The Sensitivity of Productivity Estimates: Revisiting Three Important Debates

Johannes Van Biesebroeck*
University of Toronto and NBER

August 15, 2006

[**forthcoming in *Journal of Business and Economic Statistics***]

Abstract

Researchers interested in estimating productivity can choose from an array of methodologies, each with its strengths and weaknesses. This study compares productivity estimates and evaluates the extent to which the conclusions from three important productivity debates in the economic development literature are sensitive to the choice of estimation method. Five widely used techniques are considered, two nonparametric and three parametric: index numbers, data envelopment analysis, instrumental variables estimation, stochastic frontiers, and semiparametric estimation. Using data on manufacturing firms in two developing countries, Colombia and Zimbabwe, I find that the different methods produce surprisingly similar productivity estimates when the measures are compared directly, even though the estimated input elasticities vary widely. Furthermore, the methods reach the same conclusions for two of the debates, supporting endogenous growth effects and showing that firm level productivity changes are an important contributor to aggregate productivity growth. In terms of the third debate, the parametric productivity measures provide evidence of learning-by-exporting, while the nonparametric measures that allow for a different production technology for exporters and nonexporters do not.

*University of Toronto, Institute for Policy Analysis, 140 St. George Street, Suite 707, Toronto, ON M5S 3G6, Canada, johannes.vanbiesebroeck@utoronto.ca. I would like to thank Jim Tybout, Mark Roberts, and Marcel Fafchamps for providing me with the data. The associate editor, an anonymous referee, Mel Fuss, Rob McMillan, Marc Melitz, Amil Petrin, Frank Wolak, and participants at the NBER Summer Institute, the 2003 summer workshop at the CEU, and the 2006 CEPR meeting in Dublin provided very useful comments. All remaining errors are my own. Financial support from the Connaught Fund and SSHRC is gratefully acknowledged.

1 Motivation

Productivity measurement has become ever more widespread since Solow's first decomposition of output growth into the contribution of input growth and a residual productivity term. Productivity is often used as a performance benchmark to rank firms or countries or to measure the rate of performance improvement over time.[1] Such rankings gained credibility once studies documented that productivity is positively correlated with other indicators of success such as profit, employment growth, export status, technology adoption, or mere survival. At the same time, the concept is not without ambiguity. Many different ways to measure productivity exist, each relying on some untestable assumptions. As a consequence, the reader is often left in some doubt as to how sensitive the conclusions of a given study are to the particular productivity measure used.

Given the objective of productivity measurement to identify output differences that cannot be explained by input differences, at least six issues will affect how successful various methodologies are at accomplishing this.[2] First, one has to specify whether all firms share the same production technology and input trade-off or not.[3] Second, most methods require a functional form assumption, or at least some restrictions, on the deterministic portion of the production technology.[4] Third, especially when some heterogeneity in technology is allowed, an assumption on firm behavior is required to learn about the technological differences. Fourth, when technology is assumed homogeneous across firms, it can be estimated econometrically, but one has to control for the well-known problem of endogeneity of input choices. Fifth, to distinguish productivity from other unobservable elements that affect output, one has to place some structure on the stochastic evolution of the unobserved productivity difference. Sixth, methodologies differ in sensitivity to measurement error in output or inputs. Especially in less intensively used data sets from developing countries, this is likely to be important.

This study evaluates five widely-used methodologies, which deal differently with the

[1]Examples of productivity used as a criterion to evaluate policy interventions or firms' decisions can be found in many fields of economics. In industrial economics, a large literature investigates the effect of R&D on productivity and the resulting impact on industry structure – see Griliches (1994) for an overview. In international economics, the impact of trade liberalization is now as likely to be measured by firm level productivity changes as by changes in price-cost margins – see Tybout and Westbrook (1995) for an influential example.

[2]See Coelli, Rao, and Battese (1997) for a more elaborate discussion of the different issues involved.

[3]Most parametric methods assume a homogenous production function; Mairesse and Griliches (1990) and Klette (1999) are notable exceptions.

[4]A number of studies investigate the impact of functional form assumptions, see for example Berndt and Khaled (1979) and Gagné and Ouellette (1998). I sidestep this issue here.

above issues. Index numbers are relatively flexible in the specification of technology, but do not allow for measurement error. Assuming firms minimize costs and factor and goods markets are competitive, they provide an exact measure of productivity without having to estimate the full range of input substitution possibilities. Data envelopment analysis is an entirely nonparametric method. Production plans are explicitly compared to the frontier, which is constructed as a linear combination of 'efficient' production plans in the sample. The method is deterministic, but the benchmark is intuitive from an activities analysis point of view.

The three parametric methods I consider calculate productivity from an estimated production function. Because the framework is explicitly stochastic, they are less vulnerable to measurement error, especially in output, but misspecification of the production function can be an issue. Estimators differ most importantly in the way they control for the simultaneity of productivity and input choices. The system GMM estimator – see Blundell and Bond (1998) – relies on exogeneity assumptions on lagged inputs and output to generate instruments; productivity is assumed to follow an AR(1) process with a firm-specific intercept. Stochastic frontier estimators make distributional assumptions on the unobserved productivity differences to construct a likelihood function for the observed variables; the evolution of productivity has to be modeled explicitly. The semiparametric approach – pioneered by Olley and Pakes (1996) – exploits the information on productivity differences contained in the investment decision; productivity is assumed to follow a Markov process.

To evaluate how sensitive the results are to the five measurement methodologies, I first compare the estimates directly. The parametric methods produce point estimates and standard errors for the production function parameters, which can be compared to the entire distribution of input elasticities (weights) that the nonparametric methods calculate. For the productivity level and growth estimates, I discuss the correlations between the various measures, as well as a number of summary statistics for the productivity distributions.

Secondly, I verify whether the conclusions from important debates in the productivity literature depend on the productivity measure used. I focus on three questions, which have received a great deal of attention in development economics:

- Can the often-observed positive correlation between productivity level and export status be explained entirely by self-selection of more productive firms (plants) into the export market, or is there a role for learning-by-exporting effects?

- Which variables, if any, are consistently associated with productivity growth resulting from knowledge acquisition—as modeled in the endogenous growth literature?

- How important is firm (plant) level productivity change, relative to the reallocation of output between production units, as a source of aggregate productivity growth for the industry?

Each debate concerns a different aspect of the productivity distribution. The first question compares productivity levels across firms, the second compares growth rates, and the third question depends on changes in the entire productivity distribution. Together, they allow a comprehensive overview of the impact of measurement methodology.

Two data sources are used to evaluate the robustness of the different methods: one contains the universe of textile plants in Colombia from 1977 to 1991, the second is a sample of manufacturing firms from Zimbabwe from 1993 to 1995.[5] The two data sets provide distinct case studies. The Zimbabwean firms operate in a low-income country with a relatively small and less developed manufacturing sector. Firms produce a variety of products, possibly using different technologies. The Colombian plants operate in a medium-income country that is approximately three times as large and has a longer history of manufacturing.[6] These plants are all in the textile industry, which is relatively open to international competition and the production technology is expected to be more homogenous.

Both data sets (or similar ones in the respective regions) have been used extensively, not least to study the three debates I focus on. There is little consensus in these literatures. The learning-by-exporting question is studied for Colombia in Isgut (2001) and Clerides, Lach, and Tybout (1998). The former highlights the positive correlation between exporting and productivity, while the latter finds that the correlation can be explained entirely by self-selection. In contrast, Bigsten et al. (2004) and Van Biesebroeck (2005a) find robust learning-by-exporting effects for Zimbabwe and other sub-Saharan African countries. In terms of the second debate, several studies have stressed that technology created in developed countries is likely to be inappropriate for developing countries – see Acemoglu and Ziliboti (2001) and Los and Timmer (2005). Tybout (2000) surveys studies that link openness and the acquisition of foreign knowledge to productivity growth, finding weak but inconclusive evidence for a positive effect. The studies of Tybout and Westbrook (1995) for Mexico and Handoussa, Nishimizu, and Page (1986) for Egypt are illustrative for the two regions I study here. Finally, the debate on the relative importance of firm or plant level change, called the 'within' effect, is far from settled. Some existing studies find a smaller effect for Colombia – see Tybout and Liu (1996) or Petrin and Levinsohn (2006) – than for sub-Saharan Africa –

[5] Van Biesebroeck (200X) compares different methodologies using Monte Carlo simulations.

[6] The fraction of GDP in industry (in 1991) was 34.8% in Colombia and 27.7% in Zimbabwe. For Colombia, 46.7% of non-fuel exports consisted of manufacturers, but only 32.2% for Zimbabwe.

see Van Biesebroeck (2005b) or Shiferaw (200X).[7]

Given that studies differ on many dimensions, it is difficult to evaluate the extent to which opposing conclusions merely reflect different methodological choices or genuine economic differences. As alternative conclusions from these debates can lead to different policy prescriptions, it is important to correctly identify the underlying economic phenomena that determine them.

The contribution of this study is threefold. First, I present estimators from five distinct literatures in a consistent framework. Only the general idea and crucial equations are presented to convey the distinctive features of each. Each method generates comparable productivity estimates, but does so in a radically different measurement framework. I indicate the strengths and weaknesses of each estimator and provide links to the literature for more detailed information.

The second contribution is to compare the estimates directly. While the productivity measures are surprisingly similar across methods—the correlations between the different productivity levels and growth rates are invariably high—the input elasticity estimates differ substantially, especially if returns to scale are left unrestricted. For some applications, the results will depend crucially on the choice of estimation methodology.

Third, the results advance our understanding on the three important productivity debates. To a large extent, the different productivity estimators lead to the same conclusions. Only on the first question are the results for the parametric methods noticeably different from the nonparametric results, suggesting that exporters use a different technology than nonexporters.

The remainder of the paper is organized as follows: In Section 2, I provide a brief background to productivity measurement and introduce the different methodologies. The data are introduced in Section 3, and in Section 4 the productivity estimates are compared directly. In Section 5, I verify whether the answers to the three debates vary by estimation methodology. Finally, Section 6 summarizes what the comparisons teach us about the estimation methodologies and about the economic phenomenons analyzed.

2 Estimating productivity

One firm is more productive than another if it can produce the same output with less of all inputs or can produce more output from the same inputs. Similarly, a firm has

[7] A more complete literature review for the three debates follows in the respective subsections of Section 5.

experienced positive productivity growth if output has increased more than inputs or inputs have decreased more than output. The more interesting case is to compare two production plans where one uses more of a first input and the other more of a second input. This requires the specification of a transformation function that pins down input substitution possibilities. Each productivity measure is only defined with respect to that specific technology. The production function,

$$Q_{it} \quad = \quad A_{it}\ F_{(it)}(X_{it}), \tag{1}$$

is one such representation of technology, linking inputs (X) to output (Q). A_{it} is an unobservable productivity term, which differs between firms and time periods.

Rearranging the production function as

$$\ln \frac{A_{it}}{A_{j\tau}} \quad = \quad \ln \frac{Q_{it}}{Q_{j\tau}} - \ln \frac{F_k(X_{it})}{F_k(X_{j\tau})} \tag{2}$$

underscores that productivity is intrinsically a relative concept. If the technology varies across observations one has to be explicit which technology underlies the comparison, hence the k subscript ($k \in \{ij, j\tau\}$). A multilateral comparison of productivity levels can be achieved by using average productivity across all firms in the denominator. In practice, $\ln A_{it} - \overline{\ln A_t}$ is most often used, taking the average of the logarithm, and for comparability I follow this practice.[8]

The analysis is limited in a number of ways. Only the single output case is considered and all productivity differences are Hicks-neutral. As most studies use value added or sales as output measure and often use a Cobb-Douglas production function which cannot identify factor-bias in technological change, these restrictions are ubiquitous in the literature.[9] Another limitation is the use of a revenue-based output measure (value added) rather than quantity. If there is product differentiation or any other source of market power, the productivity measures cannot be interpreted as pure efficiency differences, as they will include price effects. This limitation is shared by almost all productivity studies, because producer level prices are rarely observed. A rare exception is Foster, Haltiwanger, and Syverson (2005). They show how the widely documented selection on 'productivity' is really selection on 'profitability', as physical productivity measures are found to be negatively correlated with

[8] In a regression framework, this amounts to including industry-year fixed effects in a regression with log productivity as dependent variable.

[9] I construct output-based productivity comparisons, i.e. how much extra output does a firm produce relative to another firm, conditional on input use. Including input-based comparisons would be straightforward.

plant-level prices.[10]

The calculation of the last term in (2)—the ratio of input aggregators—distinguishes the different methods. Three radically different approaches are possible. First, index numbers impose some restrictions on the shape of the production technology and assume optimizing behavior, but obtain productivity measures without estimating any parameters. The first order conditions for input choices imply that the factor price ratio, which is observable, equals the ratio of the marginal productivities of the factors, see Section 2.1. A second, nonparametric approach constructs a piece-wise linear frontier to maximize the productivity estimate (minimize the distance to the frontier) for the unit under consideration. Observation-specific input weights are chosen optimally with as constraint that no other observation can be more than 100% efficient if the same weights are applied to it, see Section 2.2. Finally, if one is willing to make functional form assumptions, it is possible to parametrically estimate the production function. Simultaneity of productivity and input choices is the main econometric issue and I implement three estimators that control for it differently, see Section 2.3.

The methodologies are introduced briefly in the following subsections. Estimators from different literatures are presented in a unified framework. For a more detailed exposition, the reader is referred to Van Biesebroeck (2003) and references to the literature are included.

2.1 Index numbers (IN)

Index numbers provide a theoretically motivated aggregation method for inputs and outputs, while remaining fairly agnostic on the exact shape of the production technology. For example, Caves et al. (1982a) show that the Törnqvist index exactly equals the geometric mean of Malmquist productivity indices using either firm's technology if the production technology is characterized by a translog distance function. The weighting exploits information on the input trade-off contained in the factor prices.

Assuming perfect competition in output and input markets and optimizing behavior by firms, it is possible to calculate the last term in equation (2) from observables, without having to estimate the production function. It even allows for some heterogeneity in technology: only the coefficients on the second order terms have to be equal for the two units compared. While it is not strictly necessary to assume constant returns to scale, one would need outside information on scale economies to implement an adjustment. Estimating scale economies parametrically or information on the cost of capital suffices, but following the usual practice, I limit attention to the constant returns to scale case.

[10] For the index number approaches to be theoretically 'exact', output markets have to be competitive.

I use the same formula for total factor productivity growth as Solow (1957):

$$\ln A_{it}^{IN} - \ln A_{it-1}^{IN} \;=\; \ln \frac{Q_{it}}{Q_{it-1}} - (\frac{s_{it}^L + s_{it-1}^L}{2}) \ln \frac{L_{it}}{L_{it-1}} - (1 - \frac{s_{it}^L + s_{it-1}^L}{2}) \ln \frac{K_{it}}{K_{it-1}}, \qquad (3)$$

where s_{it}^L is the firm-specific fraction of the wage bill in output. For multilateral productivity level comparisons, Caves et al. (1982b) propose an index where each firm is compared to a hypothetical firm—with average log output ($\overline{\ln Q}$), labor share ($\overline{s^L}$), etc. The productivity level of firm i at time t is

$$\ln A_{it}^{IN} - \overline{\ln A}_t^{IN} = (\ln Q_{it} - \overline{\ln Q}_t) - \tilde{s}_{it}(\ln L_{it} - \overline{\ln L}_t) - (1 - \tilde{s}_{it})(\ln K_{it} - \overline{\ln K}_t) \qquad (4)$$

with $\tilde{s}_{it} = \frac{s_{it}^L + \overline{s_t^L}}{2}$. This yields bilateral comparisons that are transitive and still allows for technology that is firm-specific.

The main advantages of the index number approach are the straightforward computations (no estimation is required), the ability to handle multiple outputs and many inputs, and the flexible and heterogeneous production technology it allows. The main disadvantages are the deterministic nature and the necessary assumptions on firm behavior and market structure.[11]

2.2 Data envelopment analysis (DEA)

Data envelopment analysis (DEA) or nonparametric frontier estimation dates back to Farrell (1957). It was operationalized by Charnes et al. (1978) and an overview of the method with applications can be found in Seiford and Thrall (1990). No particular production function is assumed. Instead, productivity is defined as the ratio of a linear combination of outputs over a linear combination of inputs. Observations that are not dominated are labeled 100% efficient. Domination occurs when another firm, or a linear combination of other firms, produces more of all outputs using the same input aggregate, where inputs are aggregated using the same weights.

A linear programming problem is solved separately for each observation. Input and output weights are chosen to maximize efficiency (productivity) for the unit under consideration. In addition to sign restrictions, the efficiency of all other firms cannot exceed 100% when the same weights are applied to them. For unit 1 in the single-output case, the problem boils

[11] Adjustments exist for regulated firms, non-competitive output markets and temporary equilibrium, but they either involve estimating some structural parameters or are more data intensive.

down to

$$
\begin{aligned}
\max_{v_q, v^*, u_l, u_k} \quad & \theta_1 = \frac{v_q Q_1 + v^*}{u_l L_1 + u_k K_1} \\
\text{subject to} \quad & \frac{v_q Q_i + v^*}{u_l L_i + u_k K_i} \leq 1 \qquad\qquad i = 1 \ldots N \\
& v_q, u_l + u_k > 0, \ u_l, u_k \geq 0
\end{aligned}
\tag{5}
$$

Multiple outputs would be aggregated linearly and v^* is a complementary slack variable to allow for variable returns to scale ($v^* = 0$ for constant returns to scale). In practice, most applications solve the dual problem, where θ_1 is chosen directly.

The efficiency measure θ_i is estimated on a sample that includes all firm-years as separate observations and can be interpreted as the productivity difference between unit i and the most productive unit: $\theta_i = \frac{A_i}{A_{max}}$. Estimates of productivity levels and growth rates that are comparable to those obtained with the other methodologies can be defined as:

$$
\ln A_{it}^{DEA} - \overline{\ln A_t}^{DEA} = \ln \theta_{it} - \frac{1}{N_t} \sum_{j=1}^{N_t} \ln \theta_{jt} \tag{6}
$$

$$
\ln A_{it}^{DEA} - \ln A_{it-1}^{DEA} = \ln \theta_{it} - \ln \theta_{it-1}. \tag{7}
$$

These transformations do not change the ranking of firms, only the absolute productivity levels and growth rates.

The main advantage of DEA is the absence of functional form or behavioral assumptions. The underlying technology is entirely unspecified and allowed to vary across firms. The linear aggregation is natural in an activities analysis framework. Each firm is considered a separate process that can be combined with others to replicate the production plan of the unit under investigation. On the other hand, the flexibility in weighting has drawbacks. Each firm with the highest ratio for any output-input combination is 100% efficient. Under variable returns to scale, each firm with the lowest input or highest output level in absolute terms is also fully efficient. The most widely used implementations are not stochastic, making estimates sensitive to outliers. Because each observation is compared to all others, measurement error for a single firm can affect all productivity estimates.

2.3 Parametric estimation

The parametric methods assume the same input trade-off and returns to scale for all firms. Functional form assumptions concentrate all heterogeneity in the productivity term, but the explicitly stochastic framework is likely to make estimates less susceptible to measurement error. I follow most of the literature by estimating a Cobb-Douglas production function in logarithms,

$$q_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + \epsilon_{it}. \tag{8}$$

ω_{it} represents a productivity difference known to the firm, but unobservable to the econometrician; ϵ_{it} captures other sources of i.i.d. error.

Consistent estimation of the input parameters faces an endogeneity problem. Firms choose inputs knowing their own level of productivity and a least squares regression of output on inputs will give inconsistent estimates of the production function parameters. I implement three estimators that explicitly address the endogeneity problem. The two stochastic frontiers in Section 2.3.1 make explicit distributional assumptions on the unobserved productivity; the GMM-SYS estimator in Section 2.3.2 relies on instrumental variables; the semiparametric estimator in Section 2.3.3 inverts the investment function nonparametrically to obtain an observable expression for productivity.

Because the input aggregator is assumed constant across time and firms, productivity level comparisons and growth rates are straightforward:

$$\begin{aligned} \ln A^z_{it} - \overline{\ln A}^z_t &= (q_{it} - \overline{q_t}) - \hat{\alpha}^z_l (l_{it} - \overline{l_t}) - \hat{\alpha}^z_k (k_{it} - \overline{k_t}) && (9) \\ \ln A^z_{it} - \ln A^z_{it-1} &= (q_{it} - q_{it-1}) - \hat{\alpha}^z_l (l_{it} - l_{it-1}) - \hat{\alpha}^z_k (k_{it} - k_{it-1}), && (10) \end{aligned}$$

$z \in \{\text{SF1, GMM, OP}\}$. To obtain a clean estimate of ω_{it} one should subtract an estimate for the difference in error terms from the right-hand side. Generally this is not possible and ignored because $E(\epsilon_{it}) = 0$. For the second stochastic frontier estimator (SF2), a different formula will be used to purge the random noise (ϵ) from the productivity estimates.

2.3.1 Parametric estimation: stochastic frontiers (SF)

The stochastic frontier literature uses assumptions on the distribution of the unobserved productivity component to separate it from the random error. The method is credited to Aigner et al. (1977) and Meeusen and van den Broeck (1977) who model productivity as a stochastic draw from the negative of an exponential or half-normal distribution. Estimation is usually with maximum likelihood. In the production function (8), the term ω_{it} is weakly negative and interpreted as the inefficiency of firm i at time t relative to the best-practice production frontier. An alternative interpretation is that the firm-specific production function lies ω_{it} below best-practice.

Initially developed to measure productivity in a cross-section of firms, the model was generalized for panel data in a number of ways. Battese and Coelli (1992) provide the most straightforward, but also the most restrictive generalization, modeling the inefficiency term as

$$\omega_{it}^{SF1} \quad = \quad -e^{\eta(t-t_0)}\,\omega_i \qquad \text{with} \;\; \omega_i \;\sim\; N^{+}(\gamma,\sigma^2). \tag{11}$$

The relative productivity of each firm (ω_i) is a time-invariant draw from a truncated normal distribution. Inefficiency increases (decreases) deterministically over time if η is positive (negative) at the same rate for all firms.

A more flexible generalization of the cross-sectional stochastic frontier, by Cornwell et al. (1990), is to estimate a time-varying firm-specific effect using three coefficients per firm:

$$\omega_{it}^{SF2} = \alpha_{i0} + \alpha_{i1}t + \alpha_{i2}t^2. \tag{12}$$

Productivity still evolves deterministically, but the growth rate changes over time and varies by firm.

While it is customary to calculate technical inefficiency as $E(e^{\omega_{it}}|\hat{\omega}_{it}+\hat{\epsilon}_{it})$, for comparability with the other methods I use the expected value of log-productivity. For SF1, this boils down to the earlier formulas, (9) and (10).[12] For SF2, productivity level and growth can be calculated as

$$\ln A_{it}^{SF2} - \overline{\ln A}_{t}^{SF2} \quad = \quad (\hat{\alpha}_{i0} - \overline{\hat{\alpha}_0}) + (\hat{\alpha}_{i1} - \overline{\hat{\alpha}_1})t + (\hat{\alpha}_{i2} - \overline{\hat{\alpha}_2})t^2 \tag{13}$$

$$\ln A_{it}^{SF2} - \ln A_{it-1}^{SF2} \quad = \quad (\hat{\alpha}_{i1} - \hat{\alpha}_{i2}) + 2\hat{\alpha}_{i2}t. \tag{14}$$

An advantage of the stochastic frontiers is that the deterministic part of the production function can be generalized easily to allow more sophisticated specifications, e.g. to incorporate factor-bias in technological change. They straightforwardly generalize the popular fixed effects estimator. The two implementations trade off flexibility in the characterization of productivity with estimation precision. SF2 uses $3 \times N$ degrees of freedom and it is the only estimator where consistency relies on asymptotics in the time dimension. One might also be uncomfortable with identification coming solely from distributional assumptions, which are especially restrictive for SF1.

[12] The best estimate of $E(\omega_{it}|\hat{\omega}_{it}+\hat{\epsilon}_{it})$ is $(\hat{\omega}_{it}+\hat{\epsilon}_{it})$ if ω_{it} is independent of ϵ_{it}.

2.3.2 Parametric estimation: instrumental variables (GMM)

The general approach to estimate dynamic error component models of Blundell and Bond (1998) was first applied to production functions in Blundell and Bond (2000). The productivity term is modeled as a firm fixed effect (ω_i) plus an autoregressive component ($\omega'_{it} = \rho\omega'_{it-1} + \eta_{it}$). Quasi-differencing the production function gives the estimating equation in its dynamic representation,

$$q_{it} \;=\; \rho q_{it-1} + \alpha_l(l_{it} - \rho l_{it-1}) + \alpha_k(k_{it} - \rho k_{it-1}) + \alpha'_t + \omega'_i + \underbrace{(\eta_{it} + \epsilon_{it} - \rho\epsilon_{it-1})}_{\varepsilon_{it}}. \quad (15)$$

There is still a need for moment conditions to provide instruments, because the inputs will be correlated with the composite error ε_{it}.

Estimating equation (15) in first-differenced form takes care of the firm fixed effects. Three and more periods lagged inputs and output will be uncorrelated with $\Delta\varepsilon_{it}$ under standard exogeneity assumptions on the initial conditions.[13] Blundell and Bond (1998) illustrate theoretically and with a practical application that these instruments can be weak. If one is willing to make the additional assumption that input changes are uncorrelated with the firm fixed effects, twice lagged first differences of inputs are valid instruments for the production function in levels. The production function in first differences and levels are estimated jointly as a system with the appropriate set of instruments for each equation. Productivity is again calculated using equations (9) and (10).

The GMM-SYS method is flexible in generating instruments and one can test for overidentification. It allows for an autoregressive component to productivity, in addition to a fixed and an idiosyncratic component. Relative to the simple fixed effects estimator, it also uses the information contained in the levels, which is likely to help with measurement error, see Griliches and Mairesse (1998). The main disadvantage is the need for a long panel, at least four time periods are required. Also, if instruments are weak, the method risks underestimating the coefficients.

2.3.3 Semiparametric estimation (OP)

The last method was introduced by Olley and Pakes (1996) to estimate productivity effects of restructuring in the U.S. telecommunications equipment industry. Productivity, a state variable of the firm, is assumed to follow a Markov process unaffected by the control vari-

[13] At least three lags are necessary as $\Delta\varepsilon_{it}$ contains errors as far back as ϵ_{it-2}. Van Biesebroeck (2003) lists the exact form of the necessary assumptions and moment conditions.

ables. Investment, which is shown to be a monotonically increasing function of productivity, becomes part of the capital stock with a one period lag. Inverting the investment equation nonparametrically provides an observable expression for the productivity term that can be used to substitute it from the production function.[14]

In a first estimation step, the variable input coefficients and the joint effect of all state variables are estimated. In a second step, the coefficients on the observable state variables—just capital in our case—are identified, relying on the orthogonality of capital and the innovation in productivity. An intermediate step controls for sample selection, as firms are assumed to exit if productivity falls below a threshold, which is likely to be decreasing in capital. The probability of survival ($\hat{P}$) is predicted from a Probit regression and will enter as a second argument in the nonparametric function $\psi(.)$ in the second step. The estimating equations for the two steps are

$$\begin{aligned} q_{it} &= \alpha_0 + \alpha_l l_{it} + \phi_t(i_{it}, k_{it}) + \epsilon^1_{it} && (16) \\ q_{it} - \hat{\alpha}_l l_{it} &= \alpha_k k_{it} + \psi(\hat{\phi}_{it-1} - \alpha_k k_{it-1}, \hat{P}_{it}) + \epsilon^2_{it}. && (17) \end{aligned}$$

The functions ϕ_t and ψ are approximated nonparametrically by a fourth order polynomial or a kernel density. Productivity is calculated from (9) and (10).

An advantage of the Olley-Pakes approach is the flexible characterization of productivity, only assuming that it evolves according to a Markov process. Potential weaknesses are the nonparametric approximations. The functions that are inverted are complicated mappings from states to actions, which have to hold for all firms regardless of their size or competitive position. Ackerberg et al. (2005) also illustrate that the implicit assumptions required to identify the variable input coefficients are relatively restrictive.

3 Data

I evaluate the different methodologies using two data sets that have been used extensively in the productivity literature. The first is a panel of manufacturing plants from the Colombian Census of Manufacturers. For a detailed description of the data and variable construction, see Roberts (1996). It covers all active establishments between 1977 and 1991, but I limit the sample to plants that at some point are classified in industry ISIC (Rev. 2) 322: 'Clothing and

[14] The methodology is more general than this exposition makes appear. The basic idea is to use another decision by the firm to provide additional information on the unobserved productivity term. Alternatively, Levinsohn and Petrin (2003) invert the material input demand .

Apparel'. Plants in this industry are expected to be relatively homogeneous in technology, at least compared to other industries. The sector also has a relatively large foreign exposure, which makes it an interesting place to evaluate the different debates.[15]

The sample is further limited by only including plants that operate for at least three years, as many estimation methods need at least three observations per plant. This results in an unbalanced panel of 14348 observations from 1957 plants with nonmissing information on output, labor, capital, wages, and investment (investment is often zero). The output concept used is value added, defined as sales minus indirect costs and material input. Labor input is total employment and capital input is the reported book value of the plant and equipment. Value added is deflated with the same sectoral output deflator used in Roberts (1996). For capital, the capital goods deflator from the IMF Financial Tables is used.

The second data set contains a sample of manufacturing firms in Zimbabwe.[16] Data was collected from firm surveys for 1993, 1994, and 1995. Approximately 200 firms were interviewed in three consecutive years. For details on the sampling frame, some background information on the country, and the size distribution of firms, see Van Biesebroeck (2005b). Firms come from four broadly defined manufacturing sectors: food, textile, wood, and metal, corresponding roughly to the ISIC classification 31, 32, 33, and 38. Some firms exit the sample each year and new firms were added in later rounds to maintain the sample size.[17]

As for Colombia, only firms with nonmissing data on output, inputs, wage bill and investment (including zeros) are retained. Output is value added, sales minus indirect costs and material input, inputs are total employment (labor) and the reported replacement value of the plant and equipment (capital). Value added and capital are deflated using the manufacturing deflator from the IMF Financial Tables.[18] Table 1 contains some summary statistics. In all tables, results for Colombia will be on the left and for Zimbabwe on the right. Note that the Colombian data set covers plants and the Zimbabwean data set firms; for ease of exposition I will occasionally use the terms plants and firms interchangeably to denote observations in both data sets.

[Table 1 approximately here]

[15] In the final year of the sample, the textile industry accounts for 10% of manufacturing employment, 3% of value added, and 8% of Colombian manufacturing exports.

[16] The web site of the Centre for the Study of African Economies at http://www.csae.ox.ac.uk/ provides a link to the data and numerous published and working papers that use this and similar data for other sub-Saharan countries.

[17] The short sample period required some modifications in the estimation algorithms. The SF2 estimator only uses a linear time trend, estimating firm-specific intercepts and growth rates, but no quadratic effects; the GMM estimator does not include firm dummies.

[18] Absent more detailed indices, both variables are transformed using the same deflator.

4 Direct Comparison of Methodologies

The following table summarizes the acronyms used when discussing the results and indicates the equations used to calculate productivity levels and growth rates.[19]

method		(level) - (growth)
IN	Törnqvist Index enforcing constant returns to scale	(4) - (3)
DEA	Data Envelopment Analysis: nonparametric frontier	(6) - (7)
SF1	Stochastic frontier with time-invariant productivity ranking	(9) - (10)
SF2	Stochastic frontier with two/three sets of dummies per firm	(13) - (14)
GMM	joint estimation of prod. function in levels and first differences	(9) - (10)
OP	Semiparametric inversion of investment equation	(9) - (10)

4.1 Production function parameters

Table 2 lists the parametric estimates for the production function coefficients with standard errors. For comparison, I also include least squares estimates of the production function (OLS). For the Törnqvist index, which allows for heterogeneity in technology, the average wage bill in value added is reported in the labor column, with its standard deviation across all observations. For the DEA results, I calculate the relative weight of labor and capital in the output aggregate and show the median and standard deviation of the distribution.

[Table 2 approximately here]

One choice to make is whether to enforce constant returns to scale (CRS) or not. The CRS results for Colombia tend to be relatively similar for the different parametric methods: labor coefficient estimates range from 0.75 to 0.80 and capital coefficient estimates

[19] The index numbers calculations can be performed easily with most software packages. DEA estimation was carried out with software developed by the Operations Research and Systems Group at the Warwick Business School (Windows version 1.10). SF1 calculations are performed with the FRONTIER 4.1 program written by Tim Coelli, available online at http://www.uq.edu.au/economics/cepa/frontier.htm. SF2 estimation is by OLS; estimating the large number of coefficients is facilitated by the use of sparse matrix utilities in GAUSS or by an iterative procedure where a subset of the explanatory variables and the dependent variable are first regressed on the remainder of the explanatory variables, and the residuals are subsequently regressed on each other. Estimation of the Arellano-Bond dynamic panel model using lagged levels as instruments for the first-differenced equation is now available for STATA (command `xtabond`). For the system GMM estimator, I used the GAUSS program DPD98 written by Manuel Arellano, available online at http://www.cemfi.es/~arellano/#dpd. Finally, the semiparametric estimation routine (as implemented by Levinsohn and Petrin (2003)) is now also available for STATA (command `levpet`), but I programmed it myself to include the intermediate step, controlling for endogenous exit.

are between 0.20 and 0.29.[20] However, when returns to scale are estimated freely, all four parametric methods find them to be strongly and statistically significantly decreasing in the Colombian textile industry. For Zimbabwe, three parametric estimators find increasing returns to scale, while only the SF2 estimator points to decreasing returns to scale. DEA allows the returns to scale to vary by firm and the range of estimates is large (the standard deviation is 0.33 in Colombia and 0.32 in Zimbabwe). The median estimate is 0.88 for Colombia and 1.04 for Zimbabwe, broadly consistent with the parametric estimates.

In the absence of *a priori* evidence on scale economies, I do not impose them to be constant (except for the index numbers). While the absolute size of scale economies differs, all estimators agree that the average plant in the Colombian textile sector has exhausted all scale economies and might even be operating above efficient scale. On the other hand, all but one estimator suggests that there are moderate scale economies left to be exploited by Zimbabwean manufacturing firms, consistent with the evidence quoted in Tybout (2000).[21]

Accounting for the simultaneity of inputs and productivity lowers the labor coefficient estimate significantly for the parametric methods relative to the OLS estimates. The two nonparametric methods also calculate an average weight for labor much below the OLS estimates. The range of estimates across the different methods is extremely wide in both samples. It ranges from an implausibly low 0.21 (GMM) to 0.74 (OP) in Colombia and from 0.38 (SF2) to 0.75 (OP) in Zimbabwe. While the methods agree that OLS estimates are biased upward, there is no agreement whatsoever on the true labor coefficient. Moreover, the low standard errors convey a misleading sense of accuracy.

The capital coefficient estimate is less affected by the simultaneity correction, but the change relative to the OLS estimate can go in either direction. The SF1 estimator even finds a change in opposite directions in both data sets (relative to OLS). The range for the capital coefficient is even wider than for labor: in Colombia it ranges from 0.09 (SF2) to 0.47 (IN) and in Zimbabwe from 0.10 (SF2) to 0.64 (SF1).

While the range of estimates is large, not all results are equally reliable. One notable pattern is that estimators that include fixed effects, SF2 in both countries and GMM in Colombia, find strongly decreasing returns to scale. SF2 estimates of the coefficient on the capital stock, which tends to be relatively constant over time, are extremely low. The

[20] Full results are reported in the working paper version, Van Biesebroeck (2003). All labor coefficients are estimated below 0.83 and the capital coefficients above 0.17, the respective OLS estimates.

[21] The different unit of analysis and the oversampling of large firms in Zimbabwe, makes it difficult to compare the two samples. It is noteworthy that the median textile firm in Zimbabwe employs only one third as many workers as the median firm in the other industries in the sample, consistent with a much lower minimum efficient scale in textiles.

GMM capital coefficient estimate is higher, as lagged values of inputs are relatively strong instruments for the capital stock. Griliches and Mairesse (1998) argue that the signal-to-noise ratio in the data is much reduced if input coefficients are identified off the changes over time, with measurement error biasing the coefficient estimates downward.[22]

Of the remaining estimates, the average wage bill (IN) is below all parametrically estimated labor coefficients. For developing countries, this is not entirely surprising. To the extent that the production technology and machinery is imported from more developed countries and input substitution is limited, the capital intensity will be higher than optimal, given the low relative factor price for labor. In addition, a substantial fraction of worker compensation might be in the form of non-wage benefits. Hence, the wage bill will underestimate the share of labor in total costs. Both phenomenons are likely to be more pronounced for Zimbabwe than Colombia, which is consistent with the relative size of the average wage shares.

Finally, the SF1 assumption that all firms improve productivity at the same rate is likely to be less appropriate in Zimbabwe than in Colombia. In the first two years of the sample, half of all Colombian plants experienced labor productivity growth between -0.03 and 0.34; the equivalent range for Zimbabwean firms was -0.21 to 0.29, or 33% wider. Comparing the average firm-level growth rates over the entire sample periods in the two countries, the difference is even larger. Half of all Colombian plants averaged labor productivity growth between -0.03 and 0.11; the equivalent range for Zimbabwe is -0.17 to 0.30, more than three times as wide.

Limiting attention to the three most reliable estimates for each country, the results are much more consistent. For Colombia, the labor elasticity is estimated at 0.68 with SF1 and 0.74 with OP; the interquartile range for the relative weight of labor with DEA is 0.54–0.88. The three methods also find returns to scale to be decreasing in the range of 0.85–0.92. For Zimbabwe, the labor coefficient is estimated rather similar: 0.70 by GMM and 0.75 by OP; DEA gives an interquartile range of 0.50–0.84. The difference in capital elasticities between the two countries is larger: the average for Colombia is 0.21, while it is 0.37 for Zimbabwe. As a result, returns to scale are estimated to be increasing in Zimbabwe, with an average point estimate of 1.09, and decreasing in Colombia, on average 0.88.

It is comforting to know that the large differences in Table 2 can be understood and that the range of most trustworthy estimates is relatively narrow. However, these 'ex post'

[22]Van Biesebroeck (200X) uses simulated data to show that measurement error can be a severe problem in these models. An alternative interpretation is that there are important unobserved inputs; increasing only labor and capital inputs will not raise output proportionately.

insights when the entire range of estimates is available are not very useful from a practical perspective. Generally, only a single estimate is available and one has to decide whether to trust it or not. For each method, a number of diagnostics exist to judge how reliable the results are. For example, almost 10% of firms in Zimbabwe report a wage share higher than one, but only 4% in Colombia, providing even more reason not to rely on the IN results in Zimbabwe. For DEA, we can check what fraction of observations are deemed 100% efficient. In both data sets, 20 observations are found to be fully efficient, which represents 5.1% of all firms in Zimbabwe, but only 0.14% of plants in Colombia. On the other hand, only 4.5% of observations in Zimbabwe construct an input aggregate by putting full weight on a single input (almost always labor); this happens for 15.6% of the observations in Colombia (split almost equally between capital and labor).

For the two stochastic frontiers there are no obvious diagnostic checks as they are not identified if the distributional assumptions are violated. We do find, however, that the estimated residuals of both the SF1 and SF2 models seem to follow an AR(1) process, which violates the models' assumptions. For the GMM-SYS estimator, the restrictions on the parameters of the dynamic representation of the production function (e.g. the parameter on lagged capital should equal $-\rho\alpha_k$) are violated in both countries and we do not impose them. In Colombia, the overidentifying restrictions on the instruments are narrowly rejected and there is strong evidence of serial correlations of more than one period. In the short panel available for Zimbabwe, the GMM-SYS estimation reduces to a cross-sectional system and neither of these tests can be performed. For the OP estimation, a crude check is to verify graphically whether the monotonicity assumption between investment and productivity, conditional on capital, holds for the estimated productivity series.[23] This is the case for Colombia, but for Zimbabwe the surface curves up for low levels of investment (especially at high capital levels). A more formal test for the validity of the nonparametric inversions is to verify whether lagged labor input has any predictive power in the second stage regression. For Colombia, the sample had to be split in three periods, based on the business cycle, with separate inversions performed over each period, for the coefficient to become insignificant. In Zimbabwe, the coefficient was significant with a p-value of 0.04.

While these tests do raise a number of flags, the concerns they raise are unlikely to be sufficiently severe to reject any of the methodologies if one had only one set of results available. The large differences in the observable component of the production function has important consequences for any application that uses the production function directly as a representation of technology. For example, the potential effect of a trade policy that elimi-

[23]This can be done straightforwardly in STATA using the `gr39` command.

nates capital controls and attracts more FDI will obviously depend on the capital coefficient estimate. The estimated input coefficients allow for a decomposition of output differences in observable input differences and a residual. Whether the residuals are deemed similar or not crucially depends on the fraction of the total output variation that can be explained by input differences. This is investigated in detail in the next two sections.

4.2 Productivity level estimates

A first way to compare the productivity estimates is by the dispersion they imply. The first two columns in Table 3 contain the interquartile range for each method for Colombia and columns (1b) and (2b) contain the same statistics for Zimbabwe. The median is normalized to zero by year. The widths of the intervals are relatively similar across rows, especially in Colombia, which is remarkable because the methods rely on very different calculations and assumptions. In each country, the most narrow interval is about two thirds as wide as the widest interval. Intervals are almost 50% wider in Zimbabwe than in Colombia and the difference goes the same way for every method. It could be the result of the lower level of development or simply of the much smaller sample size.

[Table 3 approximately here]

The methods that estimate large decreasing returns to scale, SF2 in both countries and GMM in Colombia, also find the widest intervals of all methods. In Colombia, the two nonparametric methods that allow heterogeneity in technology (IN and DEA) also tend to find somewhat wider intervals, but the difference is less pronounced in Zimbabwe. In general, productivity is highly dispersed. Even in Colombia, only half of all plant have a productivity level between 45% below and 42% above the median.

Most methods find a distribution of productivity that is slightly skewed to the left in Colombia and more noticeably right-skewed in Zimbabwe. The right-skewness in Zimbabwe is consistent with the model of competitive selection in Olley and Pakes (1996): firms exit when their productivity drops below a threshold, truncating the distribution from the left.[24] The SF1 methodology that has left-skewness built-in—productivity is the sum of a symmetric normal error and an inefficiency term that follows a normal distribution truncated from the right—is the only method to find left-skewness in Zimbabwe. In Colombia, only the DEA method finds right-skewness, the opposite of all other methods.

[24] Most likely, this tendency is enhanced by the stratified sampling on firm size in Zimbabwe, oversampling large (more productive) firms.

A second way to compare the productivity estimates is to look at the correlations between the different measures directly. For Colombia, in the top left panel of Table 4, the average correlation is 0.79, and even 0.86 limiting the comparison to the parametric methods.[25] Even DEA, which leaves technology entirely unspecified, or GMM, which estimates returns to scale to be very low, produce productivity estimates that are very similar to those of the other methods. Only the correlation between the SF2 and IN productivity estimates is below 0.50. Results for Zimbabwe are broadly similar. The average correlation is lower, at 0.66, but this is largely driven by the more dissimilar results for SF2. Omitting this estimator, which estimates two firm-specific coefficients from only three years of data, raises the average correlation to 0.86.[26]

[Table 4 approximately here]

For comparison, I also added correlations between labor productivity (LP), defined as value added per worker, and the different multifactor productivity measures. Only in a single instance (correlations with the DEA estimates in Colombia) is the correlation with labor productivity lower than with all the other measures. In Colombia, correlations of either SF1 or OP with each alternative productivity measure always exceed correlations with LP, but both of these measures achieve a very high correlation with LP themselves (0.92). In Zimbabwe, correlations with LP are never the lowest and no method achieves consistently higher correlations with the others than LP.

The broad similarity of the productivity estimates across methods, in spite of the large differences in input coefficient estimates, indicates that the variation in the observable part of the production function is swamped by variation in unobservables. Even labor productivity estimates are surprisingly similar. Only the SF2 results—which purge the random errors (ϵ in equation (8)) from the productivity estimates—are noticeably different from the other results in Zimbabwe. If one is only interested in the productivity residuals, not in the input coefficients or scale economies, the choice of methodology turns out to be of secondary importance.

[25] Imposing constant returns to scale, the correlations are even higher, see Van Biesebroeck (2003): the lowest correlation is 0.79. The parametric methods that use equation (9)—GMM, SF1, and OP—become virtually indistinguishable.

[26] Spearman-rank correlations between the different methods are similar, but slightly lower for the stochastic frontier estimates. Calculating the correlations separately by year yields virtually identical results.

4.3 Productivity growth estimates

To compare the productivity growth estimates, Table 3 lists the unweighted and output-weighted averages of productivity growth across all observations in each sample. 1977 to 1991 was clearly a very successful period for Colombian textile plants. The unweighted average (real) growth rate across all methods is 5.9% per year. By any standard, this is extremely rapid multifactor productivity growth. At the same time, the differences between productivity measures are small. The one method that takes out measurement error, SF2, produces the lowest estimate, which is still 4.9%. All other methods produce an estimate in the 5.7–6.5% bracket.

In Zimbabwe, the differences are larger. The average is now -7.7% and the estimates range from -13.6% to +2.4%. The only positive average is for SF2, which estimates a constant (deterministic) growth rate per firm based on at most three years of data. Moreover, it calculates the average growth rate only on the limited sample of firms that are observed in each of the three years and one would expect survivors to be more successful. Without this outlier, the average is -9.7% and all estimates are in the -13.6% to -6.1% range, clearly a dismal period for productivity growth in Zimbabwean manufacturing.

Weighing the growth rates by plants' output level increases the average productivity growth in Colombia for each method. This is as expected: plants with high output at the end period receive a higher weight and they tend to have, *ceteris paribus*, higher productivity growth. For Zimbabwe, the same effect is consistent with a higher average growth rate for the DEA or SF2 results if output weights are used. Results for the other methods go in the opposite direction, i.e. weighting lowers the average. This can be explained by the increasing returns to scale technology that the SF1, GMM, and OP methods estimate: in a declining economy, larger firms are penalized additionally. In both countries, the differences between the methods are exacerbated by weighting and the SF2 method is now even more of an outlier. Nevertheless, the main conclusion is the same using each productivity measure: Colombian textile plants were extremely successful over the sample period and Zimbabwean manufacturing firms extremely unsuccessful.

The correlations between the different productivity growth estimates, in the bottom panels of Table 4, mirror the patterns in the level correlations. Especially correlations between the GMM, SF1, and OP results are extremely high. Enforcing constant returns to scale, see Van Biesebroeck (2003), makes them virtually identical. The much lower estimate for scale economies by GMM in Colombia does not result in less correlated productivity growth estimates. Even the nonparametric DEA and IN results are very similar to the parametric results. Except for the correlations with the SF2 measures, which are the clear outliers in both countries, the lowest correlation statistic for Colombia is 0.84 and for Zimbabwe it is 0.77. The average correlation is 0.77 for Colombia and 0.84 for Zimbabwe and even 0.92 and 0.91 omitting the SF2 results.

Finally, at the bottom of each panel in Table 4, I report the correlation statistics between productivity estimates and labor input. Such measures are sometimes calculated in macroeconomics to study the covariance between technology and inputs, see for example Basu, Fernald, and Kimball (200X). In levels, the correlations indicate that larger plants (high employment) are on average more productive, consistent with the higher wages they generally pay. The correlations are especially large for the three measures that include fixed effects and negative only for SF1 in Zimbabwe, as expected given the estimated returns to scale. More interesting are the correlations in growth rates. Several measures find a significantly negative relationship between labor input growth and concurrent productivity growth, consistent with the evidence in Basu et al. (200X) at the aggregate level for U.S. manufacturing.[27] Moreover, the correlations between lagged productivity growth and labor growth, the bottom row in Table 4, are positive and similar in size for all measures in both countries. They suggest that productivity growth, as a proxy for technology improvements, leads to future expansions, again consistent with Basu et al. (200X).

Even though the relative importance attached to the different inputs varies substantially across methods, the impact on productivity estimates is limited. The differences in input coefficients are swamped by the huge differences in output and input growth rates across firms. Productivity growth rates across the different methods are even more similar than the productivity levels. Especially the similarity between the nonparametric and parametric results is remarkable. The principal reason is that the correlation of the growth rates of capital and labor across firms exceeds the corresponding correlation for the input levels. This is highlighted by the high correlations obtained between most measures and labor productivity growth. The one exception is SF2: these estimates are similar to the others for productivity levels, but not for growth rates.

The direct comparison of productivity measures showed surprisingly similar estimates for the different methods. The second approach to evaluate the importance of measurement methodology is to verify whether the conclusions on the three productivity debates are more sensitive to the choice of productivity estimator.

[27] The point estimates vary across measures, but the positive correlation for DEA in Colombia is insignificant and for Zimbabwe the only two significant correlation coefficients are the negative estimates for SF1 and OP.

5 Three Debates

5.1 Does learning-by-exporting increase productivity?

The first question—whether firms that export are able to increase their productivity level—has been one of the most intensely researched questions in the productivity literature for a decade. While it is well established that exporters have higher productivity than nonexporters, see for example Bernard and Jensen (1995) or Aw, Chung, and Roberts (2000), causality could go either way. A first channel is the self-selection of more productive firms into the export market. Possibly, exporters do not derive any productivity gains from this activity and their productivity advantage could be fully established before they start exporting. Future exporters have been found to differ on many dimensions from nonexporters, even before they start exporting. Self-selection is certain to explain at least part of the observed correlation between export status and productivity level. Lopez (2005) surveys the literature and notes that each microeconomic study finds support for such an effect.

An additional causal effect could go in the other direction if exporters are able to increase their productivity level as they learn from their export activities. Such a learning-by-exporting effect is not mutually exclusive with self-selection, but establishing its existence has important policy implications. Trade liberalization is often promoted as a stimulus to raise productivity levels: the domestic industry will have to face foreign competition at home and, should firms choose to export, abroad. Hard evidence for such an effect was virtually nonexistent until recently. Moreover, the earliest rigorous studies looking for learning-by-exporting effects did not find any. For example, Clerides, Lach, and Tybout (1998) and Bernard and Jensen (1999) find for Colombia, Morocco, and the U.S. that the positive correlation between productivity and export status can be explained entirely by self-selection. Later studies, starting with Kraay (1999) for China, did find learning effects, but they often come with caveats: only in certain industries, only after a longer spell on the export market, or only in the first one or two years.

Lopez (2005) provides an extensive list of additional studies concluding against, e.g. in Spain, Germany, and South Korea, or in favor of the learning-by-exporting hypothesis, e.g. for sub-Saharan Africa, the U.K., or Canada. From this literature it is difficult to gauge to what extent opposing conclusions reflect methodology or genuine economic differences between the countries. I test for a learning-by-exporting effect using each of the productivity measures and two distinct approaches to control for firms self-selecting into the export market.

First, I estimate the simultaneous equation model in the seminal paper by Clerides et al. (1998). A Probit model of a firm's export decision, equation (19), is estimated jointly with equation (18) that represents the evolution of productivity.[28] Two lags of export status and productivity are included in each equation. The lagged productivity terms in equation (18) capture persistence in productivity; in equation (19) they capture self-selection of more productive firms in the export market. Lagged export status in the Probit equation captures export persistence, for example resulting from sunk costs of exporting. The parameters of interest are those on lagged export status in the productivity equation, α_{x1} and α_{x2}, which will be positive if past export experience has a beneficial effect on the current productivity level. Estimation follows Clerides et al. (1998).[29]

$$\ln A_{it} = \sum_{\tau=1}^{2} \left(\alpha_{a\tau} \ln A_{it-\tau} + \alpha_{x\tau} EX_{it-\tau}\right) + \text{controls} + \omega_{1i} + \epsilon_{1it} \tag{18}$$

$$EX_{it} = \begin{cases} 1 & \text{if } \sum_{\tau=1}^{2} \left(\beta_{a\tau} \ln A_{it-\tau} + \beta_{x\tau} EX_{it-\tau}\right) + \text{controls} + \omega_{2i} + \epsilon_{2it} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

The second approach to control for self-selection is with a matching estimator as in Wagner (2002) and De Loecker (2005). These authors find evidence of learning-by-exporting in Slovenia, but not in Germany. A firm is considered 'treated' the first year it exports (if productivity in the following year is observed). Each treated firm is matched with replacement to a control, the nonexporter with the closest propensity score, i.e. its 'nearest neighbor'. The propensity score is calculated as the predicted value from a Probit regression of the treatment dummy on lagged productivity, employment, wages, and the same control dummies as before. The productivity premium for exporters is estimated on the limited sample of treated and control firms by regressing log productivity one year post-treatment on the treatment dummy.[30] In the sample for Zimbabwe, only four treated firms can be identified; in Colombia, there are 119 treated plants.

As a benchmark, columns (1a) and (1b) in Table 5 show the productivity premium for exporters from a simple least squares regression of log productivity on lagged export status

[28] In the original paper, a cost measure is used instead of productivity.

[29] The unobservable in both equations is decomposed into a persistent component (ω) that is integrated out using Gaussian quadrature and a random component (ϵ) that is assumed to be normally distributed. Both components are allowed to be correlated across equations. I follow the original paper to deal with the initial values problem. In these and all other regressions in Table 5, employment and time, location, and industry dummies are included as controls.

[30] In Colombia, current employment and the earlier control dummies are included in the estimation of the productivity premium.

and controls. Estimates for Zimbabwe are all in a narrow range between 0.348 and 0.408 and highly significant. For Colombia, the parametric methods find similar premiums, between 0.268 and 0.399, but the nonparametric methods (IN and DEA) that allow different input elasticities by plant find productivity premiums an order of magnitude smaller and not significantly different from zero. This is consistent with some earlier studies that have shown that exporters are not only larger, but also produce with a larger capital stock per employee. Accounting nonparametrically for the higher capital intensity of exporters explains away most of the estimated productivity advantage. The difference in export participation helps explain the different results in the two countries. In the Colombian sample, fewer than 9% of the plants export and the parametrically estimated input coefficients will be more representative of the production technology of nonexporters. In the Zimbabwean sample, 54% of firms export and the production function estimates will be more appropriate for exporters. The IN results produce the lowest estimate for the productivity premium of exporters in Zimbabwe as well, although the DEA estimate is at the other end of the spectrum.

[Table 5 approximately here]

Controlling for self-selection of more productive firms into the export market is expected to diminish the impact of export status on productivity. Results for the simultaneous equations approach—the coefficient on once lagged export status in equation (18) is reported—are in columns (2a) and (2b) of Table 5. For Colombia, the point estimates for the parametric methods fall on average to one third of the OLS estimates, but they remain significantly positive. For the two nonparametric methods, estimates are similar to the OLS results and still insignificant. Under the maintained hypothesis that plants share the same production technology, one would conclude that learning-by-exporting effects are indeed present and relatively large. The productivity premium is estimated at 11% on average. However, the nonparametric results suggest that this conclusion is misleading, as they find a productivity premium for exporters of only 4% on average and the difference with nonexporters is not statistically significant.[31] For Zimbabwe, the range of estimates widens substantially. The point estimates vary from a productivity decline of -0.057 (SF1) to an increase of 0.151 (DEA), but no estimate is significantly different from zero anymore. In contrast to the Colombian results, the nonparametric estimates are at the high end of the range. The large reduction in the point estimates in both countries, relative to the OLS estimates in columns (1a) and (1b), points to important self-selection effects.

[31] Clerides et al. (1998), one of the most prominent studies to find against the learning-by-exporting hypothesis, used an estimate of average variable cost, purged from capital-intensity effects in a flexible way, as dependent variable.

Estimates of the productivity premium using the matching estimator to control for self-selection are in columns (3a) and (3b). Results are very similar to the simultaneous equations results. For Colombia, estimates are larger for the parametric than for the nonparametric methods, but only two methods still find significant learning effects. The point estimates tend to be higher with the matching estimator, but the standard errors increase even more. For Zimbabwe, only four firms with data available in the year post export can be identified as new exporters. Five of the six point estimates are positive, but the range is extremely wide and the large standard errors do not hide the imprecision.

In sum, all methods find that exporters are more productive, but controlling for self-selection reduces the difference in most cases, especially for the parametric methods, and widens the range of point estimates across methods. The results for Zimbabwe all become insignificant, although some point estimates remain large. In Colombia, there is a significant learning effect if we assume that technology is homogeneous across plants, but the size of the premium is estimated much lower using the nonparametric methods. This is one instance where an important assumption of the productivity measurement methodology crucially affects the results. A formal test for a structural break in the production function parameters between exporters and nonexporters strongly rejects that both groups operate with the same technology.[32] The nonparametric methods estimate the position of a plant relative to other plants in the industry to be similar pre and post exporting, but they use a different input substitution frontier in both instances.[33]

5.2 What brings about technological change?

The previous debate centered around differences in productivity levels. Now we turn to the question of what explains differences in productivity growth across firms. In the neoclassical growth model, long term per capita growth can only come from technological change, which is generally left unspecified, exogenous to the model. In the endogenous growth literature, a nonrival input—knowledge—enters the production function and investments by profit maximizing firms in knowledge can lead to steady state (per capita) growth. Tests of this model using country or industry level data have generally not been supportive. Using time series variation, the model predicts long-lasting effects of the level of variables that proxy for

[32]See Van Biesebroeck (2005a) for results in sub-Saharan Africa. Similar results for Colombia are available upon request.

[33]In addition, Van Biesebroeck (2005a) finds that approximately half of the learning-by-exporting effect in nine sub-Saharan African countries can be explained by the realization of scale economies. The large differences in estimated returns to scale for the different methods lead to very different conclusions on the importance of scale effects as well; see Van Biesebroeck (2003) for these results.

knowledge on growth rates. In reality, higher levels of R&D or human capital (engineers and scientists in the workforce) have not lead to higher growth rates, see for example Jones (1995).[34] Cross-sectional studies, exploiting variation across countries, have by and large also been unsupportive, see the survey by Durlauf and Quah (1999). These studies have been criticized because countries at different stage of development are lumped together and preference and technology parameters are assumed to be constant across countries and over time.

Much of the empirical growth literature focuses on variables that are constant within a country, such as institutional quality, the legal system, educational attainment, inequality, etc. The goal here is to identify variables that explain variation in productivity growth across firms within a single country. Ehrlich et al. (1994) introduced a useful extension of the endogenous growth model to explain such differences. Their model has been used extensively to study the effect of ownership or evaluate the impact of privatization on firm performance. The Colombian data does not contain information on ownership, but I can investigate the effect of other variables that plausibly shift a firm-specific asset (knowledge) that serves as an engine of productivity growth. Five proxies are considered: exporting output, importing inputs, acquiring external technology, frequent capital investments, and high levels of human capital.

Importing knowledge from more advanced economies is a first channel with the potential to raise productivity growth. The first two predictors are dummy variables taking a value of one if a firm exports some of its final output or imports intermediate inputs.[35] Exporters compete with foreign firms, have to satisfy demanding foreign clients, and are exposed to advanced technologies. Imported inputs can embody foreign knowledge directly. There is a large literature on international technology spillovers. A recent paper, Yasar and Morrison Paul (2005), looks at the effects of three such channels—exporting of output, importing of machinery, and foreign ownership—and finds each to be associated with higher productivity levels. A mismatch between knowledge created in developed countries and the production structure in developing countries, the 'inappropriate technology' phenomenon, can reduce the effects of these channels, see Los and Timmer (2005) for an illustration.

The data set in each country contains an indicator of direct technology acquisition. Colombian plants indicate whether they paid any royalties (only 2% of plants) and Zimbab-

[34] Gong, Greiner, and Semmler (2004) incorporate decreasing opportunities for technological innovations, which reduces the return to knowledge over time. They do find support for the model using time series data for the U.S. and Germany.

[35] In Colombia, exporting or importing are much less common (respectively 9% and 6% of plants) than in Zimbabwe (54% and 64% of firms).

wean firms indicate whether they invested in any advanced technologies in the preceding year (68% of firms). Given that domestic R&D in these countries is limited, it is likely that to a large extent these variables pick up further effects of imported technology. The endogenous growth model in Diao et al. (1999) explicitly models how effects of R&D can be heightened by international trade. Acemoglu and Ziliboti (2001) argue that many technologies used in developing countries are developed in OECD countries and are inappropriate for the local mix of skills. Mere access to new technologies might not suffice to improve productivity in developing countries.

A high level or superior quality of human and physical capital is the final channel. While capital accumulation alone cannot raise long term economic growth, it can increase the growth rate through embodied technical change, see De Long and Summers (1991) for an influential study at the aggregate level. Investments in the quality of capital could also reflect a high level of knowledge within the firm. For physical capital, I use investment frequency: the fraction of years a firm's capital investments exceed 5% of the capital stock. Such an effect would be consistent with the endogenous growth model of Hsieh (2001), where obsolescence of capital equipment requires continuous investments to maintain growth rates. For human capital, I use the fraction of employees that are classified as highly-skilled or technical workers, averaged over all active years.

The statistics in Table 6 are the coefficient estimates on each of the five predictors in separate regressions with average productivity growth over the entire period as dependent variable. I first discuss the Colombian results, which tend to be very consistent across the different productivity measures. The evidence on exporting and frequent investments mirrors the results on the previous debate. Both are associated with significantly higher productivity growth, but only under the assumption that all plants operate with the same technology, i.e. using a parametric productivity measure. For the two nonparametric measures, I still find a positive but insignificant effect. The point estimates for the parametric measures are large, relative to the average unweighted growth rate of 5.7%. Exporters increase their productivity at a 2.9% higher rate than nonexporters and the growth premium is 4.1% for plants that invest every year relative to plants that never invest.[36]

[Table 6 approximately here]

Importing inputs or paying royalties are not associate with a significant growth effect. With only a single exception, all point estimates are negative, but not significantly different

[36] Results in the working paper version, see Van Biesebroeck (2003), further indicate that the positive growth effects associated with investments in fixed capital are more pronounced for frequent investments than for large investments on average or for large investment spikes.

from zero. Moreover, the size of the effects are comparable for the different productivity measures, on average -0.8% for plants that import inputs and -2.8% for plants that pay royalties. This could reflect that for the sector studied in Colombia, textiles, there is little scope for technological advances to be embedded in imported inputs or for licensing more advanced production technologies.

Finally, high levels of human capital are consistently associated with lower productivity growth using each of the six productivity measures and the effect is estimated significantly different from zero in four cases. A one standard deviation increase in the human capital measure, e.g. from the average of 0.35 to 0.82, is estimated to lower productivity growth by almost 7%. Especially for the nonparametric methods, the results point towards much lower productivity growth at plants that employ skilled workers.[37]

For Zimbabwe, the data is too noisy or the sample size too small to find many effects that are precisely estimated. The only indicator that is consistently associated with large productivity growth effects is the dummy for investments in advanced technology. The point estimates are extremely large, on average 26%, similar in size for all methods, and significantly different from zero for four of the six measures. Firms that do not invest in advanced technology, approximately one third of the sample, are clearly and quickly falling behind in productivity. As in Colombia, the evidence also hints at positive effects of exporting and frequent investments on productivity growth and is suggestive of negative effects of high levels of human capital. In contrast with Colombia, importing inputs tends to be associated with somewhat higher productivity growth. The effects of exporting and importing are estimated to be of similar magnitude, except for the SF2 results. Overall, the point estimates are of very similar magnitude for all productivity measures.

While the economic effects are interesting in their own right, the uniformity of the effects across the different productivity measurement methodologies is striking. This is in line with the direct comparison of productivity growth estimates earlier. For Colombia, the sign of only 3 of the 30 coefficient estimates is the opposite of the majority finding. In each case, the anomalous result is for the index numbers and it is never significant. For Zimbabwe, even with the lower precision of the estimates, only 4 of the 30 signs indicate disagreement between measures. The conclusion on the importance of endogenous growth explanations for productivity growth are very much independent of the choice of estimation method for productivity: export status, frequent investments, and technology adoption are to varying degrees associated with higher productivity growth, while firms that employ many skilled

[37] On the other hand, plants with a large fraction of managers had on average a 2% higher productivity growth with the parametric productivity measures, see Van Biesebroeck (2003).

workers improve productivity more slowly.

5.3 What drives aggregate productivity growth?

The third debate is affected by both productivity level and growth estimates and the distribution of these among firms. The aggregate productivity level, of the entire economy or a single industry, can increase for two reasons: individual firms can improve in productivity, a within-firm effect, or the relative weight of firms with above average productivity level can increases, a between-firms effect. Reallocation of inputs or outputs can take place at the intensive margin, between firms already active, or at the extensive margin, as more productive entrants gain market share or less productive firms exit from the industry. I investigate the relative importance of the two effects using three different decompositions.[38]

Baily, Hulten, and Campbell (1992) (BHC), decomposed aggregate productivity growth as follows:

$$
\begin{aligned}
\Delta \ln A_t &= \sum_{i=s,n}^{\text{stay, enter}} \theta_{it} \ln A_{it} - \sum_{j=s,x}^{\text{stay, exit}} \theta_{jt-\tau} \ln A_{jt-\tau} = \ln\Big(\prod_i A_{it}^{\theta_{it}} \Big/ \prod_j A_{jt-\tau}^{\theta_{jt-\tau}}\Big) \\
&= \underbrace{\sum_{s}^{\text{stay}} \theta_{st-\tau}\Delta \ln A_{st}} + \underbrace{\sum_{s}^{\text{stay}} \Delta\theta_{st} \ln A_{st} + \sum_{n}^{\text{enter}} \theta_{nt} \ln A_{nt} - \sum_{x}^{\text{exit}} \theta_{xt-\tau} \ln A_{xt-\tau}} \qquad (20)
\end{aligned}
$$

Aggregate productivity is defined as the output-weighted (θ_{it}) average of log productivity of individual firms ($\ln A_{it}$).[39] The linear aggregation of log productivity implies a geometric average of productivity levels; it permits a linear decomposition into terms with an intuitive interpretation. Firms that stay in the sample from $t-\tau$ to t are indexed by s. Their contribution is split in two parts. The first term of equation (20) measures the effect of firm level productivity changes, weighted by their initial share. The second term captures the reallocation effect at the intensive margin; it sums changes in shares using productivity as weight. The last two terms capture reallocation at the extensive margin, the net contribution of firms entering or exiting the industry. Recently, Petrin and Levinsohn (2006) criticized this definition of aggregate productivity growth, in particularly because the reallocation term is large and volatile. I limit attention to the within term and investigate whether different productivity measures find it to be of similar magnitude.

[38] Eslava et al. (2004) investigate for Colombia whether and how reallocation can have an effect. I only look at the size of the within-plant effect.

[39] In these decompositions, $\ln A_{it}$ is the logarithm of productivity calculated using the different methods, omitting the normalization by the average log productivity level.

Haltiwanger (1997) introduced an improved decomposition for unbalanced panels where all firm level productivity terms are expressed as differences from aggregate productivity in $t-\tau$. In addition, he decomposed the second term into a 'pure' between effect, weighing the change in shares by the relative productivity in the initial period and a covariance term.[40] The BHC decomposition amounts to lumping the entire covariance term with the between term. An alternative, by Griliches and Regev (1995) (GR), is to modify equation (20) using $\overline{\theta_s} = (\theta_{st} + \theta_{st-\tau})/2$ as weight in the first term and similarly replace $\ln A_{st}$ in the second term by $(\ln A_{st} + \ln A_{st-\tau})/2$. This is equivalent to splitting the covariance term equally between the within and between terms.[41]

An entirely different decomposition was introduced by Olley and Pakes (1996) (OP). Aggregate productivity is defined as the average of the productivity levels, as opposed to the logarithms, and decomposed into two terms as follows:

$$A'_t = \sum_i^{N_t} \theta_{it}\, A_{it} = \underbrace{\frac{1}{N_t}\sum_i^{N_t} A_{it}}_{\overline{A'}_t} + \sum_i^{N_t}\Big(\theta_{it} - \frac{1}{N_t}\Big)(A_{it} - \overline{A'}_t) = \overline{A'}_t + \sum_i^{N_t} \Delta\theta'_{it}\Delta A'_{it} \quad (21)$$

The first term is the unweighted average productivity and the second term captures to what extent firms of above average size ($\Delta\theta'_{it} > 0$) have above average productivity ($\Delta A'_{it} > 0$). An alternative measure of the relative importance of the within-firm effect in aggregate productivity growth is obtained by comparing cumulative growth in the unweighted average ($\overline{A'}_t$) to the growth in the aggregate (A'_t).

For the U.S. manufacturing sector, Baily et al. (1992) find that in the periods 1972–77 and 1982–87, respectively 70% and 87% of aggregate productivity growth comes from the reallocation of output shares, with only a minor role for within-plant changes. In the 1977-82 period, the aggregate growth and within-plant effect are even of opposite signs. Haltiwanger (1997) finds that over the 1977-87 period 54% of the aggregate growth comes from plant level growth. Foster, Haltiwanger, and Krizan (2001) show that these findings are sensitive to the choice of sector, time period, or to the use of labor instead of multifactor productivity. Results with the GR decomposition tend to be more stable. For the telecommunications sector, Olley and Pakes (1996) find with their decomposition that the unweighted average productivity level declines almost continuously from 1974 to 1987, while the weighted average is relatively constant. As a result, the second term in their decomposition accounts for only 10% of the total initially, but for almost one third at the end of the period. Bartelsman

[40]The second term in equation (20) is decomposed as $\sum_s \Delta\theta_{st}\,(\ln A_{st-\tau} - \ln A_{t-\tau}) + \sum_s \Delta\theta_{st}\,\Delta \ln A_{st}$.
[41]See Foster, Haltiwanger, and Krizan (2001) for a more elaborate discussion.

and Dhrymes (1998) argue that it is more intuitive to use an input aggregate as weight. Their graphical decomposition is similar in spirit to OP and they find that for the U.S. manufacturing sector the simple average of plant level productivity growth is relatively constant, while the weighted average increases substantially. They also conclude that reallocation effects dominate. Results for Colombia in Tybout and Liu (1996) suggest that within-plant changes can account for virtually the entire aggregate productivity growth.[42] For a number of sub-Saharan African countries, including Zimbabwe, Van Biesebroeck (2005b) also finds that within-firm effects dominate. On the other hand, Pavcnik (2002) finds for the manufacturing sector in Chile that only one third of aggregate productivity growth is accounted for by the unweighted plant level growth.

Clearly, there is a wide range of estimates in the literature. The results in Table 7 indicate whether the way productivity is measured matters for the conclusion and whether the experience for Colombia and Zimbabwe was similar. Column (1a) contains the cumulative change in aggregate productivity over the 1981–1991 period for Colombia and column (1b) contains the cumulative change for Zimbabwe from 1993 to 1995. The aggregate is calculated according to equation (21).[43] Average growth in Colombia across the different measures was 7.2% per year or 101% cumulatively over ten years. In Zimbabwe, average growth was negative at -13.6% per year.

Columns (2a) and (2b) contain the growth in the unweighted average productivity level over the same time periods, expressed as a fraction of the aggregate growth rate. With as sole exception the SF2 measure in Zimbabwe, the growth rate of the unweighted average tracks the aggregate growth rate very closely. In Colombia, it accounts on average for 97% of aggregate growth with little variation across methods. The decline in unweighted average productivity in Zimbabwe averages 72% of the decline in the weighted average. Outliers are SF1, which finds exactly the same trend in the weighted and unweighted average, and SF2, which finds a much smaller decline in the aggregate and an even smaller decline in the unweighted average. Reallocation of output weights appears to have been fairly unimportant. The difference in the relative importance of the within-firm effect across the different productivity measures is a lot smaller than the difference in the actual growth estimates.

[Table 7 approximately here]

[42] The comparison with the U.S. results should be done cautiously as Tybout and Liu (1996) calculate year-by-year changes, while the between-plants effect, especially at the extensive margin, generally increases in importance over longer time horizons.

[43] For the index numbers, the productivity level is calculated differently than before, using the average wage share in the sample for all firms in order to calculate an aggregate growth estimate that is unit invariant. Chain-linking the input shares, as in Aw et al. (2000), is an alternative solution.

The within-plant terms in the BHC decomposition are reported in columns (3a) and (3b), and for the GR decomposition in columns (4a) and (4b). Especially for Colombia and for the GR decomposition, the size of the effects is extremely similar for all productivity measures. As in the first column, the estimates of aggregate growth differ more (not reported), which introduces some variation in the fraction of aggregate growth explained by within-plant changes: for BHC it ranges from 11% to 25% and for GR from 21% to 33%. Irrespective of the productivity measure used, the reallocation of output between plants is found to be a much more important driver of aggregate change over a 10 year period if this alternative definition of aggregate productivity is used.

For Zimbabwe, differences between the productivity measures are slightly more pronounced, but only the SF2 results are (again) a clear outlier. The within-firm terms in the linear decomposition of the geometric average, columns (3b) and (4b), provide consistent evidence for the importance of firm level productivity changes. The within term accounts on average for 95% of aggregate growth using the BHC decomposition and for 85% of aggregate growth in the GR case. The much shorter time period, only three years, makes it not surprising that the importance of within-firm changes is much larger than in Colombia. For Zimbabwe, the importance of the within term is not only consistent across the different productivity measures, but also across the three decomposition methods.

The different productivity measures come up with very uniform within-firm effects, at least comparing within each column of Table 7 across the rows. One would reach almost identical conclusions with each method to estimate productivity in Colombia and very similar conclusions in Zimbabwe. Only the SF2 estimator in the case of Zimbabwe, where two firm-specific coefficients are estimated using three years of data and where the sample is limited to survivors, produces results that deviate from the other methods. In Colombia, the conclusion on the importance of the within-plant effect does depend on the aggregation and decomposition method used. Over the ten year interval, its importance is estimated to be much larger for the arithmetic average (OP) than for the geometric averages (BHC or GR), but the results are extremely similar for the different productivity measures.[44]

[44] The use of an aggregate input weight instead of output, as proposed in Bartelsman and Dhrymes (1998) and Tybout and Liu (1996), leads to results that vary somewhat more across methods, see Van Biesebroeck (2003). In this case the weights depend directly on the input coefficient estimates, which vary more across methods than the productivity estimates, as discussed earlier.

6 Lessons

In response to the question "Does it matter which method one uses to estimate productivity?", the answer crucially depends on what one is interested in. If the main interest is in the residual—the nondeterministic part of the production function—the choice of method is of lesser importance. The fraction of output differences that cannot be explained by input differences is similar across methods and even more so for differences in output growth rates. The correlations, interquartile ranges, and averages of the productivity measures are very similar for the different methods. Even the deterministic DEA and index numbers generate results that are surprisingly similar to the parametric productivity estimates. Only the method that explicitly takes out random measurement error (SF2) produces noticeably different estimates, especially for productivity growth.[45]

When revisiting the three productivity debates, only in one instance do the conclusions depend on the productivity estimator used.

- While exporters have higher productivity levels using each method, there is evidence of learning-by-exporting only if one assumes the same production technology for all firms, i.e. using a parametrically estimated production function.
- Firm level productivity growth is robustly associated with frequent investments in physical capital, somewhat with export activity, and in Zimbabwe with the adoption of new technologies.
- In the two developing economies studied, firm level productivity growth tracks aggregate growth closely, especially if the aggregate is constructed by averaging the level of productivity (as opposed to the logarithm).

Especially for the latter two debates, the choice of estimation method for productivity is immaterial to the conclusions reached. Other methodology choices, e.g. how to control for endogeneity, how to calculate the aggregate, or how long of a time period to study, tend to be at least as important as the choice of productivity estimation method. At the same time, the results do indicate that differences between countries, even when exactly the same method is used, can be quite large.

It should be stressed that if one is interested in the observable part of the production process, the estimation method matters a great deal. The range of estimates for the capital

[45] Van Biesebroeck (2003) finds that these differences are much reduced if constant returns to scale is enforced, but the input coefficient estimates rarely support the constant returns to scale assumption.

and labor coefficients is wide and even the relative importance of the two inputs varies substantially by method. The coefficient estimates can be used directly to assess the importance of each production factor or returns to scale. The evaluation of some policy changes will also depend on these estimates. For example, the fraction of the productivity premium for exporters that can be explained by the realization of scale economies depends crucially on the estimation method. Also, if an input aggregate is used as weight, instead of output, to aggregate productivity growth, the choice of method again matters.

References

Acemoglu, D. and F. Ziliboti (2001, May). Productivity Differences. *Quarterly Journal of Economics 116*(2), 563–606.

Ackerberg, D. A., K. Caves, and G. Frazer (2005, April). Structural Identification of Production Functions. *Working paper, UCLA*.

Aigner, D., C. K. Lovell, and P. Schmidt (1977). Formulation and Estimation of Stochastic Frontier Production function Models. *Journal of Econometrics 6*, 21–37.

Aw, B. Y., S. Chung, and M. J. Roberts (2000, January). Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China). *World Bank Economic Review 14*(1), 65–90.

Baily, M. N., C. Hulten, and D. Campbell (1992). Productivity Dynamics in Manufacturing Plants. *Brookings Papers: Microeconomics 4*(1), 187–267.

Bartelsman, E. J. and P. J. Dhrymes (1998, January). Productivity Dynamics: U.S. Manufacturing Plants, 1972-1986. *Journal of Productivity Analysis 9*(1), 5–34.

Basu, S., J. G. Fernald, and M. S. Kimball (200X, forthcoming). Are Technology Improvements Contractionary? *American Economic Review*.

Battese, G. E. and T. J. Coelli (1992). Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis 3*, 153–69.

Bernard, A. B. and J. B. Jensen (1995). Exporters, Jobs, and Wages in U.S. Manufacturing, 1976-1987. *Brookings Papers on Economic Activity*, 67–119. Washington DC.

Bernard, A. B. and J. B. Jensen (1999). Exceptional Exporter Performance: Cause, Effect, or Both? *Journal of International Economics 47*(1), 1–26.

Berndt, E. R. and M. S. Khaled (1979). Parametric Productivity Measurement and Choice among Flexible Functional Forms. *Journal of Political Economy 87*(6), 1220–45.

Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Söderbom, F. Teal, and A. Zeufack (2004, February). Do African Manufacturing Firms Learn from Exporting. *Journal of Development Studies 40*(3), 115–41.

Blundell, R. W. and S. R. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics 87*, 115–43.

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

Caves, D. W., L. R. Christensen, and E. W. Diewert (1982a, November). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica 50*(6), 1393–1414.

Caves, D. W., L. R. Christensen, and E. W. Diewert (1982b, March). Multilateral Comparisons of Output, Input, and Productivity using Superlative Index Numbers. *Economic Journal 92*(365), 73–86.

Charnes, A., W. W. Cooper, and E. Rhodes (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research 2*, 429–444.

Clerides, S. K., S. Lach, and J. R. Tybout (1998). Is Learning By Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco. *Quarterly Journal of Economics 113*(3), 903–947.

Coelli, T., D. P. Rao, and G. E. Battese (1997). *An Introduction to Efficiency and Productivity Analysis*. Norwell, MA: Kluwer.

Cornwell, C., P. Schmidt, and R. C. Sickles (1990). Productivity Frontiers with Cross-sectional and Time Series Variation in Efficiency Levels. *Journal of Econometrics 46*, 185–200.

De Loecker, J. (2005, July). Do Exports Generate Higher Productivity? Evidence from Slovenia. working paper, K.U. Leuven.

De Long, B. and L. Summers (1991). Equipment Investment and Economic Growth. *Quarterly Journal of Economics 106*, 445–502.

Diao, X., T. Roe, and E. Yeldan (1999, December). Strategic Policies and Growth: An Applied Model of R&D-driven Endogenous Growth. *Journal of Development Economics 60*(2), 343–80.

Durlauf, S. N. and D. T. Quah (1999). The New Empirics of Economic Growth. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics, Volume 1A*, Chapter 4, pp. 235–303. New York: North-Holland.

Ehrlich, I., G. Gallais-Hamonno, Z. Liu, and R. Lutter (1994). Productivity Growth and Firm Ownership: An Analytical and Empirical Investigation. *Journal of Political Economy 102*(5), 1006–38.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2004). The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia. *Journal of Development Economics 75*, 333–71.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A. 120*, 253–290.

Foster, L., J. Haltiwanger, and C. Krizan (2001). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. In C. R. Hulten, E. R. Dean, and M. J. Harper (Eds.), *New Developments in Productivity Analysis*, Volume 63 of *NBER Studies in Income and Wealth*, pp. 303–63. Chicago, IL: University of Chicago Press.

Foster, L., J. Haltiwanger, and C. Syverson (2005, August). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? NBER Working Paper No. 11555.

Gagné, R. and P. Ouellette (1998, January). On the Choice of Functional Forms: Summary of a Monte Carlo Experiment. *Journal of Business and Economic Statistics 16*(1), 118–124.

Gong, G., A. Greiner, and W. Semmler (2004). Endogenous Growth: Estimating the Romer Model for the U.S. and Germany. *Oxford Bulletin of Economics and Statistics 66*(2), 147–64.

Griliches, Z. (1994, March). Productivity, R&D, and the Data Constraint. *American Economic Review 84*(1), 1–24.

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

Griliches, Z. and H. Regev (1995, January). Firm Productivity in Israeli Industry: 1979-1988. *Journal of Econometrics 65*(1), 175–203.

Haltiwanger, J. C. (1997, May–June). Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence. *Federal Reserve Bank St. Louis Review 79*(3), 55–77.

Handoussa, H., M. Nishimizu, and J. Page (1986). Productivity Change in Egyprian Public Sector Industries after the 'Opening,' 1973–1979. *Journal of Development Economics 20*(1), 53–73.

Hsieh, C.-T. (2001, October). Endogenous Growth and Obsolescence. *Journal of Development Economics 66*(1), 153–71.

Isgut, A. E. (2001, June). What's Different about Exporters? Evidence from Colombian Manufacturing. *Journal of Development Studies 37*(5), 57–82.

Jones, C. I. (1995, May). Time Series Tests of Endogenous Growth Models. *Quarterly Journal of Economics 110*, 495–525.

Klette, T. J. (1999, December). Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data. *Journal of Industrial Economics 47*(4), 451–76.

Kraay, A. (1999). Exportations et Performances Economiques: Etude d'un Panel d'Entreprises Chinoises. *Revue d'Economie Du Developpement 0*(1-2), 183–207.

Levinsohn, J. and A. Petrin (2003, April). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies 70*(2), 317–342.

Lopez, R. A. (2005). Trade and Growth: Reconciling the Macroeconomic and Microeconomic Evidence. *Journal of Economic Surveys 19*(4), 623–48.

Los, B. and M. P. Timmer (2005). The 'Appropriate Technology' Explanation of Productivity Growth Differentials: An Empirical Approach. *Journal of Development Economics 77*, 517–31.

Mairesse, J. and Z. Griliches (1990). Heterogeneity in Panel Data: Are There Stable Production Functions. In P. C. et al. (Ed.), *Essays in Honor of Edmond Malinvaud*, Volume 3, pp. 125–147. Cambridge: MIT Press.

Meeusen, W. and J. van den Broeck (1977). Efficiency Estimation from Cobb-Douglas Production functions with Composed Error. *International Economic Review 18*(2), 435–444.

Olley, G. S. and A. Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica 64*(6), 1263–97.

Pavcnik, N. (2002, January). Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants. *Review of Economic Studies 69*(1), 245–76.

Petrin, A. and J. Levinsohn (2006, April). Measuring Aggregate Productivity Growth Using Plant-level Data. mimeo, Chicago GSB.

Roberts, M. (1996). Colombia, 1977-85: Producer Turnover, Margins, and Trade Exposure. In M. Roberts and J. R. Tybout (Eds.), *Industrial Evolution in Developing Countries*, Chapter 10, pp. 227–59. New York: Oxford University Press for the World Bank.

Seiford, L. L. and R. M. Thrall (1990). Recent Developments in DEA. The Mathematical Programming Approach to Frontier Analysis. *Journal of Econometrics 46*, 7–38.

Shiferaw, A. (200X, forthcoming). Firm Heterogeneity and Market Selection in Sub-Saharan Africa: Does It Spur Industrial Progress. *Economic Development and Cultural Change*.

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *Review of Economics and Statistics 39*, 312–320.

Tybout, J. R. (2000, March). Manufacturing Firms in Developing Countries: How Well Do They Do, and Why? *Journal of Economic Literature 28*(1), 11–44.

Tybout, J. R. and L. Liu (1996). Productivity Growth in Chile and Colombia: The Role of Entry, Exit, and Learning. In M. Roberts and J. R. Tybout (Eds.), *Industrial Evolution in Developing Countries*, Chapter 4, pp. 73–103. New York: Oxford University Press for the World Bank.

Tybout, J. R. and M. D. Westbrook (1995). Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries. *Journal of International Economics 39*, 53–78.

Van Biesebroeck, J. (2003, November). Revisiting some Productivity Debates. *NBER Working paper,* No. 10065.

Van Biesebroeck, J. (2005a, December). Exporting Raises Productivity in sub-Saharan African Manufacturing Firms. *Journal of International Economics 67*(2), 373–91.

Van Biesebroeck, J. (2005b, March). Firm Size Matters: Growth and Productivity Growth in African Manufacturing. *Economic Development and Cultural Change 53*(3), 545–84.

Van Biesebroeck, J. (200X, forthcoming). Robustness of Productivity Estimates. *Journal of Industrial Economics*.

Wagner, J. (2002). The Causal Effects of Exports on Firm Size and Labor Productivity: First Evidence from a Matching Approach. *Economics Letters 77*, 287–92.

Yasar, M. and C. J. Morrison Paul (2005, March). Foreign Technology Transfer and Productivity: Evidence from a Matched Sample. mimeo, U.C. Davis.

Table 1: Summary statistics

	Colombia		Zimbabwe	
	mean	stand. dev.	mean	stand. dev.
value added	4007	30575	164478	365686
employment	66	141	337	655
capital stock (% of VA)	0.397	1.651	1.458	2.401
investment (% of VA)	0.077	0.708	0.312	0.920
wage bill (% of VA)	0.572	0.407	0.492	0.664
output growth	0.041	0.524	0.079	0.570
employment growth	-0.013	0.354	0.035	0.267
capital growth	-0.033	0.560	0.326	1.055
exports output (yes = 1)	0.087	0.282	0.542	0.499
imports inputs (yes = 1)	0.064	0.245	0.644	0.480
paid royalties? (Colombia); any technology effort? (Zimbabwe)	0.019	0.136	0.682	0.467
fraction of years with positive investments	0.352	0.478	0.403	0.491
fraction of skilled/technical employees	0.344	0.475	0.350	0.477
number of years active	9.423	4.132	2.421	0.669
unique firms/plants	1957		179	
number of observations	14348		394	

Notes: For more information on the data sets see Roberts (1996) for Colombia and Van Biesebroeck (2005b) for Zimbabwe.

Table 2: Input coefficient estimates

	Colombia			Zimbabwe		
	labor	capital	RTS	labor	capital	RTS
IN[1]	0.53	0.47	1.00	0.39	0.61	1.00
	(0.20)	(0.20)		(0.35)	(0.35)	
DEA[2]	0.72	0.28	0.88	0.68	0.32	1.04
	(0.33)	(0.33)		(0.27)	(0.27)	
OLS	0.86	0.18	1.04	0.82	0.39	1.21
	(.007)	(.004)		(.055)	(.033)	
SF1	0.68	0.17	0.85	0.70	0.64	1.34
	(.018)	(.010)		(.068)	(.064)	
SF2	0.32	0.09	0.41	0.38	0.10	0.48
	(.011)	(.009)		(.055)	(.030)	
GMM	0.21	0.24	0.45	0.70	0.43	1.13
	(.009)	(.007)		(.054)	(.033)	
OP	0.74	0.18	0.92	0.75	0.36	1.11
	(.023)	(.017)		(.061)	(.040)	

Notes: For the parametric methods the coefficient estimates are reported with standard errors in parenthesis (RTS stands for returns to scale). [1] The labor coefficient for IN is the average wage bill as a fraction of value added, the capital coefficient is 1 - labor coefficient. The standard deviations over the entire sample are in parenthesis. [2]For DEA, the median of the relative input weights and the median returns to scale are reported; standard deviations are in parenthesis.

Table 3: Productivity level and growth estimates

	Colombia				Zimbabwe			
	productivity level		**productivity growth**		**productivity level**		**productivity growth**	
	Interquartile range		Average		Interquartile range		Average	
	25^{th} %	75^{th} %	unweighted	weighted	25^{th} %	75^{th} %	unweighted	weighted
	(1a)	(2a)	(3a)	(4a)	(1b)	(2b)	(3b)	(4b)
IN	-0.413	0.382	0.062	0.088	-0.387	0.742	-0.136	-0.336
DEA	-0.498	0.522	0.065	0.106	-0.535	0.618	-0.074	-0.042
SF1	-0.360	0.340	0.060	0.121	-0.726	0.631	-0.133	-0.273
SF2	-0.566	0.457	0.049	0.054	-0.399	1.278	0.024	0.037
GMM	-0.542	0.491	0.057	0.120	-0.499	0.547	-0.080	-0.141
OP	-0.337	0.319	0.061	0.121	-0.503	0.554	-0.061	-0.101

Notes: The quartiles for productivity level are for the entire sample, pooling all plant-year (Colombia) or firm-year (Zimbabwe) observations, normalizing productivity by the median for the year. The average productivity growth statistics are also calculated over the entire sample and output weights by year) are used when weighing.

Table 4: Correlations between different productivity level and growth estimates

Colombia

Productivity level

	IN	DEA	GMM	SF1	SF2	OP
IN	1					
DEA	0.87	1				
GMM	0.66	0.87	1			
SF1	0.76	0.80	0.89	1		
SF2	0.46	0.70	0.93	0.80	1	
OP	0.78	0.77	0.83	0.99	0.73	1
LP	0.53	0.49	0.67	0.92	0.65	0.92
labor (level)	0.10	0.43	0.66	0.27	0.71	0.17

Productivity growth

	IN	DEA	GMM	SF1	SF2	OP
IN	1					
DEA	0.93	1				
GMM	0.91	0.94	1			
SF1	0.90	0.86	0.95	1		
SF2	0.45	0.47	0.54	0.50	1	
OP	0.90	0.84	0.92	1.00	0.49	1
LP	0.79	0.70	0.82	0.96	0.45	0.97
labor (t) (growth)	-0.10	0.08	0.09	-0.23	0.09	-0.27
labor (t+1) (growth)	0.04	0.04	0.05	0.10	0.06	0.10

Zimbabwe

Productivity level

	IN	DEA	GMM	SF1	SF2	OP
IN	1					
DEA	0.79	1				
GMM	0.93	0.84	1			
SF1	0.71	0.55	0.81	1		
SF2	0.38	0.62	0.34	0.01	1	
OP	0.92	0.85	0.99	0.72	0.45	1
LP	0.68	0.70	0.73	0.25	0.82	0.82
labor (level)	0.28	0.34	0.11	-0.41	0.83	0.20

Productivity growth

	IN	DEA	GMM	SF1	SF2	OP
IN	1					
DEA	0.77	1				
GMM	0.88	0.93	1			
SF1	0.90	0.89	0.98	1		
SF2	0.63	0.69	0.75	0.66	1	
OP	0.86	0.93	1.00	0.95	0.78	1
LP	0.54	0.72	0.75	0.62	0.78	0.81
labor (t) (growth)	0.03	-0.09	-0.10	-0.11	0.10	-0.11
labor (t+1) (growth)	0.01	0.04	0.09	0.12	0.10	0.08

Notes: Partial correlations statistics between the different (log) productivity measures and between productivity and labor input; correlations are calculated across all observations (plant/firm - years). The bottom row reports the correlation between labor input growth and one period lagged productivity.

Table 5: First debate: Learning-by-exporting

	Colombia			Zimbabwe		
	OLS	simultaneous equations	matching estimator	OLS	simultaneous equations	matching estimator
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
IN	0.020	0.032	0.091	0.348**	0.124	-0.061
	(.029)	(.038)	(.133)	(.164)	(.203)	(.272)
DEA	0.026	0.051	0.008	0.408***	0.151	0.557
	(.030)	(.038)	(.140)	(.141)	(.209)	(.821)
GMM	0.268***	0.126***	0.127	0.393***	0.049	0.576
	(.026)	(.035)	(.105)	(.135)	(.212)	(.505)
SF1	0.333***	0.156***	0.230**	0.386***	-0.057	0.124
	(.026)	(.035)	(.095)	(.147)	(.193)	(.808)
SF2	0.399***	0.031**	0.139	0.376***	-0.033	0.056
	(.024)	(.015)	(.091)	(.142)	(.149)	(1.061)
OP	0.326***	0.128***	0.223**	0.396***	0.061	0.695
	(.026)	(.035)	(.095)	(.134)	(.209)	(.526)

Note: The reported statistics are coefficient estimates and standard errors on the lagged export dummy in separate regressions with the different log productivity measures as dependent variable. Results in the first columns are for an OLS regression on the full sample, controlling for employment and time, location, and industry dummies. Statistics in the second column are coefficient estimates on once lagged export status in the productivity equation, estimated by the simultaneous equation model of Clerides *et al.* (1998). Estimates in the third column are from a regression similar to the first column, but on the limited sample of treated (new exporters) and matched plants, using nearest neighbor matching with replacement. The propensity score used in the match is estimated by a Probit on lagged productivity, employment, and wages and time, location, and industry dummies. *** Significant at the 1% level, ** 5%, * 10%.

Table 6: Second debate: Endogenous growth

	Colombia					Zimbabwe				
	dependent variable: productivity growth -- explanatory variable in column headings									
	export output	import inputs	pay for royalties	invest frequently	high HC	export output	import inputs	invest in technology	invest frequently	high HC
IN	-0.002	0.003	-0.019	-0.022	-0.285**	-0.014	0.002	0.197	0.045	-0.418
	(0.016)	(0.021)	(0.024)	(0.019)	(0.047)	(0.140)	(0.135)	(0.186)	(0.197)	(0.391)
DEA	0.008	-0.009	-0.021	0.009	-0.246**	0.152	0.132	0.302**	0.173	-0.404
	(0.016)	(0.021)	(0.024)	(0.018)	(0.046)	(0.110)	(0.106)	(0.142)	(0.155)	(0.309)
GMM	0.036**	-0.013	-0.029	0.055**	-0.089**	0.094	0.092	0.271**	0.050	-0.198
	(0.015)	(0.020)	(0.022)	(0.017)	(0.044)	(0.106)	(0.103)	(0.138)	(0.150)	(0.299)
SF1	0.032**	-0.004	-0.025	0.034**	-0.065	0.124	0.106	0.266**	0.106	-0.183
	(0.015)	(0.019)	(0.022)	(0.017)	(0.043)	(0.128)	(0.123)	(0.166)	(0.179)	(0.358)
SF2	0.019	-0.020	-0.047	0.047**	-0.084	-0.095	0.393**	0.260	-0.195	0.111
	(0.023)	(0.030)	(0.034)	(0.026)	(0.067)	(0.173)	(0.170)	(0.225)	(0.231)	(0.431)
OP	0.030**	-0.003	-0.025	0.027*	-0.070*	0.086	0.088	0.273**	0.032	-0.195
	(0.015)	(0.019)	(0.022)	(0.017)	(0.043)	(0.101)	(0.097)	(0.131)	(0.142)	(0.284)

Notes: Each coefficient comes from a separate regression of productivity growth, averaged over the period each plant or firm is active, on the different covariates indicated in each column, including time, industry, and location dummies as controls. ** Significant at 5% level * 10%.

Table 7: Third debate: Importance of within-plant/firm changes

	Colombia (1981 - 1991)				Zimbabwe (1993 - 1995)			
	arithmetic average		geometric average		arithmetic average		geometric average	
	aggregate growth	unweighted growth	BHC	GR	aggregate growth	unweighted growth	BHC	GR
	(1a)	(2a)	(3a)	(4a)	(1b)	(2b)	(3b)	(4b)
IN[1]	1.207	84%	0.180	0.241	-0.331	75%	-0.320	-0.315
DEA	0.986	89%	0.154	0.221	-0.227	98%	-0.173	-0.129
GMM	0.743	102%	0.120	0.234	-0.279	76%	-0.275	-0.264
SF1	1.054	109%	0.148	0.239	-0.374	100%	-0.408	-0.450
SF2	0.922	90%	0.127	0.240	-0.062	12%	-0.056	0.061
OP	1.160	107%	0.154	0.240	-0.243	71%	-0.236	-0.211

Notes: Statistics in columns (1a) and (1b) are cumulative aggregate growth rates, i.e. the percentage change over the entire period in the aggregate productivity level calculated according to equation (21). Statistics in columns (2a) and (2b) are the cumulative growth rates for the first term on the right hand side of equation (21), expressed as a percentage of the growth rate in the first columns. Columns (3a) and (3b) are the within plant effects in the decomposition of the log productivity aggregate, i.e. the first term in equation (20). Columns (4a) and (4b) contain the same statistic, but use the average output weight instead of the initial weight.

[1] The index numbers use the same input weights for each plant/firm in the respective samples (the average wage share) to construct an aggregate growth rate that is unit-invariant.