Discussion Paper No. 02-42

What's Spurious, What's Real? Measuring the Productivity Impacts of ICT at the Firm-Level

Thomas Hempell

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 02-42

What's Spurious, What's Real? Measuring the Productivity Impacts of ICT at the Firm-Level

Thomas Hempell

Download this ZEW Discussion Paper from our ftp server:

ftp://ftp.zew.de/pub/zew-docs/dp/dp0242.pdf

Das Papier entstand im Rahmen des Forschungsprojekts "Produktivitäts- und Spillovereffekte von IKT als General Purpose Technology" im Auftrag der Landesstiftung Baden-Württemberg.

The paper was written as part of the research project "Productivity and Spillover Effects of ICT as a Gerneral Purpose Technology" commissioned by the Landesstiftung Baden-Württemberg foundation.

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

Non–technical summary

The rapidly increasing investment in information and communication technologies (ICT) and the fast diffusion of the internet during the past decade have entailed widespread hopes about a 'New Economy' ensuring large productivity gains and persistent output growth. Only the more recent economic downturn and the breakdown of once highly praised businesses have put these hopes into perspective.

In order to get a more robust picture of the productivity effects of ICT, the potential insights from using aggregate statistics have turned out to be limited. The growth of real output of the most intensive ICT–using industries, like the service sector, is frequently understated by official statistics due to problems in accounting for quality changes appropriately. Moreover, aggregate statistics contain little information about complementary efforts by firms, like organizational changes and process re-engineering, which have been found to be important accompanying efforts for a productive use of the new technologies.

Consequently, the empirical literature on the productivity impacts of ICT has been increasingly focussing on evidence at the firm–level. Since the mid 1990s, most of these studies have found evidence of significant productivity contributions of ICT. The quantitative results of these studies, however, differ to a large extent. These differences are not only due to varying samples of firms and to diverse definitions of ICT capital but also due to differences in the quantitative techniques that have been employed.

In this paper, the importance of choosing the right methodological approach is explored in more detail. A variety of interfering factors like differing management abilities, measurement errors, simultaneity of input and output decisions by firms as well as business cycles may lead to distortions in the quantitative results. These effects are illustrated by applying different econometric techniques to a representative sample of observations from German service firms over the period from 1994 to 1999. The empirical analysis yields evidence that, once all the mentioned interfering influences are controlled for, ICT is found to have enhanced productivity in German services. However, these effects are substantially smaller than those obtained in various existing studies on the topic.

What's Spurious, What's Real? Measuring the Productivity Impacts of ICT at the Firm–Level

Thomas Hempell

ZEW Mannheim

first version: July 2002
this version: September 2003

Abstract

In order to assess the productivity effects of information and communication technologies (ICT), regressions based on cross–sectional firm–level data may yield unreliable results for the commonly employed production function framework. In this paper, various estimation biases and econometric strategies to overcome their sources are discussed. The effects are illustrated on the basis of a representative set of panel data for German service firms covering the period 1994 to 1999. The application of a suited SYS–GMM estimator yields evidence for significant productivity effects of ICT. However, these are substantially smaller than those suggested by cross–section estimates.

Keywords: Productivity, Information and Communication Technologies, Production Function Estimation, Panel Data, Services

JEL–Classification: C23, C81, D24, L80

Address: ZEW (Centre for European Economic Research)
Research Group of Information and Communication Technologies
P.O. Box 10 34 43
D–68034 Mannheim
Germany
Phone: +49/621/1235–233
Fax: +49/621/1235–333
E-Mail: hempell@zew.de

Acknowledgement: I would like to thank Irene Bertschek, François Laisney, Georg Licht, Werner Smolny, Kevin Stiroh, Elke Wolf and Thomas Zwick for helpful discussions and comments and the MIP–team for providing me with the data. Moreover, I would like to thank two anonymous referees for their helpful comments and suggestions. This paper was written as part of the research project "Productivity and Spillover Effects from ICT as a General Purpose Technology" commissioned by the Landesstiftung Baden–Württemberg foundation.

1 Introduction

Since the end of the 1990s, a broad variety of empirical studies has emerged exploring the productivity impacts of ICT at the firm–level.[1] Most of the studies employ a production function framework to estimate the elasticity of output with respect to ICT capital, controlling for the amount of other inputs. The quantitative results from these studies, however, vary substantially. Apart from varying definitions of ICT stocks and sample–specific variations, a substantial part of these differences may be due to differing quantitative methods and model specifications. In particular, interferences from firm–specific effects, simultaneity of input and output decisions, measurement errors, the omission of worker skills, autocorrelated productivity shocks or functional form restrictions in the underlying production function may induce biases in the empirical analysis. However, previous firm–level studies on ICT productivity address only some, if any, of these issues.[2]

The main aim of this paper is to explore the impacts of applying different quantitative approaches to firm–level data in more detail and to discuss econometric strategies that are suited to reveal the 'real' rather than 'spurious' productivity effects resulting from the use of ICT. The paper discusses why using firm–level data (as compared to more aggregate data sources) may help to control for biases arising from quality changes in output which are not accounted for by official price statistics. Moreover, calibration suggestions are derived in the paper about how existing firm–level survey data can be transformed for the purpose of production function estimates.

The empirical application illustrating the effects of applying different models and estimation techniques is based on a sample of more than 1100 firms from a representative survey in the German business–related and distribution service sector covering the period 1994 to 1999. The focus on services seems worthwhile for three main reasons. First, ICT investment has been most dynamic and most intensive in the service sector (e.g. OECD, 2000a). Second, business–related service have been the most important driver of economic growth over the last decades in industrialised countries (OECD, 2000b). Finally, assessing service quality correctly forms a particularly difficult issue in determining the productivity impacts from ICT (Griliches, 1994). Firm–level results that address this issue may be an insightful complement to findings from aggregate statistics. Beyond

[1]See for example studies by Bertschek and Kaiser (2001), Biscourp et al. (2002), Black and Lynch (2001), Bresnahan et al. (2002), Brynjolfsson and Hitt (1995, 1996, 1998, 2000, 2001), Brynjolfsson and Yang (1999), Greenan and Mairesse (1996), Greenan et al. (2001), Lehr and Lichtenberg (1999), Licht and Moch (1999), Lichtenberg (1995).

[2] Conducting a meta–analysis of results from 20 studies, Stiroh (2002b) shows that a substantial part of differing results in the literature can indeed be explained by differences in model specification, econometric techniques and underlying data sets. Moreover, he finds similar variations for alternative specifications and quantitative methods in own estimations for a single set of U.S. industry–level data.

analyzing the methodological issues, the study also aims to present evidence on the so far hardly explored productivity impacts of ICT use on German businesses.[3]

The results presented in this paper from a preferred system GMM approach yield evidence for significant productivity effects from ICT usage in German services. A one–percent increase in ICT capital is found to raise a firm's value added by 0.06 percent. This point estimate is substantially lower than values obtained from simple pooled OLS regressions and is overall robust with respect to varying parameters underlying the construction of capital stocks as well as to sample modifications. Among the various issues considered, unobserved heterogeneity between firms is found to be the most prominent interference in conventional estimates. Controlling for this interference by estimation in first differences, however, induces further problems that call for instrumental approaches.

The paper is organized as follows. Section 2 discusses the theoretical issues and introduces a basic production function framework with three extensions. Section 3 gives an overview of the employed data and describes the calibrations for constructing separate stock values for ICT and conventional capital. Section 4 discusses the econometric issues and presents empirical results. Section 5 summarizes the main findings.

2 Theoretical and Methodological Issues

In the empirical literature, the most frequently used framework for analyzing the productivity impacts of ICT has been to use a production function setup with ICT capital entering as a separate production input.[4] Many studies based on aggregate data determine the corresponding elasticities rather indirectly applying growth accounting approaches,[5] whereas firm–level (and sometimes industry–level) studies usually take advantage of the more numerous units of observations by directly estimating the elasticities in econometric approaches. In this section, some advantages of firm–level analyses are summarized. A Cobb–Douglas production function framework is then taken

[3]To the knowledge of the author, the only related studies for Germany are cross–section analyses by Licht and Moch (1999) and Bertschek and Kaiser (2001).

[4]The most frequently applied proxies for ICT capital applied are the value of computers installed, book values of office, computing and accounting machinery (OCAM) from balance sheets or investment in ICT.

[5]The growth accounting approach aims to assign the contribution of growth of different inputs to the overall growth of output. The residual in output growth that is not explained by the growth of the observed inputs is interpreted as a rise in multifactor productivity (MFP). The approach is based on the assumption of constant returns to scale and perfect competition, such that the elasticities of output with respect to the different inputs equal the income shares of the corresponding inputs. The direct growth contribution of ICT to output growth are calculated as the product of the share of ICT capital services in total income and the growth of ICT capital stock. Extending this approach, Stiroh (2002a) uses a difference–in–difference approach to assess potential spill–overs from ICT for U.S. industry data for 1983–99. He regresses productivity growth obtained from a growth accounting framework on a time dummy variable, a dummy variable denoting ICT–intensive firms and an interaction of both. His results yield little support for spill–overs from ICT.

as a reference model to discuss econometric issues and varying model specifications.

2.1 The scope of firm–level analysis

As pointed out by Brynjolfsson (1994) and Licht and Moch (1999), quality improvements — in particular improved customer service — are a prominent goal of ICT investment decisions. Similarly, Griliches (1994) suggests that the problem of unmeasured quality improvements in aggregate statistics is especially important in the case of 'unmeasurable' services like trade and F.I.R.E. (finance, insurance, real estate) where ICT investment has grown most rapidly. As a consequence, the contribution of ICT to real output growth inferred from aggregate data are likely to be understated. Suitably specified firm–level studies, by contrast, may suffer less from this measurement bias for two reasons.

First, as set out in the next section 2.2, micro–data sets allow to include time–specific industry dummy variables to make a firm's output directly comparable to its competitors. This helps to correct for potential measurement errors in industry price deflators.

Second, also variations in output quality between firms of the same industry and in the same period may be accounted for at the firm level. If a firm invests in ICT in order to improve the quality of its product and services (like extended shopping facilities or after sales support) while its competitors continue to offer their old products, the innovating firm will be able to charge a higher price for its new product and raise revenues. Brynjolfsson and Hitt (2000) argue that microeconomic studies will capture this effect and variations in output quality will contribute to measuring a higher output elasticity of ICT investment.[6] Appendix A shows that the production function estimates obtained from firm–level data may be interpreted as reduced–form estimates of coefficients for a model that implicitly takes into account productivity effects from quality improvements. Most importantly, a strong impact of ICT use on output quality will entail a higher estimate of the ICT coefficient in the production function.

However, there is also some limitations to estimating output elasticities based on firm–level data. In particular, Klette and Griliches (1996) show that varying prices at the firm–level due to imperfect competition may induce a downward bias on the estimated input elasticities.[7] This type of bias, however, affects the estimates of *all* inputs in a similar fashion such that this issue is not addressed in more detail in this paper.

Apart from these rather technical arguments, the firm–level approach offers a broad scope of insights that are much more difficult to obtain from aggregate data. Most importantly, the productivity impacts of ICT may vary between firms. Some firms

[6] This argument is backed by empirical support from a firm–level study by Brynjolfsson and Hitt (1995) who do not find any significant differences in IT productivity between "measurable" and "unmeasurable" sectors, indicating that appropriate quality measurement is mainly a problem at the aggregate level.

[7] See Appendix A for further details.

are better enabled than others to take productive advantage of new technologies. For the particular case of ICT, it has been argued that complementary factors like skills, innovations and organisational assets play a key role for ICT to unfold its benefits.[8] In industry– or country–level data, a large part of these firm–specific differences disappears in the process of aggregation and firm–level analysis is more appropriate for addressing these questions. Even though the investigation of complementary factors is beyond the scope of this paper, the issues discussed in this paper aim at contributing to finding suited methodological approaches to assess these questions.[9]

2.2 Reference framework

In the reference specification, output is assumed to be generated by a Cobb–Douglas technology with labour and two types of capital as inputs:

$$Y_{it} = F(A_{it}, L_{it}, ICT_{it}, K_{it}) = A_{it} L_{it}^{\gamma_1} ICT_{it}^{\gamma_2} K_{it}^{\gamma_3}, \tag{1}$$

where Y_{it} is value added of firm i in period t, L_{it} represents labour input, ICT_{it} and K_{it} are the corresponding amounts of ICT and conventional (non-ICT) capital respectively, and A_{it} is the multifactor productivity of firm i. After taking logs on both sides, eq. (1) can be rewritten as:

$$y_{it} = \gamma_1 l_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it} + \eta_i + \lambda_{j(i),t} + \epsilon_{it}, \tag{2}$$

where small letters denote the corresponding logarithmic values and multifactor productivity $\log(A_{it}) = \eta_i + \lambda_{j(i),t} + \epsilon_{it}$ is decomposed into a firm–specific fixed part η_i, a time–variant industry–specific part $\lambda_{j(i),t}$ (with $j(i)$ denoting the industry j that firm i is operating in), and a time–variant firm–specific residual ϵ_{it}. Firm–effect η_i captures fixed or quasi–fixed[10] factors affecting productivity, like management ability, organisational capital, branding or location. The residual ϵ_{it} comprises measurement errors (μ_{it}) and firm–specific productivity shocks (p_{it}) such that $\epsilon_{it} = \mu_{it} + p_{it}$. In this reference framework, both m_{it} and p_{it} are assumed to be serially uncorrelated and only their sum ϵ_{it} is considered.

The industry time–variant part $\lambda_{j(i),t}$ captures variations in productivity that are common to firms of a particular industry and that are left unexplained by the factors included in the model. In this sense, $\lambda_{j(i),t}$ helps to ensure that outputs of firms are more readily comparable across industries. In particular, demand fluctuations induced by industry–specific business cycles may lead to variations in the degree of factor utilisation

[8]See Bresnahan and Greenstein (1996), Brynjolfsson and Hitt (2000) and Yang and Brynjolfsson (2001).

[9]In Hempell (2002, 2003), the role of innovation and innovative experience as well as training of employees are investigated using the preferred SYS–GMM approach explored in this paper.

[10]The time span considered in the empirical analysis comprises a maximum of 6 years for each firm.

that are similar across firms of one industry. The resulting changes of productivity of firms operating in the corresponding industry are then captured by $\lambda_{j(i),t}$.

In a similar manner, $\lambda_{j(i),t}$ helps to correct for mismeasurement of prices at the industry–level. To illustrate this, define measured prices $\hat{P}_{j(i),t}$ for industry $j(i)$ as the product of true prices $P_{j(i),t}$ and an industry–specific measurement bias M_{jt} such that $\hat{P}_{j(i),t} = P_{j(i),t}M_{j(i),t}$ or $\hat{p}_{j(i),t} = p_{j(i),t} + m_{j(i),t}$ in logarithms.[11] With z_{it} denoting nominal output of firm i in t, the real output y_{it} of firm i operating in industry $j(i)$, is $y_{it} = z_{it} - p_{j(i),t} = z_{it} - \hat{p}_{j(i),t} + m_{j(i),t}$ and observed real output (i.e. output deflated with observed prices) is $\hat{y}_{it} = z_{it} - \hat{p}_{j(i),t} = y_{it} - m_{j(i),t}$. If — as argued above — ICT is most heavily used in industries for which product quality tends to be understated (and official prices are overstated consequently), ict_{it} and $m_{j(i),t}$ are positively correlated. The omission of $m_{j(i),t}$ will then lead to understating the true productivity contributions of ICT. Since this type of mismeasurement affects *all* firms of industry j at a given point in time t in the same way, the projection of output on a common dummy variable $\lambda_{j(i),t}$ helps to control for this measurement bias.

While the industry–specific component $\lambda_{j(i),t}$ will be controlled for by including time–variant industry dummies,[12] the potentially distorting effects from unobserved η_i and ϵ_{it} will be addressed by econometric techniques. In particular, I will account for the fact that both η_i and ϵ_{it} may be correlated with the inputs in general and ICT capital in particular. This may well be the case if, e.g., firms with a good management (i.e. a high η_i) are both more productive and more inclined to make use of ICT (in the following referred to as *firm effects*), or if a demand shock (high ϵ_{it}) raises both productivity as well as investments (*simultaneity issues*).

2.3 Extensions of the reference framework

In the following, the reference model (2) is extended by further aspects, by allowing for 1.) serial correlation of the errors ϵ_{it}, 2.) heterogenous labour inputs and 3.) a more flexible functional specification. At best, these issues would be considered simultaneously in the empirical analysis. Unfortunately, due to data limitations, this is not possible such that these extensions must be explored separately.

Extension 1: **Serially correlated residuals.** Potential biases in the econometric exploration of eq. (2) may arise if the productivity shocks p_{it} are serially correlated such that $p_{it} = \rho p_{i,t-1} + e_{it}$, with $e_{it} \sim i.i.d.$[13] This serial correlation may occur if, e.g.,

[11] Note that $M_{j(i),t} > 1$ and $\log(M_{j(i),t}) \equiv m_{j(i),t} > 0$ if the quality of output in industry $j(i)$ are understated such that measured prices $\hat{P}_{j(i),t}$ are higher than the true ones.

[12] Alternatively, these dummies can be conceived as interactions between time and industry dummies.

[13] This extension follows the framework investigated in Blundell and Bond (2000).

the effects from demand shocks may only be partially captured by the industry–specific control variables $\lambda_{j(i),t}$. Measurement errors μ_{it}, by contrast, are assumed to be serially uncorrelated. In order to estimate eq. (2) for this case, a dynamic or common factor representation can be obtained by subtracting $\rho y_{i,t-1}$ from both sides of eq. (2). Inserting $e_{it} = p_{it} - \rho p_{i,t-1}$ and rearranging yields:

$$\begin{aligned} y_{it} &= \rho y_{i,t-1} + \gamma_1 n_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it} \\ &\quad -\rho\gamma_1 n_{i,t-1} - \rho\gamma_2 ict_{i,t-1} - \rho\gamma_3 k_{i,t-1} \\ &\quad +\eta_{it}(1-\rho) + \lambda_{j(i),t} - \rho\lambda_{j(i),t-1} + w_{it}, \end{aligned} \tag{2a}$$

where $w_{it} = e_{it} + \mu_{it} - \rho\mu_{i,t-1}$ is MA(1). In order to obtain estimates of the structural coefficients γ_1, γ_2, γ_3 and ρ, a two–step procedure is applied. In the first step, the reduced–form model of the following form is estimated:

$$\begin{aligned} y_{it} &= \pi_1 y_{i,t-1} + \pi_2 n_{it} + \pi_3 ict_{it} + \pi_4 k_{it} \\ &\quad +\pi_5 n_{i,t-1} + \pi_6 ict_{i,t-1} + \pi_7 k_{i,t-1} \\ &\quad +\eta_{it}(1-\rho) + \lambda_{j(i),t} - \rho\lambda_{j(i),t-1} + w_{it}. \end{aligned} \tag{3}$$

In the second step, the underlying factor restrictions $\pi_1 = \rho$, $\pi_2 = \gamma_1$, $\pi_3 = \gamma_2$, $\pi_4 = \gamma_3$, $\pi_5 = -\gamma_1\gamma_2$, $\pi_6 = -\gamma_1\gamma_3$ and $\pi_7 = -\gamma_1\gamma_4$ can then be tested and imposed by a minimum–distance estimator.[14]

Extension 2: **Heterogenous labour.** In another version of model (2), heterogeneity in the quality of labour is considered. This may be important if, e.g., the use of ICT is most intensive in firms with a high share of high–skilled workers. Omitting heterogeneity in workers' skills may then lead to overstating the productivity impacts of ICT capital. A firm's workforce is decomposed into the number of employees that are high–skilled N_h (with university degree or equivalent), medium–skilled N_m (vocational training), and low–skilled N_l (no formal qualification) with $N_{it} = N_{l,it} + N_{m,it} + N_{h,it}$ denoting the total number of employees. Letting θ_h and θ_m denote the productivity differential of high and medium skilled workers compared to low–skilled workers, effective labour input L_{it} is:

$$\begin{aligned} L_{it} &= N_{l,it} + (1+\theta_m)N_{m,it} + (1+\theta_h)N_{h,it} \\ &= N_{it} \cdot (1 + \theta_m s_{m,it} + \theta_h s_{h,it}), \end{aligned} \tag{4}$$

with $s_{m,it} = N_{m,it}/N_{it}$ and $s_{h,it} = N_{h,it}/N_{it}$ denoting the shares of medium– and high–skilled employees in total workforce of the firms respectively.[15] With small values for θ_m, θ_h, $s_{m,it}$ and $s_{h,it}$, the term controlling for the skill structure may be simplified to:

[14] The details of this calculations are described in Appendix C.

[15] The main assumption underlying this approach is that qualification raises the productivity of workers by a fixed proportion. An alternative specification would be to let the three skill–groups enter the the production function as separate inputs with each having its own constant elasticity. This is equivalent to assuming that effective labour can be decomposed into $L = N_l^{\lambda_l} N_m^{\lambda_m} N_h^{\lambda_h}$. There are, however, two main drawbacks in this approach. First, from a theoretical point of view, this approach implies that

$$\log L_{it} = \log N_{it} + \log\left(1 + \theta_m s_{m,it} + \theta_h s_{h,it}\right) \approx n_{it} + \theta_m s_{m,it} + \theta_h s_{h,it}. \tag{5}$$

Inserting (5) into (2) then yields the model:

$$y_{it} = \gamma_1 n_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it} + \beta_1 s_{m,it} + \beta_2 s_{h,it} + \eta_i + \lambda_{j(i),t} + \epsilon_{it}. \tag{2b}$$

with $\beta_1 = \gamma_1\theta_m$ and $\beta_2 = \gamma_1\theta_h$.

The inclusion of skill–shares in the production function estimations as in eq. (2) is a very common way in the related literature in order to control for heterogeneity of labour quality.[16] However, anticipating some of the results and applying mean shares for s_m and s_h, the implicit products $\beta_1 s_{m,it} \cong 0.110$ and $\beta_2 s_{h,it} \cong 0.549$ yield rather high values that make the approximation very inaccurate. This measurement error is positively correlated with the skill measures and may induce a bias also in other regressors. In addition, I therefore also consider the more precise second–order Taylor approximation:

$$\begin{aligned}
\log L_{it} &= n_{it} + \log(1 + \theta_m s_m + \theta_h s_h) \\
&\approx n_{it} + \theta_m s_m + \theta_h s_h - \frac{1}{2}(\theta_m s_m + \theta_h s_h)^2 \\
&= n_{it} + \theta_m s_m + \theta_h s_h - \frac{1}{2}\theta_m^2 s_m^2 - \frac{1}{2}\theta_h^2 s_h^2 - \theta_m\theta_h s_m s_h. \qquad (5a)
\end{aligned}$$

The model resulting from inserting eq. (5a) into (2) is:

$$\begin{aligned}
y_{it} &= \gamma_1 n_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it} \\
&\quad + \beta_1 s_{m,it} + \beta_2 s_{h,it} + \beta_{11} s_{m,it}^2 + \beta_{22} s_{h,it}^2 + \beta_{12} s_{m,it} s_{h,it} + \eta_i + \lambda_{j(i),t} + \epsilon_{it}, \quad (2b')
\end{aligned}$$

where the additional parameters correspond to $\beta_{11} = -\frac{1}{2}\theta_m^2$, $\beta_{22} = -\frac{1}{2}\theta_h^2$, $\beta_{12} = -\theta_m^2\theta_h^2$. Apart from relying on a more accurate approximation of labour quality, eq. (2b') can also be used to explore the appropriateness of the underlying model for skills from eq. (4) by testing the validity of the imposed common factor restrictions for $\beta_1, \beta_2, \beta_{11}, \beta_{22}$ and β_{12}.[17]

Extension 3: **Flexible functional form.** As it is well–known, the coefficients γ_j in eq. (2) correspond to the elasticities of output with respect to the inputs j. One disadvantage of the Cobb–Douglas production function is, however, that the elasticities

each of the three inputs is regarded as an essential input for production in the sense that $Y = 0$ if $N_l = 0 \vee N_m = 0 \vee N_h = 0$. This seems to be a very restrictive assumption given that many firms (in particular small ones) produce output employing workers of only one or two of the three skill groups. By contrast, the specification of eq. (4) assumes that only the existence of one worker (independently of her qualification) is essential such that $Y = 0$ if $N_l = 0 \wedge N_m = 0 \wedge N_h = 0$. Second, from an empirical point of view, firms that do not employ workers from each of the three skill–groups would have to be excluded in the alternative approach (since the specification is in logs). For the given sample, this implies that more than half of the 578 firms for which information on skills is available would have to be excluded from the empirical analysis. This would not only lead to a much lower precision of the estimates but might also entail a serious selection bias.

[16] See, e.g., Lehr and Lichtenberg (1999), Caroli and van Reenen (2001) or Bresnahan et al. (2002).

[17] The calculations are analogue to the minimum–distance procedure described in detail in Appendix C.

of the individual inputs are restricted to be constant and the elasticity of substitution between the individual inputs is restricted to one. A more flexible specification is the translog–function (Christensen and Jorgenson, 1969) in which both the output elasticities and the elasticities of substitution may vary. The translog–extension of equation (2) is:

$$\begin{aligned} y_{it} &= \gamma_1 l_{it} + \gamma_{11} l_{it}^2 + \gamma_2 ict_{it} + \gamma_{22} ict_{it}^2 + \gamma_3 k_{it} + \gamma_{33} k_{it}^2 \\ &\quad + \gamma_{12} l_{it} ict_{it} + \gamma_{13} l_{it} k_{it} + \gamma_{23} ict_{it} k_{it} + \eta_i + \lambda_{j(i),t} + \epsilon_{it}. \end{aligned} \tag{2c}$$

To keep the model tractable for the empirics, I abstract from the skill level in this specification. The elasticity of output with respect to input j (α_j) depends on the levels of all the inputs employed. For comparability to the Cobb–Douglas framework, they may be evaluated at the means of the corresponding logarithmic values (denoted by a bar). The implicit mean elasticities are then given by:

$$\bar{\alpha}_L = \partial y_{it}/\partial l_{it} = \gamma_1 + 2\gamma_{11}\overline{l_{it}} + \gamma_{12}\overline{ict_{it}} + \gamma_{13}\overline{k_{it}} \tag{9}$$

$$\bar{\alpha}_{ICT} = \partial y_{it}/\partial ict_{it} = \gamma_2 + 2\gamma_{22}\overline{ict_{it}} + \gamma_{12}\overline{l_{it}} + \gamma_{23}\overline{k_{it}} \tag{10}$$

$$\bar{\alpha}_K = \partial y_{it}/\partial k_{it} = \gamma_3 + 2\gamma_{33}\overline{k_{it}} + \gamma_{13}\overline{l_{it}} + \gamma_{23}\overline{ict_{it}}. \tag{11}$$

3 The Data

To implement the production framework empirically, data from the *Mannheim Innovation Panel in Services* (MIP-S) are employed. This survey is conducted by the Centre for European Economic Research (ZEW) on behalf of the German Federal Ministry for Education and Research. The data have been collected annually since 1994 in a representative survey of innovation activities in the German business–related service and distribution sector and includes information from more than 2000 firms (Janz et al. **?**). It has an (unbalanced) panel structure in important key variables for the years since 1994. Among many other features, the data set contains annual data on sales, number of employees (full-time equivalents), skill structures, expenditures on gross investment and on ICT-capital (including hardware, software and telecommunication technology). Since similar information has been collected in various other existing data sets, too, some proceedings are discussed in the following of how information from other external sources may be used to suitably transform the survey data to variables that are applicable for a production function framework.

For output Y_{it}, I construct a measure of firms' value added. Alternatively, sales could be used if firm–level intermediate goods were included as an additional input. However, the data set does not contain information on firm–specific intermediate goods. Using sales for output, this might induce an omitted variable bias in the regressions since industries that operate rather at the end of the value chain (like wholesale and trade) rely more strongly on intermediate goods in quantitative terms than other industries do. In order to control for these differences and to deflate the corresponding output values, I calculate the shares of real value added in nominal gross output at the NACE two–digit industry level.[18] The firm–level data on sales are then multiplied by these industry–specific shares.[19] For labour input, the annual average of the number of employees in full-time equivalences is used.

A further issue concerns the separate construction of capital stocks for ICT capital and conventional (non–ICT) capital from investment data. For this purpose, investment on conventional capital is defined as total investment expenditures minus ICT expenditures as reported by the firms and is deflated by the producer price deflator for investment goods from the German Statistical office. For ICT goods, however, German official price statistics tend to understate the real price decline (Hoffmann, 1998). Therefore, the harmonized ICT price index for Germany calculated by Schreyer (2000) is applied. He employs official statistics on ICT prices in the U.S., which are based on hedonic techniques, as a reference and assumes that the differences between price changes for ICT and non–ICT capital goods are the same across OECD countries.

Given the deflated investments for both types of capital, the perpetual inventory method with constant, geometric depreciation is applied to construct the capital stocks for ICT and non-ICT. Accordingly, the capital stock K_{kt} of type k in period t results from investment $I_{k,t-1}$ in the following way:

$$K_{kt} = (1 - \delta_k)K_{k,t-1} + I_{k,t-1}, \tag{12}$$

with $k = 1$ for conventional (non–ICT) and $k = 2$ for ICT capital and investment and δ_k denoting the depreciation rates of the capital stocks.[20]

Since no information is available on the level of capital stocks, initial capital stocks are constructed employing the method proposed by Hall and Mairesse (1995).[21] Under the assumption that investment expenditures on capital good k have grown at a similar, constant average rate g_k in the past in all firms, equation (12) can be rewritten for period

[18] For this purpose, the time series 7711 and 7716 from the German Statistical Office are used.

[19] Let Z_{it} and Y_{it} be sales and value added of firm i in period t, and let $Z_{j(i),t}$ and $Y_{j(i),t}$ be sales and value added aggregated over all firms of the same industry $j(i)$ that firm i is operating in. Then the unknown value added of firm i is approximated by $Y_{it} \simeq Z_{it} \cdot Y_{j(i),t}/Z_{j(i),t}$.

[20] For conventional capital, the depreciation rates δ_1 by industries are calculated as the shares of capital consumption in net fixed assets evaluated at replacement prices (time series 7719 and 7735 of the German Statistical Office). The unweighed mean over all service industries amounts to 9% with a maximum in the NACE 72 (data processing) of 21% and a minimum in NACE 70 (real estate) with 2.2%. For ICT capital, a rate of $\delta_2 = 0.30$ is assumed. Relying on available data from the U.S. (Fraumeni, 1997; Moulton et al., 1999), depreciation rates for IT–hardware, software and telecommunication capital are 31.2% for IT–hardware, 55.0% for prepackaged software, 33.0% for custom and own–account software and 15.0% for telecommunication capital. Using data by EITO (2001) for the year 1999, total ICT investment expenditures in Germany consist of 47.0% for IT–hardware, 26.9% for software and 26.1% for end–user and network telecommunication equipment. Taking these market shares as weights, this yields an average depreciation rate of ICT capital of $\delta_1 = 0.312 \cdot 0.47 + (0.55 + 0.33)/2 \cdot 0.269 + 0.15 \cdot 0.261 = 0.304$.

[21] Hall and Mairesse (1995) refer to R&D stocks for which methodological problems are very similar.

$t = 1$ (1994) by backward substitution in the following way:[22]

$$\begin{aligned} K_{k1} &= I_{k0} + (1-\delta_k) I_{k,-1} + (1-\delta_k)^2 I_{k,-2} + \ldots \\ &= \sum_{s=0}^{\infty} I_{k,-s}(1-\delta_k)^s = I_{k0} \sum_{s=0}^{\infty} \left[\frac{1-\delta_k}{1+g_k}\right]^s \\ &= \frac{I_{k1}}{g_k+\delta_k}. \end{aligned} \tag{13}$$

In order to derive the initial capital stocks, assumptions about pre-period growth rates of both type of investments must be made. For non-ICT investment expenditures, I assume an annual growth rate of approximately 5% ($g_1 = 0.05$).[23] For ICT investment, no time series are available for Germany. In order to get a rough idea of the evolution of ICT investments during the last decades, U.S. data are referred to as a rough guideline. Jorgenson and Stiroh (1995) calculate an average annual growth rate of 44.3% in real computer investment and of 20.2% for OCAM (office, computing, and accounting machinery) between 1958 and 1992 for the U.S. Since the share of computers in OCAM has been rising continuously — reaching 94% in 1992 —, an annual pre-period growth rate close to the growth rate of computer investment of $g_2 = 0.4$ is assumed for ICT investment.[24] Since there are time lags between the installation and the productive contribution of capital goods, the capital stock at each period's *beginning* (or at the end of the corresponding previous period) are taken as measures for capital inputs.

In order to apply suited econometric techniques, only firms with consistent information on at least three consecutive periods available are included in the sample. The resulting unbalanced reference sample (denoted "full sample") consists of 1177 firms with a total of 4939 observations. The statistics of the sample are summarised in Table 7 in Appendix D. The majority of firms in the reference sample are small and medium–sized firms with a median of 42 employees. About 10% of the sample consists of large firms with more than 500 employees. Tables 10 and 11 show that, overall, the sample reflects industry and size structure of the German business–related and distribution services fairly well.[25] Finally, the last two columns of Table 7 report the (cross–sectional) means and medians of the firms' (longitudinal) averages of capital and output intensity (capital per employee) for the sample. The figures indicate that in the median firm of the sample, a workplace

[22]In fact, the initial value of investment for firm i $I_{ik,1}$ is replaced by the average of the observed values of investment such that $I_{ik,1} \simeq \sum_{t=1}^{T} I_{ik,t}$. With this "smoothing" it is aimed to correct for cyclical effects which might affect investments in different initial years in the unbalanced panel. Sensitivity analyses show that the results are hardly affected if true initial investments instead of 'smoothed' ones are used.

[23]Calculations on capital data provided by Müller (1998) show that gross capital stock in German services has grown on average by 4.8% annually between 1980 and 1991.

[24]The sensitivity of the empirical results with respect to the parameters choosen for g and δ is considered in the next section.

[25]The most striking exception are the undersampling of retail trade and the oversampling of traffic and postal services as well as software and telecommunication. As far as firm size is concerned, large firms are oversampled (see Table 10).

is equipped with ICT capital worth € 1,397, and with non–ICT capital worth about € 24,979. The median value added per employee is € 60,307.[26]

Estimating the first two extensions of eq. (2a) and (2b) puts substantially more requirements on the data, which reduces the corresponding samples remarkably. For estimating the dynamic specification (2a), only 708 firms for which at least four subsequent observations are available can be included in the "reduced sample". The data needs for the regressions including human capital based on eq. (2b) are even more restrictive. For 578 firms (denoted as "small sample"), consistent data on the skill–structure are available: the fraction of employees with vocational training (*Berufs- or Fachschulabschluss*) for medium–skilled, and the fraction of employees with a university degree including universities of applied sciences (*Hochschul- or Fachhochschulabschluss*) for highly–skilled workers.[27] As indicated in Table 8, the structure of the small sample differs from the full sample. In particular, the average firm size (183 employees) is only about a third of the firm size in the full sample. Therefore, estimates based on the small sample will be used mainly to explore the effects of including human capital variables into the specification.[28]

Some firms reported a share of ICT investment in total investment expenditures equal to zero for all the periods surveyed. Since the econometric specification is in logs, these firms are excluded from the full sample. However, it may seem more reasonable to assume that ICT investment in these firms is not zero, in fact, but rather very low and rounded to zero by the respondents.[29] Excluding these firms might lead to an overestimation of the real output contributions of ICT in the economy. In order to explore this potential bias, a third sample ("extended sample") is constructed. Here, the ICT stock per worker in firms that reported zero ICT investment is assumed to be equal to the corresponding industry minimum with the corresponding values being imputed. The corresponding statistics for the extended sample (see Table 9) indicate that the endowment of workplaces with ICT is slightly smaller, and the endowment with conventional capital slightly higher than in the full sample.

Independently of the specific sample used, the summary statistics indicate that the share of ICT capital in the total capital stock is very low. Comparing the medians

[26]The corresponding mean values are substantially higher than the median since some firms — in particular of real estate — display very high values for both inputs and output per employee.

[27]In one question, firms were asked to report the number of total employees and in another question to report the number of employees by skill–groups. In various cases, the sum of the latter was not equal to the former. Some 15 firms, for which the reported number of total employees deviated more than 50% from the sum of the skill groups, were excluded from the sample.

[28]Note that there are only 82 firms for which both human capital variables and at least four subsequent observations per firm are available. These data restrictions make it necessary to analyse the extension proposed in the previous section isolation, even though it would be worthwhile, of course, to obtain estimates from an approach combining approaches (2a) and (2b).

[29]Note that the definition of ICT investment as asked in the questionnaire is very broad, including expenditures for IT hardware, software and telecommunication equipment.

of ICT per worker and conventional capital per worker for the full sample (Table 7), ICT endowment amounts to 5.1% in total capital.[30] Similarly, aggregating firms' time–averages of both types of capital over all firms in the sample yields a share of aggregate ICT capital in total aggregate capital of 5% (not reported in the tables). These values are slightly higher than the share of 3% calculated by Schreyer (2000) using aggregate data for Germany in 1996 (including the less ICT–intensive manufacturing sector). As argued in Griliches (1994), the overall small shares of ICT input together with measurement errors may make it difficult to distinguish the output contributions of ICT from stochastic events and may make the identification of productivity effects of ICT resemble the search for the "needle in the haystack". In the empirical application, controlling for measurement errors will therefore be an important issue.

4 Empirical Results

This section discusses several econometric issues that need to be adressed for estimating equations (2) consistently. The best suited system GMM estimator will then be applied to explore the three extensions (2a-c). Apart from the constant and the input variables, the empirical specification includes a regional dummy for East German firms and 6 year dummies interacted with 7 industries.[31] All regressions are computed using the DPD98 programme developed by Arellano and Bond (1998) running in GAUSS. Only heteroscedasticity–consistent standard errors are reported.

4.1 Reference specification

The reference production function (2) is estimated first in a simple pooled OLS regression[32] (see first column of Table 1). The coefficients of all three inputs from the pooled OLS regression in column 1 of Table 1 are significantly different from zero at the one–percent level. While the output elasticity with respect to labour amounts to some reasonable 61%,[33] the point estimate of the ICT coefficient (24.4%) clearly exceeds the coefficient of conventional capital (14.9%) even though ICT forms only a small part of the total capital stock. Similarly high ICT elasticities have been found in cross section

[30] Taking the corresponding means, the share is even lower (1.8%).

[31] Table 10 summarises the underlying classification of industries. Since there is no output data available for banking and insurance (only the balance sheet total and insurance premiums respectively), these industries are excluded from the analysis.

[32] From an econometric point of view, a pooled regression corresponds to a simple cross–section regression except that a larger number of observations can be obtained from the inclusion of several years.

[33] Under the assumption of constant returns to scale and perfect competition, the income share of labour in an economy should equal its labour coefficient in the production function. For Germany, the average share of labour payments in national income between 1994 and 1999 amounted to 72.4% (Statistisches Bundesamt, 2001).

regressions for Germany by Bertschek and Kaiser (2001) and Licht and Moch (1999).

The high elasticity of ICT capital found in pooled or cross section OLS regressions raises serious doubts about the correctness of the applied estimation specification. Given the average share of ICT capital in value added of 6.2%, the results imply a gross rate of return to ICT investment of nearly 400%.[34] Assuming user costs of ICT of around 42% as suggested by Jorgenson and Stiroh (1995), the implied net returns are still substantially higher than 300%. For conventional capital, for which the share in value added is 258%, the results imply gross returns of only 5.8% which are close to its generally assumed user costs. The large excess returns to ICT can hardly be explained by higher user and adjustment costs of ICT capital alone which may be 'hidden' behind ICT investment. Rather, the results may be biased due to three main sources: firm effects, simultaneity issues, and omitted variables (e.g., skills). While the latter aspect is discussed with the extensions (2b), the first two involve econometric issues which are discussed with the reference specification (2). In the exploration of these issues, also interferences arising from measurement errors and the sensitivity with respect to the construction of the ICT capital stocks are discussed.

Unobserved firm characteristics ('firm effects') may bias the results if the investment strategies of highly productive firms are systematically different from their less productive competitors within the same industry.[35] It is likely that highly productive firms with a skilled and flexible management will be both more productive and tend to invest more in new technologies than other firms. This would induce an upward bias in the ICT coefficient. The highly significant autocorrelation in the errors of both first– and second–order[36] in the pooled regression further supports this conjecture.

Table 1 reports the results of the estimation in first differences.[37] The figures indicate that once unobserved heterogeneity is controlled for, the output contributions of both types of capital are no longer significantly different from zero whereas the labour coefficient remains virtually unchanged.[38] Obviously, the high coefficients of both types of capital in the pooled regression were in fact due to unobserved heterogeneity. This finding coincides with very similar findings by Brynjolfsson and Hitt (1995) and Black and Lynch (2001). Moreover, the autocorrelation in the disturbance terms found in the

[34]The marginal returns to ICT (MPI) are just the product of the output elasticity of ICT and the inverse ratio of ICT capital in output: $MPI_{it} = \partial Y_{it}/\partial ICT_{it} = \gamma_2 \cdot Y_{it}/ICT_{it}$.

[35]Productivity differences between different industries are captured by the industry dummies.

[36]See the last two rows AR(1) and AR(2) of Table 1.

[37]This means that the firms' corresponding fixed effects are eliminated by explaining output growth by the growth rates of the inputs. The results from the alternative within estimation where deviations from means are used (not reported) are very similar.

[38]Since there is no variation in the East dummy over time, this variable is excluded from the first–differences estimation.

Table 1: Results for the ICT–augmented production function

	Dependent Variable: log(value added)						
inputs	(1) OLS pooled	(2) OLS 1st dif	(3) GMM[–1] 1st diff.	(4) GMM[–2] 1st diff.	(5) SYS–GMM reference	(6) SYS–GMM not interact.	(7) SYS–GMM extended
log(labour)	0.607***	0.598***	0.555***	0.282*	0.699***	0.717***	0.686***
	(0.020)	(0.075)	(0.087)	(0.154)	(0.056)	(0.056)	(0.058)
log(ICT capital)	0.244***	-0.025	0.024	0.032	0.060*	0.022	0.049*
	(0.020)	(0.017)	(0.026)	(0.041)	(0.034)	(0.034)	(0.026)
log(non–ICT cap.)	0.149***	-0.035	0.140	0.310**	0.201***	0.213***	0.189***
	(0.015)	(0.052)	(0.119)	(0.157)	(0.036)	(0.037)	(0.036)
East–Germany	-0.127***	—	—	—	-0.386***	-0.402***	-0.384***
	(0.043)				(0.045)	(0.047)	(0.045)
observations	4939	3762	3762	3762	4939	4939	5107
firms	1177	1177	1177	1177	1177	1177	1222
R–square	0.844	0.236	0.218	0.137	0.843	0.839	0.836
Wald stat. [df]:							
inputs	24160[4]	65.1[3]	52.2[3]	17.3[3]	560[4]	561[4]	609[4]
time and ind. dummies	702[41]	133[35]	149[35]	113[35]	651[41]	550[11]	685[41]
Sargan (p–values)	—	—	0.187	0.248	0.258	0.119	0.193
errors (p–values):							
AR(1)	0.000	0.005	0.007	0.006	0.004	0.004	0.003
AR(2)	0.000	0.131	0.135	0.085	0.049	0.042	0.039

***,** and * denote significance at the 1,5 and 10 per–cent level.

All regressions contain a constant and industry interacted dummy variables for 6 years (1994–99) and 7 industries (no interaction only in regression 6). GMM[-1] and GMM[-2] refer to estimations using all lagged levels of explanatory variables $t-s$ with lag $s \geq 1$ and ≥ 2 correspondingly (see text for details). For all regressions, heteroscedasticity consistent standard errors reported.

pooled specification has vanished and was obviously due to the firm effects.[39]

The implausibly low estimates of the capital coefficients for the estimates in first differences may be caused by a second type of bias, which is due to measurement errors. Measurement errors are likely to be substantial in both types of capital stocks since both the depreciation and the pre–sample growth rates are assumed equal across firms. Deviations from this assumption will add noise — though presumably not a systematic one — in the construction of the firms' capital stocks. As pointed out by Griliches and Hausman (1986), measurement errors may induce a downward bias in the OLS estimates.

However, this distortion may be offset by a simultaneity bias. If firms determine input and output simultaneously, exogenous shocks — like demand shifts, for example — result in an increase of both input and output for the profit–maximizing firm.[40] In econometric terms, the disturbance term ϵ_{it} will be positively correlated with the input variables in

[39]Note that the observed first–order correlation of the errors is induced by the data transformation. If the errors ϵ_{it} are i.i.d. with variance σ^2 their corresponding first differences will be AR(1): $E(\Delta\epsilon_{it} \cdot \Delta\epsilon_{i,t-1}) = E((\epsilon_{it} - \epsilon_{i,t-1})(\epsilon_{i,t-1} - \epsilon_{i,t-2})) = -\sigma^2$. Therefore, the relevant test for equations in first differences is whether the corresponding errors are AR(2) or not.

[40]See Griliches and Mairesse (1998).

equation (2) causing an upward bias in the input coefficients. However, the simultaneity bias may apply in particular to factors that can be adjusted easily in the short term. This is not so much the case for capital stocks. Moreover, in the construction of the data, capital stocks at the *beginning* of the corresponding years have been used. Therefore, the (upwards) simultaneity bias is expected to be rather small for the two capital coefficients.

In order to analyse the distortions due to measurement errors and simultaneity, GMM estimates with internal instruments are applied to the production function in first differences. This approach takes advantage of the panel structure of the data by instrumenting contemporaneous inputs in differences with the corresponding values in the past and is discussed in more detail in Appendix B. More specifically, in the specification of column 3 of Table 1, the corresponding (log) levels of the available lagged inputs $x_{t-1}, x_{t-2}, ..., x_1$ are used to instrument the input in differences $\Delta x_t = x_t - x_{t-1}$ (GMM[–1]), with x denoting the inputs L, ICT and K.[41] In column 4 of Table 1, the instruments x_{t-1} are dropped to allow for simultaneity of capital stocks at the beginning of each period t and shocks arising in t (GMM[–2]).

The corresponding results from the two–step estimation[42] show that in both specifications, the point estimates for the capital coefficients increase whereas the labour elasticity decreases. This tendency is much more pronounced in the GMM[-2] specification where the coefficient of conventional capital rises to 0.310 and the labour coefficient drops to a (quite low) value of 0.285.[43] However, the capital coefficients remain insignificant from zero in both these specifications when the one–step results are considered (see Table 12 in Appendix D). Summarizing the results, these findings indicate that the measurement error bias in the capital coefficients clearly exceeds the counteracting simultaneity bias.[44] By contrast, for the case of labour input, the simultaneity bias exceeds the measurement–error bias as it was expected. For both specifications, the Sargan test[45] does not reject the validity of the instruments. Finally, like in the specification in OLS

[41]Including x_{t-1} as an instrument is based on the assumption that by taking capital stocks at the beginning of each period it is ensured that the inputs are predetermined, i.e. uncorrelated with the idiosyncratic shock ϵ_{it} of the same period since $E(x_{t-1}\Delta\epsilon_t) = 0 \Leftrightarrow E(x_{t-1}\epsilon_t) - E(x_{t-1}\epsilon_{t-1}) = 0$. The validity of this assumption can be tested (see footnote 43). In the remainder, however, this moment condition will be dropped to explicitly control for potential simultaneity of inputs and output.

[42]The one–step results are reported in Table 12 in Appendix D. Even though the two–step estimates reported in the main part are more efficient, its standard errors are less appropriate for tests of significance. As pointed out by Blundell and Bond (1998) on the basis of Monte Carlo simulations, "[i]nference based on one–step GMM estimators appears to be much more reliable when either non–normality or heteroskedasticity is suspected" (142).

[43] The results from a Sargan difference test (see Appendix B) suggest that the additional moments employed in the GMM[–1] compared to the GMM[-2] specification ($E(x_{t-1}\Delta\epsilon_t)$=0) cannot be rejected (p=0.186).

[44]These findings coincide with similar results in Black and Lynch (2001) for estimates of the production function with one type of capital only.

[45]See Appendix B for technical details.

first differences, no autocorrelation of the error term is detected.

One reason for the insignificant capital coefficients found in the GMM regressions may be the small power of the instruments used. Since capital stocks within firms are highly persistent over time, the correlation of the first differences with the second lag in levels is close to zero.[46] Blundell and Bond (1998) show that this may induce finite–sample biases of the GMM estimator in first differences. Based on an application to production function estimation, Blundell and Bond (2000) argue that in the specification in first–differences, the weak instruments will bias the GMM estimates in the direction of the within group estimation, that is towards zero. They suggest using the sytem GMM (SYS–GMM) estimator originally proposed by Arellano and Bover (1995). In this estimation strategy, both the equation in differences is instrumented by suitably lagged differences (like in the simple GMM–estimation) and the equation in levels is instrumented by suitably lagged differences additionally. These two specifications are then estimated simultaneously.[47]

The corresponding regression ("SYS–GMM reference") adds the production function equation in levels (with lagged differences of period $t-1$ as instruments) to the GMM[–2] specification. As shown in column 5 of Table 1, like conventional capital and labour, the coefficient of ICT capital turns out to be positive and highly significant ($p = 0.014$ in the one–step estimation).[48] Several features support the appropriateness of this econometric specification. The output elasticity of labour amounts to about 70% which is consistent with aggregate statistics (see footnote 33). The coefficients of ICT and non–ICT capital are 6% and 20% respectively, and the null–hypothesis of constant returns to scale cannot be rejected at the 1%–level (not reported).[49] A further robust result is that East–German firms in services are significantly less productive than their West–German counterparts. The coefficient of the East–Dummy (-0.386) implies that the multi–factor productivity in East–German firms is only about two–thirds of the West–German level. This finding coincides with aggregate statistics on productivity differentials in Germany. While the test for serial correlation of the errors is at the border of significance ($p = 0.049$), the Sargan–statistic ($p = 0.258$) does not reject the validity of the instruments at the usual significance levels.[50] Moreover, the difference Sargan statistic (44.3[12]) does not reject

[46]Formally, this can be illustrated by assuming K_{it} being AR(1): $K_{it} = \rho K_{i,t-1} + r_{it}$ with $\epsilon \sim i.i.d$ and $E(r_{it}) = 0$. If K_{it} is weakly autocorrelated ($|\rho| \ll 1$ and $\rho \neq 0$), the past levels are correlated with the contemporaneous levels. For the first available instrument $K_{i,t-2}$, this is: $E(\Delta K_{it} \cdot K_{i,t-2}) = E((K_{it} - K_{i,t-1}) \cdot K_{i,t-2}) = E(K_{it} \cdot K_{i,t-2}) - E(K_{i,t-1} \cdot K_{i,t-2}) = \rho^2 - \rho$. If the evolution of K_{it} resembles a random walk ($\rho \approx 1$), the correlation between the variable in differences and its past values in levels will disappear ($\rho^2 - \rho \approx 0$) and the instruments will therefore turn out to be weak.

[47]Appendix B gives a brief overview of the involved technical details.

[48]The less reliable p–value in the two–step estimation amounts to 0.078.

[49]This result holds for both the one–step and the two–step estimation results.

[50]Appendix B summarises the background of this test and discusses why serial correlation of the errors may be at odds with the validity of the moment conditions (see footnote 80, p. 36). In empirical terms,

the validity of the additional instruments obtained from the equation in levels (p=0.299).

Since these results stem from the preferred specification in this study, a glance at the implied rates of return appears worthwhile. Given the calculated average share of ICT capital in output of 6.2% for the firms in the sample, the results imply that € 1 invested in ICT capital yields returns of € 1.96.[51] This high value is very similar to the findings in various related studies.[52] Assuming again user costs of ICT of around 42%, the remaining excess returns to ICT of 54% may well be due to complementary investment like training of the workforce, innovation efforts or costs due to the re-structuring of organizational forms which are not accounted for as inputs in the framework employed here.

In order to further investigate the sources of potential biases in estimating the productivity of ICT, the effect of ignoring different business cycles and mismeasured output prices is analysed. To isolate the role of including interacted time and industry dummies, the SYS–GMM approach is estimated with simple (not–interacted) time and industry dummies. The corresponding results in column 6 of Table 1 show that the coefficient of ICT capital is indeed affected by this change. The corresponding point estimate decreases to roughly 2.2% and is only marginally significant.[53] By contrast, the other coefficients do not exhibit any remarkable changes compared to the specification with interacted dummies. Moreover, a Wald test of significance of the 30 additional interaction dummies included in specification (5) clearly rejects the null hypothesis of no joint significance.[54] These results suggest that it is indeed important to control for industry–specific effects in order to assess the contributions of ICT correctly.

In the last column of Table 1, results for the SYS–GMM estimation *with* interacted dummies are replicated, but now for the extended sample in which also those firms are included that reported zero ICT investment for all the periods surveyed. As detailed in section 3, this sample is extended by 46 firms that have reported zero ICT investment for all years observed, imputing the industry minimum in terms of ICT per worker. The inclusion of these firms slightly lowers the point estimate for ICT (4.9%) as compared to the values reported for the reference sample. Moreover, the ICT coefficient is significantly positive only in the two–step estimation.[55] These results appear quite reasonable if one considers that firms may differ in their output elasticities. Those firms with a lower output elasticity of ICT will be maximizing profits with a lower share of ICT capital in

the following results for the model extension 1 will shed some more light on how results may be affected by this issue.

[51]For non–ICT capital, the results imply that one Euro invested yields a much smaller return of € 1.078.

[52]See Brynjolfsson and Hitt (2000).

[53]The results from the one–step estimation (see Table 12) yield a significance level of p=0.099.

[54]The χ^2–test statistic is 189.9 with 30 degrees of freedom.

[55]The one–step estimates imply a p–value for the ICT coefficient of 0.107.

Table 2: Sensitivity analyses of the results with respect to varying parameters for the construction of the capital stocks

	varying parametrisation of δ_{ICT} (with $g_{ICT} = 0.4$)						
	0.05	0.1	0.2	**0.3**	0.4	0.5	1.0
est. coeff. of log(ICT)	0.074**	0.072**	0.067**	**0.060****	0.052**	0.046**	0.007**
mean of log(ICT)	-1.617	-1.736	-1.947	-2.131	-2.297	-2.451	-3.596

	varying parametrisation of g_{ICT} (with $\delta_{ICT} = 0.3$)					
	0.1	0.3	**0.4**	0.5	0.7	0.9
est. coeff. of log(ICT)	0.106***	0.070**	**0.060****	0.052**	0.041**	0.033**
mean of log(ICT)	-1.825	-2.052	-2.131	-2.197	-2.303	-2.385

Notes: ICT denotes ICT capital stock measured in € million. Estimated oefficients (est. coeff.) are obtained from the reference specification employed in the regressions in the paper (Table 1, col. 5). δ_{ICT} denotes the annual depreciation rate assumed for ICT capital, g_{ICT} the assumed growth rate of ICT investments in the pre–1994 periods (see text). The parameter values of $\delta_{ICT} = 0.3$ and $g_{ICT} = 0.4$ (bold letters) correspond to the preferred parametrisation used in the regressions reported in the other tables of the paper.
** and *** denote significance at the 5%– and 1%–level respectively.

output; excluding these firms might overstate the ICT coefficient.

A last exploration based on the reference model (2) concerns the sensitivity of the results with respect to the way in which capital stocks are constructed. As is obvious from eq. (12) and (13) on page 9, both the level and the evolution of the capital stocks of the firms depends on the parametrisation used for annual depreciation δ and pre–period growth rates of investment g. In order to explore to what extent the econometric results depend on the assumed values for the ICT capital stock, I subject the reference regression underlying col. 5 of Table 1 to two kinds of robustness checks. In the first, I calculate alternative ICT capital stocks using different values for depreciation rates δ_{ICT} while holding assumed growth rate g_{ICT} constant. In the second, I did the reverse, holding δ_{ICT} constant while varying g_{ICT}.

The estimates for the elasticity of ICT resulting from these variations are reported in Table 2. Most strikingly, the qualitative result of significant productivity contributions of ICT is robust to both kinds of variations. Unsurprisingly, however, the point estimate of the elasticity decreases in both parameters. For the extreme case of a complete depreciation of ICT within one year ($\delta_{ICT} = 100\%$), the point estimate is very small.[56] Moreover, lowering the assumed depreciation of ICT from an annual rate of 30% to 20% increases the estimated elasticity of ICT only modestly from 0.060 to 0.067. The

[56]This finding supports the importance of employing capital stocks instead of investments for assessing the productivity contributions correctly. Employing investments implicitly correponds to assuming capital to depreciate completely after one period.

effects of a similar variation of g_{ICT} is only slightly higher. The main message from this exercise is thus that the empirical results reported in Table 1 do not depend critically on assuming certain values for δ_{ICT} and g_{ICT}.

4.2 Results for the extended models

This subsection reports further evidence for the variations of model eq. (2) discussed in the theoretical part.

Extension 1: **Serially correlated residuals.** In Table 3, the estimated elasticities for the dynamic extension of eq. (2a) are reported. Since the employed sample differs from the one used in the previous regressions,[57] also the results for the static model are displayed in the first row. Comparing the two–step results from col. 2 to the analogue specification for the full sample (col. 5 of Table 1) shows that the change in the sample impacts the results only very little.

As indicated in the theoretical section and described in detail in Appendix C, the results for the dynamic specification reported in cols. 3 and 4 are obtained from first estimating the reduced–form model and then imposing the common factor restrictions. The reduced–form estimates are summarised in Table 13 in Appendix D. Unlike in the results for the previous regressions, Table 3 reports both the results for the one–step and the two–step SYS–GMM results because the point estimates from the minimum distance procedure depend on both the point estimates and the variances of the reduced form estimates. For the reduced–form model, variances from the one–step SYS–GMM results are preferred whereas the point estimates from the two–step findings are more efficient.

A common finding from both the one– and the two–step specification is that there is strong evidence for serial correlation in the residuals with ρ being roughly 0.77. Similarly, the estimates of the labour elasticity is substantially lower than for the static model. The estimates of the capital coefficients, however, differ substantially between one– and two–step estimates with the one–step results being substantially higher.[58] By contrast, the two–step results for the capital coefficients are not too far from the values obtained for the static model. Both capital coefficients are estimated more imprecisely for the dynamic model, however, with the ICT coefficient not being significantly different from zero.

[57]In the dynamic specification, also the coefficients of the once lagged inputs as well as the lagged dependent variable are included. The lagged difference of these variables (e.g. Δict_{t-1}) is then instrumented by the levels lagged 3 periods ($ict_{i,t-3}$). Thus, in this specification only firms can be included for which at least four subsequent periods ($t, t-1, t-2, t-3$) are available.

[58]In the dynamic one–step results, the sum of the two capital coefficients (roughly 0.48) as well as the labour coefficient correspond fairly well to comparable results by Blundell and Bond (2000) who report estimates of 0.49 for total capital and 0.48 for labour input for U.S. manufacturing firms during 1982–89.

Table 3: Structural coefficients for the static and dynamic specification

	Dep. Variable: log(value added)			
	(1)	(2)	(3)	(4)
	static[a]	static[a]	dynamic[b]	dynamic[b]
	(one–step)	(two–step)	(one–step)	(two–step)
AR(1) of error	—	—	0.774***	0.771***
			(0.112)	(0.066)
log(labour$_t$)	0.768***	0.722***	0.464***	0.487***
	(0.099)	(0.079)	(0.153)	(0.112)
log(ICT$_t$)	0.090**	0.057*	0.158	0.074
	(0.042)	(0.030)	(0.122)	(0.070)
log(non–ICT$_t$)	0.109***	0.166***	0.320**	0.216***
	(0.069)	(0.051)	(0.138)	(0.088)
Minimised distance[df]	—	—	3.923[3]	8.778[3]
Common factor restr.				
(p–values)	—	—	0.270	0.032

***, **, * = significant at the 1, 5 and 10 per cent level respectively

Results from SYS-GMM estimates with robust standard errors in parentheses.

[a] The static model corresponds to the specification underlying col. 5 in Table 1 for the full sample except for sample differences. [b] The results for the dynamic specification are obtained from applying a minimum distance procedure to the estimated coefficients reported in Table 3 in Appendix D. The test of the validity of the common factor restrictions is based on the value of the minimised distance function (see Appendix C). The underlying sample for all results consists of the "reduced sample" with 3532 observations for 708 firms covering the years 1994–1999 (see section 3 for details).

The test of validity of the imposed common factor restrictions are rejected for the two–step estimates but are not rejected for the one–step estimates. This difference in the test statistics, may be a direct consequence of the estimated standard errors of the reduced–form parameters which tend to be biased towards zero in the two–step estimation. Since the test for the validity of the factor restrictions depends on these standard errors,[59] this test is not too informative about the question whether the one–step point estimates are more reliable than the two–step results.

To sum up the evidence from the dynamic model, accounting for serial correlation yields ambiguous results compared to the static specification. On the one hand, the point estimates are higher in the dynamic model. On the other hand, the coefficients are estimated much less precisely and fail to reach statistical significance.

Extension 2: **Heterogenous labour.** A further issue in estimating the productivity of ICT is the potential bias owing to omitted variables that may be complementary to

[59]See eq. 38 in Appendix C.

the firm's use of ICT. In particular, recent studies have found that differences in the skills of the workforce play an important role in this regard (Bresnahan et al., 2002).[60] On the one hand, ignoring differences in workers' skills might lead to an overestimation of the true impacts of ICT on production. On the other hand, a firm's 'skill–mix' tends to be very persistent over time. Thus, their effect may not be distinguishable from other quasi–fixed factors which are controlled for as unobserved heterogeneity between firms. In this case, no distortions are expected.

In order to assess the role of omitting differences in workers' skills, the model is extended by the shares of employees with vocational training and with university degree as summarised in eq. (2b) and (2b'). As discussed in section 3, the resulting small sample consists of 578 firms only. The first column of Table 4 reports the results from applying the SYS–GMM reference estimation strategy (column 5 in Table 1) to the small sample. Compared to the full sample, the coefficient of labour (0.758) is slightly higher for the small sample whereas both capital coefficients are substantially smaller.[61] One reason for these changes may be that average firm size as well as average and median endowment of workplaces with ICT capital are notably lower in the small sample.[62] Moreover, the reduction in the significance levels of both capital coefficients may be a direct consequence of the loss of precision due to the reduced sample size.

The second column of Table 4 displays the effect from including the proxies for human capital in the regression. In this specification, the shares of the employees with high and medium skills (represented by '% university' and '% vocational') are treated as exogenous, i.e. these variables are instrumenting themselves. This Both the share of employees with university degree and the share of workers with vocational training are highly significant and positive.[63] As the comparison to the first column reveals, including the human capital variables slightly reduces the coefficients of labour but leaves the

[60] Other candidates for complements to ICT are investments in intangible capital goods such as training, innovation or organizational capital (Bresnahan et al., 2002; Brynjolfsson and Hitt, 1998; Hempell, 2002; Hempell, 2003). However, the investigation of the impacts of all these complements on the ICT coefficient is beyond the scope of this paper and is left for future research.

[61] Moreover, only the non–ICT coefficient is significantly different from zero in the one–step estimation. Note that values reported in brackets of Table 4 are — unlike in the previous tables — the t–values from the one–step estimates. This comprehensive manner of presentation substitutes for further tables with one–step results in the Appendix.

[62] See last columns of Tables 7 and 8 in Appendix D.

[63] The implicit values for the productivity differentials for medium– and high–skilled workers are $\theta_m = \beta_m/\gamma_1 = 0.419/0.726 = 0.577$ and $\theta_h = \beta_h/\gamma_1 = 0.970$. With competitive salaries in the labour market, these values should roughly correspond to the wage spread over the corresponding skill levels. For the service sector, Kaiser (2000) calculates wage premiums of $\theta_m^w = 0.325$ for medium–skilled workers and of $\theta_h^w = 1.025$ for high–skilled workers. This comparison indicates that the approximation in eq. 5 may lead to an overestimation of the corresponding coefficient for medium–skilled employees. Alternatively, firms may pay less than competitive wage premiums to skills.

Table 4: The effects of heterogenous labour

	Dep. Variable: log(value added)				
inputs	(1)	(2)	(3)	(4)	(5)
log(labour)	0.758***	0.726***	0.758***	0.680***	0.565***
	(4.601)	(4.872)	(6.375)	(4.633)	(3.841)
log(ICT capital)	0.016	0.027	0.006	0.017	-0.098
	(0.362)	(0.380)	(0.589)	(0.168)	(-0.969)
log(non–ICT capital)	0.146*	0.150	0.147*	0.181	0.110**
	(1.855)	(1.386)	(1.677)	(1.242)	(2.189)
% university	—	0.704***	-0.190	1.581	1.752***
		(2.626)	(-0.947)	(1.576)	(2.878)
% vocational	—	0.419**	0.099	1.373***	0.676*
		(2.534)	(-0.231)	(3.661)	(1.802)
(% university)^2	—	—	—	-0.913	—
				(-1.200)	
(% vocational)^2	—	—	—	-0.923***	—
				(-3.202)	
%univ. * % voc.	—	—	—	-0.873	—
				(-1.281)	
log(ICT) * % univ.	—	—	—	—	0.557**
					(2.026)
log(ICT) * % voc.	—	—	—	—	0.140
					(1.524)
observations	1847	1847	1847	1847	1847
number of firms	578	578	578	578	578
R–square	0.821	0.826	0.814	0.826	0.814
WALD STAT.[DF]:					
inputs	142[3]	173[5]	339[6]	294[8]	210[7]
time and ind. dummies	335[34]	341[34]	419[34]	326[34]	353[34]
Sargan (p–values)	0.677	0.794	0.385	0.772	0.806
RESIDUALS (P–VALUES):					
AR(1)	0.014	0.020	0.002	0.008	0.043
AR(2)	0.183	0.084	0.097	0.061	0.198

***,**,* = significant at the 1, 5 and 10 per cent level
Results are based on the two–step SYS–GMM estimator and contain a constant, a regional dummy for East–German firms as well as interacted time and industry dummies. T–values reported in brackets are obtained from (heteroskedasticity–robust) first–step estimation results. The signs of coefficients and t–values may therefore vary in some few cases.
Labour and capital inputs are instrumented by past values as described in the text, while % vocational and % university are treated as exogenous except in specification 3 where also past values are used as instruments.

coefficient of non–ICT capital broadly unaffected.[64] The elasticity of ICT increases slightly from 0.016 to 0.027 even though in the one–step estimates both the coefficient

[64]In a related exercise, Lehr and Lichtenberg (1999) report similar qualitative results.

and its standard error remain practically the same.[65]

Treating the skill–composition as exogenous may be justified if productivity shocks impact the *quantity* of labour but not its composition by skills and if, moreover, skill–composition is not affected by firm effects. However, these assumptions may be violated and may also impact the ICT estimates. Specification (3) is the same as (2) except that also the skill variables are now treated as endogenous by using their past values as instrument in an analogue manner to the other inputs. The skill coefficients are estimated very imprecisely with the coefficient for high–skilled workers becoming even (insignificantly) negative.[66] Independently of the way of instrumenting skills, the coefficients of the input factors labour, ICT and non–ICT capital remain broadly unaffected.

A further issue consists in the fact that the approximation of eq. (5) is very imprecise. Col. 4 reports additional results for the more accurate model (2b') with skills being instrumented by themselves again. The (insignificant) ICT coefficient is very close to the one obtained for the specification without controlling for labour heterogeneity while the coefficient of labour is notably lower and the one for non–ICT higher than in col. 1. Again, there is no indication from the results that the omission of labour quality may exert any important bias on the ICT estimate. A test of the validity of the common factor restrictions for the skill coefficients from eq. (5a) does not reject the model at the 5%–level (p–value 0.074).[67] However, the structural parameters obtained from a minimum distance procedure yield rather high values for the implied coefficients θ_m and θ_h.[68]

In order to obtain some more evidence on the link between skills and ICT, specifications (4) and (5) additionaly consider interaction terms between ICT and the skills variables. The interaction between ICT and human capital is highly significant,

[65] Further unreported regressions show that including skill groups as separate inputs (instead of adding skill shares, see comments in footnote 15) does not yield very different results. For a sample of 222 firms with non–zero number of employees for all three skill groups, both ways of considering heterogeneity of labour quality yields very similar but insignificant ICT coefficients of slightly more than 0.05 which are slightly higher than in the specification without controlling for skill structure of the employees.

[66] It is extremely difficult to trace the sources of these counterintuitive results. Finite sample biases due to poor instruments are unlikely to be the reason since further explorations show that the power of the instruments for the skill shares is even slightly higher than for the capital variables. There is neither evidence for outliers to be driving the results. Excluding firms with exceptionally high changes in the skill shares as potential outliers have no noteworthy effects on the results. Instrumenting present skill shares with lagged shares, however, yields results that are very similar to treating skill shares as exogenous.

[67] Instrumenting skill–shares and their interactions as in col. (3) yields even higher p–values for the test of the imposed common factor restrictions.

[68] The corresponding coefficients are: $\theta_m = 1.413\,(0.155)$ and $\theta_h = 1.201\,(0.181)$ with standard errors in parantheses. A peculiarity of these results is that the coefficient for medium skills is higher than the one for medium skills. A closer look at the results of Table 4 shows that this is mainly due to the very small precision of the estimates for high skills combined with a low absolute value for the interaction term % university * % vocational. Jointly, these two results push the structural parameter θ_h toward zero in the minimum distance approach.

indicating that the productivity of ICT is increasing with the share of highly educated employees. The coefficient of ICT alone becomes even negative, implying that in order to make productive use of ICT, skilled workers are even an essential prerequisite.[69]

Summing up, there is no evidence that omitted heterogeneity in labour quality leads to an overestimation of the *average* productivity impacts of ICT. This may be due to the fact that the share of high–skilled workers tends to be highly persistent over time. Human capital might thus be treated as a firm's quasi–fixed asset that is controlled for by first differencing. However, the findings suggest that ICT must be complemented by highly educated employees in order to result in positive productivity effects — a result that is in line with similar findings in Bresnahan et al. (2002).

Extension 3: **Flexible functional form.** A final issue concerns the functional form of the production technology. In particular, the Cobb–Douglas technology may be too restrictive if scale effects and complementarities between the inputs may affect the results. To assess this question, both the simplest (pooled OLS) and the best suited (SYS–GMM) estimations are applied to the translog production function of equation (2c). Unfortuately, the scope for empirical investigation of these issues is quite limited by data constraints. In particular, the data basis is too small to obtain meaningful estimates for a human–capital augmented translog function.[70]

The corresponding results and the average elasticities as of eq. (9) are reported in the first two columns of Tables (5) and (6). Like in the estimations for the Cobb–Douglas framework, pooled OLS and SYS–GMM estimates differ substantially in both the individual coefficients and the implicit average elasticities. Again, the mean output contributions are overestimated by using pooled OLS (Table 6). A striking feature of the translog function is that even for the SYS–GMM estimation, the implicit average elasticity of ICT (0.148) is much higher than in the Cobb–Douglas specification.[71]

There are two features of the results, however, that raise doubts about the reliability of the translog specification. First, the Wald statistic for the joint significance of the additional translog inputs[72] from the one–step estimation rejects the relevance of these

[69]With skills being instrumented, the interaction of ICT and skills is positive, too, but smaller in both economic and statistical significance. The estimates of the direct productivity contributions of skills, however, are very low, too, pointing to the same problems in the specification as discussed for the corresponding specification without interaction (col. 3).

[70]Including human capital into the translog specification would require to treat each skill group as a separate input in the production function. For this case, all the problems mentioned in footnote 15 apply. Moreover, the number of regressors rises exponentially with the number of inputs considered in the translog function, which leads to a further decrease in the degrees of freedom of the regressions.

[71]In a similar comparison between the Cobb–Douglas and the translog specification, Brynjolfsson and Hitt (1995) find an only slightly higher average elasticity of ICT for the translog version.

[72]These are the regressors l^2, ict^2, k^2, $l \cdot ict$, $l \cdot k$), ($ict \cdot k$ which are not included in the Cobb–Douglas specification.

Table 5: Results for the translog production function

	Dep. Variable: value added (logs)		
inputs (log)	OLS full	SYS–GMM full	SYS–GMM extended
labour	1.100***	1.178***	1.077***
	(0.104)	(0.137)	(0.144)
ICT capital	0.050	-0.045	0.006
	(0.068)	(0.085)	(0.070)
non–ICT capital	0.065	0.156***	0.169***
	(0.048)	(0.055)	(0.059)
labour^2	-0.040***	-0.049***	-0.045***
	(0.011)	(0.013)	(0.015)
ICT capital^2	0.006	0.001	0.008***
	(0.005)	(0.006)	(0.003)
non–ICT capital^2	0.031***	0.013*	0.008
	(0.005)	(0.007)	(0.007)
labour*ICT	0.059***	0.049***	0.043***
	(0.012)	(0.016)	(0.013)
labour*non–ICT	-0.020*	-0.008	-0.004
	(0.010)	(0.012)	0.0144
ICT*non–ICT	-0.042***	-0.009	-0.009
	(0.008)	(0.008)	(0.007)
East–Germany	-0.336***	-0.607***	-0.544***
	(0.042)	(0.146)	(0.145)
observations	4939	4939	5107
firms	1177	1177	1222
R–square	0.859	0.850	0.846
Wald–statistics[df]:			
all inputs	6,801[10]	7,479[10]	6,556[10]
additional inputs†	77.15[6]	22.88[6]	89.82[6]
time and ind. dummies	721.6[41]	767.1[41]	804[41]
Sargan (p–values)	—	0.144	0.080
errors (p–values)			
AR(1)	0.000	0.003	0.004
AR(2)	0.000	0.043	0.047

***,**,* = significant at the 1, 5 and 10 per cent level
The results of the second column are based on the two–step SYS-GMM and contain a constant and industry dummy variables interacted with year dummy variables. Heteroscedasticity consistent standard errors reported. †refers to additional inputs *not* included in Cobb–Douglas specification.

variables (4.96[6], p=0.549). Second, the translog estimates are highly sensitive to small changes in the sample. To illustrate this, the SYS–GMM estimator is applied to the extended sample instead of the full sample. This extension of the sample by 45 firms (3.8% of the sample) causes substantial changes in the ICT–related coefficients (see column 3 of Table 5). Moreover, the average elasticities for all three inputs change remarkably (see Table 6). By contrast, the sensitivity to sample changes was much smaller for the Cobb–

Table 6: Implicit average elasticities for the translog production function

inputs	OLS full	SYS–GMM full	SYS–GMM extended
labour	0.707	0.677	0.619
ICT capital	0.215	0.148	0.124
non–ICT capital	0.140	0.168	0.193

Douglas specification (see columns 5 and 7 of Table 1). The underlying reason may be that in particular the quadratic terms are very sensitive to potential outliers in the sample.

5 Conclusions

The use of firm–level data is gaining in importance for the analysis of productivity effects of ICT. In contrast to aggregate data, firm–level information is less dependent on the accuracy of price deflators and entails a higher variation in the factors that may determine the performance of businesses. Moreover, unlike growth accounting approaches, estimating production functions based on firm–level data does not require to assume constant returns to scale and perfect competition.

In this paper, it is shown that the empirical results on the productivity of ICT gained from a production function framework are highly contingent upon the specific econometric methods applied. The empirical analysis based on firm–level panel data from the German service sector yields evidence of various interfering influences that should be addressed econometrically. First, and most prominently, well–managed firms are likely to be intensive users of ICT. If these unobservable firm effects are not taken into account by using a first–differences or a within–estimator, the productivity impacts of ICT will be drastically overstated. Second, counteracting this effect, measurement errors in the explanatory variables may lead to an underestimation of the corresponding elasticities. This problem turns out to be particularly important for the case of ICT capital. Even though ICT investment has increased substantially over the last years, the share of ICT equipment and software in total capital is still very small. This makes it difficult to distinguish the output contributions of ICT from statistical noise. By contrast, third, the simultaneity of input and output decisions by firms, which may induce an upward bias of the output contributions of ICT, is found to be less important for the econometric specification. If panel data are available, both the measurement error bias and the simultaneity bias may be overcome by applying a GMM estimator that uses information from suitably distant previous periods to instrument the production inputs of the firm. However, when unobserved firm–effects are taken into account, too, this estimation strategy may suffer from small sample biases due to weak instruments. Therefore, the most suited approach is found to be the system GMM (SYS–GMM) strategy proposed by Arellano and Bover (1995). This approach applies the GMM estimator to the firms' production function equation in levels and first differences simultaneously and thus makes use of more powerful instruments.

Fourth, potential mismeasurement of output prices and the omission of industry–specific business cycles may understate the productivity impacts of ICT also at the firm–level. This bias may partially be addressed by including interacted time and industry dummies in the regression. Fifth, the explicit consideration of serial correlation of exogenous shocks at the firm–level in a dynamic specification of the production function yields slightly higher but also less precise estimates for the ICT coefficient. Sixth, the shares of high– and medium skilled workers have a large and significant effect on productivity. However, the omission of these variables does not lead to an overestimation of the productivity contributions of ICT once firm–specific fixed effects are taken into account. Obviously, most of the variation in the skill structure is *between* rather than *within* firms. Finally, estimates based on the more flexible translog production function yield higher ICT elasticities than the Cobb–Douglas specification. However, these estimates turn out to be much more sensitive with respect to small sample changes and yield little improvements in the explanatory power compared to the more parsimonious Cobb–Douglas specification.

What about the implications for the empirical work on the economics of ICT? From an econometric point of view, the data needs necessary to address the methodological issues raised in this paper are indeed quite demanding. In particular, a longitudinal structure of at least three observations per firm is required to apply the suited SYS–GMM estimator. On the other hand, the calibration strategies proposed in this paper for constructing appropriate input and output data may be applicable to various other existing longitudinal micro data sets, which frequently contain information on sales, employment and investment. In any case, great caution seems to be appropriate for the interpretation of cross–section results on the topic. The findings of this study indicate that a big part of such results may be due to spurious correlations that tend to dominate the real causal impacts of ICT on the productivity of businesses.

From an economic point of view, the findings of this paper point to the need of investigating particular firm characteristics and strategies in more detail. The results from the preferred system GMM estimation imply that a one–percent increase in ICT raises output by about 0.06 percent. This corresponds to a net–rate of return to ICT investment of more than 50%. These apparent excess returns are likely to be due to unobserved complementary expenses such as adjustment cost, innovation efforts, training or intangible assets, but they may also reflect differences between firms in their ability to exploit the potential benefits of ICT. The findings from this study, for example, the availability of skilled workers are a prerequisite for using ICT productively. Therefore, the exploration of adjustment costs and of relevant firm characteristics and strategies related to ICT use are important issues for future research on the productivity and welfare impacts of the 'Information Economy'.

References

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies* **58**(2): 277–297.

Arellano, M. and Bond, S. (1998). *Dynamic Panel Data Estimation Using DPD98 for GAUSS: A Guide for Users.* ftp://ftp.cemfi.es/pdf/papers/ma/dpd98.pdf.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models, *Journal of Econometrics* **68**(1): 29–51.

Bertschek, I. and Kaiser, U. (2001). Productivity Effects of Organizational Change: Microeconometric Evidence, *ZEW Discussion Paper No. 01-32*, Centre for European Economic Research, Mannheim.

Biscourp, P., Crepon, B., Heckel, T. and Riedinger, N. (2002). How Do Firms Respond to Cheaper Computers?, CREST-INSEE, mimeo.

Black, S. E. and Lynch, L. M. (2001). How to Compete: The Impact of Workplace Practices and Information Technology on Productivity, *Review of Economics and Statistics* **83**(3): 435–445.

Blundell, R. and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics* **87**(1): 115–143.

Blundell, R. and Bond, S. (2000). GMM Estimation with Persistent Panal Data: An Application to Production Functions, *Econometric Reviews* **19**(3): 321–340.

Blundell, R., Bond, S. and Meghir, C. (1996). Econometric Models of Company Investment, *in* L. Mátyás and P. Sevestre (eds), *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, Kluwer Academic Publishers, Dordrecht, pp. 685–710.

Bresnahan, T. and Greenstein, S. (1996). Technical Progress and Co-Invention in Computing and in the Uses of Computers, *Brookings Papers on Economic Activity, Microeconomics* pp. 1–77.

Bresnahan, T. F., Brynjolfsson, E. and Hitt, L. M. (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence, *Quarterly Journal of Economics* **117**(1): 339–376.

Brynjolfsson, E. (1994). Technology's True Payoff, *Informationweek, October 10th* pp. 34–36.

Brynjolfsson, E. and Hitt, L. (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.

Brynjolfsson, E. and Hitt, L. (1996). Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending, *Management Science* **42**(4): 541–558.

Brynjolfsson, E. and Hitt, L. M. (1998). Information Technology and Organizational Design, *mimeo*, MIT Sloan School of Management. http://ebusiness.mit.edu/erik/ITOD.pdf.

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond Computation: Information Technology, Organizational Transformation and Business Performance, *Journal of Economic Perspectives* **14**(4): 23–48.

Brynjolfsson, E. and Hitt, L. M. (2001). Computing Productivity: Firm-Level Evidence, *mimeo*, University of Pennsylvania, Wharton School.

Brynjolfsson, E. and Yang, S. (1999). The Intangible Costs and Benefits of Computer Investments: Evidence from the Financial Markets, *Proceedings of the International Conference on Information Systems, Atlanta, Georgia*, MIT Sloan School.

Caroli, E. and van Reenen, J. (2001). Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments, *Quarterly Journal of Economics* **116**(4): 1449–1492.

Christensen, L. R. and Jorgenson, D. W. (1969). The Measurement of U.S. Real Capital Input, 1929–1967, *Review of Income and Wealth* **15**(4): 293–320.

EITO (2001). *European Information Technology Observatory 2001*, EITO, Frankfurt/Main.

Fraumeni, B. M. (1997). The Measurement of Depreciation in the U.S. National Income and Product Accounts, *Survey of Current Business* **77**(7): 7–23.

Gouriéroux, C. and Monfort, A. (1995). *Statistics and Econometric Models*, Vol. 1, Cambridge University Press, Cambridge.

Greenan, N. and Mairesse, J. (1996). Computers and Productivity in France: Some Evidence, *Working Paper No. 5836*, National Bureau of Economic Research.

Greenan, N., Mairesse, J. and Topiol-Bensaid, A. (2001). Information Technology and Research and Development Impacts on Productivity and Skills: Looking for Correlations on French Firm Level Data, *NBER Working Paper No. 8075*, National Bureau of Economic Research.

Griliches, Z. (1994). Productivity, R&D, and the Data Constraint, *American Economic Review* **84**(1): 1–23.

Griliches, Z. and Hausman, J. A. (1986). Errors in Variables in Panel Data, *Journal of Econometrics* **31**(1): 93–118.

Griliches, Z. and Mairesse, J. (1998). Production Functions: The Search for Identification, *in* S. Strøm (ed.), *Econometrics and Economic Theory in the 20th Century — The Ragnar Frisch Centennial Symposium*, Cambridge University Press, pp. 169–203.

Hall, B. H. and Mairesse, J. (1995). Exploring the Relationship Between R&D and Productivity in French Manufacturing Firms, *Journal of Econometrics* **65**(1): 263–293.

Hempell, T. (2002). Does Experience Matter? Innovation and the Productivity of ICT in German Services, *ZEW Discussion Paper 02–43*, Centre for European Economic Research, Mannheim. (ftp://ftp.zew.de/pub/zew-docs/dp/dp0243.pdf).

Hempell, T. (2003). Do Computers Call for Training? Firm-Level Evidence on Complementarities Between ICT and Human Capital Investments, *ZEW Discussion Paper 03–20*, Centre for European Economic Research, Mannheim. (ftp://ftp.zew.de/pub/zew-docs/dp/dp0320.pdf).

Hoffmann, J. (1998). Problems of Inflation Measurement in Germany, *Discussion Paper No. 01-98*, Economic Research Centre of the Deutsche Bundesbank.

Jorgenson, D. W. and Stiroh, K. (1995). Computers and Growth, *Economics of Innovation and New Technology* **3**(3-4): 295–316.

Kaiser, U. (2000). A Note on the Calculation of Firm-Specific and Skill-Specific Labor Costs from Firm-Level Data, *ZEW Discussion Paper 00–08*, Centre for European Economic Research, Mannheim.

Klette, T. J. and Griliches, Z. (1996). The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous, *Journal of Applied Econometrics* **11**(4): 343–361.

Lehr, B. and Lichtenberg, F. (1999). Information Technology and its Impact on Productivity: Firm-Level Evidence from Government and Private Data Sources, *Canadian Journal of Economics* **32**(2): 335–362.

Licht, G. and Moch, D. (1999). Innovation and Information Technology in Services, *Canadian Journal of Economics* **32**(2): 363–383.

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

Moulton, B. R., Parker, R. P. and Seskin, E. P. (1999). A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts — Definitional and Classificational Changes, *Survey of Current Business* **79**(8): 7–20.

Müller, A. A. (1998). *Kapitalstock und Produktionspotential im privaten und öffentlichen Sektor Deutschlands*, Nomos, Baden-Baden.

OECD (2000a). *A New Economy? The Changing Role of Innovation and Information Technology in Growth*, OECD, Paris.

OECD (2000b). The Service Economy, *Business and Industry Policy Forum Series*, STI/OECD.

Schreyer, P. (2000). The Contribution of Information and Communication Technology to Output Growth: A Study of the G7 Countries, *STI Working Paper 2000/2*, OECD.

Statistisches Bundesamt (2001). *Statistical Yearbook for the Federal Republic of Germany*, Metzler-Poeschel, Stuttgart.

Stiroh, K. (2002a). Are ICT Spillovers Driving the New Economy?, *Review of Income and Wealth* **48**(1): 33–57.

Stiroh, K. J. (2002b). Reassessing the Impact of IT in the Production Function: A Meta-Analysis, *mimeo*, Federal Reserve Bank of New York.

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

Yang, S. and Brynjolfsson, E. (2001). Intangible Assets and Growth Accounting: Evidence from Computer Investments, *mimeo*, New York University.

Appendix

A ICT–induced quality improvements in a simple partial equilibrium model

Supplementing the arguments in section 2.1, this part of the appendix uses a simple partial equilibrium model to show how quality improvements induced by ICT use may implicitly be taken into account in the ICT coefficients obtained from production function approaches at the firm–level.

Consider a simplified version of the production function stated in eq. (2) for a firm operating in some given industry I for which output deflators P_{It} are available:

$$y_{it} = \gamma_1 l_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it}, \tag{14}$$

where y_{it} is real output based on industry deflators and small letters denote values in logarithms. Suppose further that firms do not follow the same pricing strategies such that aggregate industry deflators p_{It} are an imperfect measure of output at the firm level. The heterogeneity in pricing induces an aggregation error such that (imperfectly) measured output $\tilde{y}_{it}$ can be defined as a combination of true output and a measurement error:

$$\tilde{y}_{it} \equiv y_{it} + p_{it} - p_{It}, \tag{15}$$

where $p_{it} - p_{It}$ is the measurement error due to aggregation. Inserting (15) in (14) gives:

$$\tilde{y}_{it} = \gamma_1 l_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it} + p_{it} - p_{It}. \tag{16}$$

The output of firms may, however, differ not only in terms of prices but also in quality. In particular, as argued in the main text (section 2.1, p. 3), the use of new technologies may be particularly suited to improve ancillary aspects of product quality, like speed of delivery, extended shopping facilities or after sales support. To focus on the role of ICT for product quality, consider the simplest case in which product quality Q is determined by the intensity of ICT used in the production process (defined as the fraction of ICT over non–ICT capital) such that:[73]

$$\begin{aligned} Q_{it} &= = B \cdot \left(\frac{ICT_{it}}{K_{it}}\right)^{\omega} \quad \text{or} \quad q_{it} = b + \omega(ict_{it} - k_{it}) \\ \text{and} \quad Q_t &= B \cdot \left(\frac{ICT_{It}}{K_{It}}\right)^{\omega} \quad \text{or} \quad q_{It} = b + \omega(ict_{It} - k_{It}), \end{aligned} \tag{17}$$

[73]The intensity of ICT could equivalently be defined as the share of ICT input in output produced. This makes the model slightly more involved without changing the main results.

where small letters indicate logarithms and subscripts I denote the corresponding mean values at the industry level.[74] The marginal contributions of ICT to output quality are proportional to ω and, if ω is restricted to fall into the interval $0 \leq \omega \leq 1$, the marginal contributions of relative ICT input to product quality are positive and decreasing in ICT intensity.

For the demand side, I use a slightly extended version of the model proposed by Klette and Griliches (1996), denoted by KG in the remainder. The demand for goods from firm i at t is given by:

$$Y_{it}^D = Y_{It}^D \cdot \left(\frac{P_{it}}{P_{It}} \frac{Q_{It}}{Q_{it}}\right)^{\eta} \quad \text{or} \quad y_{it}^D = y_I t^D + \eta(p_{it} - p_{It} - (q_{it} - q_{It})). \tag{18}$$

That is, the demand for output produced by firm i in period t depends on total demand for output produced in the corresponding industry Y_{It}^D and the price P_{it} relative to the price level at the industry level P_{It}. The extension of the KG–model consists in the correction of prices for differentials in output quality Q_{it}/Q_{It}. This extension is based on the idea that utility–maximising consumers take heterogenous output quality into account when comparing prices. The parameter $\eta < 0$ reflects the elasticity of demand with respect to relative prices.[75]

In equilibrium with $y_{it}^D = y_{it}$ and $y_{It}^D = y_{It}$, inserting eq. (15) and (17) in (18) yields:

$$\begin{aligned} \tilde{y}_{it} &= y_{It} + (1+\eta)(p_{it} - p_{It}) + \eta\omega(q_{it} - q_{It}) \\ &= y_{It} + (1+\eta)(p_{it} - p_{It}) + \eta\omega(ict_{it} - ict_{It}) - \eta\omega(k_{it} - k_{It}). \end{aligned} \tag{19}$$

Solving (19) for $p_{it} - p_{It}$, inserting in (16) and rearranging yields:

$$\begin{aligned} \tilde{y}_{it} = \; & \frac{1+\eta}{\eta} \left[\gamma_1 l_{it} + \gamma_2 ict_{it} + \gamma_3 k_{it}\right] - \frac{1}{\eta} y_{It} \\ & -\omega(1+\eta)(ict_{it} - k_{it}) + \omega(1+\eta)(ict_{It} - k_{It}). \end{aligned} \tag{20}$$

Eq. (20) summarises the main theoretical issues of estimating the productivity ICT as discussed in section 2.1 (p. 3). The first part of the equation is basically identical to the KG–model. It shows that if prices vary between firms due to imperfectly competitive markets (with $-1 > \eta > -\infty$), the estimates of the input elasticities obtained from a production function estimation as of eq. (14) must be interpreted as reduced–form

[74] In addition, all firms are assumed to be sufficiently small such that the impact of changes in one variable in one firm has an negligible effect on industry averages.

[75] Note that strong competition is mirrored by high (absolute) values for η, such that a small price deviation from industry average causes a strong decrease in demand for goods from firm i.

estimates that underestimate the true input elasticities by the factor $\eta/(\eta+1)$.[76]

The second line of eq. (20) corresponds to the extension of the KG–model and displays the impacts of quality improvements on the estimated reduced–form elasticities. The reduced–form elasticities can be interpreted more easily by rearranging eq. (20) to:

$$\begin{aligned}\tilde{y}_{it} &= \left[\frac{1+\eta}{\eta}\gamma_1\right] l_{it} + \left[\frac{1+\eta}{\eta}\gamma_2 - \omega(1+\eta)\right] ict_{it} + \left[\frac{1+\eta}{\eta}\gamma_3 + \omega(1+\eta)\right] k_{it} \\ &\quad -\frac{1}{\eta} y_{It} + [\omega(1+\eta)]\,(ict_{It} - k_{It}) + \epsilon_{it}.\end{aligned}$$

This equation shows that the higher the impact of ICT intensity on output quality, i.e. the higher ω, the higher will be the reduced–form estimate of ICT. Even though this term does not corresponds to the output contributions of ICT in a narrow sense (measured by γ_2), this broader measure also takes into account welfare effects from improved output quality. However, this quality–impact is closely linked to the competition parameter η. The more competitive markets are (i.e. the more negative η), the stronger are the impacts of quality improvements on the reduced–form estimate of the ICT elasticity. Moreover, as pointed out by the GK–model, higher absolute values for η also imply a lower bias of the reduced–form elasticities induced by the term $(\eta+1)/\eta$.[77]

B GMM estimation of the production function

Referring to equation (2), the basis of the Generalised Method of Moments (GMM) approach employed in this paper follows basically the suggestions by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998). It consist in assuming the choice for the $k=3$ inputs in the initial period $x_{i1} = (l_{i1}, ict_{i1}, k_{i1})$ to be uncorrelated with the residuals $u_{it} = \eta_i + \epsilon_{it}$ in the subsequent periods $E[x_{i1}\epsilon_{it}] = 0$, for $t = 2, \ldots, T$, where T denotes the number of periods.[78] This assumption entails the following moment conditions:

$$E[x'_{i,t-s}\Delta u_{it}] = 0 \qquad \text{for } t = 3, \ldots, T \text{ and } t-1 \geq s \geq 2. \tag{21}$$

Note that by first–differencing of u_{it}, the fixed–effect η_i, which may be correlated with the inputs, is cancelled out. The system of equations (21) can be summarised in matrix

[76] An empirical approach to assess the size of this bias is to include industry output y_t in the regression to get an estimate of the coefficient η.

[77] An empirical strategy to obtain the parameters γ_1, γ_2, γ_3, η and ω would be to regress measured firm–level output $\tilde{y}_{it}$ on firm–level inputs l_{it}, ict_{it} and k_{it} and on industry–level data y_{It}, ict_{It} and k_{It}. The structural coefficient η could then be recovered from the coefficient of y_t. In combination with the estimate for $ict_t - k_t$, this would allow for obtaining also ω. Finally, with η and ω known, also the elasticities $\gamma_1, \gamma_2, \gamma_3$ can be deduced from the estimates. For the analyses of this paper, however, the corresponding industry–level data for Germany are not available. However, in the empirical application, interacted time and industry dummies control for the industry–specific heterogeneity of y_{It}, ict_{It} and k_{It}.

[78] Note that in eq. (2), it is assumed that ϵ_{it} are serially uncorrelated.

notation in the following way:

$$E[Z_i' D u_i] = 0 \tag{22}$$

with[79]

$$Z_i = Z_i^D = \underset{T-1 \,\times\, k(T-2)(T-1)/2}{\begin{pmatrix} x_{i1} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & x_{i1} & x_{i2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x_{i1} & x_{i2} & \cdots & x_{i,T-2} \end{pmatrix}} \tag{23}$$

$$D = \underset{T-1 \,\times\, T}{\begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}}$$

$$u_i = (u_{i1}, u_{i2}, \ldots, u_{iT})'.$$

The dimensions of Z_i^D show that eq. (22) comprises $k(T-2)(T-1)/2$ moment conditions.[80] After solving eq. (2) for u_{it} and inserting in the moment conditions (21), the residuals depend on the data (y, x) as well as the parameters $\phi = (\gamma_1, \gamma_2, \gamma_3, \lambda_{12}, \ldots, \lambda_{JT})$, where J denotes the number of industries such that eq. (22) can be written as a function:

$$E[Z_i' D u_i] = E[\psi(y, x, \phi)] = 0. \tag{24}$$

By the analogy principle, the expected value of the population is replaced by the sample mean such that we can define $b_N(\phi) = N^{-1} \sum_{i=1}^{N} \psi(y, x, \phi)$ with N denoting the number of firms in the sample. For given sample values (y, x), the GMM estimator $\hat{\phi}(A)$ associated with a matrix A is the choice of ϕ that minimises the quadratic form:

$$\hat{\phi}(A) = \arg\min_{\phi} b_N(\phi)' \, A \, b_N(\phi), \tag{25}$$

[79]Note that $x_{it} = (l_{it}, ict_{it}, k_{it})$ has $k = 3$ columns, such that also each zero entry in the Z_i–matrix represents a vector $(0, 0, 0)$. Similarly, the apparent number of columns of the matrix Z must be multiplied by $k = 3$.

[80] Note that serial correlation of the errors ϵ_{it} may be at odds with these moment conditions. To see that, suppose that $\epsilon_{it} = \rho\epsilon_{it} + e_{it}$ and insert this into the moment condition of eq. (21) for $s = 2$. It then follows that $E[\Delta\epsilon_{it} X_{i,t-2}] = (\rho - 1)E[\epsilon_{i,t-1} X_{i,t-2}] = \rho(\rho - 1)E[\epsilon_{i,t-2} X_{i,t-2}] \neq 0$ unless $\rho = 1$ or $\rho = 0$. Thus, unlike in the case of OLS, serial correlation may harm not only the efficiency but also the consistency of the estimates in the case of GMM estimation since the consistency of the GMM estimates hinges on the validity of the underlying moment conditions. However, the validity of the instruments can also be tested directly using the Sargan statistic, which is discussed further below. This test is a further measure of how strongly potential serial correlation (among other factors) impacts the moment conditions underlying the GMM estimates, and only the combination of the test for serial correlation and the Sargan statistic will give a comprehensive picture of the validity of the moment conditions.

where any choice of the $(T-1 \times T-1)$ weighting matrix A yields a consistent (though not efficient) estimator. For a linear model of the form $y_{it} = x_{it}\beta + u_{it}$ as in eq. (2), this minimisation problem is solved by:

$$\hat{\phi}(A) = \left[\left(\sum_{i=1}^{N} x_i^{*\prime} Z_i \right) A \left(\sum_{i=1}^{N} Z_i' x_i^* \right) \right]^{-1} \left[\left(\sum_{i=1}^{N} x_i^{*\prime} Z_i \right) A \left(\sum_{i=1}^{N} Z_i' y_i^* \right) \right]. \tag{26}$$

where $x_i^* = Dx_i$. The optimal choice for A, which yields an efficient $\hat{\phi}$, is given by $A^* = Var(\psi(y, x, \phi))^{-1}$. Since this variance–covariance matrix is not known, a two–step procedure can be applied: in the first step, an arbitary weighting matrix A is used[81] to obtain a consistent estimate of ϕ. The one–step coefficients are then used to calculate the estimated first–differenced residuals $\hat{v}_i^* = D\hat{v}_i$ where $\hat{v}_i$ are the estimated errors obtained from the level equation (2). A more efficient weighting matrix for the second–step estimation is then:

$$A_N = \left(N^{-1} \sum_{i=1}^{N} Z_i' \hat{v}_i^* \hat{v}_i^{*\prime} Z_i \right)^{-1}. \tag{27}$$

A convenient feature of the GMM estimator is that for the efficient weighting matrix A_N for any given Z_i, the minimised value of the distance function $b_N(\phi)' \, A_N \, b_N(\phi)$ from eq. (25) is asymptotically χ^2_{r-k}–distributed with the number of degrees of freedom equal to the number of overidentifying restrictions, i.e. the difference between the number of columns of Z_i (denoted by r^D) and the number of columns of x_i (denoted by k). Thus, the validity of the employed instruments can be tested empirically using the Sargan test–statistic:[82]

$$S = S^D = \left(\sum_{i=1}^{N} \hat{v}_i^{*\prime} Z_i \right) A_N \left(\sum_{i=1}^{N} Z_i' \hat{v}_i^{*} \right) \overset{asym}{\sim} \chi^2_{r-k}. \tag{28}$$

The SYS–GMM estimator is an extension of the GMM estimator above. The main idea is to find variables that are uncorrelated with the fixed effects η_i and that thus can be used as instruments for the equation in levels. Arellano and Bover (1995) consider the case where the covariance between the explanatory variables x_{it} and the individual effects η_i are constant over time, such that $E(x_{it}\eta_i) = E(x_{is}\eta_i) \; \forall s$.[83] Together with the moment conditions of eq. (21), this gives the $(T-2)$ additional moment conditions for the equations in levels:[84]

[81]The DPD98 programme used in this paper employs the matrix $A = DD'$ for this purpose.

[82]For the regression GMM[-2] in column (4) of Table 1, e.g., the number of moment conditions r (with $T = 6$) is $r = 3 \cdot (6-2)(6-1)/2 = 30$ such that the corresponding Sargan statistic has $r - k = 30 - 3 = 27$ degrees of freedom.

[83]As shown by Blundell and Bond (1998), the joint stationarity of the dependent and the independent variables is a sufficient, yet not necessary prerequisite for these restrictions to hold.

[84]Arellano and Bover (1995) show that, given the moment conditions of eq. (21), further moment conditions of the type $E(\Delta x'_{i,t-s} u_{it}) = 0$ are redundant since, e.g. $E(\Delta x'_{i,t-1} u_{it}) - E(\Delta x'_{i,t-1} u_{i,t-1}) =$

$$E(\Delta x'_{i,t-1} u_{it}) = 0, \quad t = 2, \ldots, T. \tag{29}$$

These additional moment restrictions can be implemented by letting $u_i^S = \binom{Du_i}{u_i}$ and:

$$Z_i^S = \underbrace{\begin{pmatrix} Z_i^D & 0 & 0 & \cdots & 0 \\ 0 & \Delta x_{i2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \Delta x_{i,T-1} \end{pmatrix}}_{T-1 \,\times\, k(T-2)(T+1)/2}, \tag{30}$$

where Z_i^D is the instrument matrix (23) for the equation in first differences. Thus, the moment conditions of the system GMM estimator are:

$$E[Z_i^{S\prime} u_i^S] = 0. \tag{31}$$

The validity of the additional instruments obtained from the orthogonality conditions (29) can be tested using a Difference Sargan test.[85] Since the set of instruments used for the equation in first differences Z_i^D is a strict subset of the set of instruments used for the system of equations in levels and in first–differences, the corresponding Difference Sargan statistic is:

$$S^\Delta = S^S - S^D \quad \overset{asym}{\sim} \chi^2_{r^S - r^D}, \tag{32}$$

where S^S and S^D are the Sargan statistics obtained for the system GMM and first difference GMM correspondingly, and r^S and r^D are the corresponding number of columns of the instrument matrices Z^S and Z^D.[86]

C Imposing and testing common factor restrictions by minimum distance

In order to obtain the structural parameters $\theta = (\rho, \gamma_1, \gamma_2, \gamma_3)$ of eq. (2a), a two–step procedure is applied. In the first step, the reduced–form equation (3) with parameters $\pi = (\pi_1, \ldots, \pi_7)$ is estimated by SYS–GMM. In the second step, then, the estimates of the parameters π_i are used to deduce the structural parameters as of eq. (2a) by testing and imposing the corresponding common factor restrictions $\pi_1 = \rho$, $\pi_2 = \gamma_1$, $\pi_3 = \gamma_2$, $\pi_4 = \gamma_3$, $\pi_5 = -\gamma_1 \cdot \gamma_2$, $\pi_6 = -\gamma_1 \cdot \gamma_3$ and $\pi_7 = -\gamma_1 \cdot \gamma_4$ using a minimum distance (or

$E(x'_{i,t-1}\Delta u_{it}) - E(x'_{i,t-2}\Delta u_{it})$ such that the first term $E(\Delta x'_{i,t-1}u_{it})$ is just a combination of the last three terms which are already implied by conditions (21) and (29).

[85]See Arellano and Bond (1991), e.g., for further details on this test.

[86]For the exemplified case, $r^S - r^D = k(T-2)$.

asymptotic least squares) procedure.

Let the function $h : \Re^4 \rightarrow \Re^7$ relate θ to π such that $\pi = h(\theta) = (\rho, \gamma_1, \gamma_2, \gamma_3, -\rho\gamma_1, -\rho\gamma_2, -\rho\gamma_3)'$. Using this function, the focus of interest are thus the structural parameters $\theta = (\rho, \gamma_1, \gamma_2, \gamma_3)'$ that minimise the norm $\pi - h(\theta)$.[87]

In order to simplify calculations, I additionally specify the function $g : \Re^7 \rightarrow \Re^7$ such that $g(\pi) = (\pi_1, \pi_2, \pi_3, \pi_4, -\pi_5/\pi_1, -\pi_6/\pi_1, -\pi_7/\pi_1)'$ which makes $g(h(\theta)) = (\rho, \gamma_1, \gamma_2, \gamma_3, \gamma_1, \gamma_2, \gamma_3)'$ linear in the components of θ. For given reduced–form estimates $\hat{\pi}$, the structural parameter estimates $\hat{\theta}$ are then imposed to minimise the quadratic distance:

$$\hat{\theta} = \arg\min_{\theta} \, [g(\hat{\pi}) - g(h(\theta))]' \, \hat{\Omega}^{-1} \, [g(\hat{\pi}) - g(h(\theta))] \tag{33}$$

$$= \arg\min_{\theta} \begin{pmatrix} \hat{\pi}_1 - \rho \\ \hat{\pi}_2 - \gamma_1 \\ \hat{\pi}_3 - \gamma_2 \\ \hat{\pi}_4 - \gamma_3 \\ -\hat{\pi}_5/\hat{\pi}_1 - \gamma_1 \\ -\hat{\pi}_6/\hat{\pi}_1 - \gamma_2 \\ -\hat{\pi}_7/\hat{\pi}_1 - \gamma_3 \end{pmatrix}' \hat{\Omega}^{-1} \begin{pmatrix} \hat{\pi}_1 - \rho \\ \hat{\pi}_2 - \gamma_1 \\ \hat{\pi}_3 - \gamma_2 \\ \hat{\pi}_4 - \gamma_3 \\ -\hat{\pi}_5/\hat{\pi}_1 - \gamma_1 \\ -\hat{\pi}_6/\hat{\pi}_1 - \gamma_2 \\ -\hat{\pi}_7/\hat{\pi}_1 - \gamma_3 \end{pmatrix}, \tag{34}$$

with $\hat{\Omega} = [\partial g(\hat{\pi})/\partial \pi'] \; \widehat{var(\hat{\pi})} \; [\partial g(\hat{\pi})/\partial \pi']'$, where $\widehat{var(\hat{\pi})}$ denotes the estimated variance covariance matrix of the unrestricted parameters $\hat{\pi}$ and

$$\partial g(\hat{\pi})/\partial \pi' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{\hat{\pi}_5}{\hat{\pi}_1^2} & 0 & 0 & 0 & \frac{-1}{\hat{\pi}_1} & 0 & 0 \\ \frac{\hat{\pi}_6}{\hat{\pi}_1^2} & 0 & 0 & 0 & 0 & \frac{-1}{\hat{\pi}_1} & 0 \\ \frac{\hat{\pi}_7}{\hat{\pi}_1^2} & 0 & 0 & 0 & 0 & 0 & \frac{-1}{\hat{\pi}_1} \end{pmatrix}. \tag{35}$$

The asymptotic variance matrix of the estimate $\hat{\theta}$ is given by:

$$var(\hat{\theta}) = \left([\partial g(h(\theta))/\partial \theta']' \, \Omega^{-1} \, [\partial g(h(\theta))/\partial \theta'] \right)^{-1} \tag{36}$$

where the Jacobian $\partial g(h(\theta))/\partial \theta'$ is just a 7×4–matrix of zeros and ones by the construction of the function $g(\bullet)$:

[87] The following specifications and derivations follow Blundell et al. (1996), Wooldridge (2002, ch. 14) and the more general discussion of asymptotic least squares by Gouriéroux and Monfort (1995, ch. 9 and 10).

$$\partial g(h(\theta))/\partial\theta' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{37}$$

Finally, the validity of the common factor restrictions that link the structural equation (2a) to the reduced–form specification (3) can be tested. For large cross–sections N, the minimised value of the distance function of eq. (33) has an asymptotic χ^2–distribution with three degrees of freedoms since:

$$[g(\hat{\pi}) - g(h(\hat{\theta}))]' \, \hat{\Omega}^{-1} \, [g(\hat{\pi}) - g(h(\hat{\theta}))] = S^{MD} \quad \overset{asym}{\sim} \quad \chi^2_{(r^{red} - r^{struct})}, \tag{38}$$

where $r^{red} = 7$ represents the number of the reduced–form parameters comprised by π and $r^{struc} = 4$ is the number of structural parameters contained in θ .

D Tables

Table 7: Detailed statistics for the full sample (4939 obs. for 1177 firms)

					percentiles			per employee	
	mean	std.	min.	max.	10%	50%	90%	mean	median
value added*	54.541	717.21	0.118	27,380	0.362	2.647	40.705	121,917	60,307
employees	614.563	9379	1	310,792	7	42	506	—	—
ICT capital*	5.058	131.25	< 0.001	6,537	0.006	0.488	0.923	3,946	1,392
non–ICT capital*	102.387	1833.645	0.001	60,340	0.061	1.107	56.360	218,492	24,979
East (dummy)	0.421	0.494	0	1	0	0	1	—	—

*measured in € million, except for values per employee

Table 8: Detailed statistics for the small sample (1847 obs. for 578 firms)

					percentiles			per employee	
	mean	std.	min.	max.	10%	50%	90%	mean	median
value added*	18.513	88.658	0.032	1,124	0.362	2.306	22.821	120,448	58,857
employees	183.673	613.885	1	7,200	7	36	300	—	—
ICT capital*	0.362	1.466	< 0.001	30.855	0.006	0.041	0.529	3,106	1,240
non–ICT capital*	24.946	102.911	0.003	1,450	0.062	0.900	41.378	228,230	24,852
East (dummy)	0.444	0.497	0	1	0	0	1	—	—
% university	0.191	0.264	0	1	0	0.061	0.667	—	—
% vocational	0.566	0.303	0	1	0.129	0.615	0.944	—	—

*measured in € million, except for values per employee

Table 9: Detailed statistics for the extended sample (5107 obs. for 1222 firms)

					percentiles			per employee	
	mean	std.	min.	max.	10%	50%	90%	mean	median
value added*	53.145	705.526	0.012	27,380	0.351	2.495	39.805	122,198	60,575
employees	596.7	9224	1	310,792	7	40	499	—	—
ICT capital*	4.892	129.075	< 0.001	6,537	0.004	0.045	0.892	3,801	1,302
non–ICT capital*	100.300	1,803	0.001	60,340	0.060	1.060	55.375	226,947	25,574
East (dummy)	0.422	0.494	0	1	0	0	1	—	—

*measured in € million, except for values per employee

Table 10: Comparison of the different samples and the population by industries

		samples						population*
		full		small		extended		
industry	nace–digit	#	%	#	%	#	%	%
wholesale trade	51	163	13.9	83	14.4	172	14.1	10.6
retail trade	50, 52	183	15.6	87	15.1	190	15.6	31.3
transport and postal services	60–63, 64.1	210	17.8	104	18.0	222	18.2	11.7
electr. processing and telecom.	72, 62.2	100	8.5	44	7.6	100	8.2	3.4
consultancies	74.1, 74.4	100	8.5	48	8.3	103	8.4	12.1
technical services	73, 74.2, 74.3	142	12.1	75	13.0	152	11.7	10.7
other business–rel. services	70, 71, 74.5-.8, 90	279	23.7	137	23.7	292	23.9	20.3
total		1177	100	578	100	1222	100	100

*German service firms with 5 and more employees in 1999.
Source: German Statistical Office, ZEW and own calculations

Table 11: Comparison of the different samples and the population by size classes

	full sample		small sample		ext. sample		population*	
size class (# employees)	#	%	#	%	#	%	% firms	% sales
5–9	189	16.1	88	15.2	205	16.8	57.6	9.4
10–19	189	16.1	105	18.2	206	16.9	24.0	9.9
20–49	246	20.9	137	23.7	254	20.8	11.7	9.7
50–99	156	13.3	87	15.1	156	12.8	3.5	6.9
100–199	167	14.2	76	13.2	168	13.8	1.6	6.0
200–499	102	8.7	48	8.3	102	8.3	1.0	7.0
500 and more	128	10.9	37	6.4	131	10.7	0.6	51.1
total	1177	100	578	100	1222	100	100	100

*German service firms with 5 and more employees in 1999.
Source: German Statistical Office, ZEW and own calculations

Table 12: One–step results for the ICT–augmented production function

	Dep. Variable: value added (logs)				
	(3)	(4)	(5)	(6)	(7)
inputs (logs)	GMM[–1]	GMM[–2]	SYS–GMM	SYS–GMM	SYS–GMM
	1st diff.	1st diff.	reference	not interact.	extended
labour	0.515***	0.247	0.707***	0.737***	0.723***
	(0.174)	(0.158)	(0.073)	(0.074)	(0.075)
ICT capital	0.053	0.069	0.114**	0.081*	0.052
	(0.043)	(0.041)	(0.046)	(0.049)	(0.032)
non–ICT capital	0.191	0.366	0.148***	0.155***	0.166***
	(0.198)	(0.208)	(0.046)	(0.049)	(0.046)
East–Germany	—	—	-0.340***	-0.343***	-0.375***
			(0.051)	(0.053)	(0.049)
observations	3762	3762	4939	4939	5107
firms	1177	1177	1177	1177	1222
R–square	0.218	0.137	0.843	0.839	0.836
Wald statistics [df]:					
inputs	14.7[3]	13.9[3]	446[4]	441[4]	494[4]
time and ind. dummies	108[35]	130[35]	586[41]	488[11]	583[41]
errors (p–values):					
AR(1)	0.010	0.000	0.002	0.003	0.002
AR(2)	0.118	0.042	0.028	0.025	0.030

***,**,*=significant on the 1,5 and 10 per cent level
Results are based on the one–step estimation corresponding to table 1. See footnotes on this table for further details.

Table 13: Dynamic specification of the production function

	Dep. Variable: log(value added)			
	(1) static (one–step)	(2) static (two–step)	(2) dynamic (one–step)	(3) dynamic (two–step)
log(value added$_{t-1}$)	—	—	0.638***	0.105***
			(0.135)	(0.084)
log(labour$_t$)	0.768***	0.722***	0.391**	0.352***
	(0.099)	(0.079)	(0.159)	(0.123)
log(ICT$_t$)	0.090**	0.057*	0.161	0.136
	(0.042)	(0.030)	(0.136)	(0.079)
log(non–ICT$_t$)	0.109	0.166***	0.294	0.194
	(0.069)	(0.051)	(0.243)	(0.179)
log(labour$_{t-1}$)	—	—	-0.066	-0.051
			(0.165)	(0.138)
log(ICT$_{t-1}$)	—	—	-0.097	-0.031
			(0.074)	(0.045)
log(non–ICT$_{t-1}$)	—	—	-0.266	-0.182
			(0.224)	(0.174)
R–square	0.829	—	0.950	—
Wald stat.[df]				
inputs	193[3]	292[3]	790[7]	1428[7]
time and ind. dummies	429[41]	530[41]	59[34]	62[34]
Sargan (p–values)	—	0.021		0.128
erros (p–values)				
AR(1)	0.001	0.003	0.003	0.001
AR(2)	0.026	0.046	0.333	0.332

***, **, * = significant at the 1, 5 and 10 per cent level respectively

SYS-GMM estimates include a constant, a regional dummy variable for East–German firms as well as interacted industry and year dummy variables. Robust standard errors reported in brackets. The underlying sample consists of an unbalanced panel with 708 firms and 3532 observations covering the years 1994–1999.

Table 14: One–step results for the translog production function

	Dep. Variable: value added (logs)	
inputs (log)	SYS–GMM full	SYS–GMM extended
labour	1.044***	1.006***
	(0.256)	(0.277)
ICT capital	0.044	0.070
	(0.156)	(0.119)
non–ICT capital	0.214**	0.224**
	(0.085)	(0.090)
labour^2	-0.041*	-0.040
	(0.024)	(0.027)
ICT capital^2	0.002	0.011**
	(0.012)	(0.005)
non–ICT capital^2	0.006*	0.004
	(0.011)	(0.011)
labour * ICT	0.047	0.042*
	(0.030)	(0.023)
labour * non–ICT	-0.020	-0.018
	(0.018)	(0.020)
ICT * non–ICT	-0.006	-0.007
	(0.015)	(0.012)
East–Germany	-0.512***	-0.513***
	(0.188)	(0.189)
[.2em] observations	4939	5107
firms	1177	1222
R–square	0.850	0.846
[.2em] Wald–statistics[df]:		
all inputs	4,784[10]	4,039[10]
additional inputs†	4.96[6]	34.17[6]
time and ind. dummies	531.5[41]	528.5[41]
[.2em] errors (p–values)		
AR(1)	0.002	0.004
AR(2)	0.008	0.014

***,**,* = significant at the 1, 5 and 10 per cent level
Results are based on the one–step SYS–GMM
corresponding to Table 5.
See comments on this table for further details.
†refers to additional inputs *not* included in Cobb–Douglas specification.