part-time employment.

3.8 Conclusion

This paper studies the evolution of the raw part-time wage gap in West Germany for the years 1990 to 2009. During this time period, the raw part-time penalty for the median wage doubled, going from 9 log-points at the beginning of the 1990s up to 18 log-points by the end of the 2000s. At the lower end of the wage distribution, the raw penalty has even reached 33 log-points.

My findings reveal that the observed difference is mostly the result of opposite selection patterns into full- and part-time employment - in particular strong and increasing selection into full-time work combined with decreasingly (positively) selected part-time employment. Human capital, especially educational achievement, appears to drive the rising selection into full-time employment over time. However, factors other than human capital are behind the pattern of selection into part-time employment. Furthermore,

[53]However, a large share of these new jobs were "second" jobs. Given that my analysis is carried out only for individuals' main job, these reforms are not expected to be driving my results.

over time the wage structure of part-time employment has become more favourable relative to that of full-time employment. As a consequence, differences in the wage structures of both kinds of employment loose their explanatory power for the widening raw part-time wage gap.

Importantly, the econometric approach used in this chapter cannot identify whether these selection patterns reflect women's choices or constraints. Nevertheless, the literature indicates that both factors are likely to play an important role. Consequently, my results point to two major policy recommendations. If the goal is to reduce the part-time wage gap, policies need to, first, reduce constraints that hinder female full-time employment and, secondly, further facilitate the transition between full- and part-time employment in both directions and for all individuals, regardless of skill level and gender.

Constraints hindering female full-time employment can be of different nature. Providing affordable high-quality full-time childcare should help to reduce labour supply constraints. Such infrastructure should cover early childcare facilities (such as kindergartens) as well as primary schools. In the last years, progress has been made on this front - so we may well see an effect in the years to come. Mitigating tax- and social security disincentives for second earners should serve the same purpose. In particular, smoothing the tax- and social security contributions schedule around the earnings threshold defining marginal employment as well as reforming the system of joint taxation towards a more individual system should lead to increases in the working hours of second earners, who are often women.

The second policy recommendation - facilitating the transition between full- and part-time employment - could be structured around a legal entitlement to work part-time as well as a legal entitlement to return to full-time work. Such measures already exist in the context of parental leave but ways to extend them to a larger circle of individuals could be explored. However, consequences for firms in terms of organisational challenges or even additional costs should be considered too.

General Conclusion

This dissertation examines the gender- and the part-time wage gap, across the distribution and over time. My findings can be summarized in the following points.

Descriptive evidence on the gender wage gap indicates that it has been decreasing over the last decades, although since the 2000s at a much lesser pace. However, the results of Chapter 2 indicate that the decreasing trend of the gender wage gap is driven by the fact that the sample of women working full-time is strongly positively selected at all points of the wage distribution. More importantly, the magnitude of the female selection into full-time employment has been increasing substantially over time. The sample of men working full-time is also slightly positively selected, but less so than it is the case for women. Nonetheless, the magnitude of selection at the lower bound of the male distribution has increased steeply from the beginning of the 1990s onwards. Second, as a result of these trends of selection into full-time employment, the selection-corrected gender wage gap is much higher than the observed gap. Moreover, while by the end of the 2000s the observed gender wage gap was smaller than 25 years before, the selection-corrected gender wage gap was not. Furthermore, the decomposition analysis shows that the unexplained gender wage gap is decreasing over time. This is particularly true for the bottom of the wage distribution. In other parts of the wage distribution, unexplained gaps are still present and increase with wages. For women earning high wages, differences in coefficients still contribute more to the total gender wage gap than differences in their characteristics.

The analysis of the part-time wage gap in Chapter 3 shows that the increasing female part-time wage gap over time is driven by opposite patterns of selection into full- and part-time employment. I find selection into full-time employment to increase over time, whereas the sample working part-time becomes less selected over time. By the end of the 2000s, selection into part-time employment is non-existent at median wages and becomes negative at the lower end of the distribution. With regard to the wage structures of full- and part-time employment, my findings suggest that they have become more equal over time and, by the end of the 2000s, the results point at a small but statistically significant corrected part-time wage premium.

Importantly, the econometric approach of this dissertation cannot identify whether these selection patterns reflect women's choices or constraints. Nevertheless, the literature indicates that both factors are likely to play an important role. Therefore, if the goal is to reduce the gender- and part-time wage gap, my dissertation's results point to three major policy recommendations.

First, policies directed towards reducing barriers that hinder female full-time employment should have a positive effect on the gender- and part-time wage gap. Such existing constraints, however, can be of different nature. Providing affordable high-quality full-time childcare should help to reduce labour supply constraints. Such infrastructure should cover early childcare facilities, such as kindergartens, as well as primary schools. In the last years, progress has been made on this front - so one may well see a positive effect in the years to come. Mitigating tax- and social security disincentives for second earners should serve the same purpose. In particular, smoothing the tax- and social security contributions schedule around the earnings threshold defining marginal employment as well as reforming the system of joint taxation towards a more individual system should lead to increases in the working hours of second earners, who are often women.

Second, policies aimed at reducing inequalities in employment careers between men and women should also contribute to reducing the gender wage gap. Recent reforms of the parental leave benefits that have shortened employment interruptions of mothers and created incentives to take parental leave for fathers are one step in this direction.

Third, policies that facilitate the transition from part-time to full-time employment in both directions and for all individuals - regardless of skill level and gender - are expected to lower both the gender- and the part-time wage gap. Such policies could be structured around a legal entitlement to work part-time as well as a legal entitlement to return to full-time employment. Such measures are already in place in the context of the parental leave and policy makers should explore ways to extend them to a larger circle of individuals. However, it should be kept in mind that such legal entitlements may pose organisational challenges as well as additional costs for firms, which should be also considered.

Appendix to Chapter 1

A.1 Tables

Table A.1: Descriptives of the wage estimation sample

Sample Means	Men	Women	Total
Age	44.214	43.700	43.958
Experience	21.196	17.306	19.261
Tenure	13.742	10.772	12.265
Public Sector	0.241	0.325	0.282
Education:			
Primary School	0.100	0.107	0.104
Sec./Midd Vocational	0.470	0.503	0.486
Upper Sec./High Voc.	0.163	0.170	0.167
University Degree	0.267	0.219	0.243
Occupation:			
Untrained Worker	0.132	0.143	0.138
Trained Worker	0.187	0.028	0.108
Foreman	0.062	0.006	0.034
Untrained Employee	0.020	0.082	0.051
Trained Employee	0.039	0.147	0.093
Qualified Professional	0.186	0.396	0.290
Highly Qualified Prof.	0.264	0.115	0.190
Lower Civil Servant	0.031	0.015	0.023
Upper Civil Servant	0.080	0.069	0.074
Industrial Branch:			
Electronics	0.192	0.075	0.134
Mining. Energy	0.018	0.007	0.013
Chemical Industry	0.069	0.042	0.055
Construction Sector	0.082	0.016	0.049
Heavy Industry	0.073	0.016	0.045
Textile Sector	0.004	0.006	0.005
Trade and Retail	0.088	0.156	0.122
Transports. Post	0.068	0.033	0.050
Public Services	0.238	0.422	0.330
Private Services	0.102	0.153	0.127
Others	0.057	0.069	0.063
Agriculture	0.010	0.005	0.007
Size of Firm:			
Up to 5 Employees	0.040	0.096	0.067
5-200 Employees	0.404	0.478	0.441
200-2000 Employees	0.247	0.203	0.225
2000+ Employees	0.309	0.223	0.266
Number of Observations	15,713	15,556	31,269

Notes: Pooled waves for years 2008 through 2012.
Source: SOEP.v29.1, own calculations.

Table A.2: Sample means by quartiles of the hourly wage distribution

Sample means	Men				Women			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Hourly Wage	9.95	15.28	19.96	32.00	7.24	11.56	15.47	25.50
Age	38.46	42.92	45.47	48.27	42.66	42.09	43.49	46.13
Experience	16.19	21.14	23.09	23.91	14.90	16.47	17.88	20.01
Tenure	6.95	13.16	16.08	17.26	6.59	8.87	11.87	15.27
Public Sector	0.14	0.23	0.32	0.26	0.13	0.23	0.43	0.48
Education:								
Primary School	0.17	0.12	0.11	0.03	0.19	0.13	0.08	0.04
Sec./Midd Vocational	0.65	0.62	0.48	0.22	0.61	0.61	0.51	0.31
Upper Sec./High Voc.	0.11	0.17	0.20	0.17	0.11	0.17	0.22	0.18
University Degree	0.07	0.10	0.22	0.57	0.09	0.09	0.19	0.47
Occupation:								
Untrained Worker	0.33	0.15	0.08	0.01	0.34	0.16	0.06	0.02
Trained Worker	0.26	0.29	0.18	0.05	0.04	0.03	0.02	0.01
Foreman	0.04	0.09	0.09	0.04	0.01	0.01	0.01	0.00
Untrained Employee	0.06	0.02	0.01	0.00	0.19	0.09	0.04	0.01
Trained Employee	0.09	0.05	0.03	0.01	0.20	0.24	0.12	0.04
Qualified Professional	0.13	0.21	0.26	0.14	0.19	0.40	0.57	0.42
Highly Qualified Prof.	0.05	0.10	0.20	0.59	0.02	0.05	0.11	0.27
Lower Civil Servant	0.03	0.05	0.04	0.01	0.00	0.01	0.02	0.02
Upper Civil Servant	0.02	0.03	0.11	0.15	0.01	0.01	0.05	0.19
Industrial Branch:								
Electronics	0.14	0.15	0.20	0.25	0.07	0.07	0.07	0.09
Mining. Energy	0.01	0.01	0.02	0.03	0.00	0.00	0.01	0.01
Chemical Industry	0.06	0.07	0.07	0.07	0.04	0.04	0.04	0.05
Construction Sector	0.11	0.13	0.07	0.03	0.02	0.02	0.02	0.01
Heavy Industry	0.07	0.08	0.08	0.06	0.01	0.02	0.02	0.02
Textile Sector	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
Trade and Retail	0.16	0.11	0.06	0.04	0.23	0.24	0.12	0.05
Transports. Post	0.10	0.08	0.05	0.04	0.03	0.04	0.03	0.03
Public Services	0.14	0.21	0.29	0.29	0.26	0.38	0.51	0.52
Private Services	0.11	0.07	0.09	0.14	0.21	0.12	0.12	0.16
Others	0.07	0.06	0.06	0.05	0.11	0.06	0.05	0.05
Agriculture	0.02	0.01	0.01	0.00	0.01	0.01	0.00	0.00
Size of Firm:								
Up to 5 Employees	0.10	0.04	0.03	0.01	0.21	0.11	0.05	0.02
5-200 Employees	0.60	0.46	0.34	0.26	0.56	0.53	0.45	0.38
200-2000 Employees	0.16	0.26	0.30	0.26	0.11	0.19	0.27	0.24
2000+ Employees	0.14	0.24	0.33	0.47	0.11	0.18	0.22	0.36

Notes: Quartiles of the male and female distributions were calculated separately. Pooled waves for years 2008 through 2012.

Source: SOEP.v29.1, own calculations.

Table A.3: Wage regression coefficients, males

	OLS	q10	q25	q50	q75	q90
Constant	1.651***	0.747***	1.509***	1.899***	2.023***	2.449***
Age	0.0153***	0.0316***	0.0111**	0.00695*	0.0159***	0.00428
Age sq	-0.0143***	-0.0378***	-0.0108**	-0.00436	-0.0132***	0.00392
Experience	0.0185***	0.0221***	0.0220***	0.0197***	0.0143***	0.0125***
Experience sq	-0.0353***	-0.0361***	-0.0427***	-0.0395***	-0.0285***	-0.0307***
Tenure	0.0162***	0.0222***	0.0168***	0.0132***	0.0114***	0.0117***
Tenure sq	-0.0248***	-0.0354***	-0.0253***	-0.0188***	-0.0160***	-0.0188***
Public Sector	-0.0376***	0.0111	-0.00854	-0.0217**	-0.0722***	-0.114***
Education:						
Primary School	omitted	omitted	omitted	omitted	omitted	omited
Sec./Midd Vocational	0.0239**	0.00426	-0.00976	0.0234**	0.0266***	0.0274
Upper Sec./High Voc.	0.0839***	0.0948***	0.0647***	0.0880***	0.0892***	0.0844***
University Degree	0.229***	0.244***	0.203***	0.226***	0.223***	0.211***
Occupation:						
Untrained Worker	omitted	omitted	omitted	omitted	omitted	omited
Trained Worker	0.199***	0.269***	0.223***	0.169***	0.135***	0.113***
Foreman	0.288***	0.372***	0.324***	0.268***	0.211***	0.198***
Untrained Employee	-0.0149	-0.0308	-0.0787***	-0.0671***	-0.0467**	0.0158
Trained Employee	0.145***	0.139***	0.161***	0.129***	0.0901***	0.0822***
Qualified Professional	0.341***	0.405***	0.369***	0.314***	0.277***	0.262***
Highly Qualified Prof.	0.603***	0.624***	0.599***	0.554***	0.547***	0.567***
Lower Civil Servant	0.204***	0.302***	0.209***	0.161***	0.124***	0.122***
Upper Civil Servant	0.422***	0.413***	0.404***	0.394***	0.395***	0.368***
Industrial Branch:						
Electronics	omitted	omitted	omitted	omitted	omitted	omited
Mining, Energy	0.00861	0.0104	-0.00174	-0.0206	-0.0183	0.0240
Chemical Industry	0.0216**	0.0189	0.00686	0.0220*	0.0279***	0.0123
Construction Sector	-0.0528***	-0.0237	-0.0555***	-0.0706***	-0.0818***	-0.117***
Heavy Industry	0.0268**	0.0143	0.0134	0.0215*	0.0297***	0.0190
Textile Sector	0.000559	-0.126*	-0.163***	-0.175***	-0.0460	0.447
Trade and Retail	-0.143***	-0.188***	-0.176***	-0.149***	-0.145***	-0.139***
Transports, Post	-0.105***	-0.179***	-0.158***	-0.128***	-0.0585***	-0.0368*
Public Services	-0.0796***	-0.0951***	-0.114***	-0.109***	-0.0677***	-0.0184
Private Services	-0.0467***	-0.0933***	-0.101***	-0.0563***	-0.00186	0.0310
Others	-0.0777***	-0.123***	-0.113***	-0.0729***	-0.0504***	-0.0401*
Agriculture	-0.270***	-0.374***	-0.318***	-0.237***	-0.202***	-0.215***
Size of Firm:						
Up to 5 Employees	omitted	omitted	omitted	omitted	omitted	omited
5-200 Employees	0.147***	0.305***	0.221***	0.136***	0.0626***	0.0242
200-2000 Employees	0.250***	0.447***	0.340***	0.232***	0.146***	0.0799***
2000+ Employees	0.303***	0.493***	0.390***	0.284***	0.206***	0.148***
Year Dummies:						
2008	-0.0438***	-0.0623***	-0.0487***	-0.0524***	-0.0544***	-0.0327**
2009	-0.0365***	-0.0669***	-0.0357***	-0.0351***	-0.0386***	-0.0314**
2010	-0.0278***	-0.0464***	-0.0259**	-0.0314***	-0.0360***	-0.0265*
2011	-0.0244***	-0.0337**	-0.0224**	-0.0260***	-0.0279***	-0.0197
2012	omitted	omitted	omitted	omitted	omitted	omitted
Observations				15,713		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: SOEP.v29.1, own calculations.

Table A.4: Wage regression coefficients, females

	OLS	q10	q25	q50	q75	q90
Constant	1.570***	0.992***	1.388***	1.634***	1.861***	1.909***
Age	0.0157***	0.0191***	0.0124***	0.0133***	0.0134***	0.0228***
Age sq	-0.0194***	-0.0302***	-0.0187***	-0.0159***	-0.0151***	-0.0236***
Experience	0.0113***	0.0143***	0.0150***	0.00972***	0.00951***	0.00560*
Experience sq	-0.0145***	-0.0122	-0.0202***	-0.0132***	-0.0130***	-0.00502
Tenure	0.0150***	0.0220***	0.0160***	0.0177***	0.0144***	0.0103***
Tenure sq	-0.0248***	-0.0386***	-0.0273***	-0.0328***	-0.0250***	-0.0156***
Public Sector	0.0548***	0.0916***	0.0804***	0.0490***	0.0437***	0.0254*
Education:						
Primary School	omitted	omitted	omitted	omitted	omitted	omited
Sec./Midd Vocational	0.0164	0.0270	0.0160	0.00589	0.0346***	0.0627***
Upper Sec./High Voc.	0.0914***	0.122***	0.117***	0.0811***	0.0889***	0.100***
University Degree	0.206***	0.175***	0.194***	0.197***	0.249***	0.293***
Occupation:						
Untrained Worker	omitted	omitted	omitted	omitted	omitted	omited
Trained Worker	0.244***	0.241***	0.239***	0.238***	0.254***	0.278***
Foreman	0.248***	0.259***	0.258***	0.233***	0.275***	0.201***
Untrained Employee	0.0455***	0.0151	0.0356*	0.0647***	0.0811***	0.0836***
Trained Employee	0.220***	0.265***	0.244***	0.250***	0.222***	0.183***
Qualified Professional	0.406***	0.456***	0.445***	0.429***	0.388***	0.346***
Highly Qualified Prof.	0.644***	0.650***	0.640***	0.645***	0.646***	0.615***
Lower Civil Servant	0.415***	0.436***	0.405***	0.382***	0.380***	0.443***
Upper Civil Servant	0.593***	0.559***	0.564***	0.607***	0.628***	0.640***
Industrial Branch:						
Electronics	omitted	omitted	omitted	omitted	omitted	omited
Mining, Energy	0.105***	0.153**	0.0863***	0.0293	0.0559	0.0890
Chemical Industry	-0.0616***	-0.211***	-0.0899***	-0.0362**	0.00226	0.00340
Construction Sector	-0.117***	-0.0190	-0.0776***	-0.134***	-0.159***	-0.219***
Heavy Industry	0.0470*	0.0677	0.110***	0.0322	0.0196	0.0297
Textile Sector	-0.180***	-0.115	-0.162***	-0.203***	-0.200***	-0.201**
Trade and Retail	-0.196***	-0.197***	-0.186***	-0.207***	-0.239***	-0.239***
Transports, Post	-0.0825***	-0.0904**	-0.0735***	-0.117***	-0.0907***	-0.0888***
Public Services	-0.111***	-0.0675***	-0.0898***	-0.136***	-0.168***	-0.156***
Private Services	-0.0890***	-0.111***	-0.107***	-0.115***	-0.0919***	-0.0444**
Others	-0.164***	-0.238***	-0.184***	-0.179***	-0.164***	-0.122***
Agriculture	-0.227***	-0.278**	-0.230***	-0.233***	-0.274***	-0.223**
Size of Firm:						
Up to 5 Employees	omitted	omitted	omitted	omitted	omitted	omited
5-200 Employees	0.159***	0.225***	0.207***	0.175***	0.147***	0.0851***
200-2000 Employees	0.259***	0.377***	0.318***	0.270***	0.217***	0.134***
2000+ Employees	0.309***	0.386***	0.354***	0.326***	0.287***	0.228***
Year Dummies:						
2008	-0.0346***	-0.0486***	-0.0463***	-0.0410***	-0.0316***	-0.0158
2009	-0.00958	-0.0200	-0.0290**	-0.0178**	-0.00693	0.0202
2010	-0.00593	-0.00503	-0.0147	-0.0165*	-0.00427	-0.00148
2011	-0.00720	-0.0378**	-0.0218**	-0.00304	-0.00486	0.0238*
2012	omitted	omitted	omitted	omitted	omitted	omited
Observations				15,556		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: SOEP.v29.1, own calculations.

Table A.5: Decomposition results

Quantiles	Overall Gap				Explained Gap				Unexplained Gap			
	Coeff.	pointwise	confidence bands		Coeff.	pointwise	confidence bands		Coeff.	pointwise	confidence bands	
		std error	lower	upper		std error	lower	upper		std error	lower	upper
0.05	.3148	.0232	.2730	.3565	.1810	.0179	.1501	.2119	.1338	.0243	.0855	.1820
0.10	.3132	.0171	.2801	.3462	.1715	.0146	.1471	.1960	.1416	.0183	.1050	.1783
0.15	.3113	.0141	.2820	.3406	.1621	.0130	.1385	.1857	.1492	.0153	.1198	.1786
0.20	.3072	.0125	.2805	.3339	.1555	.0119	.1323	.1787	.1517	.0135	.1257	.1777
0.25	.3022	.0113	.2776	.3268	.1500	.0111	.1289	.1711	.1522	.0120	.1306	.1738
0.30	.2948	.0103	.2729	.3167	.1452	.0102	.1256	.1648	.1496	.0108	.1295	.1697
0.35	.2865	.0096	.2648	.3082	.1405	.0096	.1222	.1589	.1460	.0098	.1269	.1650
0.40	.2780	.0089	.2592	.2969	.1370	.0090	.1199	.1540	.1411	.0090	.1233	.1589
0.45	.2706	.0084	.2532	.2881	.1334	.0086	.1177	.1491	.1372	.0083	.1207	.1537
0.50	.2636	.0080	.2487	.2785	.1305	.0081	.1157	.1454	.1331	.0079	.1174	.1488
0.55	.2578	.0078	.2439	.2717	.1285	.0078	.1138	.1432	.1293	.0076	.1155	.1432
0.60	.2533	.0077	.2395	.2672	.1267	.0077	.1122	.1411	.1267	.0074	.1131	.1403
0.65	.2507	.0079	.2360	.2654	.1264	.0076	.1115	.1414	.1243	.0073	.1105	.1381
0.70	.2486	.0083	.2336	.2635	.1280	.0075	.1137	.1423	.1206	.0074	.1061	.1351
0.75	.2463	.0088	.2299	.2628	.1300	.0075	.1165	.1434	.1164	.0076	.1004	.1323
0.80	.2437	.0094	.2275	.2599	.1326	.0076	.1186	.1465	.1111	.0081	.0949	.1273
0.85	.2396	.0105	.2198	.2595	.1358	.0079	.1216	.1500	.1038	.0089	.0871	.1206
0.90	.2318	.0125	.2066	.2569	.1400	.0087	.1244	.1556	.0952	.0099	.0754	.1151
0.95	.2107	.0165	.1793	.2420	.1440	.0103	.1255	.1626	.0666	.0139	.0398	.0934

Notes: Statistical inference based on bootstrap with 200 replications.

Source: SOEP.v29.1, own calculations.

Table A.6: Conditional logit coefficients

	Couples (both flex)	Singles	Couples (woman flex)
net income	23.50	-5.372 *	-22.47 **
net income^2	-0.78	.5060 ***	1.93 ***
income*leisure$_m$	-1.36 ***		
income*leisure$_f$	-0.93 **	.186	-.672
leisure$_m$	61.99 ***		
leisure$_m^2$	-7.46 ***		
leisure$_f$	25.15 **	22.612 **	24.10 **
leisure$_f^2$	-4.23 ***	-3.012 ***	-2.47 **
leisure$_{m,f}$	2.56		
dlzm dtm	14.45 *		
dlzf dtf	12.52	-.963	0.84
dlzmf dt	-3.17		
eknp dt	-19.58		
eknp2 dt	1.44		
leisure$_m$×age	-.28 **		
leisure$_m$×age^2	.38 ***		
leisure$_f$×age	.24 **	.04793	.0479
leisure$_f$×age^2	-.18	.02376	.0237
leisure$_m$×health	1.85 **		
leisure$_f$×health	2.57 **	-1.243	.94
leisure$_f$×child age 3	8.01 ***	10.44 ***	4.87 ***
leisure$_f$×child age 6	4.02 ***	3.076 ***	3.69 ***
leisure$_f$×child age 16	2.46 ***		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: SOEP.v29.1, STSM, own calculations.

A.2 Imputation of wages for the non-working

In order to estimate employment effects from the unexplained gender wage gap, we need to obtain potential wages for individuals that do not take up paid employment in the time period of our interest. Table A.9 below shows for how many observations by year and gender we need to impute wages. Note that the wording “missing wage observation” in this section exclusively refers to individuals who do not work (instead of item non-response or attrition issues) for whom we need to impute a wage.

We exploit the longitudinal dimension of the SOEP to impute wages for the non-

Table A.7: Female labour supply elasticities by household type

	Change in Hours	Change in Participation
Couples (both spouses flexible)		
without children	0.289	0.063
with children	0.357	0.110
Couples (only woman flexible)		
without children	0.479	0.120
with children	0.457	0.133
Single Households		
without children	0.229	0.049
with children	0.266	0.065

Notes: Elasticities computed for a one percent increase of gross earnings.
Source: SOEP.v29.1, own calculations.

Table A.8: Dynamic net gender pay gap by income quintiles and household type

Quintiles	Married		Single		All
	without children	with children	without children	with children	
% of household disposable income, $\widetilde{\Omega}_j$					
1	8.3	3.4	9.5	5.4	6.9
2	6.6	3.4	11.0	7.3	7.3
3	5.7	3.7	10.8	8.1	6.7
4	5.1	3.6	9.6	8.1	5.9
5	4.2	2.0	7.4	5.8	4.3
Average	5.5	3.3	9.9	6.3	6.2
absolute difference (in Euro)					
1	124	110	111	89	110
2	140	95	146	159	129
3	145	107	176	184	144
4	181	138	197	191	176
5	183	111	208	179	171
Average	162	111	164	126	146

Notes: Gaps presented as % of net household income. Quintiles of overall equivalized net household income. $\widetilde{\Omega}_j$ allows for labor supply adjustment, as described in Section 1.3.5. Weighted observations.
Source: SOEP.v29.1, own calculations.

Table A.9: Missing and non-missing wage observations, by gender and year

	Women			Men		
	Non-Missing	Missing	Missing (in%)	Non-Missing	Missing	Missing (in%)
2008	3,087	1,129	36.57	3,277	303	9.25
2009	3,028	966	31.90	3,070	311	10.13
2010	2,433	808	33.21	2,449	247	10.09
2011	3,506	1,126	32.12	3,429	364	10.62
2012	2,895	882	30.47	2,801	278	9.93
Total	14,949	4,911	32.85	15,026	1,503	10.00

Source: SOEP.v29.1, own calculations.

working. The literature suggests different methods thereto. We cannot use our conditional quantile wage model to predict wages for the non-working because we lack work-place related information for them. For this reason, we preserve the rank they had in the unconditional distribution of the hourly wage when they last worked or when they take up paid employment again. We choose the time period that is closest to the missing wage information, under the rationale that the labor market relevant characteristics of the individual should be the closest to his or her labor market characteristics at the time of the wage missing observation. This procedures allows us to impute wages for a significant chunk of the non-working (see Table A.10 below).

Table A.10: Number of successful imputations by gender and year

	Women			Men		
	Missing	Imputed	Imputed (in%)	Missing	Imputed	Imputed (in%)
2008	1,129	732	64.84	303	246	81.19
2009	966	648	67.08	311	251	80.71
2010	808	563	69.68	247	211	85.43
2011	1,126	562	49.91	364	205	56.32
2012	882	456	51.70	278	169	60.79
Total	4,911	2,961	60.29	1,503	1,082	71.99

Source: SOEP.v29.1, own calculations.

Table A.10 above indicates that we achieve to impute the most observations in year 2010. For this reason, we use data for 2010 to proceed with the analysis of the unexplained wage gap in terms of net income.

A.3 Simulation of labour supply reactions

A general description of the microsimulation model is provided by Steiner et al. (2012). The model consists of a discrete choice labour supply model and a tax and transfer simulation model.

A.3.0.0.1 Labour supply estimation Following Van Soest (1995) and Aaberge et al. (1995) we estimate a labour supply model using a static structural discrete choice model. The discrete choice approach allows to account for non-linearities in the budget constraint in a transparent and flexible way. Furthermore we can derive the model from random utility model in which decision makers are assumed to maximize utility. We assume that a household can choose among J working hours categories (in case of couple households, J comprises combinations of working hour categories for both spouses including zero hours). Thresholds between hours categories are derived from the actual distribution of working hours in the sample. The actual distribution of working hours shows well known spikes. In our model, women can chose between six categories (0,8.5,15.5,25.5,38,45.5) and men between four categories (0,12.5,36.5,47) – for couple households this results in 24 possible combinations of working hours. The values are median values of different domains of the hours distribution. The estimation results are robust to variation of thresholds of hours categories or if we choose the mean instead of the median.

Each combination of working hours, given the hourly wage, corresponds to a certain net household income. The different combinations of working hours and leisure j result in different utility levels U_{ij} of household i. The decision maker chooses alternative κ over j if and only if $U_\kappa > U_j$, $\forall \kappa \neq j$. Household utility U_{ij} depends on a systematic function V_{ij} that relates household characteristics to the level of utility and a random component ϵ_{ij}. For couple households, the utility function includes the logarithm of household net income and its square, the logarithm of leisure of the man, the logarithm of leisure of the woman, the logarithm of the joint leisure of the spouses, the interactions of log leisure of each spouse with German nationality, age, age squared, work incapacity, the interaction of joint log leisure with German nationality, and the interaction of the woman's log leisure with children in the household under 3, 6 and 16 years of age. The errors are assumed to be iid extreme value. Then we can model the probability that alternative κ is chosen by household i using a multinomial Logit model (McFadden, 1974):

$$P_{i\kappa} = Pr(U_{i\kappa} \geq U_{ij}, \forall j = 0, \ldots, J) = \frac{exp(V_{i\kappa})}{\sum_{j=0}^{J} exp(V_{ij})} \tag{A.1}$$

Based on the observed choices we can estimate the parameter of the utility function using Maximum Likelihood. We specify different models for different households: (1) both spouses can adjust their labor supply flexibly, (2) only one spouse has flexible labor supply (3) single households.

A.3.0.0.2 Tax and transfer simulation The households labor supply model described in the previous paragraph requires the simulation of net household income for different combinations of working hours. In brief, the tax and transfer calculation takes into account the whole German tax and welfare system.

Appendix to Chapter 2

B.1 Tables

Table B.1: Male person-year observations

	All	Work FT	Do Not Work FT	Ever Worked FT		Never Worked FT	
	(1)	(2)	(3)	(4)	% of (3)	(5)	% of (1)
1984	3119	2445	674	565	84%	109	3%
1985	3229	2516	713	579	81%	134	4%
1986	3123	2458	665	549	83%	116	4%
1987	3073	2458	615	511	83%	104	3%
1988	2917	2308	609	535	88%	74	3%
1989	2887	2364	523	443	85%	80	3%
1990	2780	2303	477	399	84%	78	3%
1991	2782	2269	513	434	85%	79	3%
1992	2703	2207	496	433	87%	63	2%
1993	2698	2145	553	486	88%	67	2%
1994	2593	2043	550	474	86%	76	3%
1995	2722	2124	598	512	86%	86	3%
1996	2630	2071	559	470	84%	89	3%
1997	2550	2053	497	416	84%	81	3%
1998	2683	2145	538	434	81%	104	4%
1999	2611	2173	438	340	78%	98	4%
2000	4094	3423	671	506	75%	165	4%
2001	3962	3341	621	458	74%	163	4%
2002	4096	3369	727	551	76%	176	4%
2003	3950	3206	744	564	76%	180	5%
2004	3753	3033	720	556	77%	164	4%
2005	3488	2816	672	544	81%	128	4%
2006	3477	2825	652	521	80%	131	4%
2007	3294	2734	560	435	78%	125	4%
2008	3080	2569	511	393	77%	118	4%
2009	2788	2285	503	386	77%	117	4%
2010	4137	3327	810	592	73%	218	5%
2011	4636	3883	753	468	62%	285	6%
2012	4567	3782	785	437	56%	348	8%
2013	5290	4248	1.042	498	48%	544	10%
2014	5193	4186	1.007	407	40%	600	12%
2015	4509	3729	780	304	39%	476	11%

Source: SOEP.v32, own calculations.

Table B.2: Female person-year observations

	All	Work FT	Do Not Work FT	Ever Worked FT		Never Worked FT	
	(1)	(2)	(3)	(4)	% of (3)	(5)	% of (1)
1984	3216	1044	2172	791	36%	1381	43%
1985	3312	1115	2197	809	37%	1388	42%
1986	3192	1102	2090	828	40%	1262	40%
1987	3127	1121	2006	843	42%	1163	37%
1988	2978	1044	1934	882	46%	1052	35%
1989	2900	1113	1787	819	46%	968	33%
1990	2862	1108	1754	855	49%	899	31%
1991	2859	1094	1765	897	51%	868	30%
1992	2840	1078	1762	952	54%	810	29%
1993	2813	1070	1743	963	55%	780	28%
1994	2760	1006	1754	1004	57%	750	27%
1995	2924	1071	1853	1070	58%	783	27%
1996	2862	1043	1819	1078	59%	741	26%
1997	2781	1015	1766	1076	61%	690	25%
1998	2978	1061	1917	1145	60%	772	26%
1999	2863	1086	1777	1069	60%	708	25%
2000	4661	1706	2955	1431	48%	1524	33%
2001	4553	1677	2876	1389	48%	1487	33%
2002	4768	1747	3021	1544	51%	1477	31%
2003	4621	1732	2889	1505	52%	1384	30%
2004	4400	1632	2768	1492	54%	1276	29%
2005	4153	1561	2592	1446	56%	1146	28%
2006	4265	1622	2643	1457	55%	1186	28%
2007	4012	1554	2458	1370	56%	1088	27%
2008	3706	1494	2212	1242	56%	970	26%
2009	3363	1395	1968	1137	58%	831	25%
2010	5607	1643	3964	1625	41%	2339	42%
2011	6549	2164	4385	1549	35%	2836	43%
2012	6365	2227	4138	1418	34%	2720	43%
2013	7154	2507	4647	1431	31%	3216	45%
2014	6626	2457	4169	1211	29%	2958	45%
2015	5832	2217	3615	1032	29%	2583	44%

Source: SOEP.v32, own calculations.

Table B.3: Distribution of weekly working hours in full-time employment, by gender

	P25	P50	P75
(A) Women			
1985-1989	39.0	40.0	42.0
1990-1994	38.0	40.0	42.0
1995-1999	38.0	40.0	42.0
2000-2004	37.5	40.0	42.0
2005-2009	38.0	40.0	44.0
(B) Men			
1985-1989	40.0	41.0	46.0
1990-1994	38.5	40.0	45.0
1995-1999	38.5	40.0	45.0
2000-2004	40.0	41.0	46.0
2005-2009	40.0	42.0	48.0

Source: SOEP.v32, weighting factors used, own calculations.

Table B.4: Male wages: subsample sizes for each grid of quantile regressions

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1.984	302	318	327	299	308	284	280	267	249	217	203	182	171	152	151	140	120	113	108	100	94	80	81	75	67
1.985	0	278	305	303	310	295	296	276	246	223	205	186	163	149	149	142	122	120	113	105	95	78	80	73	64
1.986	209	0	255	245	277	279	268	247	241	214	210	185	183	167	167	153	141	135	125	118	101	87	84	80	66
1.987	178	177	0	187	235	238	241	242	241	222	217	199	194	183	174	161	149	150	132	120	114	98	95	90	70
1.988	183	184	192	0	233	253	266	248	258	223	222	211	210	192	182	170	156	148	142	132	121	108	100	95	77
1.989	110	110	120	109	0	161	196	194	201	198	187	183	179	171	173	163	150	142	132	125	115	106	94	92	77
1.990	105	104	110	102	109	0	131	157	162	162	156	165	172	166	169	158	141	135	119	118	111	97	89	84	69
1.991	96	99	101	109	125	123	0	156	178	176	174	181	188	181	179	168	155	155	138	138	123	111	104	101	81
1.992	99	108	116	118	138	136	131	0	149	177	179	186	195	190	185	178	165	163	146	142	128	123	109	108	89
1.993	116	135	141	156	168	162	168	169	0	173	185	200	220	218	208	209	196	194	177	172	161	146	135	130	111
1.994	117	124	139	140	167	164	164	178	162	0	149	172	199	211	207	207	195	191	177	173	162	145	138	130	114
1.995	112	119	129	142	159	152	155	156	161	136	0	174	215	224	228	215	193	188	181	166	163	151	142	132	110
1.996	99	103	109	112	136	142	144	154	156	144	158	0	182	199	215	209	192	181	175	167	154	147	143	137	111
1.997	81	88	97	102	115	115	126	130	137	126	141	129	0	151	180	187	181	166	151	147	135	126	133	124	103
1.998	72	82	91	88	101	108	110	116	125	127	140	134	121	0	172	181	183	162	153	155	131	127	125	121	96
1.999	48	54	51	52	63	68	77	76	81	71	87	89	84	87	0	104	121	123	119	123	113	109	110	108	85
2.000	36	36	48	46	55	56	67	67	72	64	71	82	84	87	82	0	191	201	213	218	200	190	184	175	155
2.001	30	36	34	41	51	51	54	61	62	61	60	69	69	77	79	145	0	149	164	171	150	159	157	151	138
2.002	43	45	57	55	63	65	70	76	78	81	87	89	90	102	115	197	174	0	163	212	203	210	205	191	179
2.003	37	39	45	44	54	58	61	61	68	67	79	89	76	90	99	196	189	154	0	160	165	184	210	200	195
2.004	31	39	43	46	58	59	64	70	68	66	76	89	84	93	106	199	186	187	149	0	152	173	190	186	198
2.005	26	32	36	43	52	53	52	60	66	65	70	72	74	91	104	175	174	185	167	165	0	156	192	204	212
2.006	25	28	31	37	48	45	50	56	63	61	66	69	66	73	85	142	155	154	135	135	113	0	168	217	214
2.007	18	20	20	22	30	28	30	31	40	41	49	53	48	52	56	104	105	101	97	106	80	85	0	142	154
2.008	13	15	16	18	26	23	22	22	28	31	39	44	40	54	58	90	94	97	97	97	84	96	93	0	117
2.009	12	16	17	18	26	24	27	27	36	34	39	41	41	51	60	99	102	105	99	104	89	99	113	118	0
2.010	10	11	17	17	25	26	30	30	35	31	41	41	37	48	52	95	93	101	100	104	104	106	117	139	108
2.011	8	9	12	12	17	20	19	20	25	21	28	29	27	31	30	61	60	65	64	67	65	68	75	83	75
2.012	2	4	6	8	14	14	16	15	19	15	21	25	23	24	23	45	44	46	42	45	46	46	53	62	70
2.013	1	3	4	6	9	13	14	14	18	16	22	23	20	19	21	44	44	50	50	47	52	52	58	58	58
2.014	1	1	4	4	6	10	11	12	15	9	13	15	16	16	18	34	33	35	34	35	39	38	41	45	43
2.015	1	1	2	4	7	10	13	15	17	14	17	16	13	17	19	31	31	40	43	38	44	43	47	43	42

Source: SOEP.v32, own calculations.

Table B.5: Female wages: subsample sizes for each grid of quantile regressions

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1.984	225	239	276	247	276	269	268	265	267	228	209	206	190	176	172	164	154	140	122	111	93	82	69	65	53
1.985	0	205	249	231	264	262	280	261	255	231	209	203	198	184	175	169	160	138	120	107	98	86	66	64	51
1.986	191	0	211	207	257	262	275	255	265	235	215	205	200	180	184	175	162	141	132	115	113	96	84	86	65
1.987	213	175	0	164	241	247	272	264	259	231	214	206	196	182	175	169	162	141	127	114	112	95	88	88	67
1.988	233	216	192	0	209	235	254	252	248	235	216	211	200	192	187	178	168	158	146	128	123	107	99	100	83
1.989	210	202	202	132	0	158	208	215	216	208	205	199	188	180	176	182	165	155	144	126	117	105	98	95	79
1.990	208	208	216	170	155	0	174	194	217	210	200	198	202	195	191	189	182	165	156	137	134	120	118	110	92
1.991	224	233	248	201	210	178	0	162	201	198	197	194	198	191	190	197	183	169	158	144	142	124	119	111	100
1.992	219	242	255	222	259	230	165	0	161	184	210	203	208	202	209	212	190	178	160	151	154	142	128	127	119
1.993	220	248	256	231	241	230	210	165	0	134	189	193	208	197	217	226	204	194	185	168	166	152	144	131	127
1.994	239	247	269	250	269	263	252	227	183	0	164	189	210	197	205	223	197	180	172	163	156	146	140	141	122
1.995	226	239	260	245	276	267	257	247	216	150	0	157	185	191	215	245	221	212	207	190	186	174	165	155	148
1.996	222	233	256	248	265	278	265	255	222	184	174	0	129	157	197	217	218	205	192	184	185	172	160	151	145
1.997	209	221	244	231	251	270	257	248	238	209	193	140	0	139	191	232	225	211	201	182	180	173	168	155	148
1.998	199	204	232	229	259	276	258	253	261	244	230	194	164	0	192	244	245	234	230	207	209	205	186	176	171
1.999	177	189	207	214	249	266	250	248	254	223	222	195	179	143	0	168	183	182	196	180	178	186	171	172	159
2.000	150	163	181	173	210	238	226	228	230	213	218	188	175	168	128	0	242	290	351	317	309	320	312	315	293
2.001	140	143	169	169	193	225	203	209	218	194	206	197	192	185	165	216	0	200	285	265	278	289	294	291	295
2.002	136	140	164	165	188	216	204	205	214	196	220	209	210	212	201	302	236	0	261	291	313	347	331	340	337
2.003	124	130	147	155	178	204	195	197	210	198	213	200	205	206	206	335	270	204	0	206	272	308	327	338	334
2.004	123	128	147	147	170	191	171	172	186	190	197	192	186	184	200	310	275	239	209	0	224	288	320	334	333
2.005	110	118	133	135	152	178	166	168	179	185	198	194	188	204	210	338	294	276	257	196	0	201	268	304	296
2.006	100	112	121	123	146	168	159	166	162	166	183	178	173	190	194	311	281	278	264	225	170	0	231	288	300
2.007	94	105	114	115	145	159	148	152	166	169	179	177	172	177	195	298	274	258	253	241	216	183	0	232	279
2.008	79	92	100	99	118	135	128	137	138	142	155	150	154	154	176	267	239	225	235	211	205	189	169	0	202
2.009	64	75	80	82	101	120	111	119	129	127	142	131	133	138	154	243	230	212	216	198	187	189	180	154	0
2.010	61	74	78	79	93	104	99	104	109	115	131	123	124	126	142	229	214	205	207	199	188	180	188	190	165
2.011	54	60	64	64	79	92	84	86	96	98	111	112	110	110	125	200	194	192	186	176	166	166	172	176	152
2.012	46	51	56	60	70	85	75	83	84	82	99	95	94	92	113	184	168	165	170	159	157	153	174	170	164
2.013	40	44	48	52	58	69	62	70	74	69	86	81	83	78	102	176	160	155	154	145	142	148	158	162	160
2.014	38	43	45	51	56	65	62	66	73	68	79	73	81	76	93	156	149	146	153	144	140	153	156	168	166
2.015	31	36	39	46	46	58	55	60	59	59	69	70	79	75	91	152	143	133	138	131	125	147	152	155	157

Source: SOEP.v32, own calculations.

B.2 Figures

Figure B.1: Development of the observed gender wage gap over time at different quantiles of the distribution

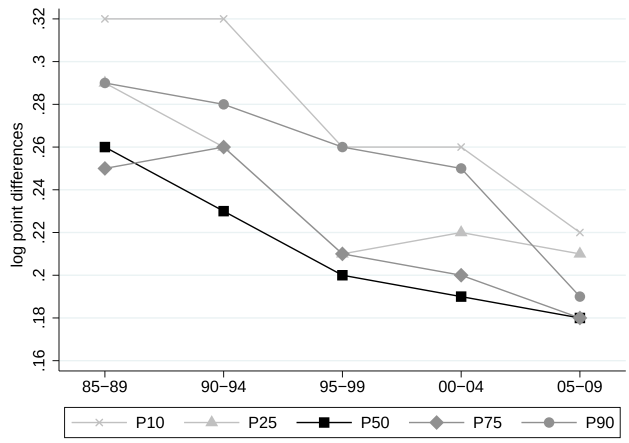


Figure B.2: Development of the selection-corrected gender wage gap over time at different quantiles of the distribution

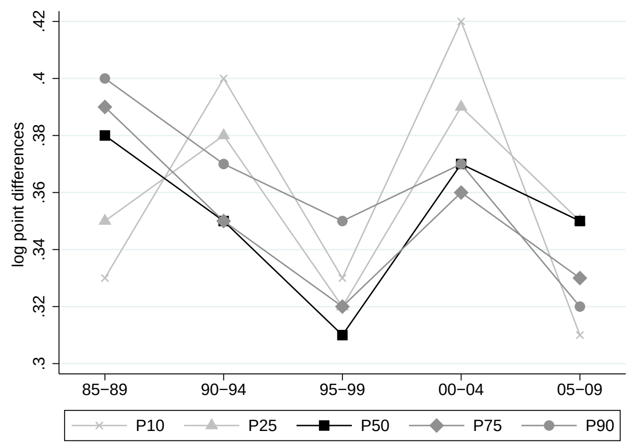
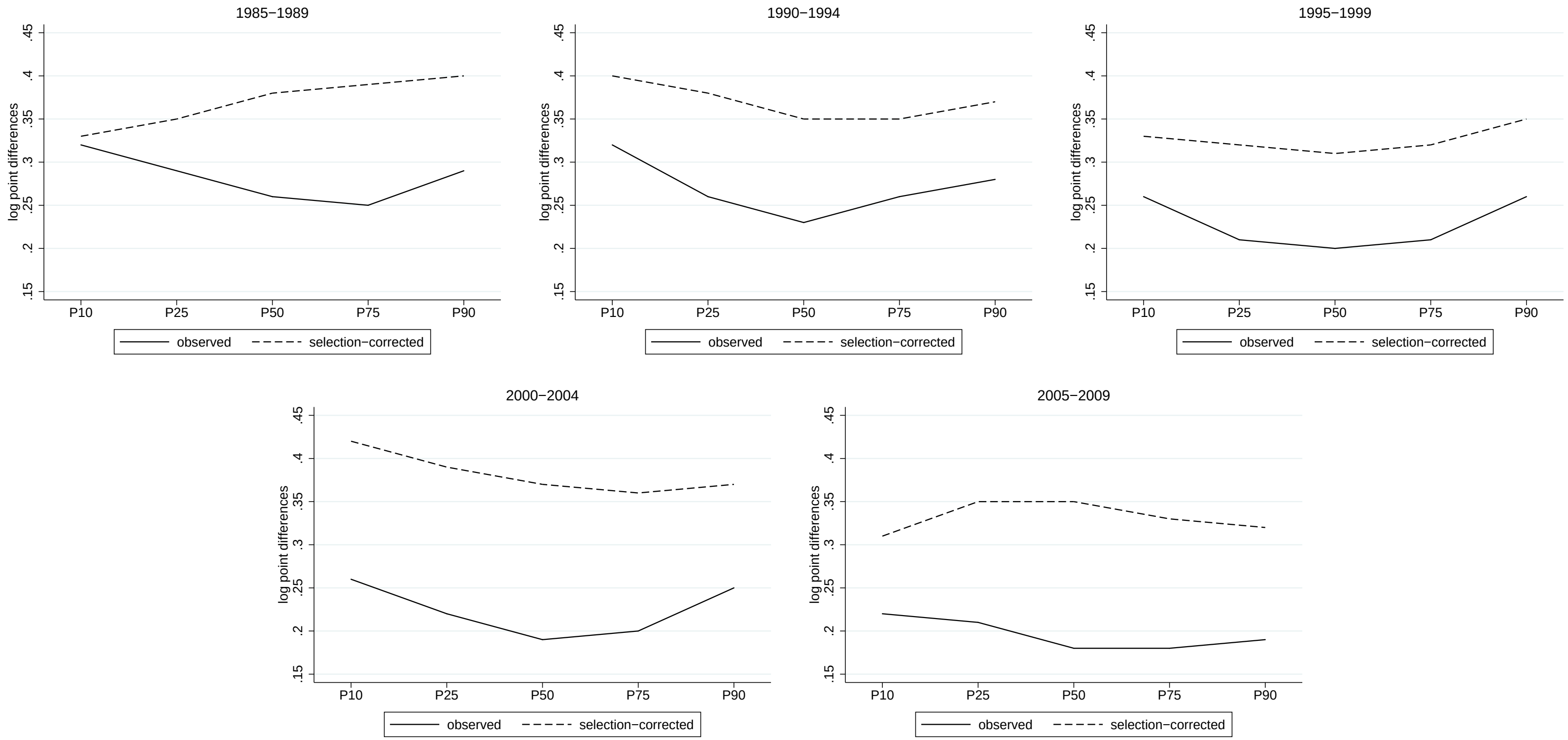


Figure B.3: Observed vs selection-corrected wage gap, over time



Source: SOEP.v32, own calculations.

Appendix to Chapter 3

C.1 Tables

Table C.1: Person-year observations, full-time status

	All	Work FT	Do Not Work FT	Ever Worked FT		Never Worked FT	
	(1)	(2)	(3)	(4)	% of (3)	(5)	% of (1)
1984	3216	974	2242	727	32%	1515	47%
1985	3312	1052	2260	740	33%	1520	46%
1986	3192	1037	2155	763	35%	1392	44%
1987	3127	1044	2083	794	38%	1289	41%
1988	2978	983	1995	817	41%	1178	40%
1989	2900	1045	1855	770	42%	1085	37%
1990	2862	1044	1818	806	44%	1012	35%
1991	2859	1031	1828	847	46%	981	34%
1992	2840	1005	1835	915	50%	920	32%
1993	2813	1013	1800	911	51%	889	32%
1994	2760	937	1823	960	53%	863	31%
1995	2924	987	1937	1035	53%	902	31%
1996	2862	959	1903	1049	55%	854	30%
1997	2781	937	1844	1044	57%	800	29%
1998	2978	976	2002	1100	55%	902	30%
1999	2863	993	1870	1044	56%	826	29%
2000	4661	1538	3123	1371	44%	1752	38%
2001	4553	1484	3069	1359	44%	1710	38%
2002	4768	1548	3220	1502	47%	1718	36%
2003	4621	1521	3100	1481	48%	1619	35%
2004	4400	1423	2977	1481	50%	1496	34%
2005	4153	1369	2784	1434	52%	1350	33%
2006	4265	1409	2856	1445	51%	1411	33%
2007	4012	1353	2659	1363	51%	1296	32%
2008	3706	1288	2418	1259	52%	1159	31%
2009	3363	1213	2150	1155	54%	995	30%
2010	5607	1361	4246	1551	37%	2695	48%
2011	6549	1751	4798	1516	32%	3282	50%
2012	6365	1825	4540	1369	30%	3171	50%
2013	7154	2027	5127	1401	27%	3726	52%
2014	6626	2003	4623	1214	26%	3409	51%
2015	5832	1793	4039	1046	26%	2993	51%
Total	127902	40923	86979	36269	42%	50710	40%

Source: SOEP.v32, own calculations.

Table C.2: Person-year observations, part-time status

	All	Work PT	Do Not Work PT	Ever Worked PT		Never Worked PT	
	(1)	(2)	(3)	(4)	% of (3)	(5)	% of (1)
1984	3216	524	2692	1078	40%	1614	50%
1985	3312	567	2745	1089	40%	1656	50%
1986	3192	542	2650	1119	42%	1531	48%
1987	3127	550	2577	1140	44%	1437	46%
1988	2978	555	2423	1119	46%	1304	44%
1989	2900	526	2374	1159	49%	1215	42%
1990	2862	552	2310	1150	50%	1160	41%
1991	2859	613	2246	1122	50%	1124	39%
1992	2840	584	2256	1164	52%	1092	38%
1993	2813	575	2238	1173	52%	1065	38%
1994	2760	585	2175	1141	52%	1034	37%
1995	2924	656	2268	1176	52%	1092	37%
1996	2862	664	2198	1141	52%	1057	37%
1997	2781	643	2138	1131	53%	1007	36%
1998	2978	686	2292	1245	54%	1047	35%
1999	2863	759	2104	1123	53%	981	34%
2000	4661	1366	3295	1658	50%	1637	35%
2001	4553	1405	3148	1562	50%	1586	35%
2002	4768	1531	3237	1609	50%	1628	34%
2003	4621	1529	3092	1532	50%	1560	34%
2004	4400	1466	2934	1487	51%	1447	33%
2005	4153	1406	2747	1430	52%	1317	32%
2006	4265	1451	2814	1467	52%	1347	32%
2007	4012	1419	2593	1350	52%	1243	31%
2008	3706	1328	2378	1232	52%	1146	31%
2009	3363	1204	2159	1149	53%	1010	30%
2010	5607	1945	3662	1909	52%	1753	31%
2011	6549	2541	4008	1848	46%	2160	33%
2012	6365	2498	3867	1724	45%	2143	34%
2013	7154	2837	4317	1714	40%	2603	36%
2014	6626	2678	3948	1529	39%	2419	37%
2015	5832	2402	3430	1328	39%	2102	36%
Total	127902	38587	89315	42798	48%	46517	36%

Source: SOEP.v32, own calculations.

Table C.3: Availability of multiple imputations

	All	Need FT Imp	Multiple FT avail.		Need PT Imp	Multiple PT avail.	
	(1)	(2)	(3)	% of (2)	(4)	(5)	% of (4)
1990	2862	1818	638	35%	2310	884	38%
1991	2859	1828	679	37%	2246	858	38%
1992	2840	1835	740	40%	2256	904	40%
1993	2813	1800	739	41%	2238	910	41%
1994	2760	1823	780	43%	2175	906	42%
1995	2924	1937	830	43%	2268	927	41%
1996	2862	1903	834	44%	2198	895	41%
1997	2781	1844	841	46%	2138	897	42%
1998	2978	2002	883	44%	2292	988	43%
1999	2863	1870	827	44%	2104	874	42%
2000	4661	3123	1069	34%	3295	1269	39%
2001	4553	3069	1034	34%	3148	1185	38%
2002	4768	3220	1162	36%	3237	1210	37%
2003	4621	3100	1138	37%	3092	1144	37%
2004	4400	2977	1152	39%	2934	1094	37%
2005	4153	2784	1118	40%	2747	1052	38%
2006	4265	2856	1119	39%	2814	1074	38%
2007	4012	2659	1075	40%	2593	983	38%
2008	3706	2418	984	41%	2378	894	38%
2009	3363	2150	889	41%	2159	824	38%

Source: SOEP.v32, own calculations.

Table C.4: Sample means (standard deviations) and sample size

	age		years of education		intermediate degree (0/1)		advanced degree (0/1)		experience full-time		experience part-time		sample size
Complete Sample													
1990-1994	36.8	(10.2)	11.3	(2.3)	0.40	(0.49)	0.08	(0.27)	8.0	(7.6)	2.9	(5.1)	14134
1995-1999	37.2	(9.7)	11.6	(2.5)	0.44	(0.50)	0.11	(0.31)	8.4	(7.8)	3.1	(4.9)	14408
2000-2004	37.9	(9.6)	11.9	(2.4)	0.48	(0.50)	0.13	(0.34)	8.5	(8.0)	3.5	(5.5)	23003
2005-2009	38.5	(9.7)	12.1	(2.5)	0.51	(0.50)	0.16	(0.36)	8.6	(8.3)	4.2	(5.6)	19499
Subsample Working Full-Time													
1990-1994	35.1	(10.2)	11.5	(2.4)	0.45	(0.50)	0.09	(0.29)	11.4	(9.0)	1.6	(3.7)	5030
1995-1999	36.4	(9.8)	11.9	(2.5)	0.45	(0.50)	0.14	(0.35)	12.3	(9.1)	1.6	(3.5)	4852
2000-2004	37.3	(9.7)	12.4	(2.5)	0.50	(0.50)	0.18	(0.38)	12.7	(9.2)	1.8	(3.9)	7514
2005-2009	38.2	(9.8)	12.7	(2.7)	0.50	(0.50)	0.22	(0.41)	12.8	(9.4)	2.4	(4.4)	6632
Subsample Working Part-Time													
1990-1994	39.9	(8.7)	11.3	(2.4)	0.36	(0.48)	0.09	(0.29)	6.8	(5.6)	7.2	(6.7)	2909
1995-1999	40.2	(8.4)	11.8	(2.5)	0.43	(0.50)	0.12	(0.33)	7.2	(6.1)	7.1	(6.0)	3408
2000-2004	40.5	(8.2)	11.8	(2.3)	0.46	(0.50)	0.12	(0.32)	7.3	(6.4)	7.2	(6.4)	7297
2005-2009	40.9	(8.5)	12.0	(2.3)	0.51	(0.50)	0.12	(0.33)	7.3	(6.6)	7.8	(6.2)	6808
Subsample Out-of-Work													
1990-1994	36.6	(10.6)	11.0	(2.3)	0.37	(0.48)	0.06	(0.24)	5.7	(6.0)	1.8	(3.9)	6195
1995-1999	36.0	(10.0)	11.2	(2.4)	0.42	(0.49)	0.08	(0.27)	6.0	(6.2)	1.9	(3.8)	6148
2000-2004	36.1	(10.1)	11.5	(2.4)	0.48	(0.50)	0.10	(0.30)	5.6	(6.3)	2.0	(4.1)	8192
2005-2009	36.3	(10.1)	11.7	(2.4)	0.51	(0.50)	0.12	(0.32)	5.4	(6.3)	2.4	(4.1)	6059

Source: SOEP.v32, weighting factors used, own calculations.

Table C.5: Descriptive statistics of wage equation regressors, alternative rule

	age	years of education	intermediate degree (0/1)	advanced degree (0/1)	experience full-time	experience part-time	sample size
Never Worked FT, working PT							
1990-1994	43.4 (8.3)	10.8 (2.0)	0.27 (0.45)	0.05 (0.22)	6.1 (4.8)	9.3 (7.4)	1342
1995-1999	42.8 (8.5)	11.4 (2.2)	0.37 (0.48)	0.09 (0.29)	6.3 (6.3)	8.6 (6.7)	1370
2000-2004	42.1 (8.2)	11.6 (2.2)	0.42 (0.49)	0.10 (0.30)	6.7 (6.3)	8.1 (7.1)	3648
2005-2009	41.9 (8.8)	11.8 (2.2)	0.50 (0.50)	0.11 (0.31)	6.0 (5.8)	9.0 (6.7)	3057
Never Worked FT, nonworking							
1990-1994	41.2 (10.1)	10.7 (2.2)	0.28 (0.45)	0.06 (0.23)	5.4 (5.5)	2.2 (4.6)	3323
1995-1999	38.7 (10.5)	11.0 (2.4)	0.35 (0.48)	0.08 (0.28)	4.8 (5.3)	2.1 (4.5)	2914
2000-2004	38.0 (10.3)	11.3 (2.3)	0.42 (0.49)	0.09 (0.29)	5.1 (5.9)	2.2 (4.6)	4647
2005-2009	37.8 (10.7)	11.3 (2.3)	0.48 (0.50)	0.09 (0.29)	4.3 (5.8)	2.5 (4.5)	3154
Never Worked PT, working FT							
1990-1994	36.1 (10.6)	11.2 (2.1)	0.45 (0.50)	0.06 (0.24)	13.7 (9.8)	0.6 (2.1)	2653
1995-1999	36.5 (10.1)	11.6 (2.3)	0.48 (0.50)	0.09 (0.29)	14.2 (9.9)	0.4 (1.4)	2558
2000-2004	37.6 (9.8)	12.2 (2.3)	0.53 (0.50)	0.16 (0.36)	14.3 (9.7)	0.8 (2.6)	4274
2005-2009	37.9 (10.1)	12.5 (2.5)	0.53 (0.50)	0.19 (0.40)	14.5 (9.9)	0.8 (2.4)	3630
Never Worked PT, nonworking							
1990-1994	39.2 (11.8)	10.7 (2.0)	0.33 (0.47)	0.04 (0.19)	6.1 (6.8)	0.9 (3.0)	2822
1995-1999	37.3 (11.2)	10.9 (2.2)	0.39 (0.49)	0.06 (0.23)	6.4 (7.2)	0.9 (2.8)	2626
2000-2004	36.7 (11.2)	11.2 (2.3)	0.45 (0.50)	0.09 (0.28)	5.6 (6.9)	1.0 (3.0)	3584
2005-2009	35.6 (11.3)	11.3 (2.3)	0.52 (0.50)	0.09 (0.28)	5.0 (6.6)	0.9 (2.3)	2433

Source: SOEP.v32, weighting factors used, own calculations.

Table C.6: Sample means (standard deviations) and sample size

	age	years of education	intermediate degree (0/1)	advanced degree (0/1)	experience full-time	experience part-time	sample size
			Basis for FT alternative imputation, working PT				
1990-1994	36.72 (7.89)	11.80 (2.58)	0.44 (0.50)	0.13 (0.33)	7.42 (6.09)	5.30 (5.22)	1,567
1995-1999	38.23 (7.73)	12.12 (2.64)	0.48 (0.50)	0.15 (0.35)	7.83 (5.83)	6.04 (5.10)	2,038
2000-2004	38.76 (7.82)	12.11 (2.45)	0.50 (0.50)	0.14 (0.35)	7.94 (6.51)	6.19 (5.34)	3,649
2005-2009	39.98 (8.02)	12.16 (2.36)	0.53 (0.50)	0.13 (0.34)	8.44 (6.97)	6.77 (5.45)	3,751
			Basis for FT alternative imputation, not working				
1990-1994	31.48 (8.68)	11.34 (2.30)	0.48 (0.50)	0.07 (0.25)	6.10 (6.46)	1.28 (2.81)	2,872
1995-1999	33.55 (8.84)	11.41 (2.32)	0.50 (0.50)	0.07 (0.26)	7.11 (6.81)	1.59 (3.10)	3,234
2000-2004	33.30 (9.08)	11.93 (2.37)	0.57 (0.50)	0.11 (0.31)	6.50 (6.77)	1.70 (3.31)	3,545
2005-2009	34.47 (9.07)	12.21 (2.51)	0.54 (0.50)	0.15 (0.36)	6.71 (6.50)	2.27 (3.61)	2,905
			Basis for PT alternative imputation, working FT				
1990-1994	34.02 (9.61)	11.79 (2.59)	0.45 (0.50)	0.13 (0.33)	8.95 (7.41)	2.56 (4.58)	2,377
1995-1999	36.24 (9.38)	12.18 (2.79)	0.42 (0.49)	0.19 (0.40)	10.18 (7.53)	2.92 (4.43)	2,294
2000-2004	36.88 (9.50)	12.56 (2.78)	0.46 (0.50)	0.21 (0.41)	10.41 (8.05)	3.21 (4.90)	3,240
2005-2009	38.77 (9.40)	12.92 (2.91)	0.46 (0.50)	0.25 (0.44)	10.37 (8.07)	4.66 (5.43)	3,002
			Basis for PT alternative imputation, not working				
1990-1994	34.61 (9.06)	11.26 (2.43)	0.40 (0.49)	0.08 (0.27)	5.45 (5.16)	2.43 (4.41)	3,373
1995-1999	35.09 (8.87)	11.43 (2.48)	0.45 (0.50)	0.10 (0.30)	5.63 (5.35)	2.58 (4.32)	3,522
2000-2004	35.59 (9.02)	11.77 (2.37)	0.50 (0.50)	0.11 (0.31)	5.70 (5.69)	2.82 (4.73)	4,608
2005-2009	36.76 (9.17)	12.03 (2.50)	0.51 (0.50)	0.14 (0.35)	5.75 (5.99)	3.45 (4.80)	3,626

Source: SOEP.v32, weighting factors used, own calculations.

Table C.7: Full-time wages: subsample sizes for each grid of quantile regressions

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1.984	247	249	242	243	202	177	170	166	149	151	143	136	117	106	88	82	68	62	52	44
1.985	246	260	241	239	203	175	173	173	157	155	149	137	119	107	86	87	70	59	52	42
1.986	239	256	233	248	210	183	177	175	159	164	156	138	120	113	93	95	83	75	67	53
1.987	242	264	250	248	209	186	178	177	159	155	148	138	119	108	91	95	83	83	71	55
1.988	219	239	228	240	215	195	192	186	172	165	158	147	133	122	102	102	86	88	77	67
1.989	144	192	192	206	185	175	178	172	157	152	158	144	135	121	106	100	88	89	75	63
1.990	0	157	166	211	192	172	176	179	169	163	161	155	135	128	108	111	97	101	82	67
1.991	166	0	140	194	179	165	167	177	164	164	167	155	140	123	111	112	96	101	85	78
1.992	219	158	0	160	166	181	181	189	176	180	183	164	151	135	116	123	110	108	100	93
1.993	225	198	151	0	116	161	165	188	170	183	191	168	155	150	131	133	122	121	100	97
1.994	256	241	215	178	0	143	171	192	168	179	186	162	146	144	129	126	122	121	109	97
1.995	258	249	238	213	142	0	144	178	172	199	216	193	174	176	151	150	142	139	122	115
1.996	274	262	255	222	182	164	0	123	148	185	191	194	173	168	147	152	144	136	118	114
1.997	265	260	253	241	210	190	130	0	128	178	207	194	178	171	151	155	148	143	120	113
1.998	271	254	252	256	234	213	179	150	0	172	218	218	208	200	179	173	166	161	139	133
1.999	266	254	249	251	225	218	191	173	133	0	155	165	169	173	154	157	158	154	136	127
2.000	248	237	241	238	216	219	191	176	166	127	0	208	255	300	276	275	287	276	260	245
2.001	236	219	219	221	197	203	196	185	180	163	209	0	187	260	232	257	266	253	247	246
2.002	229	219	216	217	201	221	208	212	211	202	287	214	0	234	263	280	307	292	294	286
2.003	217	201	209	215	204	213	205	205	207	206	325	259	194	0	180	239	275	285	287	290
2.004	207	187	182	194	193	195	193	192	192	203	310	261	237	194	0	202	265	282	289	298
2.005	191	173	174	184	187	196	195	199	210	221	335	286	258	230	177	0	180	245	265	266
2.006	177	165	170	168	173	184	186	184	198	209	316	278	269	241	208	147	0	208	257	265
2.007	166	154	155	167	171	180	184	179	187	204	303	273	258	242	231	205	171	0	203	245
2.008	145	140	147	149	149	162	161	164	167	180	278	249	236	230	212	199	190	171	0	183
2.009	133	125	132	139	142	153	147	147	150	163	252	233	211	211	195	178	178	172	140	0
2.010	119	111	116	123	132	145	143	140	135	153	243	225	219	213	202	195	189	198	189	165
2.011	106	98	99	107	107	116	121	119	114	131	220	212	206	198	192	180	184	188	174	156
2.012	95	84	89	98	94	106	108	106	93	117	196	185	180	180	174	168	162	186	167	164
2.013	81	73	80	83	80	93	92	95	82	106	182	169	161	164	154	157	165	171	161	169
2.014	72	70	75	80	77	83	81	89	80	99	172	165	164	163	156	155	168	168	169	174
2.015	63	59	68	69	68	73	76	85	77	95	163	156	148	150	142	141	163	159	158	164

Source: SOEP.v32, own calculations.

Table C.8: Part-time wages: subsample sizes for each grid of quantile regressions

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1.984	270	323	308	294	294	301	295	275	272	276	248	234	214	213	196	168	151	137	120	101
1.985	258	308	297	281	286	287	287	274	267	278	248	246	225	223	211	186	170	158	137	115
1.986	256	309	290	286	283	288	286	271	261	273	251	243	233	224	216	198	181	167	150	133
1.987	243	300	284	297	296	300	305	299	276	299	269	267	257	256	243	226	205	190	166	147
1.988	204	269	262	276	285	289	292	291	278	297	277	279	275	266	253	239	221	205	180	160
1.989	185	265	272	289	289	306	301	305	294	317	307	305	297	291	273	251	240	226	202	177
1.990	0	193	213	248	260	275	276	276	270	293	275	293	295	300	286	261	252	238	208	192
1.991	113	0	141	194	212	232	258	256	248	289	287	302	310	309	299	270	260	246	210	192
1.992	146	147	0	158	203	230	252	264	263	304	301	312	322	336	312	291	277	268	233	214
1.993	186	203	151	0	158	210	240	257	262	307	293	306	327	332	320	303	276	275	244	226
1.994	170	184	149	125	0	161	212	242	243	282	280	297	312	320	303	298	275	277	254	237
1.995	148	158	145	136	120	0	177	219	232	287	282	302	323	339	334	330	307	310	284	266
1.996	149	167	154	141	142	143	0	155	181	240	248	276	301	327	328	320	300	304	273	264
1.997	135	150	148	138	149	166	148	0	147	227	240	277	301	318	323	317	295	308	282	268
1.998	133	148	148	140	148	178	185	148	0	228	275	308	326	353	360	355	346	338	321	314
1.999	105	128	125	119	123	152	154	134	123	0	179	245	270	297	307	314	299	303	287	285
2.000	102	130	133	122	125	152	149	147	160	167	0	344	413	474	482	496	478	495	494	449
2.001	99	117	120	113	123	147	144	137	158	183	279	0	317	402	437	463	446	467	457	427
2.002	89	106	113	112	110	135	143	133	150	174	302	257	0	374	425	455	447	474	456	406
2.003	91	94	105	101	108	133	142	129	141	167	302	279	299	0	300	375	379	409	414	376
2.004	73	82	82	81	87	109	121	118	125	145	294	291	332	277	0	282	325	384	395	352
2.005	82	83	85	86	88	115	117	114	114	138	278	295	324	321	255	0	252	331	366	332
2.006	65	63	67	67	76	99	99	105	105	126	256	269	302	309	290	236	0	314	389	363
2.007	62	63	67	67	78	97	100	101	95	126	249	248	294	301	294	271	251	0	302	320
2.008	52	51	50	54	70	73	70	80	84	105	227	232	281	279	271	263	282	250	0	240
2.009	46	46	43	45	60	70	77	79	84	104	209	221	251	255	263	246	269	272	229	0
2.010	38	38	37	41	52	65	66	76	80	97	191	189	236	218	235	233	245	250	238	163
2.011	33	35	37	35	43	52	52	64	74	86	173	173	211	226	227	231	238	247	241	206
2.012	28	29	28	31	38	45	47	54	65	75	151	166	200	202	196	198	213	233	232	201
2.013	21	25	26	27	36	43	40	50	52	63	133	144	182	184	189	169	202	204	208	197
2.014	17	23	24	26	36	39	35	43	45	58	122	133	158	166	170	162	181	197	202	192
2.015	15	19	21	23	29	32	23	34	37	54	107	112	133	141	142	142	157	151	168	165

Source: SOEP.v32, own calculations.

Table C.9: Full-time wages: Subsample sizes for each grid of quantile regressions

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1990	404	453	519	583	669	808																				
1991	365	399	450	502	587	700	798																			
1992	315	344	389	437	507	602	681	780																		
1993	282	307	335	389	443	546	598	671	767																	
1994	230	243	268	320	364	448	494	558	628	738																
1995	208	221	247	290	319	388	439	489	544	626	710															
1996	193	206	232	265	289	352	387	432	477	555	612	748														
1997	170	182	210	237	266	319	353	389	427	486	529	640	736													
1998	143	156	179	202	223	272	300	337	363	409	450	541	610	687												
1999	126	135	152	179	196	240	262	294	324	356	410	484	539	603	755											
2000	129	130	144	172	188	212	239	262	287	317	358	413	468	510	619	741										
2001	114	115	134	153	163	192	211	237	260	284	322	372	394	435	516	621	1,189									
2002	97	101	114	133	139	160	179	195	220	246	271	320	345	366	439	520	972	1,082								
2003	87	90	97	119	124	147	168	186	202	213	238	281	309	334	394	466	847	933	1,189							
2004	77	84	90	106	113	131	151	162	184	191	214	256	278	303	357	418	753	834	1,031	1,152						
2005	72	74	76	89	102	123	134	150	164	173	192	229	243	263	313	361	652	719	895	997	1,078					
2006	60	66	64	76	88	106	116	126	137	150	165	200	217	231	274	320	583	641	785	873	935	1,068				
2007	54	61	59	64	76	96	101	110	124	133	144	180	192	207	243	284	501	562	689	745	794	884	1,061			
2008	52	57	55	64	74	90	97	96	104	120	127	154	173	184	209	244	429	484	599	653	701	772	909	1,013		
2009	48	54	52	64	67	89	97	91	94	107	116	135	152	163	185	217	376	413	517	567	603	673	786	856	957	
2010							75	76	74	83	88	103	118	124	148	178	313	354	425	471	505	552	662	703	774	880
2011							72	67	64	74	80	96	109	113	134	155	271	302	364	407	432	471	552	604	652	747
2012							69	66	65	68	76	90	101	107	109	129	241	260	302	336	363	394	468	488	531	607
2013							71	62	61	68	74	81	92	93	99	112	204	215	258	284	304	345	390	430	456	513
2014							63	54	52	61	63	75	90	87	92	102	186	193	233	256	275	307	351	381	391	445
2015							56	50	47	52	60	72	82	79	85	92	164	171	204	224	240	268	291	317	327	365

Source: SOEP.v32, own calculations.

Table C.10: Part-time wages: subsample sizes for each grid of quantile regressions

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1990	182	199	226	255	301	337																				
1991	177	194	224	250	293	304	396																			
1992	145	163	190	203	240	249	327	416																		
1993	131	147	168	174	207	207	262	336	387																	
1994	116	134	164	165	191	203	249	312	337	386																
1995	106	129	151	154	185	189	243	306	318	344	412															
1996	89	111	128	131	161	168	217	262	280	302	345	452														
1997	82	99	117	118	143	143	203	245	246	269	301	387	461													
1998	67	90	109	117	134	130	172	218	217	239	267	331	387	433												
1999	55	73	91	97	115	115	164	194	202	221	259	311	366	394	500											
2000	43	65	78	86	96	93	135	153	157	179	211	264	311	331	394	504										
2001	49	62	83	82	89	90	122	140	150	167	195	236	278	291	352	435	1,002									
2002	44	57	75	76	77	81	113	127	128	142	176	215	245	265	325	396	871	974								
2003	33	45	68	63	72	72	97	117	116	129	155	188	212	240	291	358	767	859	1079							
2004	29	37	56	55	64	70	92	107	114	123	157	187	202	226	278	342	683	752	938	1092						
2005	27	35	45	48	55	61	81	94	100	104	140	158	179	208	254	309	621	667	840	942	1065					
2006	22	24	34	37	44	47	67	80	84	95	119	143	164	192	226	288	564	608	748	836	922	1006				
2007	16	17	28	27	30	31	53	61	65	74	93	113	135	157	198	245	491	537	648	748	793	862	1,049			
2008	14	13	22	24	29	25	49	60	66	69	82	105	128	145	179	226	416	453	554	639	675	723	854	974		
2009	14	12	16	18	24	21	39	51	56	60	67	82	96	114	142	183	344	369	487	551	590	638	740	808	910	
2010							33	38	43	44	55	68	77	96	120	150	302	326	409	483	496	534	617	694	761	854
2011							32	37	39	43	52	69	81	95	113	136	262	275	342	391	413	444	533	594	655	710
2012							21	29	33	33	40	52	64	77	82	100	201	205	253	312	328	364	428	465	498	546
2013							18	27	31	30	39	46	62	70	81	96	181	190	234	279	286	321	376	421	454	482
2014							18	26	30	26	32	39	52	65	65	82	160	165	206	250	246	274	331	361	391	418
2015							13	22	22	22	28	33	46	52	55	73	134	143	172	201	212	228	279	325	339	371

Source: SOEP.v32, own calculations.

Table C.11: Selection-corrected full-time wage distribution

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) Nominal wages				
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .10)$	1.57	1.72	1.67	1.68
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .25)$	1.80	1.98	1.96	1.99
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .50)$	2.03	2.22	2.25	2.30
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .75)$	2.27	2.45	2.51	2.60
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .90)$	2.48	2.67	2.76	2.86
(B) Real wages (1995 Euro)				
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .10)$	1.68	1.69	1.58	1.51
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .25)$	1.89	1.96	1.87	1.82
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .50)$	2.12	2.20	2.16	2.13
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .75)$	2.35	2.43	2.42	2.43
$\widehat{F}^{-1}_{W^{FT}}$ $(\tau = .90)$	2.56	2.64	2.66	2.68

Comments: Units are log hourly wages. Wages deflated according to CPI figures provided by the German Federal Statistical Office.

Source: SOEP.v32,own calculations.

Table C.12: Selection-corrected part-time wage distribution

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) Nominal wages				
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .10)$	1.47	1.57	1.72	1.74
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .25)$	1.75	1.85	2.01	2.04
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .50)$	2.05	2.16	2.32	2.38
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .75)$	2.35	2.48	2.60	2.68
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .90)$	2.68	2.84	2.90	2.99
(B) Real wages (1995 Euro)				
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .10)$	1.57	1.54	1.63	1.56
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .25)$	1.83	1.83	1.92	1.86
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .50)$	2.13	2.14	2.23	2.20
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .75)$	2.43	2.45	2.51	2.50
$\widehat{F}^{-1}_{W^{PT}}$ $(\tau = .90)$	2.76	2.81	2.81	2.81

Comments: Units are log hourly wages. Wages deflated according to CPI figures provided by the German Federal Statistical Office.

Source: SOEP.v32,own calculations.

Table C.13: Wage dispersion within full- and part-time employment

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) Wage dispersion within part-time employment:				
Overall ($P90 - P10$)	1.29	1.34	1.31	1.31
Upper-tail ($P90 - P50$)	0.68	0.70	0.60	0.59
Lower-tail ($P50 - P10$)	0.60	0.64	0.71	0.72
(B) Wage dispersion within full-time employment:				
Overall ($P90 - P10$)	0.89	0.87	0.96	1.06
Upper-tail ($P90 - P50$)	0.42	0.41	0.44	0.49
Lower-tail ($P50 - P10$)	0.47	0.46	0.52	0.56

Comments: Units are log-point differences between the inverse cumulative distributions evaluated at $P10$, $P50$ and $P90$.

Source: SOEP.v32, own calculations. Weighting factors used.

Table C.14: Presence of part-time employment in the low- and high-wage sectors

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) $F(low\ wage)$				
part-time distribution	.22	.25	.27	.30
full-time distribution	.11	.11	.12	.13
overall distribution	.15	.17	.19	.21
(B) $1 - F(high\ wage)$				
part-time distribution	.16	.16	.13	.14
full-time distribution	.11	.12	.16	.19
overall distribution	.13	.14	.15	.16
(C) Frequence of part-time employment in:				
the low-wage sector	56%	62%	67%	68%
the high-wage sector	46%	48%	44%	39%

Source: SOEP.v32, own calculations. Weighting factors used.

Table C.15: Observed full-time wage distribution

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
		(A) Nominal wages		
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .10$)	1.71	1.91	1.96	1.98
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .25$)	1.97	2.16	2.22	2.28
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .50$)	2.18	2.38	2.47	2.55
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .75$)	2.40	2.60	2.71	2.80
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .90$)	2.61	2.79	2.92	3.04
		(B) Real wages (in 1995 Euro)		
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .10$)	1.81	1.89	1.87	1.81
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .25$)	2.07	2.13	2.13	2.10
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .50$)	2.27	2.35	2.38	2.38
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .75$)	2.48	2.57	2.61	2.62
$\widehat{F}^{-1}_{W^{FT}}$ ($\tau = .90$)	2.69	2.76	2.82	2.86

Comments: Units are log hourly wages. Wages deflated according to CPI figures provided by the German Federal Statistical Office.

Source: SOEP.v32,own calculations.

Table C.16: Observed part-time wage distribution

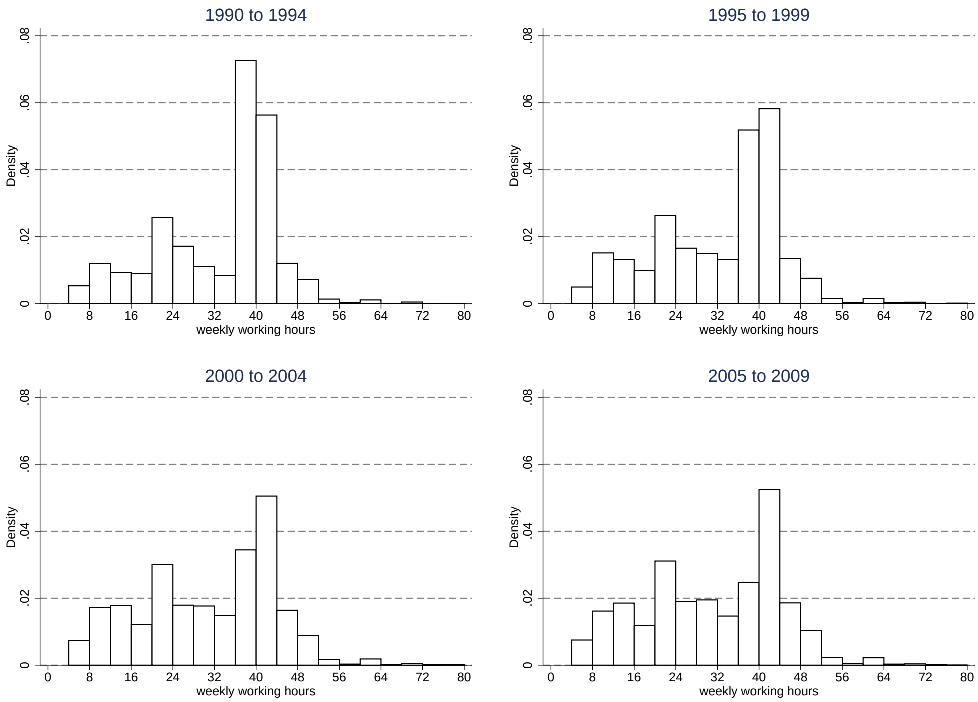
	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
		(A) Nominal wages		
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .10$)	1.49	1.66	1.61	1.65
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .25$)	1.78	1.95	2.00	2.00
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .50$)	2.10	2.30	2.32	2.36
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .75$)	2.42	2.59	2.63	2.67
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .90$)	2.78	2.99	2.92	2.96
		(B) Real wages (1995 Euro)		
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .10$)	1.59	1.62	1.53	1.48
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .25$)	1.87	1.91	1.90	1.82
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .50$)	2.19	2.27	2.22	2.18
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .75$)	2.49	2.56	2.53	2.49
$\widehat{F}^{-1}_{W^{PT}}$ ($\tau = .90$)	2.85	2.96	2.82	2.79

Comments: Units are log hourly wages. Wages deflated according to CPI figures provided by the German Federal Statistical Office.

Source: SOEP.v32,own calculations.

C.2 Figures

Figure C.1: Distribution of actual weekly working hours over time



Source: SOEP.v32, own calculations.

C.3 Goodness of fit of the conditional wage model

The imputation of non-realized wages is based on the conditional quantile regression wage model presented in Section 3.2.3. In order to evaluate the in-sample fit of the specification of the wage quantile regression, Figure C.2 contrasts observed log hourly wages (solid lines) with those produced by the conditional wage model (dashed lines) at different points of the unconditional distribution. In order to obtain unconditional wage distributions from the grid of conditional quantile regressions, I use the algorithm provided by Chernozhukov et al. (2013) which yields results equivalent to those suggested by Machado and Mata (2005).

Figure C.2: Goodness of fit of the conditional wage model

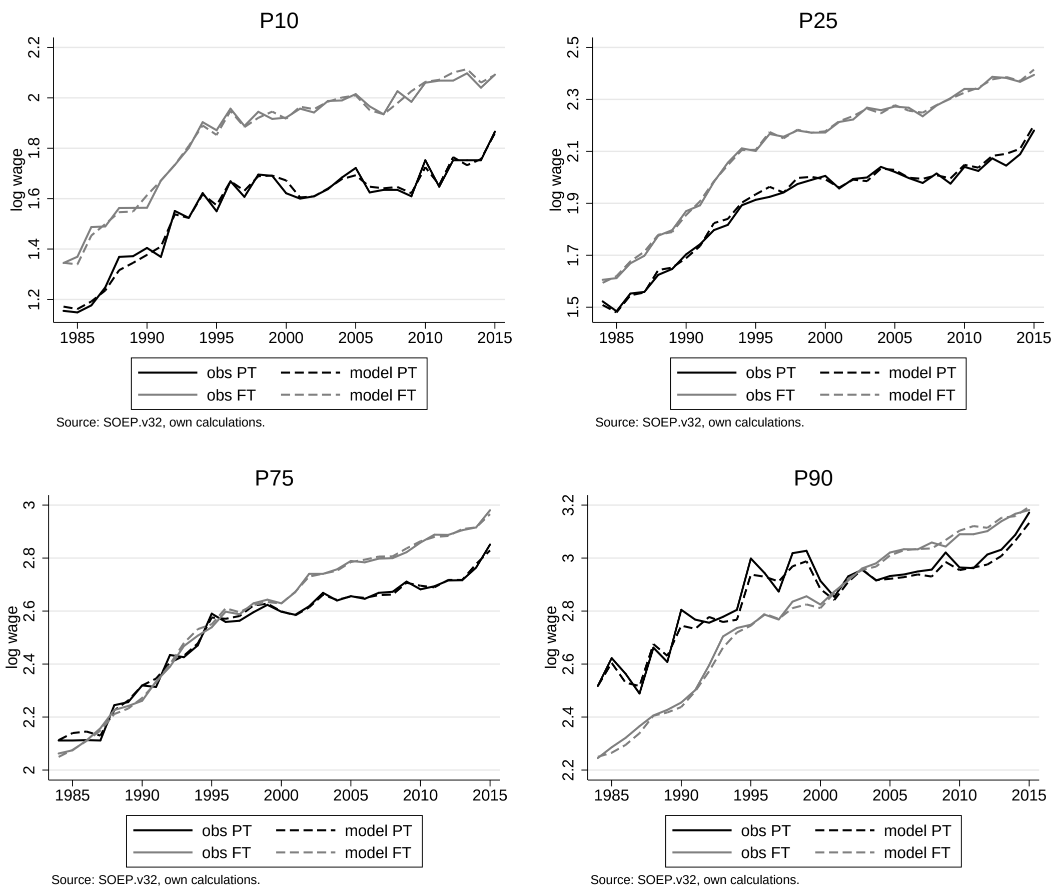


Figure C.2 shows that the conditional wage model fits very well the observed wage data at most points of the distribution, with the exception of a few peaks at the 90^{th}

percentile of the part-time distribution for the time period 1995-1999.

C.4 Robustness checks

In this section, I present results of several robustness checks that largely confirm the stability of the main results of the paper. First, I replicate the results on the corrected part-time wage gap for two alternative thresholds on weekly hours determining full- and part-time employment (Section C.4.1) set at 30 and 35 weekly working hours. In Section C.4.2 I examine the sensitivity of the results to using an alternative definition of working hours, namely contractual hours plus paid overtime. Next, I discuss possible alternative weighting rules of multiple imputations (section C.4.3) and, finally, I reflect on the sensitivity of the results to the median imputation rule used for recovering wages of those who I never observed engaging in full- or part-time work (section C.4.4).

C.4.1 Sensitivity of the results to the chosen hours cut-off

Tables C.17 and C.18 display figures on the observed and corrected part-time wage gap using two alternative hours thresholds (30 and 35 working hours per week, respectively). The first remark to be made is that results on both the observed and the adjusted part-time gap stay qualitatively the same if the hours threshold is changed to 30 or 35 working hours per week. Interestingly, adjusted part-time wage premiums become slightly higher under the 35 hours threshold, suggesting that long-hours part-time might be a driving force behind the relative more favourable pay of part-time work relative to full-time work. This is in line with the findings of Paul (2016).

Table C.17: Part-time gap, threshold at 30 hours per week

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
evaluated at the 10th percentile				
G_{obs} $(\tau = .10)$	.23	.28	.35	.35
$\widehat{G}_{corr}$ $(\tau = .10)$	.11	.11	- .08	- .03
evaluated at the 25th percentile				
G_{obs} $(\tau = .25)$	.19	.22	.23	.29
$\widehat{G}_{corr}$ $(\tau = .25)$	.05	.11	- .05	- .04
evaluated at the 50th percentile				
G_{obs} $(\tau = .50)$	.09	.09	.17	.20
$\widehat{G}_{corr}$ $(\tau = .50)$	.00	.04	- .07	- .07
evaluated at the 75th percentile				
G_{obs} $(\tau = .75)$	-.02	.01	.09	.14
$\widehat{G}_{corr}$ $(\tau = .75)$	-.07	- .05	- .10	- .07
evaluated at the 90th percentile				
G_{obs} $(\tau = .90)$	-.18	-.21	.00	.08
$\widehat{G}_{corr}$ $(\tau = .90)$	-.21	-.19	- .15	- .12

Comments: Units are log-point differences between inverse cumulative distributions. Unweighted average of multiple imputations.

Source: SOEP.v32,own calculations.

Table C.18: Part-time gap, threshold at 35 hours per week

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
evaluated at the 10th percentile				
G_{obs} ($\tau = .10$)	.21	.25	.32	.34
$\widehat{G}_{corr}$ ($\tau = .10$)	.11	.10	-.02	.01
evaluated at the 25th percentile				
G_{obs} ($\tau = .25$)	.20	.19	.21	.27
$\widehat{G}_{corr}$ ($\tau = .25$)	.04	.08	-.04	-.03
evaluated at the 50th percentile				
G_{obs} ($\tau = .50$)	.09	.06	.13	.17
$\widehat{G}_{corr}$ ($\tau = .50$)	.00	.00	-.09	-.10
evaluated at the 75th percentile				
G_{obs} ($\tau = .75$)	-.02	-.01	.07	.12
$\widehat{G}_{corr}$ ($\tau = .75$)	-.06	-.06	-.12	-.14
evaluated at the 90th percentile				
G_{obs} ($\tau = .90$)	-.18	-.18	-.00	.08
$\widehat{G}_{corr}$ ($\tau = .90$)	-.17	-.15	-.15	-.14

Comments: Units are log-point differences between inverse cumulative distributions.
Unweighted average of multiple imputations.

Source: SOEP.v32,own calculations.

C.4.2 Results using contractual- instead of actual- working hours

Table C.19 presents figures on the observed and corrected part-time wage gap using contractual working hours plus paid overtime instead of actual working hours. The concern here is that (paid) overtime hours may be systematically different for full- and part-time employment, thereby biasing my results in one or the other direction. However, Table C.19 shows that the results remain qualitatively the same regardless of the definition of working hours.

Table C.19: Part-time gap, contractual working hours

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
evaluated at the 10^{th} *percentile*				
$G_{obs}\,(\tau = .10)$	.25	.25	.37	.33
$\widehat{G}_{corr}\,(\tau = .10)$	.12	.15	-.03	-.08
evaluated at the 25^{th} *percentile*				
$G_{obs}\,(\tau = .25)$	.18	.21	.26	.29
$\widehat{G}_{corr}\,(\tau = .25)$	.08	.12	-.01	-.04
evaluated at the 50^{th} *percentile*				
$G_{obs}\,(\tau = .50)$	.07	.07	.14	.18
$\widehat{G}_{corr}\,(\tau = .50)$	.05	.04	-.05	-.07
evaluated at the 75^{th} *percentile*				
$G_{obs}\,(\tau = .75)$	-.05	-.01	.08	.15
$\widehat{G}_{corr}\,(\tau = .75)$	-.02	-.04	-.09	-.08
evaluated at the 90^{th} *percentile*				
$G_{obs}\,(\tau = .90)$	-.33	-.29	-.01	.05
$\widehat{G}_{corr}\,(\tau = .90)$	-.16	-.17	-.17	-.15

Comments: Units are log-point differences between inverse cumulative distributions. Unweighted average of multiple imputations.

Source: SOEP.v32,own calculations.

C.4.3 Weighting of multiple imputations

In this section I present robustness checks for alternative ways of weighting multiple imputations when these are available. These are available for individuals whom I observe working full-time (part-time) for several years. This is the case for 34% to 46% of all person-year observations who require imputations in each survey year[54]. If individuals would hold the same conditional rank all years they work in a particular type of employment, then multiple imputations available for an individual would be identical and there would be no need to deal with this issue. However, in the data one observes that individuals hold different conditional ranks over the years. While this does not contravene the (conditional) rank similarity assumption[55], it poses the question of how to best use the several available imputations for some individuals for a given year.

Whereas the persistence of wage shocks suffered by individuals is outside of the scope of this paper, I replicate the main results with two alternative weighting schemes in order to determine to which extent my results are driven by the choice of weights.

First, I use a closest-neighbour approach, in which the imputation which relies on information chronologically closer to the year of interest gets a weight of one and the remaining available imputations are not used at all. Formally:

$$d_{ikn}^{B} = \begin{cases} 1 & \text{if } |y_n - y_k| < |y_j - y_k| \forall n \neq j \\ 1 & \text{if } |y_n - y_k| = |y_j - y_k| \ and \ y_n > y_j \forall n \neq j \\ 0 & \text{if } |y_n - y_k| = |y_j - y_k| \ and \ y_n < y_j \forall n \neq j \\ 0 & \text{if } |y_n - y_k| > |y_j - y_k| \forall n \neq j \end{cases}$$

Secondly, I replicate the main results with an average of all available imputations which gives a higher weight to imputations relying in information chronologically closer to the year of interest than imputations relying in more distant information. Under this scheme, contrary to the closest neighbour approach, all imputations are used. The weighting function in this case reads:

---

[54] See Table C.3 in the Appendix for the exact figures.

[55] See Chernozhukov and Hansen (2005) for a discussion on why unsystematic slippages on individuals' conditional ranks do not pose a problem.

$$d_{ikn}^{C} = \frac{\sum_{n=1}^{N} (y_n - y_k)^2}{(y_n - y_k)^2} \cdot \left[\sum_{n=1}^{N} \frac{\sum_{n=1}^{N} (y_n - y_k)^2}{(y_n - y_k)^2} \right]^{-1} \tag{C.1}$$

Table C.20 shows that the results are robust to the choice of weighting rule, as there are no qualitative differences among the two. This, in turn, suggests that multiple imputations available for each individual are rather stable.

Table C.20: Sensitivity to the weighting scheme of multiple imputations

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
(A) Closest neighbor imputation				
$\widehat{G}_{corr}$ ($\tau = .10$)	.16	.19	-.02	-.02
$\widehat{G}_{corr}$ ($\tau = .25$)	.07	.16	-.04	-.02
$\widehat{G}_{corr}$ ($\tau = .50$)	.01	.08	-.07	-.07
$\widehat{G}_{corr}$ ($\tau = .75$)	-.08	-.01	-.09	-.07
$\widehat{G}_{corr}$ ($\tau = .90$)	-.23	-.19	-.14	-.15
(B) Weighted average of multiple imputations				
$\widehat{G}_{corr}$ ($\tau = .10$)	.12	.17	-.05	-.05
$\widehat{G}_{corr}$ ($\tau = .25$)	.06	.16	-.05	-.05
$\widehat{G}_{corr}$ ($\tau = .50$)	.00	.06	-.08	-.08
$\widehat{G}_{corr}$ ($\tau = .75$)	-.06	-.03	-.09	-.08
$\widehat{G}_{corr}$ ($\tau = .90$)	-.18	-.19	-.14	-.13

Comments: Units are log-point differences between inverse cumulative distributions.

*Statistical significance at the 5% level (computed by bootstrap with 200 replications).

Source: SOEP.v32,own calculations.

C.4.4 Effect of imputation for those never observed working

This section deals with the sensitivity of the results to the imputation of wages for those individuals who I never observe working and for those individuals who I only observe working in one of the two work arrangements under study. Given that this affects around 30% of all observations every year, it is important to check whether and to which extent the results are driven by the additional assumption made on the unobservable characteristics of this group.

Figure C.3: Imputation for individuals never in full- or part-time work

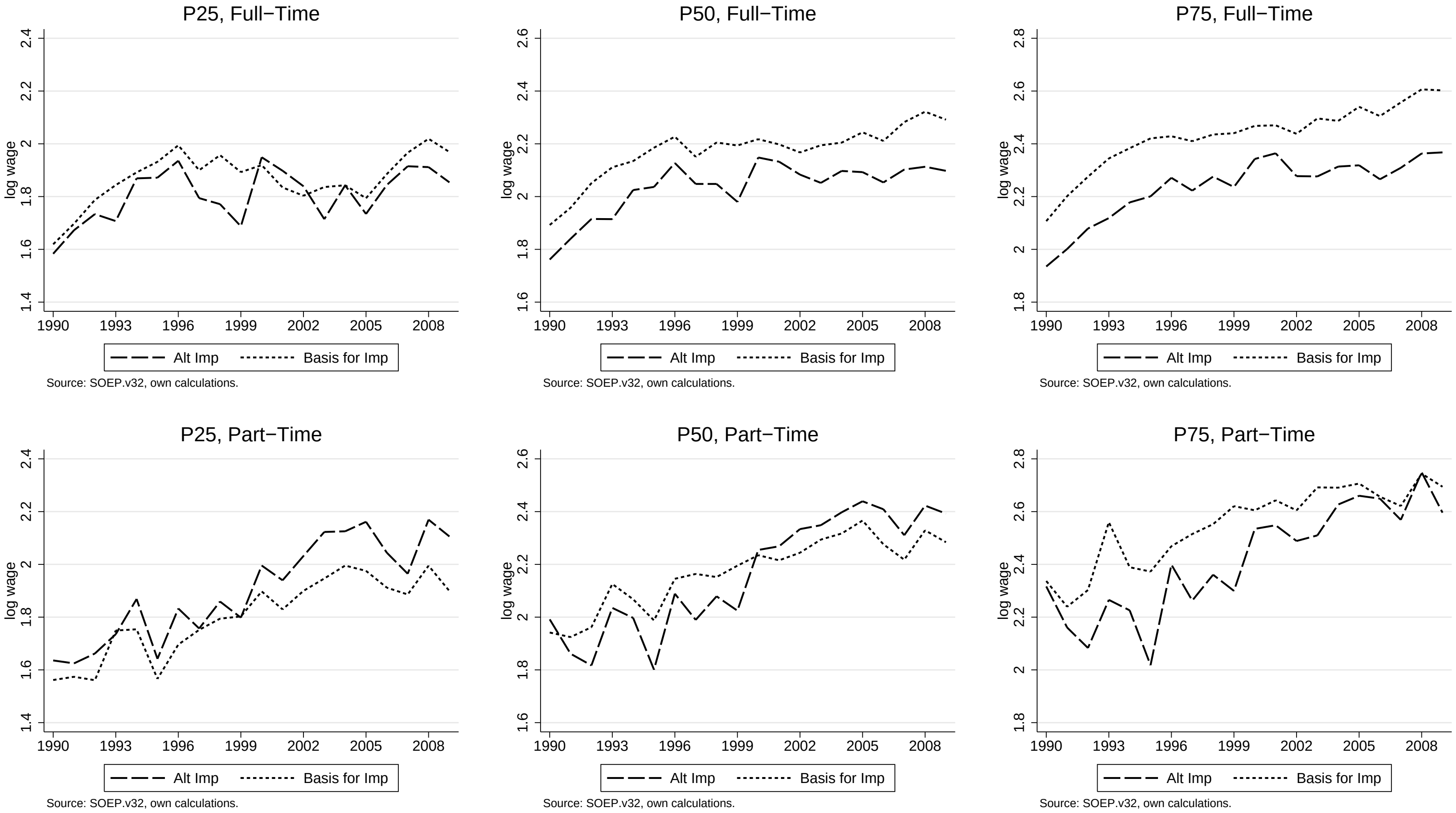
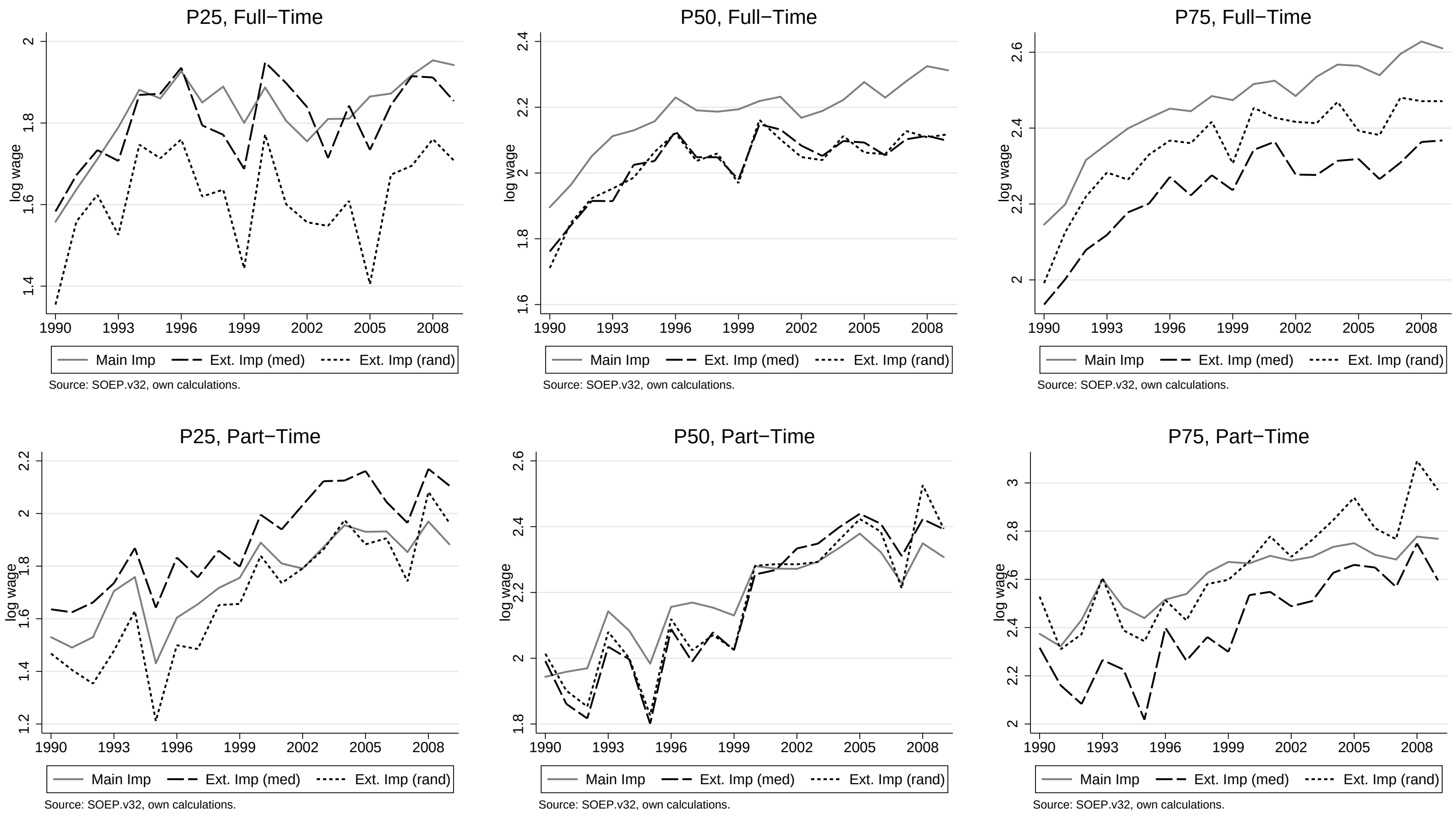


Figure C.3 presents separately the imputations made for those individuals for whom I never observe a full- and/or a part-time wage with the imputations carried out with the main model, which serve as the basis on which conditional median regressions are estimated for computing the alternative imputations.

In case of similar joint distribution of human capital characteristics among basis for imputation and alternative imputations, one would expect the two imputations to be very similar at the (unconditional) median, whereas the alternative imputation should have higher levels than the main imputation at the 25^{th} percentile and lower levels at the 75^{th} percentile, given the reduction in the dispersion of the alternative imputation using only regression coefficients from the conditional median regression. With regard to full-time wages, alternative imputations are below main imputations at the upper half of the distribution and almost at the same level for the 25^{th} percentile (see upper graphs of Figure C.3). In the case of part-time wages, both imputations yield similar results for the 25^{th} and 50^{th} percentiles. For the upper part of the distribution, the difference is wider during the 1990s but diminishes during the 2000s. These patterns are consistent with the differences in descriptive statistics on human capital stocks between the groups serving as basis for the imputation and the groups requiring imputations (see Tables C.9 and C.10 in the Appendix).

In order to determine whether the differential is driven by the human capital characteristics of these individuals or by the additional assumption made on their unobservables, I carry out the imputation for these individuals again, this time allocating them a random conditional rank. Figure C.4 offers a comparison of the imputations according to the main rule, those made under the median alternative imputation rule and those computed for random conditional ranks.

Figure C.4: Median versus random extended imputation



As one would expect, the two extended imputation rules make no difference for the unconditional mean, implying that the main results for the median are robust to the additional assumption made on the unobservables of this group. Unfortunately, this is not the case for the 25^{th} and 75^{th} percentiles. Imposing the median conditional rank for the imputation of a set of observations necessarily reduces the dispersion of the resulting distribution, which graphically implies the dashed line being above the dotted line in the 25^{th} percentile and inversely in the 75^{th} percentile. However, given that the same pattern occurs in both full- and part-time employment and the relative size of the group to which the extended imputation applies is similarly big in both groups, its effect with regard to the adjusted part-time gap should cancel out - making the results across the distribution robust to the additional assumption made for this group.

Table C.21 shows results on the corrected part-time wage gap if the random conditional rank alternative imputation is used. Results for the 10^{th}, 25^{th} and 75^{th} percentiles are qualitatively equivalent to the main results. The corrected part-time wage gap at the median almost disappears for the time period 2005-2009 (instead of the 8 log points wage premium found in the main results), and the corrected part-time wage gaps at the 90^{th} become much more amplified under this alternative imputation rule.

Table C.21: Robustness check: random alternative imputation

	1990-1994	**1995-1999**	**2000-2004**	**2005-2009**
$\widehat{G}_{corr}$ ($\tau = .10$)	0.19	0.23	-0.17	-0.06
$\widehat{G}_{corr}$ ($\tau = .25$)	0.13	0.22	-0.02	0.02
$\widehat{G}_{corr}$ ($\tau = .50$)	0.04	0.08	-0.03	-0.01
$\widehat{G}_{corr}$ ($\tau = .75$)	-0.09	-0.05	-0.11	-0.08
$\widehat{G}_{corr}$ ($\tau = .90$)	-0.32	-0.28	-0.25	-0.28

Comments: Log-point differences between the relevant inverse cumulative distributions. Unweighted average of multiple imputations. *Statistical significance at the 5% level (computed by bootstrap with 200 replications).

Source: SOEP.v32,own calculations.

C.5 Time-Invariance of the unobservables

At the core of the imputation procedure used in this paper resides the assumption that the distribution of unobservables conditional on the observables are time-invariant. This requires, for example, the distribution of unobservable factors such as innate ability or motivation for work to stay unchanged over time once one conditions on the regressors that enter the conditional wage model (in this case, age, education and working experience) and is the equivalent to the common trend assumption in a usual difference-in-difference setting. Given that in the context of conditional quantile regressions the unobservables determine the conditional rank of an observation, Melly and Santangelo (2015) suggest to test whether the following expression holds:

$$F_{w|t=l,g=0,x}\left(F^{-1}_{w|t=l,g=1,x}(\theta)\right) = F_{w|t=j,g=0,x}\left(F^{-1}_{w|t=j,g=1,x}(\theta)\right) \forall\theta, \forall x \in X \tag{C.2}$$

where $t = \{j, l\}$ are two time periods in which all individuals are in employment, $g = 1$ determines the group of individuals that work in $t = \{j, l\}$ but do not work in $t = k$ and $g = 2$ determines the group of individuals working in $t = \{j, l, k\}$. This requires that, for all individuals (both in the treatment and the control group) used in a two-year imputation, to have data on a further time period in which I observe a wage for all individuals (ie in which all individuals are non-treated). If this is the case, then the idea is simple: make use of the identifying assumption that enables the imputation of non-realised wages for the treatment group in $t = k$ to produce conditional wage distributions of the treatment group for $t = j$ and $t = l$ and compare them to each other.

Unfortunately, for any third year ($t = j$) that is added to any previous two-year imputation $t = \{l, k\}$, there are some individuals from both treatment and control groups for whom one does not observe a wage. While this renders a full-fledged test not possible, the exploration of the identifying assumption in this section should give the reader an impression on how stable is the distribution of unobservables conditional on observables over time.

In order to improve the tractability of the exercise at hand, which requires expression C.2 holding for all possible three-year combinations $\{j, l, k\}$ as well as for all θ and for all values of the covariates, I have carried out the following simplifications: (a) I examine the assumption for four selected $k = 1995, 2000, 2005, 2009$, each for three combinations of $\{j, l\}$; (b) I check the conditional cumulative distributions of expression C.2 for all covariate values of the empirical sample (excluding duplicates) and for the grid of $\theta = .1, .2, .3, .4, .5, .6, .7, .8, .9$; (c) I use a more parsimonius specification of the conditional wage model that only includes years of education and experience (linear and squared) as regressors, in order to minimize problems related to the small size of the regression samples.

The results are summarized in Table C.22 for the full-time distribution and in Table C.23 for the part-time distribution. In particular, these tables tabulate cumulative frequencies by different absolute distances between conditional log wage distributions $\left|F_{w|t=l,g=0,x}(\cdot) - F_{w|t=j,g=0,x}(\cdot)\right|$. In each Table, N stands for all combinations of covariate values of the empirical sample of the chosen years times nine, as each covariate combination is evaluated at each decile of the conditional wage distribution.

Table C.22: Absolute distance between full-time conditional cumulative distributions

	(A) Group not working full-time in 1995			(B) Group not working full-time in 2000		
	1993 vs 1994	1998 vs 1999	2000 vs 2001	2002 vs 2003	2003 vs 2004	2004 vs 2005
$\leq .05$	60%	39%	58%	53%	48%	60%
$\leq .10$	88%	67%	82%	76%	70%	78%
$\leq .20$	99%	95%	99%	94%	89%	93%
N	845	981	1278	1419	1613	1638

	(C) Group not working full-time in 2005			(D) Group not working full-time in 2009		
	2000 vs 2001	2007 vs 2008	2008 vs 2009	2003 vs 2004	2005 vs 2006	2011 vs 2012
$\leq .05$	52%	42%	49%	42%	61%	45%
$\leq .10$	81%	57%	76%	69%	79%	65%
$\leq .20$	98%	80%	94%	95%	94%	87%
N	1782	1284	1476	1251	963	1017

Comments: N stands for all combinations of the empirical values of X and the grid of θ. Conditional wage model specified as $log(wage) = constant + exp + exp^2 + years\,in\,education$.
Source: SOEP.v32, own calculations.

In particular, Table C.22 should be read as follows. If one looks at the first column of Panel A, the conditional cumulative density functions for years 1993 and 1994 are less than 5 (conditional) percentiles away for 60% of all combinations of covariate values and θ. Given that one can observe larger distances between conditional cumulative functions for very few combinations of covariate values and θ, Tables C.22 and C.23 indicate that the distribution of unobservables conditional on observables stays relatively constant over time.

Table C.23: Absolute distance between part-time conditional cumulative distributions

	(A) Group not working part-time in 1995			(B) Group not working part-time in 2000		
	1996 vs 1997	1999 vs 2000	2003 vs 2004	2003 vs 2004	2005 vs 2006	2007 vs 2008
$\leq .05$	39%	39%	56%	54%	58%	41%
$\leq .10$	70%	65%	80%	77%	82%	71%
$\leq .20$	93%	92%	95%	94%	97%	97%
N	924	1462	2071	2504	2833	2673

	(C) Group not working part-time in 2005			(D) Group not working part-time in 2009		
	2002 vs 2003	2007 vs 2008	2009 vs 2010	1992 vs 1993	1992 vs 1993	1992 vs 1993
$\leq .05$	40%	53%	55%	48%	39%	43%
$\leq .10$	61%	76%	81%	73%	69%	75%
$\leq .20$	89%	96%	96%	93%	94%	94%
N	1368	1728	1887	1410	1359	951

Comments: N stands for all combinations of the empirical values of X and the grid of θ. Conditional wage model specified as $log(wage) = constant + exp + exp^2 + years\ in\ education$.
Source: SOEP.v32, own calculations.

Summary

This dissertation analyses two persistent and pervasive labour market phenomena: the gender wage gap and the part-time pay penalty. It provides empirical evidence on the evolution of these over time and across the wage distribution. All analyses focus on Germany and results are based on the German Socio-Economic Panel (GSOEP). The main findings can be summarized in the following three points.

First, my results show that the observed convergence in male and female (full-time) wages from mid 1980s until recently can be mostly explained by increasing positive selection of women into full-time employment. Once selection into employment is controlled for, the median gender wage gap by the end of the 2000s was almost as high as twenty-five years before. Second, the sample of men working full-time is also slightly positively selected, but less so than its female counterpart. However, the magnitude of selection at the lower bound of the distribution has increased steeply from the beginning of the 1990s onwards. Third, my findings suggest that the raw female part-time wage gap - which has been increasing over the last decades - can be best explained by opposite patterns of selection into employment in the full- and part-time sectors. However, the wage structures of full- and part-time employment have become more equal over time.

From a methodological perspective, my dissertation contributes to extend an existing econometric procedure in order to examine the part-time wage gap across the distribution and over time, while controlling for selection into full- and part-time employment.

Two major policy recommendations can be derived from this dissertation. If the goal is to reduce the gender- and part-time wage gap, then policies need to (1) reduce

barriers that hinder female full-time employment and (2) further facilitate the transition between full- and part-time employment in both directions and for all individuals, regardless of skill level and gender.

German Summary

Die vorliegende Dissertation analysiert zwei anhaltende und allgegenwärtige Arbeitsmarktphänomene: die geschlechtsspezifische Lohnlcke und die Lohnlücke bei Teilzeiterwerbstätigkeit. Deren empirische Entwicklung wird sowohl im Zeitverlauf als auch über die Lohnverteilung hinweg aufgezeigt. Alle Analysen legen den Schwerpunkt auf Deutschland und die Ergebnisse basieren auf dem Sozio-Ökonomischen Panel (SOEP). Die Haupterkenntnisse lassen sich in den folgenden drei Punkten zusammenfassen: Erstens zeigen die Ergebnisse, dass die beobachtbare Konvergenz des männlichen und weiblichen (Vollzeit-) Lohnniveaus von Mitte der 1980er Jahre bis heutzutage sich in erster Linie durch die zunehmende positive Selektion von Frauen hinsichtlich Vollzeitbeschäftigung erklären lässt. Kontrolliert man statistisch für diesen Selektionseffekt im Beschftigungsverhältnis, so sind Median-Lohnunterschiede zwischen Männern und Frauen bis Ende der 2000er Jahre fast genauso hoch wie fünfundzwanzig Jahre zuvor. Zweitens ist auch die Stichprobe von vollzeitbeschäftigten Männern leicht positiv selektiert, wenn auch geringfügiger als dies bei Frauen der Fall ist. Dennoch, das Ausmaß der Selektion an der unteren Grenze der Verteilung hat seit Anfang der 1990er Jahre stark zugenommen. Drittens legen die Befunde nahe, dass sich die Lohnlücke von Frauen in Teilzeitarbeit - welche in den vergangenen Jahrzehnten gewachsen ist - am besten durch gegensätzliche Selektionseffekte beim Voll- und Teilzeitbeschäftigungssektor erklären lässt. Zugleich haben sich die Lohnstrukturen von Voll- und Teilzeitbeschäftigten im Laufe der Zeit aber angeglichen. Aus methodischer Sicht leistet die Dissertation einen Beitrag bei der Erweiterung eines bestehenden ökonometrischen Verfahrens zur Untersuchung der Lohnlücke von Teilzeitbeschäftigten über die Lohnverteilung hinweg und

im Zeitverlauf, wobei für Selektionseffekte hinsichtlich Voll- und Teilzeitarbeit kontrolliert werden kann. Zwei wesentliche Politikempfehlungen ergeben sich aus dieser Dissertation. Wenn das Ziel die Reduktion der geschlechtsspezifischen Lohnlücke sowie der Lohnlücke bei Teilzeiterwerbsttigkeit ist, so müssen politische Maßnahmen sich (1) dem Abbau von Hindernissen widmen, die Frauen eine Vollzeitbeschäftigung erschweren, und darber hinaus (2) den Übergang zwischen Voll- und Teilzeitarbeit in beiden Richtungen und für alle Personen, unabhängig vom Qualifikationsniveau und Geschlecht, erleichtern.

Bibliography

Aaberge, R., Dagsvik, J. K., and Strøm, S., 1995. Labor supply responses and welfare effects of tax reforms. *The Scandinavian Journal of Economics*, 97(4):635–659.

Aaronson, D. and French, E., 2004. The effect of part-time work on wages: Evidence from the social security rules. *Journal of Labor Economics*, 22(2):329–352.

Abowd, J. M. and Stinson, M. H., 2013. Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data. *The Review of Economics and Statistics*, 95(5):1451–1467.

Aizer, A., 2010. The gender wage gap and domestic violence. *The American Economic Review*, 100(4):pp. 1847–1859.

Albrecht, J., Björklund, A., and Vroman, S., 2003. Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21(1):145–177.

Alesina, A., Ichino, A., and Karabarbounis, L., 2011. Gender-based taxation and the division of family chores. *American Economic Journal: Economic Policy*, 3(2):1–40.

Altonji, J. G. and Blank, R. M., 1999. Race and gender in the labor market. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, Part C, chapter 48, pages 3143–3259. Elsevier.

Altonji, J. G. and Pierret, C. R., 2001. Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):pp. 313–350.

Anderson, D. J., Binder, M., and Krause, K., 2002. The motherhood wage penalty: Which mothers pay it and why? *The American Economic Review*, 92(2):pp. 354–358.

———, 2003. The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *Industrial and Labor Relations Review*, 56(2):pp. 273–294.

Angelov, N., Johansson, P., and Lindahl, E., 2016. Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579.

Anger, C. and Schmidt, J., 2010. Gender Pay Gap: Gesamtwirtschaftliche Evidenz und regionale Unterschiede. IW Trends 4, IW Kln.

Antonczyk, D., Fitzenberger, B., and Sommerfeld, K., 2010. Rising wage inequality, the decline of collective bargaining, and the gender wage gap. *Labour Economics*, 17(5):835–847.

Arellano, M. and Bonhomme, S., 2017. Quantile selection models with an application to understanding changes in wage inequality. *Econometrica*, 85(1):1–28.

Arulampalam, W., Booth, A. L., and Bryan, M. L., 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review*, 60(2):pp. 163–186.

Athey, S. and Imbens, G. W., 2006. Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.

Averett, S. L. and Hotchkiss, J. L., 1996. Discrimination in the payment of full-time wage premiums. *Industrial and Labor Relations Review*, 49(2):287–301.

Baldwin, M. and Johnson, W. G., 1992. Estimating the employment effects of wage discrimination. *The Review of Economics and Statistics*, 74(3):pp. 446–455.

Bardasi, E. and Gornick, J. C., 2008. Working for less? Women's part-time wage penalties across countries. *Feminist Economics*, 14(1):37–72.

Bargain, O., 2008. Normative evaluation of tax policies: From households to individuals. *Journal of Population Economics*, 21(2):pp. 339–371.

Bargain, O., Orsini, K., and Peichl, A., 2014. Comparing labor supply elasticities in Europe and the US – new results. *Journal of Human Resources*, 49(3):723–838.

Bastani, S., 2013. Gender-based and couple-based taxation. *International Tax and Public Finance*, 20(4):653–686.

Becker, G., 1971. *The Economics of Discrimination.* University of Chicago Press, 2 edition.

Becker, G. S., Hubbard, W. H. J., and Murphy, K. M., 2010. Explaining the worldwide boom in higher education of women. *Journal of Human Capital*, 4(3):203–241.

Bergmann, A. and Mertens, A., 2011. Job stability trends, lay-offs, and transitions to unemployment in west germany. *LABOUR*, 25(4):421–446.

Bertrand, M., 2011. New perspectives on gender. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 4, Part B, pages 1543–1590. Elsevier.

Bertrand, M., Kamenica, E., and Pan, J., 2015. Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614.

Biewen, M., Fitzenberger, B., and de Lazzer, J., 2017. Rising wage inequality in Germany: Increasing heterogeneity and changing selection into full-time work. IZA Discussion Paper No. 11072.

Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J., 2008. Gender wage disparities among the highly educated. *Journal of Human Resources*, 43(3):630–659.

Black, S. E. and Strahan, P. E., 2001. The division of spoils: Rent-sharing and discrimination in a regulated industry. *The American Economic Review*, 91(4):pp. 814–831.

Blank, R. M., 1990. Are part-time jobs lousy jobs. In *A Future of Lousy Jobs? The Brookings Institute, Washington DC.*

Blau, F. and Kahn, L., 2003. Understanding international differences in the gender pay gap. *Journal of Labor Economics*, 21(1):106–144.

———, 2007. Changes in the labor supply behavior of married women: 19802000. *Journal of Labor Economics*, 25(3):393–438.

Blau, F. D. and Kahn, L. M., 1992. The gender earnings gap: Learning from international comparisons. *The American Economic Review*, 82(2):533–538.

———, 1996. Wage structure and gender earnings differentials: An international comparison. *Economica*, 63(250):29–62.

———, 2006. The U.S. gender pay gap in the 1990s: Slowing convergence. *Industrial and Labor Relations Review*, 60(1):45–66.

———, 2013. The feasibility and importance of adding measures of actual experience to cross-sectional data collection. *Journal of Labor Economics*, 31(2):S17–S58.

———, 2017. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Blinder, A. S., 1973. Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, 8(4):pp. 436–455.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J., 2016. Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Blundell, R., Gosling, A., Ichimura, H., and Meghir, C., 2007. Changes in the distribution of male and female wages accounting for employment composition usingbounds. *Econometrica*, 75(2):323–363.

Booth, A. and Francesconi, M., 2003. Union coverage and non-standard work in britain. *Oxford Economic Papers*, 55(3):383–416.

Booth, A. L. and van Ours, J. C., 2008. Job satisfaction and family happiness: The part-time work puzzle. *The Economic Journal*, 118(526):F77–F99.

———, 2013. Part-time jobs: what women want? *Journal of Population Economics*, 26(1):263–283.

Booth, A. L. and Wood, M., 2008. Back-to-front down under? Part-time/full-time wage differentials in Australia. *Industrial Relations: A Journal of Economy and Society*, 47(1):114–135.

Borowczyk-Martins, D. and Lalé, E., 2016. Employment adjustment and part-time jobs: The U.S. and the U.K. in the great recession. IZA Discussion Paper No. 9847.

Bowlus, A. J. and Grogan, L., 2009. Gender wage differentials, job search, and part-time employment in the UK. *Oxford Economic Papers*, 61(2):275–303.

Browning, M., Bourguignon, F., Chiappori, P.-A., and Lechene, V., 1994. Income and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy*, 102(6):pp. 1067–1096.

Buchinsky, M., 1998. The dynamics of changes in the female wage distribution in the USA: A quantile regression approach. *Journal of Applied Econometrics*, 13(1):1–30.

Busch, A. and Holst, E., 2008. "Gender pay gap": in Grostdten geringer als auf dem Land. *DIW Wochenbericht*, 75(33):462–468.

Campos-Vázquez, R. M., Hincapié, A., and Rojas-Valdés, R. I., 2012. Family income inequality and the role of married females' earnings in Mexico: 1988-2010. *Latin American Journal of Economics*, 49(1):67–98.

Cancian, M. and Reed, D., 1998. Assessing the effects of wives' earnings on family income inequality. *Review of Economics and Statistics*, 80(1):73–79.

———, 1999. The impact of wives' earnings on income inequality: Issues and estimates. *Demography*, 36(2):173–184.

Card, D., Cardoso, A. R., and Kline, P., 2016. Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633–686.

Card, D., Heining, J., and Kline, P., 2013. Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Chernozhukov, V., FernndezVal, I., and Melly, B., 2013. Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.

Chernozhukov, V. and Hansen, C., 2005. An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Christofides, L. N., Polycarpou, A., and Vrachimis, K., 2013. Gender wage gaps, sticky floors and glass ceilings in Europe. *Labour Economics*, 21(C):86–102.

Connolly, S. and Gregory, M., 2009. The part-time pay penalty: earnings trajectories of British women. *Oxford Economic Papers*, 61(suppl 1):i76–i97.

———, 2010. Dual tracks: part-time work in life-cycle employment for British women. *Journal of Population Economics*, 23(3):907–931.

Crossley, T. F. and Jeon, S.-H., 2007. Joint taxation and the labour supply of married women: Evidence from the Canadian tax reform of 1988. *Fiscal Studies*, 28(3):343–365.

Del Boca, D. and Pasqua, S., 2003. Employment patterns of husbands and wives and family income distribution in italy (1977–98). *Review of Income and Wealth*, 49(2):221–245.

Del Boca, D. and Vuri, D., 2005. Labor supply and child care costs: The effect of rationing. IZA Discussion Papers 1779, Institute for the Study of Labor (IZA).

DiNardo, J., Fortin, N. M., and Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5):1001–1044.

Dustmann, C., Fitzenberger, B., Schnberg, U., and Spitz-Oener, A., 2014. From sick man of europe to economic superstar: Germany's resurgent economy. *The Journal of Economic Perspectives*, 28(1):167–188.

Dustmann, C., Ludsteck, J., and Schnberg, U., 2009. Revisiting the german wage structure. *The Quarterly Journal of Economics*, 124(2):843–881.

Ermisch, J., Francesconi, M., and Siedler, T., 2006. Intergenerational mobility and marital sorting. *The Economic Journal*, 116(513):659–679.

Ermisch, J. F. and Wright, R. E., 1993. Wage offers and full-time and part-time employment by british women. *The Journal of Human Resources*, 28(1):111–133.

Euwals, R. and van Soest, A., 1999. Desired and actual labour supply of unmarried men and women in the netherlands. *Labour Economics*, 6(1):95–118.

Figari, F., Immervoll, H., Levy, H., and Sutherland, H., 2011. Inequalities within couples in Europe: Market incomes and the role of taxes and benefits. *Eastern Economic Journal*, 37(3):344–366.

Finke, C., 2010. Verdienstunterschiede zwischen mnner und frauen.

Fitzenberger, B. and Kunze, A., 2005. Vocational training and gender: Wages and occupational mobility among young workers. *Oxford Review of Economic Policy*, 21(3):392–415.

Fortin, N., Lemieux, T., and Firpo, S., 2011. Decomposition methods in economics. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 4, Part A of *Handbook of Labor Economics*, pages 1–102. Elsevier.

Fortin, N. M., 2008. The gender wage gap among young adults in the United States: The importance of money versus people. *The Journal of Human Resources*, 43(4):884–918.

Fortin, N. M. and Lemieux, T., 1998. Rank regressions, wage distributions, and the gender gap. *The Journal of Human Resources*, 33(3):610–643.

Fouarge, D. and Muffels, R., 2009. Working Part-Time in The British, German and Dutch Labour Market: Scarring for the Wage Career? *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift f?r Wirtschafts- und Sozialwissenschaften*, 129(2):217–226.

Fuchs-Schündeln, N., Krueger, D., and Sommer, M., 2010. Inequality trends for Germany in the last two decades: A tale of two countries. *Review of Economic Dynamics*, 13(1):103 – 132. Special issue: Cross-Sectional Facts for Macroeconomists.

Garnero, A., Kampelmann, S., and Rycx, F., 2014. Part-time work, wages, and productivity. *Industrial and Labor Relations Review*, 67(3):926–954.

Goldin, C., 2014. A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Goldin, C., Katz, L. F., and Kuziemko, I., 2006. The homecoming of american college women: The reversal of the college gender gap. *The Journal of Economic Perspectives*, 20(4):133–156.

Gronau, R., 1974. Wage comparisons-a selectivity bias. *Journal of Political Economy*, 82(6):1119–43.

Groshen, E. L., 1991. The structure of the Female/Male wage differential: Is it who you are, what you do, or where you work? *Journal of Human Resources*, 26(3):457–472.

Gupta, N. D., Oaxaca, R., and Smith, N., 2006. Swimming upstream, floating downstream: Comparing women's relative wage progress in the united states and denmark. *ILR Review*, 59(2):243–266.

Haywood, L. and Neumann, M., 2017. The role of aggregate preferences for labor supply evidence from low-paid employment. DIW Discussion Paper No 1652.

Heckman, J. J., 1979. Sample selection bias as a specification error. *Econometrica*, 47(1):pp. 153–161.

Heinze, A. and Wolf, E., 2010. The intra-firm gender wage gap: a new view on wage differentials based on linked employer-employee data. *Journal of Population Economics*, 23(3):851–879.

Hellerstein, J. K., Neumark, D., and Troske, K. R., 1999. Wages, productivity, and worker characteristics: Evidence from plant level production functions and wage equations. *Journal of Labor Economics*, 17(3):409–446.

———, 2002. Market forces and sex discrimination. *The Journal of Human Resources*, 37(2):353–380.

Hersch, J. and Stratton, L., 2002. Housework and wages. *Journal of Human Resources*, 37(1):217–229.

Hijzen, A. and Martin, S., 2013. The role of short-time work schemes during the global financial crisis and early recovery: a cross-country analysis. *IZA Journal of Labor Policy*, 2(1):5.

Hijzen, A. and Venn, D., 2011. The role of short-time work schemes during the 2008-09 recession. OECD Social, Employment and Migration Working Papers, No. 115, OECD Publishing, Paris.

Hinz, T. and Gartner, H., 2005. The gender wage gap within economic sectors, occupations and firms. *Zeitschrift für Soziologie*, 34:22–39.

Hirsch, B. T., 2005. Why do part-time workers earn less? The role of worker and job skills. *Industrial and Labor Relations Review*, 58(4):525–551.

Huber, M., 2015. Causal pitfalls in the decomposition of wage gaps. *Journal of Business & Economic Statistics*, 33(2):179–191.

Huber, M. and Melly, B., 2012. A test of the conditional independence assumption in sample selection models. Working Papers 2012-11, Brown University, Department of Economics.

Hübler, O., 2005. Is there a varying unexplained gender wage gap in Germany? *Applied Economics Quarterly*, 51(1):29–48.

Immervoll, H., Kleven, H. J., Kreiner, C. T., and Verdelin, N., 2011. Optimal tax and transfer programs for couples with extensive labor supply responses. *Journal of Public Economics*, 95(11-12):1485–1500.

Jacobsen, J., Khamis, M., and Yuksel, M., 2015. *Gender Convergence in the Labor Market*, volume 41, chapter Convergences in Mens and Womens Life Patterns: Lifetime Work, Lifetime Earnings, and Human Capital Investment, pages 1–33. Emerald

Group Publishing Limited. Solomon W. Polachek , Konstantinos Tatsiramos , Klaus F. Zimmermann (eds).

Juhn, C. and McCue, K., 2017. Specialization then and now: Marriage, children, and the gender earnings gap across cohorts. *Journal of Economic Perspectives*, 31(1):183–204.

Kapteyn, A. and Ypma, J., 2007. Measurement error and misclassification: A comparison of survey and administrative data. *Journal of Labor Economics*, 25(3):513–551.

Kassenböhmer, S. C. and Sinning, M., 2014. Distributional changes in the gender wage gap. *ILR Review*, 67(2).

Koenker, R. and Bassett, G. J., 1978. Regression quantiles. *Econometrica*, 46(1):33–50.

Künn-Nelen, A., de Grip, A., and Fouarge, D., 2013. Is part-time employment beneficial for firm productivity? *Industrial and Labor Relations Review*, 66(5):1172–1191.

López-Nicolás, A., Garcia Villar, J., and Hernandez, P., 2001. How wide is the gap? An investigation of gender wage differences using quantile regression. *Empirical Economics*, 26(1):149–167.

Ludsteck, J., 2014. The impact of segregation and sorting on the gender wage gap: Evidence from german linked longitudinal employer-employee data. *ILR Review*, 67(2):362–394.

Lundberg, S. J., Pollak, R. A., and Wales, T. J., 1997. Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *The Journal of Human Resources*, 32(3):463–480.

Machado, J. A. F. and Mata, J., 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4):445–465.

Manning, A. and Petrongolo, B., 2008. The part-time pay penalty for women in Britain. *The Economic Journal*, 118(526):F28–F51.

Matteazzi, E., Pailh, A., and Solaz, A., 2014. Part-time wage penalties for women in prime age. *Industrial and Labor Relations Review*, 67(3):955–985.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Economics*. Academic Press.

Melly, B., 2005. Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4):577–590.

———, 2006. Estimation of counterfactual distributions using quantile regression. Technical report.

Melly, B. and Santangelo, G., 2014. The evolution of the gender wage gap: 1968-2008. Working Paper.

———, 2015. The changes-in-changes model with covariates. Working Paper.

Meurs, D., Pailh, A., and Ponthieux, S., 2010. Child-related career interruptions and the gender wage gap in France. *Annals of Economics and Statistics*, (99/100):pp. 15–46.

Mocan, H. N. and Tekin, E., 2003. Nonprofit sector and part-time work: An analysis of employer-employee matched data on child care workers. *The Review of Economics and Statistics*, 85(1):38–50.

Müller, K.-U., Neumann, M., and Wrohlich, K., 2017. Labour supply under participation and hours constraints. Working Paper.

Mulligan, C. B. and Rubinstein, Y., 2008. Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3):1061–1110.

Mumford, K. and Smith, P. N., 2009. What determines the part-time and gender earnings gaps in Britain: Evidence from the workplace. *Oxford Economic Papers*, 61:i56–i75.

Neal, D., 2004. The measured black-white wage gap among women is too small. *Journal of Political Economy*, 112(S1):S1–S28.

Neumark, D., 1988. Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, 23(3):279–295.

Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Oaxaca, R. L. and Ransom, M. R., 1994. On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61(1):5–21.

OECD, 2010. *Employment Outlook 2010*, chapter How Good is Part-Time Work?, pages 211–266. OECD Publishing, Paris.

———, 2012. *Closing the Gender Gap: Act Now*. OECD Publishing.

———, 2016. Employment outlook 2016. Technical report, Organisation for Economic Co-operation and Development.

———, 2017a. Labour force statistics: Incidence of involuntary part time workers. Accessed on 13 Dec 2017.

———, 2017b. *The Pursuit of Gender Equality: An Uphill Battle*. OECD Publishing.

Olivetti, C. and Petrongolo, B., 2008. Unequal pay or unequal employment? A cross country analysis of gender gaps. *Journal of Labor Economics*, 26(4):621–654.

———, 2017. The economic consequences of family policies: Lessons from a century of legislation in high-income countries. *The Journal of Economic Perspectives*, 31(1):205–230.

Pasqua, S., 2002. Wives' work and income distribution in European countries. *The European Journal of Comparative Economics*, 5(2):197–226.

Paul, M., 2016. Is there a causal effect of working part-time on current and future wages? *The Scandinavian Journal of Economics*, 118(3):494–523.

Rica, S. d. l., Dolado, J. J., and Llorens, V., 2008. Ceilings or floors? Gender wage gaps by education in Spain. *Journal of Population Economics*, 21(3):751–778.

Roy, A., 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

van Soest, A., 1995. Structural models of family labor supply: A discrete approach. *The Journal of Human Resources*, 30(1):63–88.

Steiner, V. and Wrohlich, K., 2004. Household taxation, income splitting and labor supply incentives: A microsimulation study for Germany. *CESifo Economic Studies*, 50(3):541–568.

Steiner, V., Wrohlich, K., Haan, P., and Geyer, J., 2012. Documentation of the tax-benefit microsimulation model stsm: Version 2012. Data Documentation 63, DIW Berlin, German Institute for Economic Research.

Stotsky, J. G., 1996. Gender bias in tax systems. IMF Working Paper 96/99.

Sutherland, H., 1997. Women, men and the redistribution of income. *Fiscal Studies*, 18(1):1–22.

Van Soest, A., 1995. Structural models of family labor supply: a discrete choice approach. *Journal of Human Resources*, pages 63–88.

Wagner, G. G., Frick, J. R., and Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch*, 127(1):139–170.

Wolf, E., 2002. Lower wage rates for fewer hours? a simultaneous wage-hours model for germany. *Labour Economics*, 9(5):643 – 663.

Wrohlich, K., 2011. Labor supply and child care choices in a rationed child care market. Discussion Papers of DIW Berlin 1169, DIW Berlin, German Institute for Economic Research.

Wrohlich, K. and Zucco, A., 2017. The gender pay gap varies greatly by occupation.

Declaration

Erklärung gem. 4 Abs. 2 der Promotionsordnung

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, Januar 2018

Patricia Gallego Granados

Erklärung gem. 10 Abs. 3 der Promotionsordnung

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe: Stata und MS Excel. Auf dieser Grundlage habe ich die Arbeit selbständig verfasst.

Berlin, Januar 2018

Patricia Gallego Granados