I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11634

The Role of Hours Changes for the Increase in German Earnings Inequality

Martin Biewen
Daniela Plötze

JUNE 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11634

The Role of Hours Changes for the Increase in German Earnings Inequality

Martin Biewen
University of Tübingen, LEAD Tübingen and IZA

Daniela Plötze
Wüstenrot & Württembergische and University of Tübingen

JUNE 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Role of Hours Changes for the Increase in German Earnings Inequality[1]

Using data from the German Structure of Earnings Survey (GSES), this paper studies the role of changes in working hours for the increase in male and female earnings inequality between 2001 and 2010. We provide both classic decompositions of the variance of log earnings into the variances of hours, wage rates and their covariance, and decompositions based on reweighting the conditional hours distribution. Depending on the inequality measure considered, our results suggest that between 10 and 30 percent of the increase in male earnings inequality and 37 to 47 percent of the increase in female earnings inequality can be explained by changes in working hours. In addition, a large part of the inequality increase can be accounted for by changes in the composition of person and firm characteristics.

JEL Classification: C14, J22, J31

Keywords: inequality, working hours, earnings, female labor market participation

Corresponding author:
Martin Biewen
School of Business and Economics
University of Tübingen
Mohlstraße 36
72074 Tübingen
Germany
E-mail: martin.biewen@uni-tuebingen.de

[1] This paper uses the Scientific Use Files (SUF) of the German Structure of Earnings Surveys (GSES) 2001, 2006 and 2010 provided by the German Federal Statistical Office. We thank Kristin Nowak (Statistisches Landesamt Stuttgart) for her support. We also thank Matthias Seckler, Nicolas Herault and participants of ESPE 2018 for helpful comments.

1 Introduction

The trend towards rising wage inequality observed for many countries around the world has received a lot of attention (Katz and Autor, 1999, Machin, 2008, OECD, 2011). The literature has considered a large number of potential reasons for the observed trends such as skill-biased technological progress, changes in supply and demand, institutional changes, changes in tasks, and increasing differences between firms (Juhn et al., 1993, DiNardo et al., 1996, Autor et al., 2008, Dustmann et al, 2009, Card et al., 2013). A factor that appears to have been rarely considered in these analyses is changes in working hours and working hours arrangements. It is clear that the number of working hours represent an important component of aggregated earnings such as monthly or yearly wage incomes. Consequently, hours changes may have a substantial impact on the shape and the level of inequality observed for such income measures.

A limited number of studies have explicitly considered the relationship between working times and inequality in wage outcomes. For example, Doiron and Barret (1996) examine the role of hours changes for changes in male and female earnings inequality in Canada. They find that hours changes play an important role for changes in earnings inequality over time, in addition to distinctive gender differences in how working hours influence the distribution of earnings. Juhn et al. (1993) touch the question of differences in working times in their general analysis of changes in the US wage distribution. Johnson and Kuhn (2004) provide another study of the role of hours changes for increasing wage inequality in Canada. For Germany, Fuchs-Schündeln et al. (2010) document changes in hourly wages, monthly wages and monthly working hours over time. In a cross-country setting, Blau and Kahn (2011) analyse changes in earnings inequality around the world and find that the distribution of working hours substantially influences the earnings distribution in many countries, albeit to varying degrees. Finally, Checchi et al. (2016) provide a cross-country analysis of hours changes and earnings inequality for the US, the UK, Germany

and France. Their findings also suggest that hours changes may have substantially contributed to the increases in earnings inequality observed for these countries.

In this study, we want to contribute to the above literature in two ways. First, we use a high-quality German data set that has not been used for this purpose before, the German Structure of Earnings Survey (GSES), a mandatory employer-employee survey conducted by the German Federal Statistical Office. Previous studies on wage inequality in Germany were either based on administrative data (Dustmann, 2009, Card et al., 2013) which do not contain information on working hours, or on data from the German Socio-Economic Panel (Fuchs-Schündeln, 2010, Checchi et al., 2016) which have a much smaller sample size and which may differ in its information content due to its non-mandatory character. As a second aspect, we provide an alternative to the classic decomposition of changes in the variance of log earnings into the changes of the variance of log hours, the variance of log hourly wages, and their covariance. Our decomposition considers changes in the *conditional* distribution of working hours, i.e. changes in working hours for given groups of individuals. The results of our analysis corroborate earlier findings that hours changes may have a substantial effect on the earnings distribution, especially for women. In addition, we find that considering conditional rather than unconditional hours changes leads in some cases to a smaller role of hours changes for the earnings distribution but in other cases to a similarly important one as in the classic variance decomposition.

This paper is structured as follows. In section 2, we describe the methods used by us to separate the role of hours changes on earnings inequality. In section 3, we describe the data on which our study is based. Section 4 presents and discusses our empirical results. Section 5 concludes.

2 Methods

2.1 Variance decomposition

As a classic method to describe the relationship between hours worked and the distribution of monthly earnings, we report the often used decomposition of the variance of log earnings (e.g., Burtless, 1990, Juhn et al., 1993, Johnson and Kuhn, 2004, Blau and Kahn, 2011). The decomposition based on $e = h \cdot w$ is given by

$$var(\log(e)) = var(\log(h)) + var(\log(w)) + 2cov(\log(h), \log(w)), \tag{1}$$

where e denotes monthly earnings ('monthly wage'), h the monthly hours worked, and w the hourly wage rate. The decomposition breaks down the variability of log earnings into the variability of log hours, the variability of the log hourly wage, and the covariation of both. If one analyzes changes over time, the decomposition becomes

$$\Delta var(\log(e)) = \Delta var(\log(h)) + \Delta var(\log(w)) + 2\Delta cov(\log(h), \log(w)), \tag{2}$$

where $\Delta var(\log(e)) = var(\log(e_1)) - var(\log(e_0))$ etc. describes the changes of the respective variables between a period $t = 1$ compared to a base period $t = 0$. The latter decomposition splits the overall change of earnings inequality (as measured by the variance of logs) into a contribution coming from the changes in the dispersion of working hours, changes in the dispersion of hourly wages and changes in the correlation of working hours and hourly wages.

The variance decomposition above seems restrictive for several reasons. First, it is confined to a very specific inequality measure, the variance of logs. Some authors have dismissed the variance of logs as a proper inequality measure as it violates the principle of transfers and Lorenz dominance (Foster and Ok, 1999). The variance decomposition also assumes that the impact of changes in working hours specifically work through changes in the variance of log working

hours. It is unclear to what extent a measure like the variance of log working hours is capable of picking up complex changes in the often very discrete hours distribution (an example of a changing distribution of working hours will be given below, see figure 2). Generalizations of the decomposition to other inequality measures are possible but they have similar limitations, and they have rarely been used in the literature (e.g. Doiron and Barret, 1996, Checchi et al., 2016). A second potential limitation of the decomposition above is that it focusses on changes in the *unconditional* distribution of working hours and does not consider the influence of other factors. It seems natural to think of 'changes in working hours' as changes in working hours for given groups of workers. For example, if working hours arrangements in the different sectors of the economy remain unchanged, but some sectors grow and others shrink, one may want to attribute the resulting distributional changes to sectoral change rather than to changes in working hours. In comparison, the variance decomposition provides a comparatively crude way to assess the role of working hours that does not allow one to separate it from other variables that influence the earnings distribution. We do not claim that the variance decomposition is wrong or uninformative, but it seems useful to complement it with an alternative perspective on hours changes that includes the aspects described above.

2.2 Decomposition based on reweighting

As an alternative, we compute a decomposition based on the reweighting method suggested by DiNardo/Fortin/Lemieux (DFL, 1996). The advantage of this decomposition is that it can incorporate covariates, i.e. determinants of earnings other than working hours, and that it is not confined to a specific inequality measure. For the decomposition, note that the unconditional

distribution of monthly earnings e in a given year $t = 0$ can be written as

$$f_{0000}(e) = \int_h \int_{x_f} \int_{x_p} f(e|h, x_f, x_p, t = 0) dF(h|x_f, x_p, t = 0) dF(x_f|x_p, t = 0) dF(x_p|t = 0) \tag{3}$$

where x_p are worker characteristics (age, education, occupation etc.) and x_f firm characteristics (firm size, industry, union coverage etc.). The ultimate goal of the decomposition is to see how changes in $h|x_p, x_f$, i.e. changes in working hours for groups of workers with characteristics x_p, x_f change the unconditional distribution of monthly earnings e.

Using the DLF reweighting method, one can compute all kinds of 'counterfactual' earnings distributions, in which one or several of the distributions $F(h|x_f, x_p)$, $F(x_f|x_p)$ and $F(x_p)$ are changed from their level in $t = 0$ to that in $t = 1$. For example, changing only the distribution of $h|x_f, x_p$ from its level in $t = 0$ to that in $t = 1$ is achieved by

$$\begin{aligned} f_{0100}(e) &= \int \int \int f(e|h, x_f, x_p, t = 0) dF(h|x_f, x_p, t = 1) dF(x_f|x_p, t = 0) dF(x_p|t = 0) \\ &= \int \int \int f(e|h, x_f, x_p, t = 0) \Psi(h) dF(h|x_f, x_p, t = 0) dF(x_f|x_p, t = 0) dF(x_p|t = 0) \end{aligned} \tag{4}$$

with

$$\Psi(h) = \frac{dF(h|x_f, x_p, t = 1)}{dF(h|x_f, x_p, t = 0)} = \frac{P(t = 1|h, x_f, x_p) P(t = 0|h, x_f, x_p)}{P(t = 0|h, x_f, x_p) P(t = 1|x_f, x_p)}. \tag{5}$$

Similarly, one can change only the distribution of $x_f|x_p$ by

$$f_{0010}(e) = \int \int \int f(e|h, x_f, x_p, t = 0) dF(h|x_f, x_p, t = 0) \Psi(x_f) dF(x_f|x_p, t = 0) dF(x_p|t = 0), \tag{6}$$

or that of x_p by

$$f_{0001}(e) = \int \int \int f(e|h, x_f, x_p, t = 0) dF(h|x_f, x_p, t = 0) dF(x_f|x_p, t = 0) \Psi(x_p) dF(x_p|t = 0) \tag{7}$$

with reweighting factors

$$\Psi(x_f) = \frac{dF(x_f|x_p, t=1)}{dF(x_f|x_p, t=0)} = \frac{P(t=1|x_f, x_p)P(t=0|x_p)}{P(t=0|x_f, x_p)P(t=1|x_p)},$$
$$\Psi(x_p) = \frac{dF(x_p|t=1)}{dF(x_p|t=0)} = \frac{P(t=1|x_p)P(t=0)}{P(t=0|x_p)P(t=1)}.$$

One can also change two distributions simultaneously, e.g.

$$f_{0101}(e) = \int\int\int f(e|h, x_f, x_p, t=0)\Psi(h)dF(h|x_f, x_p, t=0)dF(x_f|x_p, t=0)\Psi(x_p)dF(x_p|t=0). \tag{8}$$

The reweighting factors $\Psi(h), \Psi(x_f), \Psi(x_p)$ can be estimated using logit models for expressions like $P(t=1|h, x_f, x_p)$. Note that in $f(e|h, x_f, x_p)$, the relationship between earnings e on the one hand, and hours h, firm characteristics x_f and personal characteristics x_p on the other is left completely unrestricted. This means in particular that the influence of working hours on earnings is modeled completely general and non-parametric.

Let $\Delta = I(f_{1111}) - I(f_{0000})$ be the total change of a distributional measure $I(\cdot)$ from $t=0$ to $t=1$. In order to measure the contribution of changes in $h|x_f, x_p$, changes in $x_f|x_p$ and changes in x_p on Δ, one uses a telescopic expression like

$$\begin{aligned} I(f_{1111}) - I(f_{0000}) = \quad & (I(f_{0001}) - I(f_{0000})) \\ & + (I(f_{0011}) - I(f_{0001})) \\ & + (I(f_{0111}) - I(f_{0011})) \\ & + (I(f_{1111}) - I(f_{0111})), \end{aligned} \tag{9}$$

in which the first line describes the contribution of changes in x_p, the second line contributions of changes in $x_f|x_p$ and the third line the contribution of changes in $h|x_f, x_p$. The fourth line represents all other (=residual) contributions to the overall change $\Delta = I(f_{1111}) - I(f_{0000})$.

The sequence in which the three factors $h|x_f, x_p$, $x_f|x_p$ and x_p are introduced into decomposition

(9) is only one out of $3! = 3 \cdot 2 \cdot 1$ possibilities. For example, instead of introducing changes in personal characteristics x_p first, one could first introduce changes in firm characteristics $x_f|x_p$ and then changes in personal characteristics x_p etc. In order to make the decomposition more robust to such choices, we use the concept of the Shapley value decomposition (Shorrocks, 2013) which averages the different contributions across all possible sequences.

3 Data

Our analysis is based on data from the scientific use file of the *German Structure of Earnings Surveys (GSES)* for the years 2001, 2006 and 2010 provided by the German Federal Statistical Office. The GSES is a linked employer-employee data set for which firms with at least ten employees are asked to provide information on wages and salaries of their workers. Compared to German administrative wage data, the advantage of the GSES data is the availability of information on working hours, without which our analysis would not be possible. The advantage of the GSES data over the hourly wage data in the *German Socio-Economic Panel (GSOEP)* is the much larger sample size and the fact that participation in the GSES is compulsory under German law.[2] A considerable disadvantage of the GSES is that its coverage across the sectors of the economy was incomplete in the past. In particular, there was limited coverage of parts of the service sector in earlier waves. In order to make our samples comparable across time, we restrict our analysis to the set of sectors of the economy listed in the lower part of table A1.

The key variable of our analysis is monthly gross labor earnings reported for October of the respective year. We adjust all wage information to 2010 prices. There is some minor censoring of wages at the very top whose degree varies over years. We therefore apply a time-consistent

[2]The relatively small sample size of the GSOEP tends to make estimates of inequality trends over time quite noisy, see e.g. Biewen et al. (2017a).

inflation-adjusted censoring threshold of 12.878 Euros across all waves considered. Our variable of working hours includes regular and overtime working hours. Our measure of hourly wage is obtained by dividing monthly earnings by the number of hours worked. For plausibility reasons, we censor hourly wages at a lower threshold of three euros per hour, and we exclude observations with working hours of less than 10 and more than 360 per month.

In addition to our main dependent variables, our data contains a number of firm and personal characteristics as listed in table A1. As personal characteristics we consider age (7 age brackets), education (5 categories), tenure (4 categories) and occupation (52 categories). We restrict our analysis to individuals aged between 25 and 60 years. Our firm characteristics include region (5 categories), industry (12 categories), a dummy for (at least partial) public ownership, firm size (3 categories) and union coverage at the firm level (a dummy indicating whether the firm used a sectoral or firm level union agreement to pay their employees). The total number of observations used in our analysis ranges between 333,155 (2001) and 428,265 (2010) for men, and between 150,339 (2001) and 189,076 (2010) for women.

4 Empirical results

4.1 Changes in earnings, hours and hourly wages over time

We first give an overview of how earnings, hours and hourly wages changed across the period considered by us. Figures 1 to 3 show the development of the distribution of monthly earnings, monthly working hours and hourly wages over the years 2001, 2006 and 2010 for men. The distribution of monthly earnings as shown in figure 1 becomes more spread out between 2001 and 2010, implying increasing earnings inequality. This is confirmed in table 1 showing inequality measures for monthly earnings over the years 2001 to 2010.

— Figures 1 to 3 about here —

Among the men in our sample, only few work part-time, i.e. less than 140 hours per month. The small peak in the left-hand part of figure 1 represents men working marginal part-time. The marginal part-time regulation was such that the typical value of marginal part-time earnings were 325 Euros before 2003 and 400 Euros after 2003 (so-called 'mini-jobs', see the shift of the left-hand peak in the figure). Overall, as shown in figure 2, the change in male working hours between 2001 and 2010 was complex, with some shifts towards working hours of more than 170 hours per month but also some shifts towards working hours of less than 170 hours per month. As a consequence, mean working hours slightly declined, while there was a (moderate) increase in dispersion (table 2). Changes in the distribution of male hourly wages are shown in figure 3 and table 3.

— Tables 1 to 3 about here —

For women, we observe more pronounced changes in the distribution of monthly earnings (figure 4) leading to more extreme inequality increases compared to men (table 1). Moreover, there was stronger polarization of working hours for women than for men (figure 5 and table 2). Working hours below 75 hours (marginal part-time) gained, while working hours above 125 hours (part-time) and above 170 hours (full-time) gained. As evident from table 2, rising inequality in working hours was particularly strong at the lower end of the working hours distribution (see the 50-10 percentile ratio in table 2). Moreover, for women the distribution of hourly wages was much more distinct from the distribution of monthly earnings compared to the male case, suggesting a bigger role for changes in working hours on the distribution of monthly earnings (figure 4 vs. 6).

— Figures 4 to 6 about here —

4.2 Variance decomposition

In this section, we report the results for the conventional variance decomposition of log earnings. For men, table 4 shows that over the different years considered around 70 percent of the variance in earnings can be accounted for by differences in hourly wages. Differences in hours worked account for only around 25 percent, leaving some 5 percent for the covariance of hourly wages and hours worked. Looking at the overall *change* in log earnings variance between 2001 and 2010, around 48 percent can be accounted for by changes in the variance of log hourly wages, 37 percent by changes in the variance of log working hours and around 15 percent by changes in their covariance (last row of table 4).

— Tables 4 and 5 about here —

For women, much less of the variance in monthly earnings is accounted for by differences in hourly wages (around 30 percent), leaving a bigger role for differences in working hours (around 47 percent) and the covariation of hourly wages and hours worked (around 23 percent), see table 5. As for the *change* in the earnings variance between 2001 and 2010, the decomposition suggests that over 45 percent can be accounted for by changes in working hours, 36 percent by changes in hourly wages and 19 percent by changes in their covariance.

4.3 Decomposition based on reweighting

Table 6 shows the results our decomposition based on reweighting for the case of men. It turns out that the factors considered do not contribute much to the increase in mean and median earnings between 2001 and 2010. The contribution of changes in firm characteristics $\Delta x_f|x_p$ explain some 19 (12) percent of the change in mean (median) earnings, but their effect is partly undone by

changes in employee characteristics Δx_p and changes in working hours arrangements $h|x_f, x_p$. This leaves 92 (99) percent of the overall change in mean (median) earnings unexplained (see first six rows of table 6). This is the likely result of pure wage growth unrelated to compositional changes.

— Table 6 about here —

The results for inequality measures suggest that both changes in employee characteristics Δx_p and changes in employees' firm characteristics $\Delta x_f|x_p$ played an important role for the overall increase in male earnings inequality between 2001 and 2010. For example, for the case of the Gini each of these factors explains around 41 to 42 percent, leaving some 18 percent to changes in working hours $h|x_f, x_p$. The contribution of changes in working hours tends to be particularly high for inequality measures that focus on the lower part of the distribution (MLD, Theil, p5010). Overall, depending on the inequality measure chosen, the decomposition suggests that a modest 10 to 30 percent of the overall increase in male earnings inequality can be accounted for by changes in conditional working hours $h|x_f, x_p$. This is a lower figure than in the case of the variance decomposition.

— Table 7 about here —

The results for women are shown in table 7. It turns out that changes in working hours $h|x_f, x_p$ worked towards lower mean (median) earnings (first six rows of table 7) which is consistent with the decline in mean (median) hours worked in table 5. Turning to inequality measures, changes in employee characteristics x_p and firm characteristics $x_f|x_p$ remain as substantial as in the case of men, but changes in conditional working hours $h|x_f, x_p$ become equally important for most inequality measures. Their contributions to the overall increase in earnings inequality range from

37 to 47 percent for the Gini, MLD and Theil. With 55 and 60 percent, they are particularly high for the 90-10 percentile ratio and at the 50-10 percentile ratio, i.e. at the lower end of the earnings distribution. Note that, in the case of women, the results from the decomposition based on reweighting assigned an equally strong role to changes in working hours as the variance decomposition.

4.4 Determinants of changes in working hours

In order to shed more light on the potential determinants of changes in the distribution of working hours, we carry out Oaxaca-Blinder decompositions[3] for mean hours as well as for the within-group and the between-group variance of working hours for $t = 1$ (=2010) vs. $t = 0$ (=2001). For $x_t = (x_f, x_p)_t$ and mean working hours $E[h_t] = E[E(h_t|x_t)] = E[x_t\beta_t]$, we have

$$E[h_1] - E[h_0] = [E(x_1) - E(x_0)]\,\beta_1 + [\beta_1 - \beta_0]\,E(x_0). \tag{10}$$

The decomposition describes whether level shifts in working hours between 2010 and 2001 can be explained by shifts in the composition of worker characteristics or by the changes in the mean working hours given worker characteristics. The results shown in table 8 suggest that for men, neither shifts in characteristics nor changes in hours given characteristics led to important changes in working hours between 2001 and 2010. Rather there was a general unexplained trend towards lower working hours (see the result of -3.838 hours for the constant) which was partly offset by smaller changes in characteristics and coefficients (leading to an overall decline in working hours by -1.958, see third row of table 8). For women, we observe a slightly larger general decline in working hours of -3.593 per month. Again, most of this decline was not associated with changes in characteristics or coefficients. Rather, there was an even more pronounced general

[3]See Jann (2008). We use the version of the decomposition in which the contributions for groups of categorial dummies (e.g. age groups) are normalized to be independent of the reference category.

tendency towards lower working hours (see result for the constant of -7.657) which was partly offset by shifts towards industries with higher working hours, and towards higher working hours arrangements in some occupations.

— Table 8 about here —

For the within-group variance of working hours $E\left[E((h_t - E(h_t|x_t))^2|x_t)\right] = E\left[var(h_t|x_t)\right] = E\left[x_t\gamma_t\right]$, we have

$$E\left[var(h_1|x_1)\right] - E\left[var(h_0|x_0)\right] = \left[E(x_1) - E(x_0)\right]\gamma_1 + \left[\gamma_1 - \gamma_0\right]E(x_0) \tag{11}$$

(i.e. an Oaxaca-Blinder decomposition with dependent variable $(h_t - E(h_t|x_t))^2 = (h_t - x_t\beta_t)^2$). The decomposition describes whether changes in the hours dispersion within groups of workers with characteristics x was induced by shifts in worker characteristics in the population or by the level of hours dispersion within these groups.

The results shown in table 9 suggest that for men, the within-group dispersion of working hours generally rose (by 145.777, see third row of table 9), partly induced by shifts towards age, education and occupation groups with high within-group hours dispersion. De-unionization also resulted in a shift towards groups with higher within-group hours dispersion. However, apart from these moderate compositional influences, there was a strong general trend towards more within-group hours dispersion (see the value for the constant, 272.364), which was mitigated by declining within-group hours dispersion especially along the dimensions of education, occupation, industry and firmsize. For women, the general increase in average within-group hours dispersion was even more pronounced (see the contribution of the constant, 346.818). In addition, there were strong compositional effects of de-unionization, i.e. de-unionization increased the percentage of groups with higher within-group hours dispersion, and to a lower extent of changes in education. The lower part of table 9 shows that the overall increase in within-group hours dispersion was to

a large extent countervailed by decreases in within-group hours dispersion, especially along the dimensions occupation and industry.

— Table 9 about here —

For the between-group variance[4] of working hours $var(E(h_t|x_t)) = E\left[(E(h_t|x_t) - E(h_t))^2\right] = E\left[(x_t\beta_t - E(h_t))^2\right] = E\left[\alpha_t x_t\right]$, we have

$$E\left[(E(h_1|x_1) - E(h_1))^2\right] - E\left[(E(h_0|x_0) - E(h_0))^2\right] = \left[E(x_1) - E(x_0)\right]\alpha_1 + \left[\alpha_1 - \alpha_0\right]E(x_0) \tag{12}$$

(i.e. an Oaxaca-Blinder decomposition with dependent variable $(E(h_t|x_t) - E(h_t))^2 = (x_t\beta_t - E(h_t))^2$). The decomposition describes to what extent changing average differences in working hours *between* groups of workers with characteristics x can be explained by shifts in characteristics or by the strength of the influence of characteristics on between-group hours variance. The results in table 10 show that for both men and women, between group variability of working hours increased (for men by 24.310, for women by 58.122). Only for women, a substantial part of these overall differences can be accounted for by shifts in characteristics. In the case of women, these were shifts towards higher education groups which increase the weight of working hours that are far above the population average. For men, we observe a convergence of mean working hours across occupation (-36.380), while for women we observe some convergence over regions (-44.062). However, there remains a large unexplained *general* trend of rising between-group differences in working hours, as represented by the constant (+63.956 for men and +36.041 for women).

— Table 10 about here —

[4]Note that $var(h_t) = E(var(h_t|x_t)) + var(E(h_t|x_t))$, i.e. the within-group and the between-group variance add up to the overall variance of working hours.

Summing up the findings in this section, we find that changes in *mean* hours were hardly connected to the observed characteristics used in our analysis. Changes in within-group and between-group hours *dispersion* were to some extent associated with the observed characteristics considered by us, but there remains a large unexplained trend towards more within- and between-group hours dispersion unrelated to these characteristics. The fact that there is only a weak link between working hours and the observable determinants of the earnings distribution may be a reason why the quantitative importance of hours changes to changes in earnings inequality in our decomposition do not differ more from that in the classic variance decomposition.

5 Conclusion

This paper has examined the influence of changes in hours worked on the distribution of monthly earnings of men and women in Germany. Our findings suggest that for men, between 10 and 30 percent of the increase in monthly earnings inequality between 2001 and 2010 can be explained by changes in hours worked for given groups of workers. For women, the contribution of hours changes to rising earnings inequality was much higher, around 37 to 47 percent. The likely reason for the latter finding is that women are much more affected by the expansion of part-time and marginal part-time work arrangements which tend to introduce a large amount of inequality into the earnings distribution. We point out that, due to the fact that our sample has incomplete coverage of the service sector, our estimates of the influence of hours changes are likely to constitute a lower bound of the real effect as part-time and marginal part-time work is more widespread in the service sector. Compared to a classical decomposition of the variance of log earnings, our decomposition based on changes in the *conditional* hours distribution yields a similarly high estimate of the effect of hours changes on the earnings distribution for women, but a lower one for men. Finally, in line with other studies (Dustmann et al., 2009, Card, 2013, Biewen

et al., 2017b), our results suggest that in addition to hours changes, compositional changes in personal and firm characteristics have contributed a lot to rising earnings inequality in Germany .

6 References

Autor, D.H., L.F. Katz, M.S. Kearney, (2008): Trends in us wage inequality: Revising the revisionists. *Review of Economics and Statistics*, Vol. 90, pp. 300-323.

Biewen, M., M. Ungerer, M. Löffler (2017a): Why did income inequality in Germany not increase further after 2005? Forthcoming in: *German Economic Review*.

Biewen, M., B. Fitzenberger, J. DeLazzer (2017b): Rising Wage Inequality in Germany: Increasing Heterogeneity and Changing Selection into Full-Time Work, *IZA Discussion Paper No. 11072*, IZA Institute of Labor Economics, Bonn.

Blau, F.D., L.M. Kahn (2011). Inequality and Earnings Distribution. In: Salverda, W., B. Nolan, T. Smeeding (eds.): *The Oxford Handbook of Economic Inequality*, pp. 177-203. Oxford University Press, Oxford.

Burtless, G. (1990). *A Future of Lousy Jobs: The Changing Structure of U.S. Wages*. The Brookings Institution Press, Washington D.C.

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

Checchi, D., C. Garcia-Penalosa, L. Vivian (2016). Are changes in the dispersion of hours worked a cause of increased earnings inequality? *IZA Journal of European Labor Studies*, Vol. 5:15, pp. 1-34.

DiNardo, J., N. Fortin, T. Lemieux (1996): Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, Vol. 64, pp. 1001-1044.

Doiron, D.J., G.F. Barrett (1996). Inequality in Male and Female Earnings: The Role of Hours and Wages. *Review of Economics and Statistics*, Vol. 78, pp. 410-420.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): Revisting the German Wage Structure. *The Quarterly Journal of Economics*, Vol. 124, pp. 843-881.

Foster, J., E. Ok (1999). Lorenz Dominance and the Variance of Logarithms. *Econometrica*, Vol. 67, pp. 901-907.

Fuchs-Schündeln, N., D. Krüger, M. Sommer (2010). Inequality trends for Germany in the last two decades: A tale of two countries. *Review of Economic Dynamics*, Vol. 13, pp. 103-132.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *Stata Journal*, Vol. 8, pp. 453-479.

Juhn, C., K. Murphy, B. Pierce (1993). Wage Inequality and the Rise in Return to Skill. *Journal of Political Economy*, Vol. 101, pp. 410-442.

Johnson, S, P. Kuhn (2004). Increasing Male Earnings in Canada and the United States, 1981-1997: The Role of Hours Changes vs. Wage Changes. *Canadian Public Policy/Analyse de Politiques*, Vol. 30., pp. 155-175.

Katz, L.F., D.H. Autor (1999): Changes in the Wage Structure and Earnings Inequality. In: Ashenfelter, O., D. Card (editors), *Handbook of Labor Economics*, Volume 4b, chapter 26, pp. 1463–1166. North Holland, Amsterdam.

Machin, S. (2008): An Appraisal of Economic Research on Changes in Wage Inequality, *LABOUR*, Vol. 22, pp. 7-26.

OECD (2011): Divided we stand: *Why inequality keeps rising*. OECD Publishing. http://dx.doi.org/10.1787/9789264119536-en

Shorrocks, A. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, Vol. 11, pp. 99-126.

7 Tables and figures

Table 1 – Inequality measures monthly earnings

	Men			Women		
Index	2001	2006	2010	2001	2006	2010
Mean	3456.19	3469.65	3406.26	2186.62	2107.93	2089.01
	(23.28)	(24.77)	(22.04)	(21.35)	(17.13)	(20.932)
Median	3077.80	2093.72	3004.00	2060.64	1943.56	1885.00
	(17.99)	(23.02)	(18.76)	(20.54)	(17.519)	(21.00)
Gini	0.230	0.246	0.261	0.301	0.334	0.345
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
MLD	0.096	0.117	0.130	0.193	0.240	0.248
	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Theil	0.093	0.107	0.120	0.155	0.190	0.201
	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)
p9010	2.662	2.849	3.064	4.627	8.738	9.342
	(0.028)	(0.024)	(0.028)	(0.144)	(0.066)	(0.096)
p9050	1.749	1.778	1.840	1.783	1.915	1.982
	(0.014)	(0.012)	(0.012)	(0.014)	(0.013)	(0.016)
p5010	1.522	1.602	1.665	2.595	4.563	4.713
	(0.008)	(0.008)	(0.010)	(0.076)	(0.041)	(0.049)

Source: Structure of Earnings Surveys 2001, 2006, 2010, and own calculations.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

Table 2 – Inequality measures monthly hours worked

	Men			Women		
Index	2001	2006	2010	2001	2006	2010
Mean	166.52	166.35	164.33	135.97	131.05	129.66
	(0.234)	(0.283)	(0.234)	(0.506)	(0.468)	(0.540)
Median	167.05	169.45	168.50	156.92	152.07	152.08
	(0.245)	(0.036)	(0.451)	(1.250)	(0.134)	(0.008)
Gini	0.060	0.072	0.072	0.169	0.199	0.201
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
MLD	0.016	0.025	0.027	0.084	0.110	0.108
	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
Theil	0.012	0.018	0.019	0.062	0.081	0.080
	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)
p9010	1.209	1.228	1.200	2.500	3.361	3.342
	(0.003)	(0.003)	(0.002)	(0.079)	(0.057)	(0.038)
p9050	1.102	1.102	1.083	1.108	1.144	1.143
	(0.002)	(0.003)	(0.003)	(0.009)	(0.001)	(0.000)
p5010	1.097	1.114	1.108	2.257	2.937	2.925
	(0.002)	(0.000)	(0.003)	(0.065)	(0.065)	(0.033)

Source: Structure of Earnings Surveys 2001, 2006, 2010, and own calculations.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

Table 3 – Inequality measures hourly wages

	Men			Women		
Index	2001	2006	2010	2001	2006	2010
Mean	20.80	20.97	20.77	15.57	15.50	15.48
	(0.122)	(0.205)	(0.143)	(0.122)	(0.102)	(0.128)
Median	18.48	18.58	14.24	13.89	13.88	13.85
	(0.143)	(0.143)	(0.160)	(0.146)	(0.099)	(0.134)
Gini	0.226	0.245	0.256	0.215	0.239	0.249
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
MLD	0.083	0.097	0.106	0.077	0.093	0.101
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Theil	0.088	0.103	0.112	0.079	0.096	0.104
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
p9010	2.669	2.941	3.084	2.647	2.968	3.095
	(0.029)	(0.026)	(0.029)	(0.024)	(0.029)	(0.037)

p9050	1.742	1.789	1.847	1.631	1.746	1.770
	(0.016)	(0.016)	(0.012)	(0.012)	(0.012)	(0.014)
p5010	1.533	1.644	1.669	1.624	1.700	1.749
	(0.008)	(0.012)	(0.011)	(0.014)	(0.011)	(0.017)

Source: Structure of Earnings Surveys 2001, 2006, 2010, and own calculations.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

Table 4 – Variance decomposition log monthly earnings, men

Year	$var(\log(e))$	$var(\log(h))$	$var(\log(w))$	$2cov(\log(h), \log(w))$
2001	0.211	0.0454	0.155	0.011
	(0.007)	(0.003)	(0.003)	(0.003)
	[100]	[21.42]	[73.25]	[5.28]
2006	0.272	0.075	0.185	0.017
	(0.006)	(0.002)	(0.002)	(0.002)
	[100]	[27.18]	[66.79]	[6.00]
2010	0.306	0.080	0.201	0.025
	(0.006)	(0.003)	(0.003)	(0.003)
	[100]	[26.24]	[65.69]	[8.20]
Year	$\Delta var(\log(e))$	$\Delta var(\log(h))$	$\Delta var(\log(w))$	$\Delta 2cov(\log(h), \log(w))$
2006/2001	0.065	0.030	0.030	0.005
	(0.009)	(0.004)	(0.004)	(0.004)
	[100]	[46.00]	[45.75]	[8.25]
2010/2006	0.029	0.005	0.016	0.009
	(0.009)	(0.004)	(0.004)	(0.004)
	[100]	[17.12]	[55.81]	[29.05]
2010/2001	0.094	0.035	0.045	0.014
	(0.009)	(0.004)	(0.004)	(0.004)
	[100]	[36.98]	[48.25]	[14.77]

Source: Structure of Earnings Surveys 2001, 2010, and own calculations. Percentages in square brackets.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

e = monthly earnings, h = monthly hours, w = hourly wage

Table 5 – Variance decomposition log monthly earnings, women

Year	$var(\log(e))$	$var(\log(h))$	$var(\log(w))$	$2cov(\log(h), \log(w))$
2001	0.497	0.228	0.151	0.117
	(0.011)	(0.007)	(0.003)	(0.005)

	[100]	[45.88]	[30.38]	[23.54]
2006	0.622	0.299	0.182	0.141
	(0.014)	(0.008)	(0.003)	(0.006)
	[100]	[48.08]	[29.26]	[22.67]
2010	0.623	0.285	0.197	0.141
	(0.010)	(0.005)	(0.004)	(0.005)
	[100]	[45.75]	[31.62]	[22.63]
Year	$\Delta var(\log(e))$	$\Delta var(\log(h))$	$\Delta var(\log(w))$	$\Delta 2cov(\log(h), \log(w))$
2006/2001	0.124	0.071	0.030	0.023
	(0.018)	(0.010)	(0.004)	(0.007)
	[100]	[56.94]	[24.52]	[18.79]
2010/2006	0.002	-0.013	0.015	0.001
	(0.017)	(0.009)	(0.005)	(0.008)
	[100]	[-631.71]	[703.15]	[28.35]
2010/2006	0.126	0.057	0.045	0.024
	(0.015)	(0.008)	(0.004)	(0.007)
	[100]	[45.22]	[35.90]	[18.97]

Source: Structure of Earnings Surveys 2001, 2010, and own calculations. Percentages in square brackets.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

e = monthly earnings, h = monthly hours, w = hourly wage

Table 6 – Decomposition of change in distribution of monthly earnings 2001 to 2010, men

Index	ΔIndex	Δx_p	$\Delta x_f \vert x_p$	$\Delta h \vert x_f, x_p$	Residual
Mean	354.86	-13.76	67.76	-26.08	326.94
	(28.06)	(6.20)	(10.23)	(1.39)	(31.15)
	[100]	[-3.88]	[19.09]	[-7.35]	[92.13]
Median	315.11	-20.44	37.09	-14.73	313.19
	(22.28)	(3.97)	(7.15)	(0.82)	(23.50)
	[100]	[-6.49]	[11.77]	[-4.67]	[99.39]
Gini	0.024	0.010	0.010	0.004	-0.000
	(0.003)	(0.001)	(0.001)	(0.000)	(0.004)
	[100]	[41.90]	[41.49]	[17 .93]	[-1.50]
MLD	0.028	0.014	0.007	0.009	-0.002
	(0.003)	(0.002)	(0.001)	(0.001)	(0.004)
	[100]	[49.12]	[25.41]	[32.58]	[-7.21]
Theil	0.023	0.009	0.007	0.005	0.002
	(0.003)	(0.001)	(0.001)	(0.000)	(0.003)

	[100]	[40.00]	[32.00]	[19.40]	[8.42]
p9010	0.400	0.122	0.151	0.038	0.089
	(0.039)	(0.017)	(0.014)	(0.005)	(0.053)
	[100]	[30.50]	[37.75]	[9.60]	[22.23]
p9050	0.092	0.037	0.046	0.001	0.009
	(0.020)	(0.004)	(0.006)	(0.001)	(0.022)
	[100]	[39.72]	[49.46]	[0.94]	[9.96]
p5010	0.141	0.036	0.044	0.206	0.040
	(0.011)	(0.007)	(0.005)	(0.002)	(0.016)
	[100]	[25.32]	[31.35]	[14.61]	[28.44]

Source: Structure of Earnings Surveys 2001, 2010, and own calculations.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

Shapley decomposition. Percentages in square brackets.

Table 7 – Decomposition of change in distribution of monthly earnings 2001 to 2010, women

Index	ΔIndex	Δx_p	$\Delta x_f \vert x_p$	$\Delta h \vert x_f, x_p$	Residual
Mean	174.20	1.09	2.50	-78.49	249.09
	(22.16)	(6.68)	(10.35)	(1.97)	(23.04)
	[100]	[0.63]	[1.44]	[-45.06]	[142.99]
Median	92.72	-22.77	-28.03	-81.72	225.24
	(24.44)	(6.47)	(8.42)	(3.20)	(23.50)
	[100]	[-24.56]	[-30.24]	[-88.14]	[242.93]
Gini	0.043	0.014	0.018	0.016	-0.006
	(0.004)	(0.002)	(0.002)	(0.000)	(0.004)
	[100]	[33.10]	[42.72]	[37.09]	[-13.03]
MLD	0.054	0.021	0.025	0.025	-0.017
	(0.005)	(0.003)	(0.002)	(0.001)	(0.007)
	[100]	[39.11]	[46.00]	[46.93]	[-31.84]
Theil	0.045	0.016	0.020	0.017	-0.008
	(0.004)	(0.002)	(0.001)	(0.001)	(0.005)
	[100]	[35.71]	[45.09]	[38.17]	[-18.75]
p9010	4.492	1.510	1.590	2.474	-1.081
	(0.156)	(0.263)	(0.152)	(0.294)	(0.165)
	[100]	[33.62]	[35.40]	[55.08]	[-24.07]
p9050	0.191	0.063	0.076	0.053	0.000
	(0.021)	(0.007)	(0.008)	(0.003)	(0.022)
	[100]	[32.83]	[39.74]	[27.44]	[0.000]
p5010	2.011	0.670	0.683	1.202	-0.544

	(0.080)	(0.127)	(0.076)	(0.158)	(0.064)
	[100]	[33.32]	[33.97]	[59.78]	[-27.07]

Source: Structure of Earnings Surveys 2001, 2010, and own calculations.

Bootstrapped standard errors in parentheses (500 replications, clustered at firm level).

Shapley decomposition. Percentages in square brackets.

Table 8 – Mean working hours 2010 vs. 2001: Oaxaca-Blinder decomposition

	Men		Women	
Mean hours	Result	Std.err.	Result	Std.err
2010	165.121	.164	138.296	.286
2001	167.080	.172	141.890	.321
Difference	-1.958	.238	-3.593	.430
Groups of covariates	Explained by differences in characteristics			
Age	-.459	.032	-.439	.050
Education	-.221	.042	-.488	.101
Tenure	-.051	.014	.310	.038
Occupation	-.007	.049	.691	.140
Region	.043	.052	.244	.153
Industry	-.015	.077	1.153	.170
Public ownership	-.012	.011	-.006	.016
Firmsize	-.117	.051	.102	.037
Union coverage	.165	.031	-.462	.080
Total	-.675	.155	1.105	.313
Groups of covariates	Explained by differences in coefficients			
Age	.017	.026	-.385	.046
Education	-.087	.135	-.115	.170
Tenure	.354	.108	.753	.227
Occupation	.827	.355	1.961	.685
Region	.639	.580	.132	.818
Industry	.170	.216	-.032	.307
Public ownership	-.109	.042	-.048	.080
Firmsize	.869	.101	.565	.113
Union coverage	-.125	.029	.127	.056
Constant	-3.838	.649	-7.657	1.047
Total	-1.282	.232	-4.698	.368

Source: Structure of Earnings Surveys 2001, 2010, and own calculations.

Standard errors clustered at firm level.

Table 9 – Within-group variance working hours 2010 vs. 2001: Oaxaca-Blinder decomposition

	Men		Women	
Mean hours	Result	Std.err.	Result	Std.err
2010	623.705	9.747	1562.846	12.939
2001	477.927	10.058	1322.562	14.983
Difference	145.777	14.033	240.284	19.819
Groups of covariates	Explained by differences in characteristics			
Age	37.310	2.164	2.053	2.329
Education	31.638	4.380	29.712	5.155
Tenure	-11.458	1.514	-18.854	2.106
Occupation	20.506	4.022	-18.125	5.617
Region	-8.002	2.744	-7.526	5.539
Industry	3.962	4.450	-20.091	6.375
Public ownership	1.075	.655	.329	1.102
Firmsize	-1.369	1.282	-3.880	1.945
Union coverage	19.600	3.267	74.414	7.204
Total	93.263	9.929	38.031	14.478
Groups of covariates	Explained by differences in coefficients			
Age	-2.498	2.063	11.999	2.796
Education	-17.814	11.260	5.091	9.633
Tenure	-.348	6.897	25.288	11.527
Occupation	-114.934	25.850	-135.691	31.644
Region	9.000	30.818	-.446	41.028
Industry	-47.353	12.871	-33.078	14.803
Public ownership	-2.382	2.737	7.667	4.625
Firmsize	-52.915	7.458	-17.598	5.594
Union coverage	9.396	2.132	-7.797	3.438
Constant	272.364	41.420	346.818	51.541
Total	52.513	17.245	202.253	18.344

Source: Structure of Earnings Surveys 2001, 2010, and own calculations.

Standard errors clustered at firm level.

Table 10 – Between-group variance working hours 2010 vs. 2001: Oaxaca-Blinder decomposition

	Men		Women	
Mean hours	Result	Std.err.	Result	Std.err
2010	71.556	.777	390.668	5.201
2001	47.245	1.048	332.545	5.548
Difference	24.310	1.300	58.122	7.595
Groups of covariates	Explained by differences in characteristics			
Age	2.806	.174	6.488	.549
Education	-.057	.144	29.200	2.111
Tenure	.245	.064	.322	.477
Occupation	1.992	.544	.949	4.388
Region	.190	.595	4.860	1.467
Industry	-.661	.482	-9.455	2.517
Public ownership	.078	.046	-.187	.212
Firmsize	.129	.106	-.277	.318
Union coverage	.414	.076	1.631	1.110
Total	5.138	1.064	33.531	5.882
Groups of covariates	Explained by differences in coefficients			
Age	-.108	.086	-2.820	.678
Education	.234	.598	16.278	2.724
Tenure	-12.107	.804	7.095	2.895
Occupation	-36.380	1.689	6.064	7.654
Region	7.491	4.774	-44.062	18.783
Industry	-4.315	1.368	15.338	7.726
Public ownership	-.205	.167	-1.674	1.005
Firmsize	.654	.352	-7.483	1.748
Union coverage	-.047	.071	-.187	.219
Constant	63.956	4.789	36.041	15.753
Total	19.171	.913	24.590	5.507

Source: Structure of Earnings Surveys 2001, 2010, and own calculations.

Standard errors clustered at firm level.

Figure 1 – Distribution of monthly earnings, men

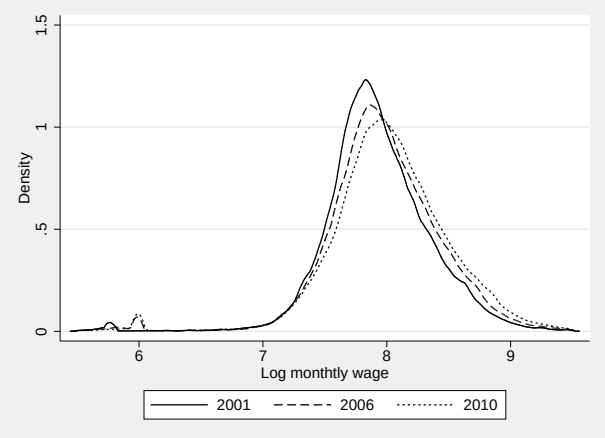


Figure 2 – Distribution of monthly hours worked, men

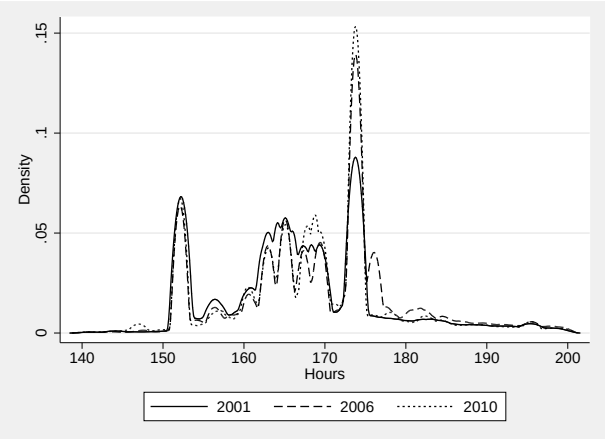


Figure 3 – Distribution of hourly wages, men

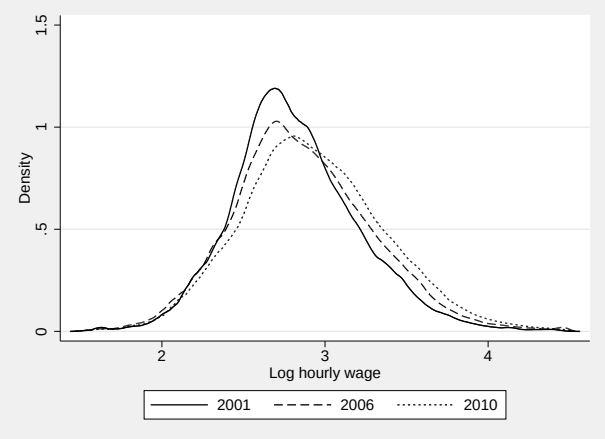


Figure 4 – Distribution of monthly earnings, women

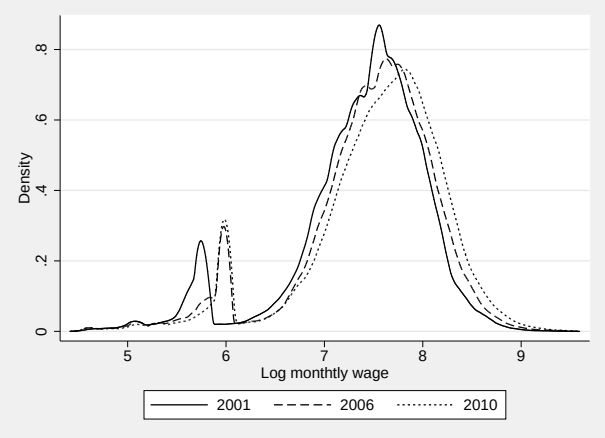


Figure 5 – Distribution of monthly hours worked, women

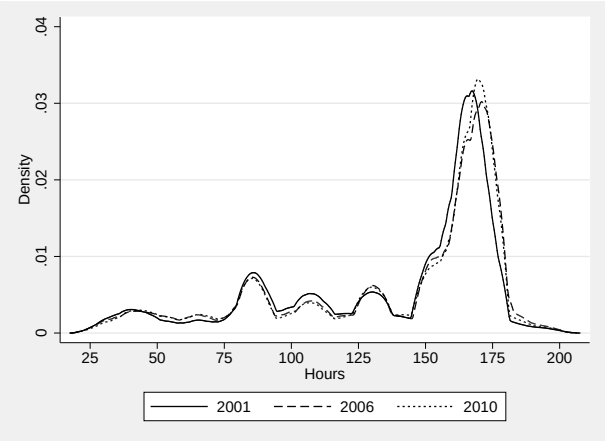
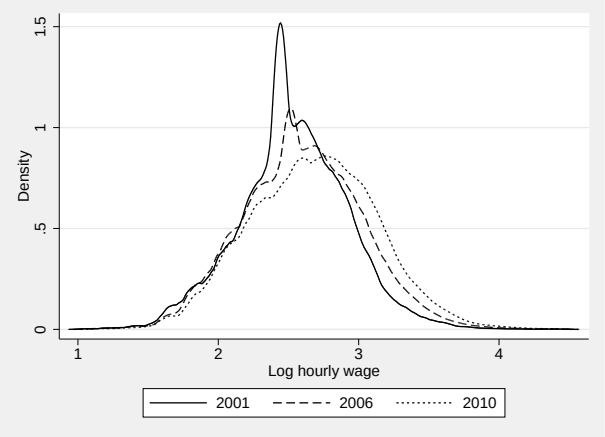


Figure 6 – Distribution of hourly wages, women



8 Appendix

Table A1 – Summary statistics (means)

	Men		Women	
Variable	2001	2010	2001	2010
Personal characteristics				
Age 25-29 years	.101	.100	.112	.107
Age 30-34 years	.163	.112	.159	.106
Age 35-39 years	.198	.126	.183	.121
Age 40-44 years	.174	.181	.170	.183
Age 45-49 years	.145	.194	.153	.192
Age 50-54 years	.122	.153	.129	.152
Age 55-60 years	.094	.131	.091	.136
Lower/Middle sec., no voc. tr.	.122	.099	.171	.118
Lower/Middle sec., voc. tr.	.671	.627	.626	.566
Upper sec., possibly voc. tr.	.045	.054	.066	.088
University degree	.100	.107	.045	.064
Education unknown	.060	.109	.089	.162
Tenure 0-9 years	.515	.471	.603	.545
Tenure 10-20 years	.283	.309	.266	.306
Tenure 21-30 years	.140	.146	.099	.101
Tenure 31-46 years	.060	.073	.030	.046
Occupation	*(52 categories)*			
Firm characteristics				
Region North (SH,HH,BR,NS)	.157	.165	.157	.172
Region Middle (NRW)	.280	.228	.263	.220
Region Middle-South (HE,RP,SL)	.137	.136	.140	.133
Region South (BW,BY)	.326	.351	.334	.349
Region East (MV,B,BR,S,SA)	.096	.118	.104	.123
Mining	.012	.007	.002	.001
Food production, tobacco	.048	.051	.102	.099
Wood	.052	.058	.056	.051
Chemical industry	.045	.042	.035	.037
Manufacturing plastic	.044	.059	.027	.038
Metal production	.204	.207	.084	.089
Production business machines	.079	.062	.076	.056
Vehicle construction	.094	.086	.025	.027
Energy and water supply	.031	.047	.015	.027

Building and civil engineering	.070	.043	.013	.009
Constructional installations	.050	.065	.018	.023
Automobile trade, repair, gas	.125	.161	.111	.148
Retail, mending of durables	.054	.046	.247	.248
Banking	.043	.042	.113	.108
Insurance	.043	.017	.068	.032
Public ownership	.058	.034	.066	.045
Firmsize up to 50 employees	.219	.231	.227	.265
Firmsize 50-249 employees	.270	.290	.262	.279
Firmsize more than 249 employees	.509	.478	.510	.454
Union agreement	.623	.524	.641	.486
Number of observations	332,155	428,265	150,339	189,076

Source: Structure of Earnings Surveys 2001, 2010. Weighted data.