New evidence on the complementarity of education and training

Regina T. Riphahn
Univ. Erlangen-Nuremberg

Parvati Trübswetter
Univ. Erlangen-Nuremberg

(October 2007)

LASER Discussion Papers - Paper No. 4

(edited by A. Abele-Brehm, R.T. Riphahn, K. Moser and C. Schnabel)

Correspondence to:

Regina T. Riphahn, Lange Gasse 20, 90403 Nuremberg, Germany, Email: regina.riphahn@wiso.uni-erlangen.de.

Abstract

We compare the incidence of training and the changing correlation of training with age for high and low skilled German workers between 1996 and 2004. Not only do highly skilled workers receive more training than low skilled workers at any point in time, also the increase in the provision of training disproportionately benefited those with high skills. Thus education and training appear to be complements and the gap in labor market performance between skill groups can be expected to widen over time. The share of training provided to older workers, particularly high skill older workers, increased substantially.

Zusammenfassung

Wir betrachten die Häufigkeit beruflicher Weiterbildungsmaßnahmen und ihre Korrelation mit dem Alter für hoch und gering qualifizierte Arbeitnehmer. Unsere Datenbasis stellen die Mikrozensuserhebungen von 1996 bis 2004 dar. Die Weiterbildungwahrscheinlichkeit ist höher für hoch qualifizierte Arbeitnehmer und sie steigt stärker als die gering qualifizierter Arbeitnehmer. Qualifikation und Weiterbildung sind komplementär und der Abstand im Arbeitsmarkterfolg zwischen unterschiedlichen Qualifikationsgruppen wächst weiter. Über den betrachteten Zeitraum hinweg ist der Anteil der Weiterbildungsmaßnahmen zugunsten älterer, insbesondere hoch qualifizierter Arbeitnehmer deutlich gestiegen.

Author note

We acknowledge the financial support of the RatSWD / BMBF Germany as well as the Frisch Stiftung at the University of Erlangen-Nuremberg. We thank Jennifer Hunt and seminar participants at MPI-Rostock, the Paris School of Economics, and at the cesifo ESP area conference for helpful comments.

1. Introduction

The question of whether formal education and workplace training are complements or substitutes has recently sparked new debates and research (Ariga and Brunello 2006, van Smoorenburg and van der Velden 2000, see also Krueger and Rouse 1998). The OECD emphasized the relevance of this issue since "…differences in the association of prior human capital investments and training could also have important implications for equity…" (OECD 1999, p.149). If only the highly educated received additional training, labor market differences by formal education would be reinforced. If, however, workplace training were a substitute for formal education existing differences in labor market performance and productivity could be bridged.

In this study we investigate the relationship between formal education and workplace training over time. Due at least in part to institutional changes in Germany and other European countries, workers have started to stay in the workforce longer.[1] In West Germany the average retirement age rose by about 2 years between 1980 and 2005 (DRV 2006).[2] When workers stay active longer, the expected returns to workplace training increase. This should render human capital investments more profitable and affect the frequency of training. We first show that indeed the incidence of workplace training has increased over time and particularly so among older employees. Then we turn to our main question and investigate whether this increase in the training incidence has equally affected high and low skilled workers. If particularly the highly skilled workers received additional training rendering training a complement to education, the equity problem pointed out by the OECD intensified. If low skilled workers benefited more from human capital investments and training is provided as a substitute to formal education, the labor

---

[1] Among these changes are reduced durations of unemployment benefits for older unemployed workers, the abolition of early retirement pathways, and the introduction of benefit discounts following early retirement (see e.g. OECD (2006) for detail).
[2] Hoffmann (2007) finds an increase by 0.8 years between 2000 and 2005 alone.

market disadvantages of the least educated workers are reduced by the recent developments.

This dynamic perspective is new to a literature that typically investigates the correlation between formal education and training at just one point in time. In an early study Lynch and Black (1998) use data on employers in the United States in 1994 to investigate both the employer- and employee-related patterns of employer-provided training. They conclude that training complements formal education and that there is a 'virtuous circle' for those who obtained schooling early on.[3] In their survey paper Bassanini et al. (2005) use data of the European Community Household Panel and confirm that education and training are positively correlated. The same conclusion is drawn by Blundell et al. (1999) in their review of the returns to the 'three main components of human capital', i.e. ability, formal education, and on the job training. Several search and matching models interested in describing equilibria with respect to training provision start out with the assumption that education and training are complements (e.g. Laing et al. 1995, Brunello and Medio 2001, or Burdett and Smith 2002).

However, the empirical evidence it not clear cut. Krueger and Rouse (1998) find a significant positive correlation between training and schooling in manufacturing but not for the service sector. Van Smoorenburg and van der Velden (2000) study the training experience of a Dutch sample of school leavers. They confirm that the more educated have a higher probability of receiving firm training. But they also show that the overeducated receive less training while the undereducated do not receive more training as would have been suggested by a matching theory argument, according to which training bridges the differences between available and required skills. The authors find substitutability for some educational degrees and types of employment, while for others education and

[3] Bellmann and Düll (1999) obtain similar results with German establishment data.

training are complements. Ariga and Brunello (2006) applied employee data from Thailand to study the relationship between education and training. The authors found both, significant substitutability as well as complementarity between formal education and training, here depending on the type of employer-provided training considered.

The issue remains unsettled. The assumption that formal education and training are complements is plausible, if employers' returns to training the well educated exceed the returns to training lesser skilled workers. This could be the case if the former have better learning skills and lower marginal training costs than those with less education. Also, the highly educated might be able to extend their active life by more than employees who may have damaged their health in a working life of predominantly manual labor.[4] On the other hand there are arguments for the substitutability of education and training: Groot et al. (1994) found that the better educated demand more compensation to participate in training and Sicherman (1991) shows empirically that employers invest less in overeducated workers because these are more likely to quit and move to more suitable jobs.

Using data from the German *Mikrozensus* 1996 – 2004 we describe the incidence of training and its change for high and low skilled workers. First, we apply two different decomposition procedures to determine whether the increase in workplace training relates to changes in worker characteristics or e.g. to changes in age-specific training probabilities. Then we study skill group-specific trends in the provision of training in multivariate regressions.

We find that high skill workers receive more training than low skill workers, which suggests a complementarity relationship of training and education in Germany. Also, the incidence of training has increased substantially over time, and the relative provision of training to older workers increased more for high than for low skill workers.

[4] Recent contributions by Cunha and Heckman (2007) also suggest complementarity relationships based on the technology of human capital acquisition, where "learning begets learning."

Changes in the training incidence are neither due to developments of the population age structure nor to other shifts in the characteristics of workers and their jobs. Instead, the provision of training increased for given worker characteristics over time. Such behavioral changes might reflect the increase in returns to training which follows from prolonged working lives of employees.

2. Data and Descriptive Evidence

The German *Mikrozensus* surveys the residents of one percent of all German dwellings. The dataset provides large, nationally representative, annual samples of over 800,000 individual observations.[5] Since 1996 a random 45 percent of the full sample was asked about training, which generates the group of individuals we look at here. Our sample considers full-time workers, aged 25-65 who have been employed over the course of the last calendar year. Excluded are apprentices, military personnel, family helpers, and the self-employed. Because our results would be sensitive to a systematic shift in labor force participation over the period of our analysis we show in Figure 1 that that the share of our sample as a fraction of the relevant *Mikrozensus* population did not change in a systematic way.[6]

We apply an indicator of vocational qualification to categorize workers as high or low skilled employees: those with tertiary (polytechnic and university) and advanced vocational degrees (e.g. master of crafts or technician) are defined as high skilled, all others are considered as low skilled. Based on this definition about one in three individuals is high skilled.

The first part of our analysis compares the training propensity for the years 1996 and 2004, when individuals were asked about their participation in training for

[5] Other national data on training such as the German Socioeconomic Panel (GSOEP) or *Berichtssystem Weiterbildung* (BMBF 2003) have substantially smaller samples and are not available on an annual basis.
[6] We also inspected in Figure 1 by sex and did not find vast shifts in labor force participation either.

professional purposes over the course of the last year.[7] After omitting observations with missing values e.g. on vocational qualification our sample contains 49,768 and 45,860 observations for 1996 and 2004 respectively.[8]

Figure 2(a) describes the incidence of training by age and skill group in 1996 and 2004. Over time the propensity to participate in continuing education increased for both groups, which, however, may in part be due to the changed wording of the question. Low skill employees have a decidedly lower training incidence. Training becomes less common with age for both skill groups in both years. Figure 2(b) describes the change in the training incidence by skill group over time.[9] For both skill groups particularly those above age 45 experienced increases in training probabilities over time, flattening the age profile of training provision.[10] Table 1 depicts the change in training incidence by skill group and industry across all age groups. The training incidence increased for all groups and always more so for high than for low skill workers.

In order to test whether these overall trends are specific to the *Mikrozensus* data we consulted evidence on the change in training incidence based on the German Socioeconomic Panel (GSOEP) data. In this survey the exact same questions on training were asked at different points in time. Table 2 presents the share of individuals indicating that they participated in any training over the last three years as well as their participation in specific types of training. The sample consists of full time employed males and females aged 25-65. As in the *Mikrozensus* the sample shares receiving training increased over the

[7] The wording of the question on training participation changed somewhat between surveys. We assume that any changes in answer behavior over time are independent of respondent age.

[8] The questionnaire explicitly states that the information on vocational qualification is provided on a voluntary basis for individuals age 51 and above. However, the share of missing values on the vocational degree does not increase vastly for older age groups. Those who do not provide the information on vocational training are more likely to be non-German and blue collar workers than those who do.

[9] To calculate these differences we first normalized the observed values for 2004. All age group-specific probabilities were adjusted by the constant ratio of the 1996 probability of training across all age groups relative to the same 2004 probability. The ratio amounts to 0.15 / 0.30 for high skill workers and to 0.06 / 0.11 for low skill workers.

[10] These patterns are similar for both sexes when considered separately.

considered period and particularly so for older workers. Our subsequent analyses are performed with *Mikrozensus* data as it provides substantially larger samples.

3. Algebraic Decomposition of Changes in Training

In order to determine the relevance of population aging for the observed developments we decompose the total change in the probability of training between 1996 and 2004, separately for the full sample and for the high and low skill subsamples. The overall probability of training at time t, $P_t(tr)$, can be described as the weighted sum of age-specific training probabilities:

$$P_t(tr) = \sum_{a=25}^{65} \left[P_t\left(tr \middle| Age_a \right) \cdot P_t\left(Age_a \right) \right] \tag{1}$$

To see how the change in training propensities between 1996 and 2004 can be decomposed into changes in the age-specific training propensities and in the population age distribution, consider the following decomposition:

$$\begin{aligned}
\Delta P(tr) &= P_{04}(tr) - P_{96}(tr) \\
&= \sum_{a=25}^{65} \left[P_{04}\left(tr \middle| Age_a \right) \cdot P_{04}\left(Age_a \right) \right] - \sum_{a=25}^{65} \left[P_{96}\left(tr \middle| Age_a \right) \cdot P_{96}\left(Age_a \right) \right] \\
&= \sum_{a=25}^{65} \left[P_{04}\left(tr \middle| Age_a \right) - P_{96}\left(tr \middle| Age_a \right) \right] P_{04}\left(Age_a \right) - \sum_{a=25}^{65} P_{96}\left(tr \middle| Age_a \right) \left[P_{96}\left(Age_a \right) - P_{04}\left(Age_a \right) \right] \\
&= \sum_{a=25}^{65} \left[\Delta P\left(tr \middle| Age_a \right) \cdot P_{04}\left(Age_a \right) \right] + \sum_{a=25}^{65} \left[\Delta P\left(Age_a \right) \cdot P_{96}\left(tr \middle| Age_a \right) \right] \\
&= \quad \textit{shift effect} \quad + \quad \textit{age structure effect}
\end{aligned} \tag{2}$$

We label the first part of this expression the "shift effect" because it reflects the share of the total change, $\Delta P(tr)$, that is independent of changes in the population age structure and due only to shifts in age-specific training probabilities. In contrast, the second part labeled "age structure effect" measures that part of the total change, $\Delta P(tr)$, that is due to changes in the population age structure and independent of behavioral changes.

Next, we can decompose the "shift effect" further, to describe changes in training probabilities for specific age groups:

$$
\begin{aligned}
shift &= \sum \left[\Delta P\left(tr \middle| Age_a\right) \cdot P_{04}\left(Age_a\right) \right] \\
&= \overline{\Delta P\left(tr \middle| Age_a\right)} + \sum \left\{ \left[\Delta P\left(tr \middle| Age_a\right) - \overline{\Delta P\left(tr \middle| Age_a\right)} \right] \cdot P_{04}\left(Age_a\right) \right\} \\
&= \overline{\Delta P\left(tr \middle| Age_a\right)} + \sum \delta_a
\end{aligned}
\tag{3}
$$

where $\overline{\Delta P\left(tr \middle| Age_a\right)} = \frac{1}{65-25} \sum_{a=25}^{65} \left[\Delta P\left(tr \middle| Age_a\right) \right]$ describes the "Average Shift" of age-specific training probabilities over time. It would also capture the effects of changes in the wording of the survey question over time. The second term of the equation, $\sum \delta_a$, sums the weighted "specific age effects", δ_a. If the training propensities had changed in exactly the same manner for all age years, then all specific age effects, δ_a, were zero. If, however, particularly older workers receive more training than before, we would expect larger "specific age effects" δ_a for these older age groups than for others.

The results are summarized in Table 3 for the two subsamples. While the changes in the training incidence are different for the two subsamples with an increase of almost 15 percentage points for the high skilled and only about 5 percentage points for the low skilled workers, the decompositions for the two subsamples yield very similar patterns: the increase in the training incidence results predominantly from a shift effect. The age structure effect is negative, which indicates that population aging *per se* would have reduced the overall training probability over time. The vast shift effects reflect considerable changes in age-specific training probabilities. Applying the decomposition of equation (3) we find that most of this change is due to an overall increase in age-specific training probabilities: the average shift alone explains just about the entire observed change, while the sum of the specific age effects is small. However, these sums hide substantial differences across age groups which are depicted in Figure 3 for the two subsamples of high and low skilled workers. We see that the age-specific changes are generally larger for the high skill groups. Also, the figure shows increases in the age-specific training incidence for the age groups 40 through 54 among low skill workers and for the age group 45 through 59 for high skill workers. This is the development we would expect if the extension of the work life provided an incentive for employers to invest more and longer in the training of the work force.

4. Regression-Based Decomposition

After finding that most of the increase in training probabilities was due to an overall increase in age-specific training probabilities, we now apply a different approach to study the determinants of changes in training. Instead of differentiating only the effects of the population *age structure* from general behavioral changes we now look at the changes of *all* potential determinants of training. In the spirit of Oaxaca-Blinder decompositions we separately evaluate the effects of changes in the values of training determinants and of changes in these determinants' association with the incidence of training.

Again, the analysis is performed separately for the high and low skill subsamples. The first two columns of Table 4 present descriptive statistics on those individual and employment characteristics of the full sample in 1996 and 2004, which are typically considered as determinants of receiving training. A comparison of the covariate means for the two years indicates that the characteristics of the sample have changed only somewhat over time: on average, workers aged, the share of blue collar workers declined, that of white collar workers increased, and the fraction of high skill workers increased slightly.

The last two columns of Table 4 provide the marginal effects of probit estimations of the determinants of individual training by skill in 2004.[11] Particularly with respect to the marginal effects of individual characteristics we observe substantial differences in the determinants of training across skill groups. The coefficient estimates of the quadratic age effect (not presented) imply that training falls with age for low skill workers above age 25

[11] The results for 1996 are not presented to save space but are available upon request.

and for high skill workers above age 34. Among the more interesting results is the finding that females receive significantly less training among low skilled workers but suffer no disadvantage in the high skill sample. To the contrary, non-Germans are particularly disadvantaged in the high skill group but less so among low skill workers. In both samples, civil servants have the highest training incidence. The pseudo R^2 value of both regressions is relatively low at 7.4 and 4.5 percent for low and high skill workers, respectively: only a small fraction of the overall change in the training probability is subject to the systematic impact of the considered determinants.

As a second step we now apply a version of the Oaxaca-Blinder decomposition to quantify the relative impact of changes in the explanatory variables and of changes in their effects for the overall development of training propensities over time. We apply the procedure developed by Fairlie (1999, 2005) to translate the Oaxaca-Blinder decomposition for a situation with a bivariate dependent variable. The effects of changes in parameters (α) and covariates (X) can be distinguished using equation (4):

$$\begin{aligned}\Delta P(tr) &= \left\{\overline{P}(\alpha_{04}, X_{04}) - \overline{P}(\alpha_{96}, X_{04})\right\} + \left\{\overline{P}(\alpha_P, X_{04}) - \overline{P}(\alpha_P, X_{96})\right\} \\ &= \quad \textit{parameter effect} \quad + \quad \textit{characteristics effect}\end{aligned} \tag{4}$$

$\overline{P}(\alpha_{04}, X_{04})$ represents the average predicted probability of receiving training, where every worker's characteristics (X) are as observed in 2004 and the parameters (α) of the probit estimation for 2004 are applied. The first term in equation (4) ("parameter effect") considers the differential in average training probabilities that results when the 2004 characteristics are used with both the 2004 and the 1996 parameter vector. We focus on the second term, the "characteristics effect." It evaluates the difference in predicted training probabilities when the individual characteristics of different years are applied to a parameter vector α that is held constant. The particular values of α can be set to those of the 1996, 2004, and the pooled regression. In our calculations we use the latter. This second term indicates the extent to which the change in training probabilities over time can be attributed to changes in worker characteristics.

An additional interesting option within this framework of analysis is to decompose the "characteristics effect" further and to measure the extent to which certain groups of covariates explain the total "characteristics effect." To measure the effect of only the group of covariates X_k we evaluate

$$\overline{P}\left(\alpha_P^k X_{04}^k + \alpha_P^{\neg k} X_{04}^{\neg k}\right) - \overline{P}\left(\alpha_P^k X_{96}^k + \alpha_P^{\neg k} X_{04}^{\neg k}\right). \qquad (5)$$

Each group of covariates k can be evaluated separately and their individual contributions add up to the total "characteristics effect" as in equation (4). The distinguishing feature of the Fairlie approach is that it calculates the average of individual predictions instead of predicting at average covariate values, which is usually done (see e.g. Shields 1998).[12] The problem of matching observations on X^k from different years is solved using a procedure akin to propensity score matching (c.f. Fairlie 2005). The standard errors are calculated using the delta method.[13]

The results of our analysis are summarized in Table 5. Again, we start with raw differences in training probabilities of 14.7 and 5.1 percentage points for the high skill and low skill sample between 1996 and 2004. When we use the parameter estimates of the pooled regression no more than about 2 percent of the raw differences for high and low skill workers can be explained by changes in characteristics over time.[14] When we investigate the main determinants of the small effect of covariate changes we obtain the results presented in the bottom part of Table 5: the change in sample age had a

---

[12] In a logit model with a constant term the average of the predicted values exactly matches the sample average, i.e. equation (4) holds exactly. This is neither the case for the probit estimator when the predicted values are calculated based on average covariate values.

[13] We apply the Stata9 algorithm "fairlie" provided by Jann (2006).

[14] We also calculated the decomposition effects when using the 1996 and 2004 coefficients as base values. However, most likely due to the small explanatory power of the changes in characteristics we obtained implausible results.

considerable effect, however, it yields a *decline* in probabilities rather than the observed increase. Instead, just about all the other significant effects contribute to explain the increase in training probabilities. Among high skill workers the other large effects derive from the change in the workforce's marital status and region of residence over time, among low skill workers the largest effect besides that of age derives from shifts between blue and white collar workers over time. Overall, we can explain only a small portion of the changes in training probabilities based on worker characteristics. This leaves more than 95 percent of the difference to be explained by either changes in employee and employer behaviors, by changes in survey design, or by other "unexplained" factors.

5. Multivariate Analysis of the Change in Training Probabilities over Time

In this section we take a closer look at the dynamics of the change in training probabilities. We explain the probability of individual training applying a probit model to pooled annual data for the years from 1996 to 2004. In order to model the dynamics of training probabilities for the different age groups we interact a linear time trend with age group indicators. We use individual year dummies to control for the main time effect. This allows to flexibly account for changes in answering behavior over time and to capture annual developments in training probabilities. Again the models are estimated separately by skill group.

The estimation results for the key indicators of training participation are presented in Table 6 for both skill subsamples. The marginal effects on the interaction terms are statistically significant. They yield a clear pattern of increasing training participation over time for all age groups. A comparison of the marginal effects for low and high skill employees yields again much larger effects for the latter: most of the marginal effects are at least twice as large for high compared to low skill employees. This suggests that the expansion of training disproportionately benefited those with better education and that training was provided as a complement to formal education.

To illustrate these developments we calculated the predicted probability of training for a married, German, male reference person, who works in Berlin, in a firm with more than 50 employees in manufacturing, separately for each age and skill group. Table 7 presents the change in the predicted probability of training for this person between 1996 and 2004 by age and skill group based on separate regressions for the two subsamples. As all entries are positive we conclude that the probability of training increased in each age and skill group in the considered period. Since the entries increase almost linearly with age it appears that the oldest workers benefited most from the expansion of training over time and more so among the high than among the low skill employees. The probability of training tripled among high skill workers above age 45.

Whereas Table 7 focuses on a comparison between the years 1996 and 2004, Figure 4 utilizes the results of the pooled probit regressions to present the development of training probabilities on an annual basis. It depicts the ratio of the predicted training probability for those aged 60-65 relative to the probability predicted for those in age group 25-30. In 1996 the probability of training of those aged 60-65 reached on average about 15 percent of the training probability of the young. This probability changed over time to about 35 percent by 2004. Therefore, even if the level of training for older workers is always much lower than that for younger workers, we see a change. Also here, the age ratio for high skill workers is above that for low skill workers and increased faster. Thus, among the high skilled, older workers always had a higher chance of receiving training relative to the young and these chances grew faster than among low skill workers. The rising training propensity for older workers over time might reflect that firms increasingly employ older workers and particularly those with high skills for a longer work-life.

6. Conclusions

This study uses German *Mikrozensus* data of the last decade (1996–2004) on high and low skill workers in order to answer the question whether education and training are complements or substitutes. Overall, the incidence of training rose over time and particularly for older workers when due to institutional changes and demographic aging the expected duration of employees' work lives went up as well. This matches closely the conclusions of Bassanini et al. (2005) who point out that the age gradient in the training incidence flattens as the employment rate of older workers increases.

We apply a decomposition analysis in order to determine the relative importance of population aging in this development: population aging would have reduced the training incidence, while we see in fact an overall increase. Most of this overall increase falls equally on all population age groups, even though workers above age 45 appear to benefit more than their younger colleagues.

In a regression-based decomposition we find that neither population aging nor other shifts in the characteristics of workers contribute substantially to explain the observed changes. More than 90 percent of the increase in the training incidence remains to be explained by other factors, such as changes in the behavior of employers and employees. Using regressions on pooled annual data we find that the positive trend in the training incidence lasted over the entire considered period, and covered all age and skill groups. The probability of training for older compared to younger workers increased substantially.

The incidence of training differs vastly by the skill level of workers. Already in 1996 high skill workers are twice as likely to report training compared to low skill workers. The increase in training provision over time did not change this pattern. The distribution of training across age groups shifted somewhat from the young to the old for all skill groups. The increase in the incidence of training disproportionately benefited the group of skilled workers. Across all age groups high skill workers experienced larger increases in the training propensity than low skill workers. This suggests that the complementarity relationship between education and training intensified over time and that the group of low skilled workers fell even further behind with respect to their stock of human capital and earnings potential in the labor market. Apparently, training is not provided in a manner that balances the existing labor market problems of low skill workers. This is a fact any active labor market policy promoting training for older workers should be aware of.

References

Ariga, Kenn and Giorigo Brunello, 2006, Are education and training always complements? Evidence from Thailand, *Industrial and Labor Relations Review* 59(4), 613-629.

Bassanini, Andrea, Alison Booth, Giorgio Brunello, Maria de Paola, and Edwin Leuven, 2005, Workplace Training in Europe, *IZA Discussion Paper* No. 1640, Bonn / Germany.

Bellmann, Lutz and Herbert Düll, 1999, Die Bedeutung des beruflichen Bildungsabschlusses in der betrieblichen Weiterbildung – Eine Analyse auf der Basis des IAB Betriebspanels 1997 für West- und Ostdeutschland, in: Bellmann Lutz and Viktor Steiner (eds.) *Panelanalysen zur Lohnstruktur, Qualifikation und Beschäftigungsdynamik*, Nürnberg, 317-352.

Blundell, Richard, Lorraine Dearden, Costas Meghir, and Barbara Sianesi, 1999, Human Capital Investment: The Returns from Education and Training to the Individual, the Firm, and the Economy, *Fiscal Studies* 20(1), 1-23.

BMBF (Bundesministerium für Bildung und Forschung), 2003, *Berichtssystem Weiterbildung 2000*, Bonn.

Brunello, Giorgio and Alfredo Medio, 2001, An explanation of international differences in education and workplace training, *European Economic Review* 45(2), 307-322.

Burdett, Ken and Eric Smith, 2002, The low skill trap, *European Economic Review* 46(8), 1439-1451.

Cunha, Flavio and James J. Heckman, 2007, The Technology of Skill Formation, *IZA Discussion Paper* No. 2550, Bonn.

DRV (Deutsche Rentenversicherung), 2006, Deutsche Rentenversicherung in Zahlen 2006, download 21.2.2007 at http://www.deutsche-rentenversicherung.de/nn_23924/SharedDocs/de/Inhalt/04__Formulare__Publikationen/03__publikationen/Statistiken/Broschueren/rv__in__zahlen__pdf,property=publicationFile.pdf/rv_in_zahlen_pdf

Fairlie, Robert W., 1999, The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment, *Journal of Labor Economics* 17(1), 80-108.

Fairlie, Robert W., 2005, An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models, *Journal of Economic and Social Measurement* 30(4), 305-316.

Groot, Wim, Joop Hartog, Hessel Oosterbeek, 1994, Costs and Revenues of Investment in Enterprise Related Schooling, *Oxford Economic Papers* 46(4), 658-675.

Hoffmann, Hilmar, 2007, Wege in den Ruhestand, *Deutsche Rentenversicherung* 2007(4-5), 298-320.

Jann, Ben, 2006, fairlie – Nonlinear decomposition of binary outcome differentials, software module available with Stata 9 (downloaded August 3, 2006).

Krueger, Alan and Cecilia Rouse, 1998, The Effect of Workplace Education on Earnings, Turnover, and Job Performance, *Journal of Labor Economics* 16(1), 61-94.

Laing, Derek, Theodore Palivos, and Ping Wang, 1995, Learning, Matching and Growth, *Review of Economic Studies* 62(1), 115-129.

Lynch, Lisa M. and Sandra E. Black, 1998, Beyond the incidence of employer-provided training, *Industrial and Labor Relations Review* 52(1), 64-81.

OECD, 1999, *Employment Outlook*, Paris.

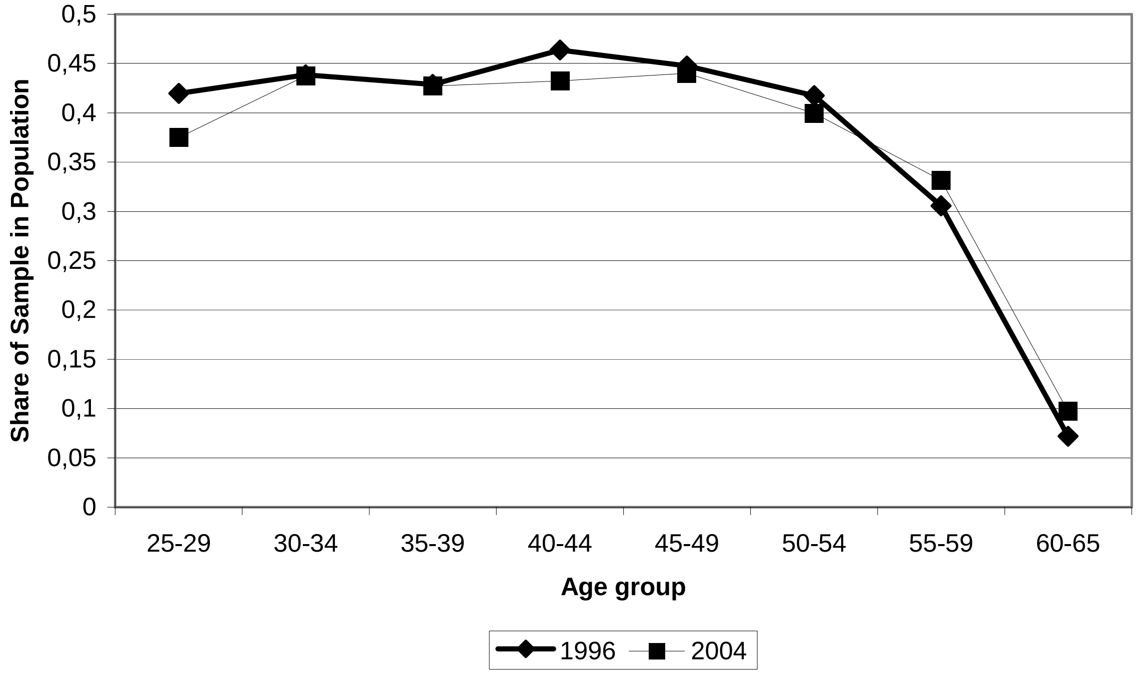
OECD, 2006, *Economic Survey of Germany 2006*, Paris.

Shields, Michel, 1998, Changes in the Determinants of Employer-Funded Training for Full-Time Employees in Britain, 1984-1994, *Oxford Bulletin of Economics and Statistics* 60(2), 189-214.

Sicherman, Nachum, 1991, 'Overeducation' in the Labor Market, *Journal of Labor Economics* 9(2), 101-122,

Van Smoorenberg, M.S.M and R.K.W. van der Velden, 2000, The training of school-leavers. Complementarity or substitution?, *Economics of Education Review* 19(2), 207-217.

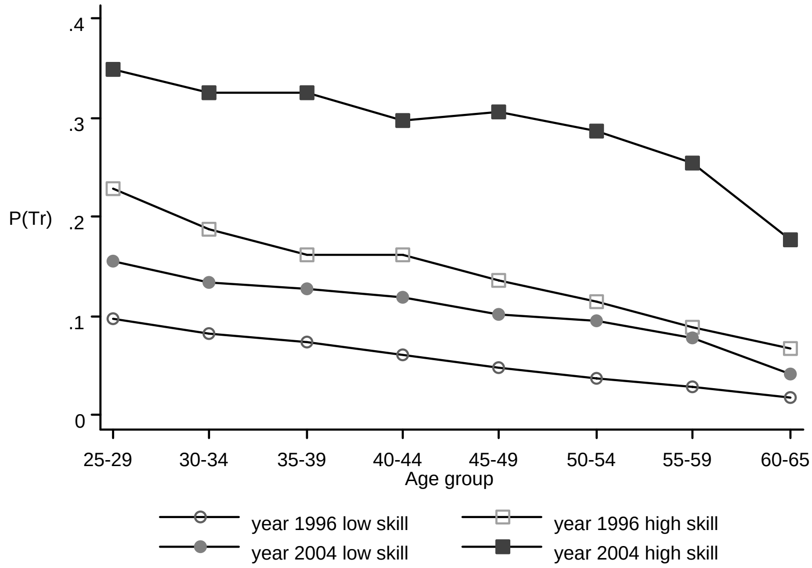
Figure 1 Population share of our sample over time



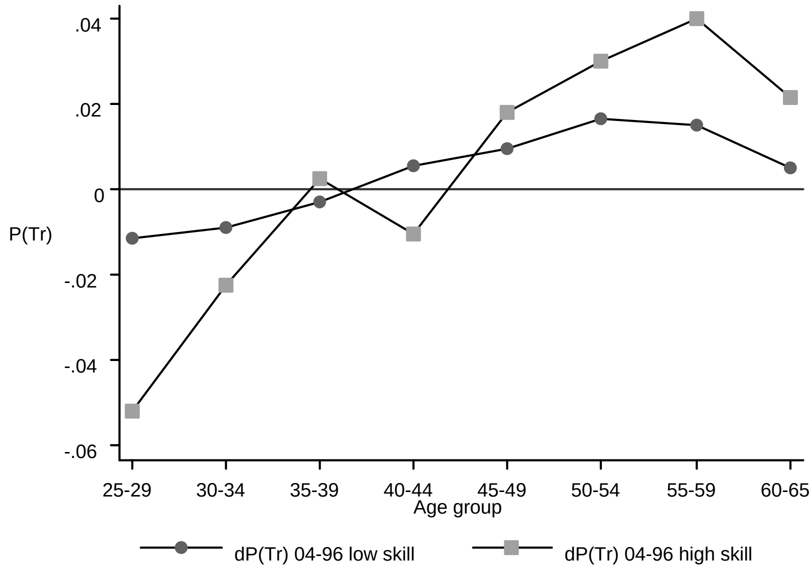
Source: Own calculations based on German Mikrozensus 1996 and 2004.

Figure 2 Training Incidence by Age Group, Skill Group, and Year

(a) Levels of Training Incidence by Age Group, Skill Group, and Year



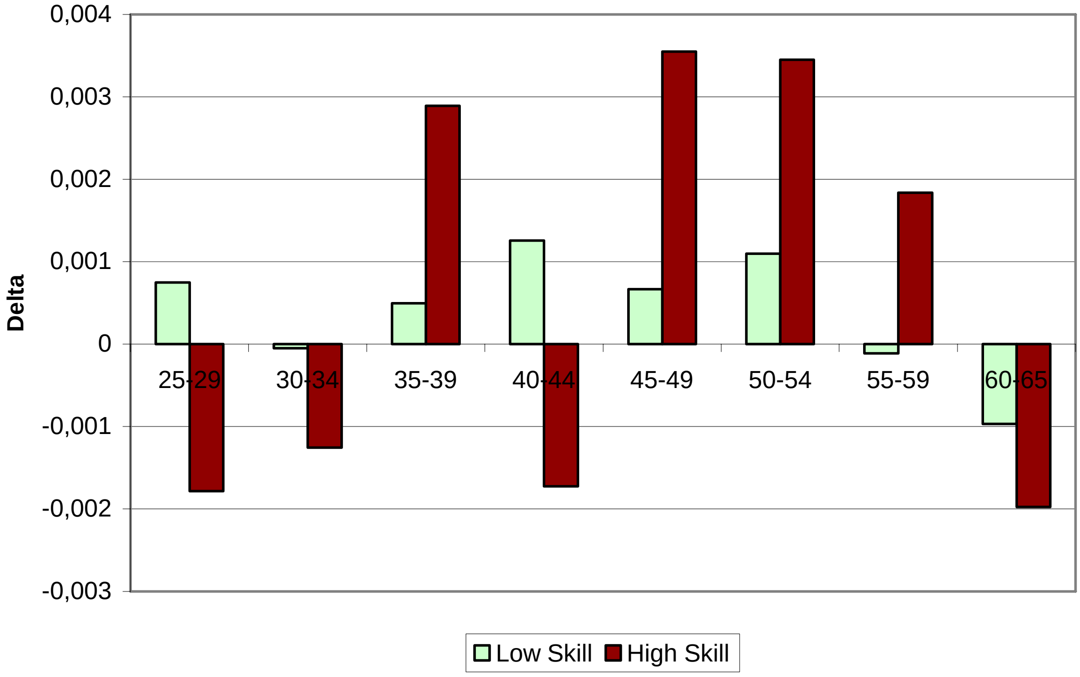
(b) Changes in the Training Incidence over Time by Age and Skill Group



Note: Figure 1(b) provides the difference of the normalized training incidence in 2004 and the observed training incidence in 1996. The 2004 data were normalized to account for the overall increase in the training incidence that might be due to changes in the questionnaire. The normalization was performed by dividing all observed age group-specific training probabilities of 2004 by the same ratio of the overall average training probability for 1996 over that of 2004, for high skill workers 0.15 / 0.30 and for low skill workers 0.07 / 0.12.

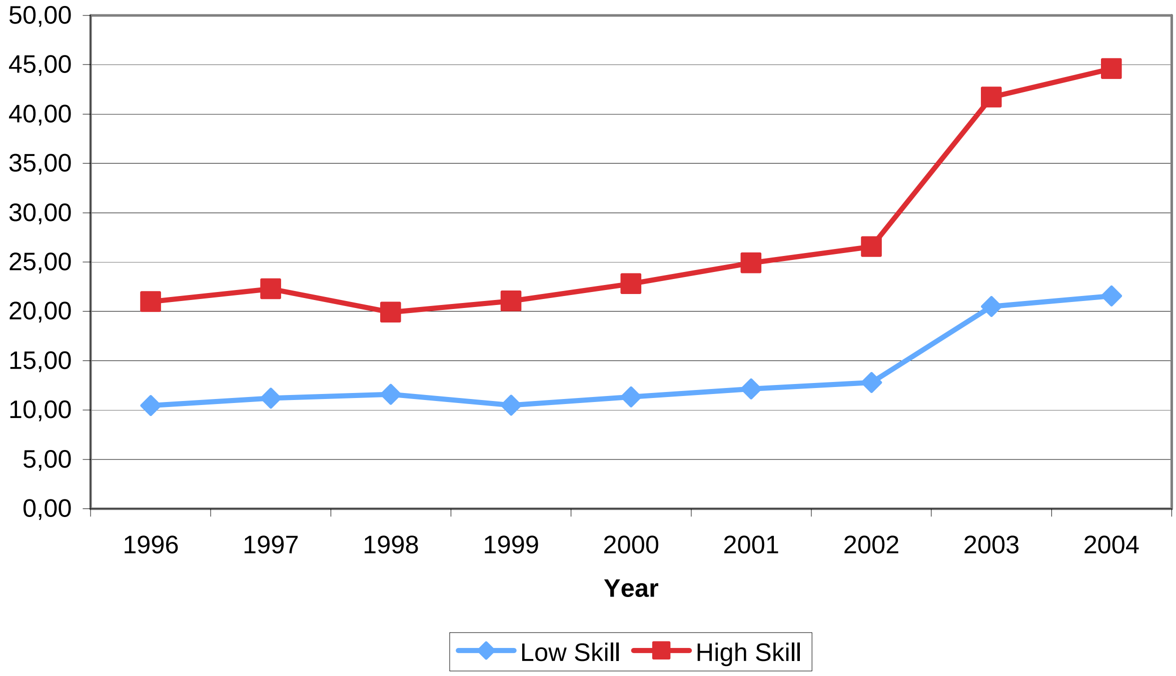
Source: Own calculations based on German Mikrozensus 1996 and 2004.

Figure 3 Age-specific Change in the Conditional Training Probability (δ)



Source: Own calculations based on German Mikrozensus 1996 and 2004.

Figure 4 Relative Predicted Training Probability of Older Workers (60-65) Compared to Younger Workers (25-30) (in percent)



Source: Own calculations based on German Mikrozensus 1996 to 2004.

Table 1 Changes in Training Incidence by Skill Level and Industry

		Low skill			High skill			Difference
Industry	**Share**	**1996**	**2004**	**Diff.**	**1996**	**2004**	**Diff.**	**Diff.Hi - Diff.Lo**
Electricity, Gas, Water	0,014	0,09	0,17	0,07	0,12	0,35	0,24	0,16
Mining	0,007	0,04	0,07	0,02	0,08	0,24	0,15	0,13
Education	0,055	0,11	0,21	0,11	0,20	0,42	0,23	0,12
Health Sector	0,083	0,11	0,18	0,07	0,22	0,41	0,19	0,12
Manufacturing	0,300	0,05	0,08	0,03	0,12	0,23	0,12	0,09
Services	0,039	0,08	0,10	0,03	0,15	0,26	0,11	0,08
Agriculture, Forestry	0,015	0,02	0,05	0,04	0,08	0,20	0,12	0,08
Hotels and Restaurants	0,016	0,02	0,05	0,03	0,09	0,19	0,11	0,08
Wholesale, Retail Trade	0,110	0,05	0,10	0,05	0,11	0,23	0,12	0,07
Finance, Insurance	0,042	0,14	0,24	0,10	0,18	0,36	0,17	0,07
Construction	0,087	0,03	0,08	0,05	0,10	0,21	0,11	0,07
Real Estate and Rental	0,056	0,08	0,14	0,06	0,15	0,26	0,11	0,06
Public Administration	0,108	0,10	0,17	0,07	0,18	0,30	0,12	0,05
Transport, Communic.	0,067	0,07	0,11	0,04	0,14	0,22	0,09	0,04
Internat. Organizations	0,001	0,10	0,23	0,12	0,09	0,23	0,14	0,02

Source: Own calculations based on German Mikrozensus 1996 and 2004.

Table 2 Training Incidence by Age over Time - GSOEP Data (in percent)

	Any training last 3 yrs		Type of Training					
			Regular Reading		Conferences&Conventions		Vocational Courses	
Agegroup	**1993**	**2004**	**1993**	**2004**	**1993**	**2004**	**1993**	**2004**
All	28.7	34.0	42.2	55.1	20.9	30.3	24.8	34.2
25-39	32.7	36.7	42.6	50.6	21.4	28.4	33.7	36.8
40-65	24.9	32.3	41.9	57.8	20.3	31.4	26.0	32.7

Source: Own calculations based on German Socioeconomic Panel (GSOEP).

Table 3 Results of the Algebraic Decomposition of Changes in Training Incidence

	Total Change	Shift Effect			Age-Structure Effect
		Total =	Average Shift +	Sum of Specific Age Effects	
High Skill	0.1474	0.1517	0.1467	0.0050	-0.0043
Low Skill	0.0507	0.0540	0.0509	0.0031	-0.0033

Source: Own calculations based on German Mikrozensus 1996 and 2004.

Table 4	Data Description and Probit Marginal Effects for the Probability of Reporting Training for the High and Low Skill Samples

	Mean (Std. dev.) Full Sample 1996	Mean (Std. dev.) Full Sample 2004	ME Probit Low Skilled 2004	ME Probit High Skilled 2004
High skilled	0.278	0.312	–	–
Low skilled	0.722	0.688	–	–
Age	41,316	42,633	0.001	0.016**
	(10.079)	(9.638)	(0.49)	(4.32)
Age squared / 100	18,086	19,105	-0.003**	-0.023**
	(8.572)	(8.372)	(2.62)	(5.55)
Sex (male=1)	0,674	0,665	0.022**	0.003
			(8.97)	(0.30)
Marital status (married=1)	0,686	0,636	-0.012**	-0.019*
			(4.68)	(2.15)
Nationality (German=1)	0,950	0,948	0.022**	0.095**
			(4.35)	(4.16)
Civil servant	0,090	0,085	ref.	ref.
White collar worker	0,516	0,566	-0.012*	-0.068**
			(2.50)	(5.62)
Blue collar worker	0,394	0,349	-0.065**	-0.212**
			(11.80)	(12.05)
Firmsize 1-10 workers	0,125	0,128	ref.	ref.
Firmsize 11-19 workers	0,096	0,099	-0.002	0.004
			(0.48)	(0.21)
Firmsize 20-49 workers	0,139	0,139	0.006	0.022
			(0.78)	(1.31)
Firmsize at least 50 workers	0,634	0,621	0.016**	0.019
			(4.65)	(1.41)
Firmsize unknown	0,007	0,011	-0.009	-0.019
			(0.63)	(0.48)
Observations	49,768	45,860	31,564	14,296
Pseudo R-squared	–	–	0.0743	0.0451

Notes: The columns entitled M.E. represent marginal effects and absolute values of z-statistic in parentheses. ** and * indicate statistical significance at the 1 and 5 percent level. Not presented are the marginal effects for 15 federal states and 10 industries. The descriptive statistics by skill groups as well as estimations for 1996 are not presented to save space.

Table 5 Results of Regression Decomposition: Effect of Changed Characteristics

	High Skill		Low Skill	
Percentage point difference to be explained:	0.147		0.051	
Share of total difference explained:	2.36 %		2.42 %	
Explained effect due to:				
Age	-147,8%	**	-499,6%	**
Sex	2,1%		115,0%	**
Marital Status	98,2%	**	111,5%	**
Nationality	-6,4%	**	13,0%	**
Region of Residence	93,7%	**	140,9%	**
Blue / White Collar / Civil Servant	9,0%		220,4%	**
Firmsize	6,1%		-18,5%	**
Industry	44,9%	**	17,6%	*

Note: ** and * indicate statistical significance at the 1 and 5 percent level. The standard errors were obtained using the delta method.

Table 6 Probit Coefficient Estimates and Marginal Effects of the Probability of Training (Pooled Data 1996 to 2004)

	M.E. (z-statistic) low skill	M.E. (z-statistic) high skill
Time indicators, omitted: 1996		
1997	-0.000	-0.006+
	(0.17)	(1.74)
1998	-0.004	-0.059**
	(0.69)	(16.40)
1999	-0.022**	-0.064**
	(13.64)	(16.23)
2000	-0.021**	-0.066**
	(12.30)	(15.12)
2001	-0.022**	-0.067**
	(11.61)	(13.46)
2002	-0.024**	-0.073**
	(11.74)	(13.33)
2003	0.034**	0.069**
	(11.24)	(7.87)
2004	0.032**	0.071**
	(9.71)	(7.17)
Age group dummies, omitted: 25-29		
30-34	-0.009**	-0.015*
	(3.66)	(2.30)
35-39	-0.019**	-0.042**
	(7.73)	(6.74)
40-44	-0.027**	-0.055**
	(10.81)	(9.00)
45-49	-0.032**	-0.071**
	(13.24)	(11.94)
50-54	-0.040**	-0.077**
	(16.01)	(12.82)
55-59	-0.043**	-0.096**
	(17.30)	(16.62)
60-65	-0.046**	-0.096**
	(10.86)	(11.24)
Interaction effects between age and linear trend		
trend * agegroup 30-34	-0.001	0.001
	(1.11)	(0.62)
trend * agegroup 35-39	0.001+	0.005**
	(1.77)	(3.88)
trend * agegroup 40-44	0.001**	0.005**
	(2.98)	(3.66)
trend * agegroup 45-49	0.002**	0.007**
	(3.40)	(5.60)
trend * agegroup 50-54	0.003**	0.008**
	(4.91)	(5.97)
trend * agegroup 55-59	0.002**	0.009**
	(3.33)	(6.28)
trend * agegroup 60-65	0.003+	0.007**
	(1.89)	(3.13)
Observations	270,343	142,166
Pseudo R squared	0.099	0.101

Notes: The column entitled M.E. represents marginal effects, absolute values of z-statistic are presented in parentheses. **, *, and + indicate statistical significance at the 1, 5, and 10 percent level. Not presented are the separate effects of indicators for 15 federal states, 10 industries, and the personal and job characteristics described in Table 3. The estimations are based on observations over 9 years (1996 to 204).

Table 7 Increase in the Predicted Probability of Training between 1996 and 2004 based on the Probit Estimations in Table 5 (in percent)

Age group	low skill	high skill
25-30	61	64
30-35	53	75
35-40	91	135
40-45	114	141
45-50	131	199
50-55	184	224
55-60	167	293
60-65	232	248

Note: The reference person is a married, male, German, blue collar worker. He works in Berlin in a firm with more than 50 workers in the manufacturing industry. The entries were calculated as ((predicted probability 2004 / predicted probability 1996) – 1) * 100 %.