Nonparametric Production Analysis with Unobserved heterogeneity in productivity

Laurens Cherchye
Department of Economics, KU Leuven

Thomas Demuynck
ECARES, Université libre de Bruxelles

Bram De Rock
ECARES, SBS-EM, Université libre de Bruxelles
Department of Economics, KU Leuven

Marijn Verschelde
IESEG School of Management, Paris
Department of Economics, KU Leuven

September, 2018

ECARES working paper 2018-25

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Nonparametric production analysis with unobserved heterogeneity in productivity*

Laurens Cherchye† Thomas Demuynck‡ Bram De Rock§
Marijn Verschelde¶

September 20, 2018

Abstract

We propose a novel nonparametric method for the structural production analysis in the presence of unobserved heterogeneity in productivity. We assume cost minimization as the firms' behavioral objective, and we model productivity on which firms condition the input demand of the observed inputs. Our model can equivalently be represented in terms of endogenously chosen latent input costs that guarantee data consistency with our behavioral assumption, and we argue that this avoids a simultaneity bias in a natural way. Our Monte Carlo simulation and empirical application to Belgian manufacturing data show that our method allows for drawing strong and robust conclusions, despite its nonparametric orientation. For example, our results pinpoint a clear link between international exposure and productivity and show that primary inputs are substituted for materials rather than for productivity enhancement.

Keywords: productivity, unobserved heterogeneity, simultaneity bias, nonparametric production analysis, cost minimisation, manufacturing

---

*We are very grateful to the National Bank of Belgium for supporting this project. In particular, we thank Catherine Fuss for her support and Jean-Marc Troch for data assistance. We also thank Jan De Loecker, Knox Lovell, Johannes Van Biesebroeck, Ruben Dewitte, Klaas Mulier, Veerle Hennebel and the participants of EWEPA 2017, EEA 2018 and EARIE 2018 for useful comments. The computational resources and services used in this work for the Monte Carlo simulation were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government - department EWI. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the National Bank of Belgium or any other institution to which the authors are affiliated.

†Department of economics, University of Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven.be. Laurens Cherchye gratefully acknowledges the European Research Council (ERC) for his Consolidator Grant 614221. Part of this research is financed by the Fund of Scientific Research Flanders (FWO-Vlaanderen).

‡ECARES, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: thomas.demuynck@ulb.ac.be.

§ECARES, Université Libre de Bruxelles, and Department of Economics, University of Leuven (KU Leuven). Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges FWO and FNRS for their financial support.

¶Department of Economics and Quantitative Methods, IÉSEG School of Management, LEM (UMR-CNRS 9221), and Department of Economics, KU Leuven. E-mail: m.verschelde@ieseg.fr.

JEL: C14, D21, D22, D24.

1 Introduction

The increasing prevalence of global sourcing (Antras and Helpman, 2004) and changing input cost shares (Autor et al., 1998, 2003, 2017) lies at the hart of the industrial policy debate.[1] Paradoxically, these phenomena are excluded by construction under the assumption of Hicks neutral technical change, which is usually made in existing methods for empirical production analysis. The few empirical production studies that do relax this assumption of Hicks neutrality typically rely on a specific parametrization of the production technology or impose a common structure on the factor bias across firms.[2] However, empirical evidence and economic theory show that there can be firm heterogeneity in factor biased technical change (see, for example, Acemoglu et al. (2015) and references therein). This makes an a priori parametrization difficult.

In a series of seminal papers, Afriat (1972), Hanoch and Rothschild (1972), Diewert and Parkan (1983) and Varian (1984) proposed an intrinsically nonparametric approach to address the identification of production functions.[3] It recovers the production possibilities directly from the data and avoids functional specification bias by not imposing any (nonverifiable) parametric structure on the production technology. Its identifying power comes from a structural specification of the firms' objectives that underlie the observed production behavior.

Despite this conceptually appealing starting point, the more recent literature on the identification and estimation of production functions has largely ignored this nonparametric alternative. We interpret this lack of attention as principally originating from the fact that the existing nonparametric methods are unable to deal with heterogeneity in unobserved productivity. The importance of effectively dealing with unobserved productivity is by now well-established in the literature (see, for example, the recent review of Syverson (2011)). Basically, incorporating unobservables in the empirical analysis is a prerequisite to account for endogeneity between input choice and unobserved productivity. This endogeneity issue was first pointed out by Marschak and Andrews (1944), and originates from the fact that a firm's productivity transmits to its optimal input choices. It implies that standard OLS-type estimation techniques will suffer from a simultaneity bias (see also Olley and Pakes (1996) and Griliches and Mairesse (1998)).[4]

[1] See, for example, OECD (2012), Dall'Olio et al. (2013) and CompNet Task Force (2014) for empirical studies of manufacturing firms with a specific focus on policy implications.

[2] See, for example, the recent study of Doraszelski and Jaumandreu (2018), which parametrizes the labor augmenting technological change next to the Hicks neutral technological change in a constant elasticity of substitution (CES) framework.

[3] We refer to Grifell-Tatjé et al. (2018) (and references therein) for a recent review of alternative approaches of productivity measurement that have been proposed in the Economics and OR/MS literature.

[4] The literature on the estimation and identification of production functions has paid considerable

The principal aim of the current paper is to re-establish the nonparametric approach as a full-fledged alternative for empirical production analysis. To this end, we present a methodology that uses minimal assumptions to address identification of unobserved productivity differences across firms. Specifically, we assume cost minimization as the firms' behavioral objective and we model unobserved heterogeneity as an unobserved productivity factor, on which we condition the input demand of observed inputs. This avoids the endogeneity bias in a natural way, by explicitly accounting for the simultaneity between productivity and input decisions in our structural specification of the firm's optimization problem. We also provide a novel and intuitive way to quantify unobserved heterogeneity in terms of endogenously chosen latent input costs. Our method allows us to analyze cost share changes of both observed and latent input costs. For example, we can investigate to what extent observed primary manufacturing inputs are substituted over time for other observed inputs and/or latent input costs. This unique feature is intrinsic to the nonparametric nature of our methodology, which avoids imposing particular functional structure on the (changing) production technology (such as Hicks neutrality).

An attractive feature of our method is its empirical applicability. The method can be operationalized through linear programming, which makes it easy to apply in practice. Next, in contrast to most production function estimators, our method is based on gross output rather than on value added. As such, our methodology follows closely the theory of the firm.[5] Furthermore, as our methodology is not based on structuring the timing of (input) decisions, it can be applied on both panel and cross-sectional data. We demonstrate this through a Monte Carlo simulation and an empirical application that studies productivity differences at the firm-year level in the Belgian manufacturing sector for the period 1997-2007. Our application shows that our method does allow for drawing strong and robust conclusions, despite its nonparametric orientation. For the period under study, we confirm the well-established connection between international exposure and productivity. Generally, the cost share of latent input costs remains constant over time, which is in accordance with the well-documented productivity slowdown in manufacturing since the early 2000s (see, for example, Syverson (2017)). Further, we document that Belgian manufacturing firms substitute labor and capital for domestic and foreign materials (i.e., outsourcing), rather than for latent input costs (i.e., technology), and that this substitution pattern is more pronounced for large firms. We also show that our results are robust for altering our revenue based definition of output to a produced value based definition that excludes servicing and carry-along trade (see Bernard et al. (Forthcoming)). We see all this as strong empirical evidence against the assumption of Hicks neutrality.[6]

---

attention to developing techniques that address this endogeneity problem. Notable examples include Olley and Pakes (1996); Levinsohn and Petrin (2003); Wooldridge (2009); Ackerberg et al. (2015); Gandhi et al. (2017). A main difference with our nonparametric approach is that the empirical implementation of these existing approaches requires a (semi)parametric specification of the production technology.

[5] See, for example, Gandhi et al. (2017) for a comparison of gross output and value added production function estimates.

[6] Our empirical application shows general patterns of both observed and latent input cost share changes,

The remainder of this paper unfolds as follows. Section 2 presents our novel methodology for nonparametric production analysis with unobserved heterogeneity. We also introduce our nonparametric measure to empirically quantify productivity, and we indicate how to bring our model to data. We end Section 2 with a Monte Carlo simulation in which we demonstrate the attractivity of our advocated methodology for noisy production data. Section 3 motivates our application to Belgian manufacturing firms, and discusses the input and output data that we use. Section 4 presents our main empirical findings. Section 5 concludes and discusses possible avenues for follow-up research.

2 Methodology

We begin this section by presenting our specification of the firm's optimization problem under productivity heterogeneity, to subsequently establish the associated nonparametric characterization of optimizing firm behavior. This will pave the way for introducing our concept of a productivity factor to empirically quantify differences in productivity between firms. We conclude by showing how to bring our model to data. We will explain how we can account for (small) deviations from "exactly" optimizing behavior in empirical applications, by using a nonparametric measure of goodness-of-fit.

2.1 Production with heterogeneity in productivity

Firms' production levels depend on observed inputs, as well as on unobserved productivity. Formally, we assume a production function F that defines

$$Q = F(\mathbf{X}, \Omega),$$

for $Q \in \mathbb{R}_+$ the output level, $\mathbf{X} \in \mathbb{R}_+^M$ a M-dimensional vector of observed inputs, and $\Omega \in \mathbb{R}_+$ a single-dimensional measure of the unobserved productivity heterogeneity in the production process across firms. The assumption that unobserved productivity differences are one-dimensional follows the standard practice in the literature (see, for example, Olley and Pakes (1996); Levinsohn and Petrin (2003); Wooldridge (2009); Ackerberg et al. (2015); Gandhi et al. (2017)). A useful implication is that it allows for a transparent empirical analysis of heterogeneity patterns, as we will demonstrate in our empirical application in Section 4.

Generally, we can interpret the unobserved Ω in two ways.[7] In the first interpretation, Ω

hereby specifically concentrating on the role of the productivity factor. Our results on observed cost share changes fall in line with those reported by Verschelde et al. (2014), who focused on changes of output elasticities over time for a closely similar dataset of Belgian manufacturing firms. More recently, Dewitte et al. (2017) provided a detailed study of heterogeneity in factor biased technological change for Belgian manufacturing firms. These authors considered firm-level changes of output elasticities by applying a nonparametric kernel regression with time-varying fixed effects.

[7] See Syverson (2011) for a general discussion on alternative interpretations of productivity differences that appeared in the literature.

falls beyond the firms' control and stands for external drivers of productivity and random productivity shocks. For example, firms with higher $\mathit{\Omega}$ have access to better technologies, thereby increasing their output $F(\mathbf{X}, \mathit{\Omega})$ for the same level of observed inputs $\mathbf{X}$. Alternatively, we can also interpret $\mathit{\Omega}$ as a latent input, which implies that it is optimally chosen by the firm. This interpretation includes all factors under the control of the firm that influence productivity, such as managerial input, information technology and R&D. Importantly, while the two interpretations are clearly distinct, we will show that the associated models of optimizing firm behavior are empirically equivalent in terms of their nonparametric testable implications. As a result, our following characterization of optimizing behavior does not depend on the specific meaning that is attached to $\mathit{\Omega}$.

Throughout, we assume that the function F is strictly monotonic, continuous and jointly concave in $(\mathbf{X}, \mathit{\Omega})$. In addition, we postulate that the production technology is characterized by constant returns-to-scale (CRS), which means that, for all numbers $t > 0$,

$$F(t\mathbf{X}, t\mathit{\Omega}) = tF(\mathbf{X}, \mathit{\Omega}).$$

In our main empirical analysis, we only impose CRS within a specific firm size category, as productivity heterogeneity is analyzed for each firm size category separately. This effectively implies that we (only) assume CRS to hold "locally" (i.e., for the given firm size), so avoiding the "global" CRS postulate, which –admittedly– may seem overly strong in many practical settings.

Usually, the CRS assumption is motivated by a replication argument: if one doubles all the inputs, one can always double the output. Implicitly, this assumes that all inputs are taken into account. From this perspective, we can effectively motivate the CRS assumption in our context by interpreting $\mathit{\Omega}$ as a latent input factor. That is, we assume constant returns to scale of the production function $F(\mathbf{X}, \mathit{\Omega})$, which takes into account both the observable and unobservable inputs. In this respect, it is worth remarking that imposing CRS on the production technology with latent input does not impose any specific returns-to-scale structure on the functional relation between observed input and output. As such, it is fully consistent with any evidence on variable (decreasing/increasing) returns-to-scale in terms of observed production factors that has been documented in the literature for particular production settings.

Finally, and importantly, our CRS assumption is crucial for obtaining nonparametric identifying restrictions in terms of observed input and output. It can be verified that our assumption of cost minimizing behavior with latent input would define vacuous conditions for rationalizable production behavior (as specified in the following Definition 1) if we put no structure on the returns-to-scale of the production technology. For example, we obtain this "negative" conclusion by setting the cost of the latent input sufficiently high, so that the observed input does not generate meaningfully testable implications associated with cost minimization for the observed output.[8] Essentially, our CRS assumption disciplines

[8]Compare with Varian (1988), who formalized a similar argument in a consumption context.

the latent cost structure in a way that excludes such a trivial rationalization of the observed production behavior.

2.2 Cost minimizing production behavior

Throughout we assume that firms are price takers in the input market and we impose no structure on the form of the output market. As shown by Carvajal et al. (2013, 2014), it is possible to impose alternative (for example, Cournot or Bertrand) structures on the output market in our advocated nonparametric framework. In our following analysis, we purposely do not impose any such assumption, so showing that our identification results are independent of the output market form.

Let $\mathbf{W} \in \mathbb{R}^M_{++}$ be the price vector for the observed inputs. Our above two interpretations of the heterogeneity factor $\mathit{\Omega}$ yield two different models of optimizing firm behavior. First, if we assume that $\mathit{\Omega}$ is beyond the firm's control, then the firm solves the optimization problem

$$\min_{\mathbf{X}} \mathbf{WX} \text{ s.t. } F(\mathbf{X}, \mathit{\Omega}) \geq Q_0 \qquad (OP.I).$$

That is, the firm's input choice $\mathbf{X}$ is conditional on the unobserved factor $\mathit{\Omega}$. Second, if $\mathit{\Omega}$ is a latent input factor that is chosen by the optimizing firm, then this firm solves

$$\min_{\mathbf{X},\mathit{\Omega}} \mathbf{WX} + \mathit{\Gamma}\mathit{\Omega} \text{ s.t. } F(\mathbf{X}, \mathit{\Omega}) \geq Q_0, \qquad (OP.II)$$

for $\mathit{\Gamma} \in \mathbb{R}_{++}$ the unobserved price of $\mathit{\Omega}$. In both scenarios, the simultaneity bias is absent by construction, because either the (observed) inputs $\mathbf{X}$ are optimally chosen conditionally on the unobserved $\mathit{\Omega}$ or, alternatively, these inputs are defined simultaneously with the latent input $\mathit{\Omega}$.

We demonstrate the empirical equivalence of optimizing behavior in terms of (OP.I) and (OP.II) by establishing the associated testable implications. To this end, we assume to observe a dataset

$$S = \{\mathbf{W}_i, \mathbf{X}_i, Q_i\}_{i \in N},$$

with $\mathbf{W}_i$ the observed input prices, $\mathbf{X}_i$ the observed input levels, and Q_i the observed output levels for a set of N firm observations. The data set can be a cross-section, a time-series or, as in our own empirical application, a panel with firm observations specified at the firm-year level. The set S contains all information on observed production behavior that is used by the empirical analyst. In principle, it is possible to integrate in our set-up extra information on indicators that are (assumed to be) correlated with the unobserved technological heterogeneity (e.g., R&D investments). Again, we intentionally restrict to our minimalistic setting to show the generality of our identification results.

The functional form of the production function F is unknown to the empirical analyst. Our nonparametric method basically checks whether there exists at least one specification of F that represents the observed firm behavior in terms of the optimization problems (OP.I) and (OP.II). If such a function exists, we say that the dataset S is rationalizable in terms of (OP.I) and (OP.II).

Definition 1. *Let $S = \{\mathbf{W}_i, \mathbf{X}_i, Q_i\}_{i \in N}$ be a given dataset. S is (OP.I)-rationalizable if there exist numbers $\Omega_i \in \mathbb{R}_+$ and a production function $F : \mathbb{R}_+^{M+1} \to \mathbb{R}_+$ such that, for all firm observations $i \in N$,*

$$\mathbf{X}_i \in \arg\min_{\mathbf{X}} \mathbf{W}_i \mathbf{X} \ \ s.t. \ \ F(\mathbf{X}, \Omega_i) \geq Q_i.$$

The dataset S is (OP.II)-rationalizable if, in addition, there exist prices $\Gamma_i \in \mathbb{R}_+$ such that, for all firm observations $i \in N$,

$$(\mathbf{X}_i, \Omega_i) \in \arg\min_{\mathbf{X}, \Omega} \mathbf{W}_i \mathbf{X} + \Gamma_i \Omega \ \ s.t. \ \ F(\mathbf{X}, \Omega) \geq Q_i.$$

In Appendix A.1 we prove that (OP.I)-rationalizability and (OP.II)-rationalizability generate exactly the same nonparametric testable implications for a given dataset S. This conclusion is summarized in the following proposition.

Proposition 1. *Let $S = \{\mathbf{W}_i, \mathbf{X}_i, Q_i\}_{i \in N}$ be a given dataset. The following statements are equivalent:*

(i) The dataset S is (OP.I)-rationalizable;

(ii) The dataset S is (OP.II)-rationalizable;

(iii) There exist $\Omega_i \in \mathbb{R}_+$ and $\Lambda_i \in \mathbb{R}_{++}$ that satisfy, for all $i, j \in N$, the inequalities

$$\frac{Q_i}{Q_j} \leq \frac{\Lambda_j \mathbf{W}_j \mathbf{X}_i + \Omega_i}{\Lambda_j \mathbf{W}_j \mathbf{X}_j + \Omega_j}.$$

To sharpen the intuition behind condition (iii) in Proposition 1, we start from the observation that the input bundle $(\mathbf{X}_i, \Omega_i)$ can produce the output Q_i. Then, it follows from our CRS assumption that the rescaled input bundle $\frac{Q_j}{Q_i} \times (X_i, \Omega_i)$ must be able to produce the output level Q_j. The cost of using this last input combination at the prices (W_j, Γ_j) equals

$$\frac{Q_j}{Q_i}(\mathbf{W}_j \mathbf{X}_i + \Gamma_j \Omega_i).$$

On the other hand, cost minimizing production behavior (as specified in (OP.II)) also requires that, at the prices $(\mathbf{W}_j, \Gamma_j)$, the input bundle $(\mathbf{X}_j, \Omega_j)$ produces the output Q_j at a lower or equal cost. Thus, we must have

$$\mathbf{W}_j \mathbf{X}_j + \Gamma_j \Omega_j \leq \frac{Q_j}{Q_i}(\mathbf{W}_j \mathbf{X}_i + \Gamma_j \Omega_i),$$

Using $\varLambda_j = 1/\varGamma_j$, this effectively obtains condition (iii) in Proposition 1.

In the next section we show how the inequalities in the third statement of Proposition 1 can be brought to the data by means of linear programming, which will allow us to specify values of $\varOmega_i$ that rationalize the dataset S. Moreover, when interpreting these numbers $\varOmega_i$ as representing latent input quantities, the associated numbers $\varLambda_i$ give the inverse of the corresponding shadow input prices ($1/\varGamma_i$). Interestingly, we can use this to nonparametrically quantify productivity heterogeneity in terms of latent input cost, which we refer to as our nonparametric estimate of productivity *"NP"*,

$$\mathrm{NP} = \varGamma\varOmega.$$

It readily follows from our above discussion that this NP measure has a direct interpretation as capturing productivity differences. All else equal, higher NP values indicate that the same output can be produced with less observed costs, which effectively reveals a higher (unobserved) productivity level. In our empirical analysis, we will not only focus on *NP*-levels, but also on the *"cost share of latent input"*,

$$\mathrm{CSLI} = \frac{\varGamma\varOmega}{\mathbf{WX} + \varGamma\varOmega},$$

which expresses the firm's latent input as a fraction of the total (observed plus latent) input cost. This measure is naturally bounded between zero and one, and a higher CSLI value indicates a greater importance of the latent input relative to the other (observed) inputs. As we will show in our empirical application, we can use the CSLI measure to investigate substitution patterns between the observed inputs and the latent input (i.e., technology).

As a concluding note, Appendix A.2 presents a numerical example that illustrates the testable implications in Proposition 1. It shows that our empirical conditions for cost minimization with unobserved productivity differences can be rejected (i.e., have empirical content) even in a minimalistic setting with only two firm observations and two observed inputs. Generally, the empirical bite of the conditions will increase with the number of observations and observed inputs.

2.3 Bringing our model to data

The rationalizability conditions in Proposition 1 are strict: either the dataset S satisfies them "exactly" or it does not. In practice, it is often useful to allow for small deviations from exactly rationalizable behavior. Such deviations may be due to (small) unanticipated shocks experienced by the firms or, alternatively, data imperfections (for example, ill-measured input/output quantities and/or input prices).[9] To include these possibilities,

[9]In fact, it is also possible to explicitly account for measurement errors in prices and quantities in our nonparametric analysis. For example, we may use the procedure suggested by Varian (1985) and Epstein and Yatchew (1985), which is fairly easily adjusted to our setting. This complies with the more standard

we define a nonparametric goodness-of-fit parameter that has an intuitive economic interpretation in terms of departures from the cost minimization hypothesis that we maintain as our core identifying assumption (see Afriat (1972) and Varian (1990)).[10] By fixing our goodness-of-fit parameter at a value close to (but different from) one, we can take account of observed behavior that is close to (but not exactly) rationalizable in the sense of Definition 1.

More precisely, we increase the right hand sides of the inequality requirements in Proposition 1 by using the goodness-of-fit parameter θ (with $0 \leq \theta \leq 1$) to specify

$$\frac{Q_i}{Q_j} \leq \frac{\mathit{\Lambda}_j \mathbf{W}_j \mathbf{X}_i + \mathit{\Omega}_i}{\mathit{\Lambda}_j \mathbf{W}_j \mathbf{X}_j + \mathit{\Omega}_j} + (1-\theta). \tag{1}$$

Obviously $\theta = 1$ obtains the exact conditions in Proposition 1, while lower values for θ weaken the rationalizability requirements. To further interpret our goodness-of-fit measure, we formally show in Appendix A.3 that adding $(1-\theta)$ is equivalent to equiproportionally contracting the inputs $(\mathbf{X}_j, \mathit{\Omega}_j)$ to $(\theta\mathbf{X}_j, \theta\mathit{\Omega}_j)$. This in turn corresponds to lowering the cost level $(\mathbf{W}_j\mathbf{X}_j + \mathit{\Gamma}_j\mathit{\Omega}_j)$ by the same degree and, therefore, implies a weaker criterion of "nearly" (instead of "exactly") optimizing behavior. In our following empirical application, our main focus will be on $\theta = 0.95$ (which, intuitively, decreases firm i's total cost level $(\mathbf{W}_j\mathbf{X}_j + \mathit{\Gamma}_j\mathit{\Omega}_j)$ by 5 percent). In Appendix C, we also check robustness of our main results for alternative θ-values.

To bring our inequalities to the data, we reformulate (1) as

$$Q_i(\mathit{\Lambda}_j \mathbf{W}_j \mathbf{X}_j + \mathit{\Omega}_j) - Q_j(\mathit{\Lambda}_j \mathbf{W}_j \mathbf{X}_i + \mathit{\Omega}_i) \leq (1-\theta) Q_j(\mathit{\Lambda}_j \mathbf{W}_j \mathbf{X}_j + \mathit{\Omega}_j). \tag{2}$$

For a fixed value of θ, this defines restrictions that are linear in the unknowns $\mathit{\Lambda}_i$ and $\mathit{\Omega}_i$. We can use simple linear programming tools to check if there exists a solution of (2) and, thus, to conclude if the dataset is exactly rationalizable (when using $\theta = 1$) or nearly rationalizable (when using $\theta < 1$).

Finally, the linear restrictions (2) will generally define a multitude of feasible specifications of $\mathit{\Lambda}_i$ and $\mathit{\Omega}_i$ (and, thus, of our productivity measures NP and CSLI). To empirically evaluate the importance of firm heterogeneity, a natural choice is to use the specification that minimizes the cost shares of latent input that are required for rationalizability (as

econometric use of a minimum distance criterion. To facilitate our exposition, we will not consider this extension in the current paper. In the current context, measurement error in the output quantity can also be interpreted as reflecting productivity shocks that are not anticipated by the firm. Our following simulation exercise will include output error and will show how our goodness-of-fit parameter θ incorporates this error.

[10] In a similar spirit, Varian (1990) argues that such a nonparametric goodness-of-fit measure can also be interpreted in terms of "economic significance" of departures from optimization, which is to be distinguished from the more standard notion of statistical significance.

characterized in (2)) or, equivalently, that maximizes the role played by the observed inputs. This corresponds to solving the program

$$\min \sum_i \mathrm{CSLI}_i = \min \sum_i \frac{\Gamma_i \Omega_i}{\mathbf{W}_i \mathbf{X}_i + \Gamma_i \Omega_i} = \min \sum_i \frac{(1/\Lambda_i)\Omega_i}{\mathbf{W}_i \mathbf{X}_i + (1/\Lambda_i)\Omega_i} \tag{3}$$

subject to the linear restrictions (2). In Appendix A.4, we discuss the technical issue that the objective function in (3) is not linear in the unkowns Λ_i and Ω_i. Given this, we replace this objective function with a linearized version, which conveniently allows us to use standard linear programming techniques in our empirical analysis.

2.4 Monte Carlo simulation

Before proceeding to our empirical application, we demonstrate the usefulness of our advocated nonparametric methodology by means of a Monte Carlo simulation analysis. This will show the proper working of our methodology for noisy production data.

Set-up. We generate data that resemble actual production data as used in empirical firm-level productivity studies, with as basis income and balance sheet information. A main focus of our following exercise will be on comparing the results of our methodology with the ones of the methodology that was recently proposed by Ackerberg et al. (2015) (ACF in what follows). Following these authors, we assume a representative firm that is dynamically optimizing, as in Van Biesebroeck (2007). We generate data that is consistent with the ACF model with a Cobb-Douglas production technology. We refer to Appendix A.5 for the technical details of our data generating process.

In what follows, we will compare our estimates of productivity with the ones obtained through the ACF methodology. We also compare with OLS estimates based on a Cobb Douglas specification of the production technology. These OLS estimates serve as benchmarks corresponding to a parametric methodology that does not control for simultaneity bias. In our analysis, we consider three different productivity indicators: the latent input cost $\Gamma\Omega$, the productivity parameter $a = ln(A)$, and the value $a + \varepsilon$, which includes both anticipated and unanticipated productivity shocks. To assess the performance of the different estimation methods in terms of recovering productivity heterogeneity, we calculate the average Spearman correlations between true and estimated values for all three productivity indicators.

One preliminary remark is in order. In contrast to proxy variable approaches, our methodology makes no use of the panel structure of empirical production data. As such, it is well applicable in cross-sectional settings. The cost of this flexibility is that, in our general set-up, we consider all inputs to be flexible, and the firm's optimization problem is modeled as static rather than dynamic. Importantly, however, under intertemporal separability of the firm's objective function, static optimization is a necessary condition for dynamic optimization. Intertemporal interdependence of input decisions trough investment (as in ACF) can then be accounted for by suitably pricing investment to smooth investment/capital cost over the consecutive time periods. Moreover, our following simulation results, as well as our empirical analysis in Sections 3 and 4, will show that our basic framework is well applicable for the analysis of production function parameters and productivity heterogeneity of manufacturing firms, even when using panel data. In principle, adding structure to tailor our methodology to specific empirical settings is well possible, and –evidently– can be expected to improve the accuracy of the estimation results when this extra structure is correctly specified.

Simulation results. We consider different levels of noise by setting σ_ε equal to 0, 0.1 and 0.3 : $\sigma_\varepsilon = 0.1$ reflects a moderately noisy data, while $\sigma_\varepsilon = 0$ corresponds to a deterministic setting and $\sigma_\varepsilon = 0.3$ to a highly noisy setting. Our main focus will be on simulated samples that consist of 1000 firm observations, which is below the minimum sample size in our empirical application that we present in Sections 3 and 4. Our simulated samples of 1000 firm observations consist of 100 firms (F=100) that we observe over 10 years (T=10). As a robustness check, we also provide simulation results for sample sizes of 500 (F=50,T=10) and 2000 (F=200,T=10) in Appendix A.6. Using these nine scenarios, we construct 1000 Monte Carlo samples (B=1000) for which we provide results below. We set the goodness-of-fit parameter $\theta = 0.95$ for our main Monte Carlo analysis, and provide results for the alternative values $\theta = 0.90$ and $\theta = 0.99$ in Appendix A.6.

Table 1 shows our Monte Carlo estimations for the production function parameters, together with Spearman correlations for our three productivity indicators. As expected, for the scenario with moderately noisy firm data (i.e., $\sigma_\varepsilon = 0.1$), the OLS procedure (ignoring simultaneity bias) obtains highly unreliable and biased estimates of the production function parameters. In contrast, the NP method (with $\theta = 0.95$) and the ACF method (with adequately specified starting values) nicely identify the production function parameters: both approaches show no bias and low standard deviations for the given DGP. All three estimation methods perform well in terms of recovering heterogeneity in $a + \varepsilon$, while the NP and ACF methods outperform the OLS method in recovering a. The correlation with $\Gamma\Omega$ is poor for the OLS-based productivity estimates, and fair for both the NP- and ACF-estimates. The NP method performs slightly worse in recovering a and $a + \varepsilon$ and slightly better in recovering $\Gamma\Omega$ than the ACF method. In sum, the NP and ACF methods both show their value in identifying production function parameters and recovering productivity heterogeneity for moderately noisy data.

In the fully deterministic setting (i.e., $\sigma_\varepsilon = 0$), both the NP and ACF methods reliably recover productivity heterogeneity. In this case the NP method slightly underestimates the CSLI value when using $\theta = 0.95$. However, in Appendix A.6 we show that the NP method with a more adequately specified goodness-of-fit parameter $\theta = 0.99$ (reflecting low noise in the data) does provide unbiased CSLI estimates. For highly noisy data (i.e., $\sigma_\varepsilon = 0.3$), the performance of both the NP and ACF methods in terms of recovering variation in $\Gamma\Omega$ and a is rather weak. Further, the ACF estimates of production function parameters show large standard deviations, and the NP procedure overestimates CSLI. However, and similar to before, when using a more correct specification of the parameter $\theta = 0.90$ (reflecting high noise in the data), the NP method is able to reliably recover the production function parameters and –to a somewhat lesser extent– to identify heterogeneity in productivity.

Table 1: Monte Carlo results: 1000 firm observations

	PRODUCTION FUNCTION PARAMETERS					
	$\alpha^\ell = 0.6$		$\alpha^k = 0.3$		$CSLI = 0.1$	
σ_ε	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP, $\theta = 0.95$			
0	0.613	0.003	0.306	0.001	0.081	0.004
0.1	0.598	0.003	0.299	0.002	0.103	0.005
0.3	0.571	0.007	0.286	0.003	0.143	0.010
			OLS			
0	0.769	0.010	0.318	0.006	–	–
0.1	0.768	0.011	0.318	0.006	–	–
0.3	0.768	0.021	0.318	0.009	–	–
			ACF			
0	0.600	0.019	0.300	0.034	–	–
0.1	0.598	0.023	0.301	0.037	–	–
0.3	0.591	0.044	0.315	0.210	–	–
	SPEARMAN CORRELATION WITH PRODUCTIVITY					
	$\Gamma\Omega$		a		$a + \varepsilon$	
σ_ε	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP			
0	0.766	0.022	0.982	0.003	0.982	0.003
0.1	0.625	0.026	0.749	0.025	0.953	0.023
0.3	0.360	0.028	0.375	0.034	0.810	0.054
			OLS			
0	0.178	0.021	0.759	0.028	0.759	0.028
0.1	0.120	0.016	0.527	0.029	0.846	0.020
0.3	0.052	0.012	0.231	0.026	0.962	0.009
			ACF			
0	0.745	0.040	0.996	0.004	0.996	0.004
0.1	0.593	0.041	0.787	0.023	0.997	0.003
0.3	0.299	0.056	0.395	0.049	0.996	0.033

Note: B=1000. The NP estimates of production function parameters are firm-year specific with an unknown distribution. For each sample $b = 1, ..., B$, we report the median total cost shares of labor, capital and latent input as, respectively, α^ℓ,α^k and CSLI. We use the ACF procedure as provided in the '*prodest*' package in R. We set as starting values for the ACF procedure the true output elasticities augmented with normally distributed noise with standard deviation 0.1.

Overall, we believe that these simulation results show the attractivity of our NP method. In particular, for a data generating process that perfectly fits the ACF method with a Cobb-Douglas production technology, we demonstrate that the NP method produces comparable results when the empirical analyst specifies the goodness-of-fit parameter θ adequately given the level of noise in the data. Given that the NP method does not rely on functional specifications of the production technology or well-chosen starting values, we can reasonably expect that the results of such a comparison may turn out even more favorable for our methodology in a setting generating data that are less perfect for the ACF method.[11] As we expect the production data in our following empirical application to be moderately noisy, we will set $\theta = 0.95$. However, our main qualitative conclusions are found to be robust for alternative θ-values.

3 Application set-up and data

We demonstrate the empirical usefulness of our novel nonparametric method by applying it to production data drawn from the Central Balance Sheet Office database, which provides annual information on the financial accounts of Belgian firms. We link this database with firm-year level international trade data of the National Bank of Belgium to include export information (dummies per export region) and import information (dummies and shares per import region) into our analysis.[12]

Before describing our empirical application in more detail, we remark that our dataset shares the characteristics and limitations of many large-scale datasets that have been used in other productivity analyses based on recently developed production function estimators (see, for example, Olley and Pakes (1996); Levinsohn and Petrin (2003); Wooldridge (2009); Ackerberg et al. (2015); Gandhi et al. (2017)). We pool single-product and multi-product firms, and we use industry-wide deflators to approximate firm-level prices. This implies that our measure of productivity (in terms of latent input) does not only include the pure technological features of the firm (for example, innovation, intangibles and managerial quality), but also potential influences from firm-level price setting behavior in the output market (Klette and Griliches, 1996; Foster et al., 2008; De Loecker, 2011; De Loecker and Warzynski, 2012), differences in accounting practices, and/or differences (e.g., across products) in production structures (Diewert, 1973; Panzar and Willig, 1981; Bernard et al., 2010, 2011; De Loecker, 2011; Dhyne et al., 2014; De Loecker et al., 2016).

For our main analysis, we include as output the deflated revenue and as inputs the number of employees in full time equivalents (FTE), deflated tangible fixed assets and deflated (domestic and foreign) materials use (i.e., raw materials, consumables, services and other goods). For the input prices, we use the price of labor, and the nace 2-digit deflators of intermediary inputs and tangible fixed assets.[13] The firm-year level price of labor is obtained from dividing labor cost by labor numbers in full time equivalents. We estimate the unobserved heterogeneity/productivity in manufacturing production at the firm-year

[11]See, for example, Mollisi and Rovigatti (Forthcoming) for a discussion on the sensitivity to starting values of the ACF method.

[12]Import shares have been computed by the National Bank of Belgium at the firm (and group of countries) level by merging data on import from the Transaction Trade dataset and data on material inputs purchases from the VAT database. No distinction is made between final and intermediate products in either database. See, for example, Mion and Zhu (2013) for a detailed discussion.

[13]Deflators are based on EU KLEMS and measured as described in Merlevede et al. (2015, p.8).

level for the eight largest nace (rev.1.1) 2-digit sectors for the time horizon of 1997 to 2007 (see Table 2 for more details about our dataset). We thus restrict the sample to before the 2008 financial crisis.[14] As explained in Section 2.1, we (only) assume that the CRS assumption holds "locally" (i.e., for the given firm size), so avoiding the more debatable "global" CRS postulate. For that reason, we split up the sample according to firm size: small firms (Labor in FTE from 10 to 50; 10,680 observations), medium firms (Labor in FTE from 50 to 250; 8,505 observations), and large firms (Labor in FTE larger than 250; 2,365 observations). We conduct our nonparametric analysis for each firm size group separately.

Table 2: Included sectors

Nace rev.1.1. sector	Obs.	Firms
Nace 15: Manufacture of food products and beverages	4,480	755
Nace 17: Manufacture of textiles, manufacture of articles of straw and plaiting materials	2,326	421
Nace 22: Publishing, printing and reproduction of recorded media	1,854	390
Nace 24: Manufacture of chemicals and chemical products	2,611	426
Nace 25: Manufacture of rubber and plastic products	1,992	337
Nace26: Manufacture of other non-metallic mineral products	2,176	370
Nace 28: Manufacture of fabricated metal products, except machinery and equipment	3,769	808
Nace 29: Manufacture of machinery and equipment n.e.c.	2,342	454
Total	21,550	3,875

Revenue as included in balance sheets not only involves in-house production of manufacturing goods, but often also includes servicing (see, for example, Pilat et al. (2006) for a policy-oriented discussion) and reselling of products that are not produced by the firm.[15] As these decisions are closely related to any make-or-source decision, we will verify whether our empirical results are robust for altering the definition of firm output to the deflated sales of produced manufacturing goods by the firm. To this end, we make use of the subsample of Belgian firms that participate to the Prodcom survey of Eurostat, which allows us to use deflated produced value as output.[16] A main motivation of Eurostat to initiate the Prodcom survey was exactly to obtain comparable statistics on manufacturing at the product level across the European Union. Participation to the Belgian Prodcom survey is

[14]To avoid extreme outliers, we limit our sample to observations of firms with at least ten employees. We changed the flows to a number of months in a book year equal to 12 and removed observations with book periods shorter than 6 months and longer than 24 months. We removed the highest and lowest percentiles of the growth rates, at the sector-year level, for the output, observed inputs, the price of labor and the share of materials in observed costs. We also removed clear erroneous reporting by limiting the sample to input-output observations with values over 1,000 euro and labor price with values over 10,000 euro. Smaller firms (either having on average less than 100 employees during the year or not exceeding two of the following three criteria: annual average of 50 employees, annual revenue of 7,300,000 euro or a balance-sheet total of 3,650,000 euro) can report their annual accounts using an abbreviated model with the possibility of no separation between gross revenue and input use. These smaller firms have a higher probability to be excluded from the analysis due to missing values.

[15] Bernard et al. (Forthcoming) document widespread exportation of manufacturing products that are not produced by the firm and label this carry-along trade (CAT). They show that CAT relates positively with productivity.

[16]We cleaned our production data by using the same criteria as for the main analysis.

mandatory for the firms that operate above a given threshold of operation size.[17] Recent studies that make use of the Belgian Prodcom database include De Loecker et al. (2014), Dhyne et al. (2014), Forlani et al. (2016) and Bernard et al. (Forthcoming).

Interestingly, by using this Prodcom database we can also show the applicability of our approach to estimate pure technological heterogeneity among single-product firms, by using output quantity data for the tightly defined sector of ready mixed concrete producers (8 digit-level product, used in numerous studies, including Syverson (2004) and Foster et al. (2008)). Syverson (2004) argues that proxy variable approaches such as the Olley and Pakes (1996) routine are not appropriate for this specific sector, as local demand states may influence input and investment decisions, which makes the assumption of a one-to-one relation between unobserved productivity and observable investment difficult to maintain. Because our routine does not rely on (semi-)parametric structuring of the simultaneity issue, it remains well applicable to such sectors that fall beyond the reach of proxy variable approaches. In Appendix B, we show that our results based on the very small sample of Belgian ready mixed concrete producers (using quantity based, revenue based and produced value based estimations) largely confirm our main conclusions on the evolution of cost shares over time.

Measured productivity differences usually relate to firm-level heterogeneity in observable characteristics (Syverson, 2011). Included firm characteristics that are expected to relate to our nonparametric measure of productivity (NP) are firm size (Haltiwanger et al., 1999; Van Biesebroeck, 2005; Forlani et al., 2016), international exposure (Bernard and Jensen, 1995; Bernard et al., 2003, 2010, Forthcoming; Melitz, 2003; Antras and Helpman, 2004; Egger et al., 2015), firm age (Wagner, 1994), and firm entry and exit (Olley and Pakes, 1996; Melitz and Polanec, 2015). Table 3 shows some descriptive statistics on these variables for the eight sectors that we consider, comprising 21,550 observations of 3,875 firms.

As Belgium is a small open economy, international exposure is usually high. In our sample of manufacturing firms, only 12 percent of the firm observations shows no exporting behavior, and 68 percent exports to non-EU countries. Export to distant countries is thus the rule rather than the exception. Production processes are generally disintegrated, with the average share of materials in observed costs amounting to 64 percent. We label this material share in observed costs as *sourcing* (see, for example, Arvantis and Loukis (2013) for a review of empirical studies that use material shares as proxies for outsourcing). The vast majority (94 percent) of observations indicate to import intermediary inputs, yet the domestic component of disintegrated activities is 2.28 times the foreign component. While 66 percent import from outside the EU, the average share of materials from outside the EU in observed costs equals only 3 percent. 16 percent of the sampled firms source from China, and this percentage is increasing over time (descriptive statistics available upon request; see also Mion and Zhu (2013) for a detailed analysis). Finally, we proxy 1 percent

[17] For our considered time period the threshold was 10 employees and a specific revenue threshold in a given year.

of our observed firms as entering firms, 6 percent as starting firms (i.e., firm age at most 5), 13 percent as young firms (i.e., firm age between 5 and 10), 80 percent as mature firms (i.e., firm age higher than 10), and 1 percent as exiting firms.[18]

Table 3: Summary statistics

	Mean	St.Dev.	Min.	25%	Med.	75%	Max.
Deflated revenue (output)	37.56	125.45	0.09	5.34	10.89	28.35	4584.31
Deflated produced value (Prodcom-based; 18,757 obs.)	29.54	107.72	0.01	4.52	9.57	23.68	4488.67
Output price	1.06	0.09	0.89	1.00	1.04	1.08	1.35
Labor in FTE	127.22	282.60	10.00	28.20	50.60	113.00	5686.00
Deflated tangible fixed assets	5.89	22.79	0.00	0.61	1.54	4.00	664.20
Deflated material costs	28.30	96.49	0.00	3.41	7.50	20.43	3492.81
Labor price	0.04	0.01	0.01	0.03	0.04	0.05	0.18
Capital price	1.12	0.08	1.01	1.07	1.11	1.16	1.38
Intermediates price	1.08	0.09	0.95	1.02	1.06	1.12	1.36
Exporting (dummy)	0.88	0.33	0.00	1.00	1.00	1.00	1.00
Exporting to Eastern Europe (dummy)	0.49	0.50	0.00	0.00	0.00	1.00	1.00
Exporting outside the EU (dummy)	0.68	0.47	0.00	0.00	1.00	1.00	1.00
Foreign sourcing (dummy)	0.94	0.24	0.00	1.00	1.00	1.00	1.00
Sourcing from outside the EU (dummy)	0.66	0.47	0.00	0.00	1.00	1.00	1.00
Sourcing from Eastern Europe (dummy)	0.30	0.46	0.00	0.00	0.00	1.00	1.00
Sourcing from China (dummy)	0.16	0.37	0.00	0.00	0.00	0.00	1.00
Sourcing (share)	0.64	0.16	0.00	0.54	0.66	0.76	0.99
Domestic sourcing (share)	0.45	0.17	0.00	0.32	0.43	0.56	0.98
Starting	0.06	0.24	0.00	0.00	0.00	0.00	1.00
Young	0.13	0.33	0.00	0.00	0.00	0.00	1.00
Mature	0.80	0.40	0.00	1.00	1.00	1.00	1.00
Entry	0.01	0.11	0.00	0.00	0.00	0.00	1.00
Exiting	0.01	0.10	0.00	0.00	0.00	0.00	1.00

Note: Deflated revenue, deflated produced value, deflated tangible fixed assets, deflated material costs and labor price are expressed in millions of euro. Eastern Europe countries: Bulgaria, Czech Republic, Cyprus, Estonia, Croatia, Hungary, Lithuania, Latvia, Malta, Poland, Romania, Slovenia, Slovakia.

4 Empirical results

In this section, we first present some descriptive statistics on our nonparametric estimates of productivity and the cost share of latent input (NP and CSLI).[19] Next, we relate our nonparametric productivity estimates to observable firm characteristics. This will demonstrate that our estimates effectively replicate stylized findings in the literature. We conclude by analyzing the evolution of cost shares (of observed inputs and latent input) over time. In particular, we assess to what extent observed primary manufacturing inputs are substituted for other observed inputs and/or unobserved technology. Our methodology allows us to address this question in a fully nonparametric fashion, without imposing a

[18] A firm is considered to enter in the first year for which employment is strictly positive, provided that the firm is not older than five years (based on its year of incorporation). Next, a firm is considered to exit in the year for which employment is no longer reported after previous year(s) with strictly positive employment, insofar the number of years to the declared exit date does not exceed five.

[19] Throughout the paper we express latent input in millions of euro. We excluded one observation ex post with a cost share of latent input above 0.999.

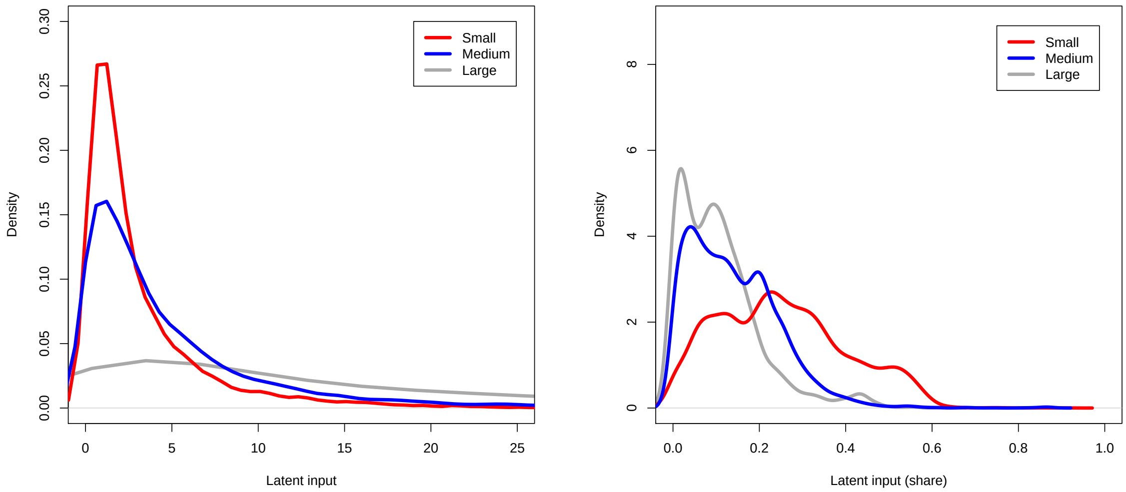
priori assumptions of Hicks neutrality or any other functional structure for the unknown production technology.

Two remarks are in order before discussing our results. First, our main analysis will be on the aggregate of all eight nace 2-digit sectors for which we solved a linear program with objective function (5), given the constraints as formulated in (2). Evidently, each sector has its own particularities (related to input use and output production), but our principal findings turn out to be robust across sectors (sector-specific results, descriptives and figures available upon request).

Second, we present two additional robustness checks in Appendices B and C. As motivated above, in Appendix B we demonstrate the possible application of our method to a sample of single-product producers (of ready mixed concrete). Next, as discussed in Section 2.4, in our empirical analysis we will use the goodness-of-fit parameter $\theta = 0.95$ to account for (small) deviations of observed firm behavior from exact rationalizability (i.e., data consistency with the strict cost minimization conditions in Proposition 1). In Appendix C, we show that our main conclusions are robust for alternative specifications of the θ-parameter.

4.1 Productivity estimates: a first look

Figure 1 depicts the distributions of our productivity (NP) estimates and cost shares of latent input (CSLI) estimates (see Table 13 in Appendix D.1 for additional descriptives). We clearly observe that accounting for technological heterogeneity is required to rationalize the observed firm behavior in terms of our cost minimization hypothesis. This provides strong nonparametric evidence against any framework that is based on a representative firm and a sector aggregate production function.



(a) NP (b) CSLI

Figure 1: Productivity estimates and firm size

Larger firms are generally characterized by higher productivity levels, which falls in line with a common finding in the literature. However, this does not mean that larger firms also have higher cost shares of latent input. For example, the cost share of latent input is above 0.24 for half of the small firms, whereas it exceeds the same cut-off level for less than ten percent of the large firms. On average, latent input accounts for approximately 25, 14 and 11 percent of the total costs of small, medium and large firms, respectively. Finally, we observe that smaller firms generally show more variation in their cost shares of latent input. This indicates that smaller firms are not only more heterogeneous in terms of their observable characteristics (summary statistics available upon request), but also in terms of their unobservable input.

Without further information, we cannot directly disentangle whether these differences across firm sizes are effectively driven by actual differences in the intra-group distributions of latent inputs or, rather, by inter-group differences in the precision of measurement of the observable characteristics. Therefore, in what follows we will analyze our productivity estimates for each firm group separately, and largely abstain from making statements that compare firm size groups. We assume that, within a given firm size group, there are no systematic differences in the precision of measurement of the observable characteristics.

In Table 4, we report correlation results that further validate our nonparametric measure of productivity NP. First, we find that NP relates strongly and positively to labor productivity as measured by dividing deflated revenues by the number of employees in FTE. The Spearman correlation increases with firm size: it equals 0.44 for small firms, and it amounts to 0.56 for large firms. Second, the correlations with a one-year lag of NP are also large and positive for the different firm size groups. The Spearman correlation over all firms equals 0.89 and is above 0.86 for all firm size groups. We conclude that our NP estimates robustly confirm the documented stylized fact of huge and persistent differences across producers in terms of measured productivity (see, for example, Syverson (2011)).

Table 4: Spearman correlations

	All	Small	Medium	Large
Labor productivity	0.46	0.44	0.52	0.56
Lagged NP	0.89	0.91	0.86	0.89

4.2 Productivity, sourcing and international exposure

As an additional validation of our interpretation of NP as representing unobserved firm productivity, we next study the relation between our nonparametric measure of productivity, sourcing (i.e., total amount of domestic and foreign material inputs) and international exposure. As argued above, the empirical and theoretical literature shows a generally positive correlation between productivity, international exposure and foreign outsourcing. This not only reveals a direct impact of internationalization on productive efficiency, it is also related to quality differences between intermediates of different origin, and to differences in the variety of intermediates used together with a taste for variety in the production process (see, for example, Goldberg et al. (2009, 2010) and Halpern et al. (2015)). Next, the literature on export behavior of firms documents a positive correlation between measured productivity and export as a stylized fact.

The left hand side of Table 5 shows the relation between (logged) NP, sourcing and international exposure within the three firm size groups. Following our discussion in Section 3, in all regressions we include nace 2-digit and year fixed effects as well as dummies controlling for firm age (starting, young, mature), entry and exit.[20] In the specific case of Belgian manufacturers, sourcing almost always implies some sort of international exposure (94 percent of the sampled firms use foreign sourcing). Thus, we can expect multicollinearity to impede disentangling the effects of foreign and domestic sourcing.

Our regression results support the widespread findings from the productivity literature. Overall, we observe a significantly positive relationship between productivity and international exposure for all firm size groups. More specifically, for small firms this significant positive relationship applies to both foreign sourcing and exporting, with the correlation being higher when sourcing is from outside the EU and exporting is to Eastern Europe. Medium firms show a significant positive relation between NP and sourcing from outside

[20]Foster et al. (2008) find that firm age, entry and exit relate to idiosyncratic demand shocks and firm-specific output prices. Therefore, we include these variables as control variables to mitigate confounding influences. Some caution is needed when interpreting our results on export, as exporting is known to imply product-specific pricing.

the EU and exporting outside the EU. Large firms show a significantly positive relation with sourcing from Eastern Europe. Further, our regressions also reveal a significantly positive correlation between NP and the share of sourcing for the three firm size groups. The significant negative relations for exporting outside the EU (small firms), exporting (medium firms) and exporting to Eastern Europe (large firms) are not robust for altering the output definition (see below). In sum, our NP estimates confirm that, more disintegrated, international production processes are positively associated with measured productivity.

Next, as discussed in Section 3, our main analysis considers (deflated) revenue of the firm as output (i.e., estimates are revenue based), pooling together multiple products, but also servicing and carry-along trade (see Bernard et al. (Forthcoming)). To verify whether our results are robust for influences of servicing and resale of out-house production, we redefined firm output as deflated sales of produced goods (reported in the Prodcom database). Summary statistics are provided in Table 13 in Appendix D.1. As for the connection between productivity, sourcing and international exposure, the right hand side of Table 5 confirms the positive relationship that we found before. In fact, when using in-house production to measure output (yielding produced value based estimates of NP and CSLI), for all firm size groups we find a significant positive relation between NP and both the foreign sourcing and exporting aspect of internationalization. Overall, we may safely conclude that our principal qualitative conclusions are largely robust to the chosen output definition.

Table 5: Productivity and international exposure: a truncated regression analysis

	Revenue based			**Produced value based**		
	Small	Medium	Large	Small	Medium	Large
Sourcing (share)	2.364***	4.088***	3.378***	3.692***	3.921***	4.224***
	(0.152)	(0.225)	(0.791)	(0.212)	(0.273)	(0.724)
Foreign sourcing (dummy)	0.465***	0.0162	0.0500	0.422***	0.0105	0.925***
	(0.0581)	(0.127)	(0.491)	(0.0707)	(0.195)	(0.348)
Sourcing from Eastern Europe (dummy)	0.0526	0.0384	0.372***	0.0789	0.0525	0.0474
	(0.0372)	(0.0463)	(0.134)	(0.0580)	(0.0518)	(0.135)
Sourcing from outside the EU (dummy)	0.0820**	0.301***	0.410	0.0190	0.167***	0.308
	(0.0371)	(0.0557)	(0.269)	(0.0501)	(0.0601)	(0.235)
Sourcing from China (dummy)	0.0231	0.00631	0.203	-0.0282	0.0700	0.193
	(0.0458)	(0.0578)	(0.141)	(0.0915)	(0.0657)	(0.155)
Exporting (dummy)	0.230***	-0.355***	-0.613	0.279***	-0.125	-0.166
	(0.0502)	(0.103)	(0.417)	(0.0742)	(0.122)	(0.333)
Exporting to Eastern Europe (dummy)	0.196***	0.107*	-0.435**	0.218***	0.116*	-0.130
	(0.0362)	(0.0559)	(0.187)	(0.0523)	(0.0647)	(0.186)
Exporting outside the EU (dummy)	-0.0930**	0.135**	0.414*	0.0130	0.206***	0.540**
	(0.0408)	(0.0645)	(0.227)	(0.0537)	(0.0783)	(0.240)
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Nace 2-digit effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-truncated Observations	10,567	8,324	2,037	8,201	7,172	1,897
Observations	10,679	8,505	2,365	8,397	7,417	2,006
Firms	2,591	1,445	349	2,031	1,279	316

Note: The dependent variable is the log of NP. Marginal effects of left-truncated regressions shown. Robust standard errors in parentheses with clustering at the firm level.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.3 Substitution between latent input and observed inputs

As motivated above, a main distinguishing feature of our methodology to identify unobserved productivity heterogeneity is that it deals with the simultaneity bias while naturally relaxing the Hicks neutrality assumption. We do not need to assume that input cost shares are constant over time, and we do not have to impose a common structure on factor biased technological change. As a final investigation, we exploit this unique aspect by considering variation in (observable and latent) input substitution patterns over time. Table 6 shows a fixed effects regression wherein we regress total cost shares of inputs on the variable '*year*'. Descriptive statistics on the evolution of cost shares over time are given in Tables 14 and 15 in Appendix D.2.

A first observation is that primary inputs gradually lose ground in large firms. Over a decade, labor loses over 1 percentage point (i.e., over 7 percent when using base year 1997) in total cost share, even though the relative price of labor is non-decreasing relative to the price of the other observed inputs (summary statistics available upon request). This confirms the well-documented loss in labor shares (OECD, 2012), now explicitly taking into account productivity differences.

Further, Figure Table 6 reveals that the total cost share of tangible fixed assets (TFA) is decreasing rather than increasing. The average TFA total cost share goes down by over 2 percentage point in a decade, corresponding to a decrease of over 14 percent. This pattern is seen for all firm size groups and is robust for using output based on produced value. It is still 14 percent in 1997, but goes down to only 12 percent in 2007. Stated differently, our within-industry estimates provide no empirical support for the argument that technological change was detrimental for labor and favorable for TFA. We find that both primary inputs are substituted for other inputs in the Belgian manufacturing sector.

By contrast, for large firms we do observe steadily increasing cost shares of materials, resulting in an increase of over 2 percentage points over a decade. Descriptive statistics show even that material total cost shares have gone up by 4 percentage points between 1997 and 2007 (i.e., from 0.58 to 0.62). This comprises an increase in both domestic and foreign materials of respectively 1 and 3 percentage points. The cost share of latent input (CSLI) remains constant over the time horizon under investigation, supporting the idea of a productivity stagnation in the manufacturing sector.

Taken together, both revenue based and produced value based regressions in Table 6 and descriptives in Table 14 and Table 15 in Appendix D.2 suggest that primary inputs are substituted for more use of materials (i.e., increased prevalence of both domestic and international disintegration) rather than for latent input (i.e., technology). This evolution is gradual as can be seen in Figure 4 in Appendix D.2. This confirms that production processes have become less integrated within firms and more international, while being characterized by a productivity stagnation.

Table 6: Fixed effects regressions of total cost shares on year

	Revenue based			Produced value based		
	Small	Medium	Large	Small	Medium	Large
Dependent variable: Total cost share of labor (times 100)						
Year	0.0420**	0.0321	-0.102**	0.0342	-0.00598	-0.140***
	(0.0194)	(0.0259)	(0.0446)	(0.0243)	(0.0253)	(0.0401)
Constant	-70.98*	-46.86	222.0**	-53.56	29.24	297.1***
	(38.91)	(51.88)	(89.19)	(48.62)	(50.65)	(80.22)
Dependent variable: Total cost share of TFA (times 100)						
Year	-0.272***	-0.224***	-0.202***	-0.255***	-0.215***	-0.212***
	(0.0318)	(0.0404)	(0.0727)	(0.0406)	(0.0418)	(0.0726)
Constant	557.2***	463.0***	419.4***	524.6***	443.2***	436.5***
	(63.67)	(80.91)	(145.6)	(81.18)	(83.73)	(145.3)
Dependent variable: Total cost share of domestic materials (times 100)						
Year	0.125***	0.0671	0.207**	0.169***	0.0618	0.206**
	(0.0448)	(0.0491)	(0.0806)	(0.0530)	(0.0497)	(0.0993)
Constant	-215.8**	-97.41	-376.9**	-300.1***	-88.39	-378.5*
	(89.58)	(98.39)	(161.3)	(106.1)	(99.47)	(198.7)
Dependent variable: Total cost share of foreign materials (times 100)						
Year	0.0979***	0.141***	0.0365	0.0750*	0.142***	0.0999
	(0.0378)	(0.0475)	(0.0785)	(0.0433)	(0.0438)	(0.0867)
Constant	-182.2**	-264.5***	-53.10	-135.5	-266.4***	-181.3
	(75.73)	(95.13)	(157.1)	(86.64)	(87.63)	(173.5)
Dependent variable: Total cost share of latent input (times 100)						
Year	0.00676	-0.0159	0.0610	-0.0232	0.0173	0.0453
	(0.0384)	(0.0408)	(0.0800)	(0.0610)	(0.0605)	(0.167)
Constant	11.70	45.82	-111.4	64.53	-17.72	-73.89
	(76.85)	(81.68)	(160.2)	(122.2)	(121.1)	(333.6)
Observations	10,679	8,505	2,365	8,397	7,417	2,006
Number of firms	2,591	1,445	349	2,031	1,279	316

Note: The dependent variable is respectively the total cost share of labor, TFA, domestic materials, foreign materials, latent input times 100. All regressions use firm-level fixed effects. Robust standard errors in parentheses.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5 Conclusion

We have developed a novel structural method for production analysis that recovers unobserved productivity in a fully nonparametric fashion. We model unobserved heterogeneity as an unobserved productivity factor on which we condition the demand of the observed inputs. Our method deals with the simultaneity bias in a natural way, and it empirically quantifies productivity differences across firms in terms of differences in latent input. Our nonparametric methodology is easy to implement as it merely requires the use of linear programming techniques. It allows for a powerful identification analysis, while avoiding (nonverifiable and often debatable) assumptions of functional form regarding the relationship between inputs and outputs (including the hypothesis of Hicks neutral technical change).

Our empirical application and a Monte Carlo analysis has shown that the method does allow for drawing strong empirical conclusions, despite its nonparametric nature. For a set of Belgian manufacturing firms, we have recovered productivity differences at the firm-year level over the period 1997-2007 for broad industry categories. Consistent with the well-established literature on international trade, we find that disintegrated firms with international sourcing are more productive. Further, we find that primary inputs (labor and tangible fixed assets) are substituted over time for (domestic and foreign) outsourcing, but usually not for greater use of technology. For large firms, this substitution is more pronounced. Overall, we provide robust empirical evidence against the assumption of Hicks neutrality for the setting at hand.

Our empirical application mainly focused on showing the usefulness of our methodology for a standard empirical production setting with output expressed in revenue terms (i.e., prices times quantities). Importantly, however, the applicability of our methodology is not merely restricted to such observational settings. As we demonstrated for the sector of ready mixed concrete producers, our method is also directly applicable when output is expressed in quantity terms. In this respect, if output price information is equally available, we may straightforwardly adapt to our setting the approach of De Loecker and Warzynski (2012) to identify markups from the available production information. A distinguishing and –in our view– particularly attractive feature of our methodology is that we can do so while abstaining from imposing any parametric structure on the production technology.[21] This suggests0. our methodology as a promising tool for empirically addressing the many questions on market power that have taken a prominent position in the empirical literature on firm behavior.

Finally, from a methodological point of view, we emphasize that we see the current paper primarily as providing a fruitful starting ground, rather than a complete toolkit for non-parametric production analysis with unobserved productivity differences. Most notably, we have focused on a single-output setting throughout. As discussed in De Loecker et al. (2016), a multiproduct framework (also involving the identification of input allocations across products) is warranted to obtain a more detailed insight into influences of exogenous trade or cost shocks. To develop this multi-output version of our methodology, a useful starting point is the study of Cherchye et al. (2014), who presented a nonparametric framework (abstracting from the input choice dependency on productivity) for the analysis of firms producing multiple products. A closely related issue concerns dealing with non-competitive output markets. In this respect, Carvajal et al. (2013, 2014) show how to analyze alternative (for example, Cournot or Bertrand) structures on output markets in the advocated nonparametric framework. In our opinion, integrating these authors' insights with our newly developed methodology may constitute another fertile avenue for

[21] Technically, under our CRS assumption we can obtain output elasticities as the nonparametrically identified input cost shares (including both observed and unobserved/latent costs). In turn, this allows us to define marginal costs of production from the first order conditions under cost minimizing behavior, which directly generates the production mark-ups.

follow-up research.

References

Acemoglu, D., Gancia, G., Zilibotti, F., 2015. Offshoring and Directed Technical Change. American Economic Journal: Macroeconomics 7 (3), 84–122.

Ackerberg, D., Caves, K., Frazer, G., 2015. Identification properties of recent production function estimators. Econometrica 83 (6), 2411–2451.

Afriat, S., 1972. Efficiency estimation of production functions. International Economic Review 13 (3), 568–598.

Antras, P., Helpman, E., 2004. Global Sourcing. Journal of Political Economy 112 (3), 552–580.

Arvantis, S., Loukis, E., 2013. Outsourcing and firm performance - a comparative study of Swiss and Greek firms. Industrial and Corporate Change 22 (3), 771–806.

Autor, D. H., Dorn, D., Katz, L. F., Patterson, C., Van Reenen, J., 2017. The Fall of the Labor Share and the Rise of Superstar Firms. NBER Working Paper 233396.

Autor, D. H., Katz, L., Krueger, A., 1998. Computing inequality: Have computers changed the labor market? The Quarterly Journal of Economics 113 (4), 1169–1213.

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

Bernard, A., Blanchard, E., Van Beveren, I., Vandenbussche, H., Forthcoming. Carry-along trade. Review of Economic Studies.

Bernard, A., Eaton, J., Jensen, J., Kortum, S., 2003. Plants and productivity in international trade. American Economic Review 93 (4), 1268–1290.

Bernard, A., Jensen, B., 1995. Exporters, Jobs, and Wages in US Manufacturing: 1976-87. Brookings Papers on Economic Activity: Microeconomics, 67–112.

Bernard, A., Redding, S., Schott, P., 2010. Multi-product firms and product switching. American Economic Review 100 (1), 70–97.

Bernard, A., Redding, S., Schott, P., 2011. Multiproduct firms and trade liberalization. The Quarterly Journal of Economics 126 (3), 1271–1318.

Carvajal, A., Deb, R., Fenske, J., Quah, K.-H., 2013. Revealed preference tests of the cournot model. Econometrica 81 (6), 2351–2379.

Carvajal, A., Deb, R., Fenske, J., Quah, K.-H., 2014. A Nonparametric Analysis of Multi-Product Oligopolies. Economic Theory 57, 253–277.

Cherchye, L., Demuynck, T., De Rock, B., De Witte, K., 2014. Nonparametric analysis of multioutput production with joint inputs. Economic Journal 124 (577), 735–775.

CompNet Task Force, 2014. Micro-based evidence of EU competitiveness: The CompNet database. ECB Working Paper Series 1634.

Dall'Olio, A., Iootty, M., Kanehira, N., Saliola, F., 2013. Productivity growth in Europe. Policy Research Working Paper 6425, Word Bank.

De Loecker, J., 2011. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. Econometrica 79 (5), 1407–1451.

De Loecker, J., Fuss, C., Van Biesebroeck, J., 2014. International competition and firm performance : Evidence from Belgium. Tech. Rep. 269, National Bank of Belgium Working Paper Series.

De Loecker, J., Goldberg, P., Khandelwal, A., Pavcnik, N., 2016. Prices, markups, and trade reform. Econometrica 84 (2), 445–510.

De Loecker, J., Warzynski, F., 2012. Markups and firm-level export status. American Economic Review 102 (6), 2437–2471.

Dewitte, R., Dumont, M., Merlevede, B., Rayp, G., Verschelde, M., 2017. Firm-heterogeneous factor-biased technological change, mimeo.

Dhyne, E., Petrin, A., Smeets, V., Warzynski, F., et al., 2014. Import competition, productivity and multi-product firms. Working Paper 268, National Bank of Belgium.

Diewert, W., 1973. Functional forms for profit and transformation functions. Journal of Economic Theory 6 (3), 284–316.

Diewert, W., Parkan, C., 1983. Linear programming tests of regularity conditions for production frontiers. In: Eichhorn, W., Henn, R., Neumann, K., Shephard, R. (Eds.), Quantitative Studies on Production and Prices. Physica-Verlag, Würzburg.

Doraszelski, U., Jaumandreu, J., 2018. Measuring the bias of technological change. Journal of Political Economy 126, 1027–1084.

Egger, H., Kreickemeier, U., Wrona, J., 2015. Offshoring domestic jobs. Journal of International Economics 97 (1), 112–125.

Epstein, L., Yatchew, A., 1985. Non-parametric hypothesis testing procedures and applications to demand analysis. Journal of Econometrics 30, 149–169.

Forlani, E., Martin, R., Mion, G., Muûls, M., 2016. Unraveling firms: Demand, productivity and markups heterogeneity. Tech. rep., National Bank og Belgium Working Paper Series, No. 293.

Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review 98 (1), 394–425.

Gandhi, A., Navarro, S., Rivers, D., 2017. On the Identification of Production Functions: How Heterogeneous is Productivity? Journal of Political Economy Conditionally Accepted.

Goldberg, P., Khandelwal, A., Pavcnik, N., Topalova, P., 2009. Trade liberalization and new imported inputs. American Economic Review 99 (2), 494–500.

Goldberg, P., Khandelwal, A., Pavcnik, N., Topalova, P., 2010. Imported intermediate inputs and domestic product growth: evidence from indianomics. The Quarterly Journal of Economics 125 (4), 1727–1767.

Grifell-Tatjé, E., Lovell, C., Sickles, R. (Eds.), 2018. The Oxford Handbook of Productivity Analysis. Oxford University Press.

Griliches, Z., Mairesse, J., 1998. Production functions: The search for identification. In: Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium. New York: Cambridge University Press, pp. 169–203.

Halpern, L., Koren, M., Szeidl, A., 2015. Imported inputs and productivity. American Economic Review 105 (12), 3660–3703.

Haltiwanger, J., Lane, J., Spletzer, J., 1999. Productivity differences across employers: The roles of employer size, age, and human capital. American Economic Review 89 (2), 94–98.

Hanoch, G., Rothschild, M., 1972. Testing assumptions of production theory: A nonparametric approach. Journal of Political Economy 80 (2), 256–275.

Klette, T. J., Griliches, Z., 1996. The inconsistency of common scale estimators when output prices are unobserved and endogenous. Journal of Applied Econometrics 11 (4), 343–361.

Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables. Review of Economic Studies 70 (2), 317–341.

Marschak, J., Andrews, W., 1944. Random simultaneous equations and the theory of production. Econometrica 12 (3-4), 143–205.

Melitz, M., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

Melitz, M., Polanec, S., 2015. Dynamic Olley-Pakes productivity decomposition with entry and exit. RAND Cooperation 46 (2), 362–375.

Merlevede, B., De Zwaan, M., Lenaerts, K., Purice, V., 2015. Multinational networks, domestic, and foreign firms in Europe. Ghent University Working Paper Series D/2015/7012/02.

Mion, G., Zhu, L., 2013. Import competition from and offshoring to china: A curse or blessing for firms? Journal of International Economics 89, 202–215.

Mollisi, V., Rovigatti, G., Forthcoming. Theory and practice of TFP estimation: The control function approach using Stata. The STATA Journal.

OECD, 2012. Employment Outlook 2012. Tech. rep., Organisation of Economic Cooperation and Development, Paris.

Olley, G. S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64 (6), 1263–1297.

Panzar, J., Willig, R., 1981. Economies of Scope. American Economic Review 71 (2), 268–272.

Pilat, D., Cimper, A., Olsen, K., Webb, C., 2006. The changing nature of manufacturing in oecd economies. Tech. rep., OECD Science, Technology and Industry Working Paper 2006 no. 9, Paris.

Syverson, C., 2004. Market structure and productivity: A concrete example. Journal of Political Economy 112 (6), 1181–1222.

Syverson, C., 2011. What determines productivity? Journal of Economic Literature 49 (2), 326–365.

Syverson, C., 2017. Challenges to mismeasurement explanations for the US productivity slowdown. Journal of Economic Perspectives 31 (2), 165–186.

Van Biesebroeck, J., 2005. Firm size matters: Growth and productivity growth in African manufacturing. Economic Development and Cultural Change 53 (3), 545–583.

Van Biesebroeck, J., 2007. Robustness of productivity estimates. Journal of Industrial Economics 55 (3), 529–569.

Varian, H., 1984. The nonparametric approach to production analysis. Econometrica 52 (3), 579–598.

Varian, H., 1985. Non-parametric analysis of optimizing behavior with measurement error. Journal of Econometrics 30 (1-2), 445–458.

Varian, H., 1988. Revealed preference with a subset of goods. Journal of Economic Theory 46 (1), 179–185.

Varian, H., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46 (1-2), 125–140.

Verschelde, M., Dumont, M., Merlevede, B., Rayp, G., 2014. A constrained nonparametric regression analysis of factor-biased technical change and TFP growth at the firm-level. Working Paper 266, National Bank of Belgium.

Wagner, J., 1994. The post-entry performance of new small firms in german manufacturing-industries. Journal of Industrial Economics 42 (2), 141–154.

Wooldridge, J. M., 2009. On estimating firm-level production functions using proxy variables to control for unobservables. Economics Letters 104 (3), 112–114.

Appendix A: Proofs and additional theoretical and simulation results

A.1 Proof of Proposition 1

Necessity of condition (iii). The CRS assumption implies that we can use Euler's theorem to obtain

$$\sum_m \frac{\partial F(\mathbf{X},\Omega)}{\partial X^m} X^m + \frac{\partial F(\mathbf{X},\Omega)}{\partial \Omega}\Omega = F(\mathbf{X},\Omega).$$

The first order conditions for the cost minimization problem (both OP.I and OP.II) imply

$$W_i^m = \lambda_i \frac{\partial F(\mathbf{X}_i,\Omega)}{\partial X^m},$$

with λ_i the associated Lagrange multipliers. From the concavity of F, we also have that,

$$F(\mathbf{X}_i,\Omega_i) - F(\mathbf{X}_j,\Omega_j) \leq \sum_m \frac{\partial F(\mathbf{X}_j,\Omega_j)}{\partial X^m}(X_i^m - X_j^m) + \frac{\partial F(\mathbf{X}_j,\Omega_j)}{\partial \Omega}(\Omega_i - \Omega_j).$$

Substituting $Q_i = F(\mathbf{X}_i,\Omega_i), Q_j = F(\mathbf{X}_j,\Omega_j)$ and using the above then gives

$$Q_i - Q_j \leq \frac{1}{\lambda_j}(\mathbf{W}_j\mathbf{X}_i + \Gamma_j\Omega_i) - Q_j,$$

where $\Gamma_j = \frac{\partial F(\mathbf{X}_j, \Omega_j)}{\partial \Omega}\lambda_j$. If $OP.I$ is used, then Γ_j is the shadow price of Ω_j, while if $OP.II$ is used then this equation follows from the first order conditions. Thus,

$$Q_i \leq \frac{1}{\lambda_j}(\mathbf{W}_j\mathbf{X}_i + \Gamma_j \Omega_i).$$

From the first order conditions (and definition of Γ_j) we also have that,

$$Q_j = \sum_m \frac{\partial F(\mathbf{X}_j, \Omega_j)}{X^m} X_j^m + \frac{\partial F(\mathbf{X}_j, \Omega_j)}{\Omega}\Omega = \frac{1}{\lambda_j}(\mathbf{W}_j\mathbf{X}_j + \Gamma_j \Omega_j).$$

Thus,

$$\frac{Q_i}{Q_j} \leq \frac{\mathbf{W}_j\mathbf{X}_i + \Gamma_j\Omega_i}{\mathbf{W}_j\mathbf{X}_j + \Gamma_j\Omega_j}.$$

Dividing numerator and denominator by Γ_j and defining $\Lambda_j = 1/\Gamma_j$ gives,

$$\frac{Q_i}{Q_j} \leq \frac{\Lambda_j\mathbf{W}_j\mathbf{X}_i + \Omega_i}{\Lambda_j\mathbf{W}_j\mathbf{X}_j + \Omega_j}.$$

Sufficiency of condition (iii). Assume that numbers Ω_i, Λ_i exist that satisfy the inequalities and define $F(\mathbf{X}, \Omega) = \min_i Q_i \frac{\Lambda_i\mathbf{W}_i\mathbf{X}+\Omega}{\Lambda_i\mathbf{W}_i\mathbf{X}_i+\Omega_i}$. It is easy to verify that his function is concave, homogeneous of degree one and continuous. Moreover the inequality conditions imply that $F(\mathbf{X}_i, \Omega_i) = Q_i$. To verify that $\mathbf{X}_i$ solves $OP.I$, assume, towards a contradiction that there is an input bundle $\mathbf{X}$ such that $\mathbf{W}_i\mathbf{X} < \mathbf{W}_i\mathbf{X}_i$ and $F(\mathbf{X}, \Omega_i) \geq Q_i$. Then, we have

$$\begin{aligned} Q_i &\leq F(\mathbf{X}, \Omega_i) \\ &\leq Q_i \frac{\Lambda_i\mathbf{W}_i\mathbf{X} + \Omega_i}{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i} \\ &< Q_i \frac{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i}{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i} \\ &= Q_i. \end{aligned}$$

Similarly, for $OP.II$ define $\Gamma_i = 1/\Lambda_i$. Now, if, towards a contradiction, there is an input bundle $(\mathbf{X}, \Omega)$ such that $\mathbf{W}_i\mathbf{X} + \Gamma_i\Omega < \mathbf{W}_i\mathbf{X}_i + \Gamma_i\Omega_i$ and $F(\mathbf{X}, \Omega) \geq Q_i$, then we have

$$\begin{aligned} Q_i &\leq F(\mathbf{X}, \Omega) \\ &\leq Q_i \frac{\Lambda_i\mathbf{W}_i\mathbf{X} + \Omega}{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i} \\ &< Q_i \frac{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i}{\Lambda_i\mathbf{W}_i\mathbf{X}_i + \Omega_i} \\ &= Q_i. \end{aligned}$$

A.2 Testability: a numerical example

The following example illustrates the testable implications in Proposition 1. It shows that these implications can be rejected even in a minimalistic setting with only two firm observations and two observed inputs.

Consider a dataset S with input prices $\mathbf{W}_1 = (1, 2)$ and $\mathbf{W}_2 = (2, 1)$ and input quantities $\mathbf{X}_1 = (1, 2)$ and $\mathbf{X}_2 = (2, 1)$. Proposition 1 requires

$$\frac{Q_1}{Q_2} \leq \frac{\Lambda_2 4 + \Omega_1}{\Lambda_2 5 + \Omega_2},$$
$$\frac{Q_2}{Q_1} \leq \frac{\Lambda_1 4 + \Omega_2}{\Lambda_1 5 + \Omega_1}.$$

Reformulating these inequalities obtains

$$(Q_1\Omega_2 - Q_2\Omega_1) \leq (4Q_2 - 5Q_1)\Lambda_2 \text{ and}$$
$$(Q_1\Omega_2 - Q_2\Omega_1) \geq (5Q_2 - 4Q_1)\Lambda_1,$$

which implies that $(5Q_2 - 4Q_1)\Lambda_1 \leq (4Q_2 - 5Q_1)\Lambda_2$. If we then assume that the (observed) output levels Q_1 and Q_2 are such that

$$\frac{4}{5} < \frac{Q_2}{Q_1} < \frac{5}{4},$$

we obtain that there can never exists strict positive Λ_1 and Λ_2 that satisfy this inequality restriction (since $4Q_2 - 5Q_1 < 0$ and $5Q_2 - 4Q_1 > 0$).

A.3 Goodness-of-fit parameter θ

We start from the rationalizability requirements in Proposition 1 and define $r_j = (\Lambda_j Q_j)/(\Lambda_j \mathbf{W}_j \mathbf{X}_j + \Omega_j)$, which allows us to rewrite the inequality restrictions as

$$Q_i - Q_j \leq r_j \left(\mathbf{W}_j(\mathbf{X}_i - \mathbf{X}_j) + \Gamma_j(\Omega_i - \Omega_j)\right).$$

We can weaken these requirements by equiproportionally contracting the inputs $(\mathbf{X}_j, \Omega_j)$, which corresponds to lowering the cost level $(\mathbf{W}_j \mathbf{X}_j + \Gamma_j \Omega_j)$ by the same degree. To do so, we use $\theta \leq 1$ and obtain

$$Q_i - Q_j \leq r_j \left(\mathbf{W}_j(\mathbf{X}_i - \theta\mathbf{X}_j) + \Gamma_j(\Omega_i - \theta\Omega_j)\right).$$

Generally, lower values of θ imply weaker rationalizability restrictions. Our optimization model provides a better (economic) fit of the dataset S if this set S satisfies the restrictions for a higher value of θ.

By using that $r_j = (\Lambda_j Q_j)/(\Lambda_j \mathbf{W}_j \mathbf{X}_j + \Omega_j)$, we can also include the goodness-of-fit measure θ in the original inequality requirements that appeared in Proposition 1. Specifically, this obtains

$$\frac{Q_i}{Q_j} \leq \frac{\Lambda_j \mathbf{W}_j \mathbf{X}_i + \Omega_i}{\Lambda_j \mathbf{W}_j \mathbf{X}_j + \Omega_j} + (1 - \theta),$$

which gives equation (1) in the main text

A.4 Reformulating the objective function in program (3)

Computing the objective $\min \sum_{i \in N} \mathrm{CSLI}_i$ in program (3) is equivalent to computing

$$\max \sum_{i \in N} (1 - \mathrm{CSLI}_i) = \max \sum_{i \in N} \frac{\mathbf{W}_i \mathbf{X}_i}{\mathbf{W}_i \mathbf{X}_i + \Gamma_i \Omega_i} = \max \sum_{i \in N} \frac{\Lambda_i \mathbf{W}_i \mathbf{X}_i}{\Lambda_i \mathbf{W}_i \mathbf{X}_i + \Omega_i}. \tag{4}$$

This objective is nonlinear in the unknowns Λ_i and Ω_i, which makes it difficult to compute. Therefore, in our empirical analysis we replace (4) by the objective

$$\max \sum_{i \in N} (\Lambda_i \mathbf{W}_i \mathbf{X}_i - \Omega_i), \tag{5}$$

which is linear in unknowns.
To see the connection between objective (5) instead of (4), let us consider

$$\begin{aligned} &\frac{\Lambda_i \mathbf{W}_i \mathbf{X}_i}{\Lambda_i \mathbf{W}_i \mathbf{X}_i + \Omega_i} \geq \rho \\ &\Leftrightarrow \Lambda_i \mathbf{W}_i \mathbf{X}_i \geq \rho(\Lambda_i \mathbf{W}_i \mathbf{X}_i + \Omega_i) \\ &\Leftrightarrow (1 - \rho) \Lambda_i \mathbf{W}_i \mathbf{X}_i - \rho \Omega_i \geq 0. \end{aligned}$$

Thus, larger differences in $(\Lambda_i \mathbf{W}_i \mathbf{X}_i - \Omega_i)$ relates to setting a higher ρ (which corresponds to a higher value of $(1 - \mathrm{CSLI}_i)$). As a result, higher values of $\sum_i (\Lambda_i \mathbf{W}_i \mathbf{X}_i - \Omega_i)$ lead to higher values of $\sum_i (1 - \mathrm{CSLI}_i)$.

A.5: Monte Carlo simulation – technical details

Data generating process. The representative firm maximizes the net present value of profits by choosing labor input and investment over time, subject to a Cobb Douglas production technology, and for a given capital accumulation equation and initial capital stock. Formally, this firm solves

$$\begin{aligned} \max_{(L_t, I_t)_{t \in \mathbb{T}}} \; & E_0 \sum_{t=0}^{\infty} \beta^t (Q_t - W_t^\ell L_t - g(I_t)), \qquad (OP.SIM) \\ \text{s.t. } & Q_t = A L^{\alpha_\ell} K^{\alpha_k}, \\ & K_{t+1} = (1 - \delta) K_t + I_t, \\ & K_0 = \overline{K}, \end{aligned}$$

with t the time period, L labor input, I investment, β the discount rate, $g(.)$ a convex investment cost function (e.g., adjustment costs), Q the value of output (i.e., the price of output is standardized to 1), W^{ℓ} the wage rate, A productivity, δ the depreciation rate of capital, α_{ℓ} the output elasticity of labor, α_k the output elasticity of capital, K capital use and $\overline{K}$ the initial capital stock. In terms of our above methodology, we can decompose A in latent input Ω and an output elasticity α_{Ω}, such that $A = \Omega^{\alpha_{\Omega}}$. Under a CRS assumption that covers both observed inputs and latent input (as discussed in Section 2.1), output elasticities correspond to total input cost shares. Therefore, $L^{\alpha_{\ell}} = \frac{W^{\ell}L}{Q}$, $L^{\alpha_k} = \frac{W^k K}{Q}$ for W^k the price of capital use, and $\alpha_{\Omega} = \frac{\Gamma \Omega}{Q}$ for Γ the price of latent input use.

Our data generating process (DGP) slightly simplifies the one of Van Biesebroeck (2007) and Ackerberg et al. (2015). In particular, we do not focus on data issues that are specific to dynamic optimization, such as the timing of labor input decisions, serially correlated wages and various forms of dynamics optimization errors. Still, as we will use that the unobserved convex investment costs $g(I_t)$ vary across firms, there is no strictly monotone relation between productivity and investment in our DGP. Therefore, our DGP is not consistent with the identification strategy of Olley and Pakes (1996). Further, as we allow wages W^{ℓ} to vary across firms, the Levinsohn and Petrin (2003) routine is not consistent with our DGP unless W^{ℓ} is included in the material input equation (see Ackerberg et al. (2015)). However, our DGP is consistent with the identification strategy of Ackerberg et al. (2015) when we assume Q to represent value added, output to be proportional to materials, and a production function that is Leontief in the material input.

More specifically, we assume that wages follow an i.i.d. distribution,

$$\ln(W_t^{\ell}) \sim i.i.d.N(0, \sigma_w^2),$$

with $\sigma_w = 0.1$, and the productivity parameter A follows an AR(1) process,

$$\ln(A_t) = (1-\rho)a + a_t, \text{ where } a_t = \rho a_{t-1} + \eta_t \text{ and } \eta_t \sim i.i.d.N(0, \sigma_a^2),$$

with $\rho = 0.7$, $a_0 = 0$ and $\sigma_a = 0.1$. Following Ackerberg et al. (2015), investment costs are given by $g(I_t) = \frac{\phi}{2} I^2$, with $1/\phi$ lognormally distributed over firms and constant over time with standard deviation 0.6. Further, we follow common practice by setting $\beta = 0.95$ and $\delta = 0.15$. Finally, our CRS technology is characterized by the following output elasticities: $\alpha_{\ell} = 0.6$ for labor, $\alpha_k = 0.3$ for capital, and $\alpha_{\Omega} = 0.1$ for latent input. We allow for unanticipated productivity shocks that cannot be modeled as latent input (i.e., measurement error, deviations from optimal conduct, etc.) by adding an i.i.d. distributed error term ε to the production function,

$$Q^{stoch} = A L^{\alpha_{\ell}} K^{\alpha_k} e^{\varepsilon}, \text{ where } \varepsilon \sim i.i.d.N(0, \sigma_{\varepsilon}^2).$$

Operationalization of OP.SIM. We start from (OP.SIM) and, for ease of exposition, we denote W^ℓ by W. Substituting I_t and Q_t gives

$$\max_{(L_t,K_{t+1})_{t\in\mathbb{N}}} \beta^t E_0 \sum_{t=0}^{\infty} (A_t L_t^{\alpha_\ell} K_t^{\alpha_k} - W_t L_t - g(K_{t+1} - (1-\delta)K_t)),$$

The first order condition with respect to L_t gives

$$\begin{aligned} A_t \alpha_\ell L_t^{\alpha_\ell - 1} K_t^{\alpha_k} - W_t = 0, \\ \leftrightarrow L_t = \left(\frac{\alpha_\ell A_t}{W_t}\right)^{\frac{1}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}}. \end{aligned}$$

If we plug this in into the optimization problem, we for the period t payoff function

$$\begin{aligned} & A_t \left(\left(\frac{\alpha_\ell A_t}{W_t}\right)^{\frac{1}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}}\right)^{\alpha_\ell} K_t^{\alpha_k} - W_t \left(\left(\frac{\alpha_\ell A_t}{W_t}\right)^{\frac{1}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}}\right) - g(K_{t+1} - (1-\delta)K_t), \\ & = \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}} - \alpha_\ell^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}} - g(K_{t+1} - (1-\delta)K_t), \\ & = (1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}} - g(K_{t+1} - (1-\delta)K_t) \end{aligned}$$

As such, we obtain the maximization problem

$$\max_{(K_{t+1})_{t\in\mathbb{N}}} \sum_{t=0}^{\infty} E_t \left[\beta^t \left((1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}} - g(K_{t+1} - (1-\delta)K_t)\right)\right],$$

Consider the Bellman equation

$$T(v(W_t, A_t, K_t)) = \max_{K_{t+1}} \left\{ \begin{array}{l} \left((1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}} - g(K_{t+1} - (1-\delta)K_t)\right) \\ +\beta E(v(W_{t+1}, A_{t+1}, K_{t+1}) | A_t, W_t) \end{array} \right\}.$$

The operationalization problem lies with the expectation, where we integrate over two stochastic variables, namely A_{t+1} and W_{t+1}.

To resolve this problem, assume that

$$\begin{aligned} & \ln(W_t) \sim i.i.d.\ N(0, \sigma_w^2), \\ & \ln(A_t) = (1-\rho)a + a_t, \text{ where, } a_t = \rho a_{t-1} + \eta_t, \\ & \eta_t \sim i.i.d.\ N(0, \sigma_a^2), \end{aligned}$$

and define

$$Z_t = (1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_t^{\frac{1}{1-\alpha_\ell}} W_t^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_t^{\frac{\alpha_k}{1-\alpha_\ell}}.$$

Then,

$$\begin{aligned}
E_t(Z_{t+1}|A_t,W_t) =& E_t\left((1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} A_{t+1}^{\frac{1}{1-\alpha_\ell}} W_{t+1}^{-\frac{\alpha_\ell}{1-\alpha_\ell}} K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}\right),\\
=&(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{(1-\rho)a+\rho\ln(A_t)}{1-\alpha_\ell}\right) E_t\left(\exp\left(\frac{\eta_{t+1}}{1-\alpha_\ell}\right)\right)\times,\\
&E_t\left(\exp\left(-\frac{\alpha_\ell\varepsilon_t}{1-\alpha_\ell}\right)\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}.
\end{aligned}$$

Now, if $\varepsilon \sim N(0,\sigma^2)$ then $E(\exp(\gamma\varepsilon)) = \exp\left(\frac{(\gamma\sigma)^2}{2}\right)$. As such,

$$\begin{aligned}
E_t(Z_{t+1}|A_t,W_t) =&(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_t}{1-\alpha_\ell}\right)\exp\left(\frac{\sigma_a^2}{2(1-\alpha_\ell)^2}\right)\times,\\
&\exp\left(\frac{(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}},\\
&= (1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_t}{1-\alpha_\ell} + \frac{\sigma_a^2+(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}.
\end{aligned}$$

Consider the operator $V(w)$ as

$$V(w(W_t,a_t,K_t)) = \max_{K_{t+1}}\left\{\begin{array}{l} \beta(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_t}{1-\alpha_\ell} + \frac{\sigma_a^2+(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}, \\ -g(K_{t+1}-(1-\delta)K_t) + \beta E(w(W_{t+1},a_{t+1},K_{t+1})|A_t,W_t) \end{array}\right\}.$$

where we made the change of variables $a_t = \ln(A_t) - a$. Observe that V is also a contraction mapping as it satisfies the Blackwell conditions.

Lemma 1. *The function v is a fixed point of the Bellman operator T if and only if w is a fixed point of the operator V, where $w(W,a,K) = v(W,A,K) - Z$. In addition, the policy functions (i.e., optimal levels of K) for both fixed points are identical.*

Proof. Let v be the fixed point of T, then,

$$\begin{aligned}
v(W_t,A_t,K_t) &= \max_{K_{t+1}}\left\{\begin{array}{l} Z_t - g(K_{t+1}-(1-\delta)K_t) \\ +\beta E(v(W_{t+1},A_{t+1},K_{t+1})|A_t,W_t) \end{array}\right\},\\
&= Z_t + \max_{K_{t+1}}\left\{\begin{array}{l} -g(K_{t+1}-(1-\delta)K_t) \\ +\beta E(v(W_{t+1},A_{t+1},K_{t+1}) - Z_{t+1}|A_t,W_t) + \beta E(Z_{t+1}|A_t,W_t) \end{array}\right\},\\
\leftrightarrow w(W_t,a_t,K_t) &= \max_{K_{t+1}}\left\{\begin{array}{l} \beta(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_t}{1-\alpha_\ell} + \frac{\sigma_a^2+(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}, \\ -g(K_{t+1}-(1-\delta)K_t) + \beta E(w(W_{t+1},a_{t+1},K_{t+1})|A_t,W_t) \end{array}\right\}.
\end{aligned}$$

□

Lemma 2. *The fixed point w of V is independent of W_t.*

Proof. Let S be the set of all value functions that are independent of W. It suffices to show that $V(S) \subseteq S$. Now, if $w \in S$, i.e., w is independent of W, then,

$$V(w(a_t, K_t)) = \max_{K_{t+1}} \left\{ \begin{array}{l} \beta(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_t}{1-\alpha_\ell} + \frac{\sigma_a^2+(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K_{t+1}^{\frac{\alpha_k}{1-\alpha_\ell}}, \\ -g(K_{t+1}-(1-\delta)K_t) + \beta E(w(a_{t+1}, K_{t+1})|A_t, W_t) \end{array} \right\}.$$

It can be seen that the right hand side value is independent of W_t, which shows that $V(w)$ will also be independent of W. □

The above two lemmata show that, in order to compute the fixed point of the Bellman equation, we can also compute the fixed point of the operator V. The optimal policy function of V will coincide with the optimal policy function of T. Additionally, in order to compute the expectation, we can ignore the randomness of W_{t+1} in the computation of the expectation operator.

Discretizing the technology process. We have that $a_t = \rho a_{t-1} + \eta_t$ where $\eta_t \sim N(0, \sigma_a^2)$. The mean of a_t is equal to

$$E(a_t) = \rho E(a_{t-1}) = \rho^2 E(a_{t-2}) = \ldots = \rho^t E(a_0) = 0.$$

Then, the variance of a_t is given by

$$\begin{aligned} E(a_t^2) &= \rho^2 E(a_{t-1}^2) + \sigma_a^2, \\ &= (1+\rho^2)\sigma_a^2 + \rho^4 E(a_{t-2}^2), \\ &= \ldots = (1+\rho^2+\rho^4+\ldots+\rho^{2t})\sigma_a^2 + \rho^{2t}E(a_0^2), \\ &\approx \frac{\sigma_a^2}{1-\rho^2}. \end{aligned}$$

Let $\tilde{a}$ be the discrete valued process to approximate a, and let $\{a_1, \ldots, a_N\}$ be the finite set of realizations of $\tilde{a}$. We choose a_N to be m times the unconditional standard deviation,

$$a_N = m\left(\frac{\sigma_a^2}{1-\rho^2}\right)^{1/2}.$$

We set $a_1 = -a_N$ and we equally space the values $\{a_2, \ldots, a_{N-1}\}$ between this min and max. Let d be the space between these successive points. Then, for $1 < k < N$, pick

$$\begin{aligned} \pi_{jk} &= \Pr\{\tilde{y}_t = y_k|\tilde{y}_{t-1} = y_j\} = \Pr\{y_k - d/2 < \rho y_j + \eta_t < y_k + d/2\}, \\ &= \Pr\{y_k - d/2 - \rho y_j < \eta_t < y_k + d/2 - \rho y_j\}, \\ &= \Phi\left(\frac{y_k + d/2 - \rho y_j}{\sigma_a}\right) - \Phi\left(\frac{y_k - d/2 - \rho y_j}{\sigma_a}\right), \end{aligned}$$

where Φ is the cdf of $N(0,1)$. For $k = 1, N$, pick

$$\pi_{j1} = \Phi\left(\frac{y_1 + d/2 - \rho y_j}{\sigma_a}\right),$$
$$\pi_{jN} = 1 - \Phi\left(\frac{y_N - d/2 - \rho y_j}{\sigma_a}\right).$$

The final program. With this discretization, we get

$$V(w(a_j, K)) = \max_{K^*}\left\{\begin{array}{l} \beta(1-\alpha_\ell)\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{a+\rho a_j}{1-\alpha_\ell} + \frac{\sigma_a^2+(\alpha_\ell\sigma_w)^2}{2(1-\alpha_\ell)^2}\right) K^{*\frac{\alpha_k}{1-\alpha_\ell}}, \\ -g(K^* - (1-\delta)K) + \sum_k \pi_{jk} w(a_k, K^*) \end{array}\right\}.$$

A.6: Monte Carlo simulation – additional results

Table 7: Monte Carlo results: 500 firm observations

	PRODUCTION FUNCTION PARAMETERS					
	$\alpha^\ell = 0.6$		$\alpha^k = 0.3$		$CSLI = 0.1$	
σ_ε	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP			
0	0.613	0.004	0.307	0.002	0.080	0.006
0.1	0.598	0.005	0.299	0.002	0.103	0.007
0.3	0.568	0.010	0.284	0.005	0.148	0.015
			OLS			
0	0.770	0.015	0.319	0.008	–	–
0.1	0.770	0.017	0.319	0.009	–	–
0.3	0.769	0.029	0.319	0.012	–	–
			ACF			
0	0.599	0.027	0.300	0.047	–	–
0.1	0.596	0.033	0.304	0.077	–	–
0.3	0.576	0.180	0.337	0.215	–	–
	SPEARMAN CORRELATION WITH PRODUCTIVITY					
	$\Gamma\Omega$		a		$a + \varepsilon$	
σ_ε	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP			
0	0.765	0.033	0.981	0.006	0.981	0.006
0.1	0.617	0.037	0.739	0.037	0.942	0.034
0.3	0.350	0.040	0.360	0.047	0.775	0.077
			OLS			
0	0.176	0.028	0.754	0.040	0.754	0.040
0.1	0.119	0.022	0.523	0.040	0.844	0.028
0.3	0.051	0.017	0.229	0.037	0.961	0.013
			ACF			
0	0.746	0.056	0.993	0.008	0.993	0.008
0.1	0.593	0.067	0.784	0.040	0.993	0.028
0.3	0.298	0.115	0.396	0.083	0.985	0.069

Note: B=1000. The NP estimates of production function parameters are firm-year specific with an unknown distribution. For each sample $b = 1, ..., B$, we report the median total cost shares of labor, capital and latent input as, respectively, α^ℓ,α^k and CSLI. We use the ACF procedure as provided in the '*prodest*' package in R. We set as starting values for the ACF procedure the true output elasticities augmented with normally distributed noise with standard deviation 0.1.

Table 8: Monte Carlo results: 2000 firm observations

	PRODUCTION FUNCTION PARAMETERS					
	$\alpha^{\ell} = 0.6$		$\alpha^{k} = 0.3$		$CSLI = 0.1$	
σ_{ε}	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP			
0	0.612	0.002	0.306	0.001	0.082	0.003
0.1	0.598	0.002	0.299	0.001	0.103	0.004
0.3	0.574	0.005	0.287	0.002	0.140	0.007
			OLS			
0	0.769	0.007	0.318	0.004	–	–
0.1	0.768	0.008	0.318	0.004	–	–
0.3	0.768	0.015	0.318	0.006	–	–
			ACF			
0	0.600	0.014	0.300	0.025	–	–
0.1	0.600	0.016	0.300	0.027	–	–
0.3	0.596	0.031	0.303	0.039	–	–
	SPEARMAN CORRELATION WITH PRODUCTIVITY					
	$\Gamma\Omega$		a		$a + \varepsilon$	
σ_{ε}	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP			
0	.768	0.016	0.983	0.002	0.983	0.002
0.1	0.631	0.019	0.757	0.018	0.959	0.016
0.3	0.368	0.020	0.388	0.023	0.841	0.035
			OLS			
0	0.178	0.015	0.760	0.020	0.760	0.020
0.1	0.121	0.012	0.528	0.021	0.847	0.015
0.3	0.052	0.009	0.232	0.019	0.962	0.007
			ACF			
0	0.746	0.029	0.998	0.002	0.998	0.002
0.1	0.593	0.029	0.789	0.016	0.998	0.002
0.3	0.298	0.030	0.394	0.033	0.999	0.001

Note: B=1000. The NP estimates of production function parameters are firm-year specific with an unknown distribution. For each sample $b = 1, ..., B$, we report the median total cost shares of labor, capital and latent input as, respectively, α^{ℓ},α^{k} and CSLI. We use the ACF procedure as provided in the '*prodest*' package in R. We set as starting values for the ACF procedure the true output elasticities augmented with normally distributed noise with standard deviation 0.1.

Table 9: Monte Carlo results: 1000 firm observations, alternative θ values

	PRODUCTION FUNCTION PARAMETERS					
	$\alpha^{\ell} = 0.6$		$\alpha^{k} = 0.3$		$CSLI = 0.1$	
σ_{ε}	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP, $\theta = 0.90$			
0	0.630	0.003	0.315	0.002	0.056	0.005
0.1	0.616	0.004	0.308	0.002	0.076	0.006
0.3	0.588	0.008	0.294	0.004	0.118	0.011
			NP, $\theta = 0.99$			
0	0.590	0.005	0.295	0.003	0.115	0.008
0.1	0.570	0.006	0.285	0.003	0.145	0.009
0.3	0.523	0.015	0.261	0.007	0.216	0.022
	SPEARMAN CORRELATION WITH PRODUCTIVITY					
	$\Gamma\Omega$		a		$a+\varepsilon$	
σ_{ε}	Mean	St.Dev.	Mean	St.Dev	Mean	St.Dev
			NP, $\theta = 0.90$			
0	0.780	0.031	0.975	0.006	0.975	0.006
0.1	0.642	0.033	0.753	0.028	0.960	0.019
0.3	0.382	0.038	0.396	0.042	0.862	0.049
			NP, $\theta = 0.99$			
0	0.755	0.037	0.976	0.016	0.976	0.016
0.1	0.572	0.048	0.671	0.066	0.843	0.076
0.3	0.293	0.044	0.265	0.055	0.440	0.129

Note: B=1000. The NP estimates of production function parameters are firm-year specific with an unknown distribution. For each sample $b = 1, ..., B$, we report the median total cost shares of labor, capital and latent input as, respectively, α^{ℓ},α^{k} and CSLI.

Appendix B: Quantity based estimates of productivity for ready mixed concrete producers

Syverson (2004) argues that proxy variable approaches, such as the Olley and Pakes (1996) routine, are not appropriate to empirically analyze the sector of ready mixed concrete producers. Local demand states may influence input and investment decisions, which makes the assumption of a one-to-one relation between unobserved productivity and observable investment difficult to maintain. Interestingly, because our routine does not rely on (semi-)parametric structuring of the simultaneity issue, it remains well applicable to this sector. More generally, as instruments to deal with the simultaneity bias are not always easily available, our methodology broadens the reach of available empirical methods to analyze productivity differences.

In our analysis of ready mixed concrete producers, we make use of (only) 118 firm-year observations on 30 small firms (see Table 10 for summary statistics). To deal with the large heterogeneity in production quantities, we set our goodness-of-fit parameter θ equal to 0.90 for this particular setting.[22] The very small sample size indicates that caution is needed

[22] We cleaned the data in a similar manner as for our main analysis, but add the output value and output quantity as variables that require cleaning. We define firm-year observations as representing single-product

when interpreting the results. Medium and large firms were not considered because of the small number of observations available. Our estimates of NP and CSLI are summarized in Table 11. The Spearman correlations with a one-year lag are above 0.77 for our three measures of NP (i.e., revenue based, produced value based and quantity based), which confirms the well established finding of persistent technological heterogeneity in narrowly defined industries. The Spearman correlation between our quantity based indicators and produced value based indicators is positive, but moderate (0.70). This demonstrates once more that value based estimation results may differ substantially from quantity based results at the level of individual firm observations. The Spearman correlation between revenue based NP and produced value based NP is 0.54 and the Spearman correlation with quantity based NP equals only 0.39. This last result reveals that a general indicator of NP captures more than the pure technological manufacturing features of the firm. On average, the quantity based and produced value based cost shares of latent input amount to respectively 11 and 17 percent. The revenue based estimates are on average 6 percent.

Figure 2 presents the evolution of the input cost shares. Due to the small sample size, the evolution patterns should be considered with caution as they are subject to changing sample compositions over the period under study. Still, all three CSLI estimates show a similar evolution over time of the input cost shares. Cost shares are evolving in favor of domestic materials and tangible fixed assets and against foreign materials and latent input, while labor cost shares are fairly constant over time. Stated differently, regardless of the how we measure latent input, also for this well defined industry we find strong empirical evidence against Hicks neutrality.

Table 10: Summary statistics – ready mixed concrete sector

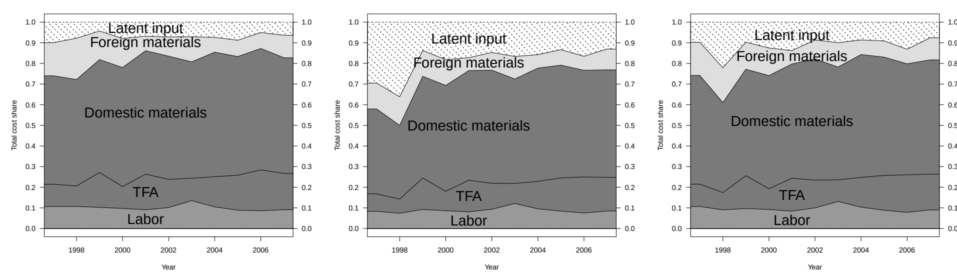
	Mean	St.Dev.	Min.	25 perc.	Median	75 perc.	Max.
Deflated revenue	7.12	3.91	0.63	4.74	6.23	8.20	22.78
Deflated produced value	5.69	2.46	0.76	4.36	5.12	7.02	13.71
Output quantity	245.26	104.39	29.70	177.16	235.80	292.14	634.04
Nace 2-digit output price	1.16	0.09	1.02	1.07	1.17	1.22	1.31
Labor in FTE	22.67	11.26	10.00	13.42	19.75	30.67	49.70
Deflated tangible fixed assets	1.35	1.23	0.04	0.61	0.98	1.64	6.83
Deflated material costs	5.95	3.46	0.66	3.84	5.29	6.72	20.14
Labor price	0.04	0.01	0.03	0.04	0.04	0.05	0.06
Capital price	1.14	0.07	1.03	1.08	1.11	1.18	1.26
Intermediates price	1.14	0.09	1.01	1.03	1.13	1.20	1.28
Sourcing (share)	0.72	0.11	0.30	0.67	0.72	0.79	0.94

Note: Deflated revenue, deflated produced value, deflated tangible fixed assets, deflated material costs and labor price are expressed in millions of euro. Output quantity is expressed in millions of kilogram.

firms if the value of one 8-digit product is over 90 percent of the production value.

Table 11: Summary statistics of productivity

	Mean	St.Dev.	Min.	25 perc.	Median	75 perc.	Max.
	Ready mixed concrete producers						
NP (quantity based)	1.18	1.86	0.00	0.16	0.48	1.62	15.29
NP (produced value based)	2.08	2.91	0.00	0.41	1.25	2.64	20.25
NP (revenue based)	0.71	0.94	0.00	0.08	0.45	0.93	5.55
CSLI (quantity based)	0.11	0.12	0.00	0.02	0.05	0.17	0.66
CSLI (produced value based)	0.17	0.13	0.00	0.06	0.15	0.25	0.66
CSLI (revenue based)	0.06	0.06	0.00	0.01	0.05	0.10	0.32



(a) Revenue based (b) Produced value based (c) Quantity based

Figure 2: Cost shares

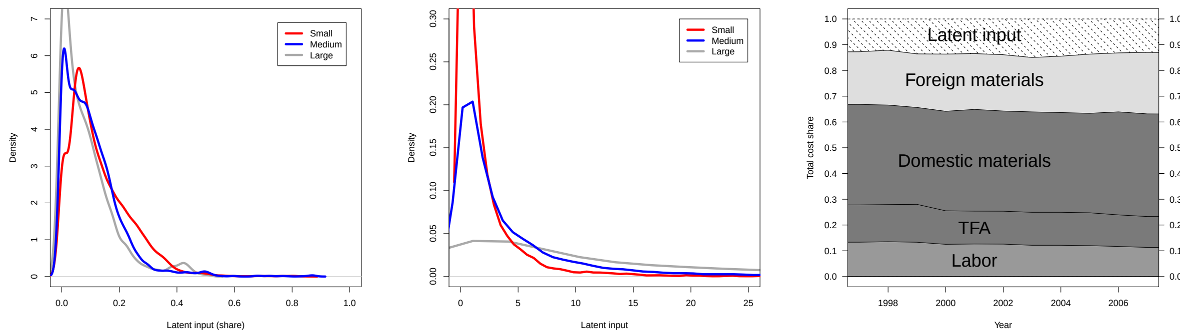
Appendix C: Alternative values for the goodness-of-fit parameter θ

Our main empirical findings are not sensitive to altering the value of the goodness-of-fit parameter θ. To show this, we replicated our complete analysis for values of θ equal to 0.925 and 0.900. Results are highly robust for altering the value of θ. Table 12 shows that the Spearman correlation between our measures of productivity with $\theta = 0.950$ and $\theta = 0.925$ is more than 0.8. Further, from Figure 3 we learn that the distribution of NP and cost shares for the three firm size groups is highly similar for different θ-values. All main findings on the evolution of input cost shares are robust for changing θ to 0.925 or 0.900. The same applies to the associations between international exposure, sourcing and productivity (results available upon request).

Table 12: Spearman correlations between NP estimates for different values of the goodness-of-fit parameter

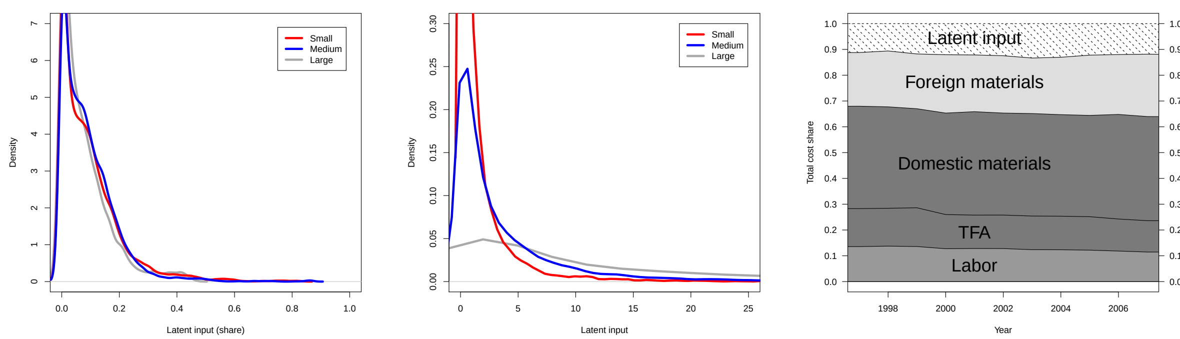
	All	Small	Medium	Large
$\theta = 0.925$	0.81	0.77	0.80	0.93
$\theta = 0.900$	0.70	0.52	0.76	0.94

$\theta = \mathbf{0.925}$



(a) CSLI (b) NP (c) Cost shares

$\theta = \mathbf{0.900}$



(d) CSLI (e) NP (f) Cost shares

Figure 3: Distribution NP and CSLI, and cost shares for alternative values of θ

Appendix D: Additional empirical results

D.1 Summary statistics of productivity

Table 13 provides summary statistics for our NP and CSLI estimates (both revenue based and produced value based).[23]

[23]Our estimates of NP (based on produced value) contain a small proportion of unrealistic values for some specific firms. Therefore, we exclude the observations that belong to the 5 percent highest values of our estimated cost share of latent input. Results available upon request show that our main results are robust for including these observations. A potential explanation for this difference may be that there is a higher level of misreporting in the Prodcom survey than in the financial accounts contained in the Central Balance Sheet Office database.

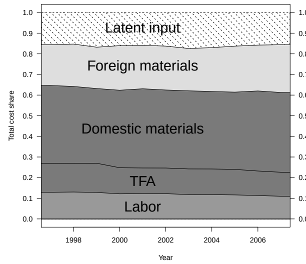
Table 13: Summary statistics of NP and CSLI

	Mean	St.Dev.	Min.	25 perc.	Median	75 perc.	Max.
				Revenue based – all sectors			
NP	8.49	38.06	0.00	1.04	2.51	6.19	1573.93
NP (small firms)	3.45	6.03	0.00	0.89	1.83	4.02	281.56
NP (medium firms)	5.89	10.97	0.00	1.15	2.99	6.69	307.79
NP (large firms)	40.60	106.94	0.00	3.51	11.15	32.02	1573.93
CSLI	0.19	0.14	0.00	0.08	0.17	0.27	0.91
CSLI (small firms)	0.25	0.15	0.00	0.13	0.24	0.35	0.91
CSLI (medium firms)	0.14	0.10	0.00	0.06	0.13	0.21	0.88
CSLI (large firms)	0.11	0.09	0.00	0.03	0.09	0.15	0.63
				Produced value based – all sectors			
NP	11.89	70.67	0.00	0.75	2.15	6.39	2835.74
NP (small firms)	2.51	5.26	0.00	0.46	1.22	2.59	156.07
NP (medium firms)	7.56	16.56	0.00	1.22	3.37	8.51	695.44
NP (large firms)	67.18	199.40	0.00	4.49	17.54	49.05	2835.74
CSLI	0.17	0.13	0.00	0.08	0.16	0.25	0.74
CSLI (small firms)	0.18	0.12	0.00	0.09	0.17	0.26	0.74
CSLI (medium firms)	0.17	0.13	0.00	0.08	0.15	0.24	0.73
CSLI (large firms)	0.17	0.15	0.00	0.04	0.14	0.25	0.73

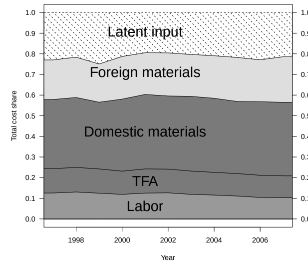
D.2 The evolution of total cost shares

Figure 4 and Tables 14 and 15 provide the evolution over time of the cost shares of latent input (again both revenue based and produced value based). Figures available upon request show that our empirical findings on the evolution of cost shares (summarized in Figure 4) are not specific to one manufacturing sector. The reported patterns are also not sensitive to including additional information on the firm's age and exporting status, or to applying a more detailed definition of the sector. Results available upon request confirm that the general picture of input cost share changes against primary inputs equally applies to mature firms, non-exporting and exporting firms.

All firms

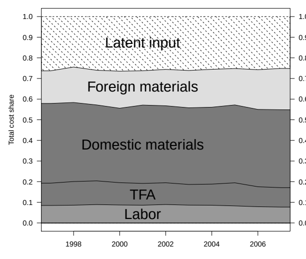


(a) Revenue based

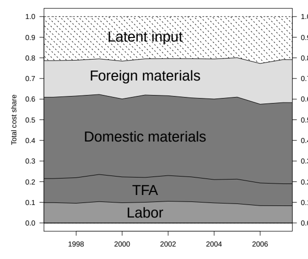


(b) Produced value based

Small firms

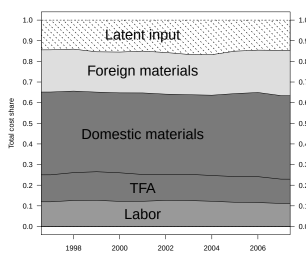


(c) Revenue based

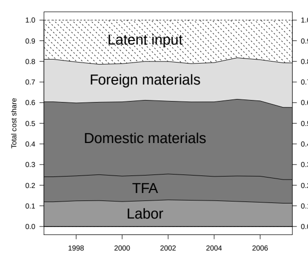


(d) Produced value based

Medium firms

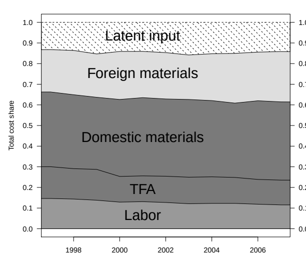


(e) Revenue based

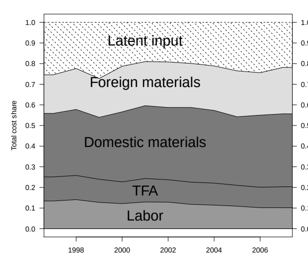


(f) Produced value based

Large firms



(g) Revenue based



(h) Produced value based

Figure 4: Total cost shares and firm size

Table 14: Revenue based total cost shares and firm size

	97	98	99	00	01	02	03	04	05	06	07
						All firms					
Labor	0.13	0.13	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.11	0.11
Tangible fixed assets	0.14	0.14	0.14	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Domestic materials	0.38	0.37	0.36	0.38	0.38	0.38	0.38	0.38	0.37	0.39	0.39
Foreign materials	0.20	0.21	0.20	0.22	0.21	0.21	0.21	0.21	0.22	0.22	0.23
Latent input	0.16	0.15	0.17	0.16	0.16	0.16	0.17	0.17	0.16	0.16	0.16
						Small firms					
Labor	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.08	0.08	0.08
Tangible fixed assets	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.11	0.10	0.09
Domestic materials	0.39	0.38	0.37	0.36	0.38	0.37	0.37	0.37	0.38	0.37	0.38
Foreign materials	0.16	0.17	0.17	0.18	0.17	0.18	0.18	0.18	0.18	0.19	0.20
Latent input	0.26	0.24	0.26	0.26	0.26	0.26	0.26	0.26	0.25	0.26	0.25
						Medium firms					
Labor	0.12	0.13	0.13	0.12	0.12	0.13	0.13	0.12	0.12	0.12	0.11
Tangible fixed assets	0.13	0.14	0.14	0.14	0.13	0.13	0.13	0.12	0.12	0.13	0.12
Domestic materials	0.40	0.39	0.39	0.39	0.40	0.39	0.39	0.39	0.40	0.41	0.40
Foreign materials	0.20	0.20	0.20	0.20	0.20	0.20	0.19	0.20	0.21	0.21	0.22
Latent input	0.14	0.14	0.15	0.15	0.15	0.16	0.17	0.17	0.15	0.15	0.15
						Large firms					
Labor	0.15	0.14	0.14	0.13	0.13	0.13	0.12	0.12	0.12	0.12	0.12
Tangible fixed assets	0.15	0.15	0.15	0.12	0.12	0.13	0.13	0.13	0.13	0.12	0.12
Domestic materials	0.36	0.36	0.35	0.37	0.38	0.37	0.38	0.37	0.36	0.38	0.38
Foreign materials	0.21	0.21	0.21	0.23	0.22	0.22	0.22	0.23	0.24	0.24	0.24
Latent input	0.13	0.14	0.15	0.14	0.14	0.15	0.16	0.15	0.15	0.14	0.14

Table 15: Produced value based total cost shares and firm size

	97	98	99	00	01	02	03	04	05	06	07
						All firms					
Labor	0.13	0.13	0.12	0.12	0.13	0.13	0.12	0.12	0.11	0.10	0.10
Tangible fixed assets	0.12	0.12	0.12	0.11	0.12	0.11	0.11	0.11	0.11	0.11	0.11
Domestic materials	0.34	0.34	0.32	0.35	0.36	0.35	0.36	0.36	0.35	0.36	0.36
Foreign materials	0.19	0.20	0.19	0.21	0.20	0.21	0.20	0.21	0.21	0.20	0.22
Latent input	0.23	0.22	0.25	0.21	0.19	0.20	0.20	0.21	0.22	0.23	0.21
						Small firms					
Labor	0.10	0.10	0.10	0.10	0.10	0.11	0.10	0.10	0.09	0.08	0.08
Tangible fixed assets	0.12	0.12	0.13	0.13	0.12	0.12	0.12	0.11	0.12	0.11	0.11
Domestic materials	0.39	0.40	0.39	0.38	0.40	0.39	0.38	0.39	0.40	0.38	0.39
Foreign materials	0.18	0.17	0.17	0.18	0.18	0.18	0.19	0.19	0.19	0.20	0.21
Latent input	0.21	0.21	0.20	0.22	0.20	0.20	0.20	0.21	0.20	0.23	0.21
						Medium firms					
Labor	0.12	0.12	0.13	0.12	0.12	0.13	0.13	0.13	0.12	0.12	0.11
Tangible fixed assets	0.12	0.12	0.13	0.12	0.12	0.13	0.12	0.12	0.12	0.13	0.11
Domestic materials	0.36	0.35	0.35	0.36	0.36	0.35	0.35	0.36	0.37	0.36	0.35
Foreign materials	0.21	0.20	0.18	0.19	0.19	0.19	0.19	0.19	0.20	0.20	0.22
Latent input	0.19	0.20	0.21	0.21	0.20	0.20	0.21	0.21	0.18	0.19	0.21
						Large firms					
Labor	0.13	0.14	0.13	0.12	0.13	0.13	0.12	0.11	0.11	0.10	0.10
Tangible fixed assets	0.12	0.12	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.10
Domestic materials	0.31	0.32	0.30	0.34	0.35	0.35	0.36	0.35	0.33	0.35	0.35
Foreign materials	0.19	0.20	0.19	0.22	0.21	0.22	0.21	0.22	0.22	0.21	0.23
Latent input	0.25	0.22	0.27	0.21	0.19	0.19	0.20	0.21	0.24	0.24	0.22