Online Appendix: Revisiting the German Wage Structure

Christian Dustmann * Johannes Ludsteck † Uta Schönberg ‡

This Version: July 2008

This appendix consists of three parts. Section 1 compares alternative methods for imputing censored wages in the IAB data. Section 2 evaluates the ability of each imputation method to recover the censored part of the wage distribution using the GSES data where wages are not top-coded. Section 3 compares trends in inequality in the GSOEP and the IABS, while Section 4 compares trends in inequality in the LIAB and IABS.

1 Imputation of Censored Wages

As in many administrative data sets, in the IABS each year between 9% and 14% of the male wage distribution is censored. Censoring is less severe among women: Here, less than 5% of the wage observations are censored each year. Censoring also varies widely across education groups. It is negligible among the low-skilled (here, less than 1% of wages are censored). For the high-skilled, in contrast, more than 50% of the wage observations are censored. Due to censoring, much of our analysis focuses on changes in the uncensored part of the wage distribution up to the 85th percentile,

*Department of Economics, University College London, Gower Street, London WC1E 6BT, England. E-mail: c.dustmann@ucl.ac.uk.

†Institut fuer Arbeitsmarkt- und Berufsforschung, Nuernberg. E-mail: Johannes.Ludsteck@iab.de.

‡Department of Economics, University College London, Gower Street, London WC1E 6BT, England. E-mail: utas@troi.cc. rochester.edu.

and imposes no distributional assumptions on the error term. However, some of our findings–in particular those in residual inequality–require distributional assumptions. We have imputed the censored part of the wage distribution under six alternative distributional assumptions. This section describes each imputation method, and compares changes in residual inequality for the different imputation methods.

Our baseline results (i.e., those reported in the paper) assume that the error term in a wage regression is normally distributed, with different variances for each education group and each age group. We label this imputation method "normal, some heteroskedasticity". In order to impute censored wages under this assumption, we estimate censored regressions (allowing for a different variance for each education and age group) separately for each year (and thereby allow the variance in each group to vary across years). We control for all possible interactions between three education and eight age categories. We then impute censored wages as the sum of the predicted wage and a random component drawn from a normal distribution with mean zero and a separate variance for each education and age group, obtained from the standard error of the forecast. We prefer to work with imputed wages rather than with censored wages because wage residuals can be computed in a straightforward manner. A comparison between OLS estimates based on imputed wages and tobit estimates based on censored wages shows that both the estimates and the standard errors are almost identical.

We conduct five robustness checks.

1. (Imputation Method 2) First, we impute wages in a similar manner, but restrict the variance to be the same across education and age groups. This imputation method is labeled "normal, no heteroskedasticity".

2. (Imputation Method 3) Second, we continue to assume that the error term is normally distributed, but allow for a separate variance for each education and age cell. This imputation method is labeled "normal, full heteroskedasticity".

3. (Imputation Method 4) As a third robustness check, we replace censored values with 1.5 times the censoring limit, similar to US studies based on the CPS. We label this imputation method "1.5 times censoring limit".

4. (Imputation Method 5) However, imputations based on the normal distribution assumption suggest that this factor is too high, and is in fact closer to 1.2. As a fourth robustness check, we replace censored values with 1.2 times the censoring limit. We label this imputation method "1.2 times censoring limit".

5. (Imputation Method 6) Fifth, we impute wages under the assumption that the upper tail of the unconditional wage distribution follows a pareto distribution. This imputation method is labeled "pareto". We implement this method as follows. In the first step, we fit (for each year separately) a Pareto distribution to the upper tail of the wage distribution. Since the empirical wage distribution is concave near the median but the pareto distribution is globally convex, we choose the 60th percentile of the wage distribution as the lower limit of this upper tail. Fitting is performed by maximizing the log-likelihood of the pareto-distribution where censoring of wages from above is accounted for by adding the complement of the value of the distribution function of censored observations to the likelihood. In the second step, we replace censored observations with draws from the fitted (truncated) pareto distribution.

Figure I* plots the standard deviation of log-wages (Panels A and C) and log-wage residuals

(Panel B and D) for men and women. In addition to the six imputation methods just described, we also plot the standard deviation when censored observations are dropped. The figure shows that the imputation method affects the *level* of the standard deviation. Not surprisingly, it is lowest when the top 15% of observations are dropped. The standard deviation is highest when censored wages are imputed under the Pareto assumption, or when censored wages are replaced by 1.5 times the censoring limit. However, all imputation methods show a steady increase in the standard deviation of log-wages and log-wage residuals for men, and an increase starting in the mid- to early 1990s for women.

Figures II and III in the paper are solely based on the uncensored part of the wage distribution, and are therefore not affected by alternative imputation methods. Our findings on polarization (Figure VII) are likewise not affected by wage censoring. The computation of the composition-constant counterfactual distribution of *overall* inequality in Table II does not impose distributional assumptions on the error term, either. However, distributional assumptions are required in order to compute the counterfacutal distribution of *residual* inequality. Table II* reports the counterfactual distribution of residual inequality for our six alternative imputation methods. Overall, results are very similar for alternative imputation methods. Note, however, that for the imputation rules that replace censored observations with 1.2 or 1.5 times the censoring limit, the composition-adjusted 85/50 wage gap slightly exceeds the observed 85/50 wage gap. Table IV* displays the counterfactual distribution of residual inequality when unionization in addition to age and education is held constant, using the LIAB data. Again, results are similar across alternative imputation methods.

Turning to between-group inequality, Figure V* plots the medium/low (Panel A) and high/medium (Panel B) wage differentials for the alternative imputation methods. All imputation methods show

a slight decline in the medium/low wage differential throughout the late 1970s and 1980s, and a substantial increase throughout the 1990s. The medium/low wage differential is slightly larger when censored wages are imputed under the Pareto assumption or when they are replaced by 1.5 times the censoring limit. In contrast, the different imputation methods yield very different trends in the high/medium wage gap during the late 1970s. Throughout the 1980s, the level of the high/medium wage gap also varies across imputation methods, with the gap being smallest when censored observations are replaced by 1.2 times the censoring limit, and highest when censored observations are imputed under the assumption that the error term is normally distributed with a different variance for each education and age cell. However, from the mid-1980s onward, all imputation methods show a similar trend in the high/medium wage gap. It is not surprising that throughout the late 1970s, the different imputation methods produce very different estimates for the high/medium wage gap. Between 1975 and 1979, 73% of the wage observations of the male high-skilled are censored, and the high-skilled make up 4.8% of the male working population. For comparison, between 2000 and 2004, in contrast, 48% of the wage observations are censored, and the high-skilled make up 13.0% of the male working population. These results highlight that trends on the evolution of the high/medium wage gap should be treated with considerable caution.

To sum up, with the exception of the high-medium wage differential in the 1970s, inequality trends (though not necessarily inequality levels) are very similar across alternative imputation methods.

2 Which Imputation Method Works Best?

This section uses the GSES (German Structure of Earnings Survey) data to evaluate which imputation method performs best at recovering the censored part of the wage distribution. We prefer the GSES over the GSOEP for this analysis because of the much larger sample size. The GSES is a survey of about 27,000 establishments, conducted by the German Federal Statistical Office. Participation in the survey is compulsory. Contrary to the IABS and LIAB, there is no wage censoring in the GSES. A scientific use file of the GSES is currently available for researchers for the year 2001. The survey covered establishments with more than ten employees. In smaller establishments with less than 50 employees, the survey covers all employees; in larger establishments, the information refers to a random sub-sample of employees. Wage data refers to October. Similar to the IABS, the education variable is missing for about 10% of the observations. Unlike the IABS, the GSES does not allow the imputation of missing observations. We therefore distinguish four education categories, with "missing" as the fourth category.

In the first step, we replace all wage observations in the GSES that are above the IABS censoring limit for the year 2001. In the second step, we impute censored wages in the GSES in the same way as in the IABS. Table A.1 compares the share of "censored" observations in the IABS and GSES by age and education group. The share of censored observations is slightly larger in the GSES and IABS. Note that in both data sets more than half of the wage observations of the high-skilled are censored. This suggests that the GSES provides useful information about the censored part of the wage distribution in the IABS.

In Table A.2, we compare the true and imputed standard deviation of log-wages (column (1)),

the standard deviation of log-wage residuals (column (2)), the 85/50 and 50/15 residual wage gaps (columns (3) and (4)) as well as the medium/low and high/medium wage differentials (columns (5) and (6)). When we compute the standard deviation of log-wage residuals or the 50/15 and 85/50 residual wage gaps, our regressions control for all possible interactions between four education categories (which includes one category for missing observations) and eight age categories. The education wage premiums are obtained simply by regressing log-wages on the education dummies.

The three imputation methods based on the normal distribution as well as the substitution of censored observations with 1.2 times the censoring limit slightly underestimate, whereas the Pareto imputation as well as the substitution of censored observations with 1.5 times the censoring limit overestimate the standard deviation of log-wages (column (1)). Note that these two imputation methods also yield the highest standard deviation in the IABS data (see Figure I*). Imputation method 1 that assume that the error term is normally distributed and allows for different variances for each age and each education group as well as imputation method 2 that restricts the variance to be the same across all age and education groups outperform the other methods. A similar picture emerges for the standard deviation of log-wage residuals (column (2)). All imputation methods are able to closely replicate the 50/15 residual log wage gap (column (4)). Turning to the 85/50 residual wage gap, imputation methods 1 and 2 again perform better than any other method. The same is true for the the medium/low and high/medium wage differential.

To summarize, imputation methods 1 and 2 outperform the other imputation methods, in particular that based on the pareto distribution. Our baseline results (i.e., those reported in the paper) are based on the more flexible method 1 that allows the variance to vary by age and education group. Our findings are similar when we use imputation method 2.

3 Comparison Between GSOEP and IABS

This section compares trends in inequality between the German Socio-Economic Panel (GSOEP) and the IABS.

Sample Selection and Variable Description The GSOEP is a representative longitudinal study of private households. It was modeled after the PSID in the United States. Once a year, all members of the households aged 16 or older are questioned. Respondents who move continue to take part in the study as long as the move is within Germany. The panel started in 1984; the last wave we use is 2003. To make the GSOEP sample comparable to our IABS sample, we impose the same sample-selection criteria. That is, we select all individuals in West Germany between ages 21and 60 who are working full-time (defined as working at least 30 hours per week) and not currently in apprenticeship training. Since the IABS only covers workers within the social security system, and thus excludes the self-employed and the civil servants, we construct a sub-sample that discards these workers. We use three alternative wage measures. Our first measure is the monthly regular salary. Our second measure includes one-time payments, such as bonuses, holiday and Christmas money. At each interview date, respondents are asked about one-time payments in the *past* year. We therefore use one-time payments from the past year to compute this year's wage measure. This means that we lose the final year (2003) from our sample. Our third measure is an hourly wage rate that includes one-time payments. Here, we divide our monthly wage measure by the monthly hours worked. The GSOEP asks about usual as well as actual weekly working hours. We use the maximum of the two, and convert weekly hours worked into monthly hours worked by multiplying by 4.33. The wage measure that is most similar to the one in the IABS is the second measure, i.e. a monthly wage that

includes one-time payments. All our results are weighted with the sampling weight provided by the GSOEP to make them representative for the West German economy.

Trends in Inequality in the GSOEP and IABS Tables A.3 (men) and A.4 (women) report the 85/50 and 50/15 wage gaps in each year, for three different specifications. Column (1) includes the self-employed and civil servants and uses the monthly wage excluding one-time payments as the wage measure. The second column refers to the same wage measure, but excludes the self-employed and civil servants. The third column also excludes the civil servants and the self-employed, but includes one-time payments in the monthly wage measure. This specification resembles that in the IABS most closely. For comparison, column (4) displays the wage gap in the IABS. Standard errors in parentheses are bootstrapped with 100 replications. We investigate the importance of using an hourly wage as opposed to a monthly wage in detail in Table A.8. To preview our results, using hourly as opposed to monthly wages does not have a large impact on measured inequality. To visualize the trends in upper- and lower-tail inequality in the GSOEP and IABS, Panels A and B of Figures A.1 (men) and A.2 (women) plot the evolution of the 85/50 and 50/15 wage gaps for the different specifications, including one based on hourly wages.

Consider first men. Table A.3 and Figure A.1 show that excluding the civil servants and the self-employed or including one-time payments in the wage measure has little impact on the evolution of the 85/50 and 50/15 wage gaps. The 50/15 wage gap started to increase sharply starting around 1993 – a finding that has also been stressed by Gernandt and Pfeiffer (2006). In fact, the rise in lower-tail inequality between 1993 and 2002 in the GSOEP *exceeds* that in the IABS. In contrast, between 1984 and 1993, the rise in the 50/15 wage gap is higher in the IABS than in the GSOEP.

The 85/50 wage gap also increased in the GSOEP; however, much of the increase is concentrated in the years 1992 to 1995. The table and figure also highlight that the 85/50 and 50/15 wage gaps are noisily estimated. Importantly, the hypothesis that the increase in 85/50 and 50/15 wage gaps between any two years is the same in the GSOEP as in the IABS can usually not be ruled out at a 5% level. For instance, using the specification that most closely resembles that in the IABS, the change in the 50/15 and 85/50 wage gap between 1984 and 1993 is -0.001 and 0.043 in the GSOEP, with a 95% confidence interval of [-0.041, 0.038] and [-0.020, 0.108]. Over the same period, the 50/15 and 85/50 wage gaps in the IABS rose by 0.029 and 0.035, respectively. Similarly, between 1993 and 2002 the 95% confidence interval for the 50/15 and 85/50 wage gap is [0.044, 0.154] and [-0.039, 0.103] in the GSOEP, while the observed change in the IABS is 0.059 and 0.058, respectively.

Just as for men, excluding the civil servants and self-employed or including one-time payments hardly affects the evolution of the 85/50 and 50/15 wage gaps for women. Table A.4 and Figure A.2 show that the 50/15 wage gap is substantially larger in the IABS than in the GSOEP, whereas the 85/50 wage gap is of similar magnitude in the two data sets. One explanation for this finding is measurement error in the full-time variable in the IABS; that is, our IABS sample may include women who in fact work part-time. This is not a problem for men, since part-time work is uncommon. This suggests that the IABS data for women, in particular at the lower end of the wage distribution, may have to be interpreted with some caution. This is another reason why our main analysis focuses on men. The *change* in both upper- and lower-tail inequality is similar in the GSOEP and IABS. Again, we are usually not able to rule out the hypothesis that the change between any two years is the same in both data sets at a 5% level. For instance, the 95% confidence interval for the change in the 50/15 wage gap between 1984 and 1993 is [-0.084, 0.071] in the GSOEP. The observed change in the IABS

is 0.008.

Tables A.5 (men) and A.6 (women), Panels C and D, compare the standard deviations of log-wages and log-wage residuals in the two data sets. The corresponding figures are Figures A.1 and A.2, Panels C and D. Our regressions control for all possible interactions between three education and eight age groups. Results for the IABS are based on imputed wages that assume that the error term is normally distributed with a different variance for each education and age group. In line with Tables A.3 and A.4, the standard deviation of log-wages and log-wage residuals is considerably larger in the IABS than in the GSOEP for women, but not for men. For both men and women, the standard deviations of log-wages and log-wage residuals evolve similarly in the two data sets, although for men the increase between 1984 and 1993 is somewhat more pronounced in the IABS than in the GSOEP. However, the change (0.032) observed for men in the IABS is once again within the 95% confidence interval [-0.018, 0.038] observed in the GSOEP.

Finally, Panels E and F in Tables A.5 and A.6 and Figures A.1 and A.2 compare the medium/low and high/medium wage differential in the IABS and GSOEP. The wage differentials are obtained from regressions that control for all possible interactions between three education and eight age groups and are computed as weighted averages of the respective premiums in each age group, where the weights are the employment-weighted worker share in each age group, averaged over the entire sample period. As before, results for the IABS are based on imputed wages. With the exception of the last three years, the high/medium wage premiums for men look similar in the two data sets. For women, the high/medium wage premiums tend to be larger in the IABS than in the GSOEP. Turning to the medium/low wage differential, the GSOEP and IABS draw rather different pictures. Contrary to the IABS, there is no evidence for an increase in the medium/low wage differential in

the 1990s in the GSOEP for men or women. Notice that the skill premiums are particularly noisily estimated in the GSOEP. For men, typical standard errors for the medium/low and high/medium wage differentials are about 0.025 and 0.045; for women, they are even larger. This makes it difficult to draw any firm conclusion about the evolution of education wage premia from the GSOEP.

To summarize, trends in the 50/15 and 85/50 wage gaps as well as in the standard deviations of log-wages and log-wage residuals are roughly similar in the GSOEP and IABS, although much of the increase in inequality in the GSOEP occurred after 1991. It is also important to stress that due to the small sample size, estimates in the GSOEP are very noisy. It is therefore not surprising that earlier studies that use data from 1984 to the mid-1990s, such as the OECD Employment Report (1996), conclude that there is little evidence for a rise in inequality in Germany. However, the change in the 50/15 and 85/50 wage gaps as well as the changes in the standard deviations of log-wages and log-wage residuals observed in the IABS is within the 95%-confidence interval of that observed in the GSOEP.

Hourly Wage versus Monthly Wage Inequality So far, our wage measure was a monthly wage. However, the hourly wage is usually thought of as a better measure for the "price of labor" than the monthly wage. It is therefore important to understand whether hourly wages or hours worked are the primary source of the rise in earnings inequality. Unfortunately, the IABS only contains information on full-time versus part-time work, but not on the number of hours worked. We now compare the rise in hourly wage and monthly wage inequality in the GSOEP.

The first sets of columns in Table A.7 compare the 85/50 and 50/15 wage gaps in the (log) monthly wage and (log) hourly wage in the two data sets. The wage measure in the GSOEP includes

one-time payments, and the sample is restricted to those workers who are likely to be subject to social security contributions. The specification based on monthly wages is the one that is most comparable to the IABS. Panel A refers to men, and Panel B to women. Figures A.1 and A.2, Panel A and B, visualize the evolution of hourly versus monthly wage gaps in the GSOEP, and compare it to that observed in the IABS. For women, wage gaps based on hourly and monthly wages are very similar and not statistically different from each other. For men, the rise in the 50/15 wage gap is slightly more pronounced when using hourly instead of monthly wages are used, while the opposite is true for the rise in the 85/50 wage gap. Note, however, that due to the small sample size in the GSOEP, wage gaps are noisily estimated.

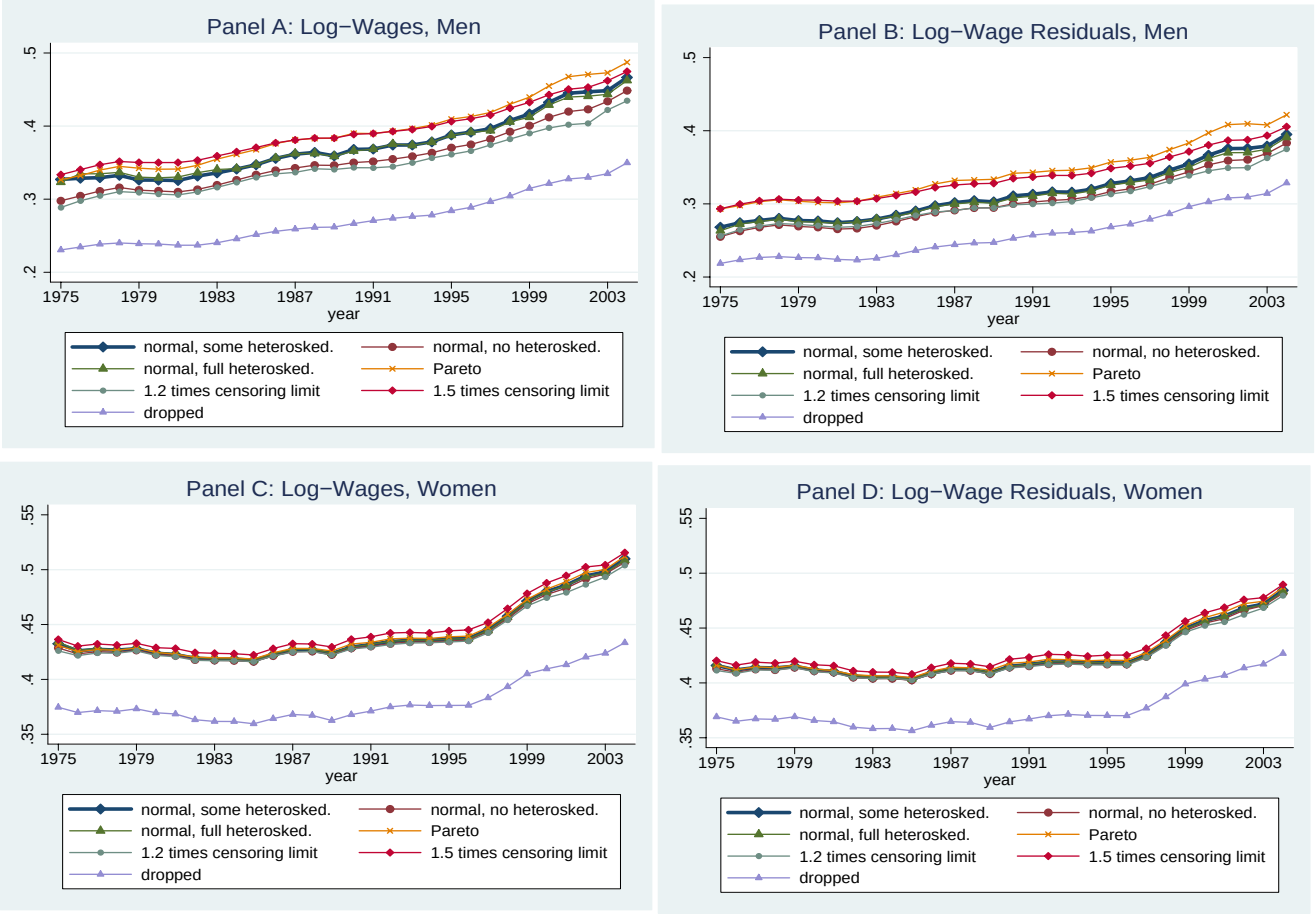
The second set of columns decompose the variance of the log monthly wage into three components: the variance in the log hourly wage, the variance of hours worked, and the covariance between the hourly wage and hours worked. The rise in the standard deviation of log monthly wages is very similar to that of log hourly wages, for both men and women. The variance of hours worked (among full-time workers) is small, and has remained roughly stable over time. The covariance between hours worked and the hourly wage is slightly negative, and has remained stable as well. Overall, these findings suggest that the rise in earnings inequality observed in the IABS is primarily due to a rise in hourly wage inequality.

4 Comparison between IABS and LIAB

In this section, we compare the median as well as the 85-50 and 50-15 wage gaps in the IABS and LIAB. Results can be found in Table A.8, and refer to men only. The wage measure in the LIAB

refers to the 1st of July. In order to facilitate the comparison between the two data sets, the wage measure in the IABS also refers to this date. Not surprisingly, the structure of wages is very similar in the two data sets. One reason for why the two data sets do not give identical results–although they are based on the same social security records–is that the weights in the LIAB were constructed to be representative for the population as a whole, not for our estimation sample of prime-aged men.

Figure I*: The Evolution of Log-Wages and Log-Wage Residuals: Alternative Imputation Methods



Note : The figures in the first column (Panels A and C) plot the evolution of log-wages, while the figures in the second column (Panels B and D) plot the evolution of log-wage residuals separately for men and women, using alternative imputation methods. Our baseline results assume that the error term in the wage regression is normally distributed, and allow for different variances by age and education group (normal, some heterosked.). As a robustness check, we also impute wages by restricting the variance to be the same for all age and education groups (normal, no heterosked.); by allowing variances to differ for each education and age cell (normal, full heterosked.); by assuming that the upper tail of the unconditional log-wage distribution follows a Pareto distribution (pareto); and by replacing censored observations with 1.2 or 1.5 times the censoring limit, respectively (1.2/1.5 times censoring limit). Finally, we drop the top 15% of observations from our sample (dropped).
Source : 2% IABS sample of full-time workers between 21 and 60 years of age.

Table II*: Observed versus Composition-Constant Residual Wage Inequality: Alternative Imputation Methods (Men)

	Normal, some hetero.		Normal, no hetero.	
	1980-1990	1990-2000	1980-1990	1990-2000
		Panel A: Δ 85/50		
observed	.039	.041	.036	.037
1980 X's	.033	.026	.032	.027
1990 X's	.033	.028	.032	.029
2000 X's	.035	.027	.034	.029
		Panel B: Δ 50/15		
observed	.026	.043	.024	.044
1980 X's	.023	.045	.024	.047
1990 X's	.023	.041	.024	.043
2000 X's	.023	.037	.025	.040
	Normal, full hetero.		**1.2 times censoring limit**	
	1980-1990	1990-2000	1980-1990	1990-2000
		Panel A: Δ 85/50		
observed	.037	.040	.030	.023
1980 X's	.032	.027	.034	.024
1990 X's	.031	.029	.037	.027
2000 X's	.033	.028	.043	.027
		Panel B: Δ 50/15		
observed	.024	.044	.023	.041
1980 X's	.021	.045	.021	.044
1990 X's	.021	.041	.020	.040
2000 X's	.022	.038	.019	.035
	1.5 times censoring limit		**pareto**	
	1980-1990	1990-2000	1980-1990	1990-2000
		Panel A: Δ 85/50		
observed	.038	.034	.035	.037
1980 X's	.036	.028	.029	.026
1990 X's	.041	.030	.030	.027
2000 X's	.048	.028	.032	.025
		Panel B: Δ 50/15		
observed	.026	.040	.026	.043
1980 X's	.022	.042	.025	.045
1990 X's	.020	.037	.027	.042
2000 X's	.018	.032	.030	.041

Note: In each panel, the first row reports the observed change in the difference between the 85th and 50th and the 50th and 15th percentile of the residual wage distribution.The next rows show the change that would have prevailed if the age and education distributions were the same as in 1980, 1990, or 2000, respectively. The residuals are obtained from an OLS regression on imputed wages that controls for three education and eight age groups as well as the interactions between these two variables. Wages are imputed in six different ways. Our baseline results assume that the error term in the wage regression is normally distributed and allow for different variances by age and by education (normal, some heterosked.). We also impute wages by restricting the variance to be the same for all age and education groups (normal, no heterosked.); by allowing variances to differ for each education and age cell (normal, full heterosked.); by assuming that the upper tail of the unconditional log-wage distribution follows a Pareto distribution (Pareto); and by replacing censored observations with 1.2 or 1.5 times the censoring limit, respectively.

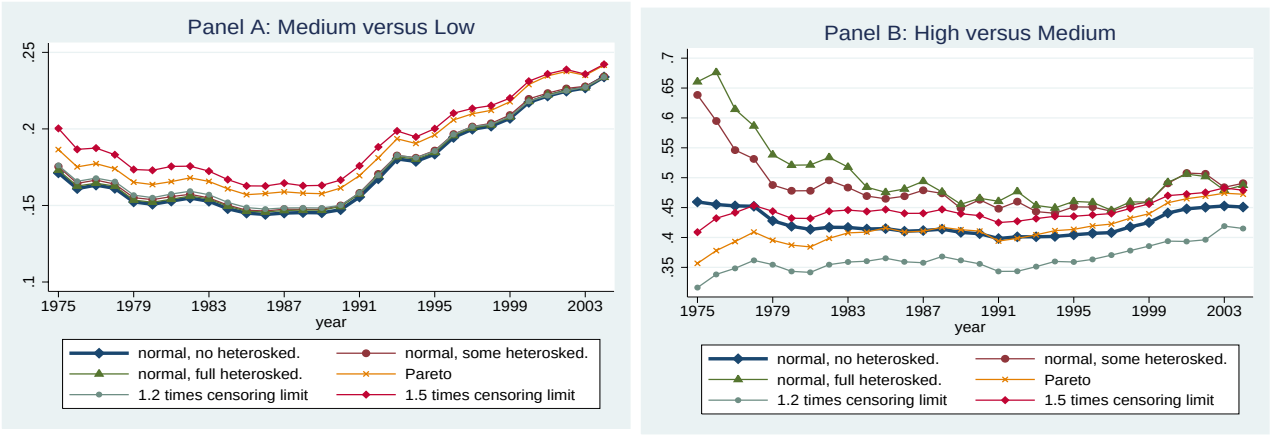
Source: 2% IABS sample for men working full-time between 21 and 60 years of age.

Table IV*: Observed versus Composition-Constant Residual Inequality: The Role of Deunionisation, Alternative Imputation Methods (1995-2004, Men)

	Normal, some hetero.		Normal, no hetero.	
	Unionization only	All	Unionization only	All
		Panel A: Δ 85/50		
observed	.046	.046	.048	.048
1995 X's	.038	.026	.040	.029
2004 X's	.035	.020	.038	.026
		Panel B: Δ 50/15		
observed	.043	.043	.045	.045
1995 X's	.034	.032	.035	.032
2004 X's	.030	.022	.032	.026
	Normal, full hetero.		**1.2 times censoring limit**	
	Unionization only	All	Unionization only	All
		Panel A: Δ 85/50		
observed	.049	.049	.041	.041
1995 X's	.040	.026	.030	.029
2004 X's	.036	.019	.029	.025
		Panel B: Δ 50/15		
observed	.041	.041	.044	.044
1995 X's	.032	.028	.035	.032
2004 X's	.029	.020	.032	.023
	1.5 times censoring limit		**Pareto**	
	Unionization only	All	Unionization only	All
		Panel A: Δ 85/50		
observed	.048	.048	052	.052
1995 X's	.033	.028	.043	030
2004 X's	.032	.027	.039	.023
		Panel B: Δ 50/15		
observed	.041	.041	.040	.040
1995 X's	.032	.027	.032	.032
2004 X's	.028	.018	.028	.025

Note : In each panel, the first row reports the observed changes in the difference between the 85th and 50th (Panel A) and the 50th and 15th (Panel B) percentiles of the residual wage distribution. Column "Unionziation only" shows the changes that would have prevailed if the unionization were the same as in 1995 or 2004, respectively. Column "All" shows the corresponding changes that would have prevailed if unionization as well as the eduation and age distribution were the same as in 1995 or 2004. Wages are imputed in six different ways. Our baseline results assume that the error term in the wage regression is normally distributed, and allow for different variances by age and education (normal, some heterosked.). We also impute wages by restricting the variance to be the same for all age and education groups (normal, no heterosked.); by allowing variances to differ for each education and age cell (normal, full heterosked.); by assuming that the upper tail of the unconditional log-wage distribution follows a Pareto distribution (Pareto); and by replacing censored observations with 1.2 or 1.5 times the censoring limit.
Source: LIAB sample (1995-2004) for men working full-time between 21 and 60 years of age.

Figure V: Education Wage Premiums: Alternative Imputation Methods (Men)**



Note : Panel A plots the fixed-weighted wage ratios of the medium- and low-skilled, while Panel B plots the fixed weighted wage ratios of the high- and medium-skilled, using alternative imputation methods. Our baseline results assume that the error term is normally distributed and restricts the variance to be the same across all age and education groups (normal, no heterosked.). As a robustness check, we also impute wages by allowing variances to differ by age and education (normal, some heterosked.); by allowing variances to differ for each education and age cell (normal, full heterosked.); by assuming that the upper tail of the unconditional log-wage distribution follows a Pareto distribution (Pareto); and by replacing censored observations with 1.2 or 1.5 times the censoring limit, respectively (1.2/1.5 times censoring limit).
Source : 2% IABS sample for men working full-time between 21 and 60 years of age.

Table A.1: Share of (Potentially) Censored Observations in the GSES and IABS, 2001 (Men)

	IABS	GSES
Low		
< 35	1.30%	1.98%
35-44	1.85%	2.94%
>44	2.46%	3.37%
all	1.81%	2.75%
Medium		
<35	3.62%	4.74%
35-44	9.85%	12.43%
>44	14.62%	17.86%
all	9.06%	11.39%
High		
<35	33.41%	32.60%
35-44	58.57%	65.06%
>44	66.67%	74.19%
all	52.98%	56.93%

Note: The first column shows the shares of censored observations in the IABS. The second column shows the shares of observations above the IABS censoring limit in the GSES.
Source: 2% IABS sample for men working full-time between 21 and 60 years of age. GSES sample of men who are working full-time between 21 and 60 years of age.

Table A.2: Which Imputation Method Works Best? Evidence from the GSES (Men)

	Standard Deviation		Residual Wage Gaps		Skill Premiums	
	(1)	(2)	(3)	(4)	(5)	(6)
	log-wages	log-wage residuals	85/50	50/15	medium/low	high/medium
true	.385	.325	.328	.275	.213	.472
normal distribution:						
no heteroskedatcisity	.364	.302	.330	.272	.206	.471
some heteroskedasticity	.367	.304	.331	.274	.207	.483
full heteroskedasticity	.318	.264	.271	.255	.168	.427
1.5 times censoring limit	.413	.348	.315	.290	.228	.515
1.2 times censoring limit	.354	.299	.297	.278	.209	.413
Pareto	.426	.367	.340	.275	.224	.496

Note: The first row compares the true standard deviations in log-wages and log-wage residuals, the true 85/50 and 50/15 residual wage gaps, and the true medium/low and high/medium skill premiums with the imputed ones. To this end, we replace all wage observations in the GSES above the 2001 censoring limit in the IABS with the censoring limit, and then impute censored wages in 6 different ways. We first assume that the error term is normally distributed, and allow for different variances for each age and education group (normal, some heterosked.). As a robustness check, we restrict the variance to be the same across all age and education groups (normal, no heterosked.). We also impute wages by allowing variances to differ for each education and age cell (normal, full heterosked.); by assuming that the upper tail of the unconditional log-wage distribution follows a Pareto distribution (Pareto); and by replacing censored observations with 1.2 or 1.5 times the censoring limit, respectively (1.2/1.5 times censoring limit).

Source: GSES sample for men working full-time between 21 and 60 years of age.

Table A.3: Comparison between the IABS and GSOEP: Upper versus Lower Tail Inequality, Men

	85/50				50/15			
	GSOEP			IABS	GSOEP			IABS
	full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.405	.360	.389	.365	.310	.275	.282	.249
	(.016)	(.018)	(.020)	(.001)	(.016)	(.019)	(.012)	(.001)
1985	.437	.405	.409	.375	.298	.266	.296	.257
	(.022)	(.018)	(.017)	(.001)	(.019)	(.016)	(.014)	(.001)
1986	.446	.416	.425	.380	.304	.256	.283	.261
	(.011)	(.025)	(.020)	(.001)	(.017)	(.014)	(.011)	(.001)
1987	.446	.405	.399	.385	.309	.288	.308	.263
	(.012)	(.031)	(.029)	(.002)	(.020)	(.015)	(.016)	(.001)
1988	.440	.416	.430	.391	.293	.278	.292	.267
	(.024)	(.017)	(.018)	(.001)	(.015)	(.013)	(.011)	(.001)
1989	.396	.357	.387	.395	.297	.297	.294	.269
	(.019)	(.018)	(.016)	(.001)	(.008)	(.013)	(.012)	(.001)
1990	.421	.405	.413	.393	.280	.255	.295	.272
	(.019)	(.019)	(.016)	(.001)	(.017)	(.017)	(.014)	(.001)
1991	.423	.370	.395	.388	.288	.270	.265	.277
	(.023)	(.017)	(.017)	(.001)	(.016)	(.018)	(.016)	(.001)
1992	.420	.405	.404	.386	.267	.255	.276	.279
	(.018)	(.019)	(.020)	(.001)	(.018)	(.014)	(.013)	(.001)
1993	.443	.437	.433	.399	.295	.272	.281	.278
	(.019)	(.020)	(.025)	(.001)	(.016)	(.014)	(.016)	(.001)
1994	.464	.430	.442	.406	.288	.248	.273	.281
	(.022)	(.030)	(.024)	(.001)	(.019)	(.018)	(.012)	(.001)
1995	.449	.435	.478	.410	.280	.261	.286	.289
	(.023)	(.030)	(.032)	(.001)	(.014)	(.018)	(.015)	(.001)
1996	.494	.462	.459	.415	.324	.302	.305	.292
	(.027)	(.033)	(.034)	(.001)	(.018)	(.016)	(.016)	(.001)
1997	.451	.448	.472	.424	.316	.302	.300	.296
	(.025)	(.030)	(.038)	(.001)	(.016)	(.019)	(.022)	(.001)
1998	.510	.463	.467	.433	.316	.307	.332	.304
	(.027)	(.030)	(.020)	(.001)	(.014)	(.016)	(.016)	(.001)
1999	.488	.467	.478	.443	.338	.324	.353	.315
	(.025)	(.027)	(.029)	(.001)	(.017)	(.019)	(.017)	(.001)
2000	.482	.470	.501	.445	.357	.357	.343	.328
	(.019)	(.020)	(.019)	(.001)	(.011)	(.013)	(.016)	(.001)
2001	.480	.470	.463	.451	.338	.329	.337	.334
	(.021)	(.026)	(.025)	(.001)	(.019)	(.011)	(.014)	(.001)
2002	.474	.454	.465	.452	.388	.381	.380	.337
	(.023)	(.029)	(.030)	(.001)	(.021)	(.023)	(.020)	(.001)

Note: The table compares the 85/50 and 50/15 wage gaps in the GSOEP and IABS. The first set of columns (full sample, no bonuses) includes the self-employed and the civil servants. The wage measure is the gross monthly wage, and does not include bonuses, holiday or Christmas money. The second set of columns (IABS sample, no bonuses) drops the self-employed and civil servants from the sample and uses the same wage measure. The third set of columns (IABS sample, plus bonuses) is based on the same sample, but includes bonuses, Christmas and holiday money in the wage measure. This sample and wage measure is most similar to those in the IABS. The fourth set of columns refers to our IABS sample. Standard errors in parentheses are bootstrapped with 100 replications.

Source: GSOEP sample for men working full-time between 21 and 60 years of age. 2% IABS sample for men working full-time between 21 and 60 years of age.

Table A.4: Comparison between the IABS and GSOEP: Upper versus Lower Tail Inequality, Women

	85/50				50/15			
	GSOEP			IABS	GSOEP			IABS
	full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.383	.310	.339	.337	.383	.318	.330	.448
	(.028)	(.025)	(.021)	(.001)	(.028)	(.029)	(.024)	(.003)
1985	.353	.330	.343	.338	.353	.363	.360	.446
	(.028)	(.027)	(.025)	(.001)	(.028)	(.029)	(.021)	(.002)
1986	.324	.361	.361	.342	.324	.302	.349	.452
	(.026)	(.024)	(.023)	(.002)	(.026)	(.027)	(.019)	(.002)
1987	.336	.377	.377	.343	.336	.288	.343	.454
	(.027)	(.024)	(.024)	(.002)	(.027)	(.025)	(.022)	(.002)
1988	.310	.336	.340	.347	.310	.310	.327	.448
	(.025)	(.020)	(.029)	(.001)	(.025)	(.026)	(.024)	(.002)
1989	.348	.347	.336	.351	.348	.332	.342	.439
	(.026)	(.018)	(.023)	(.002)	(.026)	(.022)	(.022)	(.002)
1990	.336	.326	.333	.354	.336	.300	.344	.446
	(.021)	(.025)	(.025)	(.002)	(.021)	(.029)	(.026)	(.002)
1991	.336	.328	.313	.353	.336	.336	.357	.453
	(.025)	(.023)	(.022)	(.002)	(.025)	(.019)	(.018)	(.002)
1992	.327	.334	.332	.351	.327	.310	.333	.457
	(.024)	(.022)	(.023)	(.001)	(.024)	(.023)	(.025)	(.002)
1993	.330	.363	.346	.351	.330	.330	.323	.459
	(.021)	(.021)	(.026)	(.002)	(.021)	(.026)	(.024)	(.003)
1994	.298	.310	.350	.350	.298	.302	.304	.456
	(.023)	(.021)	(.019)	(.002)	(.023)	(.021)	(.021)	(.002)
1995	.336	.309	.328	.351	.336	.322	.349	.459
	(.022)	(.020)	(.023)	(.002)	(.022)	(.027)	(.028)	(.003)
1996	.325	.319	.340	.354	.325	.315	.321	.461
	(.022)	(.026)	(.029)	(.002)	(.022)	(.024)	(.026)	(.003)
1997	.353	.321	.342	.359	.353	.315	.333	.473
	(.026)	(.030)	(.027)	(.002)	(.026)	(.028)	(.022)	(.003)
1998	.342	.336	.363	.366	.342	.336	.342	.486
	(.034)	(.032)	(.033)	(.002)	(.034)	(.033)	(.027)	(.003)
1999	.331	.370	.360	.378	.331	.309	.354	.505
	.031	.027	.026	.002	.031	.032	.028	.003
2000	.357	.361	.364	.388	.357	.333	.361	.516
	(.020)	(.028)	(.025)	(.002)	(.020)	(.018)	(.020)	(.003)
2001	.393	.360	.357	.397	.393	.357	.356	.523
	(.021)	(.027)	(.025)	(.002)	(.021)	(.021)	(.020)	(.002)
2002	.360	.357	.353	.400	.360	.336	.375	.538
	(.025)	(.023)	(.020)	(.002)	(.025)	(.016)	(.025)	(.003)

Note: The table compares the 85/50 and 50/15 wage gaps in the GSOEP and IABS. The first set of columns (full sample, no bonuses) includes the self-employed and the civil servants. The wage measure is the gross monthly wage, and does not include bonuses, holiday or Christmas money. The second set of columns (IABS sample, no bonuses) drops the self-employed and civil servants from the sample and uses the same wage measure. The third set of columns (IABS sample, plus bonuses) is based on the same sample, but includes bonuses, Christmas and holiday money in the wage measure. This sample and wage measure is most similar to those in the IABS. The fourth set of columns refers to our IABS sample. Standard errors in parentheses are bootstrapped with 100 replications.

Source: GSOEP sample for women working full-time between 21 and 60 years of age. 2% IABS sample for women working full-time between 21 and 60 years of age.

Table A.5: Comparison between the IABS and GSOEP: Standard Deviation and Education Wage Premiums, Men

	A: Standard Deviation Log-Wages				B: Standard Deviation Log-Wage Residuals			
	GSOEP			IABS	GSOEP			IABS
	full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.384	.345	.354	.341	.312	.277	.283	.285
	(.011)	(.009)	(.009)	(.001)	(.009)	(.007)	(.008)	(.001)
1985	.386	.360	.367	.347	.320	.295	.296	.290
	(.008)	(.009)	(.010)	(.001)	(.010)	(.009)	(.010)	(.001)
1986	.387	.354	.367	.356	.310	.282	.292	.298
	(.009)	(.010)	(.010)	(.001)	(.008)	(.007)	(.009)	(.001)
1987	.395	.369	.374	.362	.323	.299	.301	.302
	(.010)	(.012)	(.009)	(.001)	(.009)	(.009)	(.008)	(.001)
1988	.386	.366	.374	.364	.307	.295	.299	.304
	(.010)	(.010)	(.011)	(.001)	(.007)	(.009)	(.009)	(.001)
1989	.379	.348	.355	.359	.311	.288	.290	.303
	(.010)	(.010)	(.009)	(.001)	(.010)	(.009)	(.008)	(.001)
1990	.384	.348	.362	.368	.310	.280	.290	.311
	(.018)	(.013)	(.015)	(.001)	(.015)	(.008)	(.009)	(.001)
1991	.372	.347	.354	.369	.301	.282	.284	.313
	(.008)	(.009)	(.011)	(.001)	(.007)	(.006)	(.007)	(.001)
1992	.369	.338	.344	.374	.299	.268	.275	.316
	(.014)	(.011)	(.011)	(.001)	(.017)	(.008)	(.007)	(.001)
1993	.387	.362	.364	.374	.314	.290	.283	.316
	(.011)	(.010)	(.010)	(.001)	(.009)	(.009)	(.010)	(.001)
1994	.381	.359	.372	.378	.312	.286	.293	.320
	(.014)	(.014)	(.012)	(.001)	(.011)	(.012)	(.010)	(.001)
1995	.386	.361	.378	.388	.324	.297	.309	.328
	(.013)	(.011)	(.011)	(.001)	(.012)	(.012)	(.012)	(.001)
1996	.407	.383	.388	.392	.341	.311	.319	.331
	(.013)	(.012)	(.012)	(.001)	(.015)	(.012)	(.014)	(.001)
1997	.390	.379	.388	.397	.314	.300	.307	.336
	(.010)	(.011)	(.012)	(.001)	(.009)	(.010)	(.010)	(.001)
1998	.406	.387	.395	.408	.335	.317	.325	.346
	(.009)	(.011)	(.012)	(.001)	(.009)	(.011)	(.010)	(.001)
1999	.423	.399	.411	.417	.353	.329	.335	.355
	(.014)	(.011)	(.012)	(.001)	(.012)	(.010)	(.011)	(.001)
2000	.428	.412	.420	.433	.365	.343	.348	.367
	(.009)	(.010)	(.012)	(.001)	(.008)	(.009)	(.009)	(.001)
2001	.429	.408	.406	.445	.366	.343	.344	.376
	(.009)	(.010)	(.012)	(.001)	(.009)	(.009)	(.010)	(.001)
2002	.436	.412	.417	.447	.378	.352	.359	.376
	(.009)	(.009)	(.010)	(.001)	(.008)	(.008)	(.009)	(.001)

	C: Education Wage Premiums: Medium versus Low				D: Education Wage Premiums: High versus Medium			
	GSOEP			IABS	GSOEP			IABS
	full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.246	.190	.178	.148	.357	.418	.441	.414
	(.026)	(.016)	(.016)	(.002)	(.040)	(.041)	(.045)	(.003)
1985	.203	.187	.176	.145	.329	.410	.411	.415
	(.026)	(.026)	(.018)	(.002)	(.037)	(.044)	(.047)	(.003)
1986	.201	.177	.186	.144	.356	.398	.413	.410
	(.022)	(.019)	(.017)	(.002)	(.022)	(.033)	(.036)	(.003)
1987	.167	.154	.155	.145	.393	.413	.431	.412
	(.019)	(.022)	(.021)	(.002)	(.049)	(.038)	(.054)	(.003)
1988	.183	.158	.160	.145	.356	.389	.409	.414
	(.017)	(.017)	(.021)	(.002)	(.028)	(.038)	(.049)	(.003)
1989	.149	.148	.159	.145	.390	.415	.439	.409
	(.018)	(.019)	(.020)	(.002)	(.024)	(.032)	(.035)	(.003)
1990	.185	.194	.209	.147	.402	.415	.415	.406
	(.033)	(.031)	(.033)	(.002)	(.046)	(.054)	(.067)	(.003)
1991	.156	.179	.174	.156	.384	.418	.509	.398
	(.025)	(.022)	(.024)	(.002)	(.035)	(.051)	(.054)	(.003)
1992	.133	.146	.137	.167	.367	.390	.394	.401
	(.020)	(.017)	(.022)	(.002)	(.047)	(.055)	(.058)	(.003)
1993	.138	.145	.130	.180	.412	.452	.513	.401
	(.022)	(.023)	(.021)	(.002)	(.028)	(.044)	(.034)	(.003)
1994	.135	.154	.159	.179	.392	.442	.463	.402
	(.023)	(.019)	(.023)	(.002)	(.025)	(.038)	(.043)	(.003)
1995	.144	.166	.176	.184	.363	.387	.393	.404
	(.022)	(.021)	(.024)	(.002)	(.032)	(.033)	(.037)	(.003)
1996	.205	.216	.209	.194	.356	.398	.411	.407
	(.023)	(.022)	(.025)	(.002)	(.029)	(.030)	(.038)	(.003)
1997	.151	.160	.165	.200	.356	.414	.424	.408
	(.021)	(.023)	(.026)	(.002)	(.031)	(.043)	(.040)	(.004)
1998	.161	.175	.183	.202	.297	.416	.426	.418
	(.023)	(.026)	(.025)	(.003)	(.038)	(.043)	(.043)	(.003)
1999	.183	.188	.198	.207	.363	.435	.465	.425
	(.026)	(.027)	(.032)	(.002)	(.048)	(.038)	(.045)	(.004)
2000	.161	.160	.145	.217	.304	.348	.351	.441
	(.020)	(.021)	(.018)	(.002)	(.034)	(.039)	(.041)	(.003)
2001	.147	.150	.146	.221	.292	.305	.279	.448
	(.022)	(.023)	(.020)	(.002)	(.032)	(.037)	(.040)	(.003)
2002	.155	.156	.151	.225	.341	.371	.372	.451
	(.025)	(.023)	(.027)	(.003)	(.028)	(.031)	(.037)	(.003)

Note: The table compares the standard deviation of log-wages, the standard deviation of log-wage residuals, the medium/low wage premium and the high/medium wage premium in the GSOEP and IABS. The first set of columns (full sample, no bonuses) includes the self-employed and the civil servants. The wage measure is the gross monthly wage, and does not include bonuses, holiday or Christmas money. The second set of columns (IABS sample, no bonuses) drops the self-employed and civil servants from the sample and uses the same wage measure. The third set of columns (IABS sample, plus bonuses) is based on the same sample, but includes bonuses, Christmas and holiday money in the wage measure. This sample and wage measure is most similar to those in the IABS. The fourth set of columns refers to our IABS sample. Standard errors in parentheses are bootstrapped with 100 replications.

Source: GSOEP sample for men working full-time between 21 and 60 years of age. 2% IABS sample for men working full-time between 21 and 60 years of age.

Table A.6: Comparison between the IABS and GSOEP: Standard Deviation and Education Wage Premiums, Women

	A: Standard Deviation Log-Wages				B: Standard Deviation Log-Wage Residuals			
	GSOEP			IABS	GSOEP			IABS
	full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample, no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.373	.340	.351	.419	.324	.299	.310	.405
	(.011)	(.011)	(.011)	(.001)	(.010)	(.010)	(.011)	(.001)
1985	.368	.340	.358	.417	.320	.300	.315	.404
	(.012)	(.012)	(.011)	(.001)	(.015)	(.011)	(.012)	(.001)
1986	.384	.349	.364	.423	.337	.304	.314	.409
	(.018)	(.013)	(.012)	(.001)	(.021)	(.010)	(.011)	(.001)
1987	.378	.344	.369	.427	.325	.296	.320	.413
	(.014)	(.013)	(.014)	(.001)	(.013)	(.012)	(.013)	(.001)
1988	.364	.348	.351	.427	.304	.299	.306	.412
	(.014)	(.013)	(.013)	(.001)	(.013)	(.013)	(.012)	(.001)
1989	.360	.341	.353	.424	.301	.297	.302	.409
	(.012)	(.013)	(.012)	(.001)	(.010)	(.011)	(.012)	(.001)
1990	.359	.350	.357	.429	.302	.301	.309	.415
	(.014)	(.016)	(.014)	(.001)	(.012)	(.010)	(.013)	(.001)
1991	.373	.351	.353	.432	.315	.303	.308	.417
	(.015)	(.016)	(.015)	(.001)	(.013)	(.015)	(.011)	(.001)
1992	.383	.356	.373	.435	.337	.319	.335	.419
	(.019)	(.019)	(.020)	(.001)	(.016)	(.015)	(.014)	(.001)
1993	.371	.359	.356	.436	.328	.323	.323	.419
	(.014)	(.015)	(.014)	(.001)	(.013)	(.014)	(.013)	(.001)
1994	.355	.338	.359	.436	.312	.305	.323	.418
	(.015)	(.014)	(.017)	(.001)	(.012)	(.015)	(.014)	(.001)
1995	.371	.354	.375	.437	.327	.319	.336	.419
	(.017)	(.020)	(.023)	(.001)	(.014)	(.015)	(.018)	(.001)
1996	.363	.341	.357	.437	.317	.298	.319	.418
	(.014)	(.014)	(.015)	(.001)	(.012)	(.011)	(.014)	(.001)
1997	.389	.356	.362	.445	.340	.309	.313	.425
	(.014)	(.016)	(.017)	(.001)	(.014)	(.013)	(.012)	(.001)
1998	.376	.354	.374	.457	.336	.325	.339	.437
	(.014)	(.016)	(.013)	(.001)	(.016)	(.016)	(.019)	(.001)
1999	.369	.354	.363	.471	.323	.309	.320	.450
	(.011)	(.014)	(.012)	(.001)	(.013)	(.013)	(.013)	(.001)
2000	.394	.364	.378	.480	.346	.326	.344	.457
	(.010)	(.011)	(.013)	(.001)	(.012)	(.010)	(.013)	(.001)
2001	.420	.399	.404	.486	.381	.372	.369	.461
	(.015)	(.018)	(.020)	(.001)	(.017)	(.018)	(.017)	(.001)
2002	.383	.358	.364	.494	.338	.323	.325	.468
	(.009)	(.010)	(.008)	(.001)	(.008)	(.010)	(.007)	(.001)

	C: Education Wage Premiums: Medium versus Low				D: Education Wage Premiums: High versus Medium			
	GSOEP			IABS	GSOEP			IABS
	full sample no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses		full sample no bonuses	IABS sample, no bonuses	IABS sample, plus bonuses	
1984	.232	.223	.240	.144	.315	.237	.235	.381
	(.028)	(.026)	(.031)	(.003)	(.057)	(.073)	(.093)	(.010)
1985	.217	.232	.230	.144	.366	.320	.344	.370
	(.035)	(.033)	(.032)	(.003)	(.047)	(.065)	(.067)	(.009)
1986	.198	.213	.240	.136	.343	.272	.313	.372
	(.032)	(.032)	(.037)	(.004)	(.047)	(.059)	(.060)	(.008)
1987	.177	.202	.196	.141	.391	.323	.361	.365
	(.032)	(.032)	(.035)	(.003)	(.045)	(.041)	(.050)	(.009)
1988	.210	.202	.218	.141	.348	.309	.298	.367
	(.030)	(.031)	(.030)	(.003)	(.037)	(.060)	(.068)	(.008)
1989	.180	.169	.204	.143	.376	.342	.364	.360
	(.034)	(.030)	(.033)	(.003)	(.051)	(.046)	(.056)	(.008)
1990	.216	.213	.223	.150	.378	.370	.354	.360
	(.036)	(.034)	(.041)	(.003)	(.047)	(.057)	(.058)	(.008)
1991	.195	.201	.206	.157	.405	.408	.412	.363
	(.032)	(.031)	(.031)	(.003)	(.046)	(.071)	(.069)	(.007)
1992	.179	.173	.190	.163	.222	.241	.251	.360
	(.040)	(.037)	(.038)	(.003)	(.061)	(.088)	(.089)	(.007)
1993	.204	.198	.187	.172	.274	.291	.306	.355
	(.039)	(.038)	(.039)	(.003)	(.050)	(.076)	(.094)	(.007)
1994	.151	.145	.152	.179	.330	.300	.270	.353
	(.034)	(.034)	(.042)	(.003)	(.061)	(.067)	(.084)	(.007)
1995	.160	.156	.172	.185	.246	.241	.259	.353
	(.040)	(.039)	(.043)	(.003)	(.063)	(.069)	(.084)	(.006)
1996	.214	.210	.200	.193	.278	.310	.301	.350
	(.034)	(.034)	(.046)	(.004)	(.055)	(.073)	(.094)	(.006)
1997	.160	.157	.142	.198	.315	.357	.362	.352
	(.035)	(.039)	(.038)	(.003)	(.051)	(.071)	(.088)	(.005)
1998	.112	.100	.123	.199	.258	.278	.230	.354
	(.040)	(.034)	(.042)	(.004)	(.062)	(.053)	(.071)	(.006)
1999	.139	.155	.144	.205	.263	.292	.247	.367
	(.040)	(.038)	(.047)	(.004)	(.037)	(.052)	(.055)	(.005)
2000	.183	.185	.180	.211	.297	.249	.229	.377
	(.051)	(.045)	(.049)	(.004)	(.026)	(.030)	(.036)	(.006)
2001	.244	.241	.276	.218	.268	.227	.256	.386
	(.050)	(.050)	(.062)	(.005)	(.040)	(.047)	(.041)	(.005)
2002	.199	.205	.244	.228	.252	.264	.264	.390
	(.037)	(.036)	(.039)	(.005)	(.032)	(.032)	(.037)	(.006)

Note: The table compares the standard deviations of log-wages, the standard deviation of log-wage residuals, the medium/low wage premium and the high/medium wage premium in the GSOEP and IABS. The first set of columns (full sample, no bonuses) includes the self-employed and the civil servants. The wage measure is the gross monthly wage, and does not include bonuses, holiday or Christmas money. The second set of columns (IABS sample, no bonuses) drops the self-employed and civil servants from the sample and uses the same wage measure. The third set of columns (IABS sample, plus bonuses) is based on the same sample, but includes bonuses, Christmas and holiday money in the wage measure. This sample and wage measure is most similar to those in the IABS. The fourth set of columns refers to our IABS sample. Standard errors in parentheses are bootstrapped with 100 replications.

Source: GSOEP sample for women working full-time between 21 and 60 years of age. 2% IABS sample for women working full-time between 21 and 60 years of age.

Table A.7: Hourly Wage versus Earnings Inequality

Panel A: Men, GSOEP

	Wage Gaps								Variance					
	85/50				50/15									
	Earnings		Hourly Wage		Earnings		Hourly Wage		Earnings		Hourly Wage		Hours		Cov(Hours, Hourly Wage)	
1984	.389	(.020)	.371	(.016)	.282	(.012)	.305	(.014)	.125	(.007)	.119	(.007)	.018	(.001)	-.006	(.001)
1985	.409	(.017)	.390	(.020)	.296	(.014)	.314	(.017)	.135	(.006)	.129	(.006)	.017	(.001)	-.006	(.002)
1986	.425	(.020)	.391	(.022)	.283	(.011)	.310	(.014)	.135	(.007)	.126	(.007)	.018	(.001)	-.005	(.002)
1987	.399	(.029)	.365	(.023)	.308	(.015)	.322	(.012)	.140	(.007)	.133	(.007)	.017	(.001)	-.005	(.002)
1988	.430	(.017)	.361	(.019)	.292	(.011)	.314	(.017)	.140	(.008)	.131	(.008)	.017	(.001)	-.004	(.002)
1989	.387	(.016)	.365	(.018)	.294	(.012)	.292	(.010)	.126	(.007)	.115	(.006)	.019	(.001)	-.004	(.002)
1990	.413	(.016)	.382	(.019)	.295	(.014)	.296	(.014)	.131	(.010)	.135	(.019)	.021	(.002)	-.012	(.006)
1991	.395	(.017)	.370	(.015)	.265	(.016)	.304	(.013)	.125	(.007)	.124	(.009)	.022	(.002)	-.010	(.003)
1992	.404	(.020)	.360	(.017)	.276	(.013)	.303	(.013)	.118	(.006)	.112	(.006)	.019	(.001)	-.006	(.002)
1993	.433	(.025)	.398	(.019)	.281	(.016)	.285	(.011)	.132	(.007)	.122	(.007)	.019	(.001)	-.004	(.002)
1994	.442	(.024)	.389	(.017)	.273	(.012)	.307	(.013)	.139	(.009)	.128	(.007)	.022	(.002)	-.006	(.003)
1995	.478	(.032)	.403	(.025)	.286	(.015)	.296	(.016)	.143	(.009)	.133	(.010)	.021	(.001)	-.006	(.003)
1996	.459	(.034)	.415	(.023)	.305	(.016)	.339	(.019)	.150	(.009)	.145	(.009)	.022	(.002)	-.008	(.002)
1997	.472	(.038)	.385	(.023)	.300	(.022)	.343	(.017)	.150	(.009)	.140	(.008)	.022	(.001)	-.006	(.002)
1998	.467	(.020)	.399	(.021)	.332	(.016)	.338	(.015)	.156	(.009)	.142	(.009)	.024	(.002)	-.005	(.002)
1999	.478	(.029)	.416	(.025)	.353	(.017)	.363	(.022)	.169	(.009)	.154	(.011)	.022	(.001)	-.004	(.003)
2000	.501	(.019)	.424	(.019)	.343	(.016)	.364	(.015)	.176	(.010)	.162	(.008)	.023	(.001)	-.004	(.003)
2001	.463	(.025)	.396	(.018)	.337	(.014)	.362	(.014)	.165	(.008)	.155	(.008)	.022	(.001)	-.004	(.003)
2002	.465	(.030)	.389	(.021)	.380	(.020)	.415	(.019)	.174	(.007)	.164	(.008)	.022	(.001)	-.007	(.003)

Panel B: Women, GSOEP

	Wage Gaps								Variance					
	85/50				50/15									
	Earnings		Hourly Wage		Earnings		Hourly Wage		Earnings		Hourly Wage		Hours		Cov(Hours, Hourly Wage)	
1984	.339	(.025)	.356	(.016)	.330	(.029)	.362	(.024)	.123	(.008)	.126	(.008)	.013	(.002)	-.008	(.002)
1985	.343	(.027)	.343	(.021)	.360	(.029)	.390	(.027)	.128	(.009)	.126	(.008)	.010	(.001)	-.004	(.002)
1986	.361	(.024)	.368	(.023)	.349	(.027)	.350	(.020)	.132	(.009)	.128	(.010)	.011	(.001)	-.003	(.002)
1987	.377	(.024)	.331	(.026)	.343	(.025)	.362	(.026)	.136	(.009)	.132	(.012)	.011	(.001)	-.003	(.004)
1988	.340	(.020)	.337	(.022)	.327	(.026)	.311	(.023)	.123	(.010)	.123	(.010)	.013	(.002)	-.006	(.003)
1989	.336	(.018)	.337	(.020)	.343	(.022)	.345	(.020)	.124	(.008)	.122	(.011)	.012	(.001)	-.005	(.004)
1990	.333	(.025)	.349	(.026)	.344	(.029)	.341	(.028)	.127	(.011)	.131	(.011)	.011	(.001)	-.007	(.002)
1991	.313	(.023)	.309	(.019)	.357	(.019)	.353	(.019)	.125	(.012)	.129	(.017)	.013	(.003)	-.008	(.006)
1992	.332	(.022)	.310	(.025)	.333	(.023)	.344	(.026)	.139	(.013)	.142	(.015)	.012	(.001)	-.007	(.003)
1993	.346	(.021)	.344	(.024)	.323	(.026)	.330	(.027)	.127	(.011)	.124	(.012)	.009	(.001)	-.004	(.003)
1994	.350	(.021)	.309	(.020)	.304	(.021)	.323	(.021)	.129	(.014)	.134	(.019)	.011	(.002)	-.008	(.006)
1995	.328	(.020)	.322	(.021)	.349	(.027)	.340	(.025)	.141	(.013)	.152	(.020)	.013	(.002)	-.012	(.006)
1996	.340	(.026)	.331	(.026)	.321	(.024)	.324	(.025)	.128	(.012)	.123	(.012)	.012	(.001)	-.003	(.003)
1997	.342	(.030)	.352	(.023)	.333	(.028)	.329	(.031)	.131	(.013)	.129	(.012)	.014	(.002)	-.006	(.003)
1998	.363	(.032)	.345	(.018)	.342	(.033)	.361	(.025)	.140	(.012)	.135	(.011)	.014	(.001)	-.005	(.003)
1999	.360	(.027)	.349	(.025)	.354	(.032)	.354	(.024)	.131	(.010)	.128	(.011)	.014	(.002)	-.005	(.003)
2000	.364	(.028)	.348	(.014)	.361	(.018)	.371	(.018)	.143	(.009)	.140	(.009)	.016	(.002)	-.006	(.003)
2001	.357	(.027)	.360	(.018)	.356	(.021)	.388	(.026)	.163	(.016)	.152	(.014)	.015	(.001)	-.002	(.002)
2002	.353	(.023)	.354	(.024)	.375	(.016)	.362	(.029)	.132	(.007)	.125	(.006)	.015	(.002)	-.004	(.002)

Note: The table compares trends in hourly wage and earnings inequality in the GSOEP, for men (Panel A) and women (Panel B). Our sample excludes the self-employed and the civil servants, and our wage measure includes bonuses, holiday and Christmas money, etc. The first four columns compare the 85/50 and 50/15 earnings (i.e. the monthly wage) and hourly wage gaps (in logs). The remaining columns decmpose the variance of log-earnings into three components: the variance of the log hourly wage, the variance of hours, and the covariance between hours and the hourly wage. Standard errors in parentheses are bootstrapped with 100 replications.

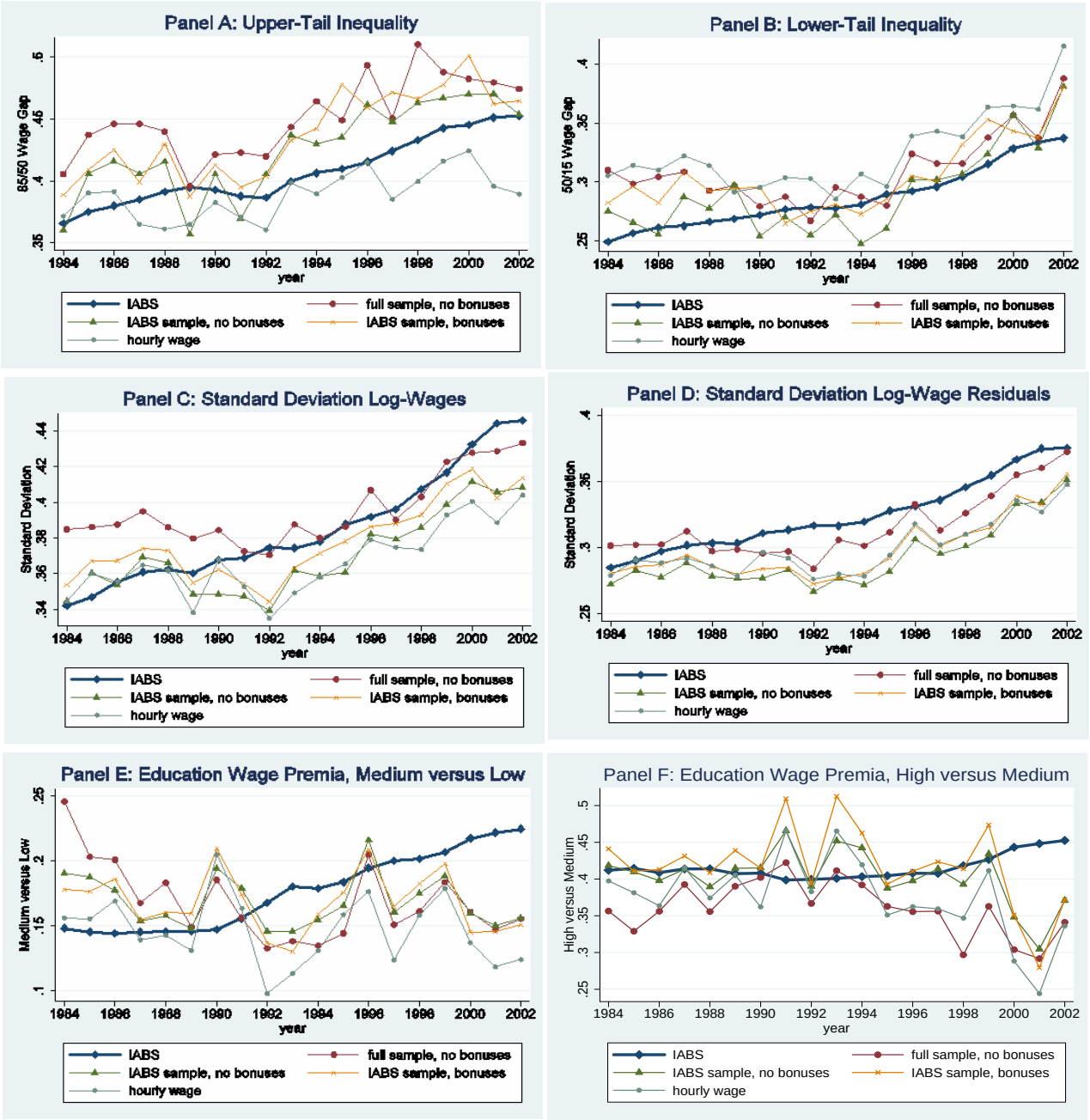
Source: GSOEP sample of full-time workers between 21 and 60 years of age.

Table A.8 Comparison Between LIAB And The IABS

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
					median					
LIAB	5.075	5.074	5.061	5.066	5.079	5.087	5.088	5.088	5.092	5.097
IABS	5.084	5.082	5.071	5.080	5.092	5.096	5.095	5.102	5.108	5.097
					85-50					
LIAB	.395	.399	.410	.417	.418	.426	.435	.441	.459	.462
IABS	.409	.414	.423	.431	.441	.443	.451	.452	.472	.477
					50-15					
LIAB	.280	.277	.291	.285	.291	.316	.313	.323	.325	.344
IABS	.287	.289	.293	.301	.311	.324	.329	.334	.342	.360

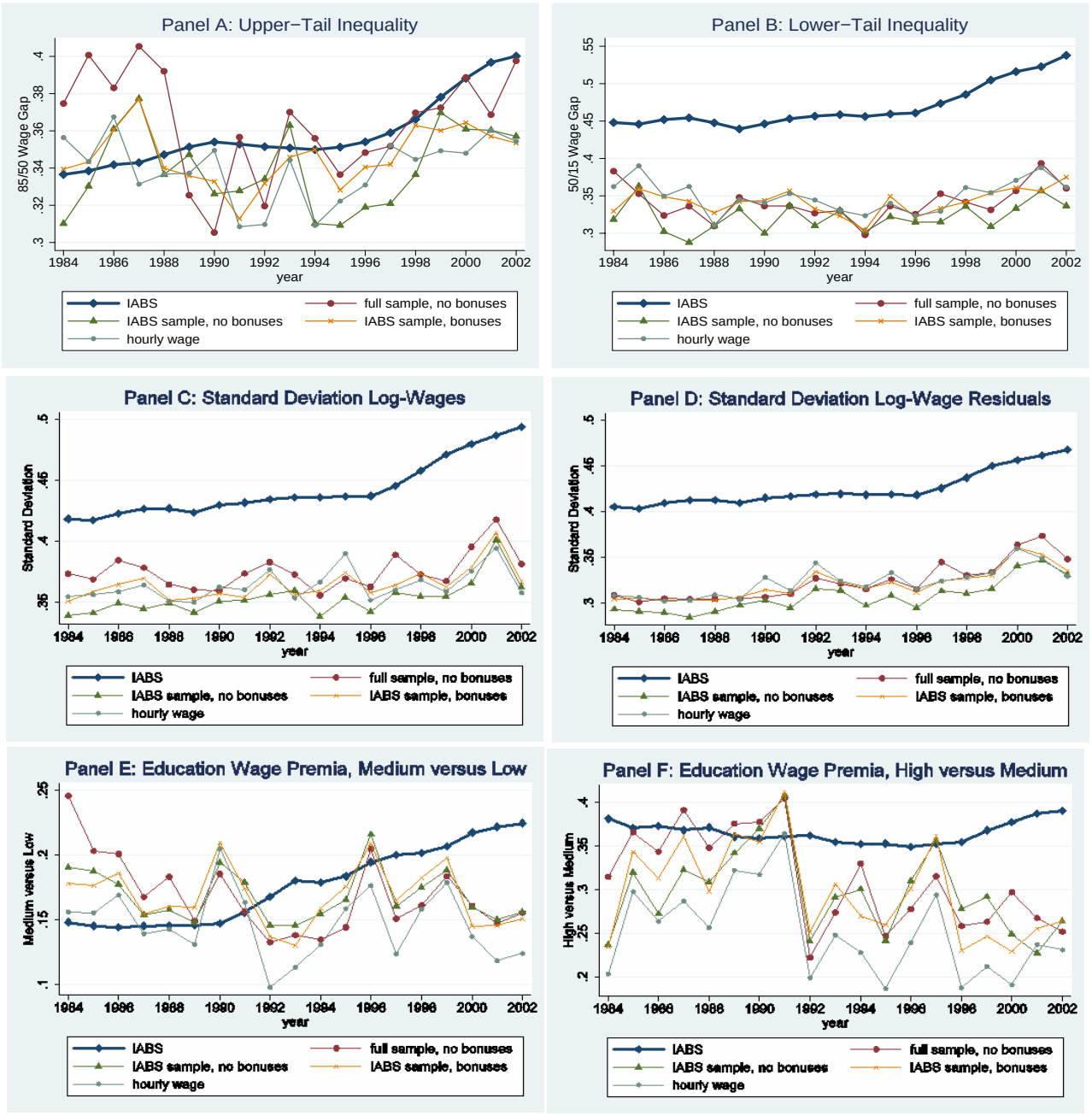
Note: The table compares the median and the 85-50 and 50-15 wage gap in the LIAB and the IABS. Wages in the LIAB data refer to the 1st of July. To facilitate the comparison between the two data sets, wages in the IABS refer to the same date.
Source: 2% IABS Sample (1995-2004) for men working full-time between 21 and 60 years of age. LIAB (1995-2004) for men working full-time men between 21 and 60 years of age.

Figure A.1: Inequality Trends in the IABS and GSOEP, Men



Note: The figure compares inequality trends in the GSOEP and IABS. Panels A and B plot the 85/50 and 50/15 wage gaps in the two data sets. Panel C and D show the evolution of the standard deviation of log-wages and log-wage residuals in the IABS and GSOEP. Panels E and F display the age-constant medium/low and high/medium wage premiums in the two data sets. In the GSOEP, we report findings for four samples: "Full sample, no bonuses" includes the civil servants and self-employed in the sample. The wage measure is the monthly wage excluding bonuses, holdiay and Christmas money. "IABS sample, no bonuses" excludes the civil servants and self-employed from the analysis. "IABS sample, bonuses" is based on the same sample, but includes bonuses, holiday and Christmas money in the wage measure. This specification is most similar to the one in the IABS. "Hourly wage" also excludes the self-employed and civil servants. The wage measure is the hourly wage including bonuses, holiday and Christmas money.
Source: GSOEP sample for men working full-time between 21 and 60 years of age. 2% IABS sample for men working full-time between 21 and 60 years of age.

Figure A.2: Inequality Trends in the IABS and GSOEP, Women



Note : The figure compares inequality trends in the GSOEP and IABS. Panels A and B plot the 85/50 and 50/15 wage gap in the two data sets. Panels C and D show the evolution of the standard deviation of log-wages and log-wage residuals in the IABS and GSOEP. Panel E and F display the age-constant medium/low and high/medium wage premium in the two data sets. In the GSOEP, we report findings for four samples: "Full sample, no bonuses" includes the civil servants and self-employed in the sample. The wage measure is the monthly wage excluding bonuses, holdiay and Christmas money. "IABS sample, no bonuses" excludes the civil servants and self-employed from the analysis. "IABS sample, bonuses" is based on the same sample, but includes bonuses, holiday and Christmas money in the wage measure. This specification is most similar to the one in the IABS. "Hourly wage" also excludes the self-employed and civil servants. The wage measure is the hourly wage including bonuses, holiday and Christmas money.
Source : GSOEP sample for women working full-time between 21 and 60 years of age. 2% IABS sample for women working full-time between 21 and 60 years of age.