NBER WORKING PAPER SERIES

COMPUTERS AND PRODUCTIVITY IN FRANCE: SOME EVIDENCE

Nathalie Greenan
Jacques Mairesse

Working Paper 5836

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1996

A very preliminary version of this work, based on only one year of data, and only for the manufacturing industries, was presented at the Charleston conference on "The Productivity Impacts of Information Technology Investments," (November 11-13, 1993). In this first stage of our work, we have benefited from the collaboration of Haja Andrianjafy and Dominique Guellec, which we gratefully acknowledge. We are also thankful for discussions with participants in various seminars since then, and in particular to Eli Berman, Ernst Berndt, Franck Briendo, Eric Brousseau, Michel Gollac, Zvi Griliches, Bronwyn Hall, Tor Klette and Frank Lichtenberg. Last, but not least, we are especially indebted to the Statistical Office of the Ministère du Travail and to INSEE for providing us with the data bases on computer use (TOTTO) and firm accounts (SUSE) which we had to match for this study. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Computers and Productivity in France:
Some Evidence
Nathalie Greenan and Jacques Mairesse
NBER Working Paper No. 5836
November 1996
JEL Nos. C20, C81, D24, L63, O33
Productivity

ABSTRACT

In this paper, we make a first attempt to explore the relationship between computer use and productivity in French manufacturing and services industries. We match information on computer utilization in the work place collected at the employee level in the years 1987, 1991 and 1993, with information on firm productivity, capital intensity and average wage available at the firm level. Being based on the answers of very few interviewed employees (only one for 75% of the firms in our samples), our measure of firm computer use is subject to important sampling errors, and hence our estimates of computer impacts are largely affected by random errors in variables downward biases. Nonetheless we find coherent and persuasive evidence that the computer impacts on productivity are indeed positive and that the returns to the firm should at least be in the same range as the returns to the other types of capital. We also show that the sampling errors in measurement biases can be assessed, and we make the general point that econometric studies of the firm can be effectively and substantially enriched by using information collected from workers, even if very few of them are surveyed per firm.

Nathalie Greenan
INSEE-DSE
15, boulevard Gabriel Peri
92245 Malakoff Cedex
FRANCE

Jacques Mairesse
INSEE-CREST
15, boulevard Gabriel Peri
92245 Malakoff Cedex
FRANCE
and NBER
mairesse@ensae.fr

1. INTRODUCTION

Although information technologies have developed and spread extremely fast in the last twenty years and computers are considered as the major innovation of our time, studies of their contributions to firm performances are not very many. There has been a number of monographs but few quantitative and statistical analyses, and to our knowledge, mainly in the U.S.. [1] One reason, perhaps the most important one, is the lack of relevant information on computer use and more generally on Information Technology (IT) related expenditures, such as might have been provided by specifically designed surveys. In a different field, however comparable in a number of respects, that of the studies on productivity and profitability of Research and Development (RD) investments, the existence of regular RD surveys in advanced economies has been fundamental for the development of econometric investigations. To paraphrase the widely quoted Solow paradox, the reason why "we do not see computers in productivity statistics" may be that we do not see them at all in statistics!

In fact, one would not truly expect to see the economic impacts of IT investments with the naked eye. This is an unrealistic hope allowing for the many other factors of firm performances and considering that the absolute magnitude of the returns to IT capital should remain in line with its size, which is still relatively small in proportion to the other forms of physical, human and knowledge capital.[2] In conjunction with case studies which can provide in-depth knowledge, statistical studies are needed to assess the existence and importance of IT contributions to productivity and other dimensions of firm performances (profits, wages or market shares).

In the present paper, we make a first attempt to explore the relationship between computer use and productivity in the French economy. To do so, we rely on the only public information available on computer use in France, which is a specific survey on the techniques and organization of work known under the acronym TOTTO. Since it is a labor force survey, performed by interviewing a sample of salaried workers, we have to be audacious in relating the available information to that pertaining to firm productivity and other firm characteristics. In spite of such audacity, the results, as we shall see, seem surprisingly coherent between themselves and altogether consistent with the view that the computer impacts on productivity are indeed positive and that the returns to the firm should at least be in the same range as the returns to the other types of capital.

In the next section of the paper, we briefly present the TOTTO survey and comment on what we learn from it on the diffusion of computer use for the three years, 1987, 1991 and 1993, when it is available, and for a breakdown of the economy in seven large manufacturing and services industries. In the third section, we try to assess the impacts of computer use on firm productivity for these three years and in these seven industries. More precisely, we consider measures of labor productivity and total factor productivity, as well as measures of the capital to labor ratio, average labor cost and gross rate of returns to capital, and compare how they differ in relation to our main indicator of computer use. In this section, we also explore in a number of ways the robustness of our estimates and investigate specifically the magnitude of the sampling errors of measurement biases affecting them.

2. EXTENT AND EVOLUTION OF COMPUTER USE IN FRENCH MANUFACTURING AND SERVICE SECTORS, FROM 1987 TO 1993

2.1 - Matching employee and firm level information

The information on computer use on which our study is based, is provided by the survey on the techniques and organization of work, "l'Enquête sur les Techniques et l'Organisation du Travail", called "TOTTO". This is to our knowledge the only public source of such information in France. TOTTO is conducted by the "Ministère du Travail" as an occasional supplement to the regular survey on employment ("l'Enquête Emploi"), applicable to all the people surveyed who have effectively been working during the year.[3] TOTTO has been performed for the first time in 1987, by interviewing about twenty thousand workers. It has been repeated on a similar scale in 1993, with a somewhat modified questionnaire. In addition, a selected subset of questions from TOTTO, and among them several on computer use, have been included in 1991 in another survey of the "Ministère de Travail", the survey on the conditions of work ("l'Enquête sur les Conditions de Travail"). We have tried to take full advantage of the relevant and comparable information thus provided in the three years 1987, 1991 and 1993.

Among other topics, TOTTO investigates whether and how "modern technologies" are used. More precisely, as concerns Information Technologies (IT) the interviewed employees are asked whether they use or not a personal computer (PC) or a computer terminal (CT). Those who do are then asked a series of additional questions on the number of hours they spend working with their PC's or CT's, on the nature of the tasks they perform with them, and on the ensuing consequences on their conditions of work, their wage levels, their career prospects, etc.[4] TOTTO is thus a rich source of knowledge on computer use in the workplace, providing estimates of the proportion of employees whose function involves working with a computer, according to different types of use, job categories or nature of tasks.[5]

TOTTO, however, as it is performed at the worker level, has a very important drawback for studies where the main level of analysis is that of the firm. For such studies as ours, trying to relate computer use to firm characteristics and performances, it is first necessary to be able to match the information pertaining to the employees to the information available from firm surveys. This can be done with TOTTO since the interviewed employees are asked the name and address of the firm in which they work . However, it is preferable to exclude the smallest firms and restrict our analysis to the firms with 20 or more employees.[6]

Then it is necessary to be able to maintain the hypothesis that the employees surveyed in a firm are "representative" of their coworkers. This may seem a very heroic assumption to be made. This would especially seem so in our case, where it most frequently happens that only one employee is surveyed per firm. Actually, as long as we can assume that the employees interviewed in the firm are chosen at random, the estimates on computer use by industry that we consider in this section are unbiased and generally quite precise since what matters for their precision is the total number of employees surveyed at the industry level (not at the firm level).[7]

The three firm samples that we obtain by matching the two TOTTO surveys proper and the complementary one to the INSEE firm data bases (for the firms with 20 or more employees) are respectively constituted of 3 190, 3 177 and 3 052 manufacturing and services firms for the three years 1987, 1991 and 1993, and they respectively correspond to 5 441, 4 897 and 4 788 interviewed employees.[8,9] As things stand, we have on average a little over 1.5 interviewed employee by firm in our three samples, and the average share of interviewed employees relative to the firm total of salaried workers is about 2%. As much as about 75% of the firms of the three samples have only one interviewed employee, 15% have two and 5% have three.

The three main indicators of computer use we consider are simply based on the proportions of the interviewed employees in the firm who, in their main task or occasionally, work with a personal computer (PC), or with a computer terminal (CT), or with either one or the other.[10] Hence, when there is one interviewed employee, these indicators can only take the two values 0 and 1; when there is two interviewed employees, they can take the values 0, 0.5 and 1; etc. In what follows, we usually rely on the overall indicator, and when we say for short "computer use" or "computer users", we mean use or user of a PC <u>or</u> a CT. As defined in the 1987 TOTTO questionnaire and again in 1991, the use of a PC explicitly excludes the use of a machine specifically dedicated to word processing. This is not the case anymore in the 1993 questionnaire, but the resulting discrepancy should remain quite small.[11,12]

In terms of number of firms, our three samples are large enough to separate the following seven manufacturing and services industries: food products, intermediate goods, equipment goods, consumer goods, commerce, services (proper), banking and insurance. We can even go further and distinguish between medium and large sized firms (from 20 to 500 employees and more than 500 employees) and three job categories : blue collar, white collar and management. The number of interviewed employees are well over a hundred for most cells at the two way classification level. It is thus possible, without being too foolhardy, to present estimates of computer use not only by sectors but also by sector-size groups and sector-job categories. One has to keep in mind, however, that our most detailed estimates may just provide orders of magnitude.

2.2 - Assessing the diffusion of computer use

The descriptive statistics on computer use by sector that we thought most interesting to present here are showed in Tables 1 to 4 and in Figures 1 and 2, and we shall comment on them in turn.

From Table 1, we see that the overall proportion of computer users in the manufacturing and services industries went up from 25% in 1987 to 38% in 1991 and 43% in 1993, thus increasing by a sizable 3% a year.[13] The corresponding figures for PC and CT taken separately show that their diffusion is roughly similar and that the rise over the period has been strong for both categories, even if the spread of PCs appears significantly faster (by an average 3% per year against an average 2% per year for CTs). The not too surprising implication is that the proportion of users working with both types of machines has also known a rapid progression, from a modest 5% in 1987 to a respectable 17% in 1993 (12% in 1991).

Banking and insurance is by far the sector where computer use is most developed, our overall indicator being already as high as 70% in 1987 and reaching 90% in 1993. The diffusion of PCs has been especially fast in this industry, along with their increasing use in complement with CTs, the proportion of PC users booming from about 30% in 1987 to 60% in 1993, and that of both PC and CT from about 20% to nearly 50%. Commerce, services and equipment goods are the industries coming next, with overall proportions of computer users of about only one fourth in 1987, but rising to 45-50 % in 1993. Food products, intermediate goods and consumer goods industries are last, with much lower ratios of computer users from about 15% in 1987 to 25-30% in 1993.

Beyond the simple average proportions of employees using computers, Figure 1 shows the distribution of the average number of hours per week spent on a computer. Among the computer users, as much as one half do work with their PC or their CT for short hours only : less than 7.5 hours a week. Another good third work for about a fourth to half of their time with their computer (from 7.5 to 22.5 hours a week). The rest of them, that is a sizable proportion of 20 to 30 %, spend most of their workday in front of their screens, a significant fraction even declaring that they do so full time. These indicators of intensity of computer use look quite consistent from one survey to the other, displaying though a marked tendency for a lengthening of the hours worked on a computer, with an overall average of 13 hours per week in 1987, 15 in 1991, 18 in 1993.[14]

Tables 2 and 3 provide other noteworthy precisions. The distinction between firms with less and more than 500 employees shows that the diffusion of computers is roughly comparable and proceeds at a similar pace for the larger firms in all industries, the main exception being that of banking and insurance which is a long way ahead. The discrepancies between industries in their use of computers thus arise, to a good extent, from the different rates of adoption of the medium sized firms. Looking separately at the frequency of computer use among blue collars, white collars and management confirms the very important spread of computers among managers, which contrast with their limited, though significant, adoption among blue collars.[15] The overall average ratio of managers using a computer thus ranges from a large 45% in 1987 to a remarkable 70 % or so in 1993, while that of the blue collar users increases from a small 5% to an appreciable 15%. It is interesting to see that the diffusion of computers among white collars is on the whole as advanced as in management.

Another instructive statistic, which helps putting our indicators of computer use into perspective, is given by the answers to a question included for the first time in the 1993 questionnaire. Interviewed employees who said that they did not use (even occasionally) a PC or a CT were asked whether other people in their company were using one. Their answers were massively a yes. It is thus easy, on the basis of these answers, to construct an indicator of "computer using firms", which tells us that nearly all firms are indeed using computers! [16] The average proportions by industry are all in the range of 85% (for food products) to 99% (for bank and insurance), and we did not think there was much point in showing them here in a Table.[17]

Finally, based on an additional question only asked to the interviewed employees working with computers (in the 1987 and 1993 TOTTO surveys), it is interesting to examine what are the tasks for which computers are more often used. Table 4 thus gives the average frequency by industry of the various tasks performed by the computer users, as indicated by them among the seven different ones listed. These estimates are for the year 1993, but they were quite comparable for 1987.[18] Documentation is the task which is most often declared with a frequency as high as 45% (possibly reflecting in part some vagueness in the nature and scope of that activity). Production, inventory and accounting come next, with a frequency of about 35% for all three functions. Cash, bank, computer aided design (CAD) and scientific computing are last, with much lower frequencies of about 10 to 20%.

It is also possible to simply characterize the degree of association between the different tasks that are computerized in the firm, by the correlations of the corresponding indicators. The pattern and magnitude of these correlations are roughly similar among industries and do not change much between 1987 and 1993. They are summarized in Figure 2. Three groups of tasks clearly appear, which are rather strongly associated within groups but more weakly so between groups. These groups are what we might have expected : first production and inventory control and management ; second, accounting, bank and cash operations ; third, documentation, scientific computing and computer aided design. They express to some extent different strategies of adoption and development of information technologies in the firm.[19]

3. MODELING THE PRODUCTIVITY OF COMPUTER USE

3.1 - A simple model

In our attempt to explore the existence and significance (and even more tentatively the order of magnitude) of the impacts of computer use on productivity at the firm level, we start from the simplest model. This model can be viewed as a standard production function (of the Cobb-Douglas form) expressed in terms of an efficient measure of labor, which itself depends on a measure of computer use. After some straightforward transformations, it directly relates the (log) productivity of the firm, to our variable of computer use, that is the proportion of employees working with a computer (either a PC or a CT). The productivity variable, as well as the other variables we consider (total labor compensation, operating income, gross book value of fixed assets, ...), are measured from the accounting information given by the firm (and its total number of employees), while the proportion of employees using a computer is estimated from the answers of those employees who have been interviewed in the firm.

To make matters even more simple and clear cut, it is possible to go one step further if we restrict ourselves to the extreme situation in which the computer use indicator is based on the answer of only one interviewed employee, and thus reduces to a dichotomous variable equal either to 1 or to 0. In this case, the estimated parameter of interest assessing the impact of computer use on productivity is nothing but the difference in the mean (log) productivity between the two subsamples of firms for which the computer use indicator is respectively 1 and 0. Loosely speaking, it is the difference in the mean (log) productivity of the firms that we can classify, with some plausibility, as being respectively "more" or "less" computerized.

Actually, since already as much as three fourth of the firms in our samples could be matched with only one interviewed employee, we thought preferable to put ourselves completely in this extreme situation. This is what we did in considering restricted samples constructed from the complete ones by keeping a unique interviewed employee per firm. That is, we took the interviewed employee if he was alone (for three fourth of the firms) and chose one at random among the interviewed employees when they were several (for the remaining one fourth of the firms). The bulk of results that we shall be presenting are based on these restricted samples. However, as we will see, we also discuss estimates obtained from even more restricted samples. In particular, we selected subsamples with exactly two and three interviewed employees, respectively constructed from the two subsets of the one fourth and one twentieth of firms with at least two and three interviewed employees.[20] These subsamples are made of firms of increasing sizes, and in order to control for potential size and competition effects, we also consider the restricted subsamples, obtained from these same two subsets of firms but keeping only one randomly chosen interviewed employee (instead of two or three).[21] Relying on the restricted samples rather than the complete ones has, of course, a cost in terms of the efficiency of estimation.[22] However, our main worry here is much more about the magnitude of the biases arising from the sampling errors of measurement of our computer use variable. Besides the appeal of simplicity, the advantage of focusing on the restricted subsamples is in allowing us to tackle more neatly this essential issue.

To be more precise in our explanations, a minimum of notations and some algebra are useful. In its simplest form, the model we rely on can be written as :

$$VA = AL^{*}(p)$$

where VA denotes our measure of the production (value added) of the firm and L^* stands for a measure of efficient labor depending on the proportion p of employees in the firm working with a computer. Assuming also the simplest expression for the measure of efficient labor, L^* can be written as :

$$L^{*} = L^{NC} + (1+\gamma)L^{C} = (L^{NC} + L^{C})\left(1 + \frac{\gamma L^{C}}{L^{NC} + L^{C}}\right) = L(1+\gamma p)$$

where L (= $L^{NC} + L^{C}$) is the total number of employees, L^C (= pL) the number of those working with a computer, L^{NC} the number of those who do not, and where the parameter γ measures the difference in the relative labor efficiency (or marginal productivity) of these two categories of employees. Dividing by L, taking logs and linearizing in terms of γp as a convenient approximation, we obtain :

$$\mathrm{Log}(VA / L) = \mathrm{Log}(1+\gamma p) + \mathrm{Log}A \approx \gamma p + c$$

where c is a constant.[23]

We are thus led to the following simple linear regression :

$$\mathrm{Log}(VA / L)_i = \gamma p_i + c + \varepsilon_i$$

for i = 1 to N, where i is the subscript for the ith firm in the samples considered (of N firms), and where ε_i denotes the usual disturbance term in the regression, summarizing all sources of "errors".

In this formulation, the parameter γ of labor efficiency is also our parameter of interest assessing the impact of computer use on the log-labor productivity VA/L. We know that it is consistently estimated by ordinary least squares, as long as we can maintain the basic hypothesis that the computer use variable p is uncorrelated (or nearly so) with the error term ε in the regression. If we cannot accept this highly problematic hypothesis, as we shall see below, the least squares estimate should not be taken as being an unbiased estimate of the "true structural" parameter γ , providing a reliable order of magnitude of the true structural impact ("other things being equal") of computer use on labor productivity. However even so, the least squares estimate remains an interesting descriptive statistic to consider, conveniently summarizing the empirical relation ("other things changing, as they do in fact") between (measured) computer use and labor productivity.

The least squares estimate of γ can be interpreted even more simply if we focus our attention (as we do) on the restricted samples of firms with only one interviewed employee. In this case, the firm i computer use variable p_i reduces to the dichotomous indicator I_i equal to 1 or 0, depending on whether the interviewed employee works with a computer or not. Our regression can then be written in terms of I_i and its complement $NI_i = 1 - I_i$ as :

$$\text{Log}(VA / L) = (\gamma + c)I_i + cNI_i + \varepsilon_i$$

Since I_i and NI_i are orthogonal indicators by construction ($\sum_{i=1}^{N}(I_i)(NI_i) = 0$), it is easy to see that the least squares estimates of $(\gamma + c)$ and (c) are the means (the simple averages) of the log-labor productivity for the two subsamples of firms such that $I_i = 1$ ($NI_i = 0$) and $NI_i = 1$ ($I_i = 0$). Thus, as already stated, the least squares estimated γ is equal to the difference between these two means.

3.2 - Controlling for capital and labor quality

There are clearly many reasons why the simple model we just presented, can be misspecified in the sense that the error term ε is correlated with our measure of the computer use variable. We shall now discuss the three reasons we think are the major ones in the present instance : first in this sub-section, the correlation with omitted capital and the complementarity with unobserved labor quality; then in the next one, the sampling errors in the measurement of computer use.[24]

The most obvious shortcoming of our model as formulated at this point is of course the omission of other relevant factors of productivity, the most influential of them, as shown by numerous econometric studies at the firm level, being the stock of (physical) capital. We can also take this factor into account here, as usually done, measuring it by the gross book value of fixed assets C, and restating the model as the usual Cobb-Douglas production function :

$$VA = AC^{\alpha} L^{*}(p)^{\beta}$$

This leads to the following regression :

$$\mathrm{Log}(VA / L)_i = \alpha \mathrm{Log}(C / L)_i + (\alpha + \beta - 1)\mathrm{Log}(L)_i + (\beta\gamma)p_i + c + \varepsilon_i$$

Although we actually estimate this regression by ordinary least squares as just written, we find it appealing to view it as directly relating an estimated log- total factor productivity to the computer use variable. With total factor productivity defined as $TFP = VA / (C^{\alpha} L^{\beta})$ the regression can be restated as simply as previously in terms of labor productivity (when ignoring capital) :

$$\mathrm{Log}(TFP)_i = \delta\ p_i + c + \varepsilon_i$$

where $\delta = (\beta\gamma)$.

Our parameter of interest δ (which is now equal to labor efficiency time labor elasticity) can thus be interpreted as measuring the impact of computer use on log-labor productivity controlling for capital intensity (C/L) and size (L), or equivalently its impact on log-total factor productivity.[25] In the case of the restricted samples with only one interviewed employee per firm, we can also note, just as previously, that the least squares estimate of δ is equal to the difference between the means of the (estimated) log-total factor productivity for the two subsamples of firms in which the one interviewed employee respectively uses a computer and does not ($I_i = 1$ and $I_i = 0$).[26]

As part of our analysis, it is interesting per se to assess the relation between log-capital intensity and computer use, by considering the simple descriptive regression between the two (which we do in presenting our results). This regression contributes to give us a better appreciation of the reliability of our computer use indicator. It can also be viewed as an "auxiliary" regression accounting for the difference in the magnitude of our two estimated parameters of interest γ or δ. A priori one would expect that computer use and capital intensity should be positively and rather strongly correlated for two reasons. First, computer equipment is part of the capital stock and its share is likely to be larger with increasing computer use. Second, it is also likely, irrespective of computer capital, that computer use is more developed in capital intensive firms.[27]

Another omitted variable type of problem which affects the specification of our model, and is particularly important in the present analysis, is that of labor quality. Computer use tends to increase with the level of education and the general skills of the workers ; it also requires from them specific knowledge and abilities developed on the job training and with experience. These two aspects of the correlation of computer use and labor quality are quite comparable to that of the correlation of computer use with capital (with capital in general and with computer capital specifically). Even if we were observing labor quality in some detail, disentangling these two aspects would be an intricate task. Since we do not, what we can do here is to make the extreme assumption that all the differences in wages are due to differences in labor quality and to use the available information on wages at the firm level to proxy for labor quality. In practice, this amounts to measuring labor in our previous regressions by total labor compensation LW (total wages and social security associated costs) instead of using the total number of employees L.[28] To put the matter as simply as before, we can also say that we relate a labor quality adjusted total factor productivity TFPA = $VA / (C^{\alpha} LW^{\beta})$ to our computer use indicator .[29]

One would *a priori* expect that the true coefficient of impact of computer use on productivity would lie in between the two estimated δ corresponding to the non adjusted and the adjusted total factor productivities (TFP and TFPA) respectively. The former will be too large to the extent that high levels of general skills largely go together with wages and productivity and are also conducive to the diffusion of computers. The latter will be too small to the extent that the specific skills, which are intrinsically needed for computer use, account for simultaneous increases in productivity and in wages. It will also be too small if the wage earners share the benefits resulting from the firm higher efficiency and competitiveness, and if these better performances are linked to more intensive computer use.

The analysis of the impact of computer use on wages is of course worth considering per se and not only in connection with productivity.[30] In parallel with our other results, we will be presenting the estimated coefficient of computer use in the simple regression of the (log) average cost of labor or average wage rate W (computed as the ratio of total labor compensation LW to the total number of employees L). In the same vein, we also show the estimated coefficient of computer use in the simple regression of the (log) average gross rate of returns to capital EBE/C (computed as the ratio to the gross book value of fixed assets C of the gross operating income EBE, obtained as value added VA less total labor compensation LW). Intuitively, the impacts of computer use on the returns to capital EBE/C thus measured and on total factor productivity adjusted for labor quality TFPA should be qualitatively similar.[31]

3.3 - Sampling errors and biases

Our most worrisome econometric problem, however, is specific to the present analysis and affects all the least squares estimates of the above regressions. It follows from the fact that our computer use variable is based at best on the answers of very few interviewed employees in the firm, and in general on that of only one. Our indicator p_i for firm i is thus an estimate of the true proportion p_i* of computers users (which in principle could exactly be known if all the firm employees were surveyed); as such it is necessarily affected by more or less severe sampling errors, which result in (downward)biases in our regression estimates of the computer impact coefficients. Fortunately, we are going to see that we are in the pure classical case of random errors in variables and that it is possible to assess the magnitude of these biases.

The persons sampled in TOTTO being randomly drawn in the population at large, we can consider that the n_i employees who turn out to be interviewed in firm i are randomly drawn among all its employees. We can further assume that their answers are independently distributed if firm i is not too small (a total number of 20 employees in the firm being enough). In other words, we know that p_i is the empirical mean of a sample of n_i independent realisations of the random variable p_{ih} with true mean p_i* and variance σ_i^2, (where h = 1 to n_i denotes the randomly drawn interviewed employees in firm i). As such p_i is an unbiased (and consistent) estimate of p_i* with variance σ_i^2/n_i. In our case, p_{ih} is the binomial variable $B(p_i^*; n_i)$ and thus $\sigma_i^2 = p_i^* (1 - p_i^*)$. More precisely, we can write for the given firm i :

$$p_i = p_i^* + e_i$$

with

$$E(p_i ; \text{if } i) = p_i^* \quad \text{or} \quad E(e_i ; \text{if } i) = 0 \text{ ,}$$

and

$$\mathrm{Var}(p_i ; \text{if } i) = \mathrm{Var}(e_i ; \text{if } i) = \sigma_i^2/n_i = p_i^* (1 - p_i^*) / n_i \text{ .}$$

Proceeding one step further and considering that our samples of firms arise at random from an underlying (large) population, we see that :

$$E(e_i) = E_i(E(e_i ; \text{if } i)) = 0 \quad \text{and} \quad \mathrm{Cov}(e_i\, p_i^*) = E(e_i\, p_i^*) = E_i(E(e_i\, p_i^* ; \text{if } i)) = 0$$

and hence :

$$E(p_i) = E(p_i^*) = \bar{p}^* \quad \text{and} \quad Var(p_i) = Var(p_i^*) + Var(e_i)$$

where $\bar{p}^*$ and $Var(p_i^*)$ the true mean and true variance of the computer use variable.

The sampling errors e_i are uncorrelated with the true values p_i^* of our computer use variable and we are thus in the pure classical (textbook) case of random (uncorrelated) errors in variables. In this case, the least squares estimate $\hat{\gamma}$ of the parameter of interest γ in our simple regression model is biased downward in proportion to the share λ of the error variance in the total measured variance :

$$p\lim(\hat{\gamma}) = (1-\lambda)\gamma$$

with $$\lambda = Var(e_i) / Var(p_i) = Var(e_i) / [Var(p_i^*) + Var(e_i)] \, .$$

An important point to be made is that, for large enough samples, the standard error of the least squares estimate $\hat{\gamma}$ is also biased downward, but in a proportion $(1-\lambda)^{1/2}$, and hence that the corresponding t-ratio is biased downward in this same proportion $(1-\lambda)^{1/2}$.[32] Thus the finding that the least squares estimate $\hat{\gamma}$ is statistically significant (i.e., statistically different from zero at a significance level of 5 percent with a t-ratio of about 2 or above) is a fortiori evidence that the true parameter γ is not zero.

Assuming for simplicity that the numbers of interviewed employees are the same: $n_i = n$ (= 1, 2 or 3, for our various restricted samples), we can write:

$$Var(e_i) = E_i\,[Var(e_i\,;\,\text{if } i)] + Var_i\,[E_i(e_i\,;\,\text{if } i)] = E_i(\sigma_i^2/n) = \sigma^2_{(w)}\,/\,n$$

where $\sigma^2_{(w)}$ can be seen as the overall within firm variance of p_{ih},

and hence:

$$Var(p_i^*) + \sigma^2_{(w)}\,/\,n = Var(p_i)$$

$$\lambda = \sigma^2_{(w)} / nVar(p_i) = \sigma^2_{(w)} / [nVar(p_i^*) + \sigma^2_{(w)}] \, .$$

Knowing that p_i is the mean of a binomial variable, it is easy to show that[33]

$$\mathrm{Var}(p_i^*) + \sigma^2_{(w)} = \bar{p}^*(1-\bar{p}^*)$$

and hence:

$$\lambda = [\bar{p}^*(1-\bar{p}^*) - \mathrm{Var}(p_i^*)]/[\bar{p}^*(1-\bar{p}^*) + (n-1)\mathrm{Var}(p_i^*)] \cdot$$

It is clear from these formulas that the relative bias λ decreases with n and the more rapidly so the higher the true variance Var (p_i*). If, for example, we consider that the true computer use variable p_i* is uniformly distributed among firms between the two extreme values 0 and 1, we have $\bar{p}_i^* = 0.50$ and $\mathrm{Var}(p_i^*) = 0.25/3$ which gives:

$$\lambda = 2/3 \text{ if } n=1, \quad \lambda = 1/2 \text{ if } n=2, \text{ and } \lambda = 1/4 \text{ if } n=3.$$ [34]

In plain words (and as could be expected), the least squares estimates of γ should be less downward biased, the larger the number n of interviewed employees per firm. We should thus be able to check that as a rule our estimated γ's are greater for the restricted samples with two interviewed employees than for those with only one, and for those with three interviewed employees than for those with two.

It can also be seen on the above formulas that using the empirical mean and empirical variance $\bar{p}$ and $\hat{\mathrm{Var}}(p_i)$, for a given n larger than 1, we can simply derive consistent estimates of the true mean $\bar{p}^*$ and true variance Var(p_i*), and hence of the relative bias λ.[35,36] We can thus retrieve a consistent estimate of the γ parameter from its biased least squares estimate $\hat{\gamma}$. In principle, we can also optimally combine the different least squares estimates $\hat{\gamma}$ for the different n (including n=1) to obtain another consistent and more efficient estimate of the true γ. This we can achieve rather simply by applying the Asymptotic Least Squares (ALS) method. For us here, however, implementing such a procedure would not change the substance of our results and the conclusions we draw from them, and we did not think it was worthwhile to do so as a mere matter of technique.[37] This will be clear from looking at our various estimates, which we can do now.

4. TENTATIVE EVIDENCE ON THE PRODUCTIVITY OF COMPUTER USE

4.1 - Main empirical results

All our estimates of the impact parameters of computer use on productivity and other firm characteristics are given in Tables 5, 6 and 7. These estimates (and their standard errors in parentheses) are shown in the same format in all three tables, that is, side by side, for the three years 1987, 1991 and 1993, and for the three (log)labor and (log)total factor productivity measures : VA/L, TFP and TFPA, and our other three (log)variables of interest : capital intensity C/L, average wage W, and average returns to capital EBE/C.

Our main results can be gathered from Table 5, which presents the estimates computed on the restricted samples with one interviewed employee, for the seven industries separately and overall. These results are obtained with our preferred (or reference) indicator of computer use, which, as we said, is based on the answers of the interviewed employees telling whether they are working with a CT or a PC, irrespective of their job status and the tasks they performed. We have found, however, that our estimates are surprisingly robust to the various other more focused indicators we can choose ; this is documented in Table 6 for the overall regressions.

As we have explained at some length, our results do suffer from downward biases due to the sampling errors in the measurement of the computer use variable (independently of its precise definition). An idea of the magnitude of the incurred biases is provided by the estimates found for the restricted samples with respectively two and three interviewed employees, as compared to the ones obtained for the restricted samples with only one interviewed employee. These estimates (plus some others to be explained) are given for the overall regressions in Table 7, and we discuss them in the next sub-section.

Considering our main results in Table 5 and putting aside the banking and insurance industry, the pattern of estimates which emerges with respect to our six outcome variables is quite comparable from one industry to another, and for the economy as a whole (i.e., across columns for the different rows). It is clear at first sight that all or nearly all estimates do not change over time in a statistically significant way ; most of them are actually quite close in the three years 1987, 1991 and 1993, and strikingly so for the overall regressions.[38] In all six industries other than banking and insurance, and overall, these estimates are positive, and in general very significant, for the firm labor productivity VA/L, as well as for the capital intensity C/L and the average wage W. The estimated impacts on the firm total factor productivity TFP, that is after controlling for the differences in capital intensity (and size), are thus markedly smaller than on labor productivity, but they mostly remain significant and positive. However, the estimates for our second measure of total factor productivity TFPA, tentatively adjusted for labor quality, that is also controlling for the differences in average wage, are all practically insignificant and negligible. This is confirmed by the estimated impacts on the average returns to capital EBE/C, which are roughly similar as expected, although they tend to lean over the negative side (even if not significantly so) and seem rather erratic.[39]

The similarity of profiles of our estimates in the six industries, other than banking and insurance, does not mean that the differences among them are small. Some of these differences can be substantial, even if they are not statistically significant. It is noteworthy that the services industry stands ahead of the others with larger coefficients of computer use on VA/L, C/L, W and TFP. The consumer goods industry comes in second, though less clearly so.[40] Such large impacts in these two industries may arise from sizable differentials in the prices of their services and products, reflecting to a large extent, if not fully, quality differences between them. These quality differences are directly or indirectly linked to the diffusion of computers in the firms and to their utilization to perform differently some important tasks, improve in various ways the work organisation and the relations with customers and suppliers, and ultimately produce better goods and services. The strict explanation in terms of productivity differences, considered in the restrictive sense of unadjusted for product quality improvements, may well play a modest role. Such an interpretation of the larger estimated impacts of computer use in the services and consumer goods industries remains of course highly hypothetical, in the absence of relevant firm level information on product prices and quality attributes.

By contrast, the special case of the banking and insurance sector, in which our estimates are insignificant for all our variables, is what it should be (and could not be otherwise). Since the diffusion of computer use has been, as we have seen, overwhelming in this industry, we would not expect to detect any differential influence in a cross-sectional sample of firms (at least using an indicator of the prevalence of computer use as ours, and not one of its actual efficacy). This is also necessarily so, estimation wise, because there is indeed very little (identifying) variance in our computer use indicator for this industry .[41]

Instead of the broad definition of computer use we have favored, we can also consider more narrowly defined indicators, on the assumption that they might perform differently in accounting for firm productivity differences, some of them being more appropriate than the others . For example, one might make the hypothesis that computers have a larger productivity impact when used in production and inventory activities than in accounting, bank or cash operations. One might similarly think that the productivity impact of computers be greater for blue collar workers than for management. One would on the other hand expect that the productivity of computer terminals (CTs) and that of personal computers (PCs) should be the same.

With enough willingness, it is possible to find in Table 6 some indication in support of such conjectures. The estimates for the PC and CT based indicators are indeed, on the whole, quite close to each other, and quite close to our reference estimates based on the overall computer use variable. The estimates for the indicators constructed from the answers of blue collars only and from those of management only are also quite comparable for the labor and total factor productivities VA/L and TFP, and the average wage rate W; however the former tend to be larger than the later as concerns the capital intensity C/L. The estimated impacts on both VA/L and TFP, as well as on C/L (not on W), appear to be consistently greater for the computerization of production and inventory tasks than for that of accounting and related tasks (even if the differences are not statistically significant).

However, in view of the very simple nature of our computer use indicators, and the potential shortcomings affecting their measurement (even if we disregard sampling errors), such detailed comments are perhaps much too far-fetched and certainly fragile. In view of these limitations, what is in fact surprising is the global concordance of the estimates based on the various more focused indicators, and the clear confirmation they give of the pattern and approximate size of impacts on our outcome variables (whether significant or not), which we already found (and commented upon) using our preferred larger definition.[42] This robustness of the results should be seen as rather convincing and encouraging evidence.

The estimates given in the last row of Table 6 also convey the same indication of robustness. Although we have cleaned our data for the obvious outliers as concerns all the variables of interest, one might still be a little suspicious that the significance and coherence of our estimates could be due to relatively few influential observations. We thus looked for such observations varying the severity of the criterion of influence and reran regressions on the samples in which the firms corresponding firms have been taken out.[43] Our estimates remain basically unchanged, as can be seen on the ones reported here, which we obtained when dropping the approximately two percent most influential observations for each of the overall regressions on our six outcome variables.[44]

4.2 - Assessing errors in variable biases and the economic significance of the estimates

The sampling errors problem in the measurement of our computer use variable is the one which worried us (and puzzled us) most, and which we would have particularly blamed for our results, if they had been inconclusive or unreasonable. We investigated it in Table 7 by considering the estimates computed on our various restricted samples and subsamples, with measures of the computer use variable resulting from the answers of either one, or two, or three, or all the interviewed employees in the firm.

To be precise, the estimates given in the second and third rows of the upper and lower panels in Table 7 (noted by Count ≥ 2, NS = 1, and Count ≥ 3 NS = 1), are computed on the restricted subsamples which we have constructed by choosing at random one interviewed employee for the firms respectively matched with two interviewed employees at least, or which three of them at least. These estimates should be of about the same magnitude than our reference estimates (recalled in the first row of the two panels, and which we could have noted Count ≥ 1, NS = 1). The estimates in the fourth row (noted Proportion) are based on the complete samples, using the proportion p of computer users for all the interviewed employees in the firm. Since as much as three fourth of the firms are matched with only one interviewed employee, these estimates should also not be too different from our reference estimates. The estimates in the fifth and sixth rows (noted by Count ≥ 2, NS = 2, and Count ≥ 3 NS = 3) are obtained from subsamples of firms respectively matched with two and three interviewed employees at least, but keeping now exactly either two or three of them. These estimates should be less downward biased than all our previous estimates, and the ones based on the subsamples with three interviewed employees even less biased than the ones based on the subsamples with two interviewed employees. More exactly, this should be so for the estimates when positive and significant, that is with respect to VA/L, C/L, W and TFP, while the other estimates, that is for TFPA and EBE/C, should remain not significantly different from zero in all cases. Last, the estimates in the seven row (noted Proportion shrunk to the sector average) are comparable to the ones in the fourth row, but using as the proportion p^s_i of computer users in the firm i the half sum of the actual proportion p_i for the firm and that p_{ind} for its industry: $p^s_i = (p_i + p_{ind}) / 2$. This is a way to reduce the variance of the computer use variable and mitigate the measurement error problem.[45] Hence, these estimates, too, should be larger than the reference ones.

All these expectations are, on the whole, remarkably fulfilled. Our estimates on the restricted subsamples with two interviewed employees (fifth row) are about 1.7 times higher on average (and ranging from 1.5 to 1.9) than the reference estimates with one interviewed employee (first row), for all the regressions on VA/L, C/L, W and TFP (and all the three years). Likewise, our estimates on the restricted subsamples with three interviewed employees (sixth row) are about 1.3 times larger on average (and varying between 1.0 and 1.5) than the previous ones for the restricted subsamples with two interviewed employees, for all the same regressions. They are thus more than two times greater on average than the reference estimates.

These substantial increases in our estimates, when we go from one to two and to three interviewed employees, are not accounted for by some features of the subsamples on which they are based, in particular the fact that they are made of firms of increasing sizes.[46] Keeping these subsamples the same, but restricting them to one interviewed employee, do provide estimates (second and third rows) which are indeed quite close to the reference ones. The increases with the number of interviewed employees found in our estimates can only be explained by decreasing biases due to smaller sampling errors variances in the computer use variable. This is also confirmed by comparing the estimates on the complete samples, using the proportion of computer users among all interviewed employees (fourth row), to the same estimates using this proportion shrunk to the industry average (seventh row). The former are about the same (slightly higher) than the reference estimates; the latter, as we expected, are much larger, by a factor of 1.7 on average, for all four types of regressions showing significant impacts (VA/L, C/L, W and TFP).

Until now, we have mainly commented on the statistical significance of our estimates and on the comparison of their relative size across regressions, by industry and according to the choice of the computer use indicator, as well as for our various restricted samples. But we have not really discussed their magnitude and economic significance, in part because we wanted first to assess the importance of the errors in variables biases affecting them. To this we now turn. As explained above, we can estimate the relative bias λ using the empirical mean and variance of the computer use variable for our restricted samples with two and three interviewed employees.[47] We thus find quite large values of λ : of about 70 percent when we have two interviewed employees and about 60 percent when we have three.[48] The consistent estimates of the coefficients of computer use that we derive by dividing the least squares estimates by $(1 - \lambda)$ vary to some extent from one year to the other. However, they are remarkably close on average over the three years, whether we compute them from the restricted samples with two interviewed employees or from the ones with three interviewed employees.[49] We thus obtain an average coefficient of 1.15 for (log) labor productivity VA/L, of 1.75 for (log) capital intensity C/L, of 0.95 for (log) average wage W, and of 0.80 for (log) total factor productivity TFP.

These corrected (consistent) estimates are indeed very large. If we take them at their face value and consider for example a cross-sectional difference of 18% in firm computer use, which is about the average value of the true dispersion $(\mathrm{Var}\ p_i^*)^{1/2}$ that we find, they would amount to accounting for about 50% of the observed cross-sectional dispersion in W, 40% in that of V/L, 30% in that of TFP, and 25% in that of C/L.[50] If we gauge them in terms of evolution, and relate them to the average computer diffusion of about 3% per year that we observe (section 2.2), they would imply yearly increases of 3.5% in VA/L, 5.2% in C/L, 2.8% in W and 2.4% in TFP. Clearly such numbers would be too high if they were to be interpreted as measuring real causal impacts. But they just express in a meaninful way what are the orders of magnitude implied by the very significant (simple) correlations we have found between computer use and the variables considered.

In order to assess the true impacts of computer use on productivity (independently of other correlated factors), we have, as previously explained, to control at least for capital intensity and labor quality. If using the firm average wage to proxy for labor quality does too much, as we also said, in the sense that it also controls for computer users specific skills and for the sharing of firm benefits due to computer related performances, then the true impacts should be within the range of the coefficients found for the non-adjusted and adjusted total factor productivities TFP and TFPA, that is less than 0.80 but more than 0. This indicates that indeed the (true) impacts of computer use on productivity should be positive at the very least. To set one's ideas, a value of 0.20 might be taken as a sensible, though perhaps conservative, estimate. This is in fact about the lowest value that we obtained for the coefficient of computer use, when we tried to (very crudely) control for labor quality in our regressions on the basis of the job status of the interviewed employees themselves (instead of the firm average wage) [51]. Such an estimate will still account for nearly 10% of the cross-sectional dispersion of TFP (for one standard deviation in firm computer use, i.e., 18%), and for a yearly increase of about 0.5% in the growth of TFP (for an average diffusion of computer use of about 3% per year).

5. CONCLUSIONS

The originality of our study is in using information on computer utilization based on the answers of one or very few interviewed employees in the firm, in order to assess whether the firm productivity is indeed significantly related to computer usage or not. Two main conclusions arise from this attempt, one of substance, one of method.

Although our indicator is both a crude and very noisy measure of computer use, we obtain surprisingly coherent and persuasive evidence. We find very significant and positive correlations between computer use and the labor productivity of the firm, as well as with its capital intensity and its average wage. By using two measures of total factor productivity, one controlling only for capital intensity and the other also for the average wage, viewed as a proxy for labor quality, we can bracket the true impacts of computer use on firm productivity. We are thus able to show that these impacts are positive, and to infer that the returns to the firm should at least be in the same range as the returns to the other types of capital.

Since the interviewed employees can be considered as randomly drawn in their firms, the sampling errors affecting our measure of firm computer use are themselves random and uncorrelated with the underlying true value. We are thus in the pure classical case of random errors in variables. Based on the firms with two interviewed employees or more, we can actually compute the variance of the sampling errors; it is equal to the within firm variance of the individual answers of the firm interviewed employees, divided by their number. We can therefore correct our (least squares) estimates of computer use impacts from their sampling errors of measurement downward biases. In the large range of possible values that we find, a value of 0.20 seems to be a sensible though perhaps conservative estimate. With such an estimate, a difference of one standard deviation in firm computer use (i.e., 18%) will account for nearly 10% of the cross-sectional dispersion of firm total factor productivity; equivalently, an average diffusion of computer use of 3% per year (which is what we observe) will result in yearly increase of about 0.5% in the growth of total factor productivity.

In our view, an important conclusion of our paper is methodological. Econometric studies of the firm can be effectively and substantially enriched by using information collected from workers, even if very few of them are surveyed per firm. It can be seen from our example that it is even possible to have only one interviewed employee for most of the firms in the sample under study, as long as one has a large enough subsample of firms with two or more interviewed employees. However, it would certainly be preferable to have a minimum of two interviewed employees for the majority of firms, in order to have a more precise and robust assessment of the sampling errors in measurement biases. One can think of implementing specific surveys of workers, in order to measure firm variables which are simply not available, or would be difficult (and extremely costly) to evaluate at the firm level, such as labor composition by education and skills, age and experience; normal hours of work and overtime; work effort and degree of motivation; ability of management, etc. The reliability of the information which can thus be collected, will of course crucially depend on the quality and adequacy of the questionnaire. In our case, for example, the question on computer use had the advantage of being easy to understand and easy to answer objectively. The resources involved in performing such complementary surveys for a sizeable sample of firms are necessarily substantial. Notwithstanding these limits, it should be possible to investigate quantitatively many various aspects of the firm behaviour that could not be otherwise. Since a major shortcoming encountered by econometric studies on firm data is the lack of relevant variables, this, we think, is an important message.

REFERENCES

BRYNJOLFSSON, Erik (1993) : "The Productivity Paradox of Information Technology" *Communications of the ACM*, Vol. 36, n° 12, 67-77.

BRYNJOLFSSON, Erik and HITT, Lorin (1995) : "Information Technology as a Factor of Production : The Role of Differences among Firms", *Economics of Innovation and New Technology*, 3(3-4), 183-199.

CUNEO, Philippe and MAIRESSE, Jacques (1984) : "Productivity and R-D at the Firm Level in French Manufacturing", in *R-D, Patents and Productivity*, ed. Z. Griliches, , The University of Chicago Press, 375-393.

ENTORF, Horst, and KRAMARZ, Francis (1994) : "The Impact of New Technologies on Wages", Document de travail du CREST, INSEE, n° 9407.

GOLLAC, Michel (1989) : "L'ordinateur dans l'entreprises reste un outil de luxe", *Economie et Statistique*, n° 224, Septembre, 17-25.

GOLLAC, Michel (1993) : "La diffusion de l'informatique au travail", *Données Sociales*, INSEE, 486-494.

GOURIEROUX, Christian and MONFORT, Alain (1995) : "*Statistics and Econometric Models*" Vol. 1, Chap. 9, Cambridge University Press, Cambridge.

GOURIEROUX, Christian, MONFORT, Alain and TROGNON, Alain (1985) : "Moindres Carrés Asymptotiques", *Annales de l'INSEE*, 58, 91-122.

GREENAN, Nathalie and GUELLEC, Dominique (1994) : "Organisation du travail, technologie et performances : une étude empirique", *Economie et Prévision*, n° 113-114, 39-56.

GRILICHES, Zvi and MAIRESSE Jacques (1995) : "Production Functions : The Search for Identification", National Bureau of Economic Research Working Paper n° 5067.

GRILICHES, Zvi and RINGSTAD, Victor (1971) : *Economies of Scale and the Form of the Production Function*, North Holland Publishing Company, Amsterdam.

KRUEGER, Alan (1993) : "How Computers Have Changed the Wage Structure : Evidence from Microdata, 1984-89", *Quarterly Journal of Economics*, 33-59.

LEHR, William and LICHTENBERG, Frank (1996): "Information Technology and Its Impact on Productivity: Evidence from Government and Private Data Sources, 1977-1995", Mimeographed, Columbia University.

LICHTENBERG, Frank (1995): "The Output Contributions of Computer Equipment and Personnel : A Firm Level Analysis", *Economics of Innovation and New Technology,* 3(3-4), 201-217.

MOATTY, Frédéric (1993) : "Qui utilise l'informatique au travail", Dossier de recherche du Centre d'Etudes de l'Emploi, n° 51.

MOATTY, Frédéric (1994) : "Les salariés et l'ordinateur", *Quatre Pages du Centre d'Etudes de l'Emploi,* Septembre.

OLINER, Stephen and SICHER, Daniel (1994) : "Computers and Output Growth Revisited : How Big is the Puzzle", *Brookings Papers on Economic Activity*, 2, 273-317.

WILSON, Diane (1992) : "IT Investment and its Productivity Effects : An Organisational Sociologist's Perspective on Directions for Future Research", *Economics of Innovation and New Technology,* 3(3-4), 235-251.

FOOTNOTES

[1] For references to these analyses, see the three papers by Erik Brynjolfsson and Lorin Hitt (1995), Frank Lichtenberg (1995) and Diane Wilson (1995), reprinted in this volume.

[2] For such computations of the possible impact (at best) of IT investments on productivity, see Stephen Oliner and Daniel Sicher (1994).

[3] The sampling scheme of l'Enquête Emploi gives a probability of 1/1000 to be selected for each individual in the French total labor force (which amounts to roughly 20 millions).

[4] These additional questions, except for the one on the time spent working on a computer, are not asked in the 1991 survey.

[5] See Michel Gollac (1989, 1993) and Frédéric Moatty (1993, 1994) for such studies.

[6] From the name and address of the firm, it is first possible to trace its so-called SIREN identification number ; we then can do the matching with the INSEE firm data files using the SIREN identification. At both stages, however, important losses occur in the numbers of interviewed employees and firms that can actually be linked. In the first stage, this is in particular the case of the self-employed persons ; and in the second stage this is mainly true for the interviewed employees working in the smaller firms. For that reason, we altogether preferred to exclude from our samples the firms with less than 20 employees and the employees interviewed in them.

[7] These sectoral estimates are computed as the simple mean (non weighted average) of the corresponding firm estimates. Although the firm estimates are extremely imprecise (being based on one respondent or at best on a few ones), they are unbiased (even if there is only one respondent and if, as a result, they can only take the values 0 or 1). Matters are less favorable in the next section, where the firm estimates on computer use are taken as measures of our explanatory variable of the firm performances in a regression model. The sampling random errors affecting them should in this case result in a downward bias in the estimated regression coefficients of impacts.

[8] The INSEE firm data bases are constructed from "SUSE" : "Système Unifié de Statistiques d'Entreprises" which combine the information of the firm annual surveys ("Enquêtes Annuelles d'Entreprises") and of the firm fiscal declarations ("Déclarations de Bénéfices

Industriels et Commerciaux"). As explained in footnote 6, we have excluded the smaller firms of our analysis, because of the difficulties involved in the matching of SUSE and TOTTO information. We have also excluded public enterprises and-non profit organizations, and we have not considered the Building, Construction, Energy, Transportation and Telecommunications industries, because of their specific features and/or because of the relatively small number of private firms concerned.

[9] Finally for our analysis of the productivity of computer use, we had also to do some data cleaning on the firm variables. We had thus to eliminate about 15 % of firms from our matched samples, either because they had missing number of employees, negative or nil value added, gross book value of fixed assets or total labor compensation, or because they were extreme outliers in terms of the ratios of value added, gross book value of fixed assets and total labor compensation, relative to the number of employees. This is not much for such firm microdata bases as ours. The numbers just given in the text are for the samples before cleaning, the corresponding numbers for the cleaned samples being : 2 815, 2 612 and 2 533 firms (and 4845, 4093 and 3987 interviewed employees) in 1987, 1991 and 1993 respectively. The regression estimates of the impact parameters of computer use on firm productivity (and other characteristics) in the next section (tables 5 to 7) are based on the cleaned samples (the overall samples and the restricted ones to be defined), while the descriptive statistics on computer use given in this section (tables 1 to 4 and figures 1 and 2) are computed from the larger samples before cleaning. The descriptive statistics computed from the cleaned samples are, however, practically identical.

Otherwise, it is also important to note that for the firms in banking and insurance, we thought appropriate to modify the standard definition of value added (sales minus purchases plus net inventory changes) by adding financial benefits and subtracting financial costs.

[10] The distinction between the workers who use a computer in their main task or occasionally was only made in the 1987 survey, these categories being lumped together in the two other surveys (1991 and 1993). In fact, the experiments we did, trying to take this distinction into account in explaining firm performances, showed that we learn little from it and that indeed it was better to aggregate the two categories than to keep the only one of main users. This is also coherent with the fact that we use the information at the firm level and not at the job level.

[11] In 1987 and 1991, based on the estimates of computer use with and without word processors, the discrepancy at the sector level would have been of about 1 % to 3 % at most.

[12] The 1987 survey (but not the 1991 and 1993 ones) also ask whether the workers use computer listings or perform data entry operations. Experimenting with these two questions gave results which were mostly in line with the ones based on the use of personal computer and computer terminal. Since these results did not add much (and since we could not replicate them for 1991 and 1993), we do not report on them.

[13] The numbers are close to the ones found in the U.S. Current Population Survey (CPS), which in some years includes a question on computer utilization in the work place. The CPS is based on a sample of approximately 55,000 households, and the question on computer usage ("Did you use a computer at work?") is comparable to what we obtain here by pooling the two questions on PC and CT from TOTTO ("Do you use a PC (even occasionally)?" and "Do you use a terminal linked to a computer (even occasionally)?"). As noted in footnote 11, however, the question concerning PC in TOTTO for 1993, contrary to 1987 and 1991, includes word processors ("a PC or a word processor"), which creates some discrepancy. The overall figure for the U.S. in 1993 is 46%, as compared to 43% for France in the same year, and 37% in 1989 as compared to 38% in 1991. In this respect, U.S. is leading France by a margin of only about two years. Note, however, that the U.S. overall estimates, contrary to ours for France, include agriculture, coal mining and petroleum, construction, transportation and telecommunications, utilities and public administration. The U.S. numbers are also quite close to the French ones in manufacturing and services: respectively 44% and 48% in 1993; they appear significantly smaller in commerce and banking and insurance: 37% and 79% in 1993 (as against 47% and 89% for France in the same year). It is also interesting to see that the figure for "management and professional" for the U.S. in 1993 is 68%, the same as the one for France (see Table 3 below). All the numbers from the U.S. Current Population Survey are taken from William Lehr and Frank Lichtenberg (1996).

[14] The distribution of the average number of hours per week working with a PC or that working with a CT are very similar, with the same overall average in 1993 of 18 hours. These distributions are also not too different across sectors. Bank and insurance, and services have the longer hours (with an overall average of respectively 22 hours and 20 hours in 1993) while

food products, intermediate goods and equipment goods have the shorter ones (with the same overall average of 16 hours in 1993).

[15] One has to be careful that this does not mean that information technology has diffused much less in production than in other areas of the firm. As a matter of fact, if information technologies take the form of PCs or CTs in office work, they are often directly incorporated in automated machines such as robots or numerically controlled machine tools on the shop floor. TOTTO also includes some questions on the use of that modern equipment, but since they are specific of the production process in manufacturing firms, we preferred, after some experiments, not to consider them in this study.

[16] This indicator is defined as being equal to 1 if there is at least 1 interviewed employee in the firm who declares either that he uses a computer himself or that he knows that other people in the firm are using them. Else, its value is 0.

[17] If we distinguish between PC's and CT's, the average proportions by sector for the PC using firms remain in the high range of 80 % to 100 %, but for the CT using firms they vary, as could be expected, in a significantly lower range : from 60% to 70%, with bank and insurance as an outlier at 95%. For the sake of curiosity, one may also note that the average proportions computed for the medium and large firms, or based on the answers of the blue collars, white collars or managers separately, all tell the same story : computers are ubiquitous, being now present in every firm or almost.

[18] One must be careful in interpreting these estimates : they measure the extent to which the different tasks are performed by the computer users in the firm, but not the extent to which these tasks are computerized in the firm. In order to derive the latter distribution from the former, one should take into account the differences in the proportions of the total numbers of employees who are working in these different tasks (whether or not they use computers). The two distributions would be close only if these numbers were of comparable magnitude.

[19] The correlations between tasks we consider here are computed at the firm level, which we prefered to do here. Actually in our case they differ very little from the correlations computed at the employee level, since we have only one interviewed employee for most of our firms. Although these correlations do not change much between the 1987 and 1993 surveys, it is worth noting that some new links seem to appear between the three groups of tasks, in particular between the two first ones relative to production and accounting. These links are

still quite frail but could be viewed as reflecting the diffusion and integration of information technologies in the firm that have been stressed in a number of case studies.

[20] As previously, the construction of these subsamples is done by randomly selecting (whenever there is a choice), the two or three interviewed employees which are kept.

[21] Since the individuals surveyed in TOTTO are randomly drawn, the probability that they would be employed in a given firm is increasing with its size. Actually, the median sizes of the firms in our restricted samples (or subsamples) with one, two or three interviewed employees are respectively about 120, 330 and 730 employees.

[22] In fact, for the restricted sample with one interviewed employee which we consider mainly (and which by construction is identical to the original complete sample for three fourth of the firms and differs only as concerns the computer use variable for the remaining one fourth), the loss of precision of the estimates appears to be negligible. This loss is, however, substantial for the restricted subsamples with two and three interviewed employees (and the corresponding ones with one interviewed employee). For that reason, as well as for the sake of parsimony, we only present the estimates based on these more restricted subsamples for the economy as a whole, and not the (much more imprecise) ones by sector. Note also that we could not realistically consider the much more restricted subsamples with four or more interviewed employees, since these samples were too small and the resulting estimates much too imprecise even for the whole economy.

[23] With the orders of magnitude involved, this approximation may be a rather poor one. However not doing it and thus performing non linear regressions, instead of linear ones, does not substantially change our results.

[24] We do not discuss here the two classical issues of misspecification usually stressed in the econometric literature on the estimation of production functions. The first is the omission of firm specific factors, say management ability, which may account for a major part of the very large heterogeneity in firm performances. When the available sample is a panel (which is not our case), the problem is usually treated by assuming that these factors are more or less "fixed", and by introducing in the model potentially correlated individual firm effects to proxy for them. For many reasons, however, this treatment may strongly exacerbate other misspecifications, such as errors of measurement, and often provide unreasonable and very

fragile estimates. The second problem is that of endogeneity of the factors of production, say labor, which may arise from the simultaneity of the firm decisions on production and factor demand (hiring and firing). It is typically dealt with by using factor lagged values and/or factor prices (lagged labor and wages) as instrumental variables in the estimation. However, this method also encounters serious difficulties in practice. For a general discussion of these issues, see Griliches and Mairesse (1995). For an example of a panel data study on computer productivity which give satisfactory estimates, even when using a specification with firm effects, see Erik Brynjolfsson and Lorin Hitt (1995) ; for another example, where these estimates were too poor to be reported, see Franck Lichtenberg (1995).

[25] Note that our first formulation ignoring capital also assumes that the elasticity of labor is unity (and that returns to scale are constant). Generalizing a bit and taking the elasticity of labor to be β, we could have written the regression : $\log(VAL/L)_i = (\beta - 1)\log L_i + \delta p_i + \varepsilon_i$ where our parameter of interest δ is also equal to $\beta\gamma$.

[26] This is to be viewed as a convenient way to interpret our results, since in fact we estimate together both total factor productivity (i.e. the factor elasticities α and β) and the impact of computer use on which we focus (i.e. the parameter $\delta = \beta\gamma$) by running the full production function regressions:

$$\log(VA/L)_i = \alpha \log(C/L)_i + (\alpha + \beta - 1)\log(L)_i + \beta\gamma p_i + \varepsilon_i$$

or

$$\log(VA/LW)_i = \alpha \log(C/LW)_i + (\alpha + \beta - 1)\log(LW)_i + \beta\gamma p_i + \varepsilon_i \ .$$

Note that we have also computed conventional TFP measures, that is assuming constant returns to scale (i.e., $\alpha + \beta = 1$) and the equality of elasticity of labor to the share of labor compensation value added (more exactly *a* equals the half sum of the firm share $(WL/VA)_i$ and the industry share $(WL/VA)_{ind}$). Our estimates of the computer use impact parameter using these measures do not differ much from the ones reported here, which result from running the full production function regressions.

[27] Note that our capital measure, to the extent that it includes computer equipment, capture in part the impact of computer use on productivity, thus raising a "double-counting" problem. While, by not controlling for capital, we overestimate the true impact of computer use (γ biased upward), by controlling for it we tend to underestimate it to some degree (δ biased

downward). These issues of biases and interpretations are very similar to that of R-D "double counting" in the econometric studies of R-D productivity (for a discussion and some empirical evidence, see Cunéo-Mairesse, 1984).

Note also that the magnitude and exact meaning of our estimated impact parameters γ or δ may be somewhat affected by the fact that our gross book value measure of physical capital is limited to computer equipment ("hardware" expenditures) and ignores the software expenditures, which are a very large part of computer capital.

[28] For more details on this approach to explore the possible labor quality biases in estimating production functions, see Griliches-Ringstad (1971).

[29] In fact as before this is mainly a convenient way of explaining and presenting our results, not the way we computed them (see footnote 26).

[30] For such a study for the U.S., see Alan Krueger (1993), and for France, see Horst Entorf and Francis Kramarz (1994). These two studies, however, are performed at the employee level (not the firm level), the one for France using also the 1987 TOTTO survey.

[31] Assuming constant shares adding up to 1 for labor and capital incomes (or costs) in value added, and with no differences in wages and product prices among firms (or equivalently with labor measured by total wages and production by value added), differences in log TFPA and in log EBE/C can be shown to be proportional: $\Delta \log(\text{TFPA}) = \alpha \Delta \log(\text{EBE/C})$ where the capital elasticity α equals the complement to one of the labor share.

[32] The corrected t-ratio, $\hat{t}$ is precisely equal to $\hat{t}/((1-\lambda)(1+s))^{1/2}$ where $\hat{t}$ is the least squares t-ratio and where $s = (\lambda\hat{t}/(1-\lambda))^2/N$, is negligibly small when the size sample N is large enough.

[33] Since p_{ih} is the binomial $B(p_i^*, n)$, $\sigma_i^2 = p_i^*(1-p_i^*)$ and thus:

$$n\mathrm{Var}(e_i) = \sigma_{(w)}^2 = E_i[p_i^*(1-p_i^*)] = \bar{p}^* - (\mathrm{Var}(p_i^*) + \bar{p}^{*2}) = \bar{p}^*(1-\bar{p}^*) - \mathrm{Var}(p_i^*) \ .$$

[34] It is easy to verify that we get the obvious results we expect in the extreme example where the true p_i^* is the same for all firms ($p_i^* = p$), and in the other extreme example where the true p_i^* is either 1 or 0 ($p_i^* = 1$ for a proportion π of the firms, and $p_i^* = 0$ for the remaining proportion $(1 - \pi)$ of them). In the first case the true variance $\mathrm{Var}(p_i^*) = 0$ and thus $\lambda = 1$, whatever n (while $\mathrm{Var}(p_i) = \mathrm{Var}(e_i) = p(1-p)/n$). In the second case, necessarily $p_i = p_i^*$ (= 0

or 1), and the observed variance is the true one, and thus $\lambda = 0$, even for $n = 1$ (with $Var(p_i) = (Var(p_i^*) = \pi(1-\pi))$.

[35] $\bar{p}$ and $V\hat{a}r(p_i)$ are consistent estimates of $\bar{p}^*$ and $Var(p_i)$ and from the formulas in the text, we have for n larger than 1:

$$Var(p_i^*) = [nVar(p_i) - \bar{p}^*(1-\bar{p}^*)]/(n-1)$$

$$\lambda = [\bar{p}^*(1-\bar{p}^*) - Var(p_i)]/[(n-1)Var(p_i)]$$

The computation of the relative bias λ given here strictly applies to the case of the simple linear regression. However, it can easily be extended to the multiple regression case. We just have to divide the λ obtained for the simple regression, say $\lambda^{(s)}$, by the ratio of the partial or net variance of p_i, which remains unexplained by the $\lambda = \lambda^{(s)}/(1 - r^2)$ other variables in the (multiple) regression considered, to its total variance; that is,, where r^2 is the multiple correlation of the coefficient of the auxiliary regression of p_i on these other variables (for us the industry dummies, and also Log C/L and Log L for the TFP regression, or Log C/LW and Log LW for the TFPA regression). Note that this assumes, which is here most likely, that the sampling errors of measurement e_i are uncorrelated with these other variables.

[36] Note also that our computation of λ given here in the special case where p_i is the mean of a binomial extends to any variable p_i, whether discrete or continuous, measured as the firm mean of an employee variable. As long as we can compute an independent estimate of the overall within firm variance $\sigma^2_{(w)}$ that is for n larger than 1, we also have one of $\lambda = \sigma^2_{(w)}/nVar(p_i)$. The only difference in the present case is that $\sigma^2_{(w)}$ can be directly expressed in terms of $\bar{p}^*$ and $Var(p_i^*)$, or $\bar{p}^*$ and $Var(p_i)$.

[37] The ALS method (also known as the minimum distance estimator method) is a two-step procedure. It seems a natural method in many applications, and straightforward enough to implement here. In the first step it would be convenient to define the so-called restricted subsamples a little differently from what we did, by simply partitioning the overall samples in non-overlapping subsamples with one, two, three (and possibly more) interviewed employees. This would insure that the "auxilliary" parameters estimated separately for these subsamples will be independent (uncorrelated). These parameters are respectively the three mean

estimates $\overline{p}_{(1)}, \overline{p}_{(2)}, \overline{p}_{(3)}$ of the true proportion $\overline{p}*$, the two variance (over)estimates $\hat{Var}(p_i)_{(2)}$, $\hat{Var}(p_i)_{(3)}$ of the true variance $Var(p_i*)$ and the three least squares (under) estimates $\hat{\gamma}_{(1)}, \hat{\gamma}_{(2)}, \hat{\gamma}_{(3)}$ of the true impact parameter γ. In the second step, the three parameters of interest $\overline{p}*$, $Var(p_i*)$ and γ would be estimated from these eight auxiliary parameters, using as "estimating equations" the relations between them. These relations can be written here as:

$$\overline{p}_{(h)} = \overline{p}* \quad \text{for} \quad h = 1, 2, 3$$

$$\hat{Var}(p_i)_{(h)} = (1 - 1/h)Var(p_i^*) + [\overline{p}*(1 - p*)]/h \quad \text{for} \quad h = 2, 3$$

$$\hat{\gamma}_{(h)} = \gamma \quad Var(p_i^*)/\left[(1 - 1/h)Var(p_i^*) + \overline{p}*(1 - p*)/h\right] \quad \text{for} \quad h = 1, 2, 3 .$$

For details on the ALS method and its implementation, see Christian Gouriéroux, Alain Monfort et Alain Trognon (1985), or chapter 9 of Christian Gouriéroux and Alain Monfort (1995).

[38] Note that our three samples for 1987, 1991 and 1993 do not practically overlap.

[39] To save space, we do not report on the elasticities of capital and of scale in the TFP and TFPA regressions. All the estimated capital elasticities are very significant, ranging from a high 0.40 in food products and banking and insurance to a low 0.15 in equipment goods (0.25 overall industries) for the TFP regressions, and from 0.35 in banking and insurance to 0.10 in equipment goods, commerce and services (0.15 overall industries) for the TFPA regressions. Nearly all the estimated scale elasticities are about equal to one, and imposing constant returns to scale would not have changed our results.

[40] The same also appears to be true for food products in 1993, but not in the two previous years 1987 and 1991.

[41] $Var(p_i)$, and a fortiori $Var(p_i*)$, are necessarily small for p_i and p_i* close to one. From the outset, it was thus to be expected that the results in banking and insurance would be poor ; one can view them, however, as providing a kind of minimum benchmark. Although we were tempted to exclude this industry from our overall regressions, we decided that our results would be more convincing if we did not. In fact, the number of observations in that industry being relatively small, the overall estimates excluding it are only marginally better than the ones including it, which are reported here in the tables. Note that our overall regressions do include industry dummies.

[42] We have experimented with other variants of the computer use indicator which are not reported in Table 6, and we found the same pattern of estimates for all of them. The indicator based on the answers of the white collars workers provides estimates which tend to be only slightly smaller than those obtained with the blue collar or management indicators. Similarly, the estimates corresponding to the use of computer in documentation, computed aided design or scientific calculation are slightly greater than those corresponding to their use in production and inventory activities (and in accounting and related tasks). Even the estimates corresponding to the information, only available in the 1987 TOTTO survey, on whether the interviewed employees use computer listings or perform data entry operations, are roughly in line with the ones considered here.

[43] This is, in fact, straightforward in simple regressions such as ours of a continuous variable y on a binary indicator I_i, which amounts to comparing the two means $\bar{y}^{(1)}$ and $\bar{y}^{(2)}$ for the two groups of observations such that I_i=1 and I_i=0. The influential observations are the ones corresponding to the most extreme values (highest and smallest) of y_i in these two groups respectively.

[44] To be precise, these regressions are ran on the restricted samples with one interviewed employee, in which we have eliminated the observations of our (log) outcomes variables deviating from their means by more than 2.3 standard deviations within industry and in the two groups of computer using and not using firms. The outcome variables being roughly (log) normally distributed, this amounts to taking off about the one percent highest and the one percent smallest observations, within industry in the two groups, for each of them, and actually about 10 percent in total for all of them (some observations being outliers for several of them). Note that for the TFP variable we cleaned on our conventional measure (see end of footnote 26). Note also that our basic cleaning of the data left out only about 1 percent of the total number original number of observations (using a cut off point of 4.4 standard deviations on both side of the industry means for our outcome variables).

[45] It is easy to see that shrinking our computer use variable p_i towards the industry average p_{ind} proportionally to the ratio of its true variance to its observed variance (instead of taking the half sum), that is :

$$p_i^{(s)} = (1-\lambda)p_i + \lambda p_{ind} \quad \text{with} \quad (1-\lambda) = \mathrm{Var}(p_i^*)/\mathrm{Var}(p_i)$$

does exactly correct for the sampling errors downward biases in our estimates. Note that $p_i^{(s)}$ so defined, although it is a biased estimator of the true p_i^*, can be considered as the best one (and a better one than p_i), in the sense of the mean square error $E(p_i^s - p_i^*)^2$.

[46] See footnote 21.

[47] More precisely, we need the empirical variance of the computer use variable conditional on the other variables in the regression considered. See section 3.3 and footnote 35.

[48] These are the averages over the three year samples of the values found for λ in the case of the impact parameters estimated in the VA/L, C/L, W and EBE/C regressions. Our estimates of λ actually vary somewhat across the three year's samples. They are precisely the following, respectively with two and three interviewed employees: 66% and 70% in 1987; 75% and 54% in 1991; 76% and 56% in 1993. The corresponding estimates for the impact parameters estimated in the TFP and TFPA regressions are slightly larger by about 5% on average, and precisely the following: 70% and 77% in 1987; 75% and 64% in 1991; 77% and 64% in 1993.

[49] They are also very close to the consistent estimates we can retrieve for our restricted samples with only one interviewed employee. These are computed from our reference (downward biased) least squares estimates, and from the corresponding relative biases computed as $1 - Var(p_i^*) / Var(p_i)$, where $Var(p_i)$ is the observed variance of the computer use indicator and $Var(p_i^*)$ is the true variance as estimated on the restricted samples with two and three interviewed employees. We thus find average values of about 0.030 for $Var(p_i^*)$ and of about 80% for λ (with little difference between the VA/L, C/L, W and EBE/C regressions and the TFP and TFPA ones).

[50] The (log) standard deviations of VA/L, C/L, W and TFP are practically constant for our three year samples and respectively equal to: 0.50, 1.15, 0.35, 0.45.

[51] In addition to knowing whether the interviewed employees are using a computer at work or not, we know their job status (see section 2.2). One can think of taking advantage of this information to control directly for the differences in the firm general skills composition. We thus experimented by adding in our TFP regressions the proportions of blue collars, white collars and management persons among the interviewed employees (or since they add up to one, two of them plus the constant). For the restricted samples with one interviewed

employee, these proportions reduce to binary variables, just as our computer use indicator. We found that the computer use coefficients were indeed decreased but remained significantly positive.

Table 1 : Average proportion of computer users by industry in 1987, 1991 and 1993

	Personal computer			Computer terminal			PC or CT		
	1987	1991	1993	1987	1991	1993	1987	1991	1993
Food products	0.12	0.17	0.18	0.07	0.15	0.16	0.16	0.24	0.23
Intermediate goods	0.11	0.19	0.26	0.10	0.21	0.21	0.17	0.29	0.34
Equipment goods	0.18	0.27	0.33	0.16	0.24	0.27	0.27	0.39	0.42
Consumer goods	0.09	0.17	0.26	0.10	0.17	0.22	0.16	0.25	0.34
Commerce	0.14	0.21	0.30	0.14	0.29	0.33	0.25	0.40	0.47
Services	0.17	0.32	0.40	0.17	0.31	0.28	0.27	0.48	0.49
Bank and insurance	0.27	0.37	0.61	0.59	0.73	0.73	0.69	0.84	0.89
Total sample	0.14	0.24	0.32	0.16	0.26	0.28	0.25	0.38	0.43

The first two indicators are the average proportions of interviewed employees who use respectively a personal computer (PC) or a computer terminal (CT). The last one is the average proportion of those who declare using a PC **or** a CT (ie. the sum of the two first ones less the average proportion of interviewed employees who declare using **both** a PC and a CT).

Figure 1 : Distribution of the average number of hours per week spent working on a computer (PC or CT) by computer users for all sectors in 1987, 1991 and 1993

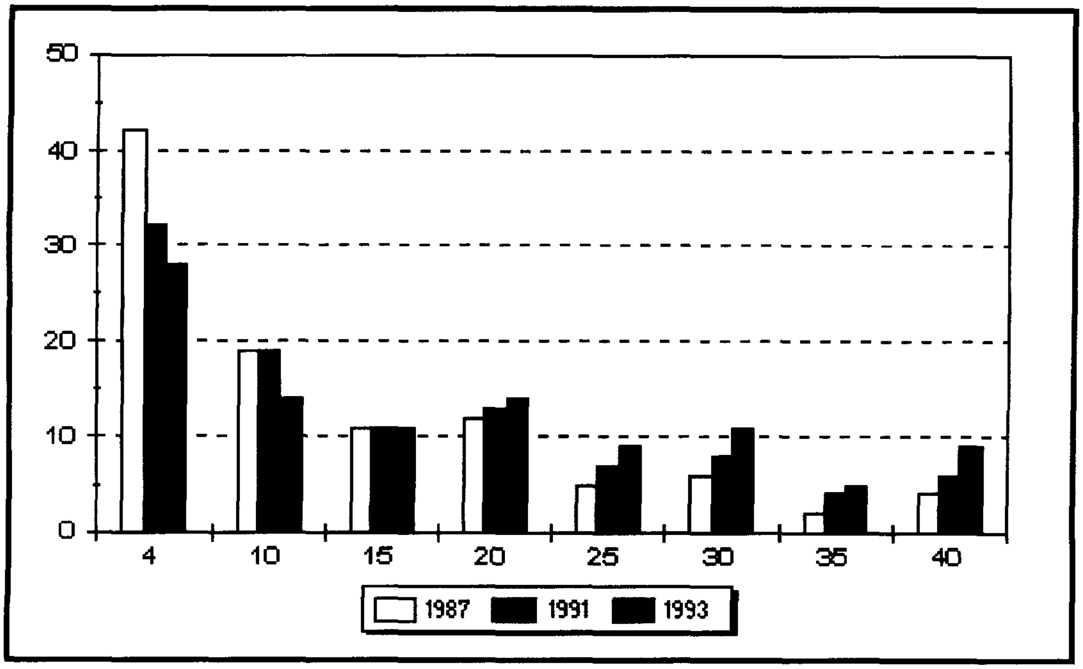


Table 2 : Average proportion of computer users (PC or CT) by industry and two size groups in 1987, 1991 and 1993

	Less than 500 employees			500 employees and more		
	1987	1991	1993	1987	1991	1993
Food products	0.11	0.20	0.14	0.26	0.35	0.40
Intermediate goods	0.14	0.26	0.30	0.28	0.40	0.48
Equipment goods	0.20	0.33	0.36	0.40	0.52	0.55
Consumer goods	0.14	0.23	0.30	0.24	0.33	0.46
Commerce	0.25	0.38	0.47	0.23	0.48	0.53
Services	0.27	0.47	0.47	0.28	0.53	0.46
Bank and insurance	0.53**	0.82*	0.89*	0.76	0.85	0.90
Total sample	0.21	0.35	0.39	0.36	0.49	0.55

* indicates that the estimated proportion is based on a number of interviewed that is less than 100 and/or a number of firms that is less than 50;
** indicates that the estimated proportion is based on number of interviewed employees that is less than 50 and/or a number of firms that is less than 25.

Table 3 : Average proportion of computer users (PC or CT) by industry and three job categories in 1987, 1991 and 1993

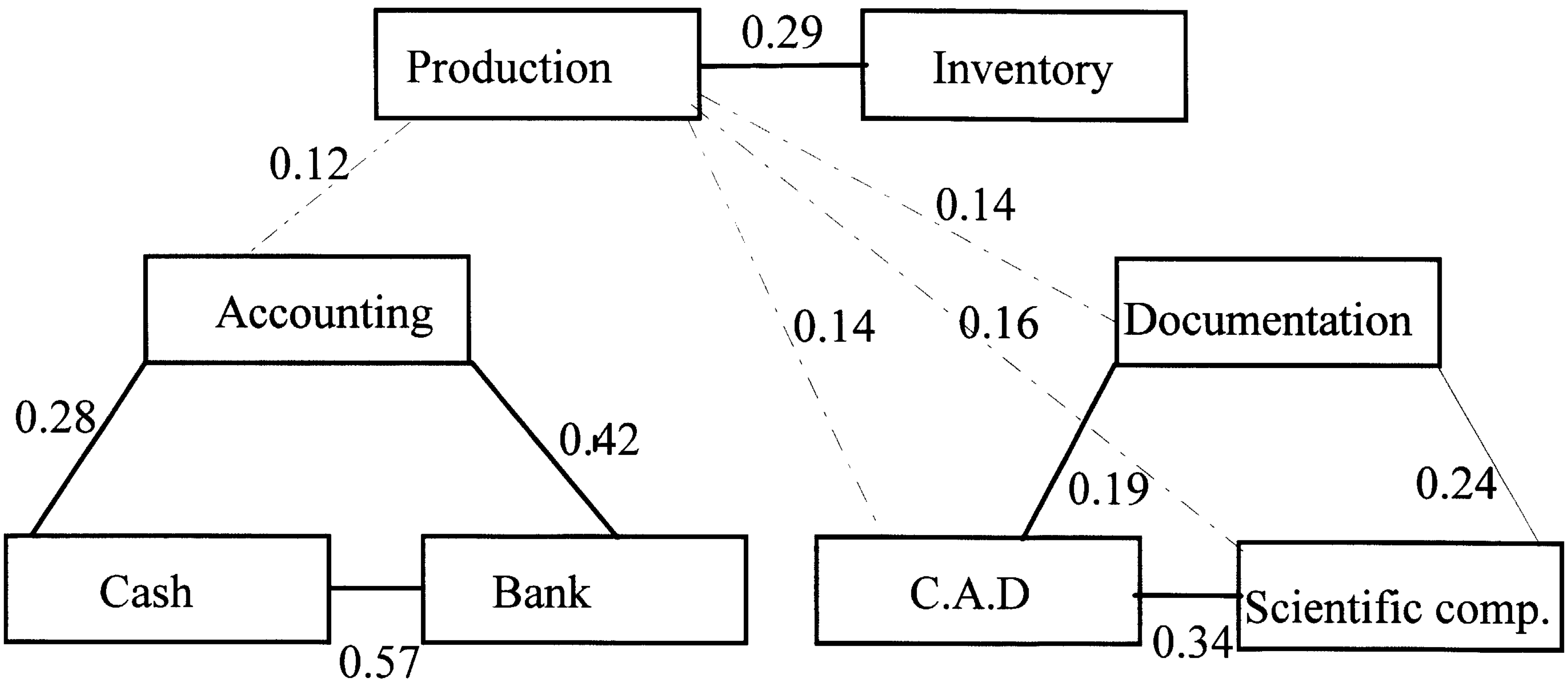
	Blue collar			White collar			Management		
	1987	**1991**	**1993**	**1987**	**1991**	**1993**	**1987**	**1991**	**1993**
Food products	0.04	0.11	0.13	0.43**	0.58**	0.34**	0.40*	0.52*	0.55*
Intermediate goods	0.06	0.13	0.17	0.58*	0.80*	0.83*	0.39	0.60	0.65
Equipment goods	0.09	0.15	0.18	0.54	0.76*	0.89*	0.53	0.71	0.70
Consumer goods	0.03	0.08	0.13	0.50*	0.60*	0.76*	0.39	0.55	0.62
Commerce	0.05	0.17	0.24	0.30	0.48	0.58	0.39	0.54	0.61
Services	0.05	0.12	0.07	0.27	0.54	0.47	0.50	0.71	0.73
Bank and insurance	0.19**	0.00**	0.00**	0.73	0.79	0.92	0.69	0.90	0.90
Total sample	0.06	0.12	0.16	0.41	0.59	0.65	0.46	0.64	0.68

*, **, see table 2.

Table 4 : Average frequency of tasks performed by computer users (PC or CT) by industry in 1993

	Production	Inventory	Accounting	Cash	Bank	Documentation	CAD	Scientific computation
Food products	0.45	0.45	0.24	0.07	0.08	0.38	0.09	0.18
Intermediate goods	0.51	0.43	0.22	0.05	0.06	0.39	0.11	0.16
Equipment goods	0.40	0.40	0.20	0.02	0.03	0.51	0.19	0.28
Consumer goods	0.37	0.33	0.32	0.07	0.10	0.40	0.12	0.15
Commerce	0.27	0.58	0.39	0.29	0.17	0.33	0.06	0.07
Services	0.29	0.17	0.41	0.09	0.11	0.48	0.17	0.16
Bank and insurance	0.31	0.10	0.53	0.25	0.56	0.59	0.06	0.15
Total sample	0.36	0.36	0.34	0.14	0.17	0.44	0.11	0.16

Figure 2 : Graph of the correlations of tasks performed by computer users (PC or CT) for all sectors in1993



The graph shows only the correlations of tasks that are significant at a 1% level of confidence. The correlations within the three main groups of tasks are represented by a full lines ; the correlations between groups are represented by dotted lines. These correlations are computed at the firm level, but do not differ pratically from the ones computed at the employee level (see footnote 18).

Table 5 : Estimated impacts of computer use on firm productivity and other characteristics by industry in 1987, 1991 and 1993 (samples restricted to one interviewed employee)

First panel

	VA/L			C/L			W		
	87	91	93	87	91	93	87	91	93
Food products (N = 178, 172, 167)	0.02 (0.13)	0.20 (0.08)	0.48 (0.09)	0.18 (0.22)	0.24 (0.17)	0.56 (0.16)	0.02 (0.08)	0.16 (0.05)	0.29 (0.05)
Intermediate goods (N = 434, 441, 572)	0.11 (0.05)	0.11 (0.04)	0.18 (0.03)	0.38 (0.12)	0.24 (0.09)	0.43 (0.08)	0.10 (0.03)	0.07 (0.03)	0.14 (0.02)
Equipment goods (N = 426, 468, 343)	0.12 (0.04)	0.16 (0.03)	0.14 (0.03)	0.24 (0.11)	0.31 (0.08)	0.29 (0.09)	0.09 (0.03)	0.14 (0.02)	0.11 (0.03)
Consumer goods (N = 470, 401, 314)	0.32 (0.07)	0.32 (0.05)	0.28 (0.06)	0.41 (0.14)	0.72 (0.11)	0.52 (0.12)	0.31 (0.05)	0.23 (0.04)	0.25 (0.04)
Commerce (N = 664, 568, 612)	0.14 (0.04)	0.10 (0.04)	0.16 (0.03)	0.17 (0.09)	0.14 (0.08)	0.09 (0.08)	0.13 (0.03)	0.08 (0.03)	0.14 (0.03)
Services (N = 514, 475, 405)	0.44 (0.06)	0.37 (0.05)	0.43 (0.05)	0.42 (0.17)	0.52 (0.12)	0.67 (0.15)	0.39 (0.05)	0.33 (0.04)	0.40 (0.04)
Bank-insurance (N = 129, 87, 120)	- 0.25 (0.13)	0.02 (0.18)	- 0.04 (0.19)	0.11 (0.15)	- 0.09 (0.32)	0.08 (0.35)	0.02 (0.03)	- 0.02 (0.08)	- 0.02 (0.07)
All industries (N= 2815, 2612, 2533)	0.18 (0.02)	0.20 (0.02)	0.24 (0.02)	0.28 (0.05)	0.34 (0.04)	0.37 (0.04)	0.18 (0.02)	0.17 (0.01)	0.20 (0.01)

Table 5 : Estimated impacts of computer use on firm productivity and other characteristics by industry in 1987, 1991 and 1993 (samples restricted to one interviewed employee)

Second panel

	EBE/C			TFP			TFPA		
	87	91	93	87	91	93	87	91	93
Food products (N = 178, 172, 167)	- 0.26 (0.23)	0.13 (0.12)	0.27 (0.14)	- 0.06 (0.09)	0.10 (0.05)	0.25 (0.07)	- 0.04 (0.07)	0.02 (0.04)	0.11 (0.05)
Intermediate goods (N = 434, 441, 572)	- 0.26 (0.10)	- 0.05 (0.08)	- 0.14 (0.08)	- 0.02 (0.04)	0.04 (0.03)	0.06 (0.02)	- 0.05 (0.03)	0.02 (0.02)	0.01 (0.02)
Equipment goods (N = 426, 468, 343)	- 0.03 (0.10)	- 0.04 (0.08)	- 0.14 (0.11)	0.07 (0.03)	0.09 (0.03)	0.09 (0.03)	0.02 (0.02)	- 0.01 (0.02)	0.01 (0.02)
Consumer goods (N = 470, 401, 314)	- 0.12 (0.14)	- 0.21 (0.10)	- 0.19 (0.11)	0.20 (0.06)	0.08 (0.04)	0.09 (0.05)	- 0.01 (0.03)	0.02 (0.03)	- 0.03 (0.03)
Commerce (N = 664, 568, 612)	0.04 (0.10)	0.06 (0.08)	0.09 (0.08)	0.11 (0.04)	0.08 (0.03)	0.14 (0.03)	0.01 (0.02)	0.01 (0.02)	0.02 (0.02)
Services (N = 514, 475, 405)	0.17 (0.12)	- 0.02 (0.09)	- 0.18 (0.11)	0.35 (0.05)	0.25 (0.04)	0.30 (0.04)	0.05 (0.02)	0.01 (0.02)	0.01 (0.02)
Bank-insurance (N = 129, 87, 120)	- 0.49 (0.19)	0.04 (0.33)	- 0.10 (0.27)	- 0.20 (0.12)	0.08 (0.15)	- 0.01 (0.12)	- 0.22 (0.11)	0.09 (0.15)	- 0.02 (0.11)
All industries (N= 2815, 2612, 2533)	- 0.06 (0.05)	-0.03 (0.04)	- 0.07 (0.04)	0.11 (0.02)	0.11 (0.02)	0.14 (0.02)	- 0.01 (0.01)	0.01 (0.01)	0.02 (0.01)

All the variables are in logs.

VA/L : Value added per employee
C/L : Capital to labor ratio
W/L : Average wage
EBE/C : Gross operating income to capital stock ratio
TFP : Total factor productivity
TFPA : Total factor productivity adjusted for labor quality
The TFP regressions control for C/L and L, and the TFPA regressions for C/LW and LW.
The regressions for all industries include seven industry dummies (6 + constant)

Table 6 : Estimated impacts of computer use on firm productivity and other characteristics for all industries in 1987, 1991 and 1993 (samples restricted to one interviewed employee) : sensitivity analysis according to different indicators of computer use

First panel

	VA/L			C/L			W		
	87	91	93	87	91	93	87	91	93
Reference (N = 2815, 2612, 2533)	0.18 (0.02)	0.20 (0.02)	0.24 (0.02)	0.28 (0.05)	0.34 (0.04)	0.37 (0.04)	0.18 (0.02)	0.17 (0.01)	0.20 (0.01)
Personal computer (N = Ref.)	0.16 (0.03)	0.19 (0.02)	0.22 (0.02)	0.15 (0.07)	0.30 (0.05)	0.30 (0.05)	0.18 (0.02)	0.16 (0.02)	0.20 (0.01)
Computer terminal (N = Ref.)	0.29 (0.03)	0.17 (0.02)	0.19 (0.02)	0,36 (0.06)	0.37 (0.05)	0.33 (0.05)	0.19 (0.02)	0.14 (0.02)	0.16 (0.02)
Blue collar (N = 1523, 1351, 1278)	0.26 (0.06)	0.20 (0.03)	0.20 (0.03)	0.73 (0.14)	0.39 (0.08)	0.51 (0.08)	0.22 (0.04)	0.16 (0.02)	0.13 (0.02)
Management (N =1 042, 1 045, 1025)	0.14 (0.03)	0.17 (0.03)	0.20 (0.03)	0.20 (0.07)	0.35 (0.06)	0.36 (0.08)	0.16 (0.02)	0.13 (0.02)	0.19 (0.02)
Production (N = Ref.)	0.17 (0.03)	–	0.13 (0.02)	0.31 (0.06)	–	0.29 (0.05)	0.15 (0.02)	–	0.10 (0.02)
Accounting (N = Ref.)	0.11 (0.03)	–	0.08 (0.03)	0.20 (0.06)	–	0.24 (0.06)	0.13 (0.02)	–	0.09 (0.02)
Excluding extreme obs. (N = 2 531, 2 375, 2276)	0.18 (0.02)	0.18 (0.02)	0.24 (0.02)	0.26 (0.04)	0.30 (0.04)	0.34 (0.04)	0.17 (0.01)	0.15 (0.01)	0.20 (0.01)

Table 6 : Estimated impacts of computer use on firm productivity and other characteristics for all industries in 1987, 1991 and 1993 (samples restricted to one interviewed employee) : sensitivity analysis according to different indicators of computer use

Second panel

	EBE/C			TFP			TFPA		
	87	91	93	87	91	93	87	91	93
Reference **(N = 2815, 2612, 2533)**	- 0.06 (0.05)	- 0.03 (0.04)	- 0.07 (0.04)	0.11 (0.02)	0.11 (0.02)	0.14 (0.02)	- 0.01 (0.01)	0.01 (0.01)	0.02 (0.01)
Personal computer **(N = Ref.)**	- 0.01 (0.06)	- 0.02 (0.04)	- 0.04 (0.04)	0.12 (0.03)	0.11 (0.02)	0.15 (0.02)	- 0.02 (0.02)	0.01 (0.01)	0.01 (0.01)
Computer terminal **(N = Ref.)**	- 0.09 (0.06)	- 0.07 (0.04)	- 0.06 (0.06)	0.11 (0.02)	0.07 (0.02)	0.10 (0.02)	- 0.02 (0.02)	- 0.01 (0.01)	0.01 (0.01)
Blue collar **(N = 1523, 1351, 1278)**	- 0.29 (0.12)	- 0.01 (0.07)	- 0.17 (0.07)	0.08 (0.04)	0.10 (0.03)	0.06 (0.02)	- 0.02 (0.03)	0.01 (0.02)	0.02 (0.01)
Management **(N =1 042, 1 045, 1025)**	- 0.08 (0.07)	- 0.07 (0.06)	- 0.11 (0.07)	0.10 (0.03)	0.10 (0.03)	0.13 (0.03)	- 0.02 (0.02)	0.02 (0.02)	- 0.01 (0.02)
Production **(N = Ref.)**	- 0.04 (0.06)	_	- 0.09 (0.05)	0.10 (0.02)	_	0.05 (0.02)	- 0.00 (0.02)	_	0.00 (0.01)
Accounting **(N = Ref.)**	- 0.04 (0.06)	_	- 0.14 (0.05)	0.06 (0.02)	_	0.02 (0.02)	- 0.03 (0.02)	_	- 0.02 (0.01)
Excluding extreme obs. **(N = 2 531, 2 375, 2276)**	- 0.04 (0.04)	- 0.02 (0.03)	- 0.02 (0.03)	0.11 (0.02)	0.10 (0.01)	0.15 (0.01)	0.00 (0.01)	0.01 (0.01)	0.02 (0.01)

Footnotes : See Table 5. The computer use indicators for production and related tasks, and for accounting and related tasks cannot be calculated in 1991.

Table 7 : Estimated impacts of computer use on firm productivity and other characteristics for all industries in 1987, 1991 and 1993 : analysis of the biases due to the sampling errors in the indicator of computer use

First panel

	VA/L			C/L			W		
	87	91	93	87	91	93	87	91	93
Reference (N = 2815, 2612, 2533)	0.18 (0.02)	0.20 (0.02)	0.24 (0.02)	0.28 (0.05)	0.34 (0.04)	0.37 (0.04)	0.18 (0.02)	0.17 (0.01)	0.20 (0.01)
Count ≥ 2, NS=1 (N=706, 554, 565)	0.20 (0.05)	0.21 (0.04)	0.23 (0.04)	0.31 (0.09)	0.31 (0.08)	0.29 (0.09)	0.20 (0.03)	0.14 (0.03)	0.17 (0.03)
Count ≥ 3, NS=1 (N=324, 251, 237)	0.14 (0.06)	0.25 (0.05)	0.21 (0.06)	0.29 (0.13)	0.39 (0.12)	0.15 (0.14)	0.14 (0.04)	0.21 (0.03)	0.20 (0.04)
Proportion (N = Ref.)	0.21 (0.02)	0.23 (0.02)	0.27 (0.02)	0.35 (0.06)	0.37 (0.04)	0.44 (0.05)	0.22 (0.02)	0.19 (0.01)	0.23 (0.01)
Count ≥ 2, NS=2 (N=706, 554, 565)	0.34 (0.06)	0.32 (0.05)	0.41 (0.05)	0.53 (0.12)	0.55 (0.11)	0.55 (0.12)	0.33 (0.04)	0.26 (0.03)	0.34 (0.03)
Count ≥ 3, NS=3 (N=324, 251, 237)	0.39 (0.09)	0.42 (0.08)	0.56 (0.09)	0.69 (0.20)	0.77 (0.17)	0.57 (0.20)	0.35 (0.06)	0.39 (0.05)	0.47 (0.05)
Proportion shrunk to sector average (N = Ref.)	0.35 (0.05)	0.40 (0.04)	0.47 (0.04)	0.57 (0.10)	0.69 (0.08)	0.75 (0.09)	0.36 (0.03)	0.33 (0.03)	0.40 (0.03)

Table 7 : Estimated impacts of computer use on firm productivity and other characteristics for all industries in 1987, 1991 and 1993 : analysis of the biases due to the sampling errors in the indicator of computer use

Second panel

	EBE/C			TFP			TFPA		
	87	91	93	87	91	93	87	91	93
Reference (N = 2815, 2612, 2533)	- 0.06 (0.05)	-0.03 (0.04)	-0.07 (0.04)	0.11 (0.02)	0.11 (0.02)	0.14 (0.02)	-0.01 (0.01)	0.01 (0.01)	0.02 (0.01)
Count ≥ 2, NS=1 (N=706, 554, 565)	- 0.05 (0.09)	0.07 (0.07)	0.03 (0.08)	0.11 (0.04)	0.13 (0.03)	0.15 (0.03)	- 0.02 (0.03)	0.05 (0.02)	0.05 (0.02)
Count ≥ 3, NS=1 (N=324, 251, 237)	- 0.22 (0.12)	- 0.05 (0.11)	0.10 (0.13)	0.07 (0.05)	0.14 (0.05)	0.16 (0.05)	- 0.02 (0.04)	0.01 (0.04)	0.02 (0.04)
Proportion (N = Ref.)	-0.10 (0.05)	-0.01 (0.04)	-0.09 (0.04)	0.13 (0.02)	0.14 (0.02)	0.16 (0.02)	-0.03 (0.06)	0.01 (0.01)	0.02 (0.01)
Count ≥ 2, NS=2 (N=706, 554, 565)	- 0.10 (0.11)	- 0.02 (0.10)	-0.03 (0.11)	0.18 (0.05)	0.18 (0.04)	0.26 (0.04)	- 0.03 (0.03)	0.03 (0.03)	0.06 (0.03)
Count ≥ 3, NS=3 (N=324, 251, 237)	- 0.30 (0.20)	- 0.15 (0.15)	0.10 (0.19)	0.22 (0.08)	0.20 (0.07)	0.39 (0.07)	- 0.01 (0.06)	- 0.03 (0.05)	0.08 (0.06)
Proportion shrunk to sector average (N = Ref.)	-0.11 (0.09)	-0.05 (0.08)	-0.14 (0.08)	0.22 (0.04)	0.23 (0.03)	0.29 (0.03)	-0.03 (0.03)	0.02 (0.02)	0.03 (0.02)

Footnotes : See Table 5.

Appendix : Number of firms and of interviewed employees by sectors, size and job categories in 1987, 1991 and 1993

	1987		1991		1993	
	Firms	Employees	Firms	Employees	Firms	Employees
Food products	199	351	212	307	203	315
Intermediate goods	471	934	509	853	651	1114
Equipment goods	502	1177	551	1098	415	788
Consumer goods	529	788	480	651	368	482
Commerce	727	1036	679	902	719	967
Services	601	749	617	747	519	704
Bank and insurance	161	406	129	339	177	418
Less than 500 employees	2379	3567	2401	2841	2236	2647
500 employees and more	811	1874	776	2056	816	2141
Blue collar	1674	2649	1594	2270	1492	2148
White collar	868	1166	820	1058	797	1058
Management	1130	1626	1196	1569	1169	1582
Total sample	3190	5441	3177	4897	3052	4788

Note that the number of firms in the case of the breakdown by job categories do not add up to their numbers in the total sample, but to larger numbers, because of the firms with two or more interviewed employees belonging to different job categories.