The Returns to Pencil Use Revisited[§]

Alexandra Spitz-Oener[*]

February 2006

Abstract

The increased diffusion of computers is one of the fundamental changes at workplaces in recent decades. While the majority of workers now spend a substantial fraction of their working day with a computer, research on the wage effect of computer use came merely to a halt after the seminal paper by DiNardo and Pischke (1997) on the wage effect from pencil usage. This paper revitalizes the discussion by showing that the pencil effect disappeared in 1998/99, whereas the computer effect is still present. Computer users have experienced a pronounced shift towards analytical and interactive tasks for which they are rewarded in the workplace.

JEL–classification: J31, C13,

Keywords: Computer wage differential, pencil wage differentials, changing skill requirements

---

[§]**Acknowledgments:** Part of the paper was written within the research project "ICT, Workplace Organization, and Wages Across and Within Skill Groups" commissioned by the Landesstiftung Baden-Württemberg foundation when the author worked in the ICT research group at ZEW Mannheim. Helpful comments by David Autor, Joshua Angrist, Irene Bertschek, Sandy Black, Bernd Fitzenberger, Julia Häring, Thomas Hempell, and Susanne Prantl are gratefully acknowledged.

[*]Humboldt University Berlin; email: alexandra.spitz-oener@wiwi.hu-berlin.de.

I Introduction

The widespread usage of computers is one of the fundamental changes at workplaces in industrial countries in recent decades. Today there is a consensus that the implementation of computer technologies has changed skill requirements and hence the demand for labor. Because evidence suggests this change is skill-biased, research also attributes part of the changes in the wage structure to recent technological change.[1] Publications on wage differentials associated with on-the-job computer usage in the end-1990s, however, sounded the death knell for this body of literature, with the most powerful and probably the most cited criticisms of the "returns from computer use" literature coming from DiNardo and Pischke (1997) who show that there is also a considerable wage effect from the use of pencils (and other "white-collar" tools) in cross-section estimates. They convincingly argue that if we don't believe that pencils changed the wage structure, why should we believe that computers did. As a result, it has since been generally accepted that the positive correlations observed in cross-sections are spurious.

Using the same data set DiNardo and Pischke used, but a more recent cross-section, I show that the positive association between pencils and wages disappeared in 1999 while it is still present between computers and wages. This result moves the discussion in the literature forward to the question of *why* one was not able to identify an independent wage effect of computer usage up to the late 1990s. While a major aim of this paper is to revitalize the discussion in this body of the literature, I also offer a speculative explanation. The explanation relies on a task-based view of technological change. As put forward by Spitz-Oener (2006), Autor et al. (2003) and Goos and Manning (2003) the "traditional" skill-biased technological change hypothesis is probably an insufficient nuanced view of changes in the workplace associated with the recent diffusion of computer technologies.

[1] Comprehensive reviews of the skill-biased technological change literature can be found in Katz and Autor (1999), Chennells and van Reenen (1999) and Acemoglu (2002). For a recent discussion see Autor, Katz and Kearney (2005). Autor, Levy and Murnane (2003) and Spitz-Oener (2006) provide evidence on how skill requirements in the workplace have changed in recent decades.

The argument is that computers substitute certain tasks - those that are expressible in rules and thus programmable (termed routine tasks)- whereas they complement workers in performing nonroutine analytic and nonroutine interactive tasks. Computer users have experienced a pronounced shift towards analytical and interactive tasks between 1979 and 1998/99 for which they are rewarded in the workplace.

Non-random assignment of computer technology to workers is the central theme in this body of the literature. Previous empirical studies have addressed the issue in one form or another. All of these studies consider observed differences between computer users and non-users, though there is a large difference with respect to the number of available observables in data sets. The methods used to account for unobserved heterogeneity range from including proxies for individual ability into the regression specification to applying panel data methods.[2] The results are mixed, however, and depend largely on the underlying assumptions. The panel methods, for example, hinge crucially on the assumption of time-invariant unobserved heterogeneity. In the presence of changing returns on unobserved skills, differencing the data will not remove the wage effect of unobservables that might be correlated with computer use. As DiNardo and Pischke (1997) already emphasized, this factor may be a plausible explanation for the differing results in panel analyses for the UK and France - because the wage structure has widened in the UK since the early 1980s but not in France. An argument that also challenges the general credibility of panel estimates in this context is that identification in panel methods is through status changers, i.e. individuals that started or stopped using a computer. In the presence of downward rigid wages, panel methods lead to an underestimation of the computer wage effect. Yet–although these arguments make clear that panel methods are not immune to producing biased estimates–these findings have generally been interpreted as corroborating the notion that cross-section estimates are spurious.

[2]See Krueger (1993), who was the first to address the question of whether workers who use computers at work are paid more as a result of their computer skills or Entorf and Kramarz (1997) and Entorf, Gollac and Kramarz (1999), Bell (1996), who use panel data for France or the UK.

This paper is organized in 5 sections. In the next section, I present the data set and the variables of interest. Section III presents and discusses the estimation results. Section IV offers a speculative explanation for the empirical finding. Section V concludes.

II Data and Empirical Framework

The analysis is based on the "Qualification and Career Survey", which is a survey of employees carried out by the German Federal Institute for Vocational Training (Bundesinstitut für Berufsbildung, BIBB) and the Research Institute of the Federal Employment Service (Institut für Arbeitsmarkt- und Berufsforschung, IAB). For most of the analysis, I use the latest cross-section, launched in 1998/99. It covers more than 30,000 individuals (men and women).[3] DiNardo and Pischke (1997), in contrast, used cross-sections from 1979, 1985/86 and 1991/92.

In addition to the main variable of interest, computer use, three types of variables are considered in the analyses: individual characteristics, company characteristics and workplace characteristics. I include variables reflecting individual characteristics in order to account for the fact that employees differ systematically with respect to characteristics that affect both computer use and wages. For example, as more highly educated workers are more likely to use computers at work and earn higher wages, I control for the level of formal education of employees, work experience and tenure with the current employer. As wages of civil servants are determined in a process that differs from the process for wages of employees in private companies, a dummy variable for civil servants is also included in the regressions.

One drawback of most estimates on individual-level data is that they generally do not provide information on employers. Employer information may however be important if

[3] I restrict the sample to West German residents with German nationality: in other words, East German residents and non-German employees are excluded from the sample. Moreover, the sample does not include self-employed, employees with agricultural occupations and employees working in the agricultural sector. Persons younger than 18 or older than 65 are also excluded from the sample.

it determines systematic effects on wages and computer use. The data set used in this study allows me to take various company characteristics into account such as company size, industry affiliation, innovation strategy and company performance. This ability is a substantial improvement over other studies in this area. Based on previous empirical research, I expect, for example, that larger companies and innovative companies pay higher wages and that they are more likely to use computer intensively.

Another feature that distinguishes this data set from others is that it includes information on the task composition of occupations.[4] These tasks describe the occupational context in which computers are introduced. In addition, this information on occupational skill requirements allows me to further reduce unobserved heterogeneity.

The variables used in the estimates are constructed as follows (Summary statistics are in Table A in the Appendix B):

Hourly Wages: The survey contains information on monthly earnings, according to 18 categories. A midpoint is assigned for each category. These midpoints are then divided by the number of hours an individual usually spends at work.[5] Compared to other data sets that are usually used in comparable analyses such as the CPS for the U.S. or the IAB-S for Germany, this data set has the advantage that earnings of highly paid workers are not censored from above. In all estimates, the logarithm of wages is used as dependent variable. On average, employees in West Germany earned about 27 German Marks in 1998/99.

Computer equipment: The data set includes detailed information on the tools and machines used by the employees in the workplace. The "computer use" variable is a dummy that takes the value 1 if the employee uses a computer, terminal or electronic data processing device on the job. 57 percent of employees have used one or more of the

[4]For a detailed discussion of task measures and an analysis of how the task composition of occupations changed in West Germany since 1979 see Spitz-Oener (2006).

[5]Comparable procedures are often used in literature, for example, by DiNardo and Pischke (1997) and by Entorf and Kramarz (1997).

above devices on the job in West Germany in 1998/99.

Individual characteristics: I distinguish three levels of formal education attained by employees. Employees with a low level of education are those with no vocational training. Employees with medium levels of education have a vocational qualification either from an apprenticeship or they have graduated from a vocational college. Employees holding a degree from a university or a technical college are classified as having a high level of education. As shown in Table A, the majority of the survey participants, 70 percent, has a medium qualification level, whereas 17 percent are highly qualified and only 12 percent have a low education level.

The survey participants also indicate their first year of work. Based on these answers, I calculate the years of (potential) work experience (1999 - first year of work). In addition, employees indicate the year when they started to work for the current employer. This information is used to calculate tenure (1999 - first year with current employer).

The data set includes information about previous unemployment spells (dummy variable: "Have you ever been unemployed before?"), marital status, gender and whether survey participants were born in East Germany. It also contains information about whether an employee is working as a civil servant. In addition, I constructed a dummy variable indicating whether employees live in a city (place of residence is larger than 100,000 inhabitants).

Workplace Characteristics: The analyses by Autor et al. (2003) and Spitz-Oener (2006) document how computer have changed the content of work towards analytical and interactive activities and away from manual and cognitive routine activities. The data set allows me to capture the content of jobs by considering task levels, and therefore, it gives a description of the context in which computers are used. Survey participants are asked what kind of activities they perform at the workplace. Based on these activities five categories are constructed, which classify the occupational skill requirements: analytic tasks, interactive tasks, repetitive cognitive tasks, repetitive manual tasks and non-repetitive manual tasks. Table 1 shows the list of activities that employees were asked for in the questionnaire and how the activities are classified in the five task categories. On the individual-level i, the task measures ($Task_{ik}$) are defined as:

$$Task_{ik} = \frac{\text{number of activities in category } k \text{ performed by } i \text{ in } 1998/99}{\text{total number of activities in category } k \text{ in } 1998/99} * 100 \qquad (1)$$

where k=1: non-routine analytic tasks; k=2: non-routine interactive tasks; k=3: routine cognitive tasks; k=4: routine manual tasks; k=5: non-routine manual tasks. For example, if the analytical task category includes 4 activities and employee i performs 2 of them, the analytical task measure for employee i is 50.

The data set also contains information about the current occupation of the employees. Occupations are grouped according to the (2-digit) classification of occupational titles by the Federal Employment Bureau in 1999, leading to 78 occupational groups.

Company characteristics: Company size has been identified as an important component of wage determination in previous studies, finding that larger companies pay higher wages to employees with similar characteristics (see, for example, Brown and Medoff, 1989, Schmidt and Zimmermann, 1991). In addition, computer usage increases in company size (see Table B in Appendix B). Company size measured as the number of employees is captured by 7 size classes. Companies with one to four employees are classified to belong to the first size bracket and companies with more than 1,000 employees to the last one. Based on these size classes, 7 dummy variables are formed. Most of the survey participants, 28 percent, belong to companies with a size class from 10 up to 49 employees, followed by the size class from 100 up to 499 employees. Companies with more than 1000 employees are represented by 12 percent of the survey participants. About 20 percent of the interviewed employees belong to small companies with less than ten employees.

The data set also includes information about the performance of companies. The survey participants were asked whether the company was doing very good, good, rather bad or bad. For each of these categories, I constructed a dummy variable. Table A shows that 18 percent of employees report to work in companies that are doing very well and 65 percent work in companies that are doing well. 17 percent of employees work in companies that are either doing rather bad or bad.

Companies are classified according to 48 detailed industry codes. Based on these codes I group companies into three sectors: manufacturing, trade, and services.[6] The inclusion of these variables accounts for inter-industry wage differentials that are not already captured by the observed individual and company characteristics (see, for example, Krueger and Summers, 1987, Dickens and Katz, 1987, Gibbons and Katz, 1992 and Abowd, Kramarz and Margolis, 1999).

III Empirical Results

The most powerful and probably the most cited criticisms of the "returns from computer use" literature comes from DiNardo and Pischke (1997) who show that there is also a considerable wage effect of the use of pencils (and other "white-collar" tools) in their cross-section estimates. Table 2, column 1-3, reports parts of their Table III. Panel A reports the coefficients for computer use when separate regressions are performed for each workplace tool. Panel B shows the results when all tools are included together in the specification. Controlling for the different workplace tools attenuates the coefficient for computer use in each period, but the effect declines over time. Including the tools reduces the estimated coefficient on computer use by around 40 percent in 1979, by around 33 percent in 1985/86 and by 26 percent in 1991/92. In 1991/92, the coefficient of computer use has a magnitude of 13 percent even after controlling for all the tools.

I reproduced their estimates for the most recent wave of the data set (Table 2, last column). Relative to the result in Panel A, the estimated coefficient for computer use decreases by around 30 percent due to the inclusion of the dummies for the different workplace tools (Panel B). The coefficient of computer use still has a magnitude of around 15 percent. In contrast to the previous years, in which the different tools have always had a significant positive impact on wages, only the coefficients for the use of calculators and

[6] I also ran regressions that included more detailed industry dummies. The results that I report in Section III are robust to this change in specification.

working while sitting remain significantly positive in 1998/99. The coefficient for using a pencil is insignificant and using a telephone at work is now even significantly negatively correlated with wages. In what follows, I analyze this pattern in more detail.

Table 3, columns (1)-(5), display the estimation results of "standard" wage regressions. Unreported results show that the raw log wage differential for computer use in Germany is 0.275 (about 31 percent) in 1998/99. This figure is slightly lower than the raw log wage differential of 0.288 that DiNardo and Pischke (1997) report for Germany based on the 1991/92 cross-section of the BIBB/IAB data. Thus, in contrast to the period between 1979 and 1991/92, when the raw log wage differentials for computer use increased steadily (although at a declining pace), as shown in the paper by DiNardo and Pischke, this differential remained stable or even declined slightly in the 1990s.

Columns (1)-(5) show the results of specifications that are successively extended with additional controls. Including individual characteristics such as the level of formal education, work experience, tenure, gender, marital status or residence in a city reduces the coefficient of computer use by more than 30 percent (column 1). The coefficients of the controls have the expected sign, therefore I will not discuss them in detail. The results can, however, be found in Table C in Appendix B. In column (2), workplace characteristics are included in the specification. The higher the measure is for non-routine cognitive activities, both analytical and interactive, the higher the wages are. By contrast, wages decrease in the measure for non-routine manual activities. Including the workplace characteristics additionally reduces the computer coefficient by 30 percent (compared to the coefficient in column 1). In column (3), company characteristics such as company size and information about the innovation strategy of the company are included in order to control, for example, for company size effects in remuneration. Industry dummies are also included to account for cross-sectoral differences in computer usage and pay. In addition, dummies indicating company performance are included. Most interestingly, the inclusion of the company characteristics hardly affect the computer coefficient. By contrast, the returns to education decreased and the dummy for employees born in East Germany as well as the dummy for employees living in the city now are insignificant. Column (4) includes 10 dummies for West German states that control for cross-state differences in wage levels owing to, for example, differing economic conditions. These variables neither affect the computer coefficient nor the coefficients of the other controls. The specification in column (5) includes 77 two-digit occupation dummies. The occupation dummies have a large impact on the estimated computer wage differential. From column (1) to column (5), the computer coefficient drops by more than 70 percent, indicating that the largest part of the raw logarithm wage differential for computer users has been due to observable differences that would have resulted in employees earning different wages even in the absence of computer. The results indicate that observable workplace characteristics such as workplace tasks and occupational affiliation account for the largest proportion of the bias in the raw logarithm wage differential. Conditional on all the controls, however, the results still suggest that employees who use computer on the job earn around 8 percent higher wages.

Ideal control variables are only those that are attributes of the assignment to computer use and the earnings process but are unaffected by the treatment itself (for example, time-invariant individual characteristics such as gender and place of birth). Some of the controls used in this study, such as work experience or the incidence of previous periods of unemployment, might be affected by the treatment. Computer users are, for example, less likely to become unemployed. Therefore, the treatment effect estimated here does not capture the indirect effects of computer use on wages (for example, through productivity).

In column (6) I extend the specification by adding a control function that includes interactions between all covariates X (previously demeaned using sample averages) and the computer use dummy D_i. Hence, in contrast to the previous regressions, the coefficient on computer use is not constrained to be homogeneous on average conditional on the observable variables. Although the previous specifications already include a large number of controls, the remaining 8 percent wage markup for computer users may still be due to characteristics that are not observable in the data set at hand if these unobservables are positively correlated with both computer use and wages. The aim of control functions is to purge the specification from the remaining covariance between unobservables and computer use.[7] The results in column (6) show that including the control function in the specification reduces the estimated coefficient of computer use to 6 percent (=average treatment effect, ATE).[8] The interaction terms are jointly significant ($F_{(112,8666)} = 8.34, p = 0.0000$) and therefore provide evidence of the presence of heterogeneous effects. Table D in Appendix B shows the results for selected interaction terms. The empirical evidence suggests, for example, that computer users with a university degree benefit particularly in terms of wages. Differences in gender, work experience or workplace tasks, by contrast, do not have a significant effect on the gains from computer use.

The estimated average treatment effect for the treated (ATT) is 0.083 based on this regression, which is significantly different to the ATE.[9] On average, treatment effect heterogeneity seems to be important. The result that the ATT is higher than the ATE suggests that there has been selection into treatment based on expected returns. If computer non-user had started to use computer instead of those who actually did, they would have enjoyed a significantly lower benefit.

As heterogeneous treatment effect appear to be important, results might differ depend-

---

[7]See Appendix A for a detailed discussion of this approach.

[8]The standard errors in column (6) are estimated using a design-matrix bootstrap approach in order to account for the generated regressors in the specification. Because of the large number of dummy variables in the specification the resamples are chosen to be twice as large as the sample in order to guarantee that all coefficients can be estimated. The estimated covariance matrix then is doubled in accordance with the rate of convergence of the estimator.

[9]See equation 15 in Appendix A. The main focus of previous studies was on the ATE, thereby neglecting the potential of parameter heterogeneity. In a recent contribution, Dolton and Makepeace (2004) provide evidence for the importance of parameter heterogeneity. By allowing for variation in the parameter values, they find in their panel data set that some individuals benefit from computer usage in the UK. My approach of estimating the ATT is more general than the approach Dolton and Makepeace (2004) were taking.

ing on the particular assumptions made and assumptions on out of the sample predictions are particularly important in this respect. As a robustness test I therefore also did matching on the propensity score that imputes the expected non-treatment outcome for a computer user with observable characteristics X by the fitted value of a nonparametric regression in the sample of computer non-users with similar X. The details on the matching estimates can be found in Appendix C. The overall conclusions are, however, that the results are invariant to the particular estimation approach taken.

I now take up the DiNardo and Pischke (1997) idea and estimate regressions that include dummies indicating the use of various workplace tools instead of the computer dummy. A subsample of these results, those for pencils use, are shown in Table 4. Unreported results show that the first-order relationship between pencil use and wages is 5.7 percent (significant at the 1-percent level). Similar to the specifications in Table 3, I successively augment the specifications with additional controls. In contrast to the computer effect, the estimated pencil effect disappears as soon as controls for individual and workplace characteristics are included in the specification (column 2).[10] The variables that have had only attenuating effects on the coefficient of computer use, now result in the pencil effect disappearing, and it takes very view controls to do so. Column (2) shows the results of the regression specification including the minimum number of controls that are necessary to eliminate the wage effect of pencil use. It is interesting to notice that the coefficient in the comparable computer equation (Table 3, column 2) still is about 12 percent.

The result of a 8 percent wage markup for computer usage appears to be small, in particular in comparison with the 10-15 percent that have typically been found in cross-section analyzes in the 1990s. The size of the effect is rather similar to results previously found in panel analyzes and seems to be more reasonable (and believable) than previous cross-section results. One also has to keep in mind that the data set does not include unemployed persons. The 8 percent wage markup for computer users therefore represents

[10]The detailed results are shown in Table E in Appendix B.

a lower bound. Analyzes of Entorf, Gollac and Kramarz (1999), for example, suggest that people who do not use computers have a higher probability of becoming unemployed. Thus, the computer non-users observed in the data set are probably already a positive selection.

IV A speculative explanation

Why am I able to identify an independent wage effect of computer usage in 1998/99? One possible explanation is that changes in occupational skill requirements brought about by workplace computerization are important in this respect.[11]

As put forward by Spitz-Oener (2006), Autor et al. (2003) and Goos and Manning (2003) the "traditional" skill-biased technological change hypothesis is probably an insufficient nuanced view of changes in the workplace associated with the recent diffusion of computer technologies. The argument is that computers substitute certain tasks - those that are expressible in rules and thus programmable (termed routine tasks)- whereas they complement workers in performing nonroutine analytic and nonroutine interactive tasks.

In order to inform about the question of interest here, I present evidence on task changes for different subgroups: all employees, computer user/pencil non-user and computer non-user/pencil user. Table 5 shows the trends in skill inputs for the different subgroups. Panel A displays the overall trend, whereas Panel B and C show the figures for the subgroups that use computers but no pencils (Panel B) and vice versa (Panel C).

The first thing worthwhile noting is the strikingly similar distribution of task inputs

[11]DiNardo and Pischke (1997) noted that the computer coefficient increases strongly over time when all tools are entered in the regression together while some of the coefficients of the other tools tend to fall over time. They concluded that this might indicate that the role of computers in the workplace is changing over time. While this view is plausible as changes in computer technology itself (in particular the convergence of information and communication technologies) have certainly altered its role in the workplace over time. The explanation that I am going to outline here, however, relies on recent research investigating changes in occupational skill requirements owing to computerization.

for all groups in 1979. Cognitive and manual routine task inputs such as calculating, bookkeeping and operating machines had the highest values in all groups. Both computer user/pencil non-user and computer non-user/pencil user had above average values of non-routine analytical and nonroutine interactive inputs (such as doing research or advising customers) whereas they had below average values of non-routine manual task inputs (such as repairing or renovating houses) in 1979. An inspection of underlying occupations shows that computer user/pencil non-user and computer non-user/pencil user belonged to a very similar set of occupations in 1979. Examples are office workers, technicians and sales clerks.

The second thing worthwhile noting is that the high routine task measures for computer user/pencil non-user and, even more pronounced, for computer non-user/pencil user in 1979 - a year relatively close to the pre-computer era - leads in the task-framework to the prediction that they should adopt computer capital relatively rapidly as its price fell.

Table 6 shows the fraction of computer users in occupations with the highest proportion of computer users/pencils non-users and vice versa. As would have been predicted by the task framework, both types of occupation groups have above average rates of computer usage in 1979. In the successive years, the pace of computer diffusion diverged between the two groups of occupations however. While "computer user/pencils non-user"-occupations have experienced above average penetration rates, "computer non-user/pencils user"-occupations have experienced below average penetration rates.

As we know from Spitz-Oener (2006) and ALM (2003), routine tasks have been substituted by computers in recent decades while rapidly computerizing occupations have witnessed a pronounced shift towards analytical and interactive tasks. Accordingly, task changes have evolved differently for the two subgroups of interest to this study. While the overall trends look very similar, the size of shifts is strikingly different. For employees who use pencils but no computers (Panel C), the shifts away from cognitive and manual routine task inputs is even more pronounced (-1.7 and -1.1 percentage points annually) than for computer users/pencil non-users while the shifts towards analytical and interactive nonroutine tasks inputs is less (0.1 and 0.6 percentage points annually). In addition, this group of employees has witnessed an average annual increase in nonroutine task inputs of 1.2 percentage points (twice as large as the average for all employees). This differential shifts in task inputs mirrors itself in an occupational segregation. While computer user/pencil non-user and computer non-user/pencil user belonged to a large extent to the same occupations in 1979, this is no longer true in 1998/99. Examples of typical occupations of computer user/pencils non-user are still office workers, technicians and sales clerks - thus the same as for both groups in 1979. Computer non-user/pencils user, in contrast, appear more frequently in crafts occupations such as mechanics, electricians and masons or in health care such as nurses and masseurs.

Difference in tasks inputs by computer usage for the two subgroups are shown in Table 7. Both the levels and changes in routine cognitive and routine manual task inputs are strikingly similar in all subgroups. Hence, the substitutive relationship between routine tasks and computer technology manifests itself even "indirectly", meaning that computer non-users have also experienced a decline in the importance of cognitive and manual routine tasks. The main difference across groups results from the difference in analytical and interactive task inputs. This evidence confirms the complementary relationship between computers and nonroutine cognitive tasks in the sense that computer users in all types of occupations perform more analytical and interactive tasks. In contrast to the routine tasks, the complementary relationship is "direct" as the shift towards analytic and interactive activities has been more pronounced for computer users than computer non-users. This feature is particularly true for the analytical task category, in which computer non-users have barely experienced an increase between 1979 and 1998/99. It is not surprising that non-routine cognitive skills are more directly attached to specific persons as they largely depend on human capital acquisition or individual traits. Analytical skills such as mathematical, logical reasoning and problem-solving skills are often associated with specific subject-matter areas defined by the various disciplines, such as mathematics, biology or physics; and interactive skills such as interpersonal and organizational skills are individual traits. They refer not only to the ability of employees to communicate effectively with others through speech and writing, but also to the capability to work with others in teams and to building networks. Routine tasks, in contrast, are nonpersonal and can be switch to others or computers.

Table 8 shows OLS regressions of log hourly wages on task inputs. The estimated coefficients may be interpreted as task prices. Column (1) shows the results without additional controls. The coefficients for analytical, interactive and routine cognitive tasks are significantly positive, whereas the coefficients for manual tasks are significantly negative. The size of coefficients drop when a dummy for computer usage is included in the specification (column 2). The task prices for analytical and interactive tasks are, however, still significantly positive, while the coefficient for the routine cognitive task category now is insignificant. Comparing these results with those in column (3) and (4) shows that the inclusion of additional controls don't affect the estimates much for the computer coefficient and the coefficients of the analytical and interactive task measure. Whereas the coefficient for the routine cognitive task measure now is insignificant. In column (5) I additionally included interaction terms between the task measures and the computer use dummy in order to test for complementarities in wages. This reduces both the level and significance of the task measures and only the coefficients for analytical and interactive activities remain significantly positive related to wages. Most of the interaction terms are insignificant, however. The exception being the interaction between the analytical task measure and computer usage, which suggests a complementary relationship between the two variables. In addition, while the level effect of non-routine manual tasks now is insignificant, computer users tend to have lower wages the larger their measure in non-routine manual tasks.

V Conclusions

Using the same data set DiNardo and Pischke (1997) used, but a more recent wave, provides opportunity to revitalize the discussion on whether computer users earn higher wages than employees who do not use computers on the job. In contrast to their findings for earlier years, I show that the positive association between wages and pencils disappeared in 1999 while it is still present for computer usage – a result challenging the generally accepted notion that the positive correlation between wages and on-the-job computer usage is spurious.

I offer a speculative explanation on why I am now able to identify a positive relation between computer usage and wages independent from the pencil effect. The explanation relies on a task-based view of technological change and focuses on changes in skill requirements brought about by the introduction of computer technologies in the workplace. The argument is that computers substitute certain tasks - those that are expressible in rules and thus programmable (termed routine tasks)- whereas they complement workers in performing nonroutine analytic and nonroutine interactive tasks. Previous work already demonstrated the pronounced shifts away from cognitive and manual routine tasks towards analytical and interactive tasks owing to computerization (Spitz-Oener, 2006, Autor et al., 2003). This study shows that computer users and pencil users performed similar tasks in the workplace in 1979 but by 1999 their activities had diverged. In contrast to pencil users, computer users have experienced a pronounced shift towards analytical and interactive tasks between 1979 and 1998/99 for which they are rewarded in the workplace.

Table 1: Assignment of Activities

Classification	Tasks
non-routine analytic	researching, evaluating and planning,
	making plans, constructing, designing, sketching
	working out rules/regulations
	using and interpreting rules
non-routine interactive	negotiating, lobbying, coordinating, organizing
	teaching or training
	selling, buying, advising customers, advertising
	entertaining or presenting
	employing or managing personnel
routine cognitive	calculating, bookkeeping
	correcting of texts/data
	measuring of length/weight/temperature
routine manual	operating or controlling machines
	setting up machines
non-routine manual	repairing or renovation houses/apartments/machines/vehicles
	restoring art/monuments
	serving or accomodating

Table 2: OLS Regressions For The Effect Of Different Tools on Pay

Dependent Variable: Log(Hourly Wages)				
	Germany 1979	Germany 1985–86	Germany 1991–92	Germany 1998–99
A. Tools entered separately				
Computer	0.112	0.157	0.171	0.204
	(0.010)	(0.007)	(0.006)	(0.006)
B. Tools entered together				
Computer	0.066	0.105	0.126	0.146
	(0.010)	(0.008)	(0.007)	(0.007)
Calculator	0.017	0.053	0.044	0.051
	(0.008)	(0.007)	(0.007)	(0.006)
Telephone	0.072	0.043	0.045	-0.019
	(0.007)	(0.008)	(0.008)	(0.006)
Pen/Pencil	0.062	0.031	0.035	0.003
	(0.007)	(0.008)	(0.008)	(0.011)
Work while sitting	0.058	0.050	–	0.065
	(0.007)	(0.007)	–	(0.006)

Standard errors are in parentheses. Source of columns 1-3: DiNardo and Pischke (1997, p. 299, Table III, 2nd part-left side). Column 4: own regressions. Similar to the specification in DiNardo and Pischke education, experience, experience squared, dummies for part-time, city, female, married, female*married, and for civil servants are included in the regressions.

Table 3: OLS REGRESSIONS FOR THE EFFECT OF COMPUTER ON WAGES

Dependent Variable: Log(Hourly Wages)						
	(1)	(2)	(3)	(4)	(5)	(6)
Computer Use	0.182***	0.123***	0.122***	0.122***	0.076***	0.061***
	(0.005)	(0.006)	(0.008)	(0.008)	(0.010)	(0.024)
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Workplace Characteristics	No	Yes	Yes	Yes	Yes	Yes
Company Characteristics	No	No	Yes	Yes	Yes	Yes
77 occupation dummies	No	No	No	No	Yes	Yes
7 company size dummies	No	No	Yes	Yes	Yes	Yes
Dummies for manufacturing & trade	No	No	Yes	Yes	Yes	Yes
10 dummies for West German states	No	No	No	Yes	Yes	Yes
$Computer * [x - E(x)]$	No	No	No	No	No	Yes
R^2	0.348	0.340	0.399	0.401	0.443	0.455
Number of observations	18547	15266	8936	8936	8897	8897

Employees with low levels of education working in large companies in the services sector are the base category. Heteroscedasticity-consistent standard errors are in parentheses in column (1)-(5). Column (6): $Computer * [x - E(x)]$ means that the specification includes interaction terms between computer use and all the level variables X, where all the X had previously been demeaned by the sample averages. Bootstrapped standard errors using 100 resamples. ***, **, *-indicate significance at the 1, 5, 10 percent level.

Table 4: OLS Regressions for the Effect of Pencil Use on Wages

Dependent Variable: Log(Hourly Wages)					
	(1)	(2)	(3)	(4)	(5)
Pencil Usage	0.033***	0.016	0.016	0.016	0.005
	(0.010)	(0.011)	(0.013)	(0.013)	(0.013)
Individual Characteristics	Yes	Yes	Yes	Yes	Yes
Workplace Characteristics	No	Yes	Yes	Yes	Yes
Company Characteristics	No	No	Yes	Yes	Yes
77 occupation dummies	No	No	No	No	Yes
7 company size dummies	No	No	Yes	Yes	Yes
dummies for manufacturing & trade	No	No	Yes	Yes	Yes
10 dummies for West German states	No	No	No	Yes	Yes
R^2	0.300	0.323	0.383	0.385	0.437
Number of observations	15951	13986	8037	8037	8001

Employees with low levels of education working in large companies in the services sector are the base category. Heteroscedasticity-consistent standard errors are in parentheses. ***, **, *-indicate significance at the 1, 5, 10 percent level.

Table 5: Skill Inputs (Annualized Changes in Percentage Points)

	Non-routine Analytic	Non-routine Interactive	Routine Cognitive	Routine Manual	Non-routine Manual
		Panel A: Overall			
1979	4.42	8.47	36.86	30.88	14.19
1998/99	13.93	33.81	22.11	17.19	26.04
avg. annual chg.	0.5	1.3	-0.7	-0.7	0.6
		Panel B: Computer User/Pencil Non-User			
1979	8.70	9.63	37.91	24.55	9.16
1998/99	13.72	33.54	26.51	17.78	21.82
avg. annual chg.	0.3	1.2	-0.6	-0.3	0.6
		Panel C: Computer Non-User/Pencil User			
1979	6.45	12.22	57.19	42.25	6.76
1998/99	7.86	24.45	23.94	20.53	31.16
avg. annual chg.	0.1	0.6	-1.7	-1.1	1.2

The sample includes workers aged 18-65 who lived in West Germany and are German nationals.

Table 6: Descriptive Figures on Computer Diffusion

Spread of Computers,Terminals, Laptops, Electronic Data-Processing Devices			
	1979	1998/99	avg. annual chg.
Overall	6.12	55.84	2.62
Computer User/Pencil Non-User	10.61	78.57	3.57
Computer Non-User/Pencil User	7.74	44.34	1.92

The sample includes workers aged 18-65 who lived in West Germany and are German nationals. Average annual changes in percentage points.

Table 7: Skill Inputs 1979 and 1998/99 for Computer Users/Non-Users in Different Subgroups

	Non-routine Analytic	Non-routine Interactive	Routine Cognitive	Routine Manual	Non-routine Manual
Occupations w/ High Fraction of Computer Non-User/Pencil User					
1979					
Computer User	10.57	11.78	55.84	37.10	8.51
Computer Non-User	4.76	10.38	52.39	47.10	8.25
1998/99					
Computer User	13.30	46.06	16.78	10.85	34.45
Computer Non-User	6.90	28.10	15.56	11.50	35.40
Annualized changes between 1979 and 1998/99					
Computer User	0.14	1.80	-2.05	-1.38	1.37
Computer Non-User	0.11	0.93	-1.93	-1.87	1.42
Occupations w/ High Fraction of Computer User/Pencil Non-User					
1979					
Computer User	9.29	9.65	54.83	34.49	7.32
Computer Non-User	5.59	7.55	58.23	43.79	8.87
1998/99					
Computer User	18.27	38.46	12.31	8.36	17.50
Computer Non-User	6.23	32.55	11.47	9.00	29.83
Annualized changes between 1979 and 1998/99					
Computer User	0.47	1.52	-2.24	-1.38	0.53
Computer Non-User	0.03	1.32	-2.46	-1.83	1.10

The sample includes workers aged 18-65 who lived in West Germany and are German nationals. Average annual changes in percentage points.

Table 8: Task Prices in 1998/99

Dependent Variable: Log(Hourly Wages)	(1)	(2)	(3)	(4)	(5)
Computer Use		0.144***		0.122***	0.146***
		(0.006)		(0.008)	(0.014)
non-routine analytic	0.239***	0.195***	0.193***	0.120***	0.058*
	(0.013)	(0.013)	(0.016)	(0.016)	(0.035)
non-routine interactive	0.357***	0.298***	0.249***	0.155***	0.147***
	(0.011)	(0.012)	(0.014)	(0.014)	(0.025)
routine cognitive	0.103***	0.075***	0.022	0.004	-0.001
	(0.014)	(0.013)	(0.014)	(0.014)	(0.018)
routine manual	-0.065***	-0.008	-0.101***	-0.037**	-0.015
	(0.018)	(0.018)	(0.019)	(0.018)	(0.024)
non-routine manual	-0.241***	-0.187***	-0.160***	-0.095***	-0.027
	(0.012)	(0.012)	(0.014)	(0.014)	(0.023)
non-routine analytic*Computer Use					0.075**
					(0.039)
non-routine interactive*Computer Use					0.012
					(0.030)
routine cognitive*Computer Use					0.014
					(0.028)
routine manual*Computer Use					-0.032
					(0.037)
non-routine manual*Computer Use					-0.103***
					(0.028)
Individual Characteristics	No	No	Yes	Yes	Yes
Company Characteristics	No	No	Yes	Yes	Yes
7 company size dummies	No	No	Yes	Yes	Yes
dummies for manufacturing & trade	No	No	Yes	Yes	Yes
10 dummies for West German states	No	No	Yes	Yes	Yes
R^2	0.112	0.136	0.342	0.363	0.365
Number of observations	15679	15679	8897	8897	8897

References

Abowd, J., Kramarz, F. and Margolis, D. (1999). High Wage Workers and High Wage Firms, *Econometrica* **67**(2): 251–333.

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

Angrist, J., Imbens, G. and Rubin, D. (1996). Identification of Causal Effects Using Instrumental Variables, *Journal of the American Statistical Association* **91**(434): 444–455.

Autor, D., Katz, L. and Kearney, M. (2005). Trends in U.S. Wage Inequality: Re-Assessing the Revisionists, *mimeo*, MIT, September.

Autor, D., Levy, F. and Murnane, R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration, *Quarterly Journal of Economics* **118**(4): 1279–1333.

Bell, B. (1996). Skill-Biased Technical Change and Wages: Evidence from a Longitudinal Data Set, *Technical report*, Nuffield College.

Blundell, R., Dearden, L. and Sianesi, B. (2003). Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS, *Working Paper 03–20*, IFS.

Borghans, L. and ter Weel, B. (2004). The Diffusion of Computers and the Distribution of Wages, *Working paper*, ROA Maastricht University.

Brown, C. and Medoff, J. (1989). The Employer-Size Wage Effect, *Journal of Political Economy* **97**(5): 1027–1059.

Chennells, L. and van Reenen, J. (1999). Has Technology Hurt Less Skilled Workers?, *Working Paper Series W99/27*, Institute for Fiscal Studies.

Dehejia, R. and Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association* **94**: 1053–1062.

Dehejia, R. and Wahba, S. (2002). Propensity Score Matching for Nonexperimental Causal Studies, *Review of Economics and Statistics* **84**: 151–161.

Dickens, W. and Katz, L. (1987). Inter-Industry Wage Differences and Industry Characteristics, *in* K. Lang and J. Leonard (eds), *Unemployment and the Structure of the Labor Market*, Basil Blackwell, New York, pp. 48–89.

DiNardo, J. and Pischke, J. (1997). The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?, *Quarterly Journal of Economics* **112**: 291–303.

Dolton, P. and Makepeace, G. (2004). Computer Use and Earnings in Britain, *Economic Journal* **114**: C117–C129.

Entorf, H. and Kramarz, F. (1997). Does Unmeasured Ability Explain the Higher Wages of New Technology Workers?, *European Economic Review* **41**: 1489–1509.

Entorf, H., Gollac, M. and Kramarz, F. (1999). New Technologies, Wages, and Worker Selection, *Journal of Labor Economics* **17**(3): 464–491.

Gibbons, R. and Katz, L. (1992). Does Unmeasured Ability Explain Inter-Industry Differentials, *Review of Economic Studies* **59**(3): 515–535.

Goos, M. and Manning, A. (2003). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain, *Technical Report*, Center for Economic Performance, LSE.

Heckman, J. and Robb, R. (1985). Alternative Methods for Evaluating the Impact of Interventions, *in* J. Heckman and B. Singer (eds), *Longitudinal Analysis of Labor Market Data*, Cambridge University Press, Cambridge, pp. 156–245.

Heckman, J., Ichimura, H. and Todd, P. (1997). Matching as a Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies* **64**(4): 605–654.

Heckman, J., Ichimura, H. and Todd, P. (1998a). Matching as a Econometric Evaluation Estimator, *Review of Economic Studies* **65**(2): 261–294.

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1998b). Characterizing Selection Bias Using Experimental Data, *Econometrica* **66**(5): 1017–1098.

Katz, L. and Autor, D. (1999). Changes in the Wage Structure and Earnings Inequality, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Elsevier Science, Amsterdam, pp. 1463–1555.

Krueger, A. (1993). How Computer have Changed the Wage Structure: Evidence from Microdata, 1984-1989, *Quarterly Journal of Economics* **108**(1): 33–60.

Krueger, A. and Summers, L. (1987). Reflections on the Inter-Industry Wage Structure, *in* K. Lang and J. Leonard (eds), *Unemployment and the Structure of the Labor Market*, Basil Blackwell, New York, pp. 17–47.

LaLonde, R. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data, *American Economic Review* **76**(4): 604–620.

Rosenbaum, P. and Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* **70**(1): 41–55.

Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Educational Psycholgy* **66**(5): 688–701.

Schmidt, C. and Zimmermann, K. (1991). Work Characteristics, Firm Size and Wages, *Review of Economic Studies* **73**(4): 705–710.

Smith, J. and Todd, P. (2004). Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics* **125**(1-2): 305–353.

Spitz-Oener, A. (2006). Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure, *forthcoming: Journal of Labor Economics.*

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Massachusetts.

Appendix A

This study focuses on the problem of estimating the wage impact of computer usage. To state this more formally, I designate wages Y. Wages are assumed to be determined by a set of exogenous variables X, and by a dummy variable D, such that $D_i = 1$ indicates that employee i uses a computer on the job, and $D_i = 0$ that employee i does not. Assuming, for the ease of illustration, that the decision to use computer was made in period s, so that, in each period t,

$$Y_{it} = \begin{cases} X_t\beta + D_i\alpha_i + U_{it} & \text{if} \quad t > s \\ X_t\beta + U_{it} & \text{if} \quad t \leq s \end{cases} \tag{2}$$

where α_i measures the heterogeneous impact of computer usage on wages. If α_i is constant across individuals, the impact is homogeneous. The set of parameters β defines the relationship between the observable variables X and the outcome Y, and U_{it} is an error term of mean zero and is assumed to be uncorrelated with X.

Most of the discussions on the wage effects of computer use evolve because the group of computer users is most probably not random. This situation arises because the decision of individual i or, more likely, the decision of an employer e to choose individual i to use computer is based on personal characteristics P, workplace characteristics W and employer characteristics E that may also affect wages Y. If this is the case, and the set of observable variables X is only a subset of characteristics P, W or E, then the error term U and the computer use variable D are very likely to be correlated. Hence, as a consequence of such a non-random assignment, computer use ($D_i = 1$) would be correlated with the error term U_{it} in the wage equation, which results in the standard econometric approach that regresses Y on X and D, producing biased results.

The decision rule may be described in a very stylized form as a latent single index function $IN_e(i)$, which specifies the gain in profitability of employer e if employee i (working at e) uses a computer. $IN_e(i)$ thus represents the difference between the productivity increases of employer e owing to the implementation of computer and the higher wages employer e has to pay to employee i for the computer usage. $IN_e(i)$ is a function of observable variables X and unobservable variables V: $IN_e(i) = X\gamma + V$. The entire set of characteristics I, which are relevant to the decision about computer usage and wages, are then the combination of variables X and V, $I = (X, V)$. Employer e decides that employee i uses computer on the job if $IN_e(i) > 0$, in which case $(D_i = 1)$ is observed. A simple behavioral model is one in which employers select employees into computer usage based on a comparison of the present value of expected profits with and without employee i using computer: $IN_e(i) = PV_e(1) - PV_e(0)$, where $PV(1)$ is the present value of expected profits if employee i uses computer on the job and $PV(0)$ without computer usage.[12]

The empirical strategy in this study is to use methods most often applied in the active labor market program evaluation literature. In the tradition of this literature, I interpret computer usage as a treatment D_i, that is, estimating the coefficient of computer usage translates into a problem of estimating an average treatment effect (ATE). The ATE is usually estimated using a counterfactual framework (Rubin, 1974), where each individual i has an outcome with $(D_i = 1)$ and without $(D_i = 0)$ treatment (conventionally denoted as Y_{1i} and Y_{0i} respectively). Since both states are never observed for one person simultaneously, different methods have been developed in literature to overcome this "omitted variable" problem depending on the specific circumstances of the question of interest.

The observed outcome may be expressed as:

$$Y_i = Y_{0i}(1 - D_i) + Y_{1i}D_i = Y_{0i} + D_i(Y_{1i} - Y_{0i}). \tag{3}$$

This body of literature distinguishes between the average treatment effect (ATE), $E(Y_{1i} - Y_{0i})$, that is, the expected effect of treatment on a person drawn randomly from the population, and the average treatment effect on the treated (ATT), $E(Y_{1i} - Y_{0i}|D_i = 1)$, that is, the expected effect of treatment of those who actually have been treated. Since $E(Y_{1i}|D_i = 1)$ is directly identifiable from the sample, the problem of estimating the ATT

[12]For a more comprehensive model of computer assignment see, for example, Borghans and ter Weel (2004).

is equivalent to the problem of estimating the counterfactual average, $E(Y_{0i}|D_i = 1)$. Thus, in order to be able to estimate the average non-treatment outcome for treated employees based on the outcomes for non-treated, identifying assumptions are needed. Estimating the ATE, on the other hand, involves estimating both counterfactual averages, $E(Y_{0i}|D_i = 1)$ for the non-treated workers in the sample and $E(Y_{1i}|D_i = 0)$ for the treated workers in the sample. In the case of computer usage, the effect on the subpopulation of treated employees (ATT) is probably of more interest than the effect on the overall population (ATE). As already discussed earlier in this paper, one of the reasons is that computers are of value only in certain jobs, for example, for computer scientists but not for artistic painters. The overall focus of the analysis will therefore be on the ATT.

The observed difference in outcomes between those treated and those untreated equals $E(Y_{1i} - Y_{0i}|D_i = 1)$ plus a bias term:

$$\begin{aligned} E(Y|D_i = 1) - E(Y|D_i = 0) &= E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0) && (4)\\ &= E(Y_{1i} - Y_{0i}|D_i = 1) && (5)\\ &+ \underbrace{[E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)]}_{\text{bias term}}. \end{aligned}$$

If computer use were randomly assigned to employees, that is, if the treatment and control group had the same distribution of both observed and unobserved characteristics, this would be an easy task: the ATE and ATT would be the difference in sample average of those using computer and those not using computer at the workplace.[13] However, this scenario is very unlikely, in particular as recent empirical studies document that computer users differ systematically from computer non-users, for example, with respect to their education. Table C1 in Appendix C shows the non-randomness of computer usage based

[13] An additional assumption is that there are no general equilibrium effects in the sense that computer use by one employee does not affect wages of other employees (so-called stable unit treatment value assumption: SUTVA; see, for example, Angrist, Imbens and Rubin, 1996). This is synonymous to stating that there are no macroeconomic effects. The analysis in this study is based on a random sample, which implies SUTVA.

on the data set at hand. The major differences are in the distribution of formal education between computer users and computer non-users. In contrast to computer being randomly assigned, employers select employees with specific favorable traits for treatment. The decision about who uses computer on the job is related to the question of how much an employer's profitability increases from computer use and, hence, to the outcomes Y_{1i} and Y_{0i}.

Since data from a social experiment or instrumental variables that allow purging of the regression from the presence of a possible omitted variable bias are not available for this research project, my strategy is based on the assumption of "strict ignorability" of treatment conditional on a set of controls (Rosenbaum and Rubin, 1983): $(Y_{1i}, Y_{0i}) \perp\!\!\!\perp D_i|X$. That is, conditional on X, D_i and (Y_{1i}, Y_{0i}) are independent, where X denotes a vector of observable covariates, D_i denotes the treatment (in our case $D_i = 1$: computer use, $D_i = 0$: no computer use) and Y_{1i} is the logarithm of hourly wages for computer users and Y_{0i} for computer non-users. This assumption implies that $E(Y_{0i}|X, D_i) = E(Y_{0i}|X)$ and $E(Y_{1i}|X, D_i) = E(Y_{1i}|X)$, referred to as the conditional mean independence assumption (CMA) in literature. That is, conditional on X, the non-treatment outcome for the treated and non-treated are comparable in expectation. For the estimators used in this study it suffices that this weaker assumption holds.

The intuition behind this assumption is that systematic differences in wages between computer users and computer non-users with the same values for the covariates X are attributable to the use of computer. Since the moments of the distribution of Y_{1i} for the treated can be estimated directly, it suffices for the assumption $E(Y_{0i}) \perp\!\!\!\perp D_i|X$ to hold in order to be able to estimate the ATT.

The present data set is particularly suited for this strategy because it contains a large number of observables that have been identified as important determinants of the computer use-wage relationship in earlier studies. In contrast to other individual-level data sets that are usually limited to information about employees, it includes information on employers. This information is important since employers may differ in a way that systematically influences both individual wages and computer use. It is, for example, very likely that a computer user would be employed at another employer in the counterfactual situation. The data set allows me to take various company characteristics into account, such as company size, industry affiliation, innovation strategy and company performance. In addition, the data set is a sample of West German employees that includes both computer non-users, the potential comparison group, and computer users, the treatment group. They answered identical questionnaires. Wages and all other characteristics are therefore measured in the same way for both groups. As the data set also includes information on the regional affiliation computer users and computer non-users operating in a common economic environment can be compared.

In spite of this merit, one must keep in mind that the CMA is a strong assumption and that, stricly speaking, the estimation strategy accounts solely for selection on observables. Differences between computer users and computer non-users in the distribution of unobserved attributes such as ability are not considered. Heckman, Ichimura, Smith and Todd (1998b) decompose the bias term of equation 4 into three components:

$$\text{bias term} = [E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)] = B_1 + B_2 + B_3 \qquad (6)$$

The first two components, B_1 and B_2, arise due to differences in the distribution of observed characteristics X whereas the last component, B_3, refers to the "classical" selection bias resulting from selection on unobservable attributes. B_1 corresponds to the bias due to differences in the support of regressors between computer users and computer non-users. B_2 corresponds to the bias owing to differences in the shape of distributions of regressors in the two groups in the region of common support.

I use two methods to overcome the bias due to B_2, regression based matching and matching on the propensity score. The two approaches differ in how they deal with the bias due to B_1. None of them is able to resolve problems due to the presence of B_3. However, Heckman, Ichimura and Todd (1997) and Heckman, Ichimura, Smith and Todd (1998b) find that the selection bias part, B_3, is the smallest of the three components of bias in their application. How important the different bias components are depends, of course, on the application. However, the richness of the data set in the present study largely reduces the scope for a potential bias owing to unobservables.

Both approaches have already been widely discussed in literature (see, for example, Heckman and Robb, 1985, Heckman, Ichimura and Todd, 1997 and 1998a, or Wooldridge, 2002, Chapter 18). I will review the basic concepts in order to demonstrate their usefulness for the estimation problem at hand.

Regression Based Matching: The aim of regression based matching (RBM) is to purge the outcome equation from potential correlation between the error term U_{it} and D_i (see equation 2). The basic logic of the RBM is analogous to the proxy solution to the omitted variable problem.[14]

To illustrate this approach it is useful to decompose the counterfactual outcomes for computer user Y_{1i} and computer non-user Y_{0i} into their means and a stochastic part:[15]

$$Y_{0i} = \mu_0 + U_{0i} \qquad E(U_{0i}) = 0 \tag{7}$$

$$Y_{1i} = \mu_1 + U_{1i} \qquad E(U_{1i}) = 0 \tag{8}$$

where $\mu_j = E(Y_{ji}), \;\; j = 0, 1$. As in the previous part of the paper, i refers to individuals and j to the treatment status. For the case of linear regressions $\mu_j = X\beta_j$. Inserting these into equation (3) of observed outcomes gives:

$$Y_i = \mu_0 + (\mu_1 - \mu_0)D_i + U_{0i} + (U_{1i} - U_{0i})D_i \tag{9}$$

Under the CMA assumption and assuming that $E(U_{1i}|X) = E(U_{0i}|X)$ it follows that standard regression methods can be used to estimate ATE. This is because equation (9) may be rewritten as:

$$\begin{aligned} E(Y_i|D_i, X) &= \mu_0 + (\mu_1 - \mu_0)D_i + E(U_{0i}|X) + \underbrace{E(U_{1i} - U_{0i}|X)}_{=0} D_i && (10) \\ &= \mu_0 + \alpha D_i + g_0(X) && (11) \end{aligned}$$

[14] The omitted variable bias can be mitigated, or even eliminated, if a proxy variable is available for an unobserved variable q. The formal requirements for a proxy variable can be found, for example, in Wooldridge (2002, p.63ff.). The intuition is that, conditional on the proxy variable, the unobservable variable q is uncorrelated to each of the observable variables X in the specification.

[15] The structure of illustration follows Wooldridge (2002), chapter 18.

where $\alpha = \mu_1 - \mu_0 = E(Y_{1i} - Y_{0i}|X) = ATE = ATT$ and $g_0(X) = E(U_{0i}|X)$. In addition, $E(U_{0i}|X)$ might be expressed as a function of some vector function $h_0(X)$, such as $E(U_{0i}|X) = \eta_0 + h_0(X)\beta_0$. Then, equation (10) translates into

$$E(Y_i|D_i, X) = \gamma_0 + \alpha D_i + h_0(X)\beta_0 \tag{12}$$

where $\gamma_0 = \mu_0 + \eta_0$. $h_0(X)\beta_0$ is an example of a control function and its purpose is to provide an approximation for $E(U_{0i}|X)$. When inserted into equation (10) and therefore implicitly subtracted from U_{0i}, it purges the equation from the covariance between D_i and U_{0i}. The idea is that the purged disturbance term $[U_{0i} - h_0(X)\beta_0]$ is orthogonal to all variables on the right-hand side of the equation. If the assumptions hold, the control function accounts for a possible self-selection bias.[16]

The previous specification rests upon the assumption that there are on average no person-specific gains conditional on X, $E(U_{1i} - U_{0i}|X) = 0$. Relaxing this assumption implies that ATE and ATT are no longer identical. However, a regression specification can still be used:

$$E(Y_i|D_i, X) = \mu_0 + \alpha D_i + g_0(X) + D_i[g_1(X) - g_0(X)] \tag{13}$$

where $\alpha = ATE$, $g_0(X) \equiv E(U_{0i}|X)$ and $g_1(X) \equiv E(U_{1i}|X)$. $g_0(X)$ and $g_1(x)$ are both operationalized by replacing them with a parametric function of X. The intuition is analogous to the approach in equation (12).

Assuming that these functions are both linear in X, equation (13) can be written as

$$E(Y_i|D_i, X) = \gamma_0 + \alpha D_i + X\beta_0 + D_i[X - E(X)]\delta \tag{14}$$

where β_0 and δ are unknown parameters (the intermediate steps are shown at the end of this Appendix). As can be seen, the control function then involves both the level

[16] This is a special case of a control function that does not rely on instruments. For a discussion on the control function approach involving instruments refer, for example, to Heckman and Robb (1985) or Blundell, Dearden and Sianesi (2003). In contrast to the instrument approach that rests upon the exclusion restriction, identification in the approach used in this study is achieved through a functional form restriction.

effect of X and the interactions of X (which have been previously demeaned by $E(X)$, approximated with the sample averages) with the treatment indicator. Based on the $\hat{ATE}$, the ATT is estimated using[17]

$$\hat{ATT} = \hat{\alpha} + (\sum_{i=1}^{N} D_i)^{-1}(\sum_{i=1}^{N} D_i(X - \bar{X})\hat{\delta}). \tag{15}$$

I take account of the fact that the demeaned variables are generated regressors by applying a bootstrap method to construct the standard errors of the estimated coefficients. All bootstrap results are based on 100 resamples.

Propensity Score Matching: The CMA assumption is also the basis of the matching technique. The traditional matching approach estimates the expected non-treatment outcome for a computer user i with observable characteristics X by the fitted value of a nonparametric regression in the sample of computer non-users with identical X. This can be represented by the following formula:

$$ATT = \frac{1}{N_1} \sum_{i \epsilon D_i=1} \{Y_{i1} - \sum_{i \epsilon D_i=0} w_{N_0} Y_{i0}\}, \tag{16}$$

where N_1 is the number of computer users and N_0 is the number of computer non-users. w_{N_0} is a weight function that weights the observations of a computer non-user according to its similarity with respect to X to computer user i. The literature discusses different matching estimators, which differ with respect to the weight function. I will discuss the estimators used in this study at the end of this section. Beforehand, I shortly refer to the so-called "curse-of-dimensionality" problem that arises when the vector of observable characteristics X is high-dimensional (as it is the case in this study). As the dimension of the data increases and, hence, the information content increases, the complexity of the estimation problem increases exponentially. The "curse-of-dimensionality" arises because in high-dimensional settings, the data requirements needed to find counterfactuals that are similar along every dimension of X increase exponentially in the dimension of X.

[17] A method to estimate the ATT directly would be to demean the X using the averages of the "treated" sample instead of the averages of the "whole" sample.

Rosenbaum and Rubin (1983) introduced the propensity score as a means to overcome the dimensionality problem that arises when the set of conditioning variables X is large. They showed that when outcomes Y_0 are independent of the treatment conditional on variables X (CMA assumption), then outcomes Y_0 are also independent of the treatment conditional on the propensity score $p(X)$, defined as $P(D_i = 1|X)$, with $0 < P(D_i = 1|X) < 1$. As can be seen from this formula, the propensity score is the probability of treatment given the covariates X. In conditional independence notation, this is $(Y_{1i}, Y_{0i}) \perp\!\!\!\perp D_i|p_i(X)$. This property allows it to use the one-dimensional propensity score $P(D_i = 1|X)$ instead of the high-dimensional vector X in the matching approach. $P(D_i = 1|X)$ instead of X now determines the similarity of treated and non-treated employees. Dehejia and Wahba (1999, 2002), among others, have drawn attention to this class of estimators because their analyses suggest that propensity score matching is a powerful tool to resolve the selection bias problem inherent in non-experimental data.[18]

To implement this estimator, one first needs an estimate for the propensity score. I choose a flexible probit model that includes various covariates (described below), their quadratics and interactions. Logit models as proposed by Rosenbaum and Rubin (1983) or non-parametric models may, however, also be used (Powell, 1994, Heckman, Ichimura and Todd, 1997). I take account of the sampling variability in the estimated propensity scores by applying a bootstrap method to estimate the standard errors of the estimated treatment effects. The bootstrap results are based on 100 resamples.

Matching estimators differ with respect to the weights they attach to members of the non-treated group. The traditional pairwise matching method, nearest neighbor matching, weights the nearest non-treated employee with 1 and all others with 0, that is, only the closest member of the comparison group is used as the comparison level for the treated

[18]Others also investigated the properties of matching methods in different settings (see, for example, Heckman, Ichimura and Todd, 1997, 1998a). However, Dehejia and Wahba (1999, 2002) explicitly make the case for matching on the propensity score in order to overcome LaLonde's criticism of non-experimental estimators (LaLonde, 1986). Criticism of Dehejia and Wahba (1999, 2002) comes from Smith and Todd (2004).

employee i. In order to guarantee that the "closest" neighbor is not too far away in terms of the propensity score, one usually sets a caliper. In the nearest neighbor matching estimations in this study the caliper is set to 0.001.

Extensions to this method are, for example, kernel or local linear matching estimators (Heckman, Ichimura and Todd, 1997 and Heckman, Ichimura, Smith and Todd, 1998), which have the advantage of reducing the asymptotic mean squared error. In this study, I present results using an Epanechnikov kernel.

Intermediate steps between equation 13 and 14:

$$
\begin{aligned}
E(Y_i|D_i,X) &= \mu_0 + \alpha D_i + g_0(X) + D_i[g_1(X) - g_0(X)] \\
g_0(X) &= \eta_0 + X\beta_0 \quad Eg_0(X) = 0 \\
g_1(X) &= \eta_1 + X\beta_1 \quad Eg_1(X) = 0 \\
E(Y_i|D_i,X) &= \mu_0 + \alpha D_i + \eta_0 + X\beta_0 + D_i[\eta_1 + X\beta_1 - \eta_0 - X\beta_0] \\
&= \mu_0 + \eta_0 + \alpha D_i + X\beta_0 + D_i[\eta_1 - \eta_0 + X(\beta_1 - \beta_0)] \\
&= \underbrace{\mu_0 + \eta_0}_{\gamma_0} + \alpha D_i + X\beta_0 + D_i[\frac{\eta_1 - \eta_0}{\beta_1 - \beta_0} + X]\underbrace{(\beta_1 - \beta_0)}_{\delta}
\end{aligned}
$$

To be shown: $\frac{\eta_1 - \eta_0}{\beta_1 - \beta_0} = -E(X)$

$$
\begin{aligned}
\frac{\eta_1 - \eta_0}{\beta_1 - \beta_0} &= \frac{1}{\beta_1 - \beta_0}(g_1(X) - X\beta_1 - g_0(X) + X\beta_0) \\
\eta_1 - \eta_0 &= g_1(X) - g_0(X) + X(\beta_0 - \beta_1) \\
X &= \frac{1}{\beta_1 - \beta_0}(g_1(X) - g_0(X) - \eta_1 + \eta_0) \\
E(X) &= \frac{1}{\beta_1 - \beta_0}(\underbrace{Eg_1(X)}_{=0} - \underbrace{Eg_0(X)}_{=0} - \eta_1 + \eta_0) \\
&= -\frac{\eta_1 - \eta_0}{\beta_1 - \beta_0}
\end{aligned}
$$

Appendix B

Table A: Summary Statistics

	Mean	Std. Deviation	Min.	Max.	Observations
computer	0.57	0.50	0	1	21816
Pencil	0.92	0.27	0	1	18775
(hourly) wages (in DM)	27.19	11.82	3.13	98.68	18561
Qualification					
high education level	0.17	0.37	0	1	21816
medium education level	0.71	0.46	0	1	21816
low education level	0.12	0.33	0	1	21816
experience	20.76	11.58	0	47	21816
tenure	11.75	9.84	0	47	21816
Workplace Characteristics:					
analytic task measure	14.01	23.80	0	100	18041
interactive task measure	30.26	28.18	0	100	21754
repetitive cognitive task measure	21.73	41.24	0	100	21813
repetitive manual task measure	17.43	30.83	0	100	21768
non-repetitive manual task measure	24.32	24.99	0	50	12319
Company Characteristics					
product innovation	0.37	0.48	0	1	20802
process innovation	0.51	0.50	0	1	20857
very good company performance	0.18	0.39	0	1	14450
good company performance	0.65	0.48	0	1	14450
rather bad company performance	0.14	0.35	0	1	14450
bad company performance	0.03	0.17	0	1	14450
Other Controls					
ever unemployed	0.30	0.46	0	1	22545
married	0.69	0.46	0	1	22677
civil servants	0.11	0.31	0	1	22677
born in East Germany	0.04	0.19	0	1	22677
woman	0.44	0.50	0	1	22677
lives in city	0.38	0.48	0	1	22677

Table B: Computer Usage by Company Size Distribution

No. of employees	Freq.	Percent	computer usage (in percent)
1 to 4	1,707	7.8	42.9
5 to 9	2,716	12.3	45.3
10 to 49	6,186	28.1	50.1
50 to 99	2,790	12.7	57.1
100 to 499	4,457	20.2	63.8
500 to 999	1,447	6.6	68.3
1000 and more	2,721	12.4	72.1
Total	22,024		

Table C: OLS Regressions for the Effect of Computer Use on Wages

Dependent Variable: Log(Hourly Wages)						
	(1)	(2)	(3)	(4)	(5)	(6)
computer	0.182***	0.123***	0.122***	0.122***	0.076***	0.061***
	(0.005)	(0.006)	(0.008)	(0.008)	(0.010)	(0.024)
Individ. Characteristics						
high educ. level	0.437***	0.379***	0.330***	0.329***	0.222***	0.089
	(0.011)	(0.014)	(0.018)	(0.018)	(0.020)	(0.068)
medium educ. level	0.126***	0.102***	0.083***	0.082***	0.052***	0.042**
	(0.009)	(0.011)	(0.014)	(0.014)	(0.014)	(0.024)
experience	0.017***	0.018***	0.018***	0.017***	0.017***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
experience^2 *(1/100)	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
tenure	0.009***	0.008***	0.006***	0.006***	0.005***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
woman	-0.092***	-0.083***	-0.098***	-0.098***	-0.088***	-0.110***
	(0.009)	(0.010)	(0.013)	(0.014)	(0.014)	(0.037)
married	0.105***	0.097***	0.088***	0.088***	0.083***	0.082***
	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)	(0.020)
woman*married	-0.120***	-0.107***	-0.104***	-0.103***	-0.105***	-0.120***
	(0.011)	(0.012)	(0.012)	(0.016)	(0.015)	(0.032)
ever unemployed	-0.022***	-0.024***	-0.037***	-0.037***	-0.036***	-0.008
	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)	(0.020)
born in East Germany	-0.045***	-0.031**	-0.021	-0.019	-0.015	0.004
	(0.013)	(0.015)	(0.018)	(0.018)	(0.014)	(0.037)
civil servant	-0.064***	-0.076***	-0.094**	-0.094**	-0.107***	0.003
	(0.009)	(0.009)	(0.041)	(0.041)	(0.045)	(0.103)
lives in city	0.016***	0.014**	0.008	0.005	0.002	-0.000
	(0.005)	(0.006)	(0.007)	(0.008)	(0.008)	(0.020)

<Table continues on next page>

<*Continued from previous page*>

Dependent Variable: Log(Hourly Wages)						
	(1)	(2)	(3)	(4)	(5)	(6)
Company Characteristics						
product innovation			0.004	0.004	0.008	0.016
			(0.008)	(0.008)	(0.007)	(0.016)
process innovation			0.036***	0.034***	0.025***	0.018
			(0.008)	(0.008)	(0.008)	(0.016)
very good company performance			0.044**	0.038*	0.045**	-0.017
			(0.026)	(0.021)	(0.022)	(0.046)
good company performance			0.007	0.003	0.010	-0.025
			(0.025)	(0.020)	(0.020)	(0.046)
rather bad company performance			-0.012	-0.016	-0.011	-0.027
			(0.021)	(0.021)	(0.021)	(0.046)
Workplace Characteristics: Measure of...						
non-routine analytic tasks		0.087***	0.120***	0.120***	0.077***	0.048
		(0.012)	(0.016)	(0.016)	(0.016)	(0.060)
non-routine interactive tasks		0.154***	0.154***	0.155***	0.150***	0.179
		(0.011)	(0.014)	(0.014)	(0.015)	(0.036)
routine cognitive tasks		0.031***	0.005	0.004	-0.001	-0.002
		(0.012)	(0.014)	(0.014)	(0.014)	(0.028)
routine manual tasks		-0.006	-0.040**	-0.037**	-0.033*	-0.043
		(0.017)	(0.018)	(0.018)	(0.020)	(0.041)
non-routine manual tasks		-0.119***	-0.096***	-0.095***	-0.083***	-0.041
		(0.011)	(0.014)	(0.014)	(0.014)	(0.036)
77 occupation dummies	No	No	No	No	Yes	Yes
7 company size dummies	No	No	Yes	Yes	Yes	Yes
dummies for manufacturing & trade	No	No	Yes	Yes	Yes	Yes
10 dummies for West German states	No	No	No	Yes	Yes	Yes
$computer * [x - E(x)]$	No	No	No	No	No	Yes
R^2	0.348	0.340	0.399	0.401	0.443	0.455
Number of observations	18547	15266	8936	8936	8897	8897

Employees with low levels of education working in large companies in the services sector are the base category. Heteroscedasticity-consistent standard errors are in parentheses in column (1)-(5). Column (6): $computer * [x - E(x)]$ means that the specification includes interaction terms between computer use and all the level variables X, where all the X had previously been demeaned by the sample averages. Bootstrapped standard errors using 100 resamples. ***, **, *-indicate significance at the 1, 5, 10 percent level.

Table D: Selected Interaction Terms

	Dependent Variable: Log(Hourly Wages)
high educ. level * computer	0.160***
	(0.077)
medium educ. level * computer	0.023
	(0.035)
woman * computer	0.040
	(0.040)
non-routine analytic tasks * computer	0.035
	(0.062)
non-routine interactive tasks * computer	-0.030
	(0.048)
routine cognitive tasks * computer	0.007
	(0.044)
routine manual tasks * computer	0.021
	(0.064)
non-routine manual tasks * computer	-0.059
	(0.046)
experience * computer	0.005
	(0.004)
experience^2 * computer	-0.000
	(0.000)
civil servant * computer	-0.154
	(0.108)
tenure * computer	-0.000
	(0.002)
born in East Germany * computer	-0.026
	(0.052)
married * computer	0.007
	(0.026)
very good company performance * computer	0.091
	(0.084)
good company performance * computer	0.053
	(0.088)
rather bad company performance * computer	0.021
	(0.080)

Table E: OLS Regressions for the Effect of Pencil Use on Wages

Dependent Variable: Log(Hourly Wages)					
	(1)	(2)	(3)	(4)	(5)
Pencil	0.033***	0.016	0.016	0.016	0.005
	(0.010)	(0.011)	(0.013)	(0.013)	(0.013)
Individ. Characteristics					
high educ. level	0.495***	0.398***	0.358***	0.357***	0.231***
	(0.013)	(0.015)	(0.020)	(0.020)	(0.021)
medium educ. level	0.151***	0.112***	0.096***	0.095***	0.057***
	(0.011)	(0.013)	(0.016)	(0.016)	(0.016)
experience	0.018***	0.019***	0.017***	0.017***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
experience^2 *(1/100)	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
tenure	0.009***	0.008***	0.005***	0.006***	0.005***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
woman	-0.079***	-0.075***	-0.088***	-0.089***	-0.095***
	(0.010)	(0.011)	(0.014)	(0.014)	(0.015)
married	0.120***	0.103***	0.091***	0.091***	0.086***
	(0.009)	(0.009)	(0.011)	(0.011)	(0.011)
woman*married	-0.119***	-0.101***	-0.095***	-0.093***	-0.098***
	(0.012)	(0.013)	(0.017)	(0.017)	(0.016)
ever unemployed	-0.036***	-0.033***	-0.046***	-0.046***	-0.044***
	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)
born in East Germany	-0.056***	-0.035**	-0.024	-0.023	-0.010
	(0.015)	(0.015)	(0.020)	(0.020)	(0.020)
civil servant	-0.065***	-0.081***	-0.105***	-0.105***	-0.107**
	(0.009)	(0.009)	(0.040)	(0.040)	(0.045)
lives in city	0.023***	0.017***	0.011	0.007	0.002
	(0.006)	(0.006)	(0.008)	(0.009)	(0.008)

<Table continues on next page>

<*Continued from previous page*>

Dependent Variable: Log(Hourly Wages)					
	(1)	(2)	(3)	(4)	(5)
Company Characteristics					
product innovation			0.002	0.002	0.005
			(0.008)	(0.008)	(0.008)
process innovation			0.057***	0.056***	0.036***
			(0.008)	(0.009)	(0.008)
very good company performance			0.053**	0.048**	0.055**
			(0.023)	(0.022)	(0.023)
good company performance			0.008	0.004	0.013
			(0.021)	(0.021)	(0.021)
rather bad company performance			-0.012	-0.016	-0.009
			(0.023)	(0.023)	(0.022)
Workplace Characteristics: Measure of...					
analytic tasks		0.110***	0.139***	0.139***	0.082***
		(0.012)	(0.017)	(0.017)	(0.017)
interactive tasks		0.185***	0.185***	0.186***	0.158***
		(0.011)	(0.014)	(0.014)	(0.015)
routine cognitive tasks		0.047***	0.023	0.022	0.008
		(0.014)	(0.016)	(0.016)	(0.016)
routine manual tasks		-0.032*	-0.067***	-0.064***	-0.038*
		(0.019)	(0.020)	(0.020)	(0.022)
non-routine manual tasks		-0.171***	-0.135***	-0.134***	-0.095***
		(0.012)	(0.015)	(0.015)	(0.015)
77 occupation dummies	No	No	No	No	Yes
7 company size dummies	No	No	Yes	Yes	Yes
dummies for manufacturing & trade	No	No	Yes	Yes	Yes
10 dummies for West German states	No	No	No	Yes	Yes
R^2	0.300	0.323	0.383	0.385	0.437
Number of observations	15951	13986	8037	8037	8001

Employees with low levels of education working in large companies in the services sector are the base category. Heteroscedasticity-consistent standard errors are in parentheses. ***, **, *-indicate significance at the 1, 5, 10 percent level.

Appendix C

The methodological properties of the matching estimation are discussed in Appendix A. Here, I focus on the estimation results. Table C1 displays the results of the matching estimations. Column (1) shows that, using nearest neighbor matching, the average wage effect of computer use for employees who use computer is 7.8 percent. The coefficient is significant on the 1-percent level. In the matching, 3,461 controls are used to estimate the potential missing outcome of computer users had they not adopted computer technology.[19] Owing to the restriction that only observations within the region of common support are used, there are 400 computer users who are not considered in the analysis. Column (2) shows the results when an Epanechnikov kernel is used in the matching process. This barely changes the magnitude of the ATT, but increases the precision of the estimate.

Table C2, last column, shows the means of the main individual characteristics of the computer non-users that are used as controls in the matching approach. It is evident that, although there is a convergence between treated and controls with respect to most of the observable characteristics, the difference in educational level is still significant. Therefore, columns (3) and (4) of Table C1 show estimates that, in addition to having close propensity scores, restrict matches to be within groups of employees with equal levels of education. This extension increases the number of computer users that are off the common support to 703. The estimate of the nearest neighbor matching declines to 6.2 percent, whereas the coefficient of the estimate that uses an Epanechnikov kernel increases to 8.2 percent. Both estimates are highly significant.

As for the ATT, the estimate using a control function (Table 3, column 6) is very close to the one estimated by the propensity score matching procedure. This may be the case because (i) there is no common support problem, (ii) there is little heterogeneity in treatment effects or all the propensity scores are small, and (iii) there is no serious

[19] Owing to missing values in the matching variables, the sample reduces to 9,240 individuals (5,779 computer user and 3,461 computer non-user) in the matching specification.

mis-specification in the no-treatment outcome (see Blundell, Dearden and Sianesi, 2003).

Figure C1 shows the kernel density estimates of the distribution of propensity scores for computer users and computer non-users. The two distributions greatly overlap. However, the results of the matching procedure reveal that 400 (703) of the 5,779 computer users are outside the region of common support. In this application imputing the values that are outside the common support by relying on a functional form assumption or discarding the computer users without similar counterfactuals from the analysis does not lead to different results.

The propensity scores in the analysis assume values between 0 and 1. The mean is 0.64 and the median is 0.73. This information does not suggest that propensity scores are particularly clustered at certain points in the distribution.

Table C1: Results of the Propensity Score Matching

Dependent Variable: Log(Hourly Wages)				
	(1)	(2)	(3)	(4)
computer	0.078***	0.077***	0.062***	0.082***
	(0.032)	(0.021)	(0.025)	(0.022)
Number of Treated	5,379	5,379	5,076	5,076
Number of Controls	3,461	3,461	3,461	3,461

(1) shows the ATT of computer use using nearest neighbor matching (random draw version, bootstrapped standard errors using 100 replications). (2) shows the ATT using kernel matching (Epanechnikov kernel, bootstrapped standard errors using 100 replications). (3) shows the ATT of computer use using nearest neighbor matching (random draw version, bootstrapped standard errors using 100 replications) with the additional restriction that treated and controls have the same level of education. (4) shows the ATT using kernel matching (Epanechnikov kernel, bootstrapped standard errors using 100 replications) with the additional restriction that treated and controls have the same level of education. Only observations that are on the common support are used. The caliper is set to 0.001. The propensity score is estimated using the level of formal education, age, age^2, work experience, work experience^2, interaction between work experience and education, workplace tasks, born in East Germany, ever unemployed, living in a city, woman, married, married woman, 8 company size dummies, 39 industry dummies, 3 dummies reflecting company performance, 79 occupation dummies, 10 dummies for West German states and a constant as regressors.

Table C2: Mean Comparison for computer Users and computer Non-Users

	computer user	computer non-user prior to matching	selected controls§
high education level	0.25	0.07*	0.17*
medium education level	0.69	0.72*	0.74*
low education level	0.06	0.21*	0.09*
age	39.99	40.29*	40.20
experience	20.04	21.52*	21.22*
tenure	12.37	10.82*	12.33
ever unemployed	0.26	0.34*	0.29
married	0.70	0.65*	0.70
woman	0.44	0.44	0.44
civil servants	0.15	0.05*	0.13
born in East Germany	0.03	0.05*	0.04

* indicates that the means differ with statistical significance of 5 percent in a two-tailed t-test between computer user and either computer non-user prior to matching (column 3) or the selected computer non-user (last column).

§ computer non-users who are selected by the matching procedure.

Figure C1: Kernel Density Estimations of Propensity Scores for computer Users and computer Non-Users