DISCUSSION PAPER SERIES

IZA DP No. 2185

Skill Wage Premia, Employment, and Cohort Effects: Are Workers in Germany All of the Same Type?

Bernd Fitzenberger
Karsten Kohn

June 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Skill Wage Premia, Employment, and Cohort Effects: Are Workers in Germany All of the Same Type?

Bernd Fitzenberger
Goethe University Frankfurt, IFS, ZEW
and IZA Bonn

Karsten Kohn
Goethe University Frankfurt
and IZA Bonn

Discussion Paper No. 2185
June 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 2185
June 2006

ABSTRACT

Skill Wage Premia, Employment, and Cohort Effects: Are Workers in Germany All of the Same Type?*

This paper studies the relationship between employment and wage structures in West Germany based on the IAB employment subsample 1975{1997. It extends the analytical framework of Card and Lemieux (2001) which simultaneously includes skill and age as important dimensions of heterogeneity. After having identified cohort effects in skill wage premia and in the evolution of relative employment measures, we estimate elasticities of substitution between employees in three different skill groups and between those of different age, taking account of the endogeneity of wages and employment. Compared to estimates in the related literature, we find a rather high degree of substitutability. Drawing on the estimated parameters, we simulate the magnitude of wage changes within the respective skill groups that would have been necessary to halve skill-specific unemployment rates in 1997. The required nominal wage reductions range from 8.8 to 12.2% and are the higher the lower the employees' skill level.

JEL Classification: J21, J31, E24

Keywords: labor demand, heterogeneity, age, skill, wage structure, employment, cohort effects, unemployment

Corresponding author:

Bernd Fitzenberger
Goethe University Frankfurt
Department of Economics
D-60054 Frankfurt am Main
Germany
E-mail: fitzenberger@wiwi.uni-frankfurt.de

* We are grateful to Gaby Wunderlich for her records of the use of German Microcensus data and to the Federal Statistical Office (*Statistisches Bundesamt*) for the provision of age-specific employment data. This paper benefited greatly from constructive comments by Michael Burda, Robert Chirinko, and David A. Jaeger. We thank Martin Biewen, David Card, Thomas Lemieux, and Alfred Garloff for fruitful discussions and Andrej Gill and Norbert Schanne for research assistance. We are also indebted to participants at the following seminars and conferences: IZA workshop \Wage Inequality, Technology and Institutions" (July 2004), EALE 2004, VfS 2004, workshop “Labour Market Imperfections, Unemployment, and Growth" of the Herbert Giersch Foundation (Oct. 2004), ESWC 2005, as well as various department seminars. All errors are our sole responsibility.

1 Introduction

Numerous empirical studies record descriptive micro-data evidence on the evolution of wages and employment measures; see the survey article of Katz and Autor (1999). To capture the heterogeneity of labor, authors usually undertake a grouping into different classes based on observed covariates like age and sex of employees or on the basis of job characteristics. Available studies typically report considerable wage dispersion both between and within adequately defined classes. Variation over time yet generates another important dimension of heterogeneity.

Particular attention is given to skill wage premia and the evolution of skill-specific employment. As a stylized fact, the unemployment rate is the higher the lower the (formal) qualificational level of the employees. In West Germany, for example, the respective rates for employees without a vocational degree, for those with, and for those with a university degree were 19.4%, 5.7%, and 2.6% in the year 2000.[1]

Rigidity of the wage structure is often referred to as a major cause for the different degrees of incidence of unemployment; compare, e. g., Fitzenberger and Franz (2001). As elaborated in the discussion about employment impacts of skill-biased technical change (SBTC; see Katz and Autor, 1999, Acemoglu, 2002), relative demand for low-skilled labor decreases faster over time than does relative supply. In line with neoclassical demand theory (Hamermesh, 1993), market clearing would in this case require an increase of qualificational wage differentials.

Despite the popularity and plausibility of this hypothesis an empirical operationalization of the interrelation between wage structures and employment that goes beyond mere descriptive evidence proves difficult due to the heterogeneity of labor, among other things. Conventional empirical analyses of qualificational labor demand typically take into account only a small number of homogeneous skill groups—mostly not more than three; cf. the surveys in Hamermesh (1993) and Katz and Autor (1999) and for Germany, e. g., the studies of Fitzenberger (1999), Steiner and Wagner (1998b), or Falk and Koebel (1999, 2002). These approaches are often justified in light of the fact that satisfactory solutions to the problem of aggregation do not exist.[2] Also, standard approaches based on cost-minimizing behavior like flexible translog systems, which allow for a larger number of factors, quickly become impracticable.

Based on US data, Katz and Murphy (1992) analyze wage differentials between high school and college graduates in the context of supply and demand effects. A CES model proves compatible with the developments of wage premia and employment over time. These are consistent with the labor market entry of young and the exit of older birth cohorts on the one hand and an increase in average educational attainment on the other. The literature interprets these trends as a race between changes in the skill structure of labor supply and that of labor demand; cf., for example, Johnson (1997), Topel (1997), and Machin (2002). However, in addition to the variation of skills between different cohorts, human capital endowments also change with age. Whereas increasing labor market experience and job

[1]Cf. Reinberg and Hummel (2002), p. 27.

[2]For discussions of the problem of aggregation in the context of labor demand estimations see, e. g., Koebel (2005) and Katz and Autor (1999).

tenure augment human capital stocks with age, skill-biased and accelerating structural change might invalidate individual endowments of older workers. Freeman (1979) and Welch (1979) thus account for imperfect substitutability between workers of different age by means of CES technologies for workers from discrete age or "career phase" groups.

Card and Lemieux's (2001) – henceforth CL – investigation using US, UK, and Canadian data reconciles the analysis of Katz and Murphy (1992) with those of Freeman (1979) and Welch (1979). In a set-up which uses the nested CES model developed by Sato (1967) the simultaneous inclusion of skill and age as dimensions of heterogeneity not only enables the separation of age, time, and cohort effects, but also facilitates the estimation of a specification with a relatively large number of different input factors. The estimation strategy undertaken in particular yields elasticities of substitution both between high school and college graduates and between workers belonging to different age classes.

The starting point of the study by CL is the observation that the college-high school gap in wages has increased strongly for younger US men whereas the gap for older men has remained nearly constant. The driving force for these observed cohort-specific changes is the slowdown in the growth of college-educated labor which did not keep up with the steady skill bias in labor demand; see also Autor, Katz, and Kearney (2005) for a recent reassessment.

Wage trends in West Germany differed from what happened in countries like the US, UK, or Canada over the last decades. In particular, wage dispersion for male workers did not increase to the same extent (Prasad, 2004). In fact, skill wage differentials decreased between workers with and those without a vocational training degree; see Fitzenberger (1999) and Fitzenberger and Wunderlich (2002). Regarding the differential between workers with a vocational degree and workers with university-type education there is conflicting evidence; see Steiner and Wagner (1998a), Möller (1999), and Fitzenberger (1999). Little evidence exists regarding age-related wage differentials. Fitzenberger (1999) and Fitzenberger and Wunderlich (2002) find that cross-sectional age profiles became somewhat steeper for male workers without and for those with a vocational training degree. According to the SBTC hypothesis, skill upgrading in employment should thus have occurred at a faster rate in Germany compared to countries like the US, the UK, or Canada and, in the spirit of CL, cohort effects are likely to be of importance in West Germany as well. There is recent concern that the necessary skill upgrading of the labor force in Germany is too slow to combat the high unemployment of the low-skilled; see the stylized facts reported in OECD (2004).

This paper broadens the scope of the nested CES framework and provides estimates based on the IAB employment sample (IABS) for Germany. While consistently reconciling the developments of relative wages and employment, our treatment extends upon the existing literature in several directions. First, we let three skill groups account for heterogeneity within the qualification dimension. This extension is necessary in light of the coexistence of vocational training and university education in Germany. Second, we treat the identification of cohort effects more rigorously. Tests for the existence of cohort effects and their separability from age and time effects (as suggested by MaCurdy and Mroz, 1995) are applied to check the validity of the specification. Third, rather than merely running regressions for skill wage differentials, we estimate a full system of skill *and* age premia implied by the nested CES model. Fourth, we take a closer look at the notions of observed employment and let instrumental variable techniques account for the endogeneity of both wages and employment. Finally, we draw on the estimated substitution parameters in order to conduct two simulation experiments: We calculate the magnitude of wage changes in the three skill groups that would have been necessary to halve skill-specific unemployment rates in 1997 (the latest period available). While allowing for relative changes between skill groups, this would have left the wage structure within skill groups unaffected. Due to the particularly high unemployment rate among low-skilled employees in Germany, the design imposes a disproportionately prominent increase in employment of this group, and thus is of high policy relevance. Alternatively, one might be interested in changes of the wage structure within skill groups, holding the structure across the respective groups constant. Here, the model set-up may provide an answer to the question how wages for employees of different age would have had to change to reduce all age-specific unemployment rates by one half.

The remainder of the paper is organized as follows: Section 2 outlines the trends in skill wage premia and skill-specific employment in the IABS between 1975 and 1997. Following an investigation into the nature of cohort effects in section 3, section 4 discusses different facets of the nested CES model which allow for the reconciliation of the stylized empirical facts, and section 5 estimates elasticities of substitution across and within skill groups. Based on the resulting parameters, the simulation experiments are presented in section 6. Section 7 concludes.

2 Descriptive Evidence

A number of recent empirical studies provides descriptive evidence for skill wage differentials in the German labor market. Among the analyses—comprising, e. g., Christensen (2003), Christensen and Schimmelpfennig (1998), Fitzenberger (1999), Fitzenberger and Wunderlich (2002), Möller (1999), Prasad (2004), Riphahn (2003), Steiner and Mohr (2000), and Steiner and Wagner (1998a)—there is some consensus that, by and large, the earnings distribution across skill groups stayed relatively stable during the 1980's and 1990's.

A closer look calls for detailed investigations which take into consideration further aspects of heterogeneity. In the tradition of Mincer (1974) work experience is an important additional determinant of individual earnings, and the effects of age—often used as a proxy for experience—are of interest themselves. Prasad (2004) and Riphahn (2003), for example, estimate year-specific Mincer equations and depict the evolution of returns to potential experience. Studies explicitly accounting for the age dimension of wage distributions examine single cross-sectional age profiles, like Fitzenberger and Reize (2003), or focus specifically on cohort analyses, as Boockmann and Steiner (2000) or Fitzenberger, Hujer, MaCurdy, and Schnabel (2001), for example. Beißinger and Möller (1998) account for the age dimension in the distribution of (un)employment for discrete years between 1980 and 1990.

Our study scrutinizes both wages and employment across the two dimensions skill and age for the time span 1975–1997. It is based on the IAB employment subsample (IABS), a 1% random draw of German employment spells subject to social insurance contributions. The

IABS covers about 80% of all employed persons, and it provides detailed information on daily wages for blue and white collar workers as well as the exact timing of employment spells. We classify employees into three skill groups and consider six age classes. An extensive description of the data and classifications used is given in appendix A.

2.1 Stylized Facts I: The Evolution of Wage Differentials

Age-specific skill wage premia or skill wage differentials $r_{sm,a,t}$ among workers of age a at time t are defined as the difference in mean *log* wage of high-skilled ($s = h$, employees with a university degree) or low-skilled workers ($s = l$, employees with neither university nor vocational training degree) and that of medium-skilled workers ($s = m$, employees with a vocational training degree). Using dummy variables $d_{s,a,t}$ for the different skill groups and possibly controlling for further influences,[3] they can be derived from regressions

$$\ln(w_{a,t}) = \text{constant}_{a,t} + r_{l,a,t} \cdot d_{l,a,t} + r_{h,a,t} \cdot d_{h,a,t} + \text{controls}_{a,t} + \epsilon_{a,t} \tag{1}$$

in the respective age-time cells. Due to the social security taxation threshold, wage data in the IABS are censored from above. Thus (1) is estimated by means of Tobit regressions. Observations are weighted by the length of the respective employment spells. Results are provided in table 4 in appendix B.

Figure 2 illustrates the evolution of age-specific wage differentials for males over time. Skill premia generally grow with age. Taking age as a proxy for experience, this corresponds to classical human capital theory (Becker, 1993). The estimated premia have evolved quite differently, though.

The education premium for high-skilled employees compared to the medium-skilled stayed roughly constant for the oldest age class until 1987 and declined by about 9 percentage points (ppoints) thereafter. The relative position of 30- to 35-year-old high-skilled, on the other hand, deteriorated by about 9 ppoints during the late 1970's, partly rose again in the first half of the 80's, and stayed constant from 1986 on.

The differential between older medium- and low-skilled workers exhibited a decline of about 5 ppoints during the eighties and recovered to an overall decline of about 2 ppoints during the nineties. In the youngest age class this wage premium exhibited an even higher volatility: Between 1975 and 1986, low-skilled workers on average gained around 6 ppoints compared to the medium-skilled. Later on, the differential increased again and even exceeded the 1975-level in 1997.

To infer the evolution of age profiles across time, we plot the wage differentials for three years against the age dimension in figure 3. Average wage differentials between high- and medium-skilled generally increase rather steeply with age: The premium grows by up to 29 ppoints. However, the shape of the profiles changes over time.

In 1975 the profile is considerably curved, showing especially a pronounced rise for young individuals. In transition to the mid-1980's, the curvature declines whilst the profile still

[3] Cf. appendix A for details on implemented specifications.

shows a similarly high increase over the entire age span: In particular for middle-aged workers the premium for higher education declines compared to 1975. Starting in the second half of the eighties, one observes a twist of the profile. Whereas the increase in the premium for higher education for workers up to their mid-thirties is much the same in 1997 as in 1986, the profile has become flatter for older employees: The relative position of older high-skilled workers has deteriorated in comparison to the situation in 1986.

In comparison, the profile of the wage differential between low- and medium-skilled workers is typically much flatter, especially for older workers. The differential declines strongly for younger workers between 1975 and 1986 and it increases again strongly between 1986 and 1997. But even though the maximum decline—roughly 8 ppoints in 1986—is found to be small relative to the one experienced by the high-to-medium-skilled differential, the picture of the developments over time is still striking. In 1975 the average education premium moderately rose with age, showing increments declining with age. Up to 1986, the profile shifted downward by about 2–6 ppoints, becoming steeper for younger age classes. In 1997, however, the profile shows a twisted shape: Whilst the differential for older workers partly recovered in a parallel kind of manner, the youngest workers now face a premium increased by 6 ppoints that renders the entire profile nearly flat.

Taking the above results together, we assert a first stylized fact:

> Between the mid-1970's and the mid-1990's, age profiles of skill wage premia have not moved in parallel fashion over time. Skill wage premia declined over time (especially between the 1970's and 1980's) in a non-uniform fashion across age groups.

Thus, the developments are not likely to be the result of pure age and time effects alone. Cohort effects, i. e., systematic differences across birth cohorts, supposedly play an additional important role. Our subsequent theoretical and empirical investigation into the development of skill wage premia hence takes account of age, time, and cohort effects.

2.2 Stylized Facts II: Trends in Relative Employment

Based on the individual spell data, we use a weighted headcount as our measure of employment: In each age-time cell, the number of skill-specific employed is summed up, weighted by the duration of the respective employment spells.

Figure 4 presents relative employment trends for the different age classes. These are the employment counts of the high- and the low-skilled relative to the employment in the medium-skill group, respectively. The measures show the skill upgrading over the past decades: For most of the sample period, both the ratio of high-skilled to medium-skilled and that of medium-skilled to low-skilled employment were the higher the younger the respective age class. Furthermore, the skill-intensity of employment has increased over time. Starting from a situation of uniform skill upgrading in all age classes, however, the increase of relative employment of the skilled slows down considerably or even comes to an end at some point in time. Beginning in the mid 1980's, this break occurs first for the youngest age group. It then works through the older classes during the following years until it affects the oldest employees in the second half of the 1990's.[4]

We record a second stylized fact:

> There is a break in the inter-cohort trend of relative employment such that younger birth cohorts do not follow the older ones towards further skill upgrading.

The empirical evidence thus suggests the existence of cohort effects in the employment dimension, too.

3 Testing for Cohort Effects

To distinguish age, cohort, and time effects in wage premia $r_{a,t}$, CL undertake a decomposition of wage premia by the following regression:

$$(2) \quad r_{a,t} = b_a + c_{t-a} + d_t + \epsilon_{a,t}$$

where b_a, c_{t-a}, and d_t denote age, cohort, and time dummies, respectively. However, one should be cautious with respect to the identification of wage premia. When separating cohort effects from pure time and age effects an identification issue arises because the cohort (defined by the individual's year of birth) is calendar time minus age.

As a first identification approach, we follow CL by estimating equations (2), setting the effects for the oldest birth-cohorts (up to 1928) equal to zero. The model is formally "identified" based on annual data by using five-year age intervals and implicitly assuming age and cohort effects to be constant within each interval.[5] A test for the existence of cohort effects is then conducted by testing for joint significance of all other cohort terms. This approach is suggestive from an economic point of view. However, it resolves the identification problem in a rather ad hoc way; see Heckman and Robb (1985) for a detailed discussion of the identification issue. We employ an alternative approach introduced by MaCurdy and Mroz (1995) and also used in Fitzenberger, Hujer, MaCurdy, and Schnabel (2001) which deals with the identification issue both more explicitly and more rigorously.

Following this approach, we formalize cohort effects as the outcome of interaction between age and time by allowing for interaction terms of different order. For identification, the

[4] Note that the approximate zero-growth of the relative employment of high-skilled in the first age class should not be over-interpreted in our context, because it likely reflects the extension of education durations and the corresponding deferments of labor market entries during the last decades; cf., for example, Reinberg and Hummel (1999).

[5] Boockmann and Steiner (2000) follow a similar identification strategy by defining their cohorts to span periods of five or ten years. In addition, the study considers actual experience rather than age.

linear cohort effect is explicitly set to zero.[6] To test for the existence of cohort effects, we estimate the following specification:

$$
\begin{aligned}
(3) \quad r_{sm,a,t} = {} & b_{sm,a} + d_{sm,t} + \sum_{i=1}^{4} \gamma_{i,sm} R_{i,a,t} + \xi_{a,t} K_{sm,\text{after}}(c_{a,t}) \\
& + (1 - \xi_{a,t}) K_{sm,\text{before}}(c_{a,t}) + \epsilon_{sm,a,t}, \quad s \in \{l, h\}, \quad \xi_{a,t} = \left\{ \begin{array}{lcl} 1 & : & c_{a,t} \geq 0 \\ 0 & : & \text{else} \end{array} \right. ,
\end{aligned}
$$

using age and time dummy variables as well as year of birth $c_{a,t}$ as cohort variable, normalized to zero for those aged 25 in the year 1975. The pure, separable cohort effects for those entering the labor market after and before 1975, respectively, are given by

$$
(4) \quad K_{sm,k}(c_{a,t}) = \delta_{k,1,sm} c_{a,t}^2 + \delta_{k,2,sm} c_{a,t}^3 + \delta_{k,3,sm} c_{a,t}^4, \quad k \in \{\text{after,before}\}, \quad s \in \{l, h\}.
$$

The terms $R_{i,a,t}$ capture polynomial interaction terms between age and cohorts in the time derivative of $r_{sm,a,t}$ as defined in MaCurdy and Mroz (1995).[7]

As a second specification, we use polynomials of order four in time instead of time dummies. In both specifications separability of age and time effects on the wage differentials holds if $\gamma_{i,sm} = 0$ for all i. Under this assumption, additive models can be valid representations. Uniform growth in wage ratios holds if additionally the pure effects for the cohorts after 1975 are equal to zero: $\gamma_{i,sm} = \delta_{\text{after},j,sm} = 0$ for all i, j. In this case, the existence of cohort effects is denied for those whose entire working life cycle falls into the observation period. Finally, one may test whether even older cohorts do not face any cohort effects: $\gamma_{i,sm} = \delta_{\text{after},j,sm} = \delta_{\text{before},h,sm} = 0$ for all h, i, j; see MaCurdy and Mroz (1995) for further details.

The approach is also applied to test for the existence of cohort effects in the employment dimension suggested by the graphical inspection in section 2.2. In this case, $r_{sm,a,t}$ in equation (3) is replaced by $\ln(L_{s,a,t}/L_{m,a,t})$.

The detailed estimation results for the cohort effects and the associated tests can be found in tables 5 and 6 in appendix B. Our major findings are that there is evidence for cohort effects in skill wage differentials as well as in relative employment measures.[8] Yet additive separability of age, time, and cohort effects in the evolution of wage differentials does not have to be rejected. Based on these results, the estimation of the structural model introduced and discussed in the subsequent section is in fact justified.

---

[6] It is natural to set the linear cohort effect to zero because in a model with separable age and time effects and only a linear cohort effect, one only observes parallel shifts of the cross-sectional age profiles over time; see Fitzenberger, Hujer, MaCurdy, and Schnabel (2001).

[7] Adapted to our notation, the integrals of interaction terms up to second order are given by $R_{1,a,t} = c_{a,t} a_{a,t}^2/2 + a_{a,t}^3/3$, $R_{2,a,t} = c_{a,t}^2 a_{a,t}^2/2 + 2a_{a,t}^3 c_{a,t}/3 + a_{a,t}^4/4$, $R_{3,a,t} = c_{a,t} a_{a,t}^3/3 + a_{a,t}^4/4$, and $R_{4,a,t} = c_{a,t}^2 a_{a,t}^3/3 + a_{a,t}^4 c_{a,t}/2 + a_{a,t}^5/5$.

[8] Since the restrictive decomposition of cohort and age effects in equation (2) following CL is rejected, we do not discuss the associated results. Though, if accepted as such, this approach suggests the existence of cohort effects as well.

4 Estimation Framework

Building on the stylized facts, we follow CL in applying a model based on the two-level CES production function developed by Sato (1967). The model treats not only workers with different educational attainment, but—well in line with the conjectures of Freeman (1979) and Welch (1979)—also similarly educated workers of different age as imperfect substitutes. Given factor remunerations according to their respective marginal products, it can be transformed into relative wage equations which permit to separate age, time, and cohort effects on the wage gaps—and therefore provides an analytical framework to link the stylized facts outlined above.

Our study extends the analysis of CL in several directions. First, we consider the three skill groups introduced above. Second, we do not only look at skill premia, but also take account of age premia implied by the model. When estimating the model, these two points require system estimation techniques. Third, we additionally allow for the possibility of cohort effects in age-specific productivity terms. Fourth, we are concerned with possible endogeneity of employment and estimate the model with and without instrumental variables.

4.1 The Two-Level CES Model

The Sato (1967) framework suggests a CES model of aggregate production y_t:

$$(5) \qquad y_t = \left(\theta_{l,t} L_{l,t}^{\rho} + \theta_{m,t} L_{m,t}^{\rho} + \theta_{h,t} L_{h,t}^{\rho}\right)^{\frac{1}{\rho}},$$

where $L_{s,t}$, the measures of employment in skill group s and period t, themselves are CES subaggregates of the skill- and time-specific employment quantities $L_{s,a,t}$ of individuals in age groups a:

$$(6) \qquad L_{s,t} = \left[\sum_a \phi_{s,a} L_{s,a,t}^{\pi}\right]^{\frac{1}{\pi}}, \quad s \in \{l, m, h\}.$$

The productivity parameters $\theta_{s,t}$ covering the usual CES distribution parameters as well as the (relative) efficiency terms of the different skill groups are allowed to vary over time to capture (skill-biased) technical change, and $\phi_{s,a}$ map the productivities of the different age classes within the skill classes.[9] $\sigma_S = 1/(1-\rho)$ and $\sigma_A = 1/(1-\pi)$ denote the elasticity of substitution between two skill groups and the elasticity of substitution between different age groups within the same skill group, respectively.

Let wages be determined by the respective marginal products:

$$(7) \qquad \frac{w_{s,a,t}}{w_{\tilde{s},\tilde{a},t}} = \frac{\frac{\partial y_t}{\partial L_{s,a,t}}}{\frac{\partial y_t}{\partial L_{\tilde{s},\tilde{a},t}}} = \frac{\theta_{s,t} \cdot L_{s,t}^{\rho-\pi} \cdot y_t^{1-\rho} \cdot \phi_{s,a} \cdot L_{s,a,t}^{\pi-1}}{\theta_{\tilde{s},t} \cdot L_{\tilde{s},t}^{\rho-\pi} \cdot y_t^{1-\rho} \cdot \phi_{\tilde{s},\tilde{a}} \cdot L_{\tilde{s},\tilde{a},t}^{\pi-1}}$$

[9] Note that the evolution of the relative efficiency terms over time captured by trends in $\theta_{s,t}$ also includes drifts in the overall efficiency of age-specific labor that are common across all age groups but may vary across skill classes. Any changes, e. g., in capital endowments affecting the productivity of the different skills implicitly enter this way.

for all $s, \tilde{s} \in \{l, m, h\}$ and $a, \tilde{a} \in \{27, ..., 52\}$. Then age-specific skill premia $r_{s\tilde{s},a,t} = \ln(w_{s,a,t}/w_{\tilde{s},a,t})$ result as

$$(8) \quad r_{s\tilde{s},a,t} = \ln\left(\frac{\theta_{s,t}}{\theta_{\tilde{s},t}}\right) + \ln\left(\frac{\phi_{s,a}}{\phi_{\tilde{s},a}}\right) - \frac{1}{\sigma_A}\ln\left(\frac{L_{s,a,t}}{L_{\tilde{s},a,t}}\right) + \left[\frac{1}{\sigma_A} - \frac{1}{\sigma_S}\right]\ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right), \quad s \neq \tilde{s}.$$

Moreover, the production technology specifies the skill-specific wage premia across age $r_{s,a\tilde{a},t} = \ln(w_{s,a,t}/w_{s,\tilde{a},t})$ as

$$(9) \quad r_{s,a\tilde{a},t} = \ln\left(\frac{\phi_{s,a}}{\phi_{s,\tilde{a}}}\right) - \frac{1}{\sigma_A}\ln\left(\frac{L_{s,a,t}}{L_{s,\tilde{a},t}}\right), \quad a \neq \tilde{a}.$$

Note that CL base their empirical analysis just on (8) whereas our study considers all information implied by both (8) and (9).

The occurrence of perfect substitutability between different age groups, i. e., $\sigma_A \to \infty$, nests the standard case of a CES with skill groups being homogeneous in the age dimension.[10] One would expect substitutability to be higher within skill groups than across, i. e., $\sigma_A > \sigma_S$. In this case both age group-specific relative employment $\ln(L_{s,a,t}/L_{\tilde{s},a,t})$ and aggregate relative employment $\ln(L_{s,t}/L_{\tilde{s},t})$ exert a negative impact on the skill premia in (8).

Moreover, rewriting equation (8) as

$$(10) \quad r_{s\tilde{s},a,t} = \ln\left(\frac{\theta_{s,t}}{\theta_{\tilde{s},t}}\right) + \ln\left(\frac{\phi_{s,a}}{\phi_{\tilde{s},a}}\right) - \frac{1}{\sigma_S}\ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right) - \frac{1}{\sigma_A}\left[\ln\left(\frac{L_{s,a,t}}{L_{\tilde{s},a,t}}\right) - \ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right)\right]$$

unveils the nature of incorporated cohort effects. If $\ln(L_{s,a,t}/L_{\tilde{s},a,t}) - \ln(L_{s,t}/L_{\tilde{s},t})$ varies over time, there are cohort effects in relative employment in the sense that age-specific relative employment evolves differently from the aggregate measure.[11] If, in addition, σ_A is finite, then differences in cohort size affect $r_{s\tilde{s},a,t}$ through the term in brackets.

To show that cohort effects identify σ_A, consider the general decomposition

$$(11) \quad \ln\left(\frac{L_{s,a,t}}{L_{\tilde{s},a,t}}\right) = \tilde{\psi}_{s\tilde{s},a} + \tilde{\mu}_{s\tilde{s},t-a} + \tilde{\nu}_{s\tilde{s},(a,t-a)}$$

such that the model involves year- (index t), age- (index a), and cohort-specific (index $t-a$) effects. Then,

$$\begin{aligned} r_{s\tilde{s},a,t} &= \ln\left(\frac{\theta_{s,t}}{\theta_{\tilde{s},t}}\right) + \left[\frac{1}{\sigma_A} - \frac{1}{\sigma_S}\right]\ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right) + \ln\left(\frac{\phi_{s,a}}{\phi_{\tilde{s},a}}\right) - \frac{1}{\sigma_A}\tilde{\psi}_{s\tilde{s},a} \\ &\quad - \frac{1}{\sigma_A}\left(\tilde{\mu}_{s\tilde{s},t-a} + \tilde{\nu}_{s\tilde{s},(a,t-a)}\right) \\ (12) \qquad &\equiv \lambda_{s\tilde{s},t} + \psi_{s\tilde{s},a} + \mu_{s\tilde{s},t-a} + \nu_{s\tilde{s},(a,t-a)}, \quad s \neq \tilde{s}. \end{aligned}$$

[10] Note that this occurrence does not preclude differences in the productivity parameters $\phi_{s,a}$ across age groups.

[11] This is what has been tested explicitly in section 3 using the MaCurdy and Mroz (1995) approach.

Observe that equations (3) and (4) used to test for cohort effects in section 3 are flexible parameterizations of (12). It is clear that the identification of σ_A depends on the existence of cohort size effects.

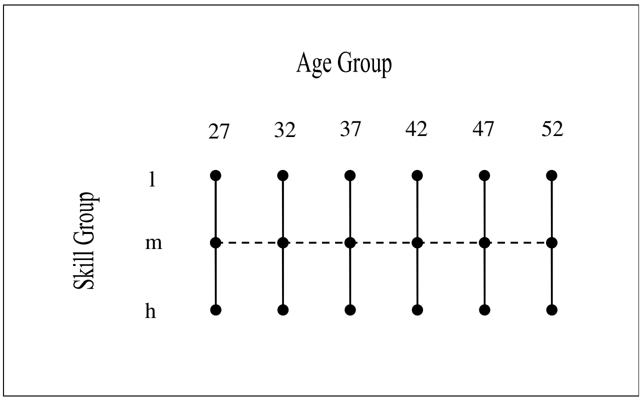
By disregarding variations of age-specific productivity $\phi_{s,a}$ over time, any cohort effects found in the skill wage premia are implicitly attributed to changes in labor quantities. This assumption is suited in light of our main focus to operationalize the relationship between relative wages and employment, and it is not contradicted by the test results of section 3; compare also Juhn, Murphy, and Pierce (1993) and Welch (1979). Yet we also show how it can be relaxed in section 4.3 below.

Apart from being operational for a large number of input factors, being consistent with a neoclassical production function is a great merit of the CES framework. The two-level CES offers the additional advantage that it accounts for an important aspect of heterogeneity within the skill groups: Workers of different age are allowed to be imperfect substitutes.[12]

4.2 Empirical Implementation

Equations (8) and (9) specify all wage ratios $r_{s\tilde{s},a\tilde{a},t}$ across skill and age groups for given t; see figure 1.[13]

Figure 1: System Structure of Wage Ratios



Knots: input factors. Solid lines: age-specific skill wage premia; dashed lines: skill-specific age premia.

The solid lines correspond to the age-specific skill wage premia in equation (8), the dashed lines to the skill-specific age premia in equation (9). With a total of $3 \times 6 = 18$ input factors

[12]Prima facie, one might judge the model's functional form restrictive. In particular, the elasticities of substitution between (identically skilled) workers of different age are restricted to be all equal, so that, say, a 55-year-old executive can be replaced by an experienced 50-year-old as well as by a 25-year-young entrant. However, the model is well-suited to tell apart the effects of the two dimensions age and time. In contrast to feasible translog systems, for example, its age$\times$time dimensioning allows to incorporate a relatively large number of input factors. For discussions on functional specification and aggregation see, e. g., Koebel (2005).

[13]One could also write down wage ratios across time based on the CES production function. However, the output in different time periods y_t would not cancel.

(knots), the system implies $6 \times 3 = 18$ age-specific skill premia plus $3 \times 5(5-1)/2 = 45$ skill-specific age premia, adding up to 63 possible equations. However, there are only 17 independent wage ratios (lines). The additional 46 ratios are redundant in the sense that they can be expressed by means of the 17 independent ones.[14] For example, the wage differential between high-skilled and medium-skilled workers on the one hand and the differential between medium-skilled and low-skilled on the other hand add up to the differential between the high-skilled and the low-skilled, i. e., $r_{hl,a,t} = r_{hm,a,t} + r_{ml,a,t}$. Analogously, age premia add up; for example, $r_{s,3727,t} = r_{s,3732,t} + r_{s,3227,t}$.

The adding-up constraints translate to cross-equation restrictions in the estimation of the system. Our basic approach is to estimate the full-rank system (17 ratios) by means of Feasible GLS (FGLS).[15] In order to achieve invariance with respect to the choice of the excluded equations, we estimate the 63-equation system by System OLS (SOLS) at the first stage. As compared to the approaches pursued in the literature so far, the inclusion of equations for age premia (9) promises more accurate estimation of, in particular, σ_A. The following paragraphs describe the estimation strategy in more detail.

The two-level CES basically entails nonlinear system equations. However, estimation can be achieved by estimating linear models in three steps:[16]

(1) Estimate the equation system

$$
\begin{aligned}
(13)\quad r_{s\tilde{s},a,t} &= b_{s\tilde{s},a} + d_{s\tilde{s},t} - \frac{1}{\sigma_A}\ln\left(\frac{L_{s,a,t}}{L_{\tilde{s},a,t}}\right) + \epsilon_{s\tilde{s},a,t}, \quad s \neq \tilde{s} \\
(14)\quad r_{s,a\tilde{a},t} &= b_{s,a\tilde{a}} - \frac{1}{\sigma_A}\ln\left(\frac{L_{s,a,t}}{L_{s,\tilde{a},t}}\right) + \epsilon_{s,a\tilde{a},t}, \quad a \neq \tilde{a}
\end{aligned}
$$

to obtain an estimate for $\frac{1}{\sigma_A}$, which is equal across equations. At this step, the first-stage SOLS contains the full set of 63 equations, but the system rank, i. e., the number of equations estimated in the second stage of FGLS, is reduced to 15: in comparison with the above number of 17 independent ratios, one additional degree is lost for each of the two estimated skill ratios due to the inclusion of the skill-specific time dummies $d_{s\tilde{s},t}$.

(2) SOLS estimation of the 18-equation system

$$
(15)\quad \ln(w_{s,a,t}) + \frac{1}{\hat{\sigma_A}}\ln(L_{s,a,t}) = d_{s,t} + \ln(\phi_{s,a}) + \epsilon_{s,a,t}
$$

then provides estimates of $\phi_{s,a}$ and allows to calculate the skill group aggregates $L_{s,t}$ defined in (6).

[14]There exist even more redundant wage ratios across both age and skill ($r_{s\tilde{s},a\tilde{a},t}$, $s \neq \tilde{s}$, $a \neq \tilde{a}$).

[15]As will become evident below, there are also cross-equation restrictions within the non-redundant part.

[16]In general, the model can be estimated in one step using nonlinear techniques. Following CL, we proceed in three steps to avoid numerical difficulties. This is the only viable alternative because we apply bootstrapping to obtain standard errors.

(3) Finally, the entire model—equations (9) and (10) extended by additive error terms—can be estimated, using the generated aggregates and taking account of the cross-equations restrictions concerning $1/\sigma_A$ and $1/\sigma_S$. Again, first-stage SOLS employs all 63 equations, but now FGLS uses 17 equations as explained above.

FGLS makes use of the covariance of the error terms across equations within a year. The relative productivity of workers over time, $\ln(\theta_{s,t}/\theta_{\tilde{s},t})$, is assumed to follow a linear time trend. This approach captures the steady demand hypothesis (Acemoglu, 2002)—the steady shift towards a higher relative demand for more highly skilled labor reflects a constant rate of SBTC.[17]

Concerning the age-productivity within skill groups, $\phi_{s,a}$, two specifications are possible: First, $\ln(\phi_{s,a}/\phi_{\tilde{s},a})$ and $\ln(\phi_{s,a}/\phi_{s,\tilde{a}})$ can be estimated freely at the third step, using age dummies in analogy to the first step (model version (a)). Alternatively, $\phi_{s,a}$ may be treated as predetermined by the estimate from the second step (model version (b)). Furthermore, a version (c) would treat σ_A at the third step as predetermined by first-step estimate, and finally, both $\phi_{s,a}$ and σ_A can be taken as predetermined from previous steps (version (d)). We compare the four versions (a) to (d) in a Monte Carlo study in appendix C. By and large, version (a) performs best in terms of closed point estimates and minimum root mean squared error. All of our estimations thus use this specification.

To account for estimation error in all steps of the estimation approach, we obtain bootstrap standard errors (details can be found in appendix D). This is crucial because the third step estimates are based on the generated regressor $L_{s,t}$.

4.3 Model Relaxations and Extensions

We consider two types of model relaxations (specification tests) of the tight specification of the production technology introduced in section 4.1. First, we allow for elasticities of substitution between age groups being different across skill groups by replacing (6) with

$$(16)\quad L_{s,t} = \left[\sum_a \phi_{s,a} L_{s,a,t}^{\pi_s}\right]^{\frac{1}{\pi_s}}, \quad s \in \{l, m, h\}.$$

The third step now estimates

$$(17)\quad r_{s\tilde{s},a,t} = b_{s\tilde{s},a} + \beta_{s\tilde{s}} t - \frac{1}{\sigma_S}\ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right) - \frac{1}{\sigma_{As}}\ln\left(\frac{L_{s,a,t}}{L_{s,t}}\right) + \frac{1}{\sigma_{A\tilde{s}}}\ln\left(\frac{L_{\tilde{s},a,t}}{L_{\tilde{s},t}}\right) + \epsilon_{s\tilde{s},a,t}$$

$$(18)\quad r_{s,a\tilde{a},t} = b_{s,a\tilde{a}} - \frac{1}{\sigma_{As}}\ln\left(\frac{L_{s,a,t}}{L_{s,\tilde{a},t}}\right) + \epsilon_{s,a\tilde{a},t}.$$

[17]See also Murphy and Welch (1992, 2001) and Card and DiNardo (2002) regarding the pros and cons of this hypothesis. We also experimented with modeling breaks in SBTC which might have resulted from German unification; compare the discussion of results in section 5.

This relaxation is quite plausible and the hypothesis $\sigma_{As} = \sigma_A$ for all $s \in \{l, m, h\}$ is easily tested.

A second relaxation, regarding additionally the uniform elasticity of substitution between skill groups, can be implemented at the third step replacing (9), (10) by

$$(19)\quad r_{s\tilde{s},a,t} = b_{s\tilde{s},a} + \beta_{s\tilde{s}} t - \frac{1}{\sigma_{Ss}} \ln(L_{s,t}) + \frac{1}{\sigma_{S\tilde{s}}} \ln(L_{\tilde{s},t}) - \frac{1}{\sigma_{As}} \ln\left(\frac{L_{s,a,t}}{L_{s,t}}\right) + \frac{1}{\sigma_{A\tilde{s}}} \ln\left(\frac{L_{\tilde{s},a,t}}{L_{\tilde{s},t}}\right) + \epsilon_{s\tilde{s},a,t}$$

$$(20)\quad r_{s,a\tilde{a},t} = b_{s,a\tilde{a}} - \frac{1}{\sigma_{As}} \ln\left(\frac{L_{s,at}}{L_{s,\tilde{a}t}}\right) + \epsilon_{s,a\tilde{a},t}$$

and testing whether $\sigma_{Ss} = \sigma_S$ for all $s \in \{l, m, h\}$. Note however that this ad hoc relaxation abandons the theoretical consistency of the model and has to be viewed as a specification test. The parameters σ_{Ss} are no longer elasticities of substitution.

The model specifications discussed so far are referred to as "benchmark version (i)". Next, we describe the model versions (ii) to (ix) which are estimated as extensions of (i) as part of our extensive sensitivity analysis. Table 7 in the appendix conveniently summarizes all estimated versions.

As discussed in section 4.1, a further type of cohort effects could arise if age-specific productivity $\phi_{s,a}$ were allowed to vary with time. This case would match an age bias in the evolution of the returns to (i. e., the price of) experience over time: There might be differential trends in the relative productivity of different age groups thus implying an age/cohort bias in technological progress (see footnote 9). The separability of these particular productivity components from the time effects captured by the educational skill measure $\theta_{s,t}$ has to rely on further functional form assumptions; cf. the discussion about the identification of cohort effects in section 3. A simple and convenient form is to assume multiplicative interaction of age or cohort and time. Then, equation (6) is replaced by

$$(21)\quad L_{s,t} = \left[\sum_a \phi_{s,a} \exp(\delta_s z t) L^{\pi}_{s,a,t}\right]^{\frac{1}{\pi}}, \quad s \in \{l, m, h\},$$

where $z = a$ (version (ii)) or $z = t - a$ (version (iii)), respectively. Then, equations (9), (10) become

$$(22)\quad r_{s\tilde{s},a,t} = b_{s\tilde{s},a} + \beta_{s\tilde{s}} t + (\delta_s - \delta_{\tilde{s}}) z t - \frac{1}{\sigma_S} \ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right) - \frac{1}{\sigma_A}\left[\ln\left(\frac{L_{s,a,t}}{L_{\tilde{s},a,t}}\right) - \ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right)\right] + \epsilon_{s\tilde{s},a,t}$$

$$(23)\quad r_{s,a\tilde{a},t} = b_{s,a\tilde{a}} + \delta_s (z - \tilde{z}) t - \frac{1}{\sigma_A} \ln\left(\frac{L_{s,a,t}}{L_{s,\tilde{a},t}}\right) + \epsilon_{s,a\tilde{a},t},$$

where $\tilde{z} = \tilde{a}$ (version (ii)) or $\tilde{z} = t - \tilde{a}$ (version (iii)), respectively. We test for significance of the additional parameters δ_s.

Further sensitivity tests are straightforward. Following CL, we estimate the limited information version (iv) which involves estimating solely equation (10) for skill wage differentials but disregards the skill-specific age premia (9). Version (v) excludes the youngest group of university graduates (25 to 29 years old) because the descriptive results in section 2.2 showed quite different trends for this group. Version (vi) examines possible breaks in the SBTC trends caused by German unification in 1990.

As comparison with a traditional CES model we estimate model version (vii) with σ_A being restricted to infinity. Focusing on the estimation of σ_S, this version still allows for productivity differences across age. Moreover, we estimate a traditional CES model (version (viii))

$$(24) \quad r_{s\tilde{s},t} = \text{constant}_{s\tilde{s}} + \beta_{s\tilde{s}} t - \frac{1}{\sigma_S} \ln\left(\frac{L_{s,t}}{L_{\tilde{s},t}}\right) + \epsilon_{s\tilde{s},t}, \quad s \neq \tilde{s},$$

again testing for uniqueness of σ_S. Here, time-specific mean wage differences $r_{s\tilde{s},t} = \ln(w_{s,t}/w_{\tilde{s},t})$ are calculated as a weighted average

$$(25) \quad r_{s\tilde{s},t} = \frac{1}{L_{s,t} + L_{\tilde{s},t}} \sum_a (L_{s,a,t} + L_{\tilde{s},a,t})(\omega_{s,a,t} - \omega_{\tilde{s},a,t}), \quad s \neq \tilde{s}$$

of time- and age-specific differences $\omega_{s,a,t}$ pre-estimated by period specific Tobit regressions

$$(26) \quad \ln(w_t) = \sum_s \sum_a \omega_{s,a,t} \cdot d_{s,a,t} + \text{controls}_t + \epsilon_t \ ,$$

where $d_{s,a,t}$ indicate dummies for the different skill$\times$age groups. Alternatively, version (ix) obtains skill differentials $r_{sm,t}$ directly from standard regressions which include separate skill and age dummies:

$$(27) \quad \ln(w_t) = r_{lm,t} \cdot d_{l,t} + r_{hm,t} \cdot d_{h,t} + \sum_a \varpi_{a,t} \cdot d_{a,t} + \text{controls}_t + \epsilon_t.$$

In contrast to the previous versions, (viii) and (ix) average out the age dimension before estimating the elasticity of substitution and $L_{s,t}$ measures aggregate employment as a headcount rather than in efficiency units. Hence, the resulting elasticities should be comparable to those found in the literature for employment in persons.

4.4 Endogeneity of Employment

The different measures of employment (headcounts for $L_{s,a,t}$ and efficiency units for $L_{s,t}$) are crucial here. The estimation approach builds on inverting labor demand and the literature (e. g., CL for the US, UK, and Canada) assumes equality of demand and (effective) supply and supply being inelastic in the short run. Market clearing is highly questionable in the case of Germany, since it disregards unemployment driving a wedge between demand and supply of labor.

Moreover, both observed wage premia, i. e., the relative price of skilled labor, as well as observed relative employment generally result as outcomes of all labor market processes—and should therefore be treated as endogenous in the empirical implementation. Endogeneity of employment follows, for instance, in wage-setting models with right-to-manage (RTM) assumption or efficient bargaining, in which wage bargaining takes account of the firms' employment decisions (McDonald and Solow, 1981, Nickell and Andrews, 1983,, or Arnsperger and de la Croix, 1990).

Under the RTM assumption—in contrast to efficient bargaining—observations on wages and employment lie on the demand curve. Then, the coefficient on relative employment $-1/\sigma$ in any of the models above represents the (negative) relationship between wage premia $r_{s\tilde{s}} = \ln(w_s/w_{\tilde{s}})$ and relative employment $\ln(L_s/L_{\tilde{s}})$ on the demand schedule. Unobserved shocks in output demand affect wages and employment in the same direction. Such shocks render relative employment endogenous and dilute the negative labor demand relation. Least squares estimation then yields (in absolute terms) downward-biased estimates of the true relationship or, put differently, upward-biased estimates for the elasticity of substitution σ.

As a remedy, we implement an instrumental variable (IV) approach. As instruments, we use measures of labor supply, which is assumed inelastic in the short run, possibly due to past human capital investment (Katz and Autor, 1999 and CL).[18] We compile measures of skill×age-specific labor force numbers $L^{\text{supply}}_{s,a,t}$ and aggregate numbers $L^{\text{supply}}_{s,t} = \sum_a L^{\text{supply}}_{s,a,t}$ from German Microcensus data available at the Federal Statistical Office and estimate a system of IV equations with equation-specific instruments as follows.

To instrument age-specific employment $L_{s,a,t}$ at the first and the third step of the estimation approach, we have the $3 \times 6 = 18$-equation system

$$(28) \quad \ln\left(L_{s,a,t}\right) = \alpha_{s,a} + d_{s,t} + \sum_{\tilde{s}} \alpha_{\tilde{s}s} \ln\left(L^{\text{supply}}_{\tilde{s},a,t}\right) + \epsilon_{s,a,t}$$

with skill×age-specific intercepts $\alpha_{s,a}$, skill-specific time dummies $d_{s,t}$, and the impact $\alpha_{\tilde{s}s}$ of the excluded instruments $L^{\text{supply}}_{\tilde{s},a,t}$.[19] The system (28) is also estimated first by SOLS and subsequently by FGLS in order to increase efficiency.[20] We then use predicted employment values $\hat{L}_{s,a,t}$ from (28) at the first and at the third step of the estimation approach; see appendix D for estimation of the covariance of this sequential estimator.

At the third step, we additionally instrument aggregate employment $L_{s,t}$ based on

$$(29) \quad \ln\left(L_{s,t}\right) = \alpha_s + \beta_s t + \sum_{\tilde{s}} \alpha_{\tilde{s}s} \ln\left(L^{\text{supply}}_{\tilde{s},t}\right) + \epsilon_{s,t}$$

[18]Accounting for the endogeneity of relative employment is also crucial as traditional demand analysis treats prices/wages as exogenous (Hamermesh, 1993).

[19]For versions (ii) and (iii), the equations of (28) also involve the terms $\alpha_s at$ and $\alpha_s(t-a)t$, respectively.

[20]Due to the cross-equation restrictions for the time dummies $d_{s,t}$, the rank of the system covariance matrix is reduced to 15, which does not allow us to estimate freely all age specific intercepts $\alpha_{s,a}$ with FGLS. Therefore, we replace the skill×age dummies $\alpha_{s,a}$ with skill-specific age polynomials of order three. The SOLS estimates for $\alpha_{s,a}$ do not differ significantly from polynomials of order three in age.

with skill-specific intercepts α_s, skill-specific linear time trends $\beta_s t$, and $\alpha_{\tilde{s}s}$ as coefficients of the excluded instruments $L^{\text{supply}}_{\tilde{s},t}$.[21] In this case, there are no cross-equation restrictions and the estimation by SOLS and FGLS is straightforward. Predicted values $\hat{L}_{s,t}$ based on (29) are used at the third step of the estimation approach.

Results for the IV equations and tests for significance of the excluded instruments are displayed in tables 11 and 12 in appendix B. The excluded instruments are jointly significant both for age-specific employment and aggregate employment, which we take as sufficient evidence in favor of our IV estimation approach. Note, however, that lack of individual significance is often the case, especially for aggregate employment.

5 Estimation Results

We estimate the model versions (i) to (ix) described in section 4.3 (see table 7 for a short description). Tables 8 to 10 comprise the estimates of the substitution parameters estimated at the first and at the third step and report the results of specification tests.

5.1 Preferred Specifications

Table 1 shows estimates of our preferred specifications (i), (ii), and (iii). Recall that version (i) is the benchmark specification and versions (ii) and (iii) allow for cohort effects in age-specific productivity.

In effect, the estimated elasticity σ_A usually proves finite, meaning that the estimated $1/\sigma_A$ is significantly positive: Employees of different age are imperfect substitutes. The structural model consistently mirrors the dimensions of cohort effects uncovered by the descriptive inspection in section 2. The preferred specifications in table 1 let the elasticity between age groups σ_{As} vary across skill classes (relaxation 1), but stick to a single elasticity of substitution across skill classes σ_S. The assumption of identical σ_A turns out overly restrictive in almost all cases; compare the test results in tables 8 and 9. However, differences in σ_{Ss} across skill groups are significant in most cases for FGLS, but no so for FGLS-IV. Moreover, the restricted FGLS-estimate for σ_S provides reasonable estimates of the average σ_{Ss}. Imposing the restriction of a uniform σ_{Ss} does not affect σ_{As} in a noticeable way. For these reasons, we will focus on the restricted estimates.

Regarding instrumentation, we find only little differences in the point estimates for σ_{As}. Yet the IV estimates for σ_S are considerably reduced. Along the reasoning of the previous section, unobserved shocks affect particularly aggregate relative employment, rendering this measure endogenous and estimation of σ_S without instruments inconsistent. Not surprisingly, though, the estimated standard errors are higher in case of IV estimation such that in some cases σ_S is not statistically different from 1.[22]

[21] In case of the model version (iv), the equations of (29) further include skill-specific breaks in the linear time trends.

[22] Large standard errors of our estimates may result for various reasons: First, the aggregate employment measures included at the third step are pre-generated regressors, the variation of which the bootstrap

Table 1: Elasticities of Substitution, Preferred Specifications of the Nested CES

model version		(i)		(ii)		(iii)	
Estimation		FGLS	FGLS-IV	FGLS	FGLS-IV	FGLS	FGLS-IV
	l	**8.58**	**10.31**	**9.18**	**9.44**	**9.20**	**8.68**
		(0.67)	(1.64)	(0.74)	(1.62)	(0.74)	(1.40)
σ_A	m	**4.81**	**5.27**	**5.23**	**6.01**	**5.22**	**5.38**
		(0.32)	(0.66)	(0.38)	(0.86)	(0.37)	(0.73)
	h	**19.52**	**20.13**	**10.36**	**8.50**	**10.15**	**8.59**
		(5.87)	(11.11)	(1.98)	(2.57)	(1.83)	(2.63)
σ_S		**9.36**	**6.15**	**9.49**	**6.97**	**6.32**	**4.91**
		(0.91)	(2.87)	(0.93)	(2.85)	(0.55)	(1.54)

Model versions: (i) benchmark model; (ii) with age×time interaction in age-specific productivity; (iii) with cohort×time interaction in age-specific productivity. Standard errors in parentheses based on 500 bootstrap repetitions. Bold numbers: Elasticities finite (inverses significant at 0.95 level). Data sources: IABS 1975–1997, German Microcensus.

Our estimates of σ_S, ranging from 4.9 to 6.9 in case of IV, imply a rather high degree of substitutability compared to findings in the related literature; cf. the surveys in Hamermesh (1993) and Katz and Autor (1999). CL report elasticities of substitution between college graduates and high school graduates for Canada, the UK, and the US between 2 and 2.5.[23] In international comparison, our high elasticities are likely to reflect the small amount of over-all wage dispersion as well as the more compressed distribution of skills in Germany; cf. Nickell and Bell (1996) and Freeman and Schettkat (2001). Comparable studies for Germany also take account of three skill types, but they find elasticities not higher than 3.6.[24] Differences in the estimates may be attributed to the selection of data or the model's functional form. We address this issue when discussing additional model

procedure takes account of. Second, FGLS instrumentation has to account for the estimation error in earlier estimation steps. Third, we lose precision by instrumentation. The labor force numbers taken to instrument employment do not differ strongly from linear time trends such that especially σ_S, the coefficient of predicted aggregate employment, is difficult to estimate precisely.

[23]Other studies quantifying elasticities for the US present σ-estimates within a similar range: Autor, Katz, and Kearney (2005), also applying a nested CES model, report elasticities around 1.6. Bound and Johnson (1992), Katz and Murphy (1992), and Krusell, Ohanian, Rios-Rull, and Violante (2000) report 1.8, 1.4, and 1.7, respectively. Ciccone and Peri (2005) prefer a span between 1.2 and 2.2, and Stapleton and Young (1988) note a value of 3.0.

[24]Fitzenberger and Franz (2001) estimate elasticities of substitution between medium- and low-skilled of 0.6–1.4 for manufacturing and of 3.0–3.6 for non-manufacturing industries, while Steiner and Wagner (1998b) and Steiner and Mohr (2000) report values for all three classes of merely 0.3–0.5 for manufacturing and 1.4 for construction and transportation. Falk and Koebel (1999, 2002) find at most substitutability between medium- and low-skilled employees, whereas Koebel, Falk, and Laisney (2003) bilaterally classify high- and medium-skilled as well as medium- and low-skilled as substitutes, but they find complementarity between low- and high-skilled employees. Entorf (1996) finds elasticities between 0.5 and 1.5 for blue and white collar workers and Beißinger and Möller (1998) of 1.8 for males and 3.3 for females.

specifications in the next section.

Employees with different skill levels are more difficult to substitute than those with identical skill levels. The substitutability across different age groups with values of σ_{As} between 5.3 and 8.6 (version (iii), IV) is lowest among the medium-skilled.[25] This finding supports the view that low-skilled employees, mainly in positions which do not require intensive training, can be substituted relatively easily. Contrary to the hypothesis that substitutability between young and old workers diminishes (monotonically) with educational attainment (Welch, 1979),[26] an analogous reasoning applies to university graduates of different age, whose education is often said to provide them with a high competence in general problem solving. Workers with a vocational degree, however, qualify for specific tasks such that, say, younger colleagues can substitute older workers less easily.

5.2 Sensitivity Analysis

Tables 8 to 10 report the results for all different model versions. First of all, table 10 reports the outcomes of models which assume perfect substitution across age classes. Estimates of the nested CES with σ_A restricted to infinity at the third step (version (vii)) as well as from the CES model (viii), which still incorporates age×skill specific intercepts, are very close to the above results. However, we get lower estimates of σ_S between 3.7 (IV) and 4.9 (no IV) from a traditional CES model (ix), which does not allow for age×skill interaction. On the one hand, the fact that the latter estimate is still relatively high in comparison with the literature suggests that prime age males are indeed a relatively homogeneous group. On the other hand, the finding that all specifications (i) to (viii) yield higher elasticities than the more restrictive version (ix) warrants the conclusion that models (including those in the literature) disregarding differences in the relative productivity of the different age classes incorrectly attribute too much variation in (relative) wages to changes in (relative) employment.

Second, disregarding the equations for age premia as in version (iv) in table 9 basically leaves σ_S unchanged, while estimates for σ_A increase. Since this approach does not use the full information content of the system, we consider these results to be less reliable.

Third, the additional interaction terms in versions (ii) and (iii) in table 8 are significant in most cases. Thus, there is evidence for cohort effects in age-specific relative productivities in addition to the cohort effects in relative employment. Yet the resulting elasticities are comparable to those obtained from the benchmark model (i). If anything, σ_{Ah} is lower in the versions with interactions (ii) and (iii).

Fourth, while we have so far assumed a constant rate of SBTC, we also estimated breaks in the linear time trends to capture a possible slowdown or an increase in SBTC resulting

[25] σ_{Am} even turns out lower than σ_S in several specifications reported in tables 8 and 9.

[26] Studies for the US report a much higher degree of substitutability between age classes within the group of high school graduates than among those with a college degree: Freeman (1979) finds elasticities of 14 and 2, respectively (even if the estimated reciprocals of both values show insignificant). Stapleton and Young (1988) note amounts of 73.6 (reciprocal insignificant) and 2.5. CL do not find any significant differences, though. They report significantly finite values of σ_A in the range of 4–6. Our higher estimates then reflect a higher degree of homogeneity within the skill groups defined for Germany, compared to that within the college and high school groups pertinently classifying Anglo-Saxon education systems.

from German unification. The corresponding breaks generally turned out insignificant; see the results for version (vi) in table 9. Finally, the results do not change notably either when we exclude the particular group of 25–29-year-old university graduates[27] from the analysis; see version (v) in table 9. We are thus somewhat confident about the results of our preferred specifications in table 1.

6 Simulation Experiments

In light of the ongoing policy debate about cures for unemployment and the creation of employment, estimates from the above structural model can be used to assess the effect of wage changes on employment by means of simulation experiments.

First, and similar to the experiment conducted by Fitzenberger and Franz (2001), we estimate the magnitude of wage changes in the three skill groups that would be necessary to induce, say, a reduction of unemployment rates by one half in all three skill groups. Due to the particularly high unemployment rate among the low-skilled,[28] this design enquires about a disproportionately sizeable increase in employment of that deprived group, and thus is of high policy relevance. The relative wage changes are assumed to be equal for all age groups within the respective skill groups: $\Delta\ln(w_{s,a}) = \Delta\ln(\bar{w}_s)$ for all a. While allowing for relative changes between skill groups, this leaves the wage structure within skill groups unaffected.

The calculations are done for the latest available year 1997. The time index t is omitted for notational simplicity. We use a first order Taylor approximation of overall employment in each skill group s as the sum of employment in the respective age groups a:

$$(30)\quad L^*_s = \sum_a L^*_{s,a} = \sum_a \left(L_{s,a} + \sum_{\tilde{s}} \sum_{\tilde{a}} \frac{\partial L_{s,a}}{\partial \ln(w_{\tilde{s},\tilde{a}})} \Delta \ln(w_{\tilde{s},\tilde{a}}) \right), \quad s \in \{l, m, h\},$$

where L^*_s, $L^*_{s,a}$ are the employment targets consistent with the goal to reduce unemployment rates by one half. Drawing on the wage elasticity of labor demand

$$(31)\quad \eta_{s\tilde{s},a\tilde{a}} = \frac{\partial L_{s,a}}{\partial w_{\tilde{s},\tilde{a}}} \frac{w_{\tilde{s},\tilde{a}}}{L_{s,a}} = \frac{\partial \ln(L_{s,a})}{\partial \ln(w_{\tilde{s},\tilde{a}})} = \frac{\partial L_{s,a}}{\partial \ln(w_{\tilde{s},\tilde{a}})} \frac{1}{L_{s,a}},$$

equation (30) can be written in terms of relative changes:

$$(32)\quad \frac{\Delta L_s}{L_s} = \frac{L^*_s - L_s}{L_s} = \sum_a \frac{L_{s,a}}{L_s} \sum_{\tilde{s}} \sum_{\tilde{a}} \eta_{s\tilde{s},a\tilde{a}} \Delta \ln(w_{\tilde{s},\tilde{a}}).$$

The relationship between wage elasticities $\eta_{s\tilde{s},a\tilde{a}}$, Allen-Uzawa elasticities of substitution $\sigma_{s\tilde{s},a\tilde{a}}$, and cost shares $S_{s,a}$ implied by cost minimizing behavior of employers is given by

$$(33)\quad \eta_{s\tilde{s},a\tilde{a}} = S_{\tilde{s},\tilde{a}} \sigma_{s\tilde{s},a\tilde{a}} + S_{\tilde{s},\tilde{a}} \eta \quad \text{for} \quad a \neq \tilde{a} \quad \vee \quad s \neq \tilde{s},$$

[27] Compare footnote 4.

[28] The skill-specific and age-specific rates of unemployment in West Germany our simulations make use of are displayed in appendix A.

where η denotes the price elasticity of product demand and

$$(34)\quad \eta_{ss,aa} = \eta - \sum_{\tilde{s}} \sum_{\tilde{a} \neq a} \eta_{s\tilde{s},a\tilde{a}} - \sum_{\tilde{s} \neq s} \eta_{s\tilde{s},aa} = S_{s,a}\eta - \sum_{\tilde{s}} \sum_{\tilde{a} \neq a} S_{\tilde{s},\tilde{a}} \sigma_{s\tilde{s},a\tilde{a}} - \sum_{\tilde{s} \neq s} S_{\tilde{s},a} \sigma_{s\tilde{s},aa};$$

see, e. g., Hamermesh (1993). Based on the nested CES production function, inter-class Allen-Uzawa partial elasticities of substitution and intra-class elasticities,[29] write

$$(35)\quad \sigma_{s\tilde{s},a\tilde{a}} = \sigma_S \quad \text{for} \quad s \neq \tilde{s}, \quad \text{and} \quad \sigma_{ss,a\tilde{a}} = \sigma_S + \frac{1}{S_s}(\sigma_A - \sigma_S) \quad \text{for} \quad a \neq \tilde{a}.$$

On principle, cost shares for the nested CES model can be derived directly from the model via Shepard's Lemma as functions of the productivity parameters θ_s and $\phi_{s,a}$ and wages $w_{s,a}$; cf., for example, Chung (1994). Yet the actual calculation fails this way due to the underidentification of the productivity parameters. Hence, we employ observed cost shares

$$(36)\quad S_{s,a} = \frac{w_{s,a} L_{s,a}}{\sum_{\tilde{s}} \sum_{\tilde{a}} w_{\tilde{s},\tilde{a}} L_{\tilde{s},\tilde{a}}} \quad \text{and} \quad S_s = \sum_a S_{s,a}.$$

The targeted relative change of employment can be inferred from the unemployment rates $ur_s = U_s/WF_s = 1 - L_s/WF_s$, where U_s and WF_s denote unemployment and work force in skill group s, respectively:

$$(37)\quad \frac{\Delta L_s}{L_s} = \frac{L_s^* - L_s}{L_s} = \frac{(0.5WF_s + 0.5L_s) - L_s}{L_s} = 0.5 \frac{ur_s}{1 - ur_s}.$$

As η we take a weighted average of the elasticities estimated by Fitzenberger and Franz (2001) separately for the manufacturing and the non-manufacturing sector, with employment ratios in the respective sectors as weights.

Since we set $\Delta \ln(w_{s,a}) = \Delta \ln(\bar{w}_s)$ for all a, the system (32) yields unique solutions for the necessary wage changes based on our estimation results. The calculation of standard errors is based on the errors of the estimated parameters.

Alternatively, one might be interested in changes of the wage structure within the skill groups, holding the structure across the respective groups constant. In this context, the model set-up allows us to answer the question how the wages for employees of different age would have to change—identically in all skill groups—to reduce all age-specific unemployment rates $ur_a = U_a/WF_a = 1 - L_a/WF_a$ by one half.

In analogy to (30), we write

$$(38)\quad L_a^* = \sum_s L_{s,a}^* = \sum_s \left(L_{s,a} + \sum_{\tilde{s}} \sum_{\tilde{a}} \frac{\partial L_{s,a}}{\partial \ln(w_{\tilde{s},\tilde{a}})} \Delta \ln(w_{\tilde{s},\tilde{a}}) \right), \quad a \in \{27, ..., 52\}.$$

[29] For model relaxation (16), σ_A in equation (35) has to be replaced by σ_{As}.

Now assuming $\Delta \ln(w_{s,a}) = \Delta \ln(\bar{w}_a)$ for all s, the system

$$(39) \quad \frac{\Delta L_a}{L_a} = \frac{L_a^* - L_a}{L_a} = \sum_s \frac{L_{s,a}}{L_a} \sum_{\tilde{s}} \sum_{\tilde{a}} \eta_{s\tilde{s},a\tilde{a}} \Delta \ln(\bar{w}_{\tilde{a}})$$

can be solved for the necessary wage changes within the skill groups.

To evaluate the respective real magnitudes of the wage changes, we calculate the price adjustments induced by the nominal wage reductions. Here, the assumption of profit maximizing behavior under monopolistic competition takes account of endogenous output effects. We consider the Amoroso-Robinson relation for the output price level p and a constant elasticity of product demand η,

$$(40) \quad \left(1 + \frac{1}{\eta}\right) p = MC, \quad \text{such that} \quad d\ln(p) = d\ln(MC),$$

with marginal costs

$$(41) \quad MC = \sum_s \sum_a w_{s,a} \frac{\partial L_{s,a}}{\partial y} = \sum_s \sum_a w_{s,a} \frac{L_{s,a}}{y} \frac{\partial L_{s,a}}{\partial y} \frac{y}{L_{s,a}} = \sum_s \sum_a \frac{w_{s,a} L_{s,a}}{y}.$$

The last equality in (41) follows from the constant returns to scale assumption. Relative price changes then arise from (40) as

$$(42) \quad d\ln(p) = \frac{\sum_s \sum_a \frac{L_{s,a} w_{s,a}}{y} d\ln(w_{s,a})}{\sum_{\tilde{s}} \sum_{\tilde{a}} \frac{L_{\tilde{s},\tilde{a}} w_{\tilde{s},\tilde{a}}}{y}} = \sum_s \sum_a \frac{L_{s,a} w_{s,a}}{\sum_{\tilde{s}} \sum_{\tilde{a}} L_{\tilde{s},\tilde{a}} w_{\tilde{s},\tilde{a}}} d\ln(w_{s,a}).$$

Now let $\Delta \ln(w_{s,a}) = \Delta \ln(\bar{w}_s)$ for all a in the first experiment. Then,

$$(43) \quad \Delta \ln(p) = \sum_s \Delta \ln(\bar{w}_s) \sum_a \frac{L_{s,a} w_{s,a}}{\sum_{\tilde{s}} \sum_{\tilde{a}} L_{\tilde{s},\tilde{a}} w_{\tilde{s},\tilde{a}}}.$$

In the second experiment, $\Delta \ln(w_{s,a}) = \Delta \ln(\bar{w}_a)$ for all s, and so

$$(44) \quad \Delta \ln(p) = \sum_a \Delta \ln(\bar{w}_a) \sum_s \frac{L_{s,a} w_{s,a}}{\sum_{\tilde{s}} \sum_{\tilde{a}} L_{\tilde{s},\tilde{a}} w_{\tilde{s},\tilde{a}}}.$$

Table 2 displays the outcome of the first simulation experiment and compares it to results obtained in Fitzenberger and Franz (2001).

Considering the employment target of reducing skill-specific unemployment rates, wages paid are too high in all skill groups, and the necessary wage reductions—ranging from 8.8 to 12.2%—are the higher the lower the skill level. This result provides evidence for wage compression across skill groups. The fact that estimated wage reductions appear rather modest may be ascribed to at least two reasons: on the one hand to the high wage

Table 2: Wage Changes for Different Skill Groups Necessary to Halve Skill-Specific Unemployment Rates in 1997 and Induced Price Change

Model	$\Delta\ln(w_l)$	$\Delta\ln(w_m)$	$\Delta\ln(w_h)$	$\Delta\ln(p)$
(i) FGLS[a]	-0.109	-.093	-0.091	-0.094
	(0.0139)	(0.0139)	(0.0139)	(0.0139)
(i) FGLS-IV[a]	-0.117	-0.092	-0.089	-0.094
	(0.0175)	(0.0135)	(0.0137)	(0.0135)
(ii) FGLS[a]	-0.109	-0.093	-0.091	-0.094
	(0.0139)	(0.0139)	(0.0139)	(0.0139)
(ii) FGLS-IV[a]	-0.114	-0.093	-0.090	-0.094
	(0.0176)	(0.0150)	(0.0151)	(0.0151)
(iii) FGLS[a]	-0.116	-0.092	-0.089	-0.094
	(0.0140)	(0.0139)	(0.0139)	(0.0139)
(iii) FGLS-IV[a]	-0.122	-0.092	-0.088	-0.094
	(0.0173)	(0.0151)	(0.0152)	(0.0151)
F/F (2001)[b]	-0.141	-0.103	-	-0.105
	(0.019)	(0.020)	(-)	(0.020)
F/F (2001)[c]	-0.342	-0.313	-	-0.314
	(0.099)	(0.020)	(-)	(0.020)

[a] Calculations based on the results displayed in table 1. Standard errors in parentheses based on 500 bootstrap repetitions.
[b] Fitzenberger and Franz (2001), specification 4; assumption of constant returns to scale; elasticities of substitution between the high-skilled on the one hand and medium- and low-skilled on the other restricted to equal 1; no changes in wages and employment for the high-skilled; results for 1995.
[c] Fitzenberger and Franz (2001), specification 3; elasticities of substitution between the high-skilled on the one hand and medium- and low-skilled on the other restricted to equal 1; no changes in wages and employment for the high-skilled; results for 1995.

elasticities resulting from the substantial elasticities of substitution, and to the assumption of constant returns to scale on the other. The latter point becomes evident by the comparison of our results to those of Fitzenberger and Franz (2001): Their specification 4, which likewise postulates constant returns to scale, yields estimates very similar to ours, whilst their unrestricted specification 3 indicates higher (nominal) reductions. The range of dispersion, however, turns out rather similar in all models.

The induced relative price changes are a weighted average of the wage reductions; compare equation (43). Thus, given our estimates of nominal wage reductions, the high-skilled experience a real wage increase, whereas the low-skilled face real losses ex constructione.

To put this result into perspective, some remarks are in order: First, the experiment models a shock to the employment decision of the firm—we do not attempt to account for supply-side reactions to the wage changes. Second, capital and other inputs are assumed to be constant, as well.[30] Third, and similarly, the simulation does not consider substi-

[30]None of the simulations reported in table 2 takes into consideration substitution effects with respect

tutability with respect to participants in different labor market segments, like women or employees in mini jobs not subject to social security contributions.

The results of the second experiment, regarding a reduction of age specific unemployment rates, are displayed in table 3.

Table 3: Wage Changes for Different Age Groups Necessary to Halve Age-Specific Unemployment Rates in 1997 and Induced Price Change

Model	$\Delta\ln(w_{27})$	$\Delta\ln(w_{32})$	$\Delta\ln(w_{37})$	$\Delta\ln(w_{42})$	$\Delta\ln(w_{47})$	$\Delta\ln(w_{52})$	$\Delta\ln(p)$
(i) FGLS[a]	-0.087	-0.087	-0.086	-0.086	-0.086	-0.087	-0.087
	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)
(i) FGLS-IV[a]	-0.087	-0.087	-0.087	-0.086	-0.086	-0.087	-0.087
	(0.0126)	(0.0125)	(0.0125)	(0.0125)	(0.0125)	(0.0125)	(0.0125)
(ii) FGLS[a]	-0.088	-0.087	-0.086	-0.086	-0.086	-0.087	-0.087
	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)
(ii) FGLS-IV[a]	-0.088	-0.087	-0.087	-0.086	-0.086	-0.087	-0.087
	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)
(iii) FGLS[a]	-0.088	-0.087	-0.087	-0.086	-0.086	-0.087	-0.087
	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)
(iii) FGLS-IV[a]	-0.088	-0.087	-0.08	-0.086	-0.086	-0.087	-0.087
	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)

Calculations based on the results displayed in table 1. Standard errors in parentheses based on 500 bootstrap repetitions.

The calculated wage reductions in the different age groups are very similar. Yet the small degree of variation comes as no surprise because the differences in unemployment rates across the age classes are rather small. In particular, it has to be recalled that very young as well as older participants close to (early) retirement age, who can be expected to face deviant labor market conditions that result in differing unemployment rates, have been excluded from the analysis. For our sample of prime age males there is no evidence of wage compression across the age distribution. As to the underlying high elasticities of substitution and concerning the interpretation of the induced price changes, the same caveats as for the first experiment apply.

7 Conclusions

The evolution of age-specific skill wage premia in the German labor market between 1975 and 1997 shows that the age profiles of skill wage differentials have not moved in parallel fashion over time, but rather experienced a twist. Accordingly, it is unlikely that these developments are associated merely with age and time effects which apply uniformly to

to intermediate inputs or capital stocks. Given capital-skill complementarities, for example, the reported numbers might overstate actual necessary wage changes. For the importance of capital issues in labor demand cf. Krusell, Ohanian, Rios-Rull, and Violante (2000).

all cohorts. Furthermore, we observe a break in the inter-cohort trend of skill- and age-specific relative employment such that young birth cohorts do not follow the path of the older ones towards further skill upgrading. The empirical evidence thus suggests the existence of cohort effects affecting the evolution of both skill wage premia and relative employment. Following the approach suggested in MaCurdy and Mroz (1995), we find such cohort effects for both relative employment and wage premia.

A coherent operationalization of wages and employment in a labor demand framework is generally difficult due to the heterogeneous nature of the input factor labor. We extend the structural approach of Card and Lemieux (2001) based on the nested CES model of Sato (1967), giving rise to a complex picture of German labor demand. On the one hand, the model consistently maps rational behavior within the framework of neoclassical production theory. On the other hand, its age×time dimensioning allows to incorporate a relatively large number of input factors. That way, we analyze wage differences between 18 types of labor.

The results are compatible with the steady demand hypothesis of a constant rate of SBTC as in Acemoglu (2002). Moreover, employees of different age are found to be imperfect substitutes—the model indeed takes account of age, time, and cohort effects. Our preferred specifications estimate the elasticity of substitution between skill groups to range between 4.9 and 6.9, and the elasticity of substitution between age groups between 5.2 and 20.1. Compared to the literature, these numbers are rather high. In international comparison, this finding reflects the fairly small amount of over-all wage dispersion in Germany as well as the relatively compressed distribution of skills. In comparison with alternative studies using different functional forms to model labor demand in Germany, we reckon that, on the one hand, our focus on prime age male employees in the IABS in fact results in a considerably homogeneous sample. On the other hand, approaches in the literature which disregard the interaction of skill and age are likely to report spuriously small elasticities.

On the basis of the estimated parameters, simulation experiments allow for policy-relevant implications. We simulate the magnitude of wage changes in the different skill groups that would have been necessary to reduce skill-specific unemployment rates in 1997 by one half. With wage changes equal for all age groups within the respective skill classes, this would have left the wage structure within skill groups unaffected. The necessary nominal wage changes range between 8.8 and 12.2% and are the higher the lower the employees' qualification. This finding provides evidence for the existence of wage compression—relative to a situation with reduced unemployment, there is too little wage dispersion across the different skill groups.

Our analysis shows the necessity to integrate different dimensions of heterogeneity into empirically meaningful models of labor demand. The nested CES approach allows to do this in a parsimonious way. However it comes at the price of strong functional form assumptions.

As a final caveat, it should be mentioned that our neoclassical production function framework fails to incorporate residual wage inequality that remains within cells defined by skill, age, and year. Residual wage inequality is a major part of total wage dispersion (Juhn, Murphy, and Pierce (1993) and Fitzenberger, Garloff, and Kohn (2003)). This should be taken account of in future research on the link between wage differences and labor demand. Yet no conceptual framework exists so far to do so.

References

Acemoglu, D. (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40, 7–72.

Arnsperger, C., and D. de la Croix (1990): "Wage Bargaining with Price-Setting Firm," *Bulletin of Economic Research*, 42, 285–298.

Autor, D. H., L. F. Katz, and M. S. Kearney (2005): "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists," Working Paper 11627, NBER.

Becker, G. S. (1993): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* The University of Chicago Press, Chicago, London, 3rd edn.

Beissinger, T., and J. Möller (1998): "Wage Flexibility and Employment Performance: A Microdata Analysis of Different Age-Education Groups for Germany," Discussion Paper 307, University of Regensburg.

Bender, S., A. Haas, and C. Klose (2000): "The IAB Employment Subsample 1975–1995," *Schmollers Jahrbuch*, 120(4), 649–662.

Bender, S., J. Hilzendegen, G. Rohwer, and H. Rudolph (1996): *Die IAB-Beschäftigtenstichprobe 1975–1990*, Beiträge zur Arbeitsmarkt- und Berufsforschung 197. Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Boockmann, B., and V. Steiner (2000): "Cohort effects and the returns to education in West Germany," Discussion Paper 00-05, ZEW Mannheim.

Bound, J., and G. Johnson (1992): "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *American Economic Review*, 82(3), 371–392.

Card, D., and J. E. DiNardo (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4), 733–783.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116, 705–746.

Christensen, B. (2003): "Die Entwicklung der qualifikatorischen Lohndifferenzierung in Deutschland," *Die Weltwirtschaft*, pp. 313–322, Heft 3.

Christensen, B., and A. Schimmelpfennig (1998): "Arbeitslosigkeit, Qualifikation und Lohnstruktur in Westdeutschland," *Die Weltwirtschaft*, pp. 177–186, Heft 2.

Chung, J. W. (1994): *Utility and Production Functions.* Blackwell Publishers, Oxford; Cambridge, MA.

Ciccone, A., and G. Peri (2005): "Long-run Substitutability Between More and Less Educated Workers: Evidence from US States 1950–1990," *Review of Economics and Statistics*, 87(4), 652–663.

Entorf, H. (1996): "Strukturelle Arbeitslosigkeit in Deutschland: Mismatch, Mobilität und Technischer Wandel," in *Arbeitslosigkeit und Möglichkeiten ihrer Überwindung*, ed. by B. Gahlen, H. Hesse, and H. J. Ramser, pp. 139–170. Mohr Siebeck, Wirtschaftswissenschaftliches Seminar Ottobeuren 25.

Falk, M., and B. Koebel (1999): "Curvature Conditions and Substitution Pattern among Capital, Energy, Materials and Heterogeneous Labour," Discussion Paper 99-06, ZEW Mannheim.

——— (2002): "Outsourcing, Imports, and Labour Demand," *Scandinavian Journal of Economics*, 104(4), 567–586.

Fitzenberger, B. (1999): *Wages and Employment Across Skill Groups: An Analysis for West Germany*. Physica, Heidelberg.

Fitzenberger, B., and W. Franz (2001): "Jobs. Jobs? Jobs! Orientierungshilfen für den Weg zu mehr Beschäftigung," in *Wirtschaftspolitische Herausforderungen an der Jahrhundertwende*, ed. by W. Franz, H. Hesse, H. J. Ramser, and M. Stadler, pp. 3–43. Mohr Siebeck, Wirtschaftswissenschaftliches Seminar Ottobeuren 30.

Fitzenberger, B., A. Garloff, and K. Kohn (2003): "Beschäftigung und Lohnstrukturen nach Qualifikationen und Altersgruppen: Eine empirische Analyse auf Basis der IAB-Beschäftigtenstichprobe," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 36(4), 509–524.

Fitzenberger, B., R. Hujer, T. E. MaCurdy, and R. Schnabel (2001): "Testing for Uniform Wage Trends in West-Germany: A Cohort Analysis Using Quantile Regressions for Censored Data," *Empirical Economics*, 26, 41–86.

Fitzenberger, B., and F. Reize (2003): "Quantilsregressionen der westdeutschen Verdienste: Ein Vergleich zwischen Gehalts- und Lohnstrukturerhebung und der IAB-Beschäftigtenstichprobe," in *Empirische Wirtschaftsforschung: Methoden und Anwendungen*, ed. by W. Franz, H. J. Ramser, and M. Stadler, pp. 255–282. Mohr Siebeck, Wirtschaftwissenschaftliches Seminar Ottobeuren 32.

Fitzenberger, B., R. Schnabel, and G. Wunderlich (2004): "The Gender Gap in Labor Market Participation and Employment: A Cohort Analysis for West Germany," *Journal of Population Economics*, 17, 83–116.

Fitzenberger, B., and G. Wunderlich (2002): "Gender Wage Differences in West Germany: A Cohort Analysis," *German Economic Review*, 3(4), 379–414.

Freeman, R., and R. Schettkat (2001): "Skill compression, wage differentials and employment: Germany vs the US," *Oxford Economic Papers*, 53(3), 582–603.

Freeman, R. B. (1979): "The Effect of Demographic Factors on Age-Earnings Profiles," *Journal of Human Resources*, 14(3), 289–318.

Hamermesh, D. S. (1993): *Labor Demand*. Princeton University Press, Princeton, NJ.

Heckman, J. J., and R. Robb (1985): "Using Longitudinal Data To Estimate Age, Period, and Cohort Effects in Earnings Equations," in *Cohort Analysis in Social Research*, ed. by W. Mason, and S. Fienberg, pp. 138–150. Springer.

Johnson, G. E. (1997): "Changes in Earnings Inequality: The Role of Demand Shifts," *Journal of Economic Perspectives*, 11(2), 41–54.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–442.

Katz, L. F., and D. H. Autor (1999): "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, chap. 26, pp. 1463–1555. Elsevier Science.

Katz, L. F., and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35–78.

Koebel, B. (2005): "Exports and Labour Demand: Searching for Functional Structure in Multi-Output Multi-Skill Technologies," Discussion Paper 1672, IZA.

Koebel, B., M. Falk, and F. Laisney (2003): "Imposing and Testing Curvature Conditions on a Box-Cox Cost Function," *Journal of Business and Economic Statistics*, 21, 319–335.

Krusell, P., L. E. Ohanian, J.-V. Rios-Rull, and G. L. Violante (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–1053.

Machin, S. (2002): "Labour Market Inequality and Changes in the Relative Demand for Skills," Paper for the Royal Statistical Society Conference on Explanations of Rising Economic Inequality.

MaCurdy, T. E., and T. Mroz (1995): "Measuring Macroeconomic Shifts in Wages from Cohort Specifications," unpublished manuscript, Stanford University and University of North Carolina.

McDonald, I. M., and R. M. Solow (1981): "Wage Bargaining and Employment," *American Economic Review*, 71, 896–908.

Mincer, J. (1974): *Schooling, Experience, and Earnings*. National Bureau of Economic Research, New York.

Möller, J. (1999): "Die Entwicklung der qualifikatorischen Lohn- und Beschäftigungsstruktur in Deutschland," *Jahrbücher für Nationalökonomie und Statistik*, 219(1), 8–31.

Murphy, K. M., and F. Welch (1992): "The Structure of Wages," *Quarterly Journal of Economics*, 107(1), 285–326.

——— (2001): "Wage Differentials in the 1990s: Is the Glass Half-Full or Half-Empty?," in *The Causes and Consequences of Increasing Inequality*, ed. by F. Welch, chap. 12, pp. 341–364. University of Chicago Press.

Nickell, S., and M. Andrews (1983): "Unions, Real Wages and Employment in Britain 1951–79," *Oxford Economic Papers*, 35, 183–206.

Nickell, S., and B. Bell (1996): "Changes in the Distribution of Wages and Unemployment in OECD Countries," *American Economic Review*, 86(2, Papers and Proceedings), 302–308.

OECD (ed.) (2004): *Education at a Glance – OECD Indicators 2004*. Organisation for Economic Co-operation and Development, Paris.

Prasad, E. S. (2004): "The Unbearable Stability of the German Wage Structure: Evidence and Interpretation," *IMF Staff Papers*, 51(2), 354–385.

Reinberg, A., and M. Hummel (1999): *Bildung und Beschäftigung im vereinigten Deutschland*, Beiträge zur Arbeitsmarkt- und Berufsforschung 226. Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

——— (2002): *Qualifikationsspezifische Arbeitslosenquoten – reale Entwicklung oder statistisches Artefakt?*, IAB-Werkstattbericht 4/2002. Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Riphahn, R. (2003): "Bruttoeinkommensverteilung in Deutschland 1984–1999 und Ungleichheit unter ausländischen Erwerbstätigen," in *Wechselwirkungen zwischen Arbeitsmarkt und sozialer Sicherung II*, ed. by W. Schmähl, vol. 294 of *Schriften des Vereins für Socialpolitik*, pp. 135–174. Duncker & Humblot, Berlin.

Sato, K. (1967): "A Two-Level Constant-Elasticity-of-Substitution Production Function," *Review of Economic Studies*, 34(2), 201–218.

Stapleton, D. C., and D. J. Young (1988): "Educational Attainment and Cohort Size," *Journal of Labor Economics*, 6, 330–361.

Statistisches Bundesamt (ed.) (1998): *Statistisches Jahrbuch 1998 für die Bundesrepublik Deutschland*. Metzler-Poeschel, Stuttgart.

Steiner, V., and R. Mohr (2000): "Industrial Change, Stability of Relative Earnings, and Substitution of Unskilled Labor in West Germany," in *Policy Measures for Low-Wage Employment in Europe*, ed. by W. Salverda, B. Nolan, and C. Lucifora, pp. 169–190. Edward Elgar.

Steiner, V., and K. Wagner (1997): "Entwicklung der Ungleichheit der Erwerbseinkommen in Westdeutschland," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 30(3), 638–641.

——— (1998a): "Has Earnings Inequality in Germany Changed in the 1980's?," *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 118(1), 29–59.

——— (1998b): "Relative Earnings and the Demand for Unskilled Labor in West German Manufacturing," in *Globalization, Technological Change, and Labor Markets*, ed. by S. W. Black. Kluwer.

Topel, R. H. (1997): "Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality," *Journal of Economic Perspectives*, 11(2), 55–74.

Welch, F. (1979): "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," *Journal of Political Economy*, 87(5), S65–S97.

A Data

Throughout the empirical investigation, we make use of the IAB employment subsample (IABS) 1975–1997, a representative 1% random draw of German employees with employment spells subject to social insurance contributions. Excluding civil servants, self-employed, and freelancers, the IABS covers about 80% of all employed persons. For an extensive description of these register-based data see Bender, Hilzendegen, Rohwer, and Rudolph (1996) and Bender, Haas, and Klose (2000). Selected data at first comprise spells of both men and women employed full-time in West-Germany, excluding parallel employment and training spells.

We restrict attention to prime-age employees between 25 and 55 years to circumvent a number of sample selection problems. Since the IABS contains no information on hours worked, we undertake a headcount to derive an employment measure, weighting each observation with the length of the respective employment spell. This procedure assumes that the number of, say, monthly hours does not change over time nor does it differ by individual, justifying the concentration on full-time employees only.

Concerning the wage data, Steiner and Wagner (1997) report a structural break between 1983 and 1984. In order not to deceivingly interpret this as increasing wage inequality across skill groups, we apply the correction procedure suggested by Fitzenberger (1999).

Observations are classified into three skill groups according to the individuals' educational attainment. The group of the low-skilled consists of employees without any vocational training. Those with a vocational training degree are considered medium-skilled, and individuals with a university or technical college degree form the group of the high-skilled. To deal with measurement error in the education information when defining the skill groups, we correct the skill information such that formal degrees an individual has once obtained are not lost later.

Stage zero of the estimation approach estimates wage differentials by means of Tobit regressions due to the censoring of wage data induced by the social security taxation threshold (*Beitragsbemessungsgrenze*). Observations are weighted by the length of the respective employment spells. As a first approach, equation (1) includes dummies for foreigners and women as control variables and further allows for possible interactions of these with the skill variables. Besides, a linear age term captures variation within the age classes. Cross terms of female and skill dummies prove significant in nearly all cells. Consequently, we base our analysis on males only. Period-specific wage differentials for the traditional CES are similarly estimated by pre-step Tobit estimations (26), using age-specific skill dummies and a dummy for foreigners.

Estimation equations at the first and at the second step include a full set of age dummies and time dummies for 1976–1997. The latter are replaced by a linear time trend at the third step.

At steps one and three we instrument observed employment measures by means of the size of the labor force obtained from the German Microcensus, a representative 1% population sample collected annually, typically via face-to-face interviews. We use representative subsamples available through the Federal Statistical Office (*Statistisches Bundesamt*). The cell-specific labor force is imputed as the sum of (male) employed and unemployed workers within the skill×age groups. For several years within our sampling period, however, individual records of educational attainment were voluntary, leading to sizable shares of missing values. We apply the procedure developed in Fitzenberger, Schnabel, and Wunderlich (2004) to assign the shares of missings to the three skill groups in each cell. For the years without any skill information in the German Microcensus, we interpolate; see also Fitzenberger (1999).

For the first simulation experiment, skill-specific unemployment rates are taken from Reinberg and Hummel (2002). Rates for low-, medium-, and high-skilled males in 1997 read 27.1%, 6.8%, and 3.0%, respectively. Age group-specific unemployment rates for the second experiment are calculated based on Statistisches Bundesamt (1998). For the six age groups (from young to old) the rates are 8.5%, 7.5%, 7.4%, 7.1%, 7.0%, and 8.1%.

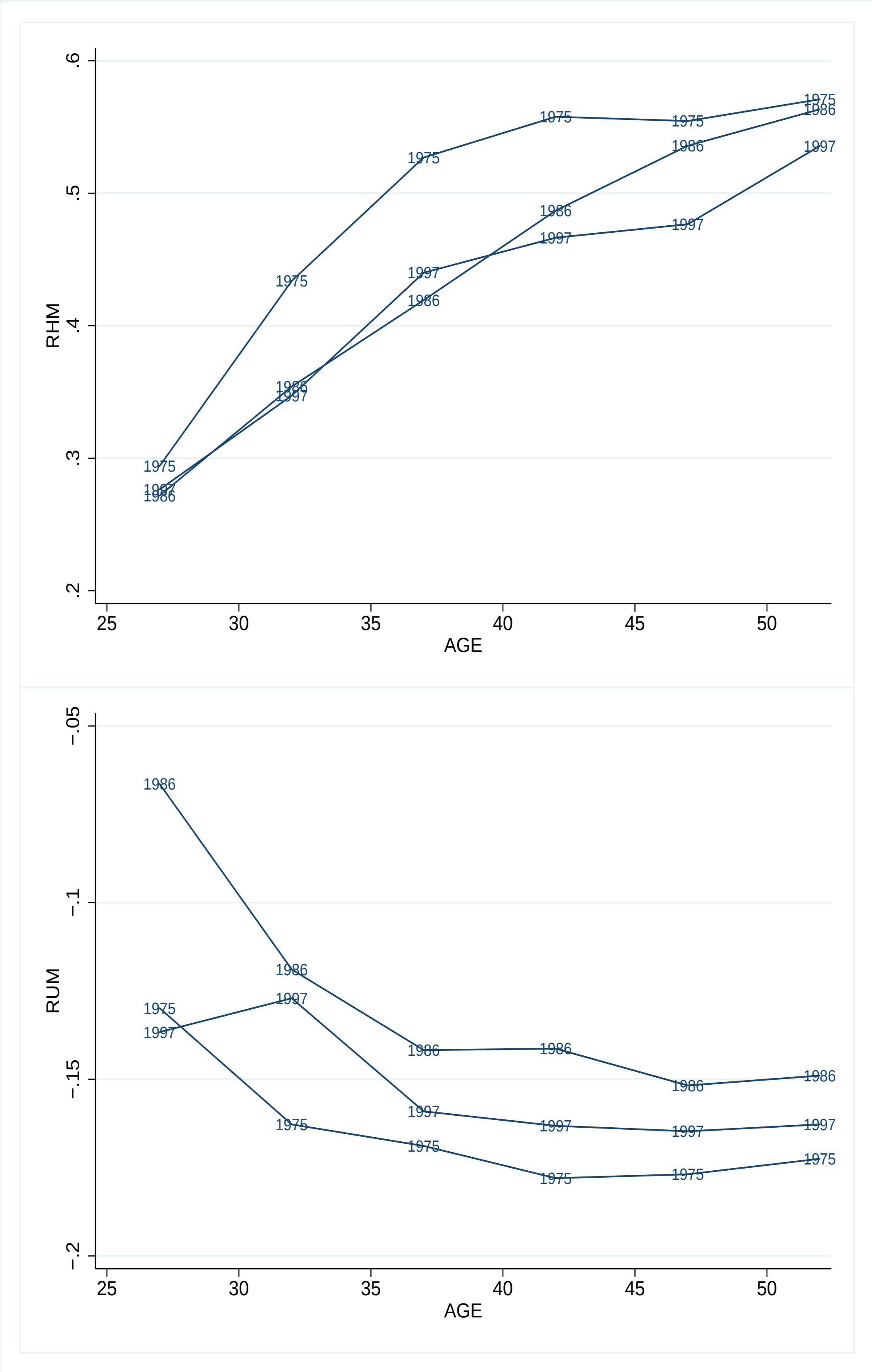
To obtain employment weights for the manufacturing and the non-manufacturing sector, we assign the IABS sector codes to the two categories as done in Fitzenberger (1999). Using the 1997-weights (0.4412 for manufacturing and 0.4746 for non-manufacturing), we calculate the price elasticity of demand, η, as a weighted average of the elasticities $\eta_{\text{man}} = -0.7994$ and $\eta_{\text{non-man}} = -0.1762$ estimated by Fitzenberger and Franz (2001).

B Tables and Figures

Figure 2: Evolution of Wage Differentials over Time

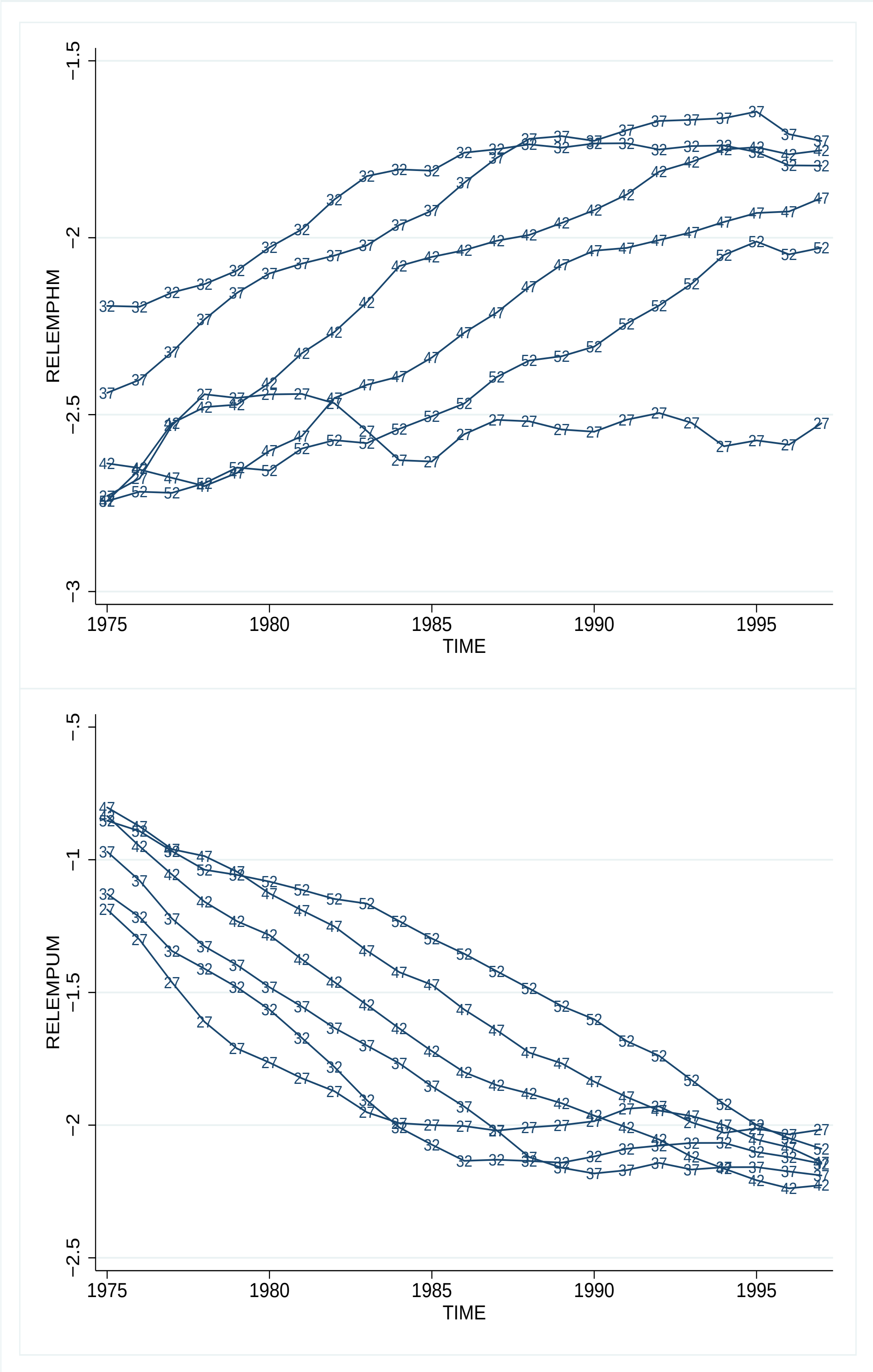
Calculations based on IABS 1975–1997. Digits within the graphs indicate the middle points of the respective age classes.

Figure 3: Age Profiles of Wage Differentials



Calculations based on IABS 1975–1997. Digits within the graphs indicate the calendar years of the respective age-time cells.

Figure 4: Trends in Relative Employment



Calculations based on IABS 1975–1997. Digits within the graphs indicate the middle points of the respective age classes.

Table 4: Estimated Wage Differentials by Age and Time

Age	25–29		30–34		35–39	
Time	$r_{l,a,t}$	$r_{h,a,t}$	$r_{l,a,t}$	$r_{h,a,t}$	$r_{l,a,t}$	$r_{h,a,t}$
1975	-0.1299	0.2942	-0.1628	0.4338	-0.1689	0.5270
	(0.0057)	(0.0111)	(0.0058)	(0.0113)	(0.0048)	(0.0140)
1976	-0.1187	0.2622	-0.1477	0.3882	-0.1564	0.5235
	(0.0057)	(0.0102)	(0.0062)	(0.0102)	(0.0050)	(0.0123)
1977	-0.1175	0.2535	-0.1467	0.3795	-0.1692	0.5013
	(0.0060)	(0.0093)	(0.0065)	(0.0097)	(0.0055)	(0.0113)
1978	-0.1151	0.2517	-0.1439	0.3638	-0.1618	0.4839
	(0.0064)	(0.0089)	(0.0067)	(0.0093)	(0.0059)	(0.0104)
1979	-0.1020	0.2482	-0.1414	0.3407	-0.1544	0.4540
	(0.0067)	(0.0090)	(0.0066)	(0.0085)	(0.0062)	(0.0098)
1980	-0.1117	0.2573	-0.1353	0.3324	-0.1498	0.4486
	(0.0068)	(0.0088)	(0.0066)	(0.0081)	(0.0068)	(0.0102)
1981	-0.1011	0.2669	-0.1387	0.3376	-0.1454	0.4222
	(0.0070)	(0.0088)	(0.0068)	(0.0078)	(0.0072)	(0.0100)
1982	-0.0990	0.2640	-0.1358	0.3473	-0.1553	0.4256
	(0.0071)	(0.0089)	(0.0071)	(0.0075)	(0.0075)	(0.0098)
1983	-0.0947	0.2537	-0.1274	0.3518	-0.1585	0.4236
	(0.0072)	(0.0090)	(0.0078)	(0.0075)	(0.0078)	(0.0096)
1984	-0.0856	0.2585	-0.1289	0.3551	-0.1583	0.4227
	(0.0073)	(0.0094)	(0.0084)	(0.0077)	(0.0081)	(0.0094)
1985	-0.0683	0.2663	-0.1269	0.3628	-0.1479	0.4251
	(0.0074)	(0.0096)	(0.0087)	(0.0078)	(0.0084)	(0.0092)
1986	-0.0664	0.2718	-0.1189	0.3541	-0.1417	0.4193
	(0.0073)	(0.0090)	(0.0089)	(0.0076)	(0.0087)	(0.0089)
1987	-0.0716	0.2917	-0.1090	0.3509	-0.1386	0.4174
	(0.0072)	(0.0087)	(0.0088)	(0.0074)	(0.0090)	(0.0086)
1988	-0.0777	0.2992	-0.1104	0.3413	-0.1317	0.4296
	(0.0072)	(0.0088)	(0.0084)	(0.0069)	(0.0094)	(0.0083)
1989	-0.0780	0.2839	-0.1039	0.3512	-0.1330	0.4230
	(0.0070)	(0.0086)	(0.0084)	(0.0070)	(0.0096)	(0.0084)
1990	-0.0866	0.2735	-0.1187	0.3597	-0.1349	0.4172
	(0.0069)	(0.0086)	(0.0081)	(0.0069)	(0.0094)	(0.0082)
1991	-0.1002	0.2661	-0.1149	0.3574	-0.1379	0.4192
	(0.0068)	(0.0084)	(0.0078)	(0.0067)	(0.0093)	(0.0081)
1992	-0.0924	0.2614	-0.1248	0.3537	-0.1340	0.4126
	(0.0066)	(0.0081)	(0.0075)	(0.0065)	(0.0089)	(0.0077)
1993	-0.0840	0.2672	-0.1212	0.3662	-0.1333	0.4190
	(0.0069)	(0.0083)	(0.0075)	(0.0065)	(0.0089)	(0.0076)
1994	-0.0846	0.2514	-0.1238	0.3651	-0.1438	0.4305
	(0.0073)	(0.0089)	(0.0075)	(0.0063)	(0.0090)	(0.0075)
1995	-0.0987	0.2373	-0.1269	0.3598	-0.1423	0.4383
	(0.0075)	(0.0091)	(0.0077)	(0.0065)	(0.0090)	(0.0075)
1996	-0.1156	0.2486	-0.1242	0.3456	-0.1544	0.4255
	(0.0079)	(0.0097)	(0.0078)	(0.0066)	(0.0091)	(0.0076)
1997	-0.1366	0.2764	-0.1271	0.3473	-0.1591	0.4400
	(0.0089)	(0.0106)	(0.0083)	(0.0069)	(0.0092)	(0.0077)

Tobit estimations, observations weighted with the length of the respective employment spells. Standard errors in parentheses. Data source: IABS 1975–1997.

Table 4: Estimated Wage Differentials by Age and Time (Continued)

Age Time	40–44 $r_{l,a,t}$	 $r_{h,a,t}$	45–49 $r_{l,a,t}$	 $r_{h,a,t}$	50–54 $r_{l,a,t}$	 $r_{h,a,t}$
1975	-0.1780	0.5577	-0.1769	0.5445	-0.1725	0.5709
	(0.0052)	(0.0189)	(0.0055)	(0.0213)	(0.0069)	(0.0248)
1976	-0.1586	0.5713	-0.1699	0.5662	-0.1736	0.5537
	(0.0053)	(0.0147)	(0.0056)	(0.0183)	(0.0065)	(0.0208)
1977	-0.1586	0.5713	-0.1699	0.6003	-0.1838	0.5739
	(0.0053)	(0.0147)	(0.0056)	(0.0185)	(0.0066)	(0.0205)
1978	-0.1640	0.5525	-0.1645	0.6142	-0.1824	0.5853
	(0.0055)	(0.0132)	(0.0059)	(0.0182)	(0.0065)	(0.0191)
1979	-0.1577	0.5185	-0.1659	0.5829	-0.1696	0.5589
	(0.0054)	(0.0118)	(0.0057)	(0.0161)	(0.0063)	(0.0169)
1980	-0.1532	0.5171	-0.1615	0.5507	-0.1649	0.5398
	(0.0054)	(0.0115)	(0.0057)	(0.0148)	(0.0060)	(0.0163)
1981	-0.1549	0.5087	-0.1599	0.5420	-0.1782	0.5675
	(0.0057)	(0.0111)	(0.0057)	(0.0143)	(0.0060)	(0.0164)
1982	-0.1585	0.4902	-0.1542	0.5475	-0.1715	0.5682
	(0.0060)	(0.0105)	(0.0057)	(0.0128)	(0.0062)	(0.0155)
1983	-0.1626	0.4905	-0.1561	0.5468	-0.1698	0.5652
	(0.0066)	(0.0101)	(0.0060)	(0.0122)	(0.0062)	(0.0149)
1984	-0.1613	0.4907	-0.1497	0.5296	-0.1674	0.5484
	(0.0073)	(0.0102)	(0.0062)	(0.0115)	(0.0066)	(0.0145)
1985	-0.1498	0.4888	-0.1504	0.5469	-0.1597	0.5655
	(0.0082)	(0.0109)	(0.0065)	(0.0116)	(0.0069)	(0.0145)
1986	-0.1413	0.4869	-0.1517	0.5359	-0.1490	0.5633
	(0.0090)	(0.0112)	(0.0067)	(0.0111)	(0.0069)	(0.0140)
1987	-0.1425	0.5013	-0.1524	0.5289	-0.1418	0.5836
	(0.0095)	(0.0117)	(0.0073)	(0.0113)	(0.0070)	(0.0138)
1988	-0.1425	0.4679	-0.1499	0.5228	-0.1437	0.5495
	(0.0096)	(0.0110)	(0.0077)	(0.0108)	(0.0071)	(0.0124)
1989	-0.1362	0.4725	-0.1465	0.5416	-0.1388	0.5427
	(0.0098)	(0.0111)	(0.0083)	(0.0115)	(0.0073)	(0.0125)
1990	-0.1438	0.4527	-0.1527	0.5045	-0.1424	0.5359
	(0.0096)	(0.0104)	(0.0092)	(0.0118)	(0.0073)	(0.0122)
1991	-0.1391	0.4578	-0.1490	0.5002	-0.1513	0.5446
	(0.0095)	(0.0102)	(0.0101)	(0.0126)	(0.0077)	(0.0124)
1992	-0.1394	0.4468	-0.1417	0.4860	-0.1562	0.5183
	(0.0095)	(0.0095)	(0.0102)	(0.0121)	(0.0080)	(0.0120)
1993	-0.1450	0.4803	-0.1472	0.5014	-0.1677	0.5274
	(0.0098)	(0.0095)	(0.0104)	(0.0121)	(0.0088)	(0.0121)
1994	-0.1552	0.4614	-0.1459	0.5024	-0.1697	0.5378
	(0.0102)	(0.0092)	(0.0106)	(0.0106)	(0.0099)	(0.0123)
1995	-0.1527	0.4641	-0.1583	0.5034	-0.1682	0.5226
	(0.0102)	(0.0091)	(0.0106)	(0.0114)	(0.0112)	(0.0130)
1996	-0.1568	0.4561	-0.1725	0.4675	-0.1575	0.4877
	(0.0105)	(0.0091)	(0.0105)	(0.0107)	(0.0120)	(0.0132)
1997	-0.1632	0.4664	-0.1647	0.4766	-0.1628	0.5356
	(0.0105)	(0.0092)	(0.0110)	(0.0106)	(0.0129)	(0.0141)

Tobit estimations, observations weighted with the length of the respective employment spells. Standard errors in parentheses. Data source: IABS 1975–1997.

Table 5: Cohort Effects in Wage Differentials?

Coefficients	Wage Differential l/m				Wage Differential h/m			
DJ32	-0.03277	(-10.72)	-0.03278	(-9.51)	0.10880	(12.00)	0.10887	(15.74)
DJ37	-0.05143	(-9.00)	-0.05145	(-8.17)	0.19881	(13.88)	0.19895	(15.78)
DJ42	-0.05798	(-5.37)	-0.05801	(-4.93)	0.25360	(9.79)	0.25384	(10.69)
DJ47	-0.06539	(-3.37)	-0.0655	(-3.10)	0.29732	(6.75)	0.29769	(7.29)
DJ52	-0.06898	(-2.19)	-0.06907	(-2.03)	0.32937	(4.87)	0.32994	(5.46)
DT76	0.01123	(3.58)			-0.00966	(-0.80)		
DT77	0.00835	(2.45)			-0.01215	(-1.09)		
DT78	0.01117	(3.50)			-0.01760	(-1.32)		
DT79	0.01841	(5.94)			-0.0428	(-3.41)		
DT80	0.02124	(5.34)			-0.05278	(-3.93)		
DT81	0.02133	(5.24)			-0.05340	(-3.73)		
DT82	0.02226	(5.37)			-0.05424	(-3.71)		
DT83	0.02336	(4.55)			-0.05669	(-3.64)		
DT84	0.02661	(4.39)			-0.06185	(-3.54)		
DT85	0.03486	(5.06)			-0.05445	(-2.89)		
DT86	0.04075	(5.36)			-0.05967	(-2.89)		
DT87	0.04317	(4.89)			-0.05404	(-2.26)		
DT88	0.04341	(4.27)			-0.0663	(-2.53)		
DT89	0.04695	(4.04)			-0.06747	(-2.30)		
DT90	0.04024	(3.01)			-0.08155	(-2.49)		
DT91	0.03848	(2.66)			-0.08363	(-2.30)		
DT92	0.03976	(2.22)			-0.09726	(-2.38)		
DT93	0.03882	(1.96)			-0.08609	(-1.89)		
DT94	0.03565	(1.62)			-0.09096	(-1.81)		
DT95	0.03279	(1.33)			-0.09749	(-1.74)		
DT96	0.02853	(1.02)			-0.11586	(-1.86)		
DT97	0.02480	(0.79)			-0.09967	(-1.40)		
TIME			0.00174	(1.08)			-0.01827	(-4.84)
TIME2			0.00032	(1.09)			0.00217	(3.16)
TIME3			0.00002	(-1.06)			-0.00012	(-2.67)
TIME4			2.16e-07	(0.47)			2.38e-06	(2.26)
R1	-0.00054	(-1.05)	-0.00054	(-0.96)	0.00059	(0.35)	0.00059	(0.32)
R2	-6.09e-06	(-0.23)	-6.14e-06	(-0.22)	-0.00009	(-0.80)	-0.00009	(-0.78)
R3	0.00017	(1.87)	0.00017	(1.68)	-3.65e-06	(-0.01)	-3.26e-06	(-0.01)
R4	2.89e-06	(0.44)	2.90e-06	(0.42)	0.00003	(1.30)	0.00003	(1.19)
COHORTA2	0.00018	(1.86)	0.00018	(1.78)	0.00071	(3.22)	0.00071	(4.60)
COHORTB2					0.00093	(3.52)	0.0009	(2.89)
COHORTA3	-0.00001	(-3.67)	-0.00001	(-3.84)	-0.00003	(-3.32)	-0.00003	(-3.62)
COHORTB3					0.00003	(1.75)	0.00003	(1.74)
CONSTANT	-0.12597	(-41.59)	-0.12178	(-35.41)	0.28023	(23.91)	0.28895	(43.45)
Tests[a]								
Separability[b]	7.83*		6.27		8.11*		9.13**	
Cohorts after 1975[c]	36.57***		34.79***		34.14***		36.00***	
Any cohort effects[d]					233.12***		232.72***	

Data source: IABS 1975–1997. White robust t-values in parentheses. Specification of equation (4): Inclusion of additional polynomial cohort terms as long as neither the respective coefficient nor those of lower orders turn insignificant.

[a] Wald tests, χ^2-values. *(**,***) Hypothesis rejected at 0.90 (0.95, 0.99) level.

[b] $H_0 : \mathrm{R}_i = 0$ for all i.

[c] $H_0 : \mathrm{R}_i = \mathrm{COHORTA}_j = 0$ for all i, j.

[d] $H_0 : \mathrm{R}_i = \mathrm{COHORTA}_j = \mathrm{COHORTB}_h = 0$ for all h, i, j.

Table 6: Cohort Effects in Relative Employment?

Coefficients	Log. Relative Employment l/m				Log. Relative Employment h/m			
DJ32	0.20288	(6.78)	0.20273	(7.04)	0.40904	(10.08)	0.40907	(10.61)
DJ37	0.43241	(9.80)	0.43213	(9.96)	0.19945	(3.48)	0.19952	(3.67)
DJ42	0.74647	(9.19)	0.74594	(9.46)	-0.19470	(-2.02)	-0.19458	(-2.13)
DJ47	1.0750	(7.67)	1.07414	(7.90)	-0.6767	(-4.19)	-0.67656	(-4.39)
DJ52	1.271	(5.76)	1.26952	(5.89)	-1.164	(-4.61)	-1.16373	(-4.84)
DT76	-0.12099	(-5.19)			0.02528	(0.69)		
DT77	-0.26984	(-12.28)			0.08666	(2.69)		
DT78	-0.38845	(-15.32)			0.13468	(3.18)		
DT79	-0.48621	(-17.13)			0.17800	(3.95)		
DT80	-0.58195	(-19.86)			0.23592	(5.19)		
DT81	-0.6860	(-22.36)			0.30340	(6.54)		
DT82	-0.79031	(-24.24)			0.3775	(8.62)		
DT83	-0.90355	(-23.41)			0.43387	(9.11)		
DT84	-1.0147	(-22.55)			0.50096	(9.02)		
DT85	-1.1114	(-21.79)			0.57136	(10.11)		
DT86	-1.2122	(-20.47)			0.67825	(11.37)		
DT87	-1.3021	(-20.57)			0.78194	(11.87)		
DT88	-1.3900	(-19.29)			0.87782	(11.65)		
DT89	-1.4639	(-17.60)			0.95953	(11.15)		
DT90	-1.5333	(-16.10)			1.0478	(10.62)		
DT91	-1.5957	(-15.04)			1.1553	(10.39)		
DT92	-1.6613	(-13.83)			1.2654	(10.11)		
DT93	-1.7531	(-12.75)			1.3681	(9.60)		
DT94	-1.8375	(-11.92)			1.473	(9.18)		
DT95	-1.9217	(-11.27)			1.5817	(8.76)		
DT96	-2.0027	(-10.59)			1.6518	(8.23)		
DT97	-2.0752	(-9.73)			1.771	(7.92)		
TIME			-0.11679	(-12.16)			0.04054	(2.86)
TIME2			-0.00076	(-0.42)			0.00105	(0.49)
TIME3			0.00020	(1.75)			0.00013	(0.99)
TIME4			-5.38e-06	(-2.04)			-4.31e-06	(-1.55)
R1	0.01396	(3.96)	0.01394	(4.10)	-0.02248	(-4.80)	-0.02248	(-5.10)
R2	0.00011	(0.62)	0.00011	(0.64)	0.00007	(0.32)	0.00007	(0.35)
R3	-0.00443	(-6.99)	-0.00443	(-7.34)	0.00455	(5.28)	0.00455	(5.63)
R4	-0.00009	(-2.05)	-0.00009	(-2.12)	0.00003	(0.66)	0.00003	(0.71)
COHORTA2	0.00484	(7.97)	0.00484	(8.05)	-0.00600	(-8.44)	-0.00600	(-8.66)
COHORTB2					-0.00326	(-4.16)	-0.00326	(-4.48)
COHORTA3	-0.00010	(-7.00)	-0.00010	(-6.91)	0.00011	(7.32)	0.00011	(7.20)
CONSTANT	-1.25039	(-40.19)	-1.26223	(-42.44)	-2.62419	(-47.43)	-2.62877	(-55.56)
Tests[a]								
Separability[b]	377.59***		355.74***		35.76***		43.66***	
Cohorts after 1975[c]	1000.3***		1035.6***		1004.0***		1034.1***	
Any cohort effects[d]					1011.1***		1043.8***	

Data source: IABS 1975–1997. White robust t-values in parentheses. Specification of equation (4): Inclusion of additional polynomial cohort terms as long as neither the respective coefficient nor those of lower orders turn insignificant.

[a] Wald tests, χ^2-values. *(**,***) Hypothesis rejected at 0.90 (0.95, 0.99) level.

[b] $H_0 : \mathrm{R}_i = 0$ for all i.

[c] $H_0 : \mathrm{R}_i = \mathrm{COHORTA}_j = 0$ for all i, j.

[d] $H_0 : \mathrm{R}_i = \mathrm{COHORTA}_j = \mathrm{COHORTB}_h = 0$ for all h, i, j.

Table 7: Summary of Model Versions (Specifications)

Label	Specification	Description
(i)	benchmark model	two-level CES
(ii)	extended benchmark	two-level CES with age×time interaction in age-specific relative productivity
(iii)	extended benchmark	two-level CES with cohort×time interaction in age-specific relative productivity
(iv)	sensitivity check	two-level CES disregarding age-premia
(v)	sensitivity check	two-level CES excluding university graduates aged 25–29 years
(vi)	sensitivity check	two-level CES with break in SBTC
(vii)	restricted benchmark	two-level CES with $1/\sigma_A = 0$
(viii)	CES with interaction	CES with age×skill interaction
(ix)	traditional CES	traditional CES

Table 8: Elasticities of Substitution, Specifications of the Nested CES

model version		(i)	(i)	(i)	(ii)	(ii)	(ii)	(iii)	(iii)	(iii)
	l		**7.10**	**7.10**		**7.24**	**7.24**		**7.24**	**7.24**
			(0.49)	(0.49)		(0.50)	(0.50)		(0.50)	(0.50)
$\sigma_{A,\text{1st step}}^{\text{FGLS}}$	m	**8.28**	**7.07**	**7.07**	**7.53**	**7.20**	**7.20**	**7.53**	**7.20**	**7.20**
		(0.55)	(0.48)	(0.48)	(0.54)	(0.50)	(0.50)	(0.54)	(0.50)	(0.50)
	h		**18.55**	**18.55**		**8.98**	**8.98**		**8.98**	**8.98**
			(5.59)	(5.59)		(1.51)	(1.51)		(1.51)	(1.51)
	l		**8.58**	**8.59**		**9.18**	**9.21**		**9.20**	**9.26**
			(0.67)	(0.69)		(0.74)	(0.75)		(0.74)	(0.76)
$\sigma_{A,\text{3rd step}}^{\text{FGLS}}$	m	**8.71**	**4.81**	**4.81**	**7.85**	**5.23**	**5.25**	**7.92**	**5.22**	**5.26**
		(0.63)	(0.32)	(0.32)	(0.60)	(0.38)	(0.39)	(0.61)	(0.37)	(0.38)
	h		**19.52**	**19.69**		**10.36**	**10.47**		**10.15**	**10.49**
			(5.87)	(5.98)		(1.98)	(2.02)		(1.83)	(1.99)
	l			**12.56**			**12.72**			**8.18**
				(1.58)			(1.75)			(0.88)
σ_{S}^{FGLS}	m	**8.97**	**9.36**	**7.15**	**9.04**	**9.49**	**6.98**	**5.65**	**6.32**	**4.74**
		(0.81)	(0.91)	(1.06)	(0.84)	(0.93)	(1.01)	(0.46)	(0.55)	(0.52)
	h			**6.81**			**6.78**			**5.36**
				(1.00)			(0.97)			(0.74)
	l		**6.87**	**6.87**		**7.23**	**7.23**		**7.23**	**7.23**
			(0.53)	(0.53)		(0.54)	(0.54)		(0.54)	(0.54)
$\sigma_{A,\text{1st step}}^{\text{FGLS-IV}}$	m	**8.11**	**6.86**	**6.86**	**7.44**	**7.20**	**7.20**	**7.44**	**7.20**	**7.20**
		(0.60)	(0.53)	(0.53)	(0.60)	(0.54)	(0.54)	(0.59)	(0.54)	(0.54)
	h		**28.95**	**28.95**		**9.85**	**9.85**		**9.86**	**9.86**
			(14.57)	(14.57)		(2.25)	(2.25)		(2.25)	(2.25)
	l		**10.31**	**11.53**		**9.44**	**10.20**		**8.68**	**9.22**
			(1.64)	(2.12)		(1.62)	(1.98)		(1.40)	(1.62)
$\sigma_{A,\text{3rd step}}^{\text{FGLS-IV}}$	m	**10.25**	**5.27**	**5.67**	**9.22**	**6.01**	**6.55**	**8.23**	**5.38**	**5.79**
		(1.52)	(0.66)	(0.82)	(1.34)	(0.86)	(1.14)	(1.19)	(0.73)	(0.91)
	h		**20.13**	21.26		**8.50**	**9.10**		**8.59**	**9.85**
			(11.11)	(12.10)		(2.57)	(3.21)		(2.63)	(3.65)
	l			9.67$^\sharp$			11.54$^\sharp$			8.03$^\sharp$
				(13.37)			(24.50)			(8.44)
$\sigma_{S}^{\text{FGLS-IV}}$	m	**8.14**	**6.15**	5.65$^\sharp$	7.94	**6.97**	6.69$^\sharp$	**5.92**	**4.91**	**4.48**$^\sharp$
		(3.11)	(2.87)	(5.72)	(4.07)	(2.85)	(7.16)	(1.65)	(1.54)	(2.16)
	h			6.8$^\sharp$			7.19$^\sharp$			6.38$^\sharp$
				(7.76)			(8.13)			(4.56)

Model versions: (i) benchmark model; (ii) with age×time interaction in age-specific productivity; (iii) with cohort×time interaction in age-specific productivity. Standard errors in parentheses based on 500 bootstrap repetitions. Bold numbers: Elasticities finite (inverses significant at 0.95 level). $^\sharp$ Respective parameters not significantly different at 0.95 level. Data sources: IABS 1975–1997, German Microcensus.

Table 9: Elasticities of Substitution, Further Specifications of the Nested CES

model version		(iv)	(iv)	(iv)	(v)	(v)	(v)	(vi)	(vi)	(vi)
	l		**15.19**	**15.19**		**6.85**	**6.85**		**7.10**	**7.10**
			(1.27)	(1.27)		(0.47)	(0.47)		(0.49)	(0.49)
$\sigma_{A,\text{1st step}}^{\text{FGLS}}$	m	**16.64**	**15.13**	**15.13**	**7.23**	**6.83**	**6.83**	**8.28**	**7.07**	**7.07**
		(1.37)	(1.26)	(1.26)	(0.50)	(0.46)	(0.46)	(0.55)	(0.48)	(0.48)
	h		**26.76**	**26.76**		**9.86**	**9.86**		**18.55**	**18.55**
			(6.37)	(6.37)		(1.74)	(1.74)		(5.59)	(5.59)
	l		**17.09**	**17.20**		**8.54**	**8.53**		**8.62**†	**8.64**
			(1.77)	(1.82)		(0.66)	(0.66)		(0.69)	(0.69)
$\sigma_{A,\text{3rd step}}^{\text{FGLS}}$	m	**19.01**	**9.21**	**9.26**	**7.65**	**4.79**	**4.79**	**8.86**†	**4.84**†	**4.84**
		(2.04)	(0.76)	(0.79)	(0.57)	(0.32)	(0.33)	(0.66)	(0.32)	(0.33)
	h		**24.53**	**25.54**		**11.19**	**11.23**		**19.19**†	**18.93**
			(6.41)	(7.00)		(2.39)	(2.33)		(5.66)	(5.46)
	l			**12.31**			**12.54**			24.55
				(1.69)			(1.70)			(14.19)
σ_{S}^{FGLS}	m	**8.94**	**8.92**	**6.83**	**8.59**	**9.24**	**6.84**	**14.42**†	**16.69**†	**13.02**
		(0.83)	(0.85)	(1.09)	(0.77)	(0.88)	(0.96)	(4.98)	(6.77)	(5.64)
	h			**5.95**			**6.46**			**7.13**
				(0.79)			(1.08)			(2.03)
	l		**14.83**	**14.83**		**6.71**	**6.71**		**6.87**	**6.87**
			(1.33)	(1.33)		(0.56)	(0.56)		(0.56)	(0.56)
$\sigma_{A,\text{1st step}}^{\text{FGLS-IV}}$	m	**15.57**	**14.78**	**14.78**	**7.02**	**6.70**	**6.70**	**8.11**	**6.86**	**6.86**
		(1.36)	(1.32)	(1.32)	(0.58)	(0.56)	(0.56)	(0.65)	(0.56)	(0.56)
	h		**21.71**	**21.71**		**9.13**	**9.13**		28.95	28.95
			(4.99)	(4.99)		(2.00)	(2.00)		(14.94)	(14.94)
	l		**15.10**	**15.03**		**8.70**	**9.83**		**10.48**†	**10.23**†
			(1.33)	(1.27)		(1.15)	(1.57)		(1.70)	(1.65)
$\sigma_{A,\text{3rd step}}^{\text{FGLS-IV}}$	m	**14.73**	**8.27**	**8.26**	**8.76**	**4.67**	**5.16**	**10.94**†	**5.31**†	**5.16**†
		(1.22)	(0.65)	(0.63)	(1.01)	(0.48)	(0.68)	(1.73)	(0.65)	(0.61)
	h		**12.15**	**11.95**		**8.40**	**11.02**		26.43†	**20.52**†
			(1.64)	(1.60)		(2.56)	(4.80)		(16.49)	(10.03)
	l			12.48$^\sharp$			13.87$^\sharp$			-26.87$^{\sharp\dagger}$
				(7.71)			(25.22)			(72.13)
$\sigma_{S}^{\text{FGLS-IV}}$	m	**8.09**	**8.72**	**6.00**$^\sharp$	**8.01**	**8.75**	6.10$^\sharp$	15.87†	-172.7†	-20.97$^{\sharp\dagger}$
		(2.98)	(3.30)	(1.64)	(1.90)	(3.43)	(3.51)	(12.29)	(2749.2)	(72.13)
	h			**9.10**$^\sharp$			6.11$^\sharp$			6.70$^{\sharp\dagger}$
				(3.32)			(7.11)			(16.96)

Model versions: (iv) excluding equations for age premia; (v) excluding high-skilled of age 25–29; (vi) with break in SBTC. Standard errors in parentheses based on 500 bootstrap repetitions. Bold numbers: Elasticities finite (inverses significant at 0.95 level). $^\sharp$ Respective parameters not significantly different at 0.95 level. † Time break in SBTC insignificant at 0.95 level. Data sources: IABS 1975–1997, German Microcensus.

Table 10: Estimates of σ_S, Assuming Perfect Substitution Between Age Classes

model version		(vii)	(vii)	(viii)	(viii)	(ix)	(ix)
	l		**12.45**		**11.95**		**10.57**
			(1.70)		(2.00)		(2.28)
σ_S^{FGLS}	m	**8.82**	**6.79**	**8.25**	**5.86**	**4.93**	**3.81**
		(0.78)	(0.97)	(0.97)	(0.94)	(0.67)	(0.45)
	h		**6.15**		**4.67**		**3.36**
			(0.80)		(0.54)		(0.24)
	l		$15.95^\sharp$		$13.80^\sharp$		198.3
			(47.95)		(8.26)		(3132.6)
$\sigma_S^{\text{FGLS-IV}}$	m	9.14	$6.93^\sharp$	**6.26**	**6.21**$^\sharp$	**3.76**	**5.15**
		(8.03)	(7.26)	(1.16)	(1.79)	(0.95)	(1.99)
	h		$8.32^\sharp$		**5.65**$^\sharp$		**4.53**
			(14.30)		(1.18)		(1.08)

Model versions: (vii) nested CES (benchmark) with $1/\sigma_A$ restricted to zero at the third step; (viii) CES model with skill differentials as employment-weighted average of age-specific premia; (ix) CES model without skill×age-interaction. Standard errors in parentheses. Bold numbers: Elasticities finite (inverses significant at 0.95 level). $^\sharp$ Respective parameters not significantly different at 0.95 level. Data sources: IABS 1975–1997, German Microcensus.

Table 11: Instrumental Variables: First Stage Results for Age-Specific Employment

Model	(i)/(iv)		(ii)		(iii)	
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
α_{uu}	0.7411***	(0.0716)	0.8091***	(0.0762)	0.8091***	(0.0762)
α_{mu}	0.5254***	(0.1204)	0.5237***	(0.1296)	0.5237***	(0.1296)
α_{hu}	-0.3894***	(0.0677)	-0.5690***	(0.0832)	-0.5690***	(0.0832)
α_{um}	0.0620	(0.0458)	0.1158**	(0.0490)	0.1158**	(0.0490)
α_{mm}	0.9232***	(0.0766)	0.9106***	(0.0814)	0.9106***	(0.0815)
α_{hm}	-0.1008**	(0.0450)	-0.2272***	(0.0572)	-0.2272***	(0.0572)
α_{uh}	-0.1603*	(0.0899)	-0.1044	(0.0947)	-0.1044	(0.0947)
α_{mh}	0.6185***	(0.1504)	0.6178***	(0.1592)	0.6178***	(0.1592)
α_{hh}	0.6484***	(0.0848)	0.4987***	(0.1081)	0.4987***	(0.1081)
χ^2	12722.7***		5150.5***		5150.5***	

Coefficients of additional instruments. See the text for a description of the instrumentation strategy. See the text and tables 8 and 9 for descriptions of the model versions (i) to (iv). χ^2: Test for joint significance of additional instruments. *(**,***) parameter significant at 0.90 (0.95, 0.99) level.
Data sources: IABS 1975–1997, German Microcensus.

Table 12: Instrumental Variables: First Stage Results for Aggregate Employment

Model	(i)		(ii)		(iii)		(iv)	
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
α_{uu}	0.3466	(0.4031)	0.3409	(0.4028)	0.3187	(0.3906)	-0.9379***	(0.2001)
α_{mu}	0.3597	(0.9445)	0.3595	(0.9436)	0.4125	(0.9151)	1.8051***	(0.4019)
α_{hu}	1.4425	(1.3686)	1.4435	(1.3673)	1.3901	(1.3260)	0.3551	(0.5588)
α_{um}	-0.2373**	(0.1074)	-0.2435**	(0.1068)	-0.3352**	(0.1424)	-0.0797	(0.1203)
α_{mm}	1.4841***	(0.2517)	1.4763***	(0.2501)	1.6945***	(0.3336)	1.3068***	(0.2417)
α_{hm}	-0.0854	(0.3647)	-0.0778	(0.3624)	-0.2980	(0.4834)	0.0480	(0.3360)
α_{uh}	-0.5004**	(0.2492)	-0.4982**	(0.2465)	-0.5909**	(0.2954)	0.1934	(0.1846)
α_{mh}	2.5600***	(0.5838)	2.5478***	(0.5774)	2.768***	(0.6921)	1.7792***	(0.3707)
α_{hh}	-1.5078*	(0.8459)	-1.4936*	(0.8367)	-1.7162*	(1.0028)	-0.9205*	(0.5154)
χ^2	235.04***		240.74***		233.86***		250.51***	

Coefficients of additional instruments. See the text for a description of the instrumentation strategy. See the text and tables 8 and 9 for descriptions of the model versions (i) to (iv). χ^2: Test for joint significance of additional instruments. *(**,***) parameter significant at 0.90 (0.95, 0.99) level.
Data sources: IABS 1975–1997, German Microcensus.

C Monte Carlo Study

We conduct a Monte Carlo Study in order to compare the following estimation approaches:

(a) Age-specific relative productivities $\phi_{s,a}/\phi_{\tilde{s},\tilde{a}}$ and elasticities σ_{As} estimated freely at the third step.

(b) $\phi_{s,a}$ at the third step taken as predetermined from the second step.

(c) σ_{As} at the third step taken as predetermined from the first step.

(d) $\phi_{s,a}$ as well as σ_{As} at the third step taken as predetermined from previous steps.

We assume the following parameter values for the benchmark model:

- elasticities $\sigma_{A_u} = 15$, $\sigma_{A_m} = 10$, $\sigma_{A_h} = 20$, and $\sigma_S = 2$.
- skill-specific linear time trends of 1% per year for $\ln(\theta_{h,t}/\theta_{m,t})$ and of 2% for $\ln(\theta_{m,t}/\theta_{l,t})$.
- age×skill-specific productivities $\phi_{s,a}$ set to appropriate values between $\exp(9.3)$ and $\exp(9.8)$.

When simulating log wages, we assume a normally distributed additive error term with standard deviation $STDDEV$, which captures residual wage dispersion. The chosen values for $STDDEV$ between 0.001 and 0.2 correspond to 90–10-percentile wage differences between 0.2% and 50%.

We then estimate the benchmark version (i) of the model. Results for the different approaches are displayed in table 13. None of the approaches strictly dominates the others in terms of minimum bias or minimum root mean squared error for all parameters. However, approach (a), which (re)estimates all parameters freely at the third step, performs best in most of the cases and, what is more, its performance is also fairly good when coming off second-best. We therefore decide to use approach (a) for the estimations throughout the paper.

Table 13: Monte Carlo Study: Average Third-Step Estimates for $1/\sigma$ and Root Mean Squared Errors

STDDEV	Parameter	Value	(a)	(b)	(c)	(d)
0.001	$1/\sigma_{Au}$	0.0667	**0.0666**	0.0763	0.0739	0.0739
			(**0.00046**)	(0.00966)	(0.00720)	(0.00720)
	$1/\sigma_{Am}$	0.1	**0.1000**	0.0976	0.0742	0.0742
			(**0.00085**)	(0.00244)	(0.02581)	(0.02581)
	$1/\sigma_{Ah}$	0.05	**0.0500**	0.0477	0.0488	0.0488
			(**0.00082**)	(0.00239)	(0.00144)	(0.00144)
	$1/\sigma_S$	0.5	**0.5002**	0.4730	0.4993	0.3208
			(**0.00139**)	(0.02704)	(0.00145)	(0.17936)
0.005	$1/\sigma_{Au}$	0.0667	**0.0666**	0.0767	0.0746	0.0746
			(**0.00230**)	(0.01023)	(0.00810)	(0.00810)
	$1/\sigma_{Am}$	0.1	**0.0999**	0.0981	0.0748	0.0748
			(0.00423)	(**0.00262**)	(0.02529)	(0.02529)
	$1/\sigma_{Ah}$	0.05	**0.0499**	0.0479	0.0496	0.0496
			(0.00409)	(**0.00357**)	(0.00385)	(0.00385)
	$1/\sigma_S$	0.5	**0.5004**	0.4758	0.4992	0.3180
			(0.00691)	(0.02530)	(**0.00659**)	(0.18496)
0.010	$1/\sigma_{Au}$	0.0667	**0.0665**	0.0771	0.0747	0.0747
			(**0.00459**)	(0.01116)	(0.00879)	(0.00879)
	$1/\sigma_{Am}$	0.1	**0.0998**	0.0986	0.0748	0.0748
			(0.00847)	(**0.00401**)	(0.02542)	(0.02542)
	$1/\sigma_{Ah}$	0.05	**0.0498**	0.0482	**0.0498**	**0.0498**
			(0.00818)	(**0.00607**)	(0.00767)	(0.00767)
	$1/\sigma_S$	0.5	**0.5006**	0.4765	0.4992	0.3130
			(0.01381)	(0.02757)	(**0.01347**)	(0.19854)
0.050	$1/\sigma_{Au}$	0.0667	**0.0658**	0.0790	0.0741	0.0741
			(0.02296)	(0.02341)	(**0.01983**)	(**0.01983**)
	$1/\sigma_{Am}$	0.1	**0.0989**	0.1020	0.0743	0.0743
			(0.04234)	(**0.01985**)	(0.03162)	(0.03162)
	$1/\sigma_{Ah}$	0.05	0.0492	0.0463	**0.0495**	**0.0495**
			(0.04088)	(**0.03095**)	(0.03842)	(0.03842)
	$1/\sigma_S$	0.5	0.5026	0.4989	**0.5009**	0.2976
			(0.06903)	(0.07197)	(**0.06884**)	(0.34478)
0.100	$1/\sigma_{Au}$	0.0667	**0.0649**	0.0795	0.0733	0.0733
			(0.04592)	(0.03973)	(**0.03737**)	(**0.03737**)
	$1/\sigma_{Am}$	0.1	**0.0978**	0.1064	0.0735	0.0735
			(0.08469)	(**0.03821**)	(0.04534)	(0.04534)
	$1/\sigma_{Ah}$	0.05	0.0484	0.0432	**0.0491**	**0.0491**
			(0.08177)	(**0.06613**)	(0.07685)	(0.07685)
	$1/\sigma_S$	0.5	0.5049	0.5578	**0.5031**	0.4805
			(0.13808)	(0.16231)	(**0.13778**)	(0.44487)
0.200	$1/\sigma_{Au}$	0.0667	**0.0632**	0.0746	0.0717	0.0717
			(0.09185)	(0.07540)	(**0.07372**)	(**0.07372**)
	$1/\sigma_{Am}$	0.1	**0.0955**	0.1087	0.0718	0.0718
			(0.16939)	(**0.07247**)	(0.07874)	(0.07874)
	$1/\sigma_{Ah}$	0.05	0.0469	0.0405	**0.0481**	**0.0481**
			(0.16356)	(**0.14403**)	(0.15371)	(0.15371)
	$1/\sigma_S$	0.5	0.5094	0.6651	**0.5073**	0.7133
			(0.27625)	(0.30859)	(**0.27571**)	(0.51902)

Simulation of the benchmark model based on 1000 resamples. RMSE in parentheses. Bold numbers: minimum bias and minimum RMSE, respectively. See the text of appendix C for a description of the estimation approaches (a) to (d).

D Calculating Standard Errors

The calculation of standard errors and test statistics for the parameters obtained from the multi-step estimation approach and in the simulation experiment has to take account of pre-step estimation variability. We therefore use bootstrapping techniques.

We resample from the distribution of $\ln(w_{s,a,t})$ estimated by the Tobit regressions at stage zero (section 2.1, equation (1)). All three subsequent estimation steps as well as the calculations for the simulation experiments are put into a single bootstrap loop. We use 500 repetitions to obtain the variance-covariance matrix of the estimated parameters from the empirical bootstrap distribution.

Standard errors for the reported elasticities σ can then be calculated by means of the Delta method relying on the estimated bootstrap distribution of $1/\sigma$. Direct bootstrapping of σ is not possible because of the discontinuity of the inverse function at the argument zero. With negative estimates of the inverse elasticities for single extreme resamples, direct calculation of the variance of σ would not be not well-defined.

In case of IV estimation the three-step approach is extended at steps one and three by the estimation of IV equations. In each iteration of the bootstrap loop we draw from the estimated distribution of IV parameters, calculate predicted values for employment $L_{s,a,t}$ and $L_{s,t}$, and estimate the three-step model. Note that while predicted employment values are used in the SOLS and FGLS estimation, the calculation of the respective FGLS weighting matrices from SOLS residuals relies on actual employment.

When drawing inference on the estimated wage changes in the simulation experiment, we assume η_{man} and $\eta_{\text{non-man}}$, the price elasticities of product demand taken from Fitzenberger and Franz (2001), to be independently normally distributed.