

The Evolution of the Early Career Gender Wage Gap

Astrid Kunze (IZA, Bonn)¹

10. December 2001

¹E-Mail address: kunze@iza.org. Address for correspondence: IZA (Institute for the Study of Labor), P.O. Box 7240, D - 53 072 Bonn. I would like to thank for helpful comments on previous versions of the paper by Michael Lechner, Christian Dustmann, Wendy Carlin, two anonymous referees and participants in presentations at the CEPR Workshop in Milan on Non-Standard Labour Market Contracts, the University of Bergen and the Aarhus Business School.

Abstract:

In this paper we investigate when the male-female wage differential arises: Does it evolve over the early career or does it exist right from entry into first employment onwards? For the analysis we use new administrative longitudinal data and focus on the early careers of skilled workers in Germany. We adopt a simple human capital theory approach. Advantages of the data for this type of analysis are that we observe complete work and wage histories, and that we observe workers' skills. Regarding entry wages we find a gap of approximately 25 percent. For the early career, i.e. up to eight years of work experience, the differential stays almost constant at this high level. We find that differences in the apprenticeship training occupation explain the main part of this gap and seem to lead to a permanent wage disadvantage throughout the early career.

JEL classification: J16, J3, J7

Key words: Male-female wage differentials, human capital, early career, sample selection, occupation, apprenticeship training.

1 Introduction

Previous research, using both cross sectional and longitudinal data, finds a male-female wage differential of 20 and 30 percent in most countries.² Taking a human capital approach of wage determination³, it has been found that differences in education, workplace characteristics, levels of work experience and time out of work are important explanatory variables for the gap.⁴ These studies, however, do not address the question of when the male-female wage differential arises. Does it evolve over the early career that is typically the period of large wage growth⁵, or does the gap exist from entry into first employment onwards? Evidence on these issues, however, can be important for the design of efficient anti-discrimination policies.

The underlying theoretical model of most gender wage gap studies is a human capital model that predicts that equally productive workers are remunerated equally. Hence, a major issue in this field of research has been to construct comparable workers, and to decompose the wage differential into an explained part and a residual, or unexplained part. The explained part measures differences in observed human capital endowments, while the unexplained part may under a number of assumptions be interpreted as an estimate of discrimination.

We use administrative longitudinal data and focus on the early careers of skilled workers in Germany.⁶ We observe complete work histories and, in contrast to most other studies, we observe skill, that make the data

²Blau and Kahn (1995).

³Becker (1964), Mincer (1974).

⁴See for example, O'Neill and Polachek (1993), Harkness (1996), Blau and Kahn (1997), Groshen (1991).

⁵Topel and Ward (1992).

⁶We define skilled as vocationally skilled through apprenticeship training schemes. Over recent decades, approximately 60 to 70 percent of each birth cohort have taken this route of training.

particularly advantageous to use for this type of analysis. Furthermore, we can measure human capital acquisition in great detail and over a long period. The data cover 15 years, 1975-1990, and post-apprenticeship employment histories can be followed for up to 12 years.

Overall evidence on the evolution of the gender wage gap over early working careers is scarce. In a few studies, it has been demonstrated that a significant entry wage gap of 11 percent exists for the U.S.⁷ and a 7 percent gap exists for graduates in the U.K.⁸. Evidence on the evolution of the male-female wage gap over the career suggests that it is first increasing with workers' time in the labour force, as well as with work experience, and later in the working life cycle again decreasing.⁹ From the literature, however, we know little about what explains this gap at entry and whether it becomes permanent over careers, and what explains its evolution.

Looking at wages conditional on work experience, we find in the German data that the male-female differential in entry wages is approximately 25 percent. Throughout the early career, it stays virtually constant at this level. This is in contrast to human capital theory which predicts zero gap between equally qualified workers. Moreover, 25 percent gap seems like a very high differential.¹⁰

In order to detail the evolution of the gap, we specify standard earnings equations allowing for a fully flexible specification of work experience, cohort and time effects. To make workers comparable within and across periods,

⁷See Loprest (1992) who uses samples of 18 to 25 year old men and women of all education groups taken from the NLS (The National Longitudinal Survey conducted) for 1978 to 1983.

⁸The data are for 1970, see Dolton and Makepeace (1986).

⁹Light and Ureta (1995) and Loprest (1992).

¹⁰By contrast to the studies on U.S. data, we focus on the vocationally skilled workers excluding unskilled workers and those with a higher degree. Hence, it is difficult to make predictions regarding the population average for young workers and how it compares to these studies.

we make use of the detailed human capital characteristics as well as the information on non-wage spells for those (female workers) that are dropping out of the wage sample when bearing and raising children. This induced type of potentially non-random sample selection has consequences for the estimation of wage regression models and wage differentials. First, one may be worried about induced bias of parameter estimates in the female sample wage regression, due to correlation of explanatory variables and the error term¹¹. Furthermore, changes in male-female differences in mean human capital characteristics may be as well affected. In case of long panels, this may lead to biased inference of wage differential changes across time. In order to take account of this potentially non-random sample selection, we suggest an extension of the Oaxaca (1973) decomposition approach¹² that uses information on participants and non-participants, and allows us to estimate the explained portion of the gap consistently under fairly general assumptions.

The remainder of the paper is organized as follows: First, we present simple wage profiles of young male and female workers. Second, we derive the extension of the Oaxaca (1973) decomposition of wage differentials. Third, we describe the data set and summary statistics. Fourth, we show estimation results for the decomposition. Finally, we conclude.

2 Wage profiles over the early career

We find a high male-female wage differential from entry into first employment and a constantly high differential across work experience levels. This is shown in Figure 1 where we plot wages as a function of actual work

¹¹Heckman (1979).

¹²See also Blinder (1973).

experience. We applied a robust non-linear smoothing technique.

figure 1 here

As can be seen, at entry into first employment a considerable differential in wages is observed.¹³ Thereafter, wage experience profiles for men and women are slightly concavely shaped and seem to develop in almost parallel fashion. The differential, accordingly, stays almost constant during this period at around 0.25.

This is in contrast with the human capital theory prediction that equally productive workers are remunerated the same. The graph shown above is based on wages for workers who are virtually homogeneous with respect to education; most have 10 years of schooling and all have undertaken apprenticeship training afterwards. Apprentices start typically at age 16 and involves 2 to 3 years of training with a firm. Apprentices have an apprenticeship contract with the firm they are trained with; wages amount to about 20-30 percent of the wage of a skilled blue or white collar worker. In order to receive a certificate about the particular occupational qualification acquired, apprentices have to pass written and oral examinations, and practical exercises in craftsmanship. Exams are unified across Germany or the Länder and are held externally by the chambers of commerce and trade and chambers of craft.¹⁴ During the period of 1975 to 1990, apprenticeships within the German dual-system apprenticeship programme could be undertaken in about 350 occupations, ranging from technical to service occupations, and in all sectors, including large or small, private or public firms of the economy.

¹³What is not shown here is that before entry into first employment - that is while workers are in apprenticeship training and while they earn only approximately 30 percent of skilled workers' wages - the mean wages are very similar for males and females.

¹⁴For a detailed description of the German dual system apprenticeship programme see Münch (1992).

In the following, we investigate this differential further. In order to make workers comparable within periods we make use of detailed human capital characteristics. Across periods, however, the sample used in Figure 1 may vary in its composition due to drop-outs. This mainly applies to young females withdrawing from work temporarily due to child bearing and rearing. Over a long time period, this may affect comparisons of estimated wage differentials.

3 Measurement of male-female wage differentials

The most common approach to measure male-female wage differentials in the literature is the Oaxaca (1973) decomposition. The Oaxaca (1973) decomposition is given by:

$$(\overline{\ln w_t^M} - \overline{\ln w_t^F}) = \hat{\beta}^M(\bar{X}_t^M - \bar{X}_t^F) + \bar{X}_t^F(\hat{\beta}^M - \hat{\beta}^F) \quad (1)$$

where the price vectors β^M and β^F are recovered after estimating $\ln w_{it}^g = X_{it}^g\beta^g + \epsilon_{it}^g$ for males ($g = M$), and females ($g = F$), respectively. We denote the natural logarithm of the wage of individual i in period t as $\ln w_{it}$, X_{it} is a vector of human capital characteristics, β the vector of prices and ϵ_{it} is a random error about which we make standard assumptions. Superbars indicate means. It follows that the difference in mean wages can be decomposed into a component explained by differences in endowments and an unexplained, residual, component due to differences in prices. Consistent estimation of each of the components, and hence their interpretation, may be inhibited by a number of problems.

First, by construction the decomposition involves the well known index number problem. This may lead, for example, to results that are sensitive

to the choice of the competitive price, i.e. β^M in the above version of the decomposition. Second, the individual human capital characteristics included in X_{it}^g may be measured with error. Third, consistent estimates of the parameters of interest in the wage regression model are required in order to weight the mean differences in endowments. This is often difficult due to endogeneity problems and the lack of instruments. Fourth, in order to draw inference across long employment history data, one must take account of changes in the population, i.e. non-random sample selection, and its impact on mean characteristics, \bar{X} . This point is particularly relevant when looking at the early careers of male and female workers, where females are typically likely to drop out of the sample.

Our decomposition approach takes account of this selection problem. Whilst the Oaxaca decomposition makes use only of information on participants, we extend this approach and exploit the information on non-participants as well. This allows for a non-biased comparison of estimates of the differentials in characteristics and the wage gap decomposition across time, under the assumption that the relevant vector of prices, β_t^M , is estimated consistently from the male sample wage regression model.

Suppose the hypothetical overall wage differential between two groups of workers can be written as a weighted average of the observed mean differential within the group of participating workers and the predicted wage differential within non-participating workers.

$$\begin{aligned} \overline{\ln w_t^M} - \overline{\ln w_t^F} &= (\rho_{pt}^M \overline{\ln w_{pt}^M} - \rho_{pt}^F \overline{\ln w_{pt}^F}) \\ &+ ((1 - \rho_{pt}^M) \overline{\ln w_{nt}^M} - (1 - \rho_{pt}^F) \overline{\ln w_{nt}^F}) \end{aligned} \quad (2)$$

where $\rho_{pt}^M = N_{pt}^M / (N_{pt}^M + N_{nt}^M)$ is the fraction of participating male work-

ers. For females ρ_{pt}^F can be written accordingly. The subscript p indexes participating individual spells, and n non-participating ones.

For illustration, assume two periods: $t-1$, t . In period $(t-1)$ everybody is working, and in period t a positive proportion of workers are participating in the labour market, whilst the remainder, $(1-\rho)$, are not. Since everybody is working in period $(t-1)$, calculation of the mean differential and the decomposition are straightforward.

For period t , however, we need to predict wages for those who are not participating in the labour market. Predictions are estimated as follows:

$$\ln w_{nit}^F = \hat{\ln} w_{nit}^F + \hat{u}_{nit}^F \Leftrightarrow \ln w_{nit}^F = X_{pit-1}^F \hat{\beta}_t^F + \hat{u}_{nit}^F \quad (3)$$

where the subscript n denotes non-participation.

If selection is only on observables, and each individual is observed at least once in the wage sample, then one can simply predict wages for each non-participating individual as follows:

$$\hat{\ln} w_{nit}^F = X_{pit-1}^F \hat{\beta}_t^F \quad (4)$$

where we use $X_{nit}^F = X_{pit-1}^F$. This is done since we only observe the characteristics as long as the individuals are participants.

If selection is also on unobservables, then one can predict the residuals for non-participants by using an individual's percentile in the residual wage distribution, q , in period $t-1$.¹⁵ That is:

$$\hat{u}_{nit}^F = E[u_{nit}^F | u_{pit-1}^F, F_{pt-1}, F_{pt}] \quad (5)$$

where F_{pt} (and F_{pt-1}) is the cumulative distribution of the error term in period t ($t-1$). Hence, by definition $\hat{u}_{nit}^F = F_{pt}^{-1}(F_{pt-1}(u_{pit-1}^F))$. An underlying

¹⁵This approach alludes to Juhn et al. (1993).

assumption is that for non-participants the position in the residual distribution, or their unobserved characteristics, do not change after the time of withdrawal from the labour market. This approach can easily be extended to more periods, and longer time lags.

It is hard to estimate the parameters in β_t^F consistently, and, thus, the hypothetical overall wage differential. In particular, strong and economically implausible assumptions would have to be made regarding the selection process that is most likely to bias the ordinary least squares estimate of β_t^F . Under less restrictive assumptions the explained part of the wage differential can be identified. We adopt this approach.

For identification, we assume that β_t^M , the vector of prices from the males sample regression, is estimated consistently. That means that we assume that non-random sample selection leaves estimates unaffected. The intuition for this assumption is that, as we are going to show in the following section, men are almost constantly participating once they have entered the labour market.¹⁶ Relaxing this assumption and considering potential upward bias of the parameter estimate of β^M due to unobserved heterogeneity leads to an upper bound estimate of the explained part of the gender wage gap.

4 The data and summary statistics

We use the IAB employment sample (IABS)¹⁷ for West-Germany. This is available for the period 1975 to 1990 and is an administrative event history data set. From the IABS we generate a sample of young workers who have undertaken vocational training within the German dual system apprenticeship programme.

¹⁶See also an application of this assumption in Blau and Kahn (1996).

¹⁷IABS abbreviates the *Institut für Arbeitsmarkt und Berufsforschung Sample*.

The IABS is a 1 percent random sample drawn from the event history data file of the social security insurance scheme, the employment statistics, collected by the German Federal Bureau of Labour.¹⁸ The IABS contains all workers in West-Germany who have had at least one employment spell eligible for the social security insurance scheme. As a result, included are all dependent employees in the private sector, i.e. about 80 percent of total employment in West-Germany.¹⁹ The event history data includes information on every change in working status distinguished into full-time work, part-time work, unemployment and interruption which captures national service and maternity - or parental - leave. The particular event history data structure implies that a unit of the data is a spell, and not necessarily a yearly spell.

A unique feature of our final data sample is that complete schooling, work and skill accumulation histories are observed, which allows precise characterization of human capital characteristics. Our sample contains only records on young full-time²⁰ workers, who have mostly graduated from school after 10 years of schooling and who are observed afterwards in apprenticeship training. In practice apprenticeship takes 2 to 3 years. We select individuals who have undertaken training for at least 450 days without interruption²¹, and have no further vocational training, no technical college or university degree. To ensure that individual employment and wage histories are observed from the beginning, we select individuals not older than

¹⁸The fact that the data was collected for administrative purposes is an obvious advantage and makes the data particularly reliable.

¹⁹Not included are: civil servants, self-employed, unpaid family workers and people who are not eligible for benefits from the social security system. For more details see Bender and Hilzendegen (1995).

²⁰This rule leads to exclusion of 3 percent of spells for males and 18.6 percent of spells for females. Inclusion of the latter will make our results presumably only stronger.

²¹This is the recommended selection rule by the IAB.

15 years in 1975. Hence, we do not have “left-censoring” of work histories problems common in labour economics. In the data individuals are followed over early careers, i.e. the oldest individuals are 30 and the mean age is 23.²² Extraction of these workers from the IABS leaves us with a sample containing approximately 15000 female and approximately 20000 male workers observed in at least one full-time working spell after completion of vocational training. The total number of spells is approximately 87000 for females and 125000 for males.

Employment Rates

As discussed in the previous section non-participants introduce problems into our analysis. This feature is most relevant for the sample of female workers. This can be seen by plotting employment rates for male and female workers, as is shown in Figures 2 and 3.

In the graph for males the first line shows employment rates for the first cohort. These workers have started apprenticeship in 1975, as indicated on the line, and the majority enter first employment 2 to 3 years after training.²³ As expected, males’ employment rates for all cohorts monotonically increase to a level of 80 to 90 percent, and stay high. By contrast, for females, as is shown in Figure 3, we find that employment rates increase first, and then decrease. This can be seen most clearly for cohort 1975 where employment rates drop below a population average, i.e. approximately 55 percent in 1988 for Germany.²⁴ Non-participation of females captures the effect of child bearing and rearing. For a longer observation window, one

²²It turns out that our sample is a sample of strongly attached to the labour market individuals. See Light and Ureta (1990) who analyse a similar sample for the U.S..

²³For a few individuals we observe wages for working in a job eligible to social security prior to apprenticeship. We drop these unskilled work wages from our analysis sample.

²⁴*Statistisches Jahrbuch*, various years, *Statistisches Bundesamt, Wiesbaden*.

would expect employment rates to go up again due to females returning to employment after periods of parental leave.

Economic intuition suggests positive selection into work if a reservation wage story holds. This implies that only those with an offered wage greater than the reservation wage take up work. On the other hand, negative selection could be plausible too if, for example, very young women with relatively short training spells are more likely to return early after giving birth due to income constraints, for example.

Figure 2 and Figure 3 here

Wages

Wages in the IABS are reported on a daily basis and are highly reliable given that they are checked by both data collectors and employees. In the estimations, we use wage spells from 1980 onwards, excluding implausibly short apprenticeships of very young workers. The wage variable is the logarithm of the daily pre-tax wage deflated by the CPI index.

Unfortunately, the IABS does not contain a detailed hours of work variable. Focusing on full-time workers does only limit the range of difference in hours of work across individuals. If women on average are working less hours per day than men, the gap in raw daily wages may be biased. In order to scrutinize this problem we use as an additional data source the German Socio-Economic Panel (GSOEP). This data set contains detailed information on weekly hours of work from 1984 onwards. Hence, we can at least look at the more recent years of our observation window. In Figure 4, average hours of work for male and female full time workers younger than 30 and 11 to 13 years of education are plotted. Females work approximately

40 hours in 1984 and 38-39 hours in 1990. Males work one to three hours more than females.²⁵ Hence, differences in hours of work can account only for less than 7 percent of the wage differential, leaving more than 18 percent unexplained in our data.

Figure 4 here

The human capital variables

Human capital characteristics are constructed from the entire records starting at entry into apprenticeship. We measure schooling before apprenticeship, age at entry - which proxies further schooling until entry into apprenticeship -, the duration of apprenticeship and the occupational qualification. We define the occupational qualification as the occupation in which apprenticeship training has been undertaken. By contrast to most other studies, we observe skill (a) because the records contain individual spells while in apprenticeship and in employment afterwards and (b) because for each of the spells information on the three-digit occupation is given. Furthermore, we can identify the firm and industry that training has been undertaken in. An additional interesting variable is the apprenticeship cohort, that is the year of entry into apprenticeship. The latter may be relevant if, for example, crowding of the baby boomers into training or other cohort-specific shocks, such as the change of quality of training schemes, have had an impact on outcomes. In the following, we will refer to this effect as the cohort effect.

General human capital acquisition during employment is measured by the variable years of actual work experience. This variable can be constructed from the individual wage spells. These spells also include details on the

²⁵We use actual hours of work, including overtime.

employer, the industry, occupation of work and a crude measure of job status. Thus, we consider the transition of human capital from apprenticeship to first employment. In order to take account of firm²⁶, occupation and industry²⁷ specific components of training we generate binary *skill match variables*. These take the value one if an individual stays and zero otherwise. Stayers with respect to occupation, for example, are defined as individuals for which the *occupational qualification* on a three digit level is the same as the *occupation of work*.

4.1 Participating males and females

Table 1 here

We first show means for the entry into first employment spell for males and females in Table (1). Workers are homogeneous with respect to education; virtually all of them have an intermediate secondary schooling degree, i.e. 10 years of schooling. They are also homogeneous with respect to type of tertiary education since all have undertaken an apprenticeship programme. Duration of the programmes varies, however, lasting on average 2.18 years for females and 2.51 years for males. It turns out that females and males are both of similar age in their first employment; women are on average 20.3 years of age whilst men are only 0.2 years older.

Despite similarities of the quantity of education and vocational training, we find striking gender differences in the type of training, i.e. *occupational qualification*. Similarly to other Western industrialized countries, females are more likely to be qualified in services, such as a *professional clerical*

²⁶Firm identifiers are given to each establishment in the IABS. Large firms are split into establishments with different firm identification numbers.

²⁷Industries are distinguished into approximately 99 groups (2-digits). The category refers to the main sector of value addition.

worker or *receptionist*, while males are more likely to do apprenticeships in manufacturing, for example, as a *motor vehicle mechanic* or *electrician*.

Occupational segregation in first employment can also be seen from the statistics on the broad measure for *job status*. Results are as expected: For example, 76.2 percent of women work in white collar jobs, whereas 64.8 percent of men work in blue collar jobs. Perhaps striking in international comparison, however, is that about 70 percent of all workers are categorised as skilled, which implies that almost 50 percent of the entire population are categorized as (occupationally) skilled at the young age of 20.²⁸

The *skill match variables* reveal quite strikingly high shares of stayers, in particular, in the *occupation of qualification*, i.e. 73 percent for females and 65 percent for males, and with the training firm, 63 and 70 accordingly. High shares of stayers may suggest that one finds positive returns for staying and losses for moving between firms, jobs (occupations) or industries due to non-transferability of human capital.²⁹ Looking at Figure 5, we see that also across time mobility with respect to occupational qualification is quite low; in particular for female workers. After 6 years of work approximately 60 percent of females and 50 percent of males are still working in the 3-digit occupation they have received their apprenticeship training in.

figure 5 here

Contrary to evidence from cross sections for the entire work force, young females are working more than males in our sample. At the mean females work 3.69 years and males 3.58, with the difference significant at the 1

²⁸To do this calculation one needs to keep in mind that about 60-70 percent of the population in Germany undertakes apprenticeships (Münch, 1992). In comparison, in the U.K. for the period 1990-1992 GHS data shows that only 27.9 percent of all male and 19.4 percent of all female aged 25-34 reached a degree or a higher educational level. See: Harkness (1996).

²⁹This is also what we find in the data.

percent significance level.³⁰

4.2 Female participants and non-participants

figure 6 and table 2 here

In order to derive evidence on the sample selection bias caused by female workers withdrawal from work, we break down summary statistics for those who are participating in two consecutive periods, $t-1$ and t , and those who are only working in $t-1$. First, in Figure 6 we document the gap in work experience that builds up over time in the labour market. Differences are monotonically increasing, indicating that a selected group stays out of the labour market for a longer period of time, while others work continuously. This difference is more pronounced for the cohorts followed longest over time, for example cohort 1975, than for the average. Second, in Table 2 we show mean characteristics for females who drop out just after completion of two years of working and those who continue working, labelled non-participants and participants respectively. Most interestingly, we find that non-participants have experienced longer spells of time out of work and they have lower levels of schooling and training. They also are more likely to be employed in blue collar jobs and they are more likely to change occupation to one that is different from their occupational qualification. All of these differences are significant at the 5 percent level. Differences stay also significant when looking at work experience levels up to 6 years. In summary, this extensive list of human capital characteristics shows that

³⁰One must note that national service is compulsory for men in Germany. It took 15 months (20 months) from 1972 until 1989 depending on whether military service or civil service was served. In Germany, the average age of mothers at first birth was 25.19 in 1980 and 26.93 in 1990 (See *Statistische Bundesamt: Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 1, 1999*). Hence, we have few individual records with an interruption due to having children in the data.

those women who drop out of the labour market have less favourable observed characteristics at the mean. This is consistent with a reservation wage argument where those with the lowest wages are most likely to drop out.

5 Estimation Results

In order to analyse the evolution of wages and the wage gap, we estimate the following simple empirical wage equation³¹:

$$\ln w_{it} = \beta_0 + ex_{it}\beta_1 + c_{it}\beta_2 + (ex * c)_{it}\beta_3 + T_t\beta_4 + Z_{it}\delta + u_{it} \quad (6)$$

where i indexes individuals and t time. The dependent variable is the logarithmic daily wage, $\ln w$. The variable ex denotes a vector of dummies for each integer year of work experience, c a vector of dummies for each apprenticeship cohort, and T contains dummies for the calendar year. Correspondingly, $(ex * c)$ are the interactions of these variables. Z_{it} is a vector of detailed human capital controls, such as age at entry, duration of apprenticeship training and occupational fixed effects. Note that a subset of these are varying across individuals but are time-invariant. The term u_{it} is idiosyncratic error. We allow for a fully flexible specification in work experience, cohort and time; hence, these coefficients are estimated non-parametrically.³² Assuming that human capital acquisition depends only on work experience and cohort-specific factors, time dummy variables enter the equation in an additive fashion. Hence, it is assumed that the time trend captures the general price of human capital level in the economy, and that only shifts of the intercept are relevant.³³

³¹The motivation for this model is the Mincer (1974) earnings equation.

³²At the same time we lose of efficiency.

³³See Dustmann and Meghir (1999).

It is interesting to identify work experience, time and cohort effects in our context. We interpret the coefficient of the work experience variable as the gain from an extra year of on the job training. Coefficients of the cohort variable capture between cohort differences due to, for example, variation in education institutions during the observation period, or variation in quality of training across cohorts. Time effects take account of general macroeconomic effects.

It is well known that identification of work experience, time and cohort effects is hard due to multicollinearity problems.³⁴ In our case, cohort measures year of entry into apprenticeship and experience measures actual years of work experience. Hence, for identification we can use the fact that within a given year we have variation in cohort, and that we have variation in time out of work periods across individuals.

5.1 Year, cohort and work experience effects

In the first set of results, we focus the estimation of year, cohort and work experience effects. We estimate equation (6) excluding Z . At this stage we neglect sample selection problems due to drop outs and do not include other detailed human capital characteristics. Both factors are, therefore, contained in the error. However, as we will show findings are fairly robust to the inclusion of these factors. Coefficients for selected cohorts of the interacted variables are reported for male workers in Table (3) and for female workers in Table (4). The omitted categories are year 1980, cohort 1975 and zero years of work experience. It is found both for males and females that

³⁴In other studies, cohort is birth cohort, work experience is potential years of employment, or age, and time are time periods. Hence, the matrix is not full rank due to the relation that birth cohort and potential work experience adds up to time period. See Fitzenberger et al. (2001) for an identification strategy.

holding other characteristics constant wages are increasing over time. Taking into account time and cohort effects, entry wages are increasing across cohorts. This can be seen from a crude summation of the corresponding marginal time and cohort effects assuming that first employment is three to four years after entry into apprenticeship. Wages are also an increasing function of work experience, which is non-linear. Furthermore, profiles are significantly different across cohorts. In particular, for cohorts 1978, 1981 and 1984, as shown in the table, profiles are steeper.

Table 3 and Table 4 here

Summarising the results on the predicted wage differentials, Figure 6 plots predictions conditional on one of the three time dimensions, holding the remaining ones constant. Means are shown as well as 95 percent confidence bands. As we have found already, the gap seems not to decrease across work experience levels. Only at 8 years of work experience a slight drop can be seen. Similarly, the gap stays constant over time. By contrast, across cohorts we find a diminishing differential.³⁵ While the mean gap is approximately 32 percent for cohort 1975, it drops to 20 percent for cohort 1986, and to 15 percent for cohort 1987.³⁶

Figure 7 here

5.2 Decomposition results

In this section we present results for the decomposition. Taking account of selection due to non-participants the explained part can be separated into

³⁵Fitzenberger et al (2001) has shown that no or only weak birth cohort effects can be found using data from 1975 to 1984. However, for the medium skilled in their analysis birth cohort effects are stronger than for other groups.

³⁶This finding remains when we plot the same graph taking means only across work experience zero to four years, and 5 to 8 years. Hence, it seems not to pick up the affect that individuals of older cohorts have been longer in the labour market than more recent ones.

three components:

$$\begin{aligned}
\beta_t^M(\overline{X_t^M} - \overline{X_t^F}) &= \beta_t^M(\rho_{pt}^M \overline{X_{pt}^M} - \rho_{pt}^F \overline{X_{pt}^F}) \\
&+ \beta_t^M((1 - \rho_{pt}^M) \overline{X_{nt}^M} - (1 - \rho_{pt}^F) \overline{X_{nt}^F}) \\
&+ (\hat{u}_t^M - \hat{u}_t^{*F})
\end{aligned} \tag{7}$$

where $\hat{u}_t^M = 0$, and \hat{u}_t^{*F} is the vector of female wage residuals at male prices of unobservables.³⁷ In more detail, the first term, neglecting the weights, corresponds basically to the Oaxaca decomposition based on participants. The second term corrects for selection on observables, and the third for selection on unobservables.³⁸ The residual, or unexplained part, can then be derived by subtracting the explained part from the total gap.

We present the results in Tables 5 to 7 broken down by work experience, year and apprenticeship cohort based on two specifications of the wage model, as formulated in equation (6). In each table results in panel A are derived from a wage equation including in addition to year, cohort and work experience fixed effects *the detailed human capital variables* as we have listed them in Table 1, yet *excluding controls for occupational qualification*. Results in panel B are based on estimates from our most extensive model *including approximately 300 dummies for each occupational qualification*.

tables 5 to 7 here

In Table 5 we see results broken down by work experience. The first column reports the total wage gap as we have seen it already. The second column lists the part that can be explained in absolute terms and as a percentage of the total wage gap. From panel A, we see that only 9.44

³⁷See Juhn et al. (1993). For estimation we split the wage residual distribution into 100 percentiles that allows a very detailed matching.

³⁸Note that equation (7) reduces to the Oaxaca decomposition substituting $\rho^M = \rho^F = 1$ and $\hat{u}_t^{*F} = \hat{u}_t^F$.

percent of the entry gap, i.e. zero years of work experience, can be explained by differences in human capital characteristics, excluding occupational fixed effects. The contribution of each component is relatively small and not reported here.³⁹ Moving to panel B and the estimate of the explained part of the entry wage differential, we see that 62 percent is explained by the occupational fixed effects. This factor also remains very important when we estimate the explained part of the gap correcting for selection on observables (column 3) and selection on unobservables (column 4). Looking at the evolution of the gap across rows one can see that while the total gap changes not very much, the explanatory power of the occupation specific effects seems to become rather stronger in comparison to the other characteristics, such as the duration of training. This result changes only when controlling for selection on unobservables. Then, the fact that females are relatively better endowed with respect to unobservables than males means that the explanatory power decreases with the increase in work experience.

Across year, work experience and cohorts, we find that differences in occupational qualification result in a permanent wage disadvantage for women. This can be seen from Panel B in Tables (5) to (7). More than 45 percent of the gap at all points in the early career, years and cohorts can be explained by this factor. In other words, this differential does not decrease due to, for example, mobility in Germany.

The strong decline across years and cohorts of the total gap as well as of the explained fraction of the gap may indicate substantial changes that have helped to improve the relative position of young skilled females in the labour market. Like in other countries, occupational segregation is high and

³⁹From Table 1, however, one can see the unweighted differences in means for the entry wage spell.

has not changed over time. On the other hand the 1970s and 1980s have seen structural changes that increase the importance of the service sector. In turn efforts have been made to improve the quality of apprenticeship training schemes. Since males are often trained in the ‘old’ craft sector that already has traditionally high quality training schemes, it is more likely that females profited more from these improvements. They are more likely to be in training schemes in new service and telecommunications jobs.⁴⁰

One may speculate how robust these findings are to the assumptions we have made. We identify the vector of parameters of interest using the male sample wage regression, following from equation (6). Since males are constantly working, the only source of endogeneity that may bias the estimates from the simple non-parametric estimation is unobserved individual specific components captured by the error term u_{it} . If, for example, more able or more highly motivated individuals are more likely to work for more years then the return to work experience is likely to be estimated with upward bias. Hence, the corresponding component of the explained part of the gap is likely to be estimated with upward bias. Accordingly, one may conjecture that more motivated individuals invest relatively more in schooling and, hence, their age at entry may be higher. Likewise, they are more likely to do *Abitur* (upper degree) and more likely to have longer apprenticeship duration, assuming that duration is positively correlated with quality. For the job status and skill match variables the relation is less straightforward. These factors, however, leave the change of the explained part of the gap across time unaffected if the bias is constant across time, work experience and cohort. Hence, we can identify the change consistently.

A major concern with these results is that occupational qualification like

⁴⁰See Münch (1992).

occupational choices in other studies is endogenous. This leads to the question why we observe a high degree of occupational segregation. The concern is that it is determined by discriminatory forces, e.g. entry barriers, that bias outcomes and implies that our estimate of the unexplained differential is downward biased. Similarly, societal rules or images pupils are taught at school and by their parents can have effects. These forces may result in females from being discouraged from going into male jobs, such as manual jobs or jobs in science. However, in order to explain male-female wage differentials one must assume that male occupations generally are more productive. While the argument in favour of discrimination may be very intuitive, though hard to measure directly, the latter is less straightforward. Another story that explains segregation of females into careers with relatively flat wage profiles, can be motivated by the human capital model.⁴¹ In short, this model shows that females who may anticipate more interruptive working careers maximize their lifetime earnings by choosing occupations with relatively low training content, and in which interruptions are penalized less heavily.⁴² Both arguments may apply to our case. This is particularly true for the very young, 16 to 18 year olds, who choose a career. Their choices may be much stronger determined by teachers, parental wishes, and societal norms, than choices at a later stage.

6 Conclusions

We have examined the male-female wage differential during the early career. For the analysis we have used German administrative data on young skilled workers that can be followed over the period 1975 to 1990. The data allow

⁴¹Polachek (1981).

⁴²The crucial assumption is that females have advantageous options outside the labour market.

us to compare a relatively homogeneous group of workers and control for a large number of human capital characteristics.

Simple descriptive statistics show an entry wage gap of approximately 25 percent that persists throughout the early career. While we find little change in the total wage gap within cohorts, we find a considerable decline of the gap between cohorts. The fraction explained by differences in human capital endowment changes within and across cohorts.

The decomposition of the gap into the explained part and the residual shows that already at entry a substantial part of the raw entry wage gap, approximately 50 percent, is due to differences in occupational qualification, with a further 10 percent due to other differences in initial human capital. This result holds under the assumption that choices of education, i.e. until age 16, and occupational qualification or apprenticeship, i.e. made approximately at age 16, are exogenous. While differences in occupational qualification seem to be a permanent component in explaining the differential over time, other differences in human capital background, such as duration of apprenticeship, school degree, seem to have diminishing effects.

These results suggest that large permanent wage disadvantages during the early career are formed by the occupational qualification while other detailed background characteristics and differences in individual work histories are only of minor importance. Loprest (1992) highlighted the importance of differences in job or firm mobility for the explanation of early career wage differentials. Correspondingly, our data show that the German youth labour market is characterised by relatively low rates of occupational career mobility and it would be interesting to investigate mobility across firms in an extended analysis of our sample.

One may object that we overestimate the explanatory power of the vari-

able *occupational qualification* because we do not model selection into apprenticeship after school. This is a point well recognized in the literature⁴³ but not easy to solve given limited information. The *occupational qualification* variable in our study can be compared to college degree used in studies for the U.S.⁴⁴ However, it measures skill in more detail and is, presumably, more highly correlated with individual productivity due to the nature of the German apprenticeship system and German labour markets that have been characterised as occupational labour markets in the literature.⁴⁵

Evidence presented here underlines the importance of further study selection into education types and occupational careers in order to measure wage effects of human capital types. This could shed more light on the permanently high wage differential during the early career that we find for the main group of young workers in Germany. For other countries where relatively small entry wage differentials and increasing male-female wage differentials over the early career have been documented, one may conjecture that selection has transitory effects on wages due to more frequent occupational changes.

⁴³Blau and Ferber (1987).

⁴⁴See e.g. Brown and Corcoran (1997).

⁴⁵Soskice (1994)

References

- Becker, Gary (1964): Human capital - A theoretical and empirical analysis with special reference to education, Chicago University Press, 3rd edition, 1993.
- Bender, Stefan, and Hilzendegen, Jürgen (1995): Die IAB-Beschäftigtenstichprobe als scientific use file, Mitt.AB 1/95, IAB, Nürnberg.
- Blau, Francine D., and Marianne A. Ferber (1987): Discrimination: Empirical evidence from the United States, *American Economic Review*, 77(2), pp. 316-320.
- Blau, Francine, and Lawrence M. Kahn (1995): The gender earnings gap: Some international evidence, in: *Differences and changes in wage structures*, Freeman and Katz, editors, National Bureau of Economic Research.
- Blau, Francine, and Kahn, Lawrence M. (1997): Swimming upstream: Trends in the gender wage differential in the 1980s, *Journal of Labor Economics*, 15(1), pp.1-42.
- Blinder, Alan S. (1973): Wage discrimination: reduced forms and structural estimates, *Journal of Human Resources*, 8(4), pp.436-55.
- Brown, Charles, and Mary Corcoran (1997): Sex-based differences in school content and the male-female wage gap, *Journal of Labor Economics*, 15(3), pp.431-465.
- Dietz, Frido (1988): Strukturwandel auf dem Arbeitsmarkt, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 21.Jg., IAB, Nürnberg.

- Dolton P.J., and Makepeace, G.H. (1986): Sample selection and male-female earnings differentials in the graduate labour market, *Oxford Economic Papers*, 38, pp.317-341.
- Dustmann, Christian, and Costas Meghir (1999): Wages, experience and seniority, Institute for Fiscal Studies, Working Paper, 99/1.
- Fitzenberger, B., R. Hujer, TE McCurdy, and R. Schnabel (2001): Testing for uniform wage trends in West-Germany: A cohort analysis using quantile regression for censored data, *Empirical Economics*, 26(1), pp.41-86
- Groshen, Erica L. (1991): The structure of the female/male wage differential: Is it who you are, what you do, or where you work?, *Journal of Human Resources*, 26 (3), pp.457-472.
- Harkness, Susan (1996): The gender earnings gap: Evidence from the U.K., *Fiscal Studies*, 17(2), pp.1-36.
- Heckman, James (1979): Sample selection bias as a specification error, *Econometrica*, 47(1), pp.153-61.
- Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce (1993): Wage inequality and the rise in returns to skill, *Journal of Political Economy*, 101(3), pp. 410-442.
- Light, Audrey, and Ureta, Manuelita (1990): Gender differences in wages and job turnover among continuously employed workers, *American Economic Review*, 80 (2), pp.293-297.
- Light, Audrey, and Ureta, Manuelita (1995): Early-career work experience and gender wage differentials, *Journal of Labor Economics*,

13(1), pp.121-154.

- Loprest, Pamela J. (1992): Gender differences in wage growth and job mobility, *American Economic Review*, 82(2), pp.526-532.
- Mincer; J. (1974): Schooling, experience and earnings, New York: Columbia University.
- Münch, Joachim (1992): Das Berufsbildungssystem in der Bundesrepublik Deutschland, European Centre for the Development of Vocational Training (CEDEFOP), Berlin.
- Oaxaca, Ronald (1973): Male-female wage differentials in urban labor markets, *International Economic Review*, 14(3), pp.693-709.
- O'Neill, June, and Polachek, Solomon (1993): Why the gender gap in wages narrowed in the 1980s, *Journal of Labor Economics*, 11(1), pp.205-228.
- Polachek, Solomon W. (1981): Occupational self-selection: A human capital approach to sex differences in occupational structure, *Review of Economics and Statistics*, 63(1), pp.60-69.
- Soskice, David (1994): Reconciling markets and institutions: The German apprenticeship system, in Lynch, L., editor, *Training and the private sector - International comparisons*, University of Chicago Press for NBER, Chicago, pp.25-60.
- Topel, Robert H., and Michael P. Ward (1992): Job mobility and the careers of young men, *Quarterly Journal of Economics*, 107(2), pp.439-479.

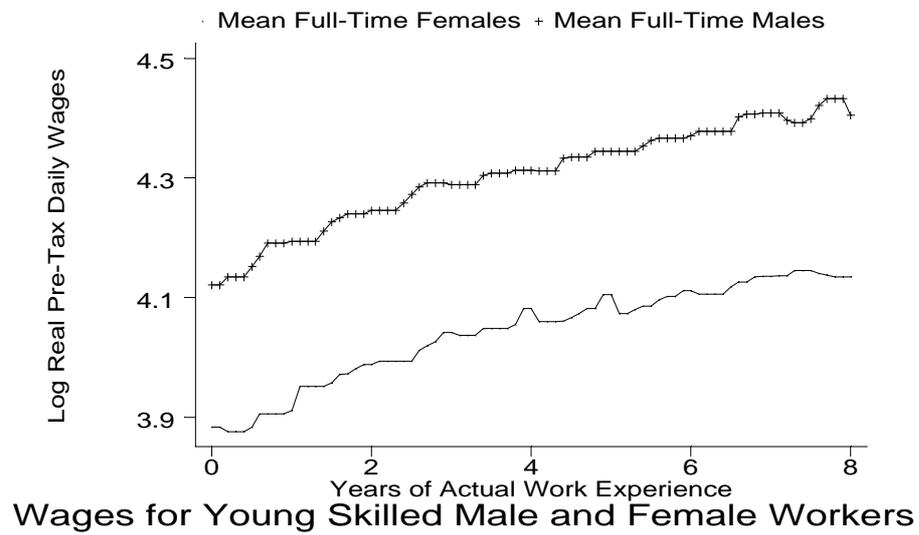


Figure 1:

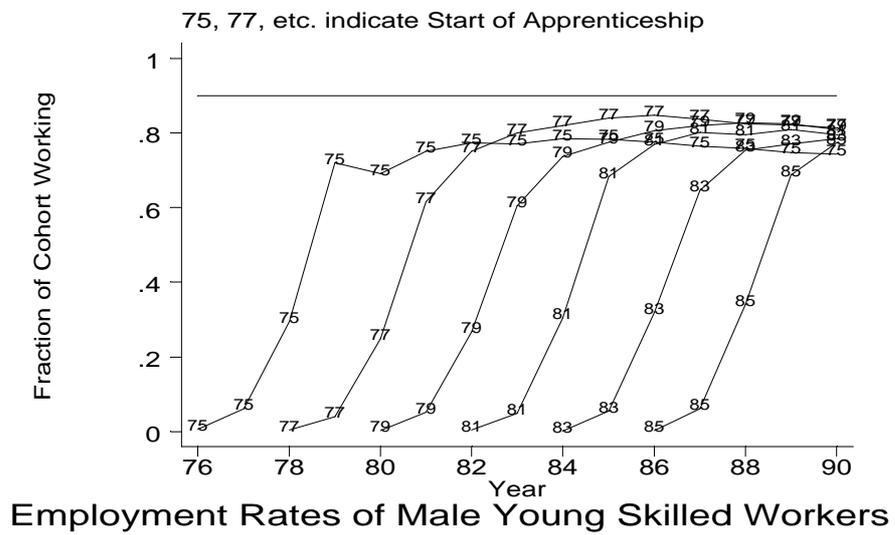


Figure 2:

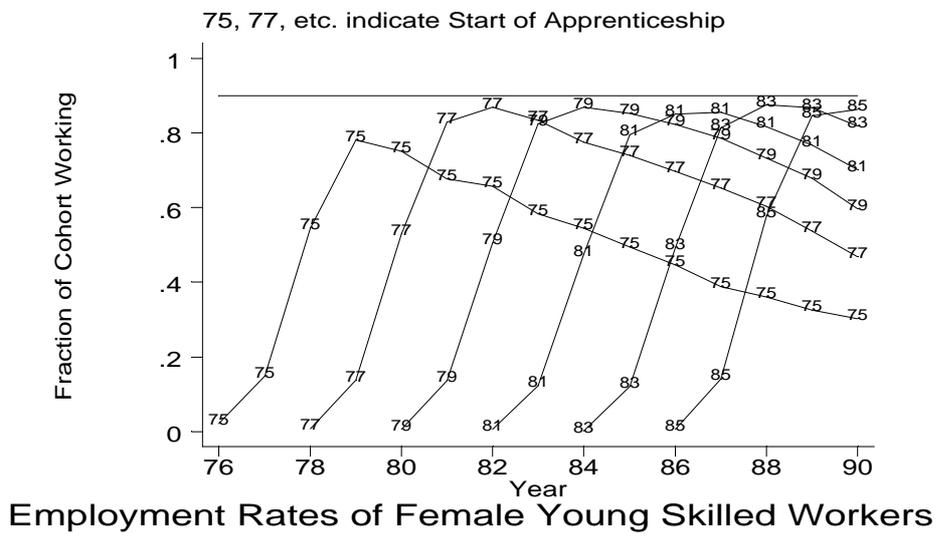
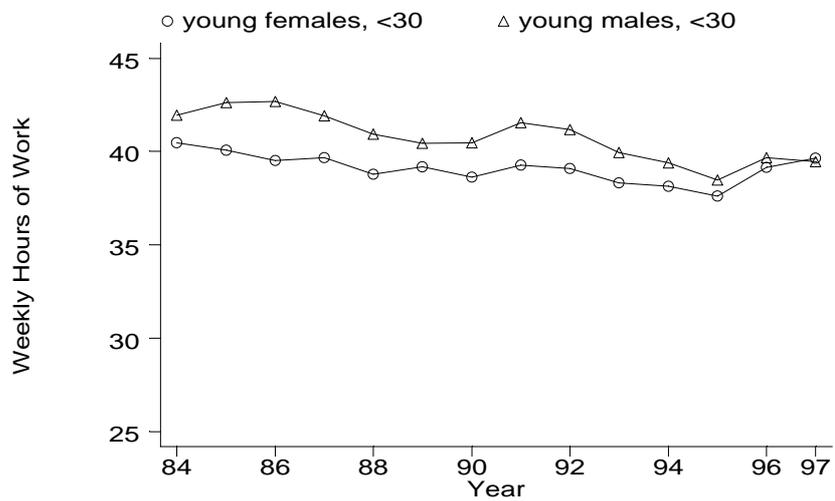


Figure 3:



Hours of work, young workers, 11-13 years education

Figure 4:

Table 1: Means for males and females at entry wage spell

	female		male		t-test for H_0 : Equality of means
	mean	(std.)	mean	(std.)	
	<i>education</i>				
1 if interm. degree	.9547	(.2078)	.9814	(.1350)	-14.37
1 if upper degree	.0453	(.2078)	.0185	(.1350)	14.37
apprentice. duration	2.18	(.7318)	2.51	(.7418)	-40.96
age at entry into apprenticeship	20.3416	(1.5725)	20.5000	(1.5815)	-9.18
	<i>occupational qualification*</i> :				
Natural products production	.0187	.1356	.0280	(.1649)	-5.53
Extraction of natural resources	.0	(.0)	.0090	(.0946)	-11.5
Investment goods production	.0138	(.1166)	.0846	(.2783)	-28.88
Consumer goods production	.0636	(.2441)	.0887	(.2844)	-8.58
Construction	.0054	(.0739)	.1702	(.3758)	-52.15
Installment of technical machines	.0247	(.1552)	.3548	(.4784)	-80.20
Services	.8542	(.3528)	.2105	(.4076)	152.87
Infrastructure services	.0193	(.1378)	.0539	(.2258)	-16.35
	<i>skill related variables</i>				
	<i>job status:</i>				
unskilled	.0896	(.2856)	.1786	(.3831)	-23.62
skilled blue collar	.1481	(.3552)	.6483	(.4775)	-106.47
other (foreman)	.0004	(.0219)	.0009	(.0310)	-1.60
skilled white collar	.7617	(.4260)	.1720	(.3774)	135.31
	<i>skill match variables**:</i>				
1 if Qual.stayer	.7367	(.4404)	.6551	(.4753)	16.20
1 if Firm stayer	.6368	(.4809)	.7015	(.4576)	-12.64
1 if Firm+qual.stayer	.5301	(.4991)	.5295	(.4991)	.11
1 if Industry stayer	.7950	(.4036)	.7983	(.4012)	-.7345
# of individuals	14563		19710		

Notes: * For calculations, the occupation of qualification classifications of the last spell in apprenticeships are used. Groups are constructed according to Dietz (1988). ** Definition of *skill match variables*: Qual. stayer: stayer in occupation of qualification (apprenticeship) measured at 3-digit level. Firm stayer: stayer with training firm. Firm + qual. stayer: stayer in occupation of qualification and training firm. Industry stayer: stayer in industry measured at 2-digit level.

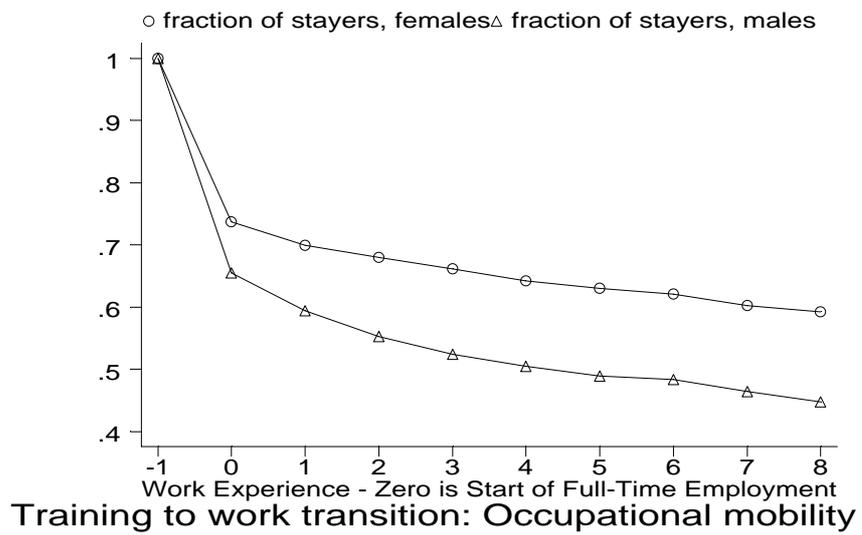
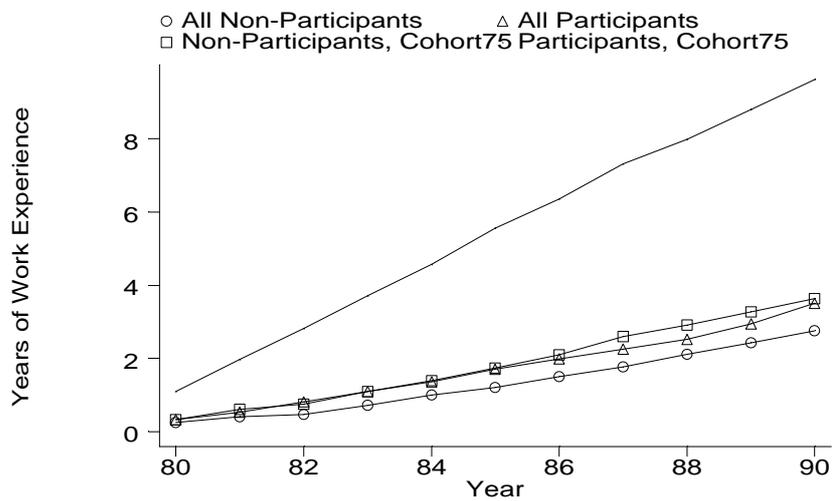


Figure 5:



Work Experience, Female Participants/Non-Participants

Figure 6:

Table 2: Conditional Means for Working Females: $E[X \mid \text{Years of Work Experience} = 2, \text{wage} > 0]$

X	Non-Participating		Participating		t-test for H_0 : Equality of means
	mean	(std.)	mean	(std.)	
age	22.7578	(1.6905)	22.8037	(1.4498)	-1.51
time out of work (until last wage)	.5710	(1.1064)	.2300	(.5585)	22.83
1 if non-zero years of time out	.5613	(.4962)	.4429	(.4967)	11.91
			<i>education</i>		
age at entry into apprenticeship	16.4714	(1.2086)	16.8310	(1.2547)	14.46
1 if interm. degree	.9809	(.1366)	.9574	(.2019)	6.30
1 if upper degree	.0190	(.1366)	.0425	(.2019)	-6.30
apprentice. duration	2.077	(.6761)	2.1066	(.7083)	-2.07
			<i>skill related variables</i>		
<i>job status:</i>					
unskilled	.1376	(.3445)	.0992	(.2990)	6.15
skilled blue collar	.1543	(.3613)	.1130	(.3166)	6.27
other (foreman)	.0020	(.0452)	.0009	(.0310)	1.52
skilled white collar	.7060	(.4556)	.7867	(.4096)	-9.55
<i>skill match variables**:</i>					
1 if Qual.stayer	.6067	(.4885)	.6355	(.4812)	-2.98
1 if Firm stayer	.6026	(.4894)	.6723	(.4693)	-7.33
1 if Firm+qual.stayer	.4202	(.4936)	.4706	(.4991)	-5.06
1 if Industry stayer	.7888	(.4081)	.8029	(.3978)	-1.75
# of individuals	3415		9300		

Notes: For more details see table 1.

Table 3: Wages and Year, Cohort and Experience Effects - Males, Age 20-30

Year	Coef (s.e.)	Cohort	Coef. (s.e.)	Years of Work Experience	Coef. (s.e.)
	(β_4)		(β_2)		(β_1)
81	.0553* (.0102)	76	.0084 (.0172)	1	.0114 (.0271)
82	.0943* (.0099)	77	-.0642* (.0165)	2	.0267 (.0285)
83	.1234* (.0101)	78	-.1136* (.0166)	3	.0387 (.0285)
84	.1676* (.0102)	79	-.1335* (.0167)	4	.0500** (.0298)
85	.1991* (.0103)	80	-.1657* (.0168)	5	.0830* (.0293)
86	.2468* (.0103)	81	-.1718* (.0170)	6	.0803* (.0308)
87	.2785* (.0104)	82	-.1868* (.0171)	7	.1263* (.0323)
88	.3034* (.0104)	83	-.2039* (.0171)	8	.1642* (.0334)
89	.3206* (.0105)	84	-.2132* (.0171)	9	.1958* (.0503)
90	.3557* (.0106)	85	-.2211* (.0173)	10	.1209 (.1029)
		86	-.2426* (.0180)		
		87	-.2740* (.0237)		
<i>Selected Coef. of Interacted Variables (β_3)</i>					
	Cohort=76	Cohort=78	Cohort=81	Cohort=84	Cohort=87
Years of Work Experience	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)
1	-.0209 (.0291)	-.0688 (.2876)	.0701* (.0280)	.0977* (.0287)	-.0702 .0763
2	-.0244 (.0305)	.0567* (.0279)	.1002* (.0295)	.1188* (.0350)	
3	-.0187 (.0305)	.0767* (.0293)	.1305* (.0296)	.1102 (.1134)	
4	-.0318 (.0317)	.0936* (.0293)	.1521* (.0313)		
5	-.0430 (.0314)	.1074* (.0306)	.1499* (.0323)		
6	-.0124 (.0329)	.1067* (.0303)	.2385* (.0424)		
7	-.0355 (.0344)	.1374* (.0322)	.3034 (.1244)		
8	-.0548 (.0361)	.1242 (.0351)			
9	-.1293 (.0549)	.1394 (.0468)			
10	.0150* (.1108)	.0954 (.1965)			

Notes: * significant at 5 % level, ** at 10 % level. Omitted group: 1980, cohort 1975 and zero integer years of work experience.

Table 4: Wages and Year, Cohort and Experience Effects - Females, Age 20-30

Year	Coef. (s.e.)	Cohort	Coef. (s.e.)	Years of Work Experience	Coef. (s.e.)
	(β_4)		(β_2)		(β_1)
81	.0376* (.0129)	76	-.0783** (.0454)	1	.0381 (.0719)
82	.0464* (.0137)	77	-.0812** (.0436)	2	.1065 (.0695)
83	.0579* (.0143)	78	-.0918* (.0433)	3	.1268 (.0786)
84	.0711* (.0146)	79	-.1101* (.0435)	4	.1606 (.0801)
85	.0975* (.0149)	80	-.1036* (.0436)	5	.1450** (.0834)
86	.1399* (.0151)	81	-.1112* (.0437)	6	.1903** (.0986)
87	.1688* (.0152)	82	-.1118* (.0438)	7	.2035** (.1126)
88	.1761* (.0153)	83	-.1141* (.0438)	8	.2848* (.1275)
89	.1783* (.0155)	84	-.1015* (.0439)	9	.2050* (.1531)
90	.1852* (.0157)	85	-.1054* (.0439)	10	.6211* (.1472)
		86	-.0943* (.0442)		
		87	-.0593* (.0468)		
		<i>Selected Coef. of Interacted Variables (β_3)</i>			
	Cohort=76	Cohort=78	Cohort=81	Cohort=84	Cohort=87
Years of Work Experience	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)
1	.0155 (.0751)	.0561 (.0725)	.0630 (.0725)	.0767 (.0724)	.0856 (.1152)
2	.0129 (.0732)	.0251 (.0702)	.0502 (.0703)	.0876 (.0702)	
3	-.0038 (.0821)	.0505 (.0792)	.0979 (.0793)	.1263* (.0801)	
4	.0226 (.0837)	.0603 (.0806)	.1204 (.0808)	.1014 (.1123)	
5	.0634 (.0871)	.1116 (.0839)	.1886* (.0843)		
6	.0553 (.1019)	.1286 (.0991)	.1862** (.1005)		
7	.0793 (.1156)	.1758 (.1132)	.2037 (.1460)		
8	.0345 (.1309)	.1024 (.1294)			
9	.1502 (.1568)	.0716 (.2016)			
10	-.1795 (.1549)				

Notes: See table 3.

Predicted Wage Differential: Mean and 95 % Confidence Intervals

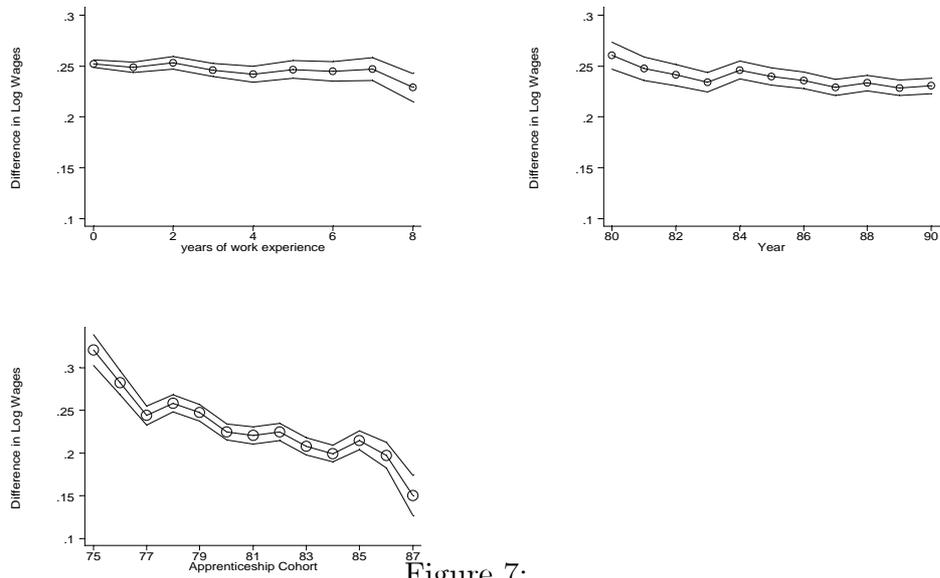


Figure 7:

Table 5: Decomposition Estimation Results, by year of work experience

Work Experience	(1)	(2)	(3)	(4)
	Total wage gap $(\ln w_{pt}^M - \ln w_{pt}^F)$	Explained part, uncorrected for selection $\hat{\beta}_t^M (\bar{X}_{pt}^M - \bar{X}_{pt}^F)$	Explained part, corrected for selection on observables $\hat{\beta}_t^M (\bar{X}_t^M - \tilde{X}_t^F)$	Explained part, corrected for selection on observables and unobservables
<i>Panel A: X includes human capital variables, excluding occupation fixed effects</i>				
0	.2520	.0238 (9.44 %)	.0220 (8.73%)	0.0233 (9.25%)
1	.2485	.0151 (6.07 %)	.0132 (5.31%)	0.0222 (8.94%)
2	.2530	.0168 (6.64 %)	.0139 (5.49%)	0.0212 (8.39%)
3	.2460	.0162 (6.58 %)	.0144 (5.85%)	0.0127 (5.18%)
4	.2418	.0158 (6.53 %)	.0144 (5.95%)	0.0081 (3.35%)
5	.2464	.0145 (5.88 %)	.0137 (5.56%)	0.0052 (2.13%)
6	.2448	.0115 (4.69 %)	.0115 (4.69%)	-0.0081 (-3.33%)
7	.2470	.0126 (5.10 %)	.0126 (5.10%)	0.0006 (0.26%)
8	.2287	.0130 (5.68 %)	.0136 (5.94%)	-0.0008 (-0.38%)
<i>Panel B: X includes human capital variables, including occupation fixed effects</i>				
0	0.2520	0.1573 (62.41%)	0.1583 (62.80%)	0.1559 (61.89%)
1	0.2485	0.1453 (58.49%)	0.1488 (59.89%)	0.1516 (61.01%)
2	0.2530	0.1445 (57.12%)	0.1471 (58.15%)	0.1471 (58.14%)
3	0.2460	0.1453 (59.07%)	0.1465 (59.56%)	0.1374 (55.89%)
4	0.2418	0.1413 (58.45%)	0.1429 (59.10%)	0.1323 (54.74%)
5	0.2464	0.1412 (57.31%)	0.1438 (58.36%)	0.1262 (51.23%)
6	0.2448	0.1400 (57.20%)	0.1403 (57.31%)	0.1122 (45.86%)
7	0.2470	0.1410 (57.09%)	0.1453 (58.82%)	0.1209 (52.89%)
8	0.2287	0.1379 (60.28%)	0.1392 (60.87%)	0.1148 (50.23%)

Notes: *Panel A:* The vector X^M for males, and X^F for females, includes in addition to the intercept the variables years of work experience specific effects, cohort specific effects, cohort correlated with work experience fixed effects, time fixed effects, age at entry, school degree (intermediate or upper), apprenticeship duration, a dummy for each job status, and skill match variables. *Panel B:* In addition, occupation fixed effects are included.

Table 6: Decomposition Estimation Results, by Year

Year	(1) Total wage gap $(\ln w_{pt}^M - \ln w_{pt}^F)$	(2) Explained part, uncorrected for selection $\hat{\beta}_t^M (\bar{X}_{pt}^M - \bar{X}_{pt}^F)$	(3) Explained part, corrected for se- lection on ob- servables $\hat{\beta}_t^M (\bar{X}_t^M - \bar{X}_t^F)$	(4) Explained part, corrected for se- lection on observables and unobservables $\hat{\beta}_t^M (\bar{X}_t^M - \bar{X}_t^F)$
<i>Panel A: X includes human capital variables, excluding occupation fixed effects</i>				
80	.2602	.0237 (9.10%)	.0235 (9.03%)	0.0038 (1.47%)
81	.2473	.0200 (8.08%)	.0196 (7.92%)	0.0080 (3.26%)
82	.2411	.0122 (5.06%)	.0118 (4.89%)	0.0009 (0.39%)
83	.2339	.0061 (2.60%)	.0051 (2.18%)	7.4E-0 (0.00%)
84	.2458	.0023 (0.93%)	.0011 (0.44%)	0.0006 (0.28%)
85	.2397	-.0004 (-0.16%)	-.0015 (-0.62%)	-0.0012 (-0.51%)
86	.2359	.0008 (0.33%)	-.0009 (-0.38%)	0.0015 (0.65%)
87	.2288	.0027 (1.18%)	.0001 (0.04%)	0.0024 (1.06%)
88	.2333	.0057 (2.44%)	.0014 (0.60%)	0.0031 (1.34%)
89	.2285	.0029 (1.26%)	-.0015 (-0.65%)	-0.0006 (-0.29%)
90	.2304	-.0045 (-1.95%)	-.0085 (-3.68%)	-0.0064 (-2.79%)
<i>Panel B: X includes human capital variables, including occupation fixed effects</i>				
80	.2602	.1865 (71.67%)	.1869 (71.82%)	0.1679 (64.53 %)
81	.2473	.1714 (69.30%)	.1717 (69.42%)	0.1557 (62.96 %)
82	.2411	.1575 (65.32%)	.1578 (65.45%)	0.1439 (59.70 %)
83	.2339	.1449 (61.94%)	.1460 (62.41%)	0.1366 (58.41 %)
84	.2458	.1387 (56.42%)	.1406 (57.20%)	0.1348 (54.86 %)
85	.2397	.1309 (54.60%)	.1333 (55.61%)	0.1279 (53.36 %)
86	.2359	.1280 (54.26%)	.1307 (55.40%)	0.1260 (53.43 %)
87	.2288	.1279 (55.90%)	.1299 (56.77%)	0.1243 (54.35 %)
88	.2333	.1281 (54.90%)	.1293 (55.42%)	0.1235 (52.94 %)
89	.2285	.1262 (55.22%)	.1268 (55.49%)	0.1206 (52.80 %)
90	.2304	.1197 (51.95%)	.1202 (52.17%)	0.1132 (49.15 %)

Notes: See table (5) for further details.

Table 7: Decomposition Estimation Results, by Cohort

Cohort	(1) Total wage gap $(\ln w_{pt}^M - \ln w_{pt}^F)$	(2) Explained part, uncorrected for selection $\hat{\beta}_t^M (\bar{X}_{pt}^M - \bar{X}_{pt}^F)$	(3) Explained part, corrected for se- lection on ob- servables $\hat{\beta}_t^M (\bar{X}_t^M - \bar{X}_t^F)$	(4) Explained part, corrected for se- lection on observables and unobservables
<i>Panel A: X includes human capital variables, excluding occupation fixed effects</i>				
75	.3208	.0798 (24.87%)	.0510 (15.89 %)	0.1027 (32.02 %)
76	.2825	.0438 (15.50%)	.0278 (9.84 %)	0.0552 (19.55 %)
77	.2439	.0283 (11.60%)	.0172 (7.05 %)	0.0164 (6.76 %)
78	.2585	.0218 (8.43 %)	.0119 (4.60 %)	0.0216 (8.38 %)
79	.2474	.0037 (1.49 %)	-.0031 (-1.25 %)	0.0062 (2.51 %)
80	.2247	-.0036 (-1.60 %)	-.0096 (-4.27 %)	-0.0154 (-6.87 %)
81	.2206	-.0095 (-4.30%)	-.0122 (-5.53 %)	-0.0225 (-10.21 %)
82	.2247	-.0117 (-5.20%)	-.0134 (-5.96 %)	-0.0127 (-5.67 %)
83	.2078	-.0149 (-7.17 %)	-.0162 (-7.79 %)	-0.0278 (-13.38 %)
84	.1992	-.0151 (-7.58 %)	-.0161 (-8.08 %)	-0.0336 (-16.90 %)
85	.2148	-.0096 (-4.46 %)	-.0105 (-4.88 %)	-0.0282 (-13.16 %)
86	.1974	-.0129 (-6.53 %)	-.0128 (-6.48 %)	-0.0347 (-17.59 %)
87	.1502	.0003 (0.19 %)	-.0000 (0 %)	-0.1193 (-79.43 %)
<i>Panel B: X includes human capital variables, including occupation fixed effects</i>				
75	0.3208	0.2486 (77.47%)	0.2268 (70.68%)	0.2458 (76.63%)
76	0.2825	0.2094 (74.13%)	0.2003 (70.91%)	0.1888 (66.84%)
77	0.2439	0.1672 (68.55%)	0.1634 (66.99%)	0.1478 (60.63%)
78	0.2585	0.1516 (58.63%)	0.1439 (55.66%)	0.1488 (57.59%)
79	0.2474	0.1327 (53.63%)	0.1282 (51.83%)	0.1333 (53.90%)
80	0.2247	0.1178 (52.43%)	0.1135 (50.53%)	0.1096 (48.78%)
81	0.2206	0.1199 (54.36%)	0.1181 (53.52%)	0.1024 (46.45%)
82	0.2247	0.1019 (45.34%)	0.1040 (46.30%)	0.1138 (50.66%)
83	0.2078	0.1114 (53.60%)	0.1110 (53.40%)	0.0941 (45.32%)
84	0.1992	0.1112 (55.82%)	0.1098 (55.11%)	0.0887 (44.56%)
85	0.2148	0.1192 (55.48%)	0.1181 (54.99%)	0.0920 (42.85%)
86	0.1974	0.0956 (48.41%)	0.0963 (48.78%)	0.0858 (43.49%)
87	0.1502	0.0769 (51.20%)	0.0768 (51.13%)	-0.0044 (-2.94%)

Notes: See table (5) for further details.