Age and Skill Biased Technological Change: A Multiple Treatment Approach Using a Linked Employer Employee Dataset

Reinhard Hujer and Dubravko Radić*

October 26, 2005

Abstract

Although there is a large literature on the skill biased nature of technological change, studies on its impact on the age structure of labor demand are rather scarce. Due to knowledge depreciation and lower incentives to invest in human capital, however, older employees within a certain skill group might be more affected by technological change than younger ones. In this paper we present empirical evidence for Germany using a unique linked employer employee dataset and applying a multiple treatment framework. We find empirical evidence for both the skill and age biased technological change hypothesis. The age bias, however, becomes weaker once we control for the skill structure. Additionally, we find that older low skilled employees are less affected relative to younger low skilled employees, while on the other hand older high skilled employees benefit less relative to younger high skilled employees. The existence of such an asymmetry of skills has important policy implications in an ageing society and might serve as an explanation for the early retirement decision of older employees.

Keywords: Age and Skill Biased Technological Change, Multiple Treatment
JEL Classification: J23, O33, C14

*Reinhard Hujer is Professor of Statistics and Econometrics at the Institute of Statistics and Econometrics, J.W. Goethe-University of Frankfurt and Research Affiliate of the IZA, Bonn and the ZEW, Mannheim, e-mail: hujer@wiwi.uni-frankfurt.de. Dubravko Radić is Research Assistant at the Department of Service Management, University of Leipzig, e-mail: radic@wifa.uni-leipzig.de. Thanks to Lutz Bellmann from the Institute for Employment Research for data support and to Marija Ugarković for helpful comments. Financial support of the German Scientific Foundation is gratefully acknowledged. The usual disclaimer applies.

1 Introduction

There is a broad consensus among researchers about the skill biased nature of technological and organizational change.[1] And although it is a natural idea to assume that an omnipresent skill biased technological change will affect all employees within a given skill group in the same way, there are a couple of reasons why older employees might be even more affected:

– Older workers received their education longer ago and due to knowledge depreciation their skills may have become obsolete in the meantime.
– Older workers have less incentives to invest in new knowledge since their time horizon until retirement is shorter as compared to younger employees.
– In contrast to seniority wage payment schemes, individual job performance is found to decrease with increasing age.[2]
– There are prejudices on behalf of human resource managers about the willingness to learn and flexibility of older employees.[3]

As a consequence, an increasing technological and organizational change which requires more flexibility and new skill structures might harm older employees more than younger ones. Establishments will reduce their demand for older employees or older employees may decide to withdraw from the labor market and retire. On the other hand, older employees are also more experienced than younger ones and legally protected against age discrimination.[4] The impact of technological change on the age structure of labor demand is thus not clear cut and needs to be addressed empirically.

Despite a large literature on skill biased technological and organizational change, studies on their impact on the age structure of labor demand are rather scarce. For the U.S., Friedberg (1999) analyzes the impact of computer use on workers of different ages. She finds, after controlling for possible endogeneity effects, that the probability for employees older than 50 years to use a computer at their job decreases dramatically and at the same time they are 28% more likely to retire than their computer using counterparts. Using

[1]For an overview see e.g. Chennels and van Reenen (2002).

[2]For some gerontological findings see e.g. Avolio and Waldman (1994) or Skirbekk (2003).

[3]Boockmann and Zwick (2004) present some empirical findings for Germany. In their paper, they analyze the determinants of labor demand for older employees.

[4]An example for the U.S. is the Age Discrimination in Employment Act of 1967 (ADEA) which protects individuals who are 40 years of age or older from employment discrimination based on age.

the Health and Retirement Study, Ahituv and Zeira (2001) show that retirement and unemployment of older workers are positively related to technological progress. Bartel and Sicherman (1993), finally, focus on the impact of technological change on the retirement decision of older workers between 1966 and 1983 in the U.S. using the National Longitudinal Surveys of Older Men. They find that costs of training associated with unanticipated technological change led workers to retire earlier.

Heywood, Ho, and Wei (1999) conducted a survey among Hong Kong establishments and demonstrate that the presence of greater skill requirements and seniority payment schemes are both strongly associated with reduced hiring opportunities for older employees. For France, Aubert, Caroli, and Roger (2004) show that the wages of older employees are lower in innovative firms and that the opposite holds for younger employees. Analyzing the inflow and outflow of employment reveals that older employees are hurt by reducing hiring opportunities due to new technologies while on the other hand, organizational changes increase the outflow of older employees.

A recent study for Germany was conducted by Beckmann (2005). Using the IAB Establishment Panel Dataset he shows that technological and organizational changes decrease the demand for older employees while increasing it for younger ones. However, due to data limitations, this study is not able to consider different skill levels of employees in different age groups. Another shortcoming, again due to data restrictions, is not taking account of unobserved heterogeneity.

In the following we present additional empirical evidence for Germany. Using a unique linked employer employee dataset for 1993 to 1997 and applying a multiple treatment framework we analyze the impact of technological and organizational change on the age and skill structure of labor demand. In the next section we introduce the dataset and offer some first descriptive results. Section three discusses the econometric issues while section four contains the estimation results. Finally, section five derives some stylized facts and concludes.

2 Dataset and First Descriptive Results

In the following we briefly describe the dataset and variables used to test the skill and age biased technological change hypothesis. We are able to use an unique linked employer employee dataset provided by the German Federal Employment Office, namely the Linked IAB Establishment Panel dataset (LIAB) which is available through the combination of the IAB Establishment Panel containing detailed information about the employer side and the Employment Statistics Register covering the employees.

The IAB Establishment Panel is conducted by the German Federal Employment Office.[5] Its population are all firms employing at least one employee subject to the compulsory social security system. The panel started out in 1993 with 4,365 West German establishments. Since 1996, East German establishments were also included in the dataset which contained about 15,526 establishments in 2000. The IAB Panel is organized in a modular form. Some topics, such as employment policy, business volume and investment, are covered annually while other topics are only covered irregularly, e.g. information about innovations and organizational changes.

Employee information come from the Employment Statistics Register at the Federal Employment Office where all employers are obliged to report information about their staff which are subject to the compulsory social security system. Notifications include information about sex, age, qualification, occupation and wages of the employees. Employment groups such as public servants e.g. are not included in the dataset but still about 85% of all employees in Germany are covered.[6] Both datasets can be merged by an unique firm identification number yielding an unique linked employer employee dataset for Germany.[7]

The relevant treatment variables reflecting the innovation behavior of establishments are taken from the IAB Establishment Panel. We apply a broad innovation concept which contains product as well as process innovations but also organizational changes as a form of disembodied innovations.[8] The following treatments are considered:

– **Product:** Product innovations introduced before 1993 (yes/no), taken from the 1993

[5] More details on the IAB Establishment Panel can be found in Bellmann (1997).

[6] See e.g. Bender, Haas, and Klose (2000) for detailed information about this dataset.

[7] The availability of such datasets has been rather limited until recently when virtually an explosion in the use of such datasets initiated. See e.g. Abowd and Kramarz (1999).

[8] For a discussion of the notion of innovation and measurement issues see e.g. Griliches (1990) or Radić (2005).

wave of the IAB Establishment Panel. Product innovations comprise the introduction of a product/service which was new for the own establishment (local innovation) or even new for the market (global innovation).

- **Process:** Process innovations introduced in 1992 (yes/no), taken from the 1993 wave of the IAB Establishment Panel. The dummy variable Process equals one if investments were undertaken which aimed at a reduction of labor and operations costs.
- **Orga**: Organizational changes introduced before 1993 (yes/no), taken from the 1995 wave of the IAB Establishment Panel. Orga is equal to one if an establishment has introduced one of the following measures: Flattening of hierarchical levels, decentralization of decision authority, introduction of group workplaces with own responsibilities, introduction of units with own costing and result calculation and introduction of just-in-time production.

The outcome variables are taken from the Employment Statistics Register at the Federal Employment Office for the years 1993 to 1997. We consider the following employment groups and combinations between them:

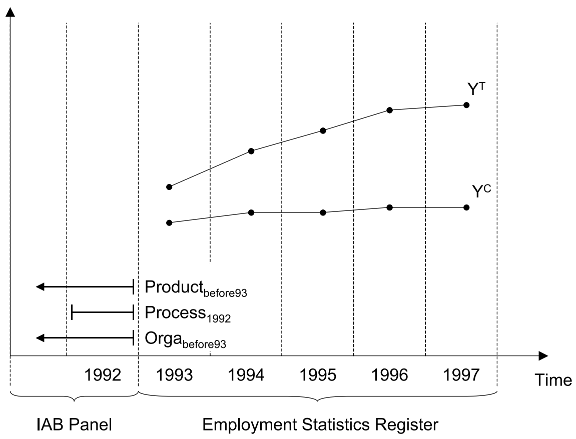
- **Total:** Total number of employees.
- **High Skilled:** Share of employees with an university or equivalent degree.
- **Medium Skilled:** Share of employees with an apprenticeship.
- **Low Skilled:** Share of employees with no formal educational attainment.
- **Young Employees:** Share of employees younger than 30 years.
- **Middle Employees:** Share of employees between 30 and 50 years old.
- **Older Employees:** Share of employees older than 50 years.

Since we are especially interested in possible discrimination effects against low skilled and older employees, we leave out medium skilled employees and those which are between 30 and 50 years old in the following analyses. Note that we are working with employment shares which sum up to one so that these shares can be deduced from the remaining information.

Combining all establishments which participated in the IAB Establishment Panel in 1993 and 1995 with employee information from 1993 to 1997 yields a total number of

2,785 West German establishments.[9] Excluding establishments from the agricultural and public sector reduces the sample size to 2,429 establishments.[10] The following Figure 1 summarizes the assumed dependency structure where Y^T illustrates an exemplary development of the outcome variable (e.g. labor demand for high skilled employees) for treated and Y^C for control establishments.

Figure 1: Dependency Structure



Before discussing appropriate econometric methods which can be used to assess the causal effects of innovation activities on labor demand we present some simple descriptive results. Table 1 contains the share of establishments which have either introduced a product innovation, process innovation, organizational changes or combinations of these. One can see that most of the establishments (31%) have neither introduced a product nor a process innovation nor an organizational change. 34% of all establishments only introduced a single measure. A considerable number of establishments (25%), however, introduced two and 10% even all three measures simultaneously.

Table 2 contains the means and standard deviations of the outcome variables for 1993 to 1997. The figures reveal that total employment has declined during this time from 704 to 565 employees. On the other hand, the skill and age composition has remained fairly stable for high skilled and older employees and decreased by one percentage point for low skilled employees and by even six percentage points for younger employees.

[9] Note that East-German establishments were not included until 1996 in the IAB Establishment Panel.

[10] Several comments are in order regarding this dataset. Employee information are available for the years 1993 to 1998. Due to a structural break in the data collection methods in 1998 (cf. Neidert (1998)), employment level information from this year are not directly comparable with previous years. We decided therefore to exclude them from the further analysis. Questions regarding innovation behavior were only

Table 1: Distribution of the Treatment Variable

Treatment	Number of Establishments	Share
0 Nothing	749	0.31
1 Product	389	0.16
2 Process	245	0.10
3 Orga	189	0.08
4 Product & Process	269	0.11
5 Product & Orga	198	0.08
6 Process & Orga	138	0.06
7 Product, Process & Orga	252	0.10

Table 2: Descriptive Results on Total Employment and Employment Structure

	1993		1994		1995		1996		1997	
Variable	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Total	704	1,762	625	1,477	618	1,462	580	1,400	565	1,382
High	0.06	0.11	0.06	0.11	0.07	0.11	0.06	0.11	0.06	0.11
Low	0.27	0.23	0.27	0.23	0.26	0.22	0.26	0.22	0.26	0.23
Young	0.29	0.19	0.27	0.18	0.26	0.18	0.24	0.18	0.23	0.17
Old	0.22	0.16	0.22	0.15	0.22	0.15	0.22	0.15	0.23	0.16

In the following we will show how these outcome variables differ for establishments which have introduced product innovations, process innovations and organizational changes as compared to establishments which were not innovative at all. Figure 2 shows first of all that innovative establishments are on average larger than their non-innovative counterparts. In the following years both groups have experienced declining employment. However, innovative establishments were affected more strongly by this decline as can be seen by the dashed line: Between 1993 and 1997 innovative establishments on average reduced employment by about 300 employees as compared to non-innovative ones.

The graphs in Figure 3 give a first impression about the impact of innovations (measured again by the simultaneous introduction of product innovations, process innovations and organizational changes) on the labor structure. The upper left (right) figure reveals that innovations increase (reduce) the demand for high (low) skilled employees. Simple mean comparisons therefore indicate that technological change is skill biased. Regarding different age groups the patterns are not so clear cut. We find that both groups of young

included in the IAB survey in 1993 and then again in 1998 so that we are using those from 1993.

and older employees are set off by technological change. It is the task of the next sections to verify whether these simple mean comparisons also reflect the true causal relationships or if these findings are simply caused by sample selection effects.

Figure 2: Simple Mean Comparison for Total Employment (Number of Employees)

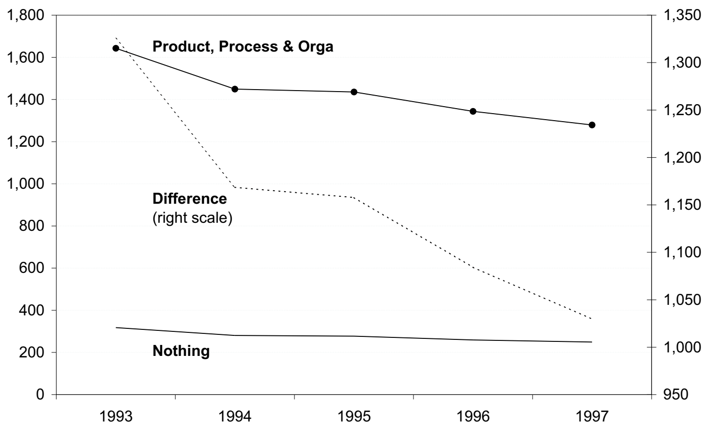
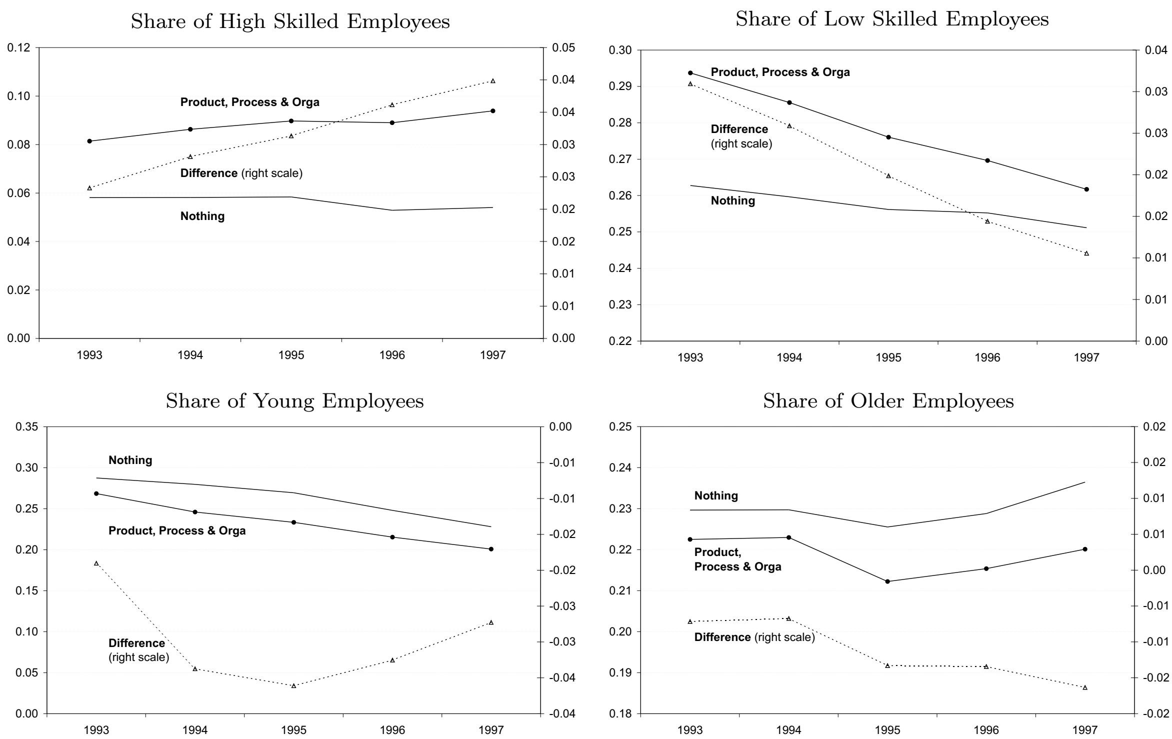


Figure 3: Simple Mean Comparison for Different Skill and Age Groups



3 A Multiple Treatment Framework

The basic binomial treatment framework, which has become the working horse in the evaluation literature in the meantime, can be used to estimate the causal effect of a certain treatment over no treatment on a certain outcome variable. In our application, however, establishments decide whether to introduce a product innovation, a process innovation, an organizational change or to do nothing. This yields a total number of $2^3 = 8$ disjunctive treatment decisions. Table 1 which was presented previously contains the distribution of this treatment variable and makes clear that a large fraction of establishments (35%) has introduced more than one measure simultaneously.

Since it is reasonable to assume that the introduction of a product innovation has other employment effects than a process innovation or even the introduction of both, it makes no sense to aggregate all possible combinations into one single treatment.[11] The simple binomial setting therefore needs to be extended to a multiple treatment framework which can take this possible impact heterogeneity into account. Figure 4 illustrates the difference between a binomial and a multiple treatment setting. In the upper part of the figure, an individual experiences a treatment, i.e. $D_i = 1$. In the post-treatment period we observe for this individual $Y_i(D_i = 1)$ and the counterfactual outcome is given by $Y_i(D_i = 0)$. If more than one treatment is possible, depending on the treatment, we observe one of the following outcomes: $Y_i(D_i = 1), \ldots, Y_i(D_i = 7)$. The fundamental evaluation problem occurs since the other counterfactual outcomes are not observable.[12]

Let $D_i \in \{0, 1, \ldots, M\}$ indicate participation in a particular treatment. Similar to the binomial case, different pair-wise treatment effects can be defined (cf. Lechner (2001, 2002)):

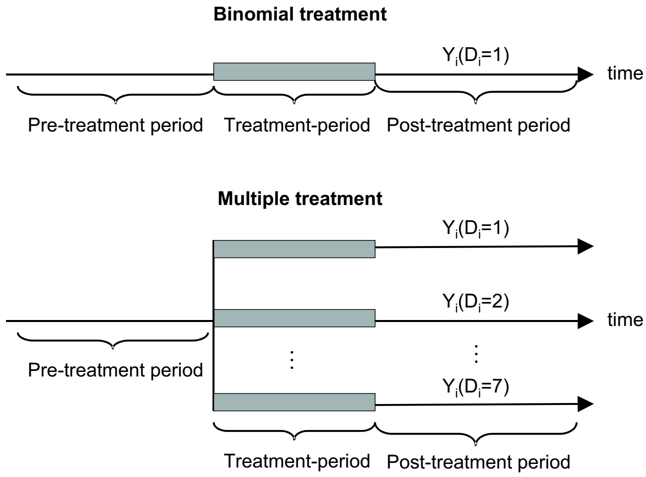
$$\begin{aligned}
\Delta^{ml} &= E_i[Y_i(m) - Y_i(l)] \\
\Delta^{ml|m,l} &= E_i[Y_i(m) - Y_i(l)|D_i = m, l] \\
\Delta^{ml|m} &= E_i[Y_i(m) - Y_i(l)|D_i = m].
\end{aligned}$$

Δ^{ml} gives the average treatment effect of treatment m compared to treatment l for a randomly selected individual, whereas $\Delta^{ml|m,l}$ and $\Delta^{ml|m}$ condition this effect on partici-

[11] See Plewis (2002) for a more thorough discussion on impact heterogeneity. See Spiezia and Vivarelli (2002) for an overview of the employment effects of technological change.

[12] More details can be found in Lechner (2001, 2002, 2002b) and Gerfin and Lechner (2002).

Figure 4: Binomial versus Multiple Treatments



pants randomly selected from the group of participants in treatment m or l and only m, respectively.

In the binomial case with one treatment and one non-treatment category, only one possible pairwise comparison arises, namely treatment as compared to non-treatment. Given $M+1$ different treatments including the non-treatment case on the other hand implies that $\binom{M+1}{2}$ different pairwise comparisons can be estimated. For $m=1$ and $l=0$ e.g., $\Delta^{ml|m}$ measures the employment impact for an establishment which has introduced a product innovation as compared to the non-treatment case, i.e. it answers the question whether product innovations create or destroy jobs.

Identification of $\Delta^{ml|m} = E_i[Y_i(m) - Y_i(l)|D_i = m]$ requires only identification of $E_i[Y_i(l)|D_i = m]$, whereas for the identification of $\Delta^{ml|m,l}$ one needs also the expression $E_i[Y_i(m)|D_i = l]$ according to the following decomposition:

$$\begin{aligned}
\Delta^{ml|m,l} &= E_i[Y_i(m) - Y_i(l)|D_i = m, l] \\
&= p(D_i = m|D_i = m, l)E_i[Y_i(m) - Y_i(l)|D_i = m] \\
&\quad + (1 - p(D_i = m|D_i = m, l))\, E_i[Y_i(m) - Y_i(l)|D_i = l] \\
&= p(D_i = m|D_i = m, l)\Delta^{ml|m} + (1 - p(D_i = m|D_i = m, l))\, \Delta^{ml|l}
\end{aligned}$$

with $p(D_i = m|D_i = m, l) = \frac{p(D_i=m)}{p(D_i=l)+p(D_i=m)}$.

Identification of Δ^{ml} requires even the identification of all counterfactuals of $Y_i(m)$

and $Y_i(l)$ since Δ^{ml} can be decomposed according to:

$$\Delta^{ml} = \sum_{j=0}^{M} \left(E_i[Y(m)|D_i = j] - E_i[Y(l)|D_i = j]\right) p(D_i = j) = \sum_{j=0}^{M} \Delta^{ml|j} p(D_i = j).$$

Hence, if $\Delta^{ml|j}$ is identified for all $j = 0, 1, \ldots, M$ all other effects are defined as well.

The conditional independence assumption of the binomial framework $Y_i(0), Y_i(1) \perp D_i | X_i$ can be extended to the multiple framework.[13] It is given by:

$$Y_i(0), Y_i(1), \ldots, Y_i(M) \perp D_i | X_i$$

and can also be generalized to the following independence assumption conditional on the propensity score:

$$Y_i(0), Y_i(1), \ldots, Y_i(M) \perp D_i | \left(p^1(X_i), p^2(X_i), \ldots, p^M(X_i)\right).$$

In the binomial case the propensity score is the probability of an individual to participate in a treatment. Since in case of the multiple framework we have $M+1$ different treatments the propensity score now becomes a vector of M independent probabilities $p^j(X_i) = p(D_i = j|X_i)$.[14]

There are a set of conditional independence assumptions which identify some of the possible pairwise treatment effects. The first one is the strongest and states that:

$$Y_i(m), Y_i(l) \perp D_i | \left(p^1(X_i), p^2(X_i), \ldots, p^M(X_i)\right).$$

This form of conditional independence implies that all counterfactuals are identified since for all $j = 0, 1, \ldots, M$:

$$\begin{aligned} E_i[Y_i(m)|D_i = m, p(X_i)] &= E_i[Y_i(m)|D_i = j, p(X_i)] \\ E_i[Y_i(l)|D_i = l, p(X_i)] &= E_i[Y_i(l)|D_i = j, p(X_i)] \end{aligned}$$

with $p(X_i) = \left(p^1(X_i), p^2(X_i), \ldots, p^M(X_i)\right)$. Therefore, all average treatment effects $\Delta^{ml|m}$, $\Delta^{lm|l}$, $\Delta^{ml|m,l}$ and Δ^{ml} can be identified.

The second less restrictive form of conditional independence assumption is relaxed to hold only for the subpopulation of individuals participating either in treatment m or l:

$$Y_i(m), Y_i(l) \perp D_i | \left(p(D_i = l|D_i = m, l)\right), D_i \in \{m, l\}.$$

[13] Cf. Imbens (2000).

[14] Thus if the number of elements in X is less than M it makes no sense to condition on the propensity vector since the dimensionality even increases.

Under this assumption only the counterfactuals $E_i[Y_i(m)|D_i = l, p(X_i)]$ and $E_i[Y_i(l)|D_i = m, p(X_i)]$ are identified using $E_i[Y_i(m)|D_i = m, p(X_i)]$ and $E_i[Y_i(l)|D_i = l, p(X_i)]$, respectively. Hence only $\Delta^{ml|m}$, $\Delta^{lm|l}$ and $\Delta^{ml|m,l}$ can be estimated.

The least restrictive conditional independence assumption, however, is obtained if we further relax the second one to hold only for the potential outcome $Y_i(l)$, i.e.

$$Y_i(l) \perp D_i | \left(p(D_i = l|D_i = m, l)\right), D_i \in \{m, l\}.$$

In this case only $E_i[Y_i(l)|D_i = m, p(X_i)]$ can be identified using the observable counterpart $E_i[Y_i(l)|D_i = l, p(X_i)]$. As a conclusion only $\Delta^{ml|m}$ is estimable.

In the following we will concentrate on these treatment effects $\Delta^{ml|m}$ for $l = 0$ and $m = 1, 2, \ldots, 7$ thereby considering various outcome variables like e.g. employment shares for high and low skilled and young and elder employees, respectively (see also Table 3). A comparison of the different treatment effects can be used to reveal complementarities. Introducing process innovations in combination with appropriate organizational changes e.g. might have a larger employment effect than introducing both measures in isolation. In this case the sum of $\Delta^{2,0|2}$ and $\Delta^{3,0|3}$ will not in general be equal to $\Delta^{6,0|6}$. It is in fact the existence of such complementarities which makes the application of a multiple treatment framework necessary.

Table 3: Considered Treatment Effects

Treatment Effect	Interpretation
$\Delta^{1,0|1} = E_i[Y_i(1) - Y_i(0)|D_i = 1]$	Isolated impact of only product innovations
$\Delta^{2,0|2} = E_i[Y_i(2) - Y_i(0)|D_i = 2]$	Isolated impact of only process innovations
$\Delta^{3,0|3} = E_i[Y_i(3) - Y_i(0)|D_i = 3]$	Isolated impact of only organizational changes
$\Delta^{4,0|4} = E_i[Y_i(4) - Y_i(0)|D_i = 4]$	Impact of Product & Process Innovations
$\Delta^{5,0|5} = E_i[Y_i(5) - Y_i(0)|D_i = 5]$	Impact of Product Innovations & Organizational Changes
$\Delta^{6,0|6} = E_i[Y_i(6) - Y_i(0)|D_i = 6]$	Impact of Process Innovations & Organizational Changes
$\Delta^{7,0|7} = E_i[Y_i(7) - Y_i(0)|D_i = 7]$	Impact of Product-, Process & Organizational Innovations

Estimating the average treatment effect on the treated in the multiple treatment framework builds on the basic idea for the binomial case. Once the conditional identifying assumption bas been imposed, the binomial matching estimator is given by:

$$\widehat{\Delta}^{PMAT} = E_{p(X_i)|D_i=1}\left[E_i[Y_i(1)|D_i = 1, p(X_i)] - E_i[Y_i(0)|D_i = 0, p(X_i)]\right] = \Delta^{ATE|T}$$

and amounts to finding for every participating individual one non-participating individual with almost the same propensity score $p(X_i)$.

In the multiple framework this basic idea is extended. To calculate e.g. $\Delta^{ml|m} = E_i[Y_i(m) - Y_i(l)|D_i = m]$, i.e. the average treatment effect of treatment m compared to treatment l on a randomly picked participant in treatment m, it is sufficient to impose the weakest form of conditional independence assumption, i.e. $Y_i(l) \perp D_i| \, (p(D_i = l|D_i = m, l))$, $D_i \in \{m, l\}$. Note that we have to condition only on an one-dimensional scalar $p(D_i = l|D_i = m, l)$ instead of a M-dimensional vector of all probabilities $p^1(X_i), \ldots, p^M(X_i)$. The appendix contains some practical issues how to construct an adequate control group and how to conduct the matching for multiple treatments.

4 Estimation Results

The first step of every matching protocol consists of estimating the vector of M different independent probabilities $p^j(X_i) = p(D_i = j|X_i)$ for $j = 1, 2, \ldots, 7$. In our context, various multinomial models can be applied to this aim. Since the multinomial Logit model is based on the unrealistic 'independence of irrelevant alternatives'-assumption, a more flexible model would be the multinomial Probit. For a large number of alternatives, however, the estimation is computationally burdensome. As a remedy, Lechner (2001) suggests to estimate a series of conditional binomial models which are sufficient to identify the treatment effects, $\Delta^{m0|m}$, considered in this paper:[15]

$$\begin{aligned} p(D_i = 0|D_i = 1, 0) &= 1 - p(D_i = 1|D_i = 1, 0) \\ &\vdots \\ p(D_i = 0|D_i = 7, 0) &= 1 - p(D_i = 7|D_i = 7, 0). \end{aligned}$$

For the estimation we have included a variety of control variables which are expected to have an impact on innovation activities.[16] The Schumpeter size hypothesis was considered by the number and squared number of employees in 1992 (EMPL and EMPL^2). Cash flow was approximated by the variable PROFIT and CAPITAL which measure profitability in 1993 (1: Very good, ..., 5: Insufficient) and the fact whether the considered establishment is a capital company or not.[17] Internal technological capabilities are measured by the following variables: STATE (state of the technological equipment, 1: Up-to-date, ..., 5: Out-of-date), ICT (investment in ICT in 1992), R&D (existence of a R&D department) and MARKET (existence of a market research department). The demand situation is reflected by the variable DEVP which measures the business development between 1992 and 1993 (1: Decreasing, 2: Constant, 3: Increasing). Due to insufficient data, technological opportunities and appropriability conditions can only be approximated by the use of industry dummies.[18] Additionally, we also included a dummy variable indicating the existence of a workers council (WORKER).[19]

---

[15] Lechner (2001) compares the performance of the multinomial Probit model and this approach and finds little difference.

[16] For an economic justification of these variables see e.g. Cohen (1995).

[17] Unfortunately we had no information about business development in 1992.

[18] We considered the following industries: Electricity and mining, primary industry, investment goods industry, consumer goods industry, construction, commerce, transportation and communication, finance

Table 4: Estimation Results for the Conditional Probit Models

Variable	Product Innovations Parameter	Product Innovations t-Value	Process Innovations Parameter	Process Innovations t-Value	Organizational Changes Parameter	Organizational Changes t-Value
Constant	-1.40	-5.82	-1.55	-6.31	-1.66	-6.8
EMPL	0.00	0.97	0.00	3.80	0.00	4.77
EMPL^2	0.00	-0.16	0.00	-2.57	0.00	-3.04
PROFIT	0.01	0.38	0.10	2.92	0.06	1.69
CAPITAL	0.13	1.93	0.20	2.76	0.38	5.04
STATE	-0.02	-0.51	-0.06	-1.39	0.08	1.83
ICT	0.31	4.57	0.59	8.00	0.36	4.85
R&D	0.36	4.07	0.27	2.97	0.00	0.04
MARKET	0.24	3.10	0.14	1.79	0.04	0.57
DEVP	0.08	1.89	-0.01	-0.14	0.02	0.41
WORKER	0.07	0.97	0.21	2.64	0.38	4.68
	χ^2-Value	p-Value	χ^2-Value	p-Value	χ^2-Value	p-Value
LR-Industry	55.31	0.00	58.12	0.00	28.31	0.00
LR-Overall	249.21	0.00	501.90	0.00	328.62	0.00
Pseudo R^2	0.09		0.19		0.13	
Observations	910		795		750	

For illustrative purposes, we report the estimation results for the first three complementary conditional probabilities in table 4. All estimated models are highly significant according to overall likelihood ratio tests and show pseudo R^2 values ranging from 0.09 for the product innovation equation to 0.19 for process innovations. Size of the establishment is significant for the process and organizational change decision. For these equations we also find non-linear albeit small effects. As expected, a better cash flow situation has a positive impact on innovation activities. Investment in ICT have a significant and positive impact in all three equations whereas state of technology is found to have only an impact on organizational changes. The existence of a R&D and market research department increases product and process innovations. Better demand conditions as indicated by an increasing business development spur product innovations. Finally, industrial affiliation turns out to be significant in all estimated models.

Table 5 contains the appropriate marginal effects. The existence of a R&D department is the dominant factor for product innovation activities and increases the probability by 14% given that all other variables take on their mean values. Equally important are investments in ICT (12%) and the existence of a market research department (9%). The

and insurance, consumer services.

[19] Regional dummies as well as interaction effects turned out to be insignificant in the estimation.

Table 5: Estimated Conditional Marginal Effects

Variable	Product Innovations Parameter	t-Value	Process Innovations Parameter	t-Value	Organizational Changes Parameter	t-Value
EMPL	0.00	0.97	0.00	3.79	0.00	4.74
EMPL^2	-0.00	-0.16	-0.00	-2.57	-0.00	-3.04
PROFIT	0.004	0.38	0.04	2.92	0.02	1.69
CAPITAL	0.05	1.94	0.08	2.80	0.13	5.25
STATE	-0.008	-0.51	-0.02	-1.39	0.03	1.83
ICT	0.12	4.64	0.21	8.57	0.12	5.05
R&D	0.14	4.10	0.11	2.95	0.001	0.04
MARKET	0.09	3.11	0.05	1.78	0.02	0.57
DEVP	0.03	1.89	-0.002	-0.14	0.01	0.41
WORKER	0.02	0.97	0.08	2.67	0.13	4.77

dummy variable indicating a capital company is also significant and increases the probability to introduce product innovations by 5%. Finally, the demand push hypothesis is also supported by the empirical evidence: An increasing business development increases the innovation probability by 3%.

Regarding the decision to introduce process innovations the most important and significant factors are investment in ICT (marginal effect: 21%), existence of a R&D department (11%), capital company (8%), market research (5%) and profitability (4%). As it was the case for the product innovation decision, internal technological capabilities play also the dominant role for process innovations. On the other hand, cash flow considerations and the demand situation have only a minor impact. Finally, organizational changes are mainly influenced by the CAPITAL dummy and the existence of a workers council which both increase their probability by 13%.

The appendix contains the matching protocol used for the estimation of the pair-wise treatment effects given in Table 3. Before presenting the results, several comments are in order. Imposing the common support condition ensures that the effects are only estimated in regions of the attribute space where observations from two treatment groups could be observed with a similar participation probability. Since we are only interested in the pair-wise treatment effects $\Delta^{ml|m}$ with $m \in \{1, 2, \ldots, 7\}$ and $l = 0$, it is sufficient to restrict the sample to the overlap for each pair.[20] Table 6 contains in the third (fourth) column the share of treated (control) establishments which were dropped due to the common support

[20]Cf. Gerfin and Lechner (2002), p. 868.

condition. The figures range from 13% to 30% for the group of treated and from 24% to 57% for the group of non-treated establishments.[21]

Another point concerns the fact that the matching is conducted with replacement. In a multiple treatment framework this is necessary because each treatment group will at the same time serve as a control group and hence it is not guaranteed that the number of treated will always be larger than the number of non-treated individuals. A potential consequence of matching with replacement is an increase in the variance. Looking at columns five and six of bale 6 reveals, however, that only a small fraction of control establishments are used more than two times as a matching partner. These figures are similar in their magnitude to other studies e.g. Lechner (2002a, 2002b). Hence, matching with replacement does not seem to pose special problems.

Table 6: Some Matching Diagnostics

Treatment	Matched Pairs	Common Support* Treated	Common Support* Controls	Share of Controls used ... two times	Share of Controls used ... > two times	ABIAS (%) Before	ABIAS (%) After
1	302	0.22	0.40	0.16	0.05	22.09	4.35
2	212	0.13	0.24	0.15	0.08	39.19	5.05
3	133	0.30	0.24	0.11	0.04	27.27	8.57
4	210	0.20	0.26	0.11	0.10	58.91	12.45
5	142	0.22	0.38	0.09	0.08	42.25	11.76
6	110	0.20	0.57	0.11	0.10	72.39	14.88
7	88	0.25	0.54	0.08	0.07	78.12	14.51

* Share of dropped establishments due to the common support condition.

The last point concerns the matching quality. Columns seven and eight in Table 6 contain the average absolute standardized percentage bias (ABIAS) before and after the matching. These figures show that the differences between treated and non-treated establishments are considerably reduced after the matching.

In the following we will present the estimated treatment effects for different outcome variables Y: Total number of employees, share of high versus low skilled employees and share of young versus older employees. Finally, we will take a closer look at the share of

[21] Especially the figures for the non-treated group seem quite large and raise doubts whether differences in the treatment effects are only due to differences in the common support. Lechner (2000) proposes several robustness checks in this context, e.g. estimating the treatment effects with and without imposing the common support condition. If both effects are nearly the same, as it is the case in our application, concentrating only on the subpopulation with common support would identify the true treatment effect. Cf. Lechner (2000), p. 11.

young and older employees differentiated by skill. The treatment variables, i.e. introduction of product innovations, process innovations, organizational changes or combinations of these, refer to a time period prior to 1993 (see also Figure 1 on page 5). Information for the outcome variables on the other hand are available for the time period from 1993 to 1997.

The estimated treatment effects are defined as follows:

$$\begin{aligned}
\Delta_{94}^{m0|m} &= E_i\left[\left(Y_i^{94}(m) - Y_i^{93}(m)\right) - \left(Y_i^{94}(0) - Y_i^{93}(0)\right)|D_i = m\right] \\
\Delta_{95}^{m0|m} &= E_i\left[\left(Y_i^{95}(m) - Y_i^{93}(m)\right) - \left(Y_i^{95}(0) - Y_i^{93}(0)\right)|D_i = m\right] \\
\Delta_{96}^{m0|m} &= E_i\left[\left(Y_i^{96}(m) - Y_i^{93}(m)\right) - \left(Y_i^{96}(0) - Y_i^{93}(0)\right)|D_i = m\right] \\
\Delta_{97}^{m0|m} &= E_i\left[\left(Y_i^{97}(m) - Y_i^{93}(m)\right) - \left(Y_i^{97}(0) - Y_i^{93}(0)\right)|D_i = m\right]
\end{aligned}$$

for $m = 1, 2, \ldots, 7$ and where $Y_i^t(m)$ measures the outcome variable for establishment i at year t under treatment m.

Note that the outcome variables are measured relative to the base year 1993 instead of using level information. Applying the matching protocol on such differences has several advantages: Comparing treated with non-treated establishments for the same years ensures that changing economic conditions which have a common impact on both groups cancel out. Matching on the propensity scores takes account of sample selection due to observable establishment specific covariates and avoids making some functional form assumptions. Using more than one year allows us to trace the dynamic structure of the employment effects of technological change since it can be expected that the adaption to innovations will take more than one year. And finally, by concentrating on differences of the outcome variables and neglecting level information we control for possible unobservable establishment specific effects.[22]

Matching Results for Total Employment

Table 7 contains the estimated average treatment effects on total employment. A simple comparison between treated and non-treated establishments presented previously indi-

[22] For clarification assume that $Y_i^t(m)$ is an additive separable function of an observable time and individual specific component x_i^t and an unobservable individual specific component ν_i: $Y_i^t(m) = x_i^t + \nu_i$. Taking differences adjusts for this unobservable heterogeneity: $Y_i^t(m) - Y_i^{t'}(m) = \left(x_i^t - x_i^{t'}\right) + (\nu_i - \nu_i) = x_i^t - x_i^{t'}$. For a similar proceeding see e.g. Hujer, Caliendo, and Radić (2001).

cated that the simultaneous introduction of product, process and organizational innovations has a negative impact on the total number of employees (see Figure 2, p. 7).

Taking account of sample selection due to observables by employing a multiple treatment framework changes these simple descriptive findings. Introducing product innovations in combination with organizational changes has now a significant employment enhancing effect.[23] Between 1993 and 1995 total employment increased by 13%-points in innovative establishments compared to their matched non-innovative counterparts. This treatment effect remains at the same level in the following years. For all other treatment effects, however, total employment evolves identically in innovative and non-innovative establishments.

Table 7: Matching Results for Changes in Total Employment (in %-Points)

	1993-1994		1993-1995		1993-1996		1993-1997	
Treatment	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value
Product	-0.06	-0.89	-0.11	-1.44	-0.06	-0.86	-0.09	-0.97
Process	0.01	0.28	-0.04	-0.62	0.00	-0.04	-0.03	-0.45
Orga	0.01	0.24	0.00	-0.07	-0.05	-0.96	0.00	-0.04
Product/Process	0.05	0.87	0.02	0.26	0.04	0.53	-0.01	-0.09
Product/Orga	0.07	1.56	**0.13**	2.44	**0.13**	2.11	**0.12**	1.85
Process/Orga	0.03	0.71	0.06	1.22	0.17	1.46	0.05	0.93
Prod/Proc/Orga	0.02	0.21	0.02	0.14	0.03	0.27	0.07	0.31

Matching Results for the Skill Biased Hypothesis

The estimation results contained in Table 8 can be used to test the skill biased technological change hypothesis. The figures show that organizational changes have a negative impact on the share of low skilled employees. Between 1993 and 1994 treated establishments employed 2%-points less low skilled employees than their matched non-treated counterparts. One year later this treatment effect increases to 3%-points and finally, between 1993 and 1996 it amounts to 4%-points. All other treatment effects, however, do not alter the evolvement of low skilled employees.

Organizational changes are also one of the main drivers of high skilled employment. Between 1993 and 1996 they increased the share of high skilled employees by 1%-point

[23] In all following tables bold figures indicate significance on a 5%-level.

compared to establishments which have not introduced any organizational changes. Process innovations have a similar quantitative impact, albeit this impact unfolds one year earlier. Process and organizational innovations introduced simultaneously, have a stronger but delayed impact on high skilled employees. Between 1993 and 1996 (1997) they increase their share by 1%-point (2%-points) compared to non-treated establishments.

All other treatment effects are not significant and thus do not change the evolvement of employment of high and low skilled employees. We therefore find support for the skill biased technological change hypothesis. The main driving force behind this increasing wedge are organizational changes within establishments. Additionally, we find slight evidence for a complementary relation between process and organizational innovations which enfold a stronger impact when introduced simultaneously.

Table 8: Matching Results for the Share of Low and High Skilled (in %-Points)

	1993-1994		1993-1995		1993-1996		1993-1997	
Treatment	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value	$\Delta^{m,0\|m}$	t-Value
				Low Skilled Employees				
Product	0.00	-0.24	0.00	-0.06	0.01	0.67	0.03	1.51
Process	0.01	0.89	0.00	0.04	0.00	0.12	0.01	0.38
Orga	**-0.02**	-1.62	**-0.03**	-2.23	**-0.04**	-2.78	-0.01	-0.97
Product/Process	0.02	0.95	0.02	0.61	0.03	1.10	0.02	0.81
Product/Orga	0.00	0.37	0.00	-0.05	0.00	-0.28	0.00	0.04
Process/Orga	0.01	0.62	-0.01	-0.67	0.00	0.12	0.00	0.05
Prod/Proc/Orga	0.00	0.12	-0.01	-0.64	0.00	-0.12	0.00	-0.17
				High Skilled Employees				
Product	0.00	0.11	0.00	0.90	0.00	0.68	0.00	0.43
Process	0.00	1.45	**0.01**	1.67	0.01	1.08	0.01	0.93
Orga	0.00	0.37	0.00	0.04	**0.01**	2.31	0.01	1.48
Product/Process	0.00	-0.32	0.00	0.71	0.01	0.63	0.00	0.70
Product/Orga	0.00	1.07	0.01	1.26	0.01	0.77	0.01	0.91
Process/Orga	0.00	0.83	0.01	1.30	**0.01**	1.87	**0.02**	1.76
Prod/Proc/Orga	0.00	0.97	0.00	0.64	0.00	0.76	0.01	1.30

Matching Results for the Age Biased Hypothesis

Besides the assumption that technological change discriminates against low skilled employees, another often raised presumption is that older employees are more affected by technological change than younger ones. The first part of Table 9 contains the treatment effects on employees younger than 30 years. Looking at the t-values reveals that not a single treatment has a significant impact on the development of the share of young employees. It thus seems that the labor demand for younger employees is not affected by technological progress.

Table 9: Matching Results for the Share of Young and Older Employees (in %-Points)

Treatment	1993-1994 $\Delta^{m,0\|m}$	1993-1994 t-Value	1993-1995 $\Delta^{m,0\|m}$	1993-1995 t-Value	1993-1996 $\Delta^{m,0\|m}$	1993-1996 t-Value	1993-1997 $\Delta^{m,0\|m}$	1993-1997 t-Value
				Young Employees				
Product	-0.01	-0.97	-0.02	-1.33	-0.02	-1.07	-0.01	-0.32
Process	0.01	0.95	0.00	0.10	-0.01	-0.31	0.00	0.09
Orga	0.00	0.37	-0.01	-0.74	-0.02	-1.25	0.02	0.89
Product/Process	-0.01	-0.54	-0.02	-0.71	-0.01	-0.41	0.01	0.31
Product/Orga	0.01	0.40	0.00	-0.27	0.00	-0.24	0.00	0.09
Process/Orga	0.01	0.55	0.00	0.12	0.00	0.18	0.01	1.13
Prod/Proc/Orga	-0.01	-0.38	-0.01	-0.32	0.00	0.08	0.01	0.32
				Older Employees				
Product	**-0.02**	-1.86	-0.02	-1.39	**-0.03**	-1.73	**-0.04**	-2.29
Process	**-0.03**	-2.20	**-0.02**	-1.68	-0.02	-1.16	**-0.03**	-1.67
Orga	0.00	-0.12	0.01	0.39	0.01	1.04	**0.03**	1.74
Product/Process	0.01	0.27	0.02	0.87	0.01	0.74	0.01	0.34
Product/Orga	-0.02	-1.11	-0.01	-0.91	0.02	1.37	**0.02**	2.03
Process/Orga	-0.01	-0.90	**-0.05**	-2.43	0.00	-0.19	0.00	0.15
Prod/Proc/Orga	0.00	0.04	0.00	-0.24	0.00	-0.11	0.00	0.05

The picture changes, however, for employees older than 50 years. We find evidence that product innovations reduce the share of older employees in innovative establishments compared to non-innovative ones. The treatment effect amounts to 2%-points for the time period 1993-1994 and grows steadily to 4%-points for 1993-1997. For process innovations we find a quantitatively similar effect: Within 1993-1994 establishments which introduced a process innovation reduced the share of older employees by 3%-points compared to non-innovative establishments. The treatment effect of process innovations remains at this level for the following years.

As opposed to product and process innovations which both reduce the employment prospects of older employees, organizational changes have a positive impact, at least in the long run. When introduced together with product innovations, they are able to turn the negative impact of product innovations into a positive effect. On the other hand, we find empirical evidence that organizational changes even increase the negative impact of process innovations when introduced simultaneously.

Let us summarize: Technological change not only discriminates against low skilled and favors high skilled employees but is also age biased. Young employees are not affected by technological change whereas the age structure changes considerably at the expense of employees older than 50 years. The only exception is the introduction of organizational changes which have a positive impact on older employees.

Detailed Matching Results for Subgroups of Employees

In the previous sections we found that technological change increases the share of high skilled employees at the expense of low skilled ones. On the other hand, technological change has no impact on young and a mixed impact on older employees: Product innovations, process innovations and process innovations in combination with organizational changes have a negative impact while organizational changes introduced solely and together with product innovations increase the share of older employees.

However, if one takes into account that older employees are on average less skilled than younger ones, the results on the age biased technological change hypothesis may itself be biased: An increase in the relative demand for high skilled employees would automatically put older employees under pressure.[24] Thus, in the following we will take a closer look at the following four subgroups of employees: low and high skilled young employees and low and high skilled older employees.

Looking at the figures in Table 10 reveals the following: Young employees with a high qualification are not affected by technological change while the share of young and low skilled employees is reduced by organizational changes.Hence, the skill biased technological change hypothesis holds within the age group of young employees. The findings are more mixed regarding the group of older employees: Older high skilled employees benefit from technological change while the introduction of product innovations reduces the share of low skilled older employees. However, introducing product innovations together with organizational changes reverses this effect from negative to positive. Thus, within the group of older employees we only find weak empirical support for the skill biased technological change hypothesis.

[24] In Germany e.g. the share of individuals who finished secondary school amounted to 25% in 1965. In the mid-90s this share climbed to 37% (Engeln et al. (2003)). For a similar argument see also Rodriguez and Zavodny (2003).

Table 10: Matching Results for Subgroups of Employees (in %-Points)

Treatment	1994-1993 $\Delta^{m,0\|m}$	t-Value	1995-1993 $\Delta^{m,0\|m}$	t-Value	1996-1993 $\Delta^{m,0\|m}$	t-Value	1997-1993 $\Delta^{m,0\|m}$	t-Value
				Young & High Skilled Employees				
Product	0.00	-0.87	0.00	-1.59	0.00	-1.32	**0.00**	-1.63
Process	0.00	0.16	0.00	-0.01	0.00	-1.06	0.00	-0.76
Orga	0.00	1.08	0.00	-1.04	0.00	-0.73	0.00	-0.51
Product/Process	0.00	-1.04	0.00	-0.24	0.00	0.01	0.00	0.28
Product/Orga	0.00	-0.83	0.00	0.01	0.00	-0.29	0.00	0.15
Process/Orga	0.00	-0.09	0.00	-0.07	0.00	0.34	0.00	0.80
Prod/Proc/Orga	0.80	0.32	0.00	0.04	0.00	-0.12	0.00	0.35
				Young & Low Skilled Employees				
Product	0.01	0.45	0.00	0.08	0.00	0.02	0.01	0.33
Process	0.02	1.46	0.01	0.65	0.01	0.57	0.01	0.86
Orga	-0.01	-1.02	**-0.02**	-1.99	**-0.04**	-2.92	-0.01	-1.02
Product/Process	0.02	0.93	0.02	0.81	0.03	1.20	0.03	1.34
Product/Orga	0.00	-0.10	0.00	-0.28	-0.01	-0.99	0.00	-0.27
Process/Orga	0.01	1.37	0.00	-0.38	0.00	-0.10	0.00	0.29
Prod/Proc/Orga	0.00	0.05	0.00	-0.29	0.00	-0.04	0.00	-0.01
				Older & High Skilled Employees				
Product	0.00	0.03	0.00	0.18	0.00	-0.34	0.00	-0.12
Process	0.00	-0.31	0.00	0.30	0.00	1.19	0.00	0.62
Orga	0.00	-0.56	0.00	-0.79	0.00	0.42	0.00	-0.74
Product/Process	0.00	-0.73	0.00	0.54	0.00	-0.06	0.00	0.34
Product/Orga	0.00	1.02	**0.00**	2.29	**0.00**	1.60	0.00	0.55
Process/Orga	0.00	0.08	0.00	0.78	**0.00**	1.62	**0.01**	1.86
Prod/Proc/Orga	0.00	0.89	0.00	1.19	0.00	1.09	0.00	1.37
				Older & Low Skilled Employees				
Product	**-0.01**	-1.86	-0.01	-1.59	0.00	0.01	-0.01	-0.66
Process	-0.01	-1.59	-0.01	-0.90	0.00	-0.64	-0.01	-1.16
Orga	-0.01	-1.05	0.00	0.05	0.00	-0.32	0.00	0.27
Product/Process	0.00	1.19	0.00	0.68	0.01	1.23	0.01	0.81
Product/Orga	0.00	0.98	0.01	1.32	**0.01**	2.56	**0.01**	1.70
Process/Orga	-0.01	-1.20	0.00	-0.60	0.00	-0.28	0.00	-0.94
Prod/Proc/Orga	0.00	0.29	0.00	-0.77	0.00	-0.46	0.00	-0.59

A similar pattern arises for the age biased technological change hypothesis after controlling for the skill structure of employees. For high skilled employees we find that technological change leaves all age groups unaffected and thus reject the age biased technological change hypothesis. For the group of low skilled employees the findings are similar: Young low skilled employees are hurt by technological change. On the other hand, the share of older low skilled employees is reduced by product innovations while product innovations introduced together with organizational changes improve their employment prospects.

Thus, within the group of low skilled employees it seems that technological change even favors older employees.

Another interesting result can be derived if one looks at the magnitude of the age bias for different skill groups: The share of young low skilled employees is reduced by 4%-points compared to treatment effects for older low skilled employees which range between -1% and +1%-points. The corresponding figures for high skilled employees are: -0.3%-points for young employees and +0.2% and +1%-points for older employees. It thus seems that older low skilled employees are less affected *relative* to young low skilled employees while on the other hand older high skilled employees benefit less *relative* to young high skilled employees. Or stated equivalently: Although skill is a 'shield' against technological change, it has a stronger impact for young employees than for older employees.

To summarize: The age bias becomes weaker once we control for the skill structure, i.e. the conjecture that skill and age bias coexist as it was documented in Aubert, Caroli, and Roger (2004) is not supported by our empirical findings. It rather seems that the age bias of technological change is in fact driven the age composition of different skill groups. However, we also find empirical evidence for an asymmetry of skills, i.e. older employees benefit less from their education than younger ones.

5 Summary and Conclusions

In the following we will try to derive some stylized facts from the variety of estimates that we presented previously. Figure 5 contains the significant treatment effects separated for two skill and two age groups and also the appropriate aggregated effects. Organizational changes were found to be the major cause for a skill biased technological change. They reduce the share of low skilled and at the same time increase the share of high skilled employees. Technological change not only discriminates against low skilled but also against older employees as can be seen by the aggregate figures. A major exception is the introduction of organizational changes, solely and in combination with product innovations, which has a positive impact on older employees.

Figure 5: Treatment Effects for Different Skill and Age Groups

Skill \ Age	Young	Middle	Older	Σ_{Skill}
Low	**- 4%-points** Orga		**- 1%-point** Product **+ 1%-point** Prod/Orga	**- 4%-points** Orga
Medium				
High	**No** impact		**+ 1%-point** Proc/Orga **+ 0.2%-points** Proc/Orga	**+ 1%-point** Process **+ 1%-point** Orga **+ 2%-points** Proc/Orga
Σ_{Age}	**No** impact		**- 4%-points** — Product **- 3%-points** — Process **+3%-points** — Orga **+2%-points** — Prod/Orga **- 5%-points** — Proc/Orga	

Looking at the development within separate groups, however, reveals that the coexistence between skill and age bias becomes weaker once we control for age and skill simultaneously. 'Winners' of technological change are high skilled employees regardless of their age. Skill is therefore definitely a 'shield' against possible negative impacts of technological change. On the other hand, especially low skilled younger employees are the 'losers' of technological change whereas the impact on older low skilled employees is not clear cut. We also find evidence for an asymmetry of skills in the sense that younger high skilled employees benefit more *relative* to younger low skilled employees than older high skilled employees *relative* to older low skilled employees.

Katz and Murphy (1992) document in their study on changes in relative wages between 1963 and 1987 in the U.S. a pattern similar to our results. They found that older low skilled employees suffered on average less than young low skilled employees regarding their wage premium for experience. On the other hand, older high skilled employees benefited less than younger high skilled employees. An explanation for this asymmetry of skills might be skill obsolescence. Older high skilled employees received their education a long time ago which does not match the skill requirements of an accelerating technological change anymore. Given flexible labor markets like in the U.S. this qualification mismatch is reflected in a decreasing relative wage premium for older high skilled employees while in Germany what we observe are decreasing relative employment prospects.

The existence of such a skill obsolescence bears also policy relevant implications. In Germany the average retirement age decreased considerably in recent years with dramatic consequences for the social systems.[25] Accordingly, the labor force participation rate for employees between 55 and 64 in 2001 amounted to 36.8% compared to the OECD average of 48.4%.[26] Looking at the population development in Germany it is clear that the share of older employees will increase in the following years. Börsch-Supan (2002) considers different population projections and estimates that the old-age dependency ratio (the ratio of individuals over 60 years to individuals between 20 and 59) will increase from 38.6% in 1998 to values ranging from 76.7% to 88.4% in 2030.

Under such circumstances it is inevitable to encourage later retirement and therefore imperative to learn more about the determinants of labor demand for older employees. If one of the determinants is skill obsolescence, changes in terms of social security which aim at an increase of the retirement age will only have a limited effect. Instead, vocational training especially designed for older employees to cope with the technological progress might be in place.

[25] In 1973 e.g. the average retirement age was 62 years while in 2001 it fell below 60 years. See e.g. Arnds and Bonin (2002) who also discuss potential determinants of early retirement.

[26] Cf. OECD, Labor Force Statistics, 2002.

A A Multiple Treatment Matching Protocol

The most simple matching protocol for the estimation of $\Delta^{ml|m} = E_i[Y_i(m) - Y_i(l)|D_i = m]$ includes the following steps:[27]

1. Specifying and estimating a multinomial choice model

As a result one obtains $\hat{p}^0(X_i), \hat{p}^1(X_i), \ldots, \hat{p}^M(X_i)$. The appropriate model depends on the problem at hand. If e.g. the treatments are inherently ordered, an ordered probit model applies whereas if there is no ordering, multinomial logit or probit models should be estimated. If one is only interested in $\Delta^{ml|m}$, an alternative is to estimate directly the conditional probabilities $p(D_i = l|D_i = m, l)$ using a binomial logit or probit model.

In order to improve the matching quality one can restrict the sample to the common support. In this case for every combination of l and m with $l, m \in \{0, 1, 2, \ldots, M\}$ all observations in the treatment group $D_i = l$ are deleted where the probabilities $p^l(X_i)$ and $p^m(X_i)$ are larger than the smallest maximum and smaller than the largest minimum in the control group $D_i = m$. If one is merely interested in $\Delta^{ml|m}$ it is sufficient to restrict only this subgroup to the common support.[28]

2. Estimating the conditional expected outcome variables

a. Choose one individual out of the group of individuals which participated in treatment m and delete it from that pool. Calculate for this individual either $\{p(D_i = m|X_i), p(D_i = l|X_i)\}$ or directly $p(D_i = l|D_i = m, l)$.
b. Find one individual in the sub-sample of participants in treatment l that is as close as possible to the chosen treated individual in terms of his probabilities $\{p(D_i = m|X_i), p(D_i = l|X_i)\}$ or in terms of $p(D_i = l|D_i = m, l)$. In the first case a Mahalanobis distance can be calculated and additionally other covariates included. Do not delete this individual, i.e. conduct a matching with replacement. The reason is that unlike in the binomial case where we typically have more non-treated individuals than treated, in the multiple treatment model each group will act as a treated as well as a control group. Therefore requiring the number of comparisons to be larger than the number of treated individuals makes no sense.

[27] Cf. Lechner (2001, 2002, 2002b) and Gerfin and Lechner (2002).
[28] Cf. Gerfin and Lechner (2002), p. 868.

c. Repeat step a. and b. until no treated participant m is left.

d. The mean difference of the outcome-variables for both groups may serve as an estimate of $\Delta^{ml|m} = E_i[Y_i(m) - Y_i(l)|D_i = m]$.

e. Having obtained an estimate of $\Delta^{ml|m} = E_i[Y_i(m) - Y_i(l)|D_i = m]$ and noting that any other effect $\Delta^{mj|j} = E_i[Y_i(m) - Y_i(j)|D_i = j]$ with $j \neq m$ can be obtained analogously, all other treatment effects can be obtained using the following relations:

$$\begin{aligned}
\Delta^{ml|m,l} &= E_i[Y_i(m) - Y_i(l)|D_i = m, l] \\
&= p(D_i = m|D_i = m, l) E_i[Y_i(m) - Y_i(l)|D_i = m] \\
&\quad + (1 - p(D_i = m|D_i = m, l))\, E_i[Y_i(m) - Y_i(l)|D_i = l] \\
&= p(D_i = m|D_i = m, l)\Delta^{ml|m} + (1 - p(D_i = m|D_i = m, l))\, \Delta^{ml|l}
\end{aligned}$$

$$\Delta^{ml} = \sum_{j=0}^{M} \left(E_i[Y(m)|D_i = j] - E_i[Y(l)|D_i = j]\right) p(D_i = j) = \sum_{j=0}^{M} \Delta^{ml|j} p(D_i = j).$$

The variance of these estimators can be calculated according to the binomial case and additionally the same diagnostic checks can be conducted.

References

Abowd, J. M., and F. Kramarz (1999): “Econometric Analyses of Linked Employer-Employee Data,” *Labour Economics*, 6, 53–74.

Ahituv, A., and J. Zeira (2001): “Technical Progress and Early Retirement,” Working Paper.

Arnds, P., and H. Bonin (2002): “Frühverrentung in Deutschland: Ökonomische Anreize und institutionelle Strukturen,” IZA Discussion Paper No. 666, Institute for the Study of Labor.

Aubert, P., E. Caroli, and M. Roger (2004): “New Technologies, Workplace Organization and the Age Structure of the Workforce: Firm-Level Evidence,” Working Paper, Institut National de la Statistique et des Études Économiques.

Avolio, B. J., and D. A. Waldman (1994): “Variations in Cognitive, Perceptual and Psychomotor Abilities Across the Working Life Span: Examining the Effects of Race, Sex, Experience, Education and Occupational Type,” *Psychology and Aging*, 9(3), 430–442.

Bartel, A., and N. Sicherman (1993): “Technological Change and Retirement Decisions of Older Workers,” *Journal of Labor Economics*, 11, 162–183.

Beckmann, M. (2005): “Age-Biased Technological and Organizational Change: Firm-Level Evidence for West Germany,” Working Paper, University of Munich.

Bellmann, L. (1997): “The IAB Establishment Panel with an Exemplary Analysis of Employment Expectations,” IAB Topics 20, Institute for Employment Research.

Bender, S., A. Haas, and C. Klose (2000): “IAB Employment Subsample 1975-1995: Opportunities for Analysis Provided by the Anonymised Subsample,” IZA Discussion Paper No. 117, Institute for the Study of Labor.

Boockmann, B., and T. Zwick (2004): “Betriebliche Determinanten der Beschäftigung älterer Arbeitnehmer,” *Zeitschrift für Arbeitsmarktforschung*, 1, 53–63.

Börsch-Supan, A. (2002): “Labor Market Effects of Population Aging,” MEA Discussion Paper, Mannheim Research Institute for the Economics of Aging.

Chennels, L., and J. van Reenen (2002): “Technical Change and the Structure of Employment and Wages: A Survey of the Microeconometric Evidence,” in *Productivity,* *Inequality, and the Digital Economy*, ed. by N. Greenan, Y. L'Horty, and J. Mairesse, pp. 175–223. Cambridge.

Cohen, W. (1995): "Empirical Studies of Innovative Activity," in *Handbook of the Economics of Innovation and Technological Change*, ed. by P. Stoneman, pp. 182–264. Oxford and Cambridge.

Engeln, J., et al. (2003): "Indikatoren zur Ausbildung im Hochschulbereich," Studien zum Innovationssytem Deutschlands, Centre for European Economic Research.

Friedberg, L. (1999): "The Impact of Technological Change on Older Workers: Evidence from Data on Computers," Discussion Paper, University of California.

Gerfin, M., and M. Lechner (2002): "A Microeconometric Evaluation of the Active Labour Market Policy in Switzerland," *The Economic Journal*, 112, 855–893.

Griliches, Z. (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28, 1661–1707.

Heywood, J. S., L.-S. Ho, and X. Wei (1999): "The Determinants of Hiring Older Workers: Evidence from Hong Kong," *Industrial and Labor Relations Review*, 52(3), 444–459.

Hujer, R., M. Caliendo, and D. Radić (2001): "Estimating the Effects of Wage Subsidies on the Labour Demand in West Germany Using the IAB Establishment Panel," *ifo Studien*, 47, 163–197.

Imbens, G. W. (2000): "The Role of the Propensity Score in Estimating Dose-Response Functions," *Biometrika*, 87, 706–710.

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.

Lechner, M. (2000): "A Note on the Common Support Problem in Applied Evaluation Studies," Discussion Paper, University of St. Gallen.

——— (2001): "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption," in *Econometric Evaluation of Labour Market Policies*, ed. by M. Lechner, and F. Pfeiffer, pp. 43–58.

——— (2002a): "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies," *The Review of Economics and Statistics*, 84, 205–220.

——— (2002b): "Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods," *Journal of the Royal Statistical Society. Series A (General)*, 165, 59–82.

Neidert, A. (1998): "Neues Meldeverfahren zur Sozialversicherung ab 1999," *Deutsche Rentenversicherung*, 5, 315–330.

Plewis, I. (2002): "Modelling Impact Heterogeneity," *Journal of the Royal Statistical Society. Series A (General)*, 165, 31–38.

Radić, D. (2005): *Econometric Analysis of Innovation Activities*. Frankfurt et. al.

Rodriguez, D., and M. Zavodny (2003): "Changes in the Age and Education Profile of Displaced Workers," *Industrial and Labor Relations Review*, 56(3), 498–510.

Skirbekk, V. (2003): "Age and Individual Productivity: A Literature Survey," Working Paper, Max-Planck Institue for Demographic Research.

Spiezia, V., and M. Vivarelli (2002): "Innovation and Employment: A Critical Survey," in *Productivity, Inequality, and the Digital Economy*, ed. by N. Greenan, Y. L'Horty, and J. Mairesse, pp. 101–131. Cambridge.