ISSN 2042-2695

CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1402

Revised October 2022
(Replaced January 2016 version)

Unraveling Firms: Demand, Productivity and Markups Heterogeneity

Emanuele Forlani
Ralf Martin
Giordano Mion
Mirabelle Muûls

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

ESRC ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
We develop a novel framework that simultaneously allows recovering heterogeneity in demand, quantity TFP and markups across firms while leaving the correlation between the three dimensions unrestricted. We accomplish this by explicitly introducing demand heterogeneity and systematically exploiting assumptions used in previous productivity estimation approaches. In doing so, we provide an exact decomposition of revenue productivity in terms of the underlying heterogeneities, thus bridging the gap between quantity and revenue productivity estimations. We use Belgian firms production data to quantify TFP, demand and markups and show how they are correlated with each other, across time and with measures obtained from other approaches. In doing so, we find quantity TFP and demand to be strongly negatively correlated with each other so suggesting a trade-off between the quality of a firm's products and their production cost. We also show how our framework provides deeper and sharper insights on the response of firms to increasing import competition from China. In particular, we find that changes in revenue productivity materialise as the outcome of complex, and sometimes offsetting, changes in quantity TFP, demand, markups and production scale.

Keywords: demand, productivity, markups, production function estimation, import competition, China
JEL codes: D24; L11; L25; F14

This paper was produced as part of the Centre’s Growth Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We thank Andrea Ariu, Andrew Bernard, Ingo Borchert, Rosario Crino, Jan De Loecker, Catherine Fuss, Marc Melitz, Luca David Opromolla, Chad Syverson, Ilke Van Beveren and John Van Reenen as well as participants at various seminars and workshops and an anonymous referee from the National Bank of Belgium. Most of this research has been carried out while Giordano Mion was affiliated to the National Bank of Belgium. The views expressed in this paper are our own and do not necessarily reflect those of the National Bank of Belgium
Emanuele Forlani, Department of Economics and Management, University of Pavia, Italy. Ralf Martin, Imperial College Business School, Imperial College, London and Senior Research Associate at Centre for Economic Performance, London School of Economics. Giordano Mion, Department of Economics, University of Sussex and Associate at Centre for Economic Performance, London School of Economics. Mirabelle Muûls, Imperial College Business School and Grantham Institute and Associate at Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Economists are interested in estimating firm-level productivity in a range of fields. These estimates are often used as inputs in applications such as the firm size distribution, firm survival and growth, self-selection of firms into exporters and non-exporters or FDI activities, just to name a few. The most common approach in the literature to measure productivity involves estimating a production function by regressing output quantity on input quantity and using the resulting residual shock as a productivity index typically referred to as Total Factor Productivity (TFP). This raises at least three issues.

First, most studies do not have output quantity data available at the firm-level so that regressions are fitted using revenue data, i.e. price times quantity. Such revenue-based measures of TFP look quite different from quantity-based ones (Foster et al., 2008). A second well known issue is the endogeneity of production factors used as explanatory variables (Olley and Pakes, 1996). Third, and more importantly, firms could be heterogeneous in dimensions other than TFP. In this respect the IO literature on demand systems (Ackerberg et al., 2007) points to substantial heterogeneity in both markups and consumers' willingness to pay for the products sold by different firms. For example, the presence of vertical and horizontal product differentiation means that firms selling otherwise similar products face rather different demands. At the same time, market power variations, due to product quality or technical efficiency, could substantially affect the markup that firms can charge. Being able to account for these different dimensions and their interconnections is important for several reasons.

One of these, is that it is crucial in order to correctly measure TFP. In this respect, higher measured TFP is typically seen as welfare improving. However, conventional measures of TFP conflate actual TFP with demand and markup heterogeneity which may lead to different welfare implications. In addition, being able to actually quantify dimensions other than TFP matters from both a welfare and a policy point of view. From a welfare perspective it is, for example, of great value to be able to assess the impact on firm markups of a trade integration episode or market size expansion. Some recent theoretical papers have indeed revisited the relationship between market size, markups and welfare and questioned the pervasiveness of the so called "pro-competitive effects" (Zhelobodko et al., 2012). Furthermore, being able to disentangle demand heterogeneity from efficiency is important for policy making, in particular to understand where the competitiveness of a firm or an industry comes from and then target interventions accordingly. In this respect, policy changes may affect efficiency and demand/quality in opposite directions, so examining effects on revenue-based TFP may be misleading and not reflect changes in underlying physical productivity.

This paper's contribution is to address these issues in a comprehensive way by making use of quantity and price data at the firm-level while developing and estimating a model where firms are heterogeneous with respect to their quantity TFP, markups and demand. More specifically, while the literature already provides models allowing to quantify heterogeneity in quantity TFP and markups (De Loecker et al., 2016), we explicitly introduce heterogeneity in demand across firms within a general framework where heterogeneity in quantity TFP, markups and demand can be simultaneously measured. In doing so, we thus allow measuring demand heterogeneity across firms without resorting to demand system models (Ackerberg et al., 2007) or to the restrictive assumptions imposed by the methodology developed in Foster et al. (2008). At the same time, we also depart from the standard proxy variable approach and develop an alternative estimation procedure for the parameters of the underlying production function.

We apply our framework to Belgian manufacturing firms and use information on both the quantity and the value of production over the period 1996-2007 to quantify our model. We first document that demand factors display at least as much variability across firms as quantity TFP. We further show that productivity and demand heterogeneity are very strongly and negatively correlated. This finding is suggestive of a trade-off between the appeal/perceived quality of a firm's products (our measure of demand) and their production cost (linked to quantity TFP) as indeed suggested in the demand system literature (Ackerberg et al., 2007).[1] Another pattern worth noting is that differences in prices and markups across firms are related to differences in demand and productivity in the way one would expect. More specifically, we find markups to be increasing in quantity while more productive firms and/or firms selling more appealing products charge higher markups. At the same time, more productive firms charge lower prices while firms selling more appealing products charge higher prices. When comparing our measure of demand with the one developed in Foster et al. (2008), we find the two measures to be mostly orthogonal to each other. We rationalise this finding in the light of the two key restrictions imposed by Foster et al. (2008), while also providing evidence that our demand measure correlates well with a measure obtained using demand elasticity estimates borrowed from Broda and Weinstein (2006). We further show how, when correctly measured, revenue TFP exactly decomposes into the underlying dimensions of heterogeneity so bridging the gap between quantity TFP estimations and revenue TFP estimations.

We finally assess how and to what extent these heterogeneities allow gaining deeper and

[1] The negative correlation we unveil can be rationalised in several ways. For example, one could reasonably argue that technology is such that higher quality products require more and/or more expensive inputs, i.e. lower quantity TFP. On the other hand, even if quantity TFP and demand were uncorrelated from a technology point of view, a negative correlation between the two will arise after selection has taken place and only firms with high enough quantity TFP and/or high enough demand survive. We provide more insights into this issue later on in the paper.

sharper insights into a productivity-related question: firm response to increasing import competition from China. Numerous studies have explored the many, besides the well-documented negative effects on employment (Autor et al., 2013), impacts of the spectacular rise of Chinese trade. With respect to productivity, Bloom et al. (2016) provide evidence supporting the claim that import competition from China caused an increase in revenue TFP for European firms. Bloom et al. (2016) rationalise these effects via a number of channels relating competition to innovation and X-inefficiencies. Building on our data and framework, we show how changes in Belgian firms revenue productivity spurred by import competition from China materialise as the outcome of complex changes in quantity TFP, product appeal, markups and production scale. More specifically, we find that quantity TFP increases and product appeal decreases while markups are little affected. The two opposing effects roughly cancel each other out and so the observed increased in revenue TFP essentially comes from the reduction in firm operations/scale. This application highlights how our framework allows to better understand firm behaviour and margins of adjustment under competitive pressure.

Our paper is related to the literature on firm TFP measurement on which the Olley and Pakes (1996) proxy variable approach to tackle the issue of endogeneity has had a deep impact. This proxy variable approach has been further developed in Levinsohn and Petrin (2003), Wooldridge (2009), Ackerberg et al. (2015) and De Loecker et al. (2016). Our interest in demand heterogeneity is common to both Foster et al. (2008) and De Loecker (2011). De Loecker (2011) introduces demand heterogeneity in a revenue-based production function model while relying on standard CES preferences and a common markup across varieties. This allows substituting for prices and getting a tractable expression for firm revenue as a function of inputs, TFP and demand heterogeneity. Compared to our framework, De Loecker (2011) does not allow for different markups across varieties while needing some adequate proxies for demand shocks. By contrast Foster et al. (2008), which is the most closely related paper to ours in terms of both data and aims, use data on both the quantity and the value of a firm's production in order to disentangle quantity TFP from demand heterogeneity. More specifically, they recover production function coefficients from industry average cost shares and subsequently estimate a demand system featuring demand heterogeneity measured as regression residuals while instrumenting firm price with firm TFP. Therefore, the key identifying assumption allowing them to disentangle productivity from demand is, besides imposing constant markups, that they are uncorrelated. In our framework we do not impose such assumptions and find productivity and demand heterogeneity to be very strongly correlated with each other.

The rest of the paper is organised as follows. Section 2 provides our model while Section 3 develops the estimation strategy. Section 4 presents our data while Section 5 contains our estimation results as well as descriptive statistics and correlations. In Section 6 we compare our demand measure with measures obtained using other approaches. In Section 7 we show how our revenue productivity decomposition can be used to obtain deeper insights into firm response to increasing import competition from China. Section 8 concludes. Additional details, results, Figure and Tables are provided in the Appendix.

2 The MULAMA model

We label our model MULAMA because of the names we give to the 3 heterogeneities we allow for: markups **MU**, demand **LAM**bda and quantity productivity **A**. While the literature already provides models allowing to quantify firm-level TFP and markups (De Loecker et al., 2016), our key contribution is to further allow measuring demand heterogeneity across firms without resorting to demand system models (Ackerberg et al., 2007) or to the restrictive assumptions imposed by the methodology developed in Foster et al. (2008). In doing so, we depart from the standard proxy variable approach and develop an alternative estimation procedure for the parameters of the underlying production function. In what follows, we focus on the case of single-product firms while extending the model to multi-product firms in Appendix C.

2.1 Production

We index firms by i and time by t. We consider a Cobb-Douglas production technology with 3 production factors: labour (L), materials (M) and capital (K). In line with the existing literature we assume capital to be a dynamic input that is predetermined in the short-run, i.e. current capital has been chosen in the past and cannot immediately adjust to current period shocks.[2] We further assume, as standard in the literature, that materials are a variable input, free of adjustment costs. In the case of labour, we could either assume it is a variable input, free of adjustment costs, or we could assume it is, very much like capital, predetermined in the short-run as in De Loecker et al. (2016). We could also assume, following Ackerberg et al. (2015), that it is a semi-flexible input.[3] In light of the features of the Belgian labour

[2]As described in Ackerberg et al. (2015), capital is often assumed to be a dynamic input subject to an investment process with the period t capital stock of the firm actually determined at period $t-1$. Intuitively, the restriction behind this assumption is that it takes a full period for new capital to be ordered, delivered, and installed.

[3]More precisely, in the semi-flexible case, L_{it} is chosen by firm i at time $t-b$ $(0<b<1)$, after K_{it} being chosen at $t-1$ but prior to M_{it} being chosen at t. In this case, one should expect L_{it} to be correlated with productivity shocks in t. Yet labour would not adjust as fully to such shocks as materials do. The choice between predetermined and semi-flexible for L_{it} does not change the structure of the model and estimation procedure we provide below but only affects the set of moments used in the estimation. We highlight any

market, we opt for the predetermined case.

We further assume firms minimise costs while taking the price of materials (W_{Mit}) as given. Labour and capital cannot adjust to current period shocks but firms can still adapt materials. Therefore, at any given point in time, each firm i is dealing with the following short-run cost minimisation problem:[4]

$$\min_{M_{it}} \{M_{it}W_{Mit}\} \text{ s.t. } Q_{it} = A_{it}L_{it}^{\alpha_L}M_{it}^{\alpha_M}K_{it}^{\gamma-\alpha_L-\alpha_M},$$

where the capital coefficient is $\alpha_K = \gamma - \alpha_L - \alpha_M$, γ characterises returns to scale, and A_{it} is quantity TFP, which is observable to the firm (and influences her choices) but not to the econometrician. In what follows we refer to the Cobb-Douglas production technology as the quantity equation and denote with lower case the log of a variable (for example a_{it} denotes the natural logarithm of A_{it}). The quantity equation can thus be written as:

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + (\gamma - \alpha_L - \alpha_M)k_{it} + a_{it}. \tag{1}$$

First order conditions to the firm's cost minimisation problem imply that:

$$W_{Mit} = \chi_{it}\frac{Q_{it}}{M_{it}}\alpha_M \tag{2}$$

where χ_{it} is a Lagrange multiplier.[5]

We can thus write the short-run cost function as:

$$C_{it} \equiv M_{it}W_{Mit} = \chi_{it}Q_{it}\alpha_M = W_{Mit}\left(\frac{Q_{it}}{A_{it}}\right)^{\frac{1}{\alpha_M}} L_{it}^{-\frac{\alpha_L}{\alpha_M}} K_{it}^{-\frac{\gamma-\alpha_L-\alpha_M}{\alpha_M}}. \tag{3}$$

Marginal cost thus satisfies the following property:

$$MC_{it} \equiv \frac{\partial C_{it}}{\partial Q_{it}} = \frac{1}{\alpha_M}\frac{C_{it}}{Q_{it}}. \tag{4}$$

By combining equations (2), (3) and (4) one obtains that the markup $\mu_{it} \equiv P_{it}/MC_{it}$ can be computed, in line with De Loecker et al. (2016), as the ratio of the output elasticity of material to the share of materials' expenditure in revenue:

$$\mu_{it} = \frac{\alpha_M}{s_{Mit}}. \tag{5}$$

differences later on.

[4] To simplify notation we ignore components that are constant across firms in a given time period as they will be controlled for by time dummies.

[5] $\chi_{it} = \frac{W_{Mit}}{\alpha_M}Q_{it}^{\frac{1}{\alpha_M}-1}A_{it}^{-\frac{1}{\alpha_M}}L_{it}^{-\frac{\alpha_L}{\alpha_M}}K_{it}^{-\frac{\gamma-\alpha_L-\alpha_M}{\alpha_M}}$.

2.2 Demand heterogeneity

We consider a monopolistic competition framework[6] in which a representative consumer chooses among a continuum of differentiated varieties and maximises at each point in time t a differentiable utility function $U(.)$ subject to budget B_t:

$$\max_{Q} \left\{ U\left(\tilde{Q}\right) \right\} \text{ s.t. } \int_i P_{it} Q_{it} \mathrm{d}i - B_t = 0,$$

where $\tilde{Q}$ is a vector of elements $\Lambda_{it} Q_{it}$ and the vector Λ (with generic element Λ_{it}) is given to the representative consumer. Therefore, while the representative consumer chooses quantities Q while paying prices P, quantities Q enter into the utility function as $\tilde{Q}$ and Λ_{it} can be interpreted as a measure of the perceived quality/appeal of a particular variety i. For example, in the standard symmetric (with respect to $\tilde{Q}$) varieties case, the representative consumer would be indifferent between having one more unit of a variety i with $\Lambda_{it} = \overline{\Lambda}$ or $\overline{\Lambda}$ more units of a variety j with $\Lambda_{jt} = 1$.

Each firm chooses quantity to maximise profits, while taking Λ_{it} and market aggregates as given, so implying the usual relationship between the markup and demand elasticity ($\eta_{it} \equiv -\frac{\partial q_{it}}{\partial p_{it}}$):

$$\mu_{it} = \frac{\eta_{it}}{\eta_{it} - 1}. \tag{6}$$

At this point it is important to note that demand elasticity, the markup and the price are functions of both the quantity chosen by firm i (Q_{it}) and the perceived quality/appeal of its variety i (Λ_{it}). While quantity Q_{it} is directly observable in production data like ours, Λ_{it} is not, but can be inferred from the data based on a linear approximation result. First, in Appendix B we show that utility maximisation conditions imply:

$$\frac{\partial p_{it}}{\partial q_{it}} = -\frac{1}{\eta_{it}} = \frac{\partial p_{it}}{\partial \lambda_{it}} - 1, \tag{7}$$

where $\lambda_{it} = \log(\Lambda_{it})$. In other words, the elasticity of the price with respect to quantity differs from the elasticity of the price with respect to product appeal by one. The intuition behind (7) is straightforward. In practice, everything works as if firms were selling quality-adjusted quantities $\tilde{Q}_{it}$ while charging quality-adjusted prices $\tilde{P}_{it} = P_{it}/\Lambda_{it}$ so generating revenues $R_{it} = \tilde{P}_{it}\tilde{Q}_{it} = P_{it}Q_{it}$. Therefore, a change in Λ_{it} or Q_{it} should have the same impact on the

[6] In Appendix A, we show that our approach for demand heterogeneity, and in particular equations (8) and (9) below, applies in exactly the same way in the oligopolistic competition structure developed in Atkeson and Burstein (2008) and further refined in Hottman et al. (2016).

quality-adjusted price $\tilde{P}_{it}$: $\frac{\partial \tilde{p}_{it}}{\partial q_{it}} = \frac{\partial \tilde{p}_{it}}{\partial \lambda_{it}}$. In this respect noting that $\tilde{p}_{it} = p_{it} - \lambda_{it}$:

$$\frac{\partial \tilde{p}_{it}}{\partial q_{it}} = \frac{\partial p_{it}}{\partial q_{it}} - \frac{\partial \lambda_{it}}{\partial q_{it}} = \frac{\partial p_{it}}{\partial q_{it}},$$

because $\frac{\partial \lambda_{it}}{\partial q_{it}} = 0$. On the other hand:

$$\frac{\partial \tilde{p}_{it}}{\partial \lambda_{it}} = \frac{\partial p_{it}}{\partial \lambda_{it}} - \frac{\partial \lambda_{it}}{\partial \lambda_{it}} = \frac{\partial p_{it}}{\partial \lambda_{it}} - 1,$$

because $\frac{\partial \lambda_{it}}{\partial \lambda_{it}} = 1$. Combining the two above equations gives (7).

Moving forward, log revenue equals log price plus log quantity ($r_{it} = p_{it} + q_{it}$). Considering a first order linear approximation of log revenue around the profit maximising solution, as well as equations (6) and (7), we have:

$$r_{it}(q_{it}, \lambda_{it}) \simeq \frac{\partial r_{it}}{\partial q_{it}} q_{it} + \frac{\partial r_{it}}{\partial \lambda_{it}} \lambda_{it} = \left(\underbrace{\frac{\partial p_{it}}{\partial q_{it}}}_{-\frac{1}{\eta_{it}}} + \underbrace{\frac{\partial q_{it}}{\partial q_{it}}}_{+1} \right) q_{it} + \left(\underbrace{\frac{\partial p_{it}}{\partial \lambda_{it}}}_{-\frac{1}{\eta_{it}}+1} + \underbrace{\frac{\partial q_{it}}{\partial \lambda_{it}}}_{0} \right) \lambda_{it} = \underbrace{(1 - \frac{1}{\eta_{it}})}_{\frac{\eta_{it}-1}{\eta_{it}} = \frac{1}{\mu_{it}}} (q_{it} + \lambda_{it}),$$

and so to sum up:

$$r_{it} \simeq \frac{1}{\mu_{it}} (q_{it} + \lambda_{it}). \tag{8}$$

Equation (8) implies that using data on the actual quantity and revenue sold by firm i, which will be observable in our production data, and the profit-maximising markup μ_{it}, which we will measure using production function parameters and (5), one can recover the firm-specific demand measure λ_{it} as:

$$\lambda_{it} \simeq \mu_{it} r_{it} - q_{it}. \tag{9}$$

Two points are worth noting at this stage. First, if one considers the limit case of identical, in terms of quantity TFP, firms charging a common markup equal to one, equation (9) implies that $\lambda_{it} \simeq p_{it}$, i.e. our measure of product appeal/perceived quality corresponds to the firm price.[7] However, with heterogeneity in quantity TFP and markups, prices no longer reflect only underlying differences in product appeal/perceived quality. Second, considering the standard symmetric CES utility case, i.e. $U = \left(\int_{i \in I_t} (\Lambda_{it} Q_{it})^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right)^{\frac{\eta}{\eta-1}}$ where I_t denotes the

[7]Unit values obtained from international trade data have been indeed used as a measure of quality in a number of contributions including Schott (2008) and Bloom et al. (2020).

set of varieties, demand is given by:

$$Q_{it} = \Lambda_{it}^{\eta-1}\left(\frac{P_{it}}{P_t}\right)^{-\eta}\frac{B_t}{P_t},$$

where P_t is the usual CES price index and the elasticity of demand is constant and equal to η. From this expression, it appears how our measure of perceived quality/product appeal boils down to a firm-specific quantity shifter $\Lambda_{it}^{\eta-1}$ conditional on the individual price ${P_{it}}^{-\eta}$. Writing in logs, while getting rid of P_t and B_t that are constant across firms, we have:

$$q_{it} = (\eta-1)\lambda_{it} - \eta p_{it},$$

from which (given that the markup is constant and given by $\mu_{it} = \mu = \frac{\eta}{\eta-1}$):

$$\mu_{it} r_{it} - q_{it} = \frac{\eta}{\eta-1}\underbrace{((\eta-1)\lambda_{it}(1-\eta)p_{it})}_{r_{it}} - \underbrace{((\eta-1)\lambda_{it} - \eta p_{it})}_{q_{it}} = \lambda_{it},$$

i.e. equation (9) holds as an equality. When moving away from the CES case, the markup μ_{it} will vary across firms and firm revenue r_{it} needs to be 'weighted' by the firm-specific markup before subtracting quantity q_{it} to compute λ_{it} from (9). At the same time, equation (9) would now only hold as a linear approximation of the log revenue function. In this respect, we provide in Appendix B some examples of log revenue functions obtained from preferences used in the IO and trade literatures supporting the log-linear approximation.[8]

2.3 Revenue productivity

Being explicit about underlying differences in demand across firms not only allows measuring such differences, but also enables bridging the gap between revenue-based TFP and quantity-based TFP.[9] First note that the quantity equation (1) can be written as $q_{it} = \bar{q}_{it} + a_{it}$, where $\bar{q}_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it}$ is an index of inputs use that we label "scale" in the remainder. Second, by defining revenue-based TFP as $TFP_{it}^R \equiv r_{it} - \bar{q}_{it}$ and using equation (8) while

[8] Despite the fact that demand functions can differ wildly in terms of their shapes and properties, the corresponding revenue functions are reasonably log-linear in quantity. Intuitively, prices vary much less than quantities for firms and so changes in log revenue are essentially driven by changes in log quantity.

[9] The idea of decomposing revenue TFP into different underlying elements is not new and it is present in the literature since at least Klette and Griliches (1996). However, most decompositions have remained only theoretical in nature due to the lack of suitable data to operationalise them. The availability of price and quantity data, along with our assumptions on demand heterogeneity, allow us to actually implement our revenue TFP decomposition in the data.

substituting we get:[10]

$$TFP_{it}^R = \frac{a_{it} + \lambda_{it}}{\mu_{it}} + \frac{1 - \mu_{it}}{\mu_{it}} \bar{q}_{it}, \tag{10}$$

meaning that TFP_{it}^R is a (non-linear) function of quantity-based TFP a_{it}, product appeal λ_{it}, the markup μ_{it} and production scale $\bar{q}_{it}$.[11] (10) can also be made linear by considering markup-adjusted quantity TFP, product appeal and scale: $\tilde{a}_{it} = \frac{a_{it}}{\mu_{it}}$, $\tilde{\lambda}_{it} = \frac{\lambda_{it}}{\mu_{it}}$, $\tilde{\bar{q}}_{it} = \frac{(1-\mu_{it})\bar{q}_{it}}{\mu_{it}}$:

$$TFP_{it}^R = \tilde{a}_{it} + \tilde{\lambda}_{it} + \tilde{\bar{q}}_{it}. \tag{11}$$

As such, changes across time within a firm of TFP_{it}^R, or differences between firms at a given point in time in TFP_{it}^R can be decomposed as the sum of changes/differences in $\tilde{a}_{it}$, $\tilde{\lambda}_{it}$ and $\tilde{\bar{q}}_{it}$. This in turn enables gaining deeper and sharper insights into productivity questions. For example, we show below in Section 7 that changes in firm revenue productivity spurred by import competition from China materialise as the outcome of complex changes in quantity TFP, demand, markups and production scale. This in turns allows to better understand firm behaviour and margins of adjustment under competitive pressure and learn important lessons that can be applied to other contexts.

3 Closing the model

The last step to close the model involves introducing some additional assumptions. These are needed to estimate the parameters of the production function, which then are used to obtain quantity TFP a_{it} as a residual from the quantity equation (1), while markups are pinned down by (5) and product appeal is measured using (9). One readily available approach to estimate the production function, that is consistent with the underlying presence of heterogeneity in markups and demand, is provided in De Loecker et al. (2016). This methodology relies on

[10]From now onwards we use the equality, rather than the approximation, when employing equations (8) and (9). Given that later on we will be computing λ_{it} using (9), the equivalent equation (8) holds as an equality and so do the revenue TFP decompositions (10) and (11).

[11]Equation (10) indicates that revenue-based TFP increases with both quantity-based TFP and product appeal at a rate given by one over the profit-maximizing markup, which is the local slope of the log revenue function. The markup μ_{it} is to be considered endogenous with respect to both a_{it} and λ_{it} and in particular we show later on that markups appear to increase with both quantity-based TFP and product appeal. Therefore, as a_{it} and/or λ_{it} increases, so does the markup and this translates into smaller increases of TFP_{it}^R because the log revenue function gets flatter (see examples of log revenue functions in Appendix B). At the same time, increasing/decreasing scale decreases/increases revenue-based TFP (for given a_{it} and λ_{it}) simply because the firm is using more/less inputs. In this respect, it is important to note that scale is to be considered endogenous with respect to both a_{it} and λ_{it}. In particular, as a_{it} and/or λ_{it} increases, so does scale because the firm will be selling more and using more inputs. The negative indirect effect on revenue-based TFP is mediated by markups and gets stronger with higher μ_{it}.

the popular proxy variable approach.[12]

In what follows we depart from the proxy variable approach and propose a new estimation framework that builds upon our explicit assumptions about demand heterogeneity. We label this estimation framework FMMM.[13] The key insight of the FMMM estimation is that we use both the log revenue function (8) and the quantity equation (1) to recover technology parameters. This is in contrast with the common practice of using only the quantity equation and is made possible by the fact that we are sufficiently specific about demand to be able to explicitly write the log revenue function in terms of observables and the heterogeneities we allow for, as well as model parameters.

We thus end up with a simultaneous system of two equations. In order to deal with this we take advantage of two insights. First, as in Wooldridge (2009), we impose a simple AR(1) process for productivity meaning that we can substitute a_{it} with its time lag and an uncorrelated component while further replacing a_{it-1} with the lag of quantity and inputs from the quantity equation (1). We do something similar for product appeal by specifying again an AR(1) and replacing λ_{it-1} with components from the log revenue function (8). Second, as in Grieco et al. (2016), we exploit the first order conditions of the firm's profit maximization problem, and in particular the markup equation (5), to substitute for the output elasticity of material parameter α_M. Thanks to these two insights we are able to greatly simplify the estimation of the simultaneous system of two equations. In particular, the estimation can be performed separately and with simple linear methods (OLS for the revenue equation and IV for the quantity equation) without resorting to a more complex, and potentially numerically unstable, joint non-linear estimation. We provide Monte Carlo evidence of the performance of the FMMM estimation framework in Appendix F.

The time process of quantity TFP a_{it} is characterised by an AR(1):

$$a_{it} = \phi_a a_{it-1} + \nu_{ait} \tag{12}$$

where ν_{ait} are iid and uncorrelated with past values of productivity. As for product appeal,

[12]In particular, starting from the conditional input demand for materials, a number of observables (prices and market shares in particular) are added to proxy for unobservables (markups and demand heterogeneity in our framework) while imposing the usual assumption of invertibility.

[13]In order to allay concerns over the robustness of our results, we provide in Appendix D key results obtained employing the De Loecker et al. (2016) methodology (that we label DGKP) to estimate the parameters of the production function, and using such parameters to compute an alternative set of TFP, product appeal, markups and revenue TFP measures. We find qualitatively, and to a large extent also quantitatively, very similar results between the FMMM and DGKP approaches.

we also assume:[14]

$$\lambda_{it} = \phi_\lambda \lambda_{it-1} + \nu_{\lambda it} \tag{13}$$

where $\nu_{\lambda it}$ are iid and uncorrelated with past values of product appeal.[15] At the same time we allow ν_{ait} in (12) and $\nu_{\lambda it}$ in (13) to be correlated with each other. Note that, by allowing ν_{ait} and $\nu_{\lambda it}$ to be correlated with each other, we allow demand and productivity λ_{it} and a_{it} to be correlated with each other.[16]

We start by manipulating the log revenue function. More specifically, by substituting q_{it} with the formula of the Cobb-Douglas we can transform (8) further as:

$$r_{it} = \frac{\alpha_L}{\mu_{it}}(l_{it} - k_{it}) + \frac{\alpha_M}{\mu_{it}}(m_{it} - k_{it}) + \frac{\gamma}{\mu_{it}} k_{it} + \frac{1}{\mu_{it}}(a_{it} + \lambda_{it}).$$

Furthermore, by using the markup equation (5), we get:

$$LHS_{it} \equiv \frac{r_{it} - s_{Mit}(m_{it} - k_{it})}{s_{Mit}} = \frac{\alpha_L}{\alpha_M}(l_{it} - k_{it}) + \frac{\gamma}{\alpha_M} k_{it} + \frac{1}{\alpha_M}(a_{it} + \lambda_{it}). \tag{14}$$

where LHS_{it} is made out of observables only.

We then build upon our assumptions on the time processes for a_{it} and λ_{it}: (12) and (13). However, before substituting (12) and (13) into (14) we need to find a convenient way to express a_{it-1} and λ_{it-1}. By using (5) and (8) we have:

$$\lambda_{it-1} = r_{it-1}\mu_{it-1} - q_{it-1} = r_{it-1}\frac{\alpha_M}{s_{Mit-1}} - q_{it-1}. \tag{15}$$

At the same time plugging (15) into (14) and re-arranging yields:

$$a_{it-1} = \alpha_M LHS_{it-1} - \alpha_L(l_{it-1} - k_{it-1}) - \gamma k_{it-1} - \left(r_{it-1}\frac{\alpha_M}{s_{Mit-1}} - q_{it-1}\right). \tag{16}$$

[14] λ_{it} captures consumers' perception of the quality and appeal a firm's products; something that arguably does not change much from one year to another. It takes years of effort and costly investments for firms to establish their brand and build their customers' base on the one hand, and to put in place and develop an efficient production process for their products on the other. In our view, there are profound similarities between the processes of productivity and product appeal and so if the former can be approximated by a Markov process so can the latter.

[15] More precisely, we posit that a_{it} and λ_{it} are jointly described by a VAR(1) process meaning that ν_{ait} and $\nu_{\lambda it}$ are both uncorrelated with past values of productivity and product appeal.

[16] Table B-1 in Appendix B shows the estimated correlation between product appeal shocks ($\nu_{\lambda it}$) and TFP shocks (ν_{ait}). Our estimates are consistent with the idea that the strong measured negative correlation between the two shocks drives the strong measured negative correlation between λ_{it} and a_{it}.

Finally, by combining (12), (13), (15) and (16) into (14) we obtain:

$$
\begin{aligned}
LHS_{it} &= \frac{\gamma}{\alpha_M}k_{it} + \frac{\alpha_L}{\alpha_M}(l_{it} - k_{it}) + \phi_a LHS_{it-1} - \phi_a \frac{\gamma}{\alpha_M}k_{it-1} - \phi_a \frac{\alpha_L}{\alpha_M}(l_{it-1} - k_{it-1}) \\
&+ (\phi_\lambda - \phi_a)\left(\frac{r_{it-1}}{s_{Mit-1}} - \frac{q_{it-1}}{\alpha_M}\right) + \frac{1}{\alpha_M}(\nu_{ait} + \nu_{\lambda it}). \qquad (17)
\end{aligned}
$$

The advantage of the transformed log revenue function (17) with respect to the original formulation (8) is that we remove the unobservable (to the econometrician) markups and product appeal. Indeed, besides the idiosyncratic productivity and demand shocks ν_{ait} and $\nu_{\lambda it}$, the transformed log revenue function (17) contains only observables and useful parameters. There are various ways of estimating (17) and here we use perhaps the simplest one. More specifically, we rewrite (17) as the following linear regression:

$$LHS_{it} = b_1 z_{1it} + b_2 z_{2it} + b_3 z_{3it} + b_4 z_{4it} + b_5 z_{5it} + b_6 z_{6it} + b_7 z_{7it} + u_{it}, \qquad (18)$$

where $z_{1it}{=}k_{it}$, $z_{2it}{=}(l_{it} - k_{it})$, $z_{3it}{=}LHS_{it-1}$, $z_{4it}{=}k_{it-1}$, $z_{5it}{=}(l_{it-1} - k_{it-1})$, $z_{6it}{=}\frac{r_{it-1}}{s_{Mit-1}}$, $z_{7it}{=}q_{it-1}$, $u_{it}{=}\frac{1}{\alpha_M}(\nu_{ait} + \nu_{\lambda it})$ as well as $b_1{=}\frac{\gamma}{\alpha_M}$, $b_2{=}\frac{\alpha_L}{\alpha_M}$, $b_3{=}\phi_a$, $b_4{=}-\phi_a\frac{\gamma}{\alpha_M}$, $b_5{=}-\phi_a\frac{\alpha_L}{\alpha_M}$, $b_6{=}(\phi_\lambda - \phi_a)$, $b_7{=}-(\phi_\lambda - \phi_a)\frac{1}{\alpha_M}$. Given our assumptions, the error term u_{it} in (18) is uncorrelated with current capital and labour as well as with lagged inputs use, quantity and revenue.[17] Therefore, z_{1it} to z_{7it} are uncorrelated to u_{it} and (18) can be estimated by OLS. Operationally, we augment (18) with a full battery of 8-digit product dummies, as well as year dummies, and perform estimations separately for each two-digit industry. We then set $\widehat{\frac{\gamma}{\alpha_M}}{=}\hat{b}_1$, $\widehat{\frac{\alpha_L}{\alpha_M}}{=}\hat{b}_2$ and $\hat{\phi}_a{=}\hat{b}_3$ without exploiting parameters' constraints in the estimation.[18]

We now turn to estimating γ from the quantity equation in a second step. Combining the

[17] Our assumptions are in line with the widespread practice in the literature (Olley and Pakes, 1996; Wooldridge, 2009; De Loecker et al., 2016) of imposing that productivity follows a Markov process (to which we add that also product appeal follows a Markov process) while at the same time exploiting the idea that some factors of production (capital and labour in our case) are predetermined in t while others (materials in our case) are endogenous in t. From these assumptions it immediately follows that contemporaneous and lagged values (lagged values) of predetermined (endogenous) variables are uncorrelated with innovations in t. If one allows labour to be a semi-flexible input then labour becomes an endogenous variable and so $E\{\nu_{ait}l_{it}\} = E\{\nu_{\lambda it}l_{it}\} = 0$ will not hold. In this case, OLS cannot be applied to equation (18) because z_{2it} is endogenous. However, z_{2it} could be instrumented with, for example, z_{2it-2} as well as the lags of order 2 of materials, capital, revenue and quantity.

[18] In principle, parameters' constraints in equation (18) could be used to obtain an estimate of α_M, and more specifically the ratio $-b_6/b_7$, and so fully recover technology parameters without further need of the quantity equation. However, this hinges on the difference between ϕ_λ and ϕ_a being significantly different from zero, i.e., on parameters b_6 and b_7 being significantly different from zero. Operationally, we very often find that b_6 and b_7 are not significantly different from zero. This is consistent with the estimates of ϕ_a and ϕ_λ (very close to each other in each of the 9 industries we consider) obtained from our two-step procedure and reported in Table B-2 in Appendix B.

quantity equation (1) and the markup equation (5) we have:

$$q_{it} = \mu_{it} s_{Mit} \left(m_{it} - k_{it}\right) + \alpha_L \left(l_{it} - k_{it}\right) + \gamma k_{it} + a_{it}. \tag{19}$$

Further using $\alpha_M = \frac{\gamma}{b_1}$ as well as $\alpha_L = \frac{\gamma b_2}{b_1}$ and we get:

$$q_{it} = \frac{\gamma}{\hat{b}_1} \left(m_{it} - k_{it}\right) + \frac{\gamma \hat{b}_2}{\hat{b}_1} \left(l_{it} - k_{it}\right) + \gamma k_{it} + a_{it}, \tag{20}$$

where we replace b_1 and b_2 with their estimates $\hat{b}_1$ and $\hat{b}_2$ coming from (18). Finally, using the autoregressive formulation (12) as well as (16), we substitute a_{it} with observables and idiosyncratic productivity shocks ν_{ait} and obtain:

$$\begin{aligned} q_{it} &= \frac{\gamma}{\hat{b}_1} \left(m_{it} - k_{it}\right) + \frac{\gamma \hat{b}_2}{\hat{b}_1} \left(l_{it} - k_{it}\right) + \gamma k_{it} + \gamma \frac{\hat{\phi}_a}{\hat{b}_1} LHS_{it-1} - \frac{\gamma \hat{b}_2 \hat{\phi}_a}{\hat{b}_1} \left(l_{it-1} - k_{it-1}\right) \\ &- \gamma \hat{\phi}_a k_{it-1} - \hat{\phi}_a \left(r_{it-1} \frac{\gamma}{\hat{b}_1 s_{Mit-1}} - q_{it-1} \right) + \nu_{ait}. \end{aligned} \tag{21}$$

Note that the only unobservable in (21) is the idiosyncratic productivity shock ν_{ait} while the only parameter left to identify is γ. We can more compactly write (21) as the following linear regression:

$$\overline{LHS}_{it} = b_8 z_{8it} + \nu_{ait} \tag{22}$$

where:

$$\begin{aligned} \overline{LHS}_{it} &= q_{it} - \hat{\phi}_a q_{it-1} \\ z_{8it} &= \frac{1}{\hat{b}_1} \left(m_{it} - k_{it}\right) + \frac{\hat{b}_2}{\hat{b}_1} \left(l_{it} - k_{it}\right) + k_{it} + \frac{\hat{\phi}_a}{\hat{b}_1} LHS_{it-1} \\ &- \frac{\hat{b}_2 \hat{\phi}_a}{\hat{b}_1} \left(l_{it-1} - k_{it-1}\right) - \hat{\phi}_a k_{it-1} - \frac{\hat{\phi}_a r_{it-1}}{\hat{b}_1 s_{Mit-1}} \end{aligned}$$

as well as $b_8 = \gamma$. Materials m_{it} in z_{8it} are endogenous to contemporaneous productivity shocks ν_{ait} and so OLS cannot be applied to estimate (22). However, one can use several moment conditions for identification including $E\left\{\nu_{ait} k_{it}\right\} = E\left\{\nu_{ait} l_{it}\right\} = E\left\{\nu_{ait} l_{it-1}\right\} = E\left\{\nu_{ait} m_{it-1}\right\} = E\left\{\nu_{ait} k_{it-1}\right\} = E\left\{\nu_{ait} q_{it-1}\right\} = E\left\{\nu_{ait} r_{it-1}\right\} = 0$.[19] Again, we augment (22) with year and 8-digit product dummies and perform estimations separately for each two-digit industry. The

[19] If one allows labour to be a semi-flexible input then $E\left\{\nu_{ait} l_{it}\right\} = 0$ will not hold. However, all of the other moment conditions will hold and there are enough to choose from. In particular, we use in our estimations $E\left\{\nu_{ait} l_{it-1}\right\} = E\left\{\nu_{ait} m_{it-1}\right\} = E\left\{\nu_{ait} k_{it-1}\right\} = E\left\{\nu_{ait} q_{it-1}\right\} = E\left\{\nu_{ait} r_{it-1}\right\} = 0$.

IV estimation of (22) provides an estimate of γ that, together with $\widehat{\frac{\gamma}{\alpha_M}}$ and $\widehat{\frac{\alpha_L}{\alpha_M}}$ coming from the first stage revenue equation, uniquely delivers production function parameters ($\hat{\alpha}_L$, $\hat{\alpha}_M$ and $\hat{\gamma}$).

Last but not least, for all our estimates and results we correct for the presence of measurement error in output (quantity and revenue) and/or unanticipated (to the firm) productivity shocks using the methodology described in De Loecker et al. (2016) and on which we provide some highlights in Appendix B.

4 Data, descriptives and additional variables

4.1 Basic data

Our primary data consists of firm-level production data for Belgian manufacturing firms coming from the Prodcom database and provided by the National Bank of Belgium. Prodcom is a monthly survey of industrial production established by Eurostat for all EU countries in order to improve the comparability of production statistics across the EU by the use of a common product nomenclature called Prodcom (8-digit codes whose first four digits come from NACE codes). Prodcom covers production of broad sectors C and D of NACE Rev. 1.1 (Mining and quarrying and manufacturing), except for sections 10 (Mining of coal and lignite), 11 (Extraction of crude petroleum and natural gas) and 23 (Manufacture of coke and refined petroleum products). During our sample period, each Belgian firm with 10 employees or more - or with a revenue greater than a certain threshold in a given year - had to fill out the survey.[20] Firms in the survey cover more than 90% of Belgian manufacturing production and the raw data is aggregated from the plant-level to the firm-level.

This gives us a sample of about 6,000 firms a year over the period of 1995 to 2007. Data is organised by product-year-month-firm. We use information on quantity (the unit of measurement depending on the specific product) and value (Euros) of production sold. We aggregate the data at the firm-year-product level. The same data has been previously used in Bernard et al. (2018) in their analysis of carry along trade as well as by De Loecker et al. (2014) for their study of the links between international competition and firm performance. There are about 4,500 distinct 8-digit products within the Prodcom classification. The level of detail is such that, for example, we are able within the "Meat and Meat products" industry (NACE code 151) to look at specific products such as "Sausages not of liver" (Prodcom code 15131215) and "Fresh or chilled cuts of geese; ducks and guinea fowls" (Prodcom code 15121157) while

[20]Rules are somewhat different for other EU countries. In particular some EU countries only surveyed firms with 20 or more employees. The 10 employees threshold has been recently increased to 20 in Belgium as well.

within the NACE 212 "Article of Papers" industry, we can distinguish between "Envelopes of paper or paperboard" (Prodcom code 21231230) and "Wallpaper and other wall coverings; window transparencies of paper" (Prodcom code 21241190). While not as detailed as product categories available with bar-code data for retail products, our dataset has the advantage of spanning across the entire manufacturing sector.

We also make use of more standard balance sheet data to get information on firms' inputs. We build on annual firm accounts from the National Bank of Belgium. For this study, we selected those companies that filed a full-format or abbreviated balance sheet between 1996 and 2007 and with at least one full-time equivalent employee. Variables include FTE employment, total wage bill (our preferred measure of the labour input), material costs, capital stock and turnover. There are more than 15,000 manufacturing firms per year displaying non-missing values for these variables.

Besides, we use standard EU-type micro trade data at the product-country-firm-month level over the period 1995-2008 provided by the National Bank of Belgium. From this data we simple borrow information on firm import and export status. The combined balance sheet and trade data has been previously used in Behrens et al. (2013), Mion and Zhu (2013) and Muûls (2015) among others and is representative of the Belgian economy. The three datasets are matched by the unique firm VAT identifier.

4.2 Additional variables and descriptives

In the analysis of the impact of import competition from China on revenue productivity and its components, that we report in Section 7, we further use additional trade and import quota data. The trade data comes from the Comtrade database provided by the United Nations. We use EU-15 and US imports data over the period 1995 to 2007 at the HS6-digit level to construct a measure of Chinese imports penetration in these two markets. We first build on a concordance between the HS6-digit classification and the CPA6-digit classification, where the latter dictates the first 6 digits of Prodcom codes, to measure imports at the CPA6-exporting country-year level.[21] We then construct the following measure of Chinese imports penetration, in either the EU15 or the US market, for each CPA6-digit product and time t:

$$IPC^{mkt}_{CPA6,t} = \frac{IMP^{mkt}_{CPA6,China,t}}{\sum_c IMP^{mkt}_{CPA6,c,t}}. \tag{23}$$

[21]The concordance between HS6 and CPA6 is quite straightforward and we have used suitable tables provided by the RAMON EU website. The same does not apply to the concordance between HS6 and Prodcom 8-digit. This is the reason why we have decided to work at the CPA6 level. The CPA 6-digit still represents a very detailed breakdown of products. For example, there are 1,370 distinct CPA6 products corresponding to about 4,500 Prodcom 8-digit products.

where $IMP^{mkt}_{CPA6,c,t}$ are imports by market $mkt = \{EU15, US\}$ of products belonging to a specific $CPA6$ product from country c at time t while $IMP^{mkt}_{CPA6,China,t}$ represents imports from China by either the EU15 or the US of products belonging to a specific $CPA6$ product at time t. A similar measure has been used in Mion and Zhu (2013) for Belgium, as well as by a number studies for other countries, for the analysis of the various economic impacts of import competition from China. In this respect, we believe that the EU15 is the relevant market for the highly export-oriented Belgian firms over our time frame. Therefore, $IPC^{EU15}_{CPA6,t}$ is our preferred measure of the import competition faced by Belgian firms producing products belonging to a specific $CPA6$ code at time t. At the same time we allow, as in Autor et al. (2013), for the presence of unobserved demand/technology shocks at the $CPA6$-time level characterising the EU15 market, and correlated with $IPC^{EU15}_{CPA6,t}$, by instrumenting $IPC^{EU15}_{CPA6,t}$ with $IPC^{US}_{CPA6,t}$. More specifically, we match in each year firms to our $CPA6$-time measure[22] and regress firm revenue productivity on $IPC^{EU15}_{CPA6,t}$. While allowing for firm fixed effects and time dummies, we subsequently instrument $IPC^{EU15}_{CPA6,t}$ with $IPC^{US}_{CPA6,t}$. We then do the same for the different components of revenue productivity (a_{it}, λ_{it}, etc.).

As a complementary attempt to identify the impact of import competition from China on revenue productivity and its components we focus on a specific industry ("Textile and Apparel") and exploit detailed HS6-level information on import quotas. These quotas were imposed at the EU-level on Chinese imports, as well as on imports from other non-WTO countries, and affected some products within the industry but not others. As a consequence of China joining the WTO, these quotas were removed during our period of analysis. To provide some context, when these quotas were abolished this generated a 240% increase in Chinese imports on average within the affected product groups. The data and estimation strategy are borrowed from Bloom et al. (2016). We compute, for each 6-digit CPA product category, the proportion of 6-digit HS products that were covered by a quota, weighting each HS6 product by its share of EU15 imports value computed over the period 1995-1997. We label this $QUOTA_{CPA6}$ and focus on the period 1998-2007 to analyse firm behaviour. More specifically, we match in each year firms to $QUOTA_{CPA6}$ and run a simple diff-in-diff specification where the time-change in, for example, firm revenue productivity is regressed on $QUOTA_{CPA6}$. We then do the same for the different components of revenue productivity.

With the exception of the quota analysis where we consider the sub-sample 1998-2007, we focus our study on the period 1996-2007 for which all datasets are available and during which there has not been any major change in data collection and data nomenclatures. [23]

[22]As explained below we focus in our analysis on firm producing a single Prodcom 8-digit product. Therefore, we match firms to $IPC^{EU15}_{CPA6,t}$ based on the first 6 digits of the unique Prodcom 8-digit product produced by firm i at time t.

[23]The NACE and the CN nomenclatures changed considerably in 2008. Note however that, as reported in

We choose not to analyse multi-product firms in this paper and focus on single-product firms, i.e. firms producing a single 8-digit Prodcom product. This is for a variety of reasons. First, we could include multi-product firms in our analysis by drawing on the extension of the MULAMA model presented in Appendix C. Yet, in doing so we would not be able to tell whether our findings are specific to the MULAMA model or whether they are driven by the additional assumptions we need to make in order to solve the inputs assignment problem for multi-product firms. Second, in order to estimate the MULAMA model on multi-product firms we first need, as in De Loecker et al. (2016), to estimate the parameters of the production function for single-product firms: these are more than half of firm-year pairs in the data. Third, in focusing on single-product firms we improve upon previous analyses of the impact of Chinese imports competition by being able to very precisely match a measure of import competition with what the firm actually produces. For example, Mion and Zhu (2013) and Bloom et al. (2016) combine information on the primary 4-digit industry code of a firm with some of the available secondary codes to account for the fact that some firms are active in many industries while the measure they use is specific to one industry. In this respect our information is more detailed, at the 6-digit CPA, and at the same time the CPA-specific measure of Chinese import competition we attach to a firm covers all of its production.

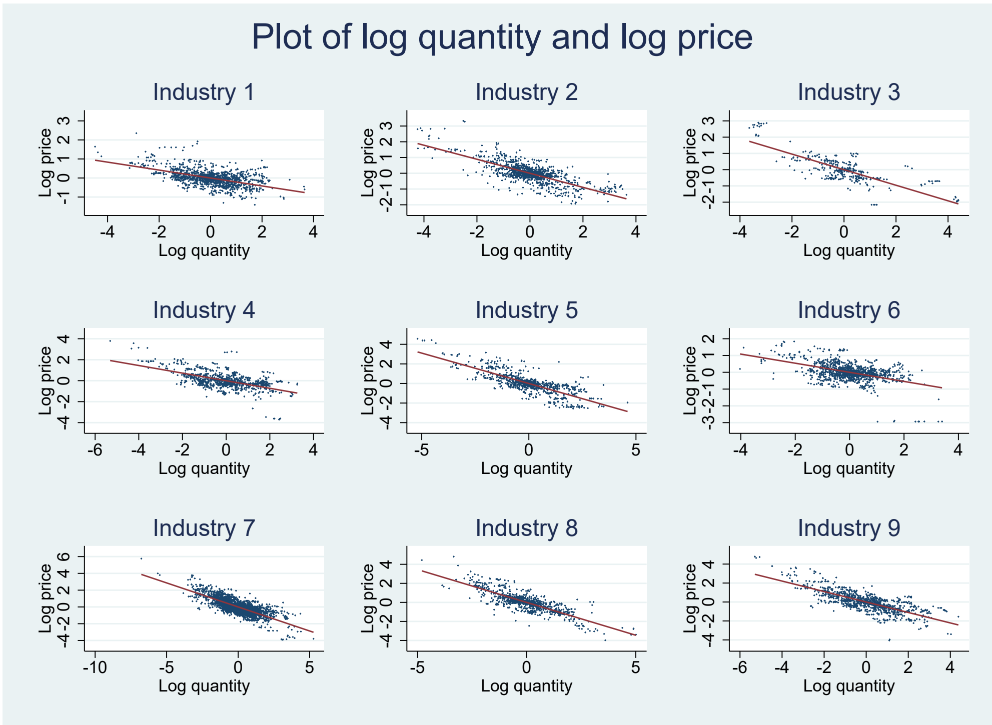
As in previous studies using either revenue or quantity data our estimations are run at a more aggregate level, that we label "industry" g, rather than at the finest available classification (8-digit products). This is needed to have a sufficiently large number of observations to estimate production function parameters in a consistent way. We estimate production function parameters separately for each industry by pooling together firms producing 8-digit products belonging to a given industry while at the same time adding time and 8-digit product dummies. This implies assuming that technology parameters are the same across products within an industry. Yet, we use actual quantities (and revenues) corresponding to the specific 8-digit product produced by a firm.

In terms of data cleaning, besides getting rid of missing and/or inconsistent observations, we exclude from the analysis firms that in a given year report different sales (± 15%) in the Prodcom and the Balance sheet data. Indeed, for some firms, manufacturing is only one part

Bernard et al. (2018), there have been some minor changes in 8-digit level Prodcom codes during our sample period. The first 6 digits of Prodcom codes have remained virtually unchanged from 1996 to 2007 because they correspond to the CPA classification and this has barely changed over our sample period. Therefore, most code changes involved the last two digits of Prodcom product codes. Whenever a new code is introduced the old code is not re-assigned to another product. Rather than attempting a complicated, and potentially imprecise, Prodcom 8-digit time concordance exercise we have decided to use the original codes: if a product changes code in year t, we will be using two dummies in the estimations; one for the code prior to t and one for the code from t. The same principle applies to one-to-many, many-to-one and many-to-many codes changes. Albeit conservative, we believe this is the best solution in this case.

of their operations. We then apply a 1% top and bottom trimming based on the following variables: (i) value added over revenue; (ii) materials expenditure over revenue; (iii) capital stock over labour expenditure; (iv) price within an 8-digit product.

Figure 1: The importance of heterogeneity in demand across firms



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict log quantity and log price, corresponding to each firm-year pair in our sample, and are constructed after demeaning both log quantity and log price by 8-digit product codes.

Table B-3 in Appendix B provides our industry breakdown as well as some basic summary statistics (mean, standard deviation, 5th and 95th percentiles) and the number of observations for the estimation sample. Figure 1 plots, after demeaning for each 8-digit product code, the log quantity and log price corresponding to each firm-year pair in our sample. We provide a plot for each industry and, as one can appreciate, there is indeed a negative correlation (within 8-digit products) between prices and quantities. However, the correlation is far from perfect with many instances of firms selling higher quantities than others for the same price and vice versa. These differences are substantial: prices and quantities in Figure 1 are in log units. Overall, this points to a fair amount of heterogeneity in demand across firms in the data.

5 Estimation results

In this Section we provide a number of results including our estimations, our measures of TFP, product appeal and markups and examine how the three dimensions of heterogeneity correlate with each other.

Table 1 provides estimates of Cobb-Douglas production function parameters obtained with the FMMM estimation procedure.[24] Coefficients are in line with expectations for a 3 inputs production function and there seems to be overall some support for constant returns to scale ($\gamma = 1$). Capital coefficients are on the low side, as is usually the case in the literature, likely due to measurement error particularly plaguing this variable, and they look very similar to those reported in De Loecker et al. (2016) using quantity data for India.[25]

Table 1: Estimates of the Cobb-Douglas production function parameters obtained with the FMMM procedure

Industry	Description	Labour	Materials	Capital	γ
1	Food products, beverages and tobacco	0.397^a	0.728^a	0.045^a	1.169^a
		(0.029)	(0.040)	(0.014)	(0.061)
2	Textiles and leather	0.325^a	0.636^a	0.020^c	0.981^a
		(0.020)	(0.019)	(0.012)	(0.014)
3	Wood except furniture	0.340^a	0.632^a	0.026	0.998^a
		(0.050)	(0.049)	(0.021)	(0.058)
4	Pulp, paper, publishing and printing	0.427^a	0.629^a	0.070^a	0.986^a
		(0.065)	(0.092)	(0.017)	(0.141)
5	Chemicals and rubber	0.328^a	0.648^a	0.034^c	1.010^a
		(0.040)	(0.052)	(0.019)	(0.071)
6	Other non-metallic mineral products	0.316^a	0.622^a	0.047^a	0.985^a
		(0.039)	(0.051)	(0.015)	(0.078)
7	Basic metals and fabric. metal prod.	0.338^a	0.629^a	0.024^a	0.991^a
		(0.015)	(0.012)	(0.008)	(0.005)
8	Machinery, electric. and optical equip.	0.347^a	0.630^a	0.026^b	1.004^a
		(0.033)	(0.023)	(0.011)	(0.008)
9	Transport equipment and n.e.c.	0.313^a	0.636^a	0.025	0.974^a
		(0.032)	(0.031)	(0.016)	(0.039)

Notes: γ denotes returns to scale. Bootstrapped standard errors in parenthesis (200 replications).
^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

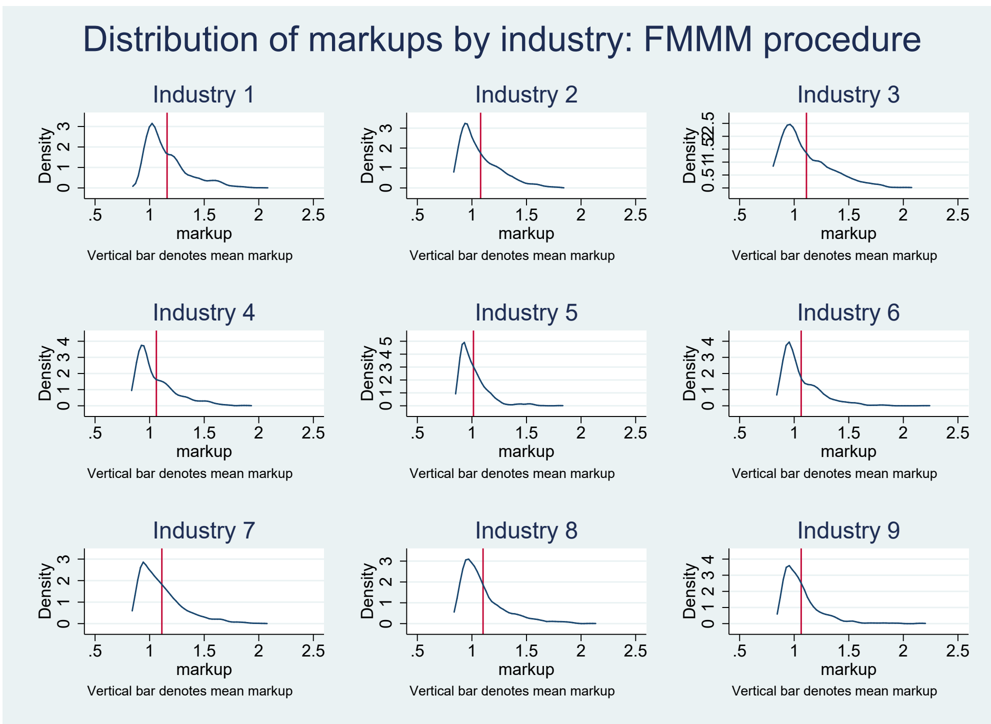
Moving to markups, the average across all observations is 1.091 which is in line with

[24]Table D-1 in Appendix D provides estimates of the mean and standard deviation of output elasticities for the Translog production function obtained with the DGKP estimation procedure. Translog production function coefficients estimates are provided in Table D-2 of Appendix D. The correlation between quantity TFP obtained with the two estimation procedures is extremely high: 0.998 across all observations and 0.987 once demeaning both TFP measures by 8-digit product codes. See Appendix D for additional comparisons of results obtained with the two estimation procedures.

[25]It is important to note that we estimate a three-inputs (capital, labour and materials) production function instead of the somewhat more common two-inputs (capital and labour with value added as a measure of output) production function. In the latter case the literature typically finds output elasticities for capital around 0.2-0.3 (Olley and Pakes, 1996; Levinsohn and Petrin, 2003). However, in the former case, the literature typically finds output elasticities for capital around 0.05-0.2 when using either quantity or revenue as a measure of output (Harris and Robinson, 2003; De Loecker et al., 2016; Pozzi and Schivardi, 2016).

numbers reported in, for example, De Loecker and Warzynski (2012). We further provide the density distribution of markups across observations separately for each industry in Figure 2 along with the corresponding mean (red vertical line). In this respect, Figure 2 points out how markups vary considerably across firms within each industry. As far as product appeal is concerned averages are, as in the case of quantity TFP, of little value. What is interesting is the variation in its values. In this respect, Table 2 reports the standard deviation of 8-digit product code demeaned values of a_{it} and λ_{it}, as well as raw values of μ_{it}.

Figure 2: Distribution of markups by industry



Notes: The figure provides the density distribution of markups across observations separately for each of the 9 industries we consider. The vertical bar denotes the mean markup.

The key finding stemming from Table 2 is that, within products, there is as much variation in product appeal as there is variation in quantity TFP, confirming the first-hand impression stemming from raw data on prices and quantities plotted in Figure 1. This suggests that heterogeneity in product appeal is a key component of firm idiosyncracies, at least as sizeable as heterogeneity in productivity, and so a potentially powerful key to unlock patterns in the data.

Moving to correlations, Table 3 provides correlations between quantity TFP, product appeal, markups and log prices. Again, we demean quantity TFP, product appeal and log prices because these measures do not compare much across 8-digit products. The main feature emerging from Table 3 is the strong negative correlation between quantity TFP and product

Table 2: Standard deviation of quantity TFP, product appeal and markups by industry

Industry	Description	TFP	product appeal	markups
1	Food products, beverages and tobacco	0.416	0.477	0.154
2	Textiles and leather	0.604	0.671	0.130
3	Wood except furniture	0.843	0.914	0.180
4	Pulp, paper, publishing and printing	0.775	0.843	0.152
5	Chemicals and rubber	0.952	0.970	0.079
6	Other non-metallic mineral products	0.520	0.607	0.123
7	Basic metals and fabric. metal prod.	0.860	0.896	0.169
8	Machinery, electric. and optical equip.	0.917	0.925	0.139
9	Transport equipment and n.e.c.	1.021	1.020	0.151

Notes: TFP and product appeal are demeaned by 8-digit Prodcom codes.

Table 3: Correlations between quantity TFP, product appeal, markups and log prices

	TFP	λ	markups	prices
TFP	1			
λ	-0.968^a	1		
markups	-0.079^a	0.173^a	1	
prices	-0.994^a	0.967^a	0.081^a	1

Notes: quantity TFP, product appeal and (log) prices are demeaned by 8-digit Prodcom codes. ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

appeal. This is robust to refining the correlation analysis by industry that we accomplish in Figure 3. Indeed, Figure 3 shows a strong within-product negative correlation between a_{it} and λ_{it} in each of the 9 industries we consider. More specifically, the correlation ranges between -0.868 and -0.988 across industries and it is always significant at the 1% level. This is also robust to using the DGKP estimation procedure, as shown by Figure D-2 in Appendix D, as well as to narrowing down the analysis to an allegedly homogeneous product like 'Ready-mixed concrete' as shown by Figure 4 below.[26] Furthermore, Table B-1 in Appendix B

[26]Contrary to common perception, ready-mixed concrete is far from being a homogeneous product. Indeed, there are many varieties of ready-mixed concrete that are differentiated both horizontally and vertically. For example, there is home-use concrete, farm-use concrete, coated concrete (specific for corrosion resistance), piling concrete (specific for pillars), early-strength concrete (useful in the first stages of construction), fibre-reinforced concrete (for higher durability), heat-conducting concrete, underwater concrete, etc. Besides differences in physical properties and attributes, concrete sold by different firms is, like any product, also differentiated in terms of elements such as the capacity to be delivered on time, the quality of customer service,

does suggest that the strong negative correlation between λ_{it} and a_{it} is driven by a strong negative correlation between product appeal shocks ($\nu_{\lambda it}$) and TFP shocks (ν_{ait}). At the same time, Table B-4 in Appendix B shows that a strong negative correlation also applies when considering (time) first-differences of product appeal and TFP so indicating that the correlation between λ_{it} and a_{it} is not simply driven by some level effects.

These results, which are in line with findings from the demand systems literature (Ackerberg et al., 2007), can be rationalised in several ways.[27] For example, one could reasonably argue that technology is such that higher quality products require more and/or more expensive inputs, i.e. lower quantity TFP. On the other hand, even if a_{it} and λ_{it} were uncorrelated from a technology point of view, a negative correlation between the two will arise after selection has taken place and only firms with high enough a and/or high enough λ survive. In Appendix E we focus on the selection mechanism and provide highlights of a Melitz-type model, based on Behrens et al. (2014), that is in line with the MULAMA framework and that delivers a negative correlation between quantity TFP and product appeal.[28]

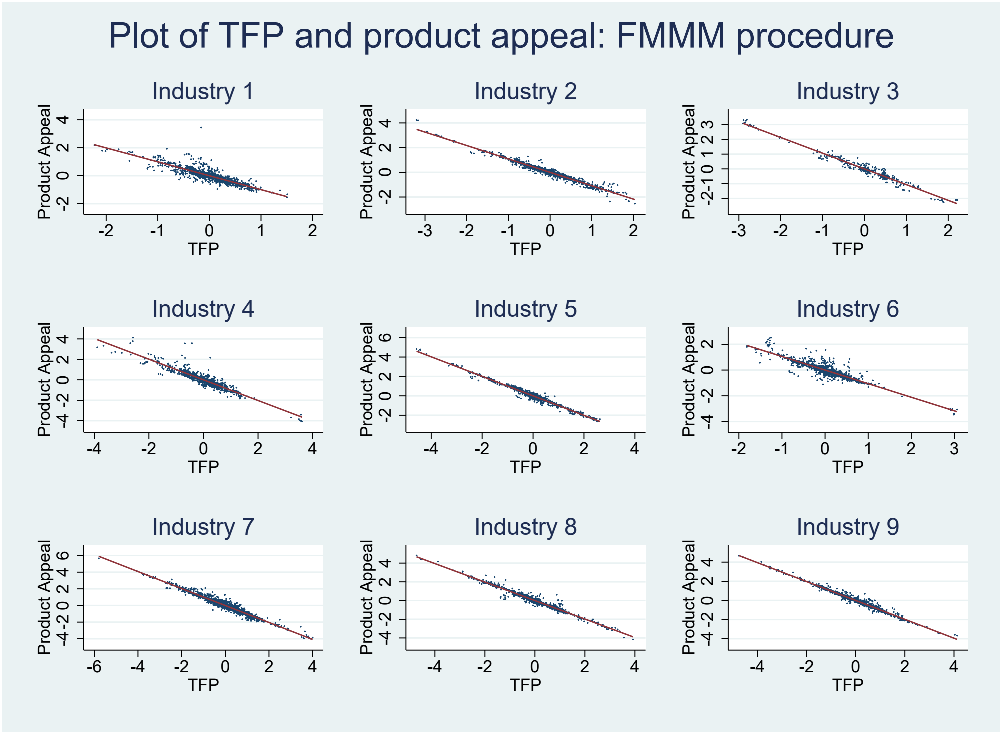
To provide further insights and intuition about the negative correlation between TFP and product appeal, as well as about why it is important to observe both measures, we find the following real world example useful. One of the most productive car plants in Europe is the Nissan factory located in Sunderland in the UK. In terms of sheer productivity, measured as cars per employee, it is nearly 100% more productive than a state of the art Mercedes plant near Rastatt in Germany. However, this hardly reflects a problem with the Mercedes plant. Rather, Mercedes and Nissan face very different demands from their clients leading to products of different quality and characteristics. Both plants are profitable and perhaps generate a very similar revenue productivity. However, their business model is different – Nissan is high a and low λ while Mercedes is low a and high λ – and with data on quantity and prices, along with our framework, we are actually able to distinguish their business models. Furthermore, it is important to distinguish product appeal from quantity TFP because Nissan and Mercedes could behave very differently when facing the same shock precisely because of their different

the availability of different financing solutions, etc.

[27]Jaumandreu and Yin (2017) provide a framework allowing for the presence of correlated demand and TFP heterogeneity. However, they do not observe quantities and so lay down a number of assumptions under which the two heterogeneities can be recovered from standard revenue and inputs data as well as data on demand shifters. Interestingly, they also find TFP and demand heterogeneity to be negatively correlated. Our approach requires more data but less assumptions and provides, among others, what we believe is a more direct and compelling evidence about the negative correlation between TFP and demand heterogeneity.

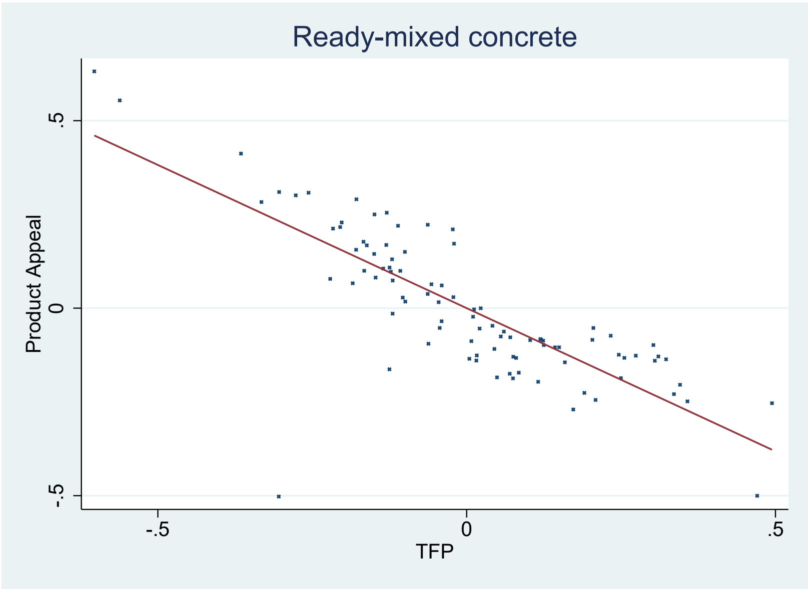
[28]More specifically, upon paying a fixed cost firms take a random draw in the (a, λ) space and only firms with 'quality adjusted' productivity $a+\lambda$ high enough are able to survive. Therefore, even if TFP and product appeal draws are independent, a negative correlation will arise in the data because only surviving firms are observed. Intuitively, when comparing two sub-samples of surviving firms with different average product appeal λ, the group with the higher average λ needs less higher values of a to satisfy the cutoff survival rule so generating the negative correlation.

Figure 3: Within 8-digit products correlation between quantity TFP and product appeal by industry



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict quantity TFP and product appeal, corresponding to each firm-year pair in our sample, and are constructed after demeaning both quantity TFP and product appeal by 8-digit product codes.

Figure 4: Correlation between quantity TFP and product appeal for ready-mixed concrete producers



Notes: The figure provides a scatter plot (as well as a best fit line) depicting quantity TFP and product appeal corresponding to each firm-year pair in the sub-sample of producers of ready-mixed concrete (Prodcom code 26631000). Both quantity TFP and product appeal have been demeaned using the corresponding averages.

business model. For example, one could assume that the availability and customisability of intermediate components is key for producing high quality cars. In that case, fostering stable long-term relationships with input suppliers, Mercedes would be less responsive than Nissan in terms of changing its sourcing strategies in the wake of changes to international trade costs.

Going back to Table 3, there are other strong correlations emerging but, in order to obtain more structured insights, one needs to go beyond pairwise correlations. This is achieved in Tables 4 and 5 where we look more in detail at prices and markups. Considering prices in Table 4 we should, in light of our assumptions and theoretical results, expect prices to depend upon both quantity and product appeal. More specifically we expect, given equation (7), that for a given product appeal λ, log prices should be decreasing with log quantity at a rate given by the inverse of the elasticity of demand. Furthermore, the elasticity of prices with respect to λ should be equal to one minus the inverse of the elasticity of demand. In the first column of Table 4, we regress log prices on log quantities and λ across all observations while demeaning all variables by 8-digit Prodcom codes and adding year dummies. In order to have more of a causal interpretation of parameters, we estimate the regression in first-differences, i.e, we regress changes in log prices on changes in log quantity and changes in λ. The parameter corresponding to log quantity is negative and significant and indicates an average, across firms and products, price elasticity of demand of 4.149=1/0.241. Furthermore, the coefficient corresponding to λ is also highly significant and roughly equivalent to one plus the coefficient corresponding to log quantity.

Table 4: Analysis of demeaned log prices

quantity	-0.241^a	
	(0.035)	
TFP		-0.887^a
		(0.009)
λ	0.700^a	0.105^a
	(0.040)	(0.009)
capital		-0.005^c
		(0.002)
Estimation method	First differences	
Year dummies	Yes	Yes
N Obs	7,768	7,768
R^2	0.940	0.993

Notes: The dependent variable is the demeaned log price. (Log) prices, quantity TFP, product appeal, (log) capital and (log) quantity are demeaned by 8-digit Prodcom codes. Bootstrapped standard errors in parenthesis (200 replications). ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Table 5: Analysis of markups

quantity	0.116^a (0.010)	
TFP		0.372^a (0.026)
λ	0.075^a (0.013)	0.371^a (0.026)
capital		0.004 (.003)
Estimation method	First differences	
Year dummies	Yes	Yes
Prod dummies	Yes	Yes
N Obs	7,768	7,768
R^2	0.218	0.397

Notes: The dependent variable is the markup. Quantity and capital are expressed in logs while TFP indicated quantity TFP. Bootstrapped standard errors in parenthesis (200 replications). ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Column 2 provides instead insights into the relationship between log prices and measures of demand and costs. More specifically, in a world in which the fundamental drivers of heterogeneity across firms are a and λ, we should expect prices to vary across firms only to the extent that a, λ and predetermined inputs like capital (with the latter contributing to determine short-term marginal costs) vary across firms. In column 2 of Table 4 we regress log prices on a, λ and log capital k across all observations while demeaning all variables by 8-digit Prodcom codes and adding year dummies. Again, in order to have more of a causal interpretation of parameters, we estimate the regression in first-differences. Coefficients indicate that more productive firms charge lower prices while firms selling more appealing products charge higher prices. In addition firms with a higher capital stock, and so with a lower short-term marginal cost, charge lower prices. In terms of magnitudes, the quantity TFP coefficient indicates that a 10% increase in productivity translates into about a 8.9% price reduction, i.e. a 0.89 average cost pass-through elasticity.[29]

[29]Our average cost pass-through elasticity might seem high compared to existing macro evidence (Campa and Goldberg, 2005). However, by looking at detailed product-level price and quantity data on French exporters, Berman et al. (2012) provide evidence that standard macro measures of pass-through elasticity are plagued by aggregation bias, while using the detailed information available in their study they find an average pass-through elasticity of 0.83 and further show the pass-through elasticity is even higher for smaller and less productive exporters. Using similar data for Belgium, Amiti et al. (2014) find an average pass-through elasticity of 0.80 for Belgian exporters with small exporters displaying a near complete pass-through. Therefore, considering our sample is comprised of relatively small single-product firms half of which do not export, a 0.89 average cost pass-through elasticity seems in line with the evidence provided in Berman et al.

Moving to markups, we perform in Table 5 very similar regressions to those of Table 4. In Table 5 the dependent variable is now the markup μ_{it} while neither the markup nor the covariates are demeaned anymore (because the level of markups can be compared across products) and a full battery of 8-digit Prodcom codes dummies is added to the regressions. Furthermore, as in Table 4, regressions are performed in first-differences. Considering column one, we find that markups significantly increase with quantity as well as with product appeal λ. Although intuitive, a positive relationship between markups and quantity is not a property of any preferences structure and it points into the direction of preferences featuring increasing relative love for variety (also called sub-convexity or Marshall's second law of demand) from which pro-competitive effects come from (Mrázová and Neary, 2017). Column 2 shows that more productive firms and/or firms producing more appealing products charge higher markups. This points again to preferences featuring increasing relative love for variety. Finally, a key difference with respect to Table 4 is that the R^2 is considerably lower. This might reflect model misspecification and/or measurement error with respect to markups, or also the presence of richer differences in demand across firms. In this respect, our framework does allow for markups that are not fully determined by a, λ and predetermined inputs and these findings might suggest markups are not simply a residual dimension of heterogeneity in the data.[30]

6 Comparison to alternative methodologies for retrieving demand

Our key contribution is to allow measuring demand heterogeneity across firms without resorting to demand system models (Ackerberg et al., 2007) or to the restrictive assumptions imposed by the methodology developed in Foster et al. (2008). Demand system models have very rich structures and allow for consumer- and product-specific elasticities of demand. However, they require detailed information on product and consumer characteristics as well as suitable instruments (like cost shifters) for identification. The high data requirements of these models are such that their application is usually limited to specific industries and contexts.[31] By contrast, our simpler and more parsimonious framework only requires information on

(2012) and Amiti et al. (2014).

[30]Within our framework nothing prevents preferences from being characterised by asymmetries across varieties, i.e. by more parameters than λ_{it}. For example, CARA-style preferences of the type $U\left(\tilde{Q}\right) = \int_{I_t}\left[1 - \mathrm{e}^{-\alpha(Q_{it}\Lambda_{it})^{\beta_{it}}}\right]\mathrm{d}i$, where $0 < \beta_{it} < 1$, are well behaved preferences falling into our general case and leading to (8). Therefore, (9) can be used to back out λ_{it}. Yet, markups will also depend upon β_{it}.

[31]See Aw-Roberts et al. (2020) for an application of a demand system model to international trade data.

product prices, quantities and inputs and does not need any additional instrument.

Concerning Foster et al. (2008), they use production data of US manufacturing firms, containing information on both value and physical quantity, to estimate quantity-based TFP as well as a measure of heterogeneity in demand, that they label demand shocks. More specifically, they measure demand as the residual of a regression where log quantity is regressed on log price, i.e. their demand measure is a firm-specific quantity shifter for given prices. The log price in their regression is further instrumented with quantity TFP which is obtained using industry costs shares to measure production function parameters. The key identifying assumption in their framework is thus that productivity is uncorrelated with demand heterogeneity.

In light of our framework, the Foster et al. (2008) approach is thus problematic for at least two reasons:

1. Markups are heterogeneous across firms: this means that the log price coefficient (demand elasticity) in their regression should be firm-specific. Within our framework we do not need to estimate those firm-specific elasticities because, based on our assumptions, they equal $\eta_{it} = \frac{\mu_{it}}{\mu_{it}-1}$ where η_{it} is the (perceived) elasticity of demand of firm i.[32]

2. Demand heterogeneity and TFP are strongly correlated with each other in our data: this means that their IV strategy would violate exclusion restrictions.[33] Within our framework we do not need to take a prior stand on the correlation between demand and productivity shocks.

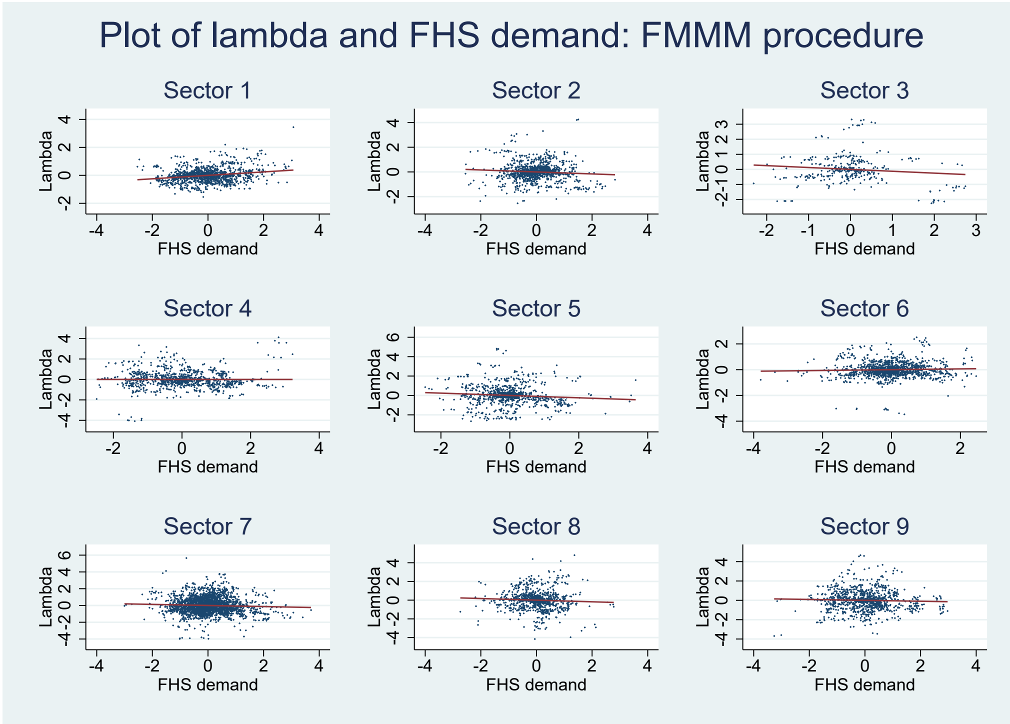
In order to gain insights into the differences between our approach and the Foster et al. (2008) approach we compute their demand measure (FHS demand measure) as the residual of a regression where log quantity is regressed on log price and the latter is instrumented with quantity TFP. Quantity TFP is constructed using industry-specific cost shares for the parameters of the production function as in Foster et al. (2008).[34] Figure 5 shows a plot of λ and the FHS demand measure for each industry obtained after demeaning both measures by

[32]Interestingly, Pozzi and Schivardi (2016) build upon a demand framework similar to Foster et al. (2008) but obtain measures of the firm-specific elasticity of demand from managers' assessment of the impacts of an hypothetical price increase of their firm.

[33]In a recent paper, Eslava and Haltiwanger (2018) improve upon Foster et al. (2008) by using as instrument the residual of a regression where quantity TFP is regressed on its time lag, i.e. productivity shocks rather than productivity are used as instruments.

[34]Foster et al. (2008) also control for a set of demand shifters, including a set of year dummies as well as the average income in the plant's local market where local markets are defined based on the Bureau of Economic Analysis' Economic Areas. We also include in our regressions year and 8-digit product dummies. Given the small size of Belgium, we did not include any control for the plant's local market income. Our IV estimations, available upon request, deliver highly (1%) significant coefficients for the log price coefficient in all nine industries.

Figure 5: Scatter plot of product appeal and the FHS demand measure



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict product appeal and the FHS demand measure, corresponding to each firm-year pair in our sample, and are constructed after demeaning both product appeal and the FHS demand measure by 8-digit product codes.

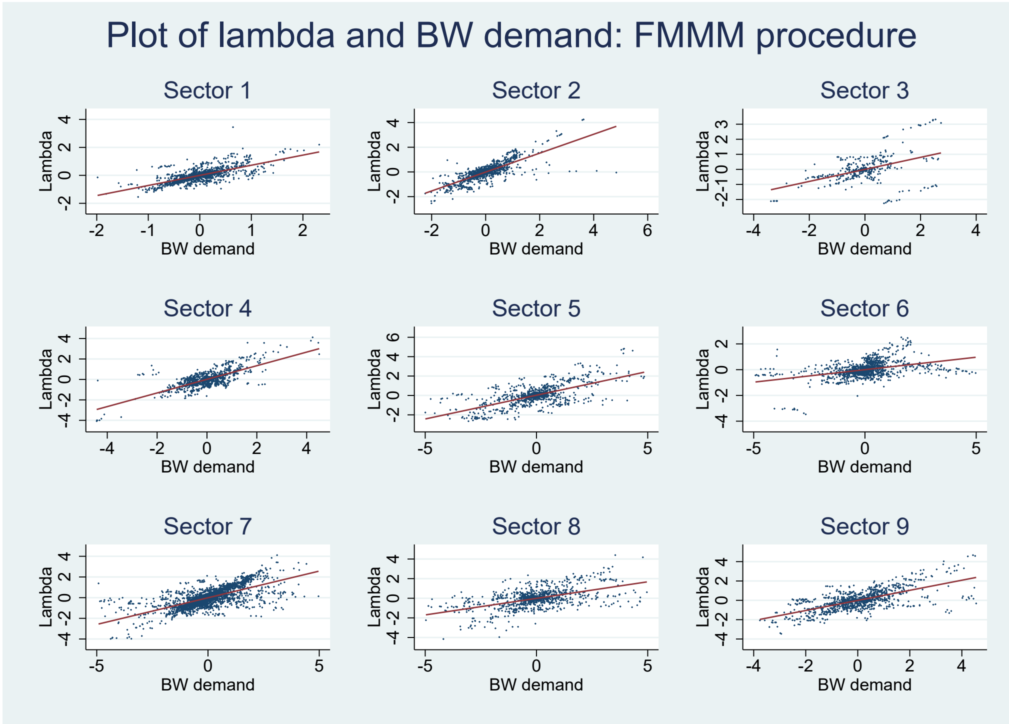
8-digit product codes. It is fair to say the two measures are mostly orthogonal to each other with correlation ranging between -0.114 and 0.211 across industries, while being significant in 4 (1) out of 9 cases at the 5% (1%) level. Furthermore, a strikingly similar picture emerges when: i) using production function coefficients from the DGKP procedure to compute product appeal; ii) not using cost shares to compute the FHS demand measure.[35]

As an additional step towards comparing our framework with other approaches for measuring demand heterogeneity we also consider the elasticity measurement exercise proposed by Broda and Weinstein (2006) and further developed in Hottman et al. (2016). More specifically, we borrow from Broda and Weinstein (2006) measures of the product-specific elasticity of demand obtained from international trade data.[36] Equipped with these elasticity measures,

[35]Figure D-3 in Appendix D shows a plot of λ obtained with the DGKP procedure and the FHS demand measure. Furthermore, Figure B-3 in Appendix B shows a plot of λ (obtained with the FMMM procedure) and an alternative FHS demand measure obtained using FMMM estimates of the production function coefficients, rather than cost shares, to compute TFP and then instrument log price.

[36]Broda and Weinstein (2006) build on a monopolistic competition constant elasticity/markup framework to estimate their elasticities. We use elasticities disaggregated at the SITC Rev.3 4-digit level (roughly 1,000 products) and referring to the period 1990-2001 for the US. We use the HS classification to bridge the original SITC Rev.3 4-digit classification to the CPA 2002 6-digit classification (first 6 digits of the Prodcom classification).

Figure 6: Scatter plot of product appeal and the BW demand measure



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict product appeal and the BW demand measure, corresponding to each firm-year pair in our sample, and are constructed after demeaning both product appeal and the BW demand measure by 8-digit product codes.

we recompute product appeal while imposing the same elasticity/markup across firms within a product category p (BW demand measure), i.e. $\lambda_{it} = \mu_p r_{it} - q_{it}$ where the common markup is $\mu_p = \frac{\eta_p}{\eta_p - 1}$ and η_p is the product-specific elasticity of demand borrowed from Broda and Weinstein (2006). Figure 6 shows a plot of λ and the BW demand measure for each industry obtained after demeaning both measures by 8-digit product codes. Interestingly, the two demand measures correlate reasonably well (correlation coefficients ranging between 0.404 and 0.791 while being always significant at the 1% level) and certainly better than product appeal and the FHS demand measure.

7 Decomposing revenue productivity: an application to Chinese imports competition

The decompositions provided in equations (10) and (11) allow us to shed new light on a well studied question. More specifically, besides the well-documented negative effects on employment (Autor et al., 2013), numerous studies have explored the many impacts of the spectacular rise of Chinese trade, that started well before China joined the WTO in 2001, on both developed and developing countries firms and workers. One particular aspect we are interested in here is how China has affected the productivity of European firms and in particular Belgian firms. In this respect, Bloom et al. (2016) provide evidence supporting the claim that import competition from China caused an increase in technical change, as well as an increase in revenue TFP, for European firms selling products most affected by rising imports from China. Bloom et al. (2016) rationalise these effects via a number of channels relating competition to innovation and X-inefficiencies.

Bloom et al. (2016) deal with the presence of unobserved demand and/or supply shocks potentially correlated with Chinese imports competition patterns by focusing on a specific industry ("Textile and Apparel") and exploiting detailed information on import quotas. These quotas were imposed at the EU-level on Chinese imports, as well as on imports from other non-WTO countries, and affected some 6-digit products within the industry but not others. As a consequence of China joining the WTO, these quotas were removed over the time frame of our analysis. To provide some context, when these quotas were abolished this generated a 240% increase in Chinese imports on average within the affected product groups. The underlying identifying assumption of this strategy is that unobserved demand/technology shocks are uncorrelated with the strength of quotas to non-WTO countries (like China) in 2000. Since these quotas were built up from the 1950s, and their phased abolition negotiated in the late 1980s was in preparation for the Uruguay Round, Bloom et al. (2016) conclude that this seems a plausible assumption.

We first start by replicating some key findings of Bloom et al. (2016) and other papers; namely that employment decreased and revenue productivity increased for firms more affected by import competition. We thus match the product-level quota measure $QUOTA_{CPA6}$ to firms in the "Textile and Apparel" industry and run a regression where the time change in either log firm-level labour expenditure or log revenue productivity is used as outcome variable. Results are reported in columns 1 and 2 of Table 6. The negative (positive) and significant coefficient for labour (revenue TFP) indicates that, on average, labour expenditure growth (revenue productivity growth) has been 3.6% lower (0.7% higher) per year over our time frame for firms affected by the quota removal as compared to non-affected firms.

Within our framework we can ask a deeper question and in particular how the increase in revenue productivity has materialised. Indeed, Bloom et al. (2016) provide evidence that import competition from China caused an increase in innovation while claiming this increased innovation went into reducing production costs, i.e. boosting quantity TFP. However, Bloom et al. (2016) are only able to measure revenue TFP and so are ultimately unable to distinguish the underlying impacts of innovation on quantity TFP, demand and markups. More specifically an alternative, to Bloom et al. (2016), scenario is one in which the increased innovation was primarily directed towards product quality improvements rather than towards increasing quantity TFP/reducing production costs. In order to get insights into this we use our decompositions and in particular markups-adjusted TFP ($\tilde{a}$), product appeal ($\tilde{\lambda}$) and scale ($\tilde{\bar{q}}$) as additional y variables in columns 3 to 5 of Table 6. In this respect note that, by construction, the sum of the 3 coefficients equals the coefficient of revenue TFP (0.007).

Table 6: Disentangling the impact of Chinese imports competition on revenue productivity in terms of markups-adjusted TFP ($\tilde{a}$), product appeal ($\tilde{\lambda}$) and scale ($\tilde{\bar{q}}$): quota analysis on the "Textile and Apparel" industry

Outcome measure	Labour	Rev.TFP	$\tilde{a}$	$\tilde{\lambda}$	$\tilde{\bar{q}}$
$Quota_{CPA6}$	-0.036[c]	0.007[b]	0.109[c]	-0.116[b]	0.014[b]
	(0.021)	(0.003)	(0.059)	(0.057)	(0.006)
Observations	700	700	700	700	700
R-squared	0.005	0.005	0.003	0.004	0.008

Notes: $Quota_{CPA6}$ denotes the share of products within a CPA6 code belonging to the "Textile and Apparel" industry affected by a removal of quota on Chinese imports. Firm-level clustered standard errors in parenthesis. [a] p<0.01, [b] p<0.05, [c] p<0.1.

The overall picture emerging from looking at coefficients is suggestive of the following scenario. First, in the light of our model a quota removal is a negative demand shock that should impact product appeal. Indeed, the coefficient of markups-adjusted product appeal $\tilde{\lambda}$ is significantly negative and quite large. This suggests that, if there has been a product quality innovation effort, it was not large enough to countervail the reduction in demand caused by the quota removal. At the same time, the increase in innovation documented in Bloom et al. (2016) seems to have been successful in increasing markups-adjusted quantity TFP $\tilde{a}$ and reducing production costs as indicated by the related positive and large coefficient. Incidentally, the two opposing effects roughly cancel each other out and so the observed increased in revenue TFP essentially comes from the reduction in firm operations/scale, i.e. from the increase in the markups-adjusted scale $\tilde{\bar{q}}_{it} = \frac{(1-\mu_{it})\bar{q}_{it}}{\mu_{it}}$.[37]

We reach the very same conclusions using a completely different regression design in Table 7. More specifically, building on Autor et al. (2013) we consider all industries and construct a time-varying 6-digit product-specific measure of Chinese imports penetration in the EU15 market ($IPC^{EU15}_{CPA6,t}$) based on import shares. In order to deal with the presence of unobserved demand/technology shocks at the product-time level characterising the EU15 market, and

[37]Markups $\mu_{it} > 1$ imply that, everything else equal, markups-adjusted scale increases when scale decreases.

correlated with $IPC^{EU15}_{CPA6,t}$, we then instrument this measure with the equivalent Chinese imports penetration measure in the US market $IPC^{US}_{CPA6,t}$. In doing so we also allow for firm fixed effects and year dummies.

Table 7: Disentangling the impact of Chinese imports competition on revenue productivity in terms of TFP (a), product appeal (λ), scale ($\bar{q}$) and markups (μ): Chinese import penetration analysis

Outcome measure	Labour	Rev.TFP	a	λ	$\bar{q}$	μ
$IPC^{EU15}_{CPA6,t}$	-0.739^a	0.185^a	1.321^a	-1.073^a	-0.848^a	0.049
	(0.241)	(0.042)	(0.403)	(0.414)	(0.258)	(0.065)
Observations	10,161	10,161	10,161	10,161	10,161	10,161
R-squared	0.174	0.029	0.017	0.011	0.214	0.037
Firm FE and year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap rk LM statistic under-id	146.71	146.71	146.71	146.71	146.71	146.71
P-value	0	0	0	0	0	0
Kleibergen-Paap rk Wald F statistic weak id.	226.34	226.34	226.34	226.34	226.34	226.34

Notes: Instrumental variable estimator with firm fixed effects implemented. Chinese import penetration in the EU15 market ($IPC^{EU15}_{CPA6,t}$) is instrumented with Chinese import penetration in the US market ($IPC^{US}_{CPA6,t}$). Firm-level clustered standard errors in parenthesis. ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1. The Kleibergen-Paap rk LM statistic tests for under-identification. P-value reported. The Kleibergen-Paap rk Wald F statistic tests for weak identification. The corresponding Stock-Yogo weak ID test critical value is 16.38 for a 10% maximal IV size bias.

Coming back to Table 7, where we provide results based on non-markups-adjusted measures as well as firm markups to complete the picture, one can appreciate in columns 3 and 4 the same two countervailing effects of Chinese imports competition on quantity TFP (positive) and products appeal (negative). At the same time production scale is negatively and significantly affected (column 5) while markups increased a little but not in a significant way (column 6). Overall, this adds up again to an increase in revenue productivity stemming from higher import penetration (column 2) while the impact on firm employment is negative as in previous studies (column 1). Last but not least, test statistics suggest $IPC^{US}_{CPA6,t}$ is a strong instrument for $IPC^{EU15}_{CPA6,t}$.

8 Conclusions

We provide a novel framework that simultaneously allows recovering heterogeneity in productivity, demand and markups across firms while leaving the correlation among the three unrestricted. In doing so, we provide an exact decomposition of revenue productivity in terms of the underlying heterogeneities, so bridging the gap between quantity and revenue productivity estimations. We accomplish this by explicitly introducing demand heterogeneity and by systematically exploiting assumptions of previous firm-level productivity estimation approaches. We apply our econometric framework to Belgian manufacturing firms and quantify productivity, markups and demand heterogeneity. We show how these heterogeneities are correlated among them, across time as well as with measures obtained from other approaches. We finally assess how and to what extent our three dimensions of heterogeneity allow gaining deeper and sharper insights on firm response to increasing import competition from China.

Our methodology is rich enough to be applied to markets where products have some features of both horizontal and vertical differentiation. At the same time, our framework is parsimonious enough to allow retrieving productivity, demand, and markups heterogeneity with relatively little information compared to demand systems models. It also builds upon firm-level data on physical production that is becoming increasingly available to researchers (Belgium, Brazil, Chile, Denmark, France, India, UK and the US to name a few countries). Both elements provide a wide scope of applications of our framework.

Our analysis has policy implications both at the micro and macro level. At the micro level, it makes a significant difference to know that some firms or industries lack in competitiveness because of poor physical TFP (due for example to low expenditure in process R&D) or poor product quality (due for example to low expenditure in product R&D). At the macro level, our framework allows analysing aggregate revenue productivity cycles, such as the severe downturn of EU countries' revenue productivity since the financial crisis, not only in terms of changes in some underlying production capacity of the economy, but also as changes in markups and demand. These are the objects of ongoing research.

References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of econometrics*, 6:4171–4276.

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

Amiti, M., Itskhoki, O., and Konings, J. (2014). Importers, exporters, and exchange rate disconnect. *American Economic Review*, 104(7):1942–78.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *The American Economic Review*, 98(5):1998–2031.

Autor, D., Dorn, D., and Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the US. *American Economic Review*, 103(6):2121–68.

Aw-Roberts, B. Y., Lee, Y., and Vandenbussche, H. (2020). Consumer taste in trade. *CEPR Discussion Papers*, 14941.

Behrens, K., Corcos, G., and Mion, G. (2013). Trade crisis? What trade crisis? *Review of economics and statistics*, 95(2):702–709.

Behrens, K., Mion, G., Murata, Y., and Südekum, J. (2014). Trade, wages, and productivity. *International Economic Review*, 55(4):1305–1348.

Berman, N., Martin, P., and Mayer, T. (2012). How do different exporters react to exchange rate changes? *The Quarterly Journal of Economics*, 127(1):437–492.

Bernard, A. B., Blanchard, E. J., Van Beveren, I., and Vandenbussche, H. (2018). Carry-along trade. *The Review of Economic Studies*, 86(2):526–563.

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

Bloom, N., Draca, M., and Van Reenen, J. (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *The Review of Economic Studies*, 83(1):87–117.

Bloom, N., Manova, K., Van Reenen, J., Sun, S. T., and Yu, Z. (2020). Trade and management. *Review of Economics and Statistics*, pages 1–44.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585.

Campa, J. M. and Goldberg, L. S. (2005). Exchange rate pass-through into import prices. *The review of Economics and Statistics*, 87(4):679–690.

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., and Van Biesebroeck, J. (2014). International competition and firm performance: Evidence from Belgium. NBB Working Papers 269.

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

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–71.

Eslava, M. and Haltiwanger, J. (2018). The Life-Cycle Growth of Plants: The Role of Productivity, Demand and Distortions. University of Maryland, Mimeo.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425.

Grieco, P. L., Li, S., and Zhang, H. (2016). Production function estimation with unobserved input price dispersion. *International Economic Review*, 57(2):665–690.

Haltiwanger, J., Kulick, R., and Syverson, C. (2018). Misallocation measures: The distortion that ate the residual. *NBER Working Paper*, 14491.

Harris, R. and Robinson, C. (2003). Foreign ownership and productivity in the United Kingdom estimates for UK manufacturing using the ARD. *Review of Industrial organization*, 22(3):207–223.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Jaumandreu, J. and Yin, H. (2017). Cost and product advantages: A firm-level model for the chinese exports and industry growth. CEPR Discussion Paper 11862.

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

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

Mion, G. and Zhu, L. (2013). Import competition from and offshoring to China: A curse or blessing for firms? *Journal of International Economics*, 89(1):202–215.

Mrázová, M. and Neary, J. P. (2017). Not so demanding: Demand structure and firm behavior. *American Economic Review*, 107(12):3835–74.

Muûls, M. (2015). Exporters, importers and credit constraints. *Journal of International Economics*, 95(2):333 – 343.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):l263–l297.

Pozzi, A. and Schivardi, F. (2016). Demand or productivity: What determines firm growth? *The RAND Journal of Economics*, 47(3):608–630.

Schott, P. K. (2008). The relative sophistication of chinese exports. *Economic policy*, 23(53):6–49.

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

Zhelobodko, E., Kokovin, S., Parenti, M., and Thisse, J.-F. (2012). Monopolistic competition: Beyond the constant elasticity of substitution. *Econometrica*, 80(6):2765–2784.

Appendix

A Product appeal in the Atkeson and Burstein (2008) oligopoly model

The key property needed for λ_{it} to be interpreted as a measure of perceived quality/product appeal is that the elasticity of revenue with respect to quantity is equal to the elasticity of revenue with respect to product appeal at the profit maximising solution. We provide here an example of how (8) holds within an oligopoly model based on Atkeson and Burstein (2008) and further refined in Hottman et al. (2016) for multi-product firms. The key ingredient is the same as in the monopolistic competition case, namely that quantities enter into preferences as $\tilde{Q}_{it} = \Lambda_{it} Q_{it}$. Atkeson and Burstein (2008) consider a nested CES model of quantity competition à la Cournot in which firms sell differentiated varieties and are large enough to perceive their impact on industry aggregates while charging markups that depend upon their market share. More specifically, a finite number of single-product firms operates within each industry j where preferences are characterised by a CES demand with parameter η_j. At each point in time final consumption is produced by a competitive firm using the output of a continuum of industries $\mathbb{Q}_{jt}$ for $j \in [0, 1]$ as inputs subject to a CES production function with parameter η and $1 < \eta < \eta_j$, i.e. varieties within an industry are more substitutable with each other than industry outputs Q_{jt} across industries. Contrary to the monopolistic competition case, firm i operating in industry j does recognise that sectoral prices and quantities vary when it changes its quantity. Introducing product appeal within this framework is quite straightforward.

First, industry j aggregate output at time t is:

$$\mathbb{Q}_{jt} = \left(\sum_{i \in I_{jt}} (Q_{ijt} \Lambda_{ijt})^{\frac{\eta_j - 1}{\eta_j}} \right)^{\frac{\eta_j}{\eta_j - 1}}, \tag{A-1}$$

where I_{jt} is the set of varieties (firms) available within industry j at time t, Q_{ijt} is firm i output in industry j and Λ_{ijt} is product appeal. The inverse demand corresponding to varieties within an industry is:

$$\frac{P_{ijt}}{\mathbb{P}_{jt}} = \left(\frac{Q_{ijt}}{\mathbb{Q}_{jt}} \right)^{-\frac{1}{\eta_j}} \Lambda_{ijt}^{\frac{\eta_j - 1}{\eta_j}}, \tag{A-2}$$

where $\mathbb{P}_{jt}$ is the CES price index for industry j and and P_{ijt} is firm i price in industry j.

Firm revenue is thus $R_{ijt} = P_{ijt}Q_{ijt}$. The inverse demand corresponding to industry outputs is instead:

$$\frac{\mathbb{P}_{jt}}{\mathbb{P}_t} = \left(\frac{\mathbb{Q}_{jt}}{\mathbb{Q}_t}\right)^{-\frac{1}{\eta}}, \tag{A-3}$$

where $\mathbb{P}_t$ and $\mathbb{Q}_t$ are the CES price and quantity indexes for the whole economy and in particular the latter is:

$$\mathbb{Q}_t = \left(\sum_j (\mathbb{Q}_{jt})^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}.$$

In choosing the optimal quantity, firms realise the impact of their choices on industry aggregates and in particular on $\mathbb{Q}_{jt}$. They have instead a zero measure at the aggregate level and so take $\mathbb{P}_t$ and $\mathbb{Q}_t$ as given. Combining (A-2) and (A-3) the relevant demand is:

$$\frac{P_{ijt}}{\mathbb{P}_t} = \left(\frac{Q_{ijt}}{\mathbb{Q}_{jt}}\right)^{-\frac{1}{\eta_j}} \Lambda_{ijt}^{\frac{\eta_j-1}{\eta_j}} \left(\frac{\mathbb{Q}_{jt}}{\mathbb{Q}_t}\right)^{-\frac{1}{\eta}}. \tag{A-4}$$

Using the properties of CES demand, the market share of firm i in industry j ($s_{Rijt} \equiv R_{ijt}/(\mathbb{P}_{jt}\mathbb{Q}_{jt})$) equals the elasticity of the industry quantity index with respect to firm quantity:

$$\frac{\partial \mathbb{Q}_{jt}}{\partial Q_{ijt}} \frac{Q_{ijt}}{\mathbb{Q}_{jt}} = s_{Rijt}. \tag{A-5}$$

Symmetrically we have:

$$\frac{\partial \mathbb{Q}_{jt}}{\partial \Lambda_{ijt}} \frac{\Lambda_{ijt}}{\mathbb{Q}_{jt}} = s_{Rijt}. \tag{A-6}$$

From (A-4) and (A-5), the elasticity of firm price P_{ijt} with respect to firm quantity Q_{ijt} is thus $-\frac{1}{\eta_j}(1 - s_{Rijt}) - \frac{1}{\eta}s_{Rijt}$. Therefore, the elasticity of demand is:

$$\epsilon_{it} = -\frac{\partial q_{ijt}}{\partial p_{ijt}} = \left(\frac{1}{\eta_j}(1 - s_{Rijt}) + \frac{1}{\eta}s_{Rijt}\right)^{-1}, \tag{A-7}$$

while from profit maximisation the markup μ_{ijt} is related to the elasticity of demand in the usual way:

$$\mu_{ijt} = \frac{\epsilon_{ijt}}{\epsilon_{ijt} - 1}. \tag{A-8}$$

Multiplying both sides of (A-4) by quantity delivers the revenue equation:

$$\frac{R_{ijt}}{\mathbb{P}_t} = (Q_{ijt}\Lambda_{ijt})^{\frac{\eta_j-1}{\eta_j}} \mathbb{Q}_{jt}^{\frac{1}{\eta_j}} \left(\frac{\mathbb{Q}_{jt}}{\mathbb{Q}_t}\right)^{-\frac{1}{\eta}}, \tag{A-9}$$

from which the elasticity of revenue with respect to quantity is equal to the elasticity of revenue with respect to product appeal and equal to one over the profit maximising markup:

$$\frac{\partial r_{ijt}}{\partial q_{ijt}} = \frac{\partial r_{ijt}}{\partial \lambda_{ijt}} = \frac{\eta_j - 1}{\eta_j} + \frac{1}{\eta_j} s_{Rijt} - \frac{1}{\eta} s_{Rijt} = \frac{1}{\mu_{ijt}}. \quad \text{(A-10)}$$

B Additional results

B.1 Some analytical results

The first order conditions of the utility maximisation problem imply:

$$\frac{\partial U}{\partial Q_{it}} = \frac{\partial U}{\partial \tilde{Q}_{it}} \frac{\partial \tilde{Q}_{it}}{\partial Q_{it}} = \frac{\partial U}{\partial \tilde{Q}_{it}} \Lambda_{it} = \kappa_t P_{it},$$

where κ_t is a Lagrange multiplier and $\frac{\partial \tilde{Q}_{it}}{\partial Q_{it}} = \Lambda_{it}$. Taking logs we have:

$$\ln \frac{\partial U}{\partial \tilde{Q}_{it}} + \lambda_{it} = \ln \kappa_t + p_{it}. \quad \text{(B-1)}$$

Differentiating both sides of (B-1) with respect to q_{it} yields:

$$\frac{\partial p_{it}}{\partial q_{it}} \equiv -\frac{1}{\eta_{it}} = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial q_{it}} = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial \tilde{q}_{it}} \frac{\partial \tilde{q}_{it}}{\partial q_{it}} = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial \tilde{q}_{it}}, \quad \text{(B-2)}$$

where $\frac{\partial \tilde{q}_{it}}{\partial q_{it}} = 1$ and η_{it} is the elasticity of demand. On the other hand, keeping in mind that $\frac{\partial \tilde{q}_{it}}{\partial \lambda_{it}} = 1$, differentiation of both sides of (B-1) with respect to λ_{it} gives:

$$\frac{\partial p_{it}}{\partial \lambda_{it}} = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial \lambda_{it}} + 1 = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial \tilde{q}_{it}} \frac{\partial \tilde{q}_{it}}{\partial \lambda_{it}} + 1 = \frac{\partial \ln \frac{\partial U}{\partial \tilde{Q}_{it}}}{\partial \tilde{q}_{it}} + 1 = 1 - \frac{1}{\eta_{it}} = 1 + \frac{\partial p_{it}}{\partial q_{it}}, \quad \text{(B-3)}$$

i.e. the elasticity of the price with respect to quantity differs from the elasticity of the price with respect to product appeal by one.

B.2 Examples of log revenue functions

In this Section we provide examples of log revenue functions obtained from HARA preferences (Haltiwanger et al., 2018) and CARA preferences (Behrens et al., 2014) supporting the log-linear approximation. In the case of CARA preferences, the underlying utility behind heterogeneity in product appeal across firms would be:

$$U\left(\tilde{Q}\right) = \int_{I_t} \left[1 - \mathrm{e}^{-\alpha Q_{it} \Lambda_{it}}\right] \mathrm{d}i,$$

where I_t is the set of varieties available at time t. The inverse demand function corresponding to such preferences is:

$$P_{it} = \kappa_t^{-1} \frac{\partial U\left(\tilde{Q}\right)}{\partial Q_{it}} = \kappa_t^{-1} \alpha \Lambda_{it} \mathrm{e}^{-\alpha Q_{it} \Lambda_{it}},$$

where κ_t is a Lagrange multiplier and so the log revenue function is:

$$r_{it} = \log(P_{it} Q_{it}) = -\log(\kappa_t) + \log(\alpha) + q_{it} + \lambda_{it} - (\alpha \mathrm{e}^{q_{it}+\lambda_{it}}),$$

from which it is clear that $\frac{\partial r_{it}}{\partial q_{it}} = \frac{\partial r_{it}}{\partial \lambda_{it}}$ and, because of (8), equal to $\frac{1}{\mu_{it}}$ at the profit maximising solution.

Figure B-1: CARA log revenue function examples

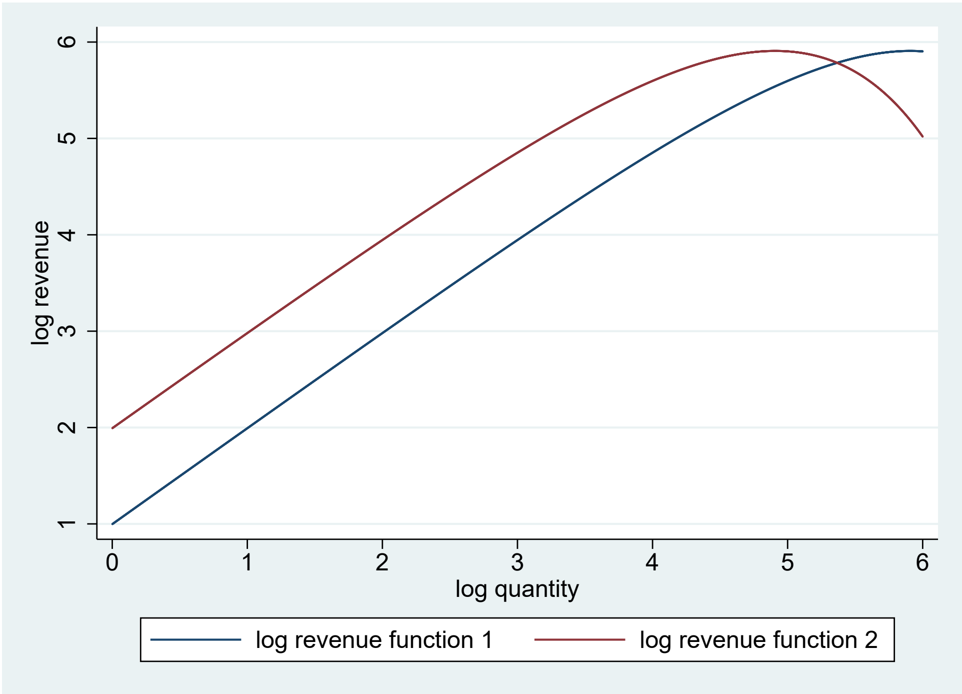
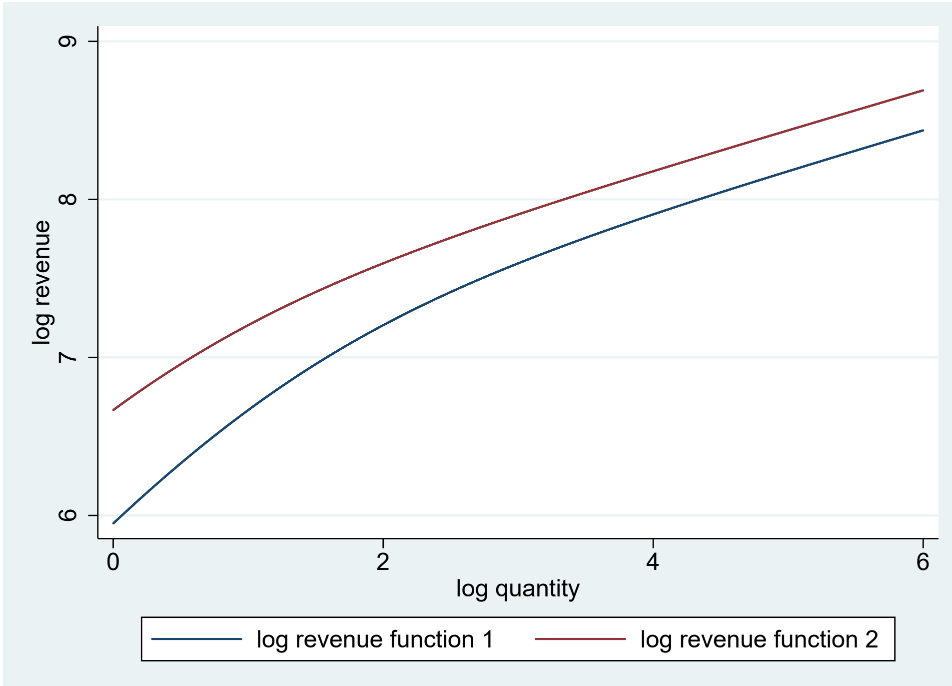


Figure B-1 plots two CARA log revenue functions obtained using two different values for product appeal: λ_{it}=1 for log revenue function 1 and λ_{it}=2 for log revenue function 2. The other parameters are α=0.001 and κ_t=0.001. As can be appreciated from Figure B-1, a linear approximation looks both reasonable and accurate for the relevant part of the two log revenue functions, i.e. within the range where log revenue (and revenue) is increasing because the marginal revenue is positive and the demand is elastic.

Figure B-2: HARA log revenue function examples



We provide further evidence about the linear approximation in Figure B-2, which is the equivalent of Figure B-1 for HARA preferences. With HARA preferences the relevant utility function is $U\left(\tilde{Q}\right) = \int_{I_t} \frac{\left(\frac{Q_{it}\Lambda_{it}}{1-\rho}+\alpha\right)^{\rho}-\alpha^{\rho}}{\frac{\rho}{1-\rho}} \mathrm{d}i$. Parameters used in Figure B-2 are, besides λ_{it}=1 for log revenue function 1 and λ_{it}=2 for log revenue function 2, α=10, ρ=0.25 and the Lagrange multiplier κ_t=0.001. Again, one can appreciate that a linear approximation looks both reasonable and accurate.

B.3 Measurement error in output and unanticipated shocks

One issue we need to account for is the presence of measurement error in output (quantity and revenue) and/or unanticipated productivity shocks. In the former case, instead of q_{it}, the econometrician might be observing q'_{it}=$q_{it} + e_{it}$ where e_{it} is standard measurement error. Another interpretation of the same equation is that e_{it} represents productivity shocks unanticipated by the firm. (1) thus becomes:

$$q'_{it} = \alpha_L l_{it} + \alpha_M m_{it} + (\gamma - \alpha_L - \alpha_M)k_{it} + a_{it} + e_{it}.$$

The approach suggested by the literature (Ackerberg et al., 2015; De Loecker et al., 2016) to deal with measurement error in output and/or unanticipated shocks e_{it} is based on the proxy variable framework and a semi-parametric implementation. We follow this approach and, building on the same logic of equation (19) in De Loecker et al. (2016), we estimate:

$$q'_{it} = poly(l_{it}, m_{it}, p_{it}, k_{it}) + e_{it}, \tag{B-4}$$

where q'_{it} is (log) quantity as reported in the data and $poly(.)$ is a third-order polynomial in l_{it}, m_{it}, p_{it} and k_{it}.[38] We run (B-4) separately for each two-digit industry and, following De Loecker et al. (2016), we also consider importer and exporter status dummies as additional proxies while adding a full set of 8-digit product dummies and year dummies to (B-4). We then use the OLS prediction of q'_{it}, that we label $\hat{q'}_{it}^{OLS}$, as quantity in the rest of the analysis.

We also use the same approach for revenue and consider:

$$r'_{it} = poly(l_{it}, m_{it}, p_{it}, k_{it}) + \bar{e}_{it}, \tag{B-5}$$

where $\bar{e}_{it}$ now contains measurement error in both quantity and prices, as well as unobserved productivity shocks, and use the OLS prediction of r'_{it}, that we label $\hat{r'}_{it}^{OLS}$, as revenue in the rest of the analysis.[39] We run (B-5) separately for each two-digit industry and consider importer and exporter status dummies as additional proxies while adding a full set of 8-digit product dummies and year dummies to (B-5). Also note that, by purging revenue from measurement error and using $\hat{r'}_{it}^{OLS}$ instead of r'_{it}, we obtain a more reliable measure of the share of materials in revenue (s_{Mit}) that is needed to compute markups from (5) as in De Loecker et al. (2016).

Last but not lest our key findings are unaffected by accounting for measurement error with this approach but the reliability and precision of technology parameters (especially capital) does benefit from it. The R^2 of regressions (B-4) and (B-5) are indeed very high (about 0.98 or higher).

B.4 Additional Graphs and Tables

[38]The logic behind using (B-4) to purge quantity from measurement error and unanticipated shocks is straightforward. From the quantity equation (1) q_{it} is a function of l_{it}, m_{it}, k_{it} and a_{it}. Using prices p_{it} as a proxy for a_{it}, while assuming invertibility, one can then write a_{it} as a function of l_{it}, m_{it}, p_{it} and k_{it}. Overall, q_{it} is thus a function of l_{it}, m_{it}, p_{it} and k_{it} that can be semi-parametrically approximated by a polynomial function. Crucially, measurement error and/or unanticipated shocks do influence a firm's choices and so are not part of the polynomial approximation but rather the residual of equation (B-4).

[39]From the revenue equation (8) r_{it} is a function of q_{it} μ_{it} and λ_{it}. We already know q_{it} is a function of l_{it}, m_{it}, p_{it} and k_{it}. Profit maximising prices will be a function of l_{it}, m_{it}, k_{it}, μ_{it} and λ_{it}. Building again on invertibility, one can thus express both μ_{it} and λ_{it} as a function of l_{it}, m_{it}, p_{it} and k_{it}. Overall, r_{it} will thus be a function of l_{it}, m_{it}, p_{it} and k_{it} than can be semi-parametrically approximated by a polynomial function. Again, measurement error and/or unanticipated shocks influence a firm's choices and so are not part of the polynomial approximation but rather the residual of equation (B-5).

Table B-1: OLS regression of product appeal shocks ($\nu_{\lambda it}$) on TFP shocks (ν_{ait}) by industry

Industry	1	2	3	4	5	6	7	8	9
					Product appeal shock				
TFP shock	-0.9147^a	-1.026^a	-0.8932^a	-0.9294^a	-1.019^a	-0.8555^a	-0.9876^a	-0.9641^a	-0.9931^a
	(0.0380)	(0.0214)	(0.0359)	(0.1136)	(0.0075)	(0.0847)	(0.0102)	(0.0091)	(0.0066)
N Obs	901	843	232	710	692	867	2,000	785	738
R^2	0.6458	0.8830	0.8178	0.5858	0.9746	0.3055	0.9583	0.9388	0.9766

Notes: Quantity TFP and product appeal have been demeaned by 8-digit Prodcom codes before computing shocks $\nu_{\lambda it}$ and ν_{ait}. Time dummies are included in estimations but are not reported here. Bootstrapped standard errors in parenthesis (200 replications). ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Table B-2: OLS regression of TFP and product appeal on their time lag by industry

Industry	1	2	3	4	5	6	7	8	9
					TFP				
lag TFP	0.9743^a	0.9718^a	0.9865^a	0.9577^a	0.8715^a	0.9711^a	0.8665^a	0.7482^a	0.8332^a
	(0.0155)	(0.0126)	(0.0157)	(0.0231)	(0.0264)	(0.0114)	(0.0245)	(0.0579)	(0.0355)
N Obs	901	843	232	710	702	867	2,000	785	738
R^2	0.8742	0.8785	0.9620	0.8986	0.7867	0.9371	0.7035	0.5986	0.7300
					product appeal				
lag product appeal	0.9654^a	0.9688^a	0.9886^a	0.9532^a	0.8737^a	0.9503^a	0.8769^a	0.7465^a	0.8292^a
	(0.0136)	(0.0136)	(0.012)	(0.0323)	(0.0265)	(0.0153)	(0.0247)	(0.059)	(0.0413)
N Obs	901	843	232	710	702	867	2,000	785	738
R^2	0.8763	0.8825	0.9688	0.8710	0.7813	0.8942	0.7216	0.6001	0.7239

Notes: Quantity TFP and product appeal are demeaned by 8-digit Prodcom codes. Time dummies are included in estimations but are not reported here. Bootstrapped standard errors in parenthesis (200 replications). ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Table B-3: Basic summary statistics of the estimation sample

Industry	Industry description	Nace	N Obs	Statistic	Revenue	Quantity	Labour	Materials	Capital
1	Food prod, beverages and tobacco	15+16	1,317						
				mean	1.431	14.311	-0.220	0.979	-0.024
				st. dev.	1.165	1.741	1.003	1.240	1.336
				p5	-0.277	11.786	-1.626	-0.944	-2.376
				p95	3.574	17.458	1.689	3.205	2.209
2	Textiles and leather	17 to 19	1,225						
				mean	1.047	12.256	-0.359	0.578	-1.035
				st. dev.	1.129	1.938	0.995	1.227	1.549
				p5	-0.696	9.029	-1.789	-1.284	-3.650
				p95	3.147	15.677	1.589	2.793	1.535
3	Wood except furniture	20	348						
				mean	1.284	11.359	-0.166	0.812	-0.394
				st. dev.	1.354	2.827	1.169	1.457	1.598
				p5	-0.314	7.738	-1.505	-0.951	-2.962
				p95	4.579	15.778	2.586	4.325	2.721
4	Pulp, paper, publish. and print.	21+22	975						
				mean	1.916	14.875	0.396	1.447	-0.299
				st. dev.	1.240	1.679	1.190	1.296	1.719
				p5	0.124	12.087	-1.181	-0.530	-3.353
				p95	4.227	17.728	2.735	3.829	2.610
5	Chemicals and rubber	24+25	1,043						
				mean	1.865	14.060	0.145	1.453	0.206
				st. dev.	1.121	2.378	1.085	1.162	1.267
				p5	0.272	10.062	-1.336	-0.308	-1.870
				p95	3.852	17.819	2.249	3.481	2.273
6	Other non-metallic mineral prod.	26	1,215						
				mean	1.792	16.002	0.295	1.298	0.239
				st. dev.	1.039	2.850	1.111	1.049	1.312
				p5	0.310	10.795	-1.258	-0.225	-2.021
				p95	3.774	19.318	2.447	3.121	2.491
7	Metals and fabric. metal prod.	27+28	2,814						
				mean	1.114	12.563	-0.160	0.588	-0.827
				st. dev.	0.980	2.301	0.875	1.067	1.262
				p5	-0.162	8.151	-1.246	-0.909	-2.883
				p95	3.036	16.109	1.543	2.668	1.360
8	Machin., electr. and optic. equip.	29 to 33	1,108						
				mean	1.514	8.828	0.233	1.002	-0.693
				st. dev.	1.169	3.365	1.031	1.259	1.439
				p5	-0.070	3.989	-1.078	-0.759	-3.049
				p95	3.506	14.374	2.105	3.127	1.738
9	Transport equipment and n.e.c.	34 to 36	1,055						
				mean	1.244	9.552	-0.110	0.762	-0.830
				st. dev.	1.130	2.935	1.072	1.206	1.419
				p5	-0.276	5.361	-1.489	-0.951	-3.411
				p95	3.481	16.102	1.981	3.158	1.371

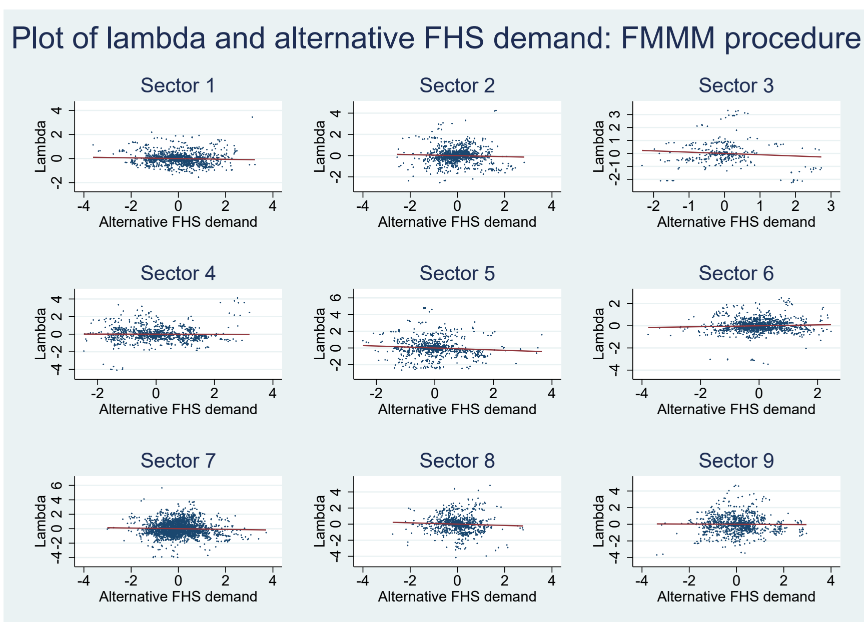
Notes: Revenue denotes log revenue, quantity denotes log quantity in the unit specific to a product, labour denotes log of labour expenditure, materials denotes log of materials expenditure, capital denotes log capital stock. All monetary values are expressed in current million euros.

Table B-4: OLS regression of changes in product appeal on changes in TFP by industry

Industry	1	2	3	4	5	6	7	8	9
					Changes in product appeal				
Changes in TFP	-0.9158[a]	-1.026[a]	-0.8921[a]	-0.9254[a]	-1.0170[a]	-0.8503[a]	-0.9847[a]	-0.9663[a]	-0.9953[a]
	(0.0345)	(0.0236)	(0.0343)	(0.1298)	(0.0079)	(0.0865)	(0.0100)	(0.0095)	(0.0067)
N Obs	901	843	232	710	692	867	2,000	785	738
R^2	0.6490	0.8827	0.8285	0.5908	0.9772	0.3237	0.9577	0.9411	0.9788

Notes: Quantity TFP and product appeal have been demeaned by 8-digit Prodcom codes before computing growth rates. Time dummies are included in estimations but are not reported here. Bootstrapped standard errors in parenthesis (200 replications). [a] p<0.01, [b] p<0.05, [c] p<0.1.

Figure B-3: Scatter plot of product appeal and an alternative FHS demand measure



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict product appeal (obtained with the FMMM procedure) and an alternative FHS demand measure (computed using FMMM estimates of the production function coefficients, rather than cost shares, to obtain TFP and instrument log price), corresponding to each firm-year pair in our sample, and are constructed after demeaning both product appeal and the alternative FHS demand measure by 8-digit product codes.

C Multi-product firms

There are several issues related to multi-product firms. We focus here on the issue of the assignment of inputs to outputs. Produced quantities and generated revenues may be observable for the different products of each firm in databases like ours. However, information on inputs used for a specific product is typically not available. We propose here an extension of our baseline model to solve the problem of assigning inputs to outputs for multi-product firms. In doing so we assume, as in De Loecker et al. (2016), there is a limited role for economies (or diseconomies) of scope on the cost side. However, contrary to De Loecker et al. (2016), we do not impose multi-product firms to be characterised by a common productivity across the different products they produce. We also allow for firm-product-time specific markups but impose product appeal/perceived quality to be common across products within a firm. This corresponds to a setting where firms can be distinguished into those consistently selling high perceived quality products and those consistently selling low perceived quality products. Yet firms are allowed to be more or less efficient in the production of a specific product and charge different markups. The assumptions we lay down below and the related estimation procedure are consistent with both a monopolistically competitive market structure, like the one developed in Bernard et al. (2011), and the Cournot competition version of the model developed in Hottman et al. (2016) that we discuss in Appendix A.

As usual we denote a firm by i and time by t. A firm i produces in t one or more products indexed by p and the number of products produced by the firm is denoted by I_{it}. In our data p is an 8-digit Prodcom product code but in other data, like the bar-code data used in Hottman et al. (2016), can be much more detailed. We assume product appeal is firm-time specific (λ_{it}) while we allow markups (μ_{ipt}) and productivity (a_{ipt}) to be firm-product-time specific. The production function for product p produced by firm i is given by:

$$Q_{ipt} = C_p C_t A_{ipt} L_{ipt}^{\alpha_{Lg}} M_{ipt}^{\alpha_{Mg}} K_{ipt}^{\gamma_g - \alpha_{Mg} - \alpha_{Lg}}, \tag{C-1}$$

where C_p and C_t are innocuous product and time constants we disregard in what follows and g identifies a product group/industry. Production function coefficients are the same for products within a product group because a certain level of data aggregation is needed to deliver enough observations to estimate parameters. (C-1) means we allow for technology (α_{Lg}, α_{Mg}, γ_g) to differ across the different products p produced by a multi-product firm. At the same time productivity is allowed to vary across products within a firm and information coming from single-product firms need to be used to infer the technology of multi-product firms, i.e. we rule out physical synergies in production but allow for some of the economies (diseconomies) of scope discussed in De Loecker et al. (2016). Furthermore, we assume firm i to maximise profits and choose (for each product p) the amount of labour L_{ipt} and materials M_{ipt} in order to minimise short-term costs while taking capital K_{ipt}, as well as productivity a_{ipt} and product appeal λ_{it} as given. We make use of (8) and so assume:

$$r_{ipt} \simeq \frac{1}{\mu_{ipt}}(q_{ipt} + \lambda_{it}). \tag{C-2}$$

Profit maximisation implies:

$$P_{ipt} = \mu_{ipt} \frac{\partial C_{ipt}}{\partial Q_{ipt}}, \tag{C-3}$$

where marginal cost is equal to[40]

$$\frac{\partial C_{ipt}}{\partial Q_{ipt}} = A_{ipt}^{-\frac{1}{\alpha_{Lg}+\alpha_{Mg}}} Q_{ipt}^{\frac{1-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}} K_{ipt}^{\frac{\gamma_g-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}} \tag{C-4}$$

Firms minimise costs and so markups are such that:

$$\mu_{ipt} = \frac{\alpha_{Mg}}{s_{Mipt}} \tag{C-5}$$

where s_{Mipt} is the expenditure share of materials for product p at time t in firm revenue for product p at time t. Finally, we assume that both a_{ipt} and λ_{it} evolve over time as linear stochastic Markov processes:

$$\begin{aligned} a_{ipt} &= \phi_{ag}\ a_{ipt-1} + \nu_{aipt} \\ \lambda_{it} &= \phi_{\lambda}\lambda_{it-1} + \nu_{\lambda it} \end{aligned}$$

where ν_{aipt} and $\nu_{\lambda it}$ can be correlated with each other.

As far as single-product firms are concerned the above assumptions are such that the parameters of the production function, as well as single-product firms' productivity, product appeal and markups, can be obtained using a variant of the MULAMA procedure. More specifically, if also labour is chosen at t (17) becomes:

$$\begin{aligned} LHS_{ipt} &= \frac{\gamma_g}{\alpha_{Mg}} k_{ipt} + \phi_{ag} LHS_{ipt-1} - \phi_{ag} \frac{\gamma_g}{\alpha_{Mg}} k_{ipt-1} \\ &+ (\phi_\lambda - \phi_{ag}) \left(\frac{r_{ipt-1}}{s_{Mipt-1}} - \frac{1}{\alpha_{Mg}} q_{ipt-1} \right) + \frac{1}{\alpha_{Mg}} \left(\nu_{aipt} + \nu_{\lambda it} \right), \end{aligned} \tag{C-6}$$

[40] We omit the innocuous product-time constant $\left(\frac{W_{Lpt}}{\alpha_{Lg}}\right)^{\frac{\alpha_{Lg}}{\alpha_{Lg}+\alpha_{Mg}}} \left(\frac{W_{Mpt}}{\alpha_{Mg}}\right)^{\frac{\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}}$

where $LHS_{ipt} \equiv \frac{r_{ipt}-s_{Lipt}(l_{ipt}-k_{ipt})-s_{Mipt}(m_{ipt}-k_{ipt})}{s_{Mipt}}$ and s_{Lipt} is the share of labour in revenue ($s_{Lipt} \equiv \frac{W_{Lpt}L_{ipt}}{R_{ipt}}$). One can rewrite (C-6) as the following linear regression:

$$LHS_{ipt} = b_{1g}z_{1ipt} + b_{2g}z_{2ipt} + b_{3g}z_{3ipt} + b_{4g}z_{4ipt} + b_{5g}z_{5ipt} + u_{ipt} \tag{C-7}$$

where z_{1ipt}=k_{ipt}, z_{2ipt}=LHS_{ipt-1}, z_{3ipt}=k_{ipt-1}, z_{4ipt}=$\frac{r_{ipt-1}}{s_{Mipt-1}}$, z_{5ipt}=q_{ipt-1}, u_{ipt}=$\frac{1}{\alpha_{Mg}}\left(\nu_{aipt}+\nu_{\lambda it}\right)$ as well as b_{1g}=$\beta_g \equiv \frac{\gamma_g}{\alpha_{Mg}}$, b_{2g}=ϕ_{ag}, b_{3g}=$-\phi_{ag}\beta_g$, b_{4g}=$(\phi_\lambda - \phi_{ag})$ and b_{5g}=$-\left(\phi_\lambda - \phi_{ag}\right)\frac{1}{\alpha_{Mg}}$. Given our assumptions, the error term u_{ipt} in (C-7) is uncorrelated with all of the regressors. Therefore (C-7) can be estimated via simple OLS. After doing this we set $\hat{\beta}_g$=$\hat{b}_{1g}$ and $\hat{\phi}_{ag}$=$\hat{b}_{2g}$ and do not exploit parameters' constraints in the estimation.

We now turn to estimating γ_g. Within this setting we have $\alpha_{Lg} = \mu_{ipt}s_{Lipt}$ and $\alpha_{Mg} = \mu_{ipt}s_{Mipt}$. (21) becomes:

$$\begin{aligned} q_{ipt} &= \frac{\gamma_g}{\hat{\beta}_g}\frac{s_{Lipt}}{s_{Mipt}}\left(l_{ipt}-k_{ipt}\right) + \frac{\gamma_g}{\hat{\beta}_g}\left(m_{ipt}-k_{ipt}\right) + \gamma_g k_{ipt} \\ &+ \hat{\phi}_{ag}\frac{\gamma_g}{\hat{\beta}_g}LHS_{ipt-1} - \hat{\phi}_{ag}\gamma_g k_{ipt-1} - \hat{\phi}_{ag}\left(r_{ipt-1}\frac{\gamma_g}{\hat{\beta}_g s_{Mipt-1}} - q_{ipt-1}\right) + \nu_{aipt.} \end{aligned} \tag{C-8}$$

(C-8) can be rewritten in a linear way:

$$\overline{LHS}_{ipt} = b_{6g}z_{6ipt} + \nu_{aipt} \tag{C-9}$$

where:

$$\overline{LHS}_{ipt} = q_{ipt} - \hat{\phi}_{ag}q_{ipt-1}$$

$$z_{6ipt} = \frac{1}{\hat{\beta}_g}\frac{s_{Lipt}}{s_{Mipt}}\left(l_{ipt}-k_{ipt}\right) + \frac{1}{\hat{\beta}_g}\left(m_{ipt}-k_{ipt}\right) + k_{ipt} + \frac{\hat{\phi}_{ag}}{\hat{\beta}_g}LHS_{ipt-1} - \hat{\phi}_{ag}k_{ipt-1} - r_{ipt-1}\frac{\hat{\phi}_{ag}}{\hat{\beta}s_{Mipt-1}}$$

as well as b_{6g}=γ_g and z_{6ipt} can instrumented with with k_{ipt} as well as past inputs, revenue and quantity. We set $\hat{\gamma}_g$=$\hat{b}_{6g}$ and are in turn able to estimate productivity as:

$$\hat{a}_{ipt} = q_{ipt} - \frac{\hat{\gamma}_g}{\hat{\beta}_g}\frac{s_{Lipt}}{s_{Mipt}}\left(l_{ipt}-k_{ipt}\right) - \frac{\hat{\gamma}_g}{\hat{\beta}_g}\left(m_{ipt}-k_{ipt}\right) - \hat{\gamma}_g k_{ipt}, \tag{C-10}$$

while product appeal and markups are computed as in the baseline procedure from (C-2) and (C-5).

Estimations need to be carried on single-product firms separately for each product group g. Turning to multi-product firms we impose, as in De Loecker et al. (2016), that the same technology parameters coming from single-product producers extend to the products of the former. Yet, in order to quantify multi-product firms' productivity, markups and product appeal we still need to solve the issue of how to assign inputs to outputs and we do so by building on the above assumptions. As far as materials are concerned, we need to assign the observable total firm material expenditure M_{it} across the I_{it} products produced by firm i at time t, i.e. we need to assign values to M_{ipt} such that $\sum_{p=1}^{I_{it}} M_{ipt} = M_{it}$. We can use this condition along with (C-5) and (C-2) to operate this assignment. Substituting (C-5) into (C-2) and adding $\sum_{p=1}^{I_{it}} M_{ipt} = M_{it}$ provides a system of $I_{it} + 1$ equations in $I_{it} + 1$ unknowns; the I_{it} inputs expenditures M_{ipt} plus λ_{it}. Indeed, at this stage we have data on r_{ipt}, q_{ipt}, α_{Mg} and M_{it}. Operationally, one can actually proceed in two stages. Combining the above equations one has $\sum_{p=1}^{I_{it}} \frac{\alpha_{Mg} r_{ipt} R_{ipt}}{q_{ipt}+\lambda_{it}} = M_{it}$. This equation can be solved for each firm and delivers λ_{it}. With this at hand one can then obtain materials expenditure from $M_{ipt} = \frac{\alpha_{Mg} r_{ipt} R_{ipt}}{q_{ipt}+\lambda_{it}}$. By recovering inputs expenditures M_{ipt} we can subsequently compute materials expenditure shares in revenues s_{Mipt} and so use (C-5) to recover our firm-product-time specific markups μ_{ipt}. Since labour is a variable input a condition analogous to (C-5) holds for this input and so we can use the computed markups μ_{ipt} and information on α_{Lg} to derive labour expenditure: $L_{ipt} = \frac{\alpha_{Lg} R_{ipt}}{\mu_{ipt}}$.[41] In the data, this is not guaranteed to satisfy the constraint $\sum_{p=1}^{I_{it}} L_{ipt} = L_{it}$ for each firm and so the L_{ipt} need to be re-scaled for each firm.

So far, the above procedure yields markups and product appeal, as well as information on labour and materials use, for each of the products of a multi-product firm. However, in order to recover productivity a_{ipt} we still need values for capital K_{ipt}. To do this one can proceed as follows. Combining the marginal cost, profit maximisation and quantity equations one gets:

$$K_{ipt} = \left(\frac{P_{ipt}}{\mu_{ipt} Q_{ipt}^{a+b} L_{ipt}^{-a\alpha_{Lg}} M_{ipt}^{-a\alpha_{Mg}}} \right)^{\left(\frac{1}{c-a\alpha_{Kg}}\right)}$$

where $a = -\frac{1}{\alpha_{Lg}+\alpha_{Mg}}$, $b = \frac{1-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}$, $c = \frac{\gamma_g-\alpha_{Lg}-\alpha_{Mg}}{\alpha_{Lg}+\alpha_{Mg}}$ and $\alpha_{Kg} = \gamma_g - \alpha_{Mg} - \alpha_{Lg}$ is the capital coefficient. Again values need to be re-scaled for each firm to meet the constraint $\sum_{p=1}^{I_{it}} K_{ipt} = K_{it}$ and even further refined by running an estimation where the computed K_{ipt} is regressed on R_{ipt}, M_{ipt}, L_{ipt} as well as total firm expenditure on materials and labour plus the capital stock and product dummies.

[41] As a matter of fact in this variant of the model we do not impose α_{Lg} to be the same across firms. From every single-product firm, using the computed markups and the observed labour expenditure share in revenue, equation (C-5) applied to labour delivers a different α_{Lg}. One can compute the mean value of these coefficients across firms producing products belonging to g to get a unique α_{Lg}.

D Key results obtained with the DGKP estimation procedure

For the DGKP procedure, we estimate a Translog production function and perform estimations separately for each two-digit industry while considering a full battery of 8-digit product dummies as well as year dummies. The average, across all observations, markup from the DGKP procedure is 1.158. The correlation between product appeal obtained with the FMMM and DGKP procedures is extremely high: 0.983 across all observations and 0.908 once demeaning both product appeal measures by 8-digit product codes.

Table D-1: Estimates of the mean and standard deviation of output elasticities for the Translog production function obtained with the DGKP procedure

Industry	Description	Statistic	Labour	Materials	Capital	γ
1	Food products, beverages and tobacco	Mean	0.180	0.638	0.039	0.857
		St. Dev.	0.110	0.121	0.060	0.094
2	Textiles and leather	Mean	0.221	0.776	0.018	1.016
		St. Dev.	0.082	0.069	0.017	0.022
3	Wood except furniture	Mean	0.268	0.670	0.016	0.953
		St. Dev.	0.120	0.208	0.018	0.096
4	Pulp, paper, publishing and printing	Mean	0.215	0.817	0.017	1.049
		St. Dev.	0.100	0.154	0.064	0.070
5	Chemicals and rubber	Mean	0.168	0.708	0.037	0.913
		St. Dev.	0.067	0.152	0.039	0.142
6	Other non-metallic mineral products	Mean	0.186	0.762	0.023	0.971
		St. Dev.	0.075	0.081	0.042	0.050
7	Basic metals and fabric. metal prod.	Mean	0.310	0.636	0.030	0.976
		St. Dev.	0.137	0.142	0.019	0.015
8	Machinery, electric. and optical equip.	Mean	0.289	0.675	0.018	0.981
		St. Dev.	0.093	0.098	0.015	0.008
9	Transport equipment and n.e.c.	Mean	0.184	0.780	0.025	0.988
		St. Dev.	0.220	0.215	0.033	0.039

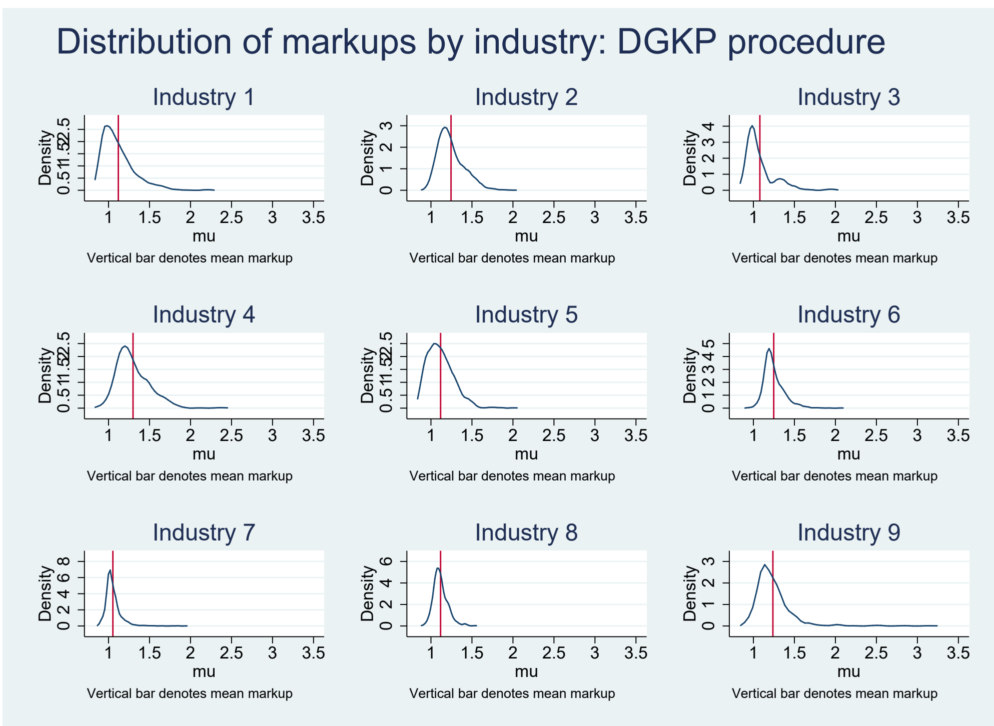
Notes: γ denotes returns to scale.

Table D-2: Estimates of the Translog production function parameters with the DGKP procedure

Industry	1	2	3	4	5	6	7	8	9
α_L	0.395^a	0.379^a	0.461^a	0.414^a	0.327^a	0.305^a	0.485^a	0.432^a	0.498^a
	(0.025)	(0.023)	(0.014)	(0.019)	(0.029)	(0.019)	(0.004)	(0.014)	(0.008)
α_M	0.577^a	0.671^a	0.395^a	0.507^a	0.458^a	0.634^a	0.438^a	0.539^a	0.423^a
	(0.138)	(0.056)	(0.032)	(0.058)	(0.140)	(0.047)	(0.010)	(0.026)	(0.036)
α_K	-0.009	0.010	0.044^a	0.124^a	-0.027	0.068	0.048^a	0.001	0.063^b
	(0.082)	(0.026)	(0.015)	(0.037)	(0.116)	(0.045)	(0.007)	(0.016)	(0.025)
α_{LL}	0.167^a	0.157^a	0.182^a	0.137^a	0.109^a	0.129^a	0.260^a	0.177^a	0.385^a
	(0.026)	(0.043)	(0.018)	(0.014)	(0.027)	(0.018)	(0.005)	(0.021)	(0.006)
α_{MM}	0.022	0.103	0.285^a	0.246^a	0.177^c	0.132^a	0.262^a	0.180^a	0.391^a
	(0.178)	(0.078)	(0.031)	(0.035)	(0.095)	(0.045)	(0.009)	(0.023)	(0.027)
α_{KK}	-0.016	-0.009	0.008	0.027^b	-0.022	0.041^b	0.021^a	-0.014^b	0.030^a
	(0.049)	(0.010)	(0.008)	(0.011)	(0.061)	(0.017)	(0.003)	(0.007)	(0.010)
α_{KM}	0.051	0.008	-0.027^c	-0.072^a	0.046	-0.038	-0.005	0.005	-0.021
	(0.074)	(0.023)	(0.016)	(0.021)	(0.087)	(0.036)	(0.007)	(0.010)	(0.016)
α_{KL}	0.012	0.015^b	0.016^b	0.014^b	0.004	-0.020^c	-0.014^a	0.006	-0.021^a
	(0.011)	(0.006)	(0.007)	(0.007)	(0.008)	(0.011)	(0.002)	(0.006)	(0.003)
α_{ML}	-0.181^a	-0.148^a	-0.193^a	-0.172^a	-0.121^a	-0.117^a	-0.246^a	-0.180^a	-0.379^a
	(0.020)	(0.026)	(0.012)	(0.014)	(0.020)	(0.018)	(0.005)	(0.018)	(0.007)

Notes: Firm-level clustered standard errors in parenthesis. ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Figure D-1: Distribution of markups obtained with the DGKP procedure by industry



Notes: The figure provides the density distribution of markups across observations separately for each of the 9 industries we consider. The vertical bar denotes the mean markup. Markups are obtained with the DGKP estimation procedure.

Table D-3: Standard deviation of quantity TFP, product appeal and markups by industry (DGKP procedure)

Industry	Description	TFP	product appeal	markups
1	Food products, beverages and tobacco	0.430	0.545	0.207
2	Textiles and leather	0.588	0.637	0.111
3	Wood except furniture	0.842	0.905	0.110
4	Pulp, paper, publishing and printing	0.769	0.975	0.129
5	Chemicals and rubber	0.946	0.984	0.114
6	Other non-metallic mineral products	0.500	0.575	0.093
7	Basic metals and fabric. metal prod.	0.855	0.861	0.063
8	Machinery, electric. and optical equip.	0.914	0.916	0.046
9	Transport equipment and n.e.c.	1.017	1.047	0.170

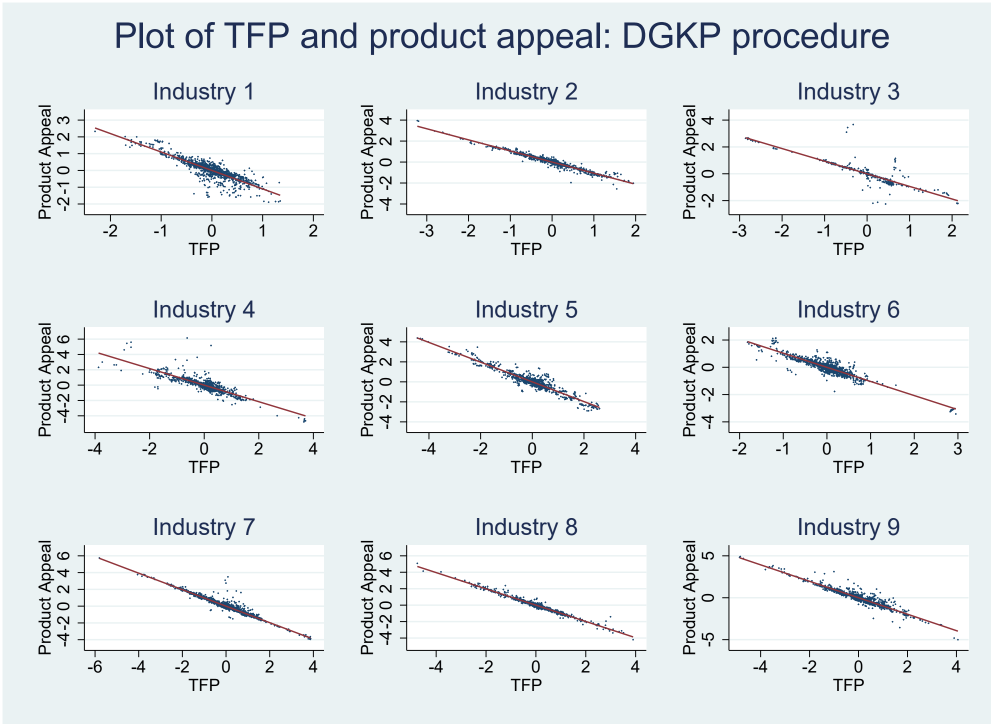
Notes: TFP and product appeal are demeaned by 8-digit Prodcom codes.

Table D-4: Correlations between quantity TFP, product appeal, markups and log prices (DGKP procedure)

	TFP	λ	markups	prices
TFP	1			
λ	-0.943^a	1		
markups	-0.024^c	0.130^a	1	
prices	-0.995^a	0.935^a	0.017	1

Notes: quantity TFP, product appeal and (log) prices are demeaned by 8-digit Prodcom codes. ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1.

Figure D-2: Within 8-digit products correlation between quantity TFP and product appeal by industry (DGKP procedure)



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict quantity TFP and product appeal, corresponding to each firm-year pair in our sample, and are constructed after demeaning both quantity TFP and product appeal by 8-digit product codes. Quantity TFP and product appeal are obtained with the DGKP estimation procedure.

Table D-5: Analysis of demeaned log prices (DGKP procedure)

quantity	-0.374[a]	
	(0.055)	
TFP		-0.934[a]
		(0.015)
λ	0.582[a]	0.061[a]
	(0.065)	(0.015)
capital		-0.003[c]
		(0.002)
Estimation method	First differences	
Year dummies	Yes	Yes
N Obs	7,768	7,768
R^2	0.954	0.996

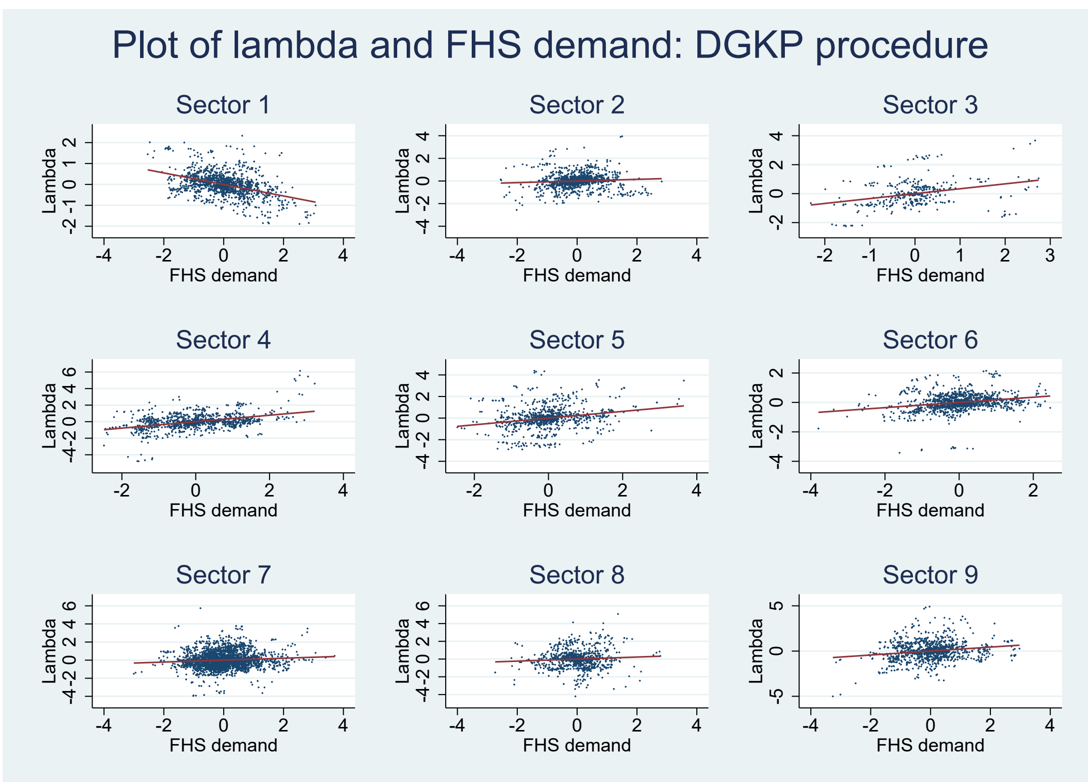
Notes: The dependent variable is the demeaned log price. (Log) prices, quantity TFP, product appeal, (log) capital and (log) quantity are demeaned by 8-digit Prodcom codes. Bootstrapped standard errors in parenthesis (200 replications). [a] p<0.01, [b] p<0.05, [c] p<0.1.

Table D-6: Analysis of markups (DGKP procedure)

quantity	0.076[a]	
	(0.011)	
TFP		0.257[a]
		(0.027)
λ	0.090[a]	0.258[a]
	(0.012)	(0.027)
capital		0.004
		(0.003)
Estimation method	First differences	
Year dummies	Yes	Yes
Prod dummies	Yes	Yes
N Obs	7,768	7,768
R^2	0.271	0.548

Notes: The dependent variable is the markup. Quantity and capital are expressed in logs while TFP indcated quantity TFP. Bootstrapped standard errors in parenthesis (200 replications). [a] p<0.01, [b] p<0.05, [c] p<0.1.

Figure D-3: Scatter plot of product appeal and the FHS demand measure (DGKP procedure)



Notes: The figure provides a scatter plot (as well as a best fit line) for each of the 9 industries we consider. Scatter plots depict product appeal (computed using the DGKP procedure) and the FHS demand measure, corresponding to each firm-year pair in our sample, and are constructed after demeaning both product appeal and the FHS demand measure by 8-digit product codes.

E A model of selection with underlying heterogeneity in both TFP and product appeal

In what follows we provide key highlights of a Meltitz-type selection model with underlying heterogeneity in both TFP and product appeal based on Behrens et al. (2013). The model delivers a cutoff rule for surviving firms based on 'quality adjusted' productivity $\widetilde{tfp} = a + \lambda$ rather than just productivity a. The reader with limited interest in the details of the model might directly jump to Section E.4 where, conditional on the existence of a cutoff survival rule based on $a + \lambda$, we show that, even if draws of a and λ are independent, a and λ will be negatively correlated in the sample of surviving firms. Intuitively, when comparing two sub-samples of surviving firms with different average product appeal λ, the group with the higher average λ needs less higher values of a to satisfy the cutoff survival rule so generating the negative correlation.

Consider an economy with $\mathcal{L}$ identical workers/consumers. Labour is the only factor of production.

E.1 Preferences and demands

There is a final consumption good provided as a continuum of differentiated varieties. More specifically, consumers have identical CARA preferences displaying 'love of variety' and giving rise to demands with variable elasticity. Let $P(i)$ and $Q(i)$ denote the price and the per capita consumption of variety i. The utility maximisation problem of the representative consumer is given by:

$$\max_{Q(j),\, j\in\Omega} \; U \equiv \int_{\Omega} \left[1 - \mathrm{e}^{-\alpha \Lambda(j) Q(j)}\right] \mathrm{d}j \qquad \text{s.t.} \quad \int_{\Omega} P(j)Q(j)\mathrm{d}j = E, \tag{E-1}$$

where Ω denotes the endogenously determined set of varieties, E denotes consumption expenditure and $\Lambda(j) > 0$ is perceived quality/product appeal. The representative consumer buys quantity $Q(j)$, while paying a price $P(j)$, that effectively translates into a quality-adjusted quantity $\tilde{Q}(j) = \Lambda(j)Q(j)$ in the utility function. Indeed, because $\Lambda(j)$ is given to the representative consumer, the utility maximisation problem (E-1) is equivalent to:

$$\max_{\tilde{Q}(j),\, j\in\Omega} \; U \equiv \int_{\Omega} \left[1 - \mathrm{e}^{-\alpha \tilde{Q}(j)}\right] \mathrm{d}j \qquad \text{s.t.} \quad \int_{\Omega} \tilde{P}(j)\tilde{Q}(j)\mathrm{d}j = E, \tag{A-1'}$$

where the representative consumer chooses quality-adjusted quantities $\tilde{Q}(j)$ and the price $\tilde{P}(j) = P(j)/\Lambda(j)$ is the quality-adjusted price of variety j with $\tilde{P}(j)\tilde{Q}(j) = P(j)Q(j)\ \forall j$.

The utility maximisation problem (A-1') now involves perfectly symmetric varieties and so the maths in Behrens et al. (2014) can be applied to (A-1') to derive a number of result. More specifically, solving (A-1') yields the following demand functions:

$$\tilde{Q}(i) = \frac{E}{N^c \overline{\tilde{P}}} - \frac{1}{\alpha}\left\{ \ln\left[\frac{\tilde{P}(i)}{N^c \overline{\tilde{P}}}\right] + \eta \right\}, \quad \forall i \in \Omega, \tag{E-2}$$

where N^c is the mass of varieties consumed and

$$\overline{\tilde{P}} \equiv \frac{1}{N^c}\int_{\Omega} \tilde{P}(j)\mathrm{d}j \quad \text{and} \quad \eta \equiv -\int_{\Omega} \ln\left[\frac{\tilde{P}(j)}{N^c \overline{\tilde{P}}}\right]\frac{\tilde{P}(j)}{N^c \overline{\tilde{P}}}\mathrm{d}j$$

denote the average quality-adjusted price and the differential entropy of the quality-adjusted price distribution, respectively. Since marginal utility at zero consumption is bounded, the demand for a variety need not be positive. Formally,

$$\tilde{Q}(i) > 0 \iff \tilde{P}(i) < \tilde{P}^d$$

where $\tilde{P}^d \equiv N^c \overline{\tilde{P}} \mathrm{e}^{\alpha E/(N^c \overline{\tilde{P}}) - \eta}$ is a reservation price that depends on the quality-adjusted price aggregates $\overline{\tilde{P}}$ and η. The definition of the reservation price allows to express the demands for varieties concisely as follows:

$$\tilde{Q}(i) = \frac{1}{\alpha}\ln\left[\frac{\tilde{P}^d}{\tilde{P}(i)}\right]. \tag{E-3}$$

Going back to the original prices and quantities delivers:

$$Q(i) = \frac{1}{\alpha\Lambda(i)}\ln\left[\frac{\Lambda(i)\tilde{P}^d}{P(i)}\right]. \tag{A-3'}$$

E.2 Technology and market structure

Prior to production, firms engage in research and development. The labour market is perfectly competitive, so that all firms take the wage rate W as given. Entry requires a fixed amount F of labour paid at the market wage. Each firm i discovers its marginal labour requirement $C(i) = 1/A(i) \geq 0$, i.e. the inverse of quantity TFP, and the perceived quality of its variety $\Lambda(i) \geq 0$ only after making this irreversible entry decision. We assume that $C(i)$ and $\Lambda(i)$ are drawn from two known, continuously differentiable distributions G^C and G^Λ. The above assumptions mean that each individual firm faces uncertainty over its productivity and the perceived quality/product appeal of its product. However, in the aggregate productivity and product appeal are predictable and in particular it is possible to derive the average firm productivity and the average product appeal.

In what follows it is not necessary to make any strong assumptions regarding the relationship between G^C and G^Λ. Indeed, all is needed is that the quality-adjusted marginal labour requirement ratio $\tilde{C}(i) = C(i)/\Lambda(i)$ is a non-degenerate random variable that we assume to be characterised by a continuously differentiable distributions $G^{\tilde{C}}$. In particular, in order to facilitate deriving analytical results, we assume that quality-adjusted productivity $\widetilde{TFP}(i) = 1/\tilde{C}(i) = A(i)\Lambda(i)$ is Pareto distributed implying that $G^{\tilde{C}}(\tilde{C}) = \left(\tilde{C}/\tilde{C}^{\max}\right)^k$, where $\tilde{C} \in (0, \tilde{C}^{\max}]$ while $\tilde{C}^{\max} > 0$ and $k \geq 1$ are the upper bound and the shape parameter, respectively.

Since entry costs are sunk, firms will survive/operate provided they can charge prices $P(i)$ above marginal costs $C(i)W$. The operating profit of a firm i is as follows:

$$\Pi(i) = LQ(i)\left[P(i) - C(i)W\right], \tag{E-4}$$

where demand $Q(i)$ is given by (A-3'). Each surviving firm maximises (E-4) with respect to its price $P(i)$. Since there is a continuum of firms, no individual firm has any impact on $\tilde{P}^d$, so that the first-order conditions for (operating) profit maximisation are given by:

$$\ln\left[\frac{\Lambda(i)\tilde{P}^d}{P(i)}\right] = \frac{P(i) - C(i)W}{P(i)}, \quad \forall i \in \Omega. \tag{E-5}$$

A price distribution satisfying (E-5) is called a *price equilibrium.* Equations (A-3') and (E-5) imply that $Q(i) = (1/\alpha\Lambda(i))[1 - C(i)W/P(i)]$. The minimum output that a firm may sell is clearly given by $Q(i) = 0$ at $P(i) = C(i)W$. This, by (E-5), implies that $P(i)/\Lambda(i) = \tilde{P}(i) = \tilde{P}^d$. Overall this means that, in order to find it profitable to operate, firms need to have a cost draw $C(i)$ and a product appeal draw $\Lambda(i)$ such that their marginal cost $C(i)W$ is lower or equal to their price while at the same time their quality-adjusted price $P(i)/\Lambda(i)$ should be lower or equal to the choke quality-adjusted price $\tilde{P}^d$. This ultimately implies a cutoff condition in terms of quality-adjusted marginal labour requirement ($C(i)W = P(i) = \tilde{P}^d\Lambda(i)$ or $\tilde{C}(i) = \tilde{P}^d/W$) such that only firms with quality-adjusted marginal labour requirement lower or equal to $\tilde{C}^* = \tilde{P}^d/W$ find it profitable to sell.

Given the cutoff $\tilde{C}^*$ and a mass of entrants N^E, only $N^c = N^E\, G^{\tilde{C}}\left(\tilde{C}^*\right)$ firms survive, namely those which are productive enough and/or have products of high enough quality to sell. Since all firms differ only by their quality-adjusted marginal labour requirements, we can express all firm-level variables in terms of $\tilde{C}$. In this respect, it should be noted that by multiplying and dividing by $\Lambda(i)$, equations (E-4) and (E-5) can be rewritten as:

$$\Pi(i) = \mathcal{L}\tilde{Q}(i)\left[\tilde{P}(i) - \tilde{C}(i)W\right], \tag{A-4'}$$

and

$$\ln\left[\frac{\tilde{P}^d}{\tilde{P}(i)}\right] = \frac{\tilde{P}(i) - \tilde{C}(i)W}{\tilde{P}(i)}, \quad \forall i \in \Omega. \tag{A-5'}$$

Equations (E-3), (A-4') and (A-5') are identical to those reported in Behrens et al. (2014) once prices, quantities and marginal costs are replaced with their quality-adjusted counterparts. In other words, our model is isomorphic to Behrens et al. (2014) and so we can directly apply a number of their results.

E.3 Equilibrium

In what follows we normalise the wage W to one to simplify the exposition. In equilibrium, aggregate profits are zero and the labour market should clear. Using the Pareto assumption for the distribution of quality-adjusted productivity $\widetilde{TFP}$, we obtain, parallel to Behrens et al. (2014), the following closed-form solutions for the equilibrium cutoff and the mass of entrants:

$$\tilde{C}^* = \left(\frac{\tilde{T}^{\max}}{\mathcal{L}}\right)^{\frac{1}{k+1}} \quad \text{and} \quad N^E = \frac{\kappa_2}{\kappa_1 + \kappa_2}\frac{\mathcal{L}}{F}, \tag{E-6}$$

where $\tilde{T}^{\max} \equiv \left[\alpha F(\tilde{C}^{\max})^k\right]/\kappa_2$ while κ_1 and κ_2 are positive constants that solely depend on k. Finally, the indirect utility can be expressed as:

$$U = \alpha\left[\frac{1}{(k+1)(\kappa_1 + \kappa_2)} - 1\right]\frac{1}{\tilde{C}^*}, \tag{E-7}$$

where the term in square brackets is, by construction, positive for all $k \geq 1$.

E.4 Relationship between TFP and product appeal

Suppose that TFP A_i and product appeal Λ_i are independent. This implies that a_i and λ_i are also independent, i.e. $G^{a,\lambda}(\mathcal{X},\mathcal{Y})$=$G^a(\mathcal{X})G^\lambda(\mathcal{Y})$ where $G^{a,\lambda}(.,.)$ is the joint cumulative distribution of a_i and λ_i, $G^a(.)$ and $G^\lambda(.)$ are the univariate cumulative distribution of a_i and λ_i, and $\mathcal{X}$ and $\mathcal{Y}$ are two generic numbers.

Because of this, the expectation of a_i conditional on λ_i is not a function of λ_i: $\mathbb{E}(a_i|\lambda_i) = \mathbb{E}(a_i)$. However, the expectation of a_i conditional on λ_i and the cutoff survival rule is a function of λ_i. To show this note that the cutoff survival rule imposes that only firms with quality-adjusted marginal labour requirement lower or equal to $\tilde{C}^*$, i.e. firms with (log) quality-adjusted productivity higher or equal to $\widetilde{tfp}^* = \log(1/\tilde{C}^*) = -\tilde{c}^*$, will find it profitable to operate. We indicate this by an indicator function s_i taking value 1 if the cutoff survival rule is satisfied and zero otherwise: $s_i = 1[a_i + \lambda_i \geq \widetilde{tfp}^*]$, where $a_i + \lambda_i = \widetilde{tfp}_i$. Using the definition of conditional probability, the probability that a_i is lower or equal to a given value $\mathcal{X}$ conditional on λ_i and the selection rule s_i being satisfied is:

$$P(a_i \leq \mathcal{X}|\lambda_i, s_i = 1) = \frac{P(a_i \leq \mathcal{X}, s_i = 1|\lambda_i)}{P(s_i = 1|\lambda_i)}. \tag{E-8}$$

The denominator of (E-8) is:

$$P(s_i = 1|\lambda_i) = P(a_i+\lambda_i \geq \widetilde{tfp}^*|\lambda_i) = P(a_i \geq \widetilde{tfp}^*-\lambda_i|\lambda_i) = P(a_i \geq \widetilde{tfp}^*-\lambda_i) = 1-G^a(\widetilde{tfp}^*-\lambda_i),$$

where the penultimate equality comes from the independence hypothesis. The numerator of (E-8) is instead:

$$P(a_i \leq \mathcal{X}, s_i = 1|\lambda_i) = P(\widetilde{tfp}^*-\lambda_i \leq a_i \leq \mathcal{X}|\lambda_i) = P(\widetilde{tfp}^*-\lambda_i \leq a_i \leq \mathcal{X}) = G^a(\mathcal{X})-G^a(\widetilde{tfp}^*-\lambda_i)$$

where, again, the penultimate equality comes from the independence hypothesis. Putting the above two results together provides:

$$P(a_i \leq \mathcal{X}|\lambda_i, s_i = 1) = \frac{P(a_i \leq \mathcal{X}, s_i = 1|\lambda_i)}{P(s_i = 1|\lambda_i)} = \frac{G^a(\mathcal{X}) - G^a(\widetilde{tfp}^* - \lambda_i)}{1 - G^a(\widetilde{tfp}^* - \lambda_i)}. \tag{E-9}$$

Taking the derivative of (E-9) with respect to $\mathcal{X}$ delivers the density function:

$$g^{a|\lambda,s=1}(\mathcal{X}) = \frac{g^a(\mathcal{X})}{1 - G^a(\widetilde{tfp}^* - \lambda_i)}.$$

Therefore, the expectation of a_i conditional on λ_i and the selection rule s_i being satisfied is:

$$\mathbb{E}(a_i|\lambda_i, s_i = 1) = \int \mathcal{X} g^{a|\lambda,s=1}(\mathcal{X}) d\mathcal{X} = \frac{\mathbb{E}(a_i)}{1 - G^a(\widetilde{tfp}^* - \lambda_i)},$$

which is a decreasing function of λ_i, i.e. the higher is λ_i the lower is the expected value of a_i for observation belonging to the surviving sample of firms.

F Monte Carlo Analysis

We implement the FMMM estimation framework on artificially generated data.[42] For this we use the following setup: we draw a data set with N firms which we observe over T periods. Markups are defined as

$$\mu_{it} = 1 + \alpha_{\mu,i} + \sigma_{\nu_\mu}\nu_{\mu,it}$$

where $\alpha_{\mu,i} = \gamma + \sigma_{\alpha_\mu} \times \nu_{\alpha,\mu,i}$ is a component that is fixed over time and $\nu_{\mu,it}$ is a time varying component. We draw $\nu_{\alpha,\mu,i}$ and $\nu_{\mu,it}$ from the uniform distribution. γ is the scale parameter and our formulation ensures that $\mu > \gamma$ which is a requirement for nonzero output.

We assume that both (the log of) demand λ_{it} and technology shocks a_{it} evolve as AR(1) processes:

$$a_{it} = \phi_a a_{it-1} + \nu_{a,it}$$

and

$$\lambda_{it} = \phi_\lambda \lambda_{it-1} + \nu_{\lambda,it}$$

where we draw $\nu_{\lambda,it}$ and $\nu_{a,it}$ from the normal distributions $N(0, \sigma_{\nu,a})$ and $N(0, \sigma_{\lambda,a})$. For the initial values we assume that $a_{i0} = \lambda_{i0} = 0$. Normalising the economy wide variables to 1 we can write demand as

$$Q_{it} = \Lambda_{it}^{\eta_{it}-1} P_{it}^{\eta_{it}}$$

where $\eta_{it} = \frac{\mu_{it}}{\mu_{it}-1}$ is the price elasticity of demand. We can invert the demand function as

$$P_{it} = \Lambda_{it}^{\frac{\eta_{it}-1}{\eta_{it}}} Q_{it}^{-\frac{1}{\eta_{it}}}$$

In the remainder we drop firm and time indices to avoid notational clutter. The short run firm level profits maximisaiton problem becomes $V(K, L) = \max_M \left\{ \Lambda^{\frac{1}{\mu}} Q^{\frac{1}{\mu}} - MW_M \right\}$ subject to the production function $Q = AK^{\alpha_K} L^{\alpha_L} M^{\alpha_M}$ where $\alpha_K = \gamma - \alpha_M - \alpha_L$ First order conditions require:

$$M = \frac{\alpha_M Q^{\frac{1}{\mu}}}{\mu W_M} \Lambda^{\frac{1}{\mu}}$$

We can plug this into the production function $Q = AK^{\alpha_K} L^{\alpha_L} \left(\frac{\alpha_M Q^{\frac{1}{\mu}}}{\mu W_M} \Lambda^{\frac{1}{\mu}} \right)^{\alpha_M}$ and solve for Q:

$$Q^* = \left[AK^{\alpha_K} L^{\alpha_L} \left(\frac{\alpha_M}{W_M} \right)^{\alpha_M} \frac{\Lambda^{\frac{\alpha_M}{\mu}}}{\mu^{\alpha_M}} \right]^{\frac{\mu}{\mu - \alpha_M}} \tag{F-1}$$

[42]We provide an interactive version of this under https://mondpanther.shinyapps.io/MulamaRoulette/

Also note that if capital and labour were flexible in the short run we would have

$$L = \frac{\alpha_L Q^{\frac{1}{\mu}}}{\mu W_L} \Lambda^{\frac{1}{\mu}} \tag{F-2}$$

$$K = \frac{\alpha_K Q^{\frac{1}{\mu}}}{\mu W_K} \Lambda^{\frac{1}{\mu}} \tag{F-3}$$

where W_K is the user cost of capital and $\alpha_K = \gamma - \alpha_L - \alpha_M$. With this we can work out optimal output Q^{FLEX} - the amount of output if all factors were flexible - as $Q^{FLEX} = \left[A \left(\frac{\alpha_L}{W_L} \right)^{\alpha_L} \left(\frac{\alpha_M}{W_M} \right)^{\alpha_M} \left(\frac{\alpha_K}{W_K} \right)^{\alpha_K} \frac{\Lambda^{\frac{\alpha_K+\alpha_L+\alpha_M}{\mu}}}{\mu^{\alpha_K+\alpha_L+\alpha_M}} \right]^{\frac{\mu}{\mu-\alpha_L-\alpha_M-\alpha_K}}$ We assume that labor and capital evolve converging to the level consistent with flexible levels but with a random deviations ν_L and ν_K so that

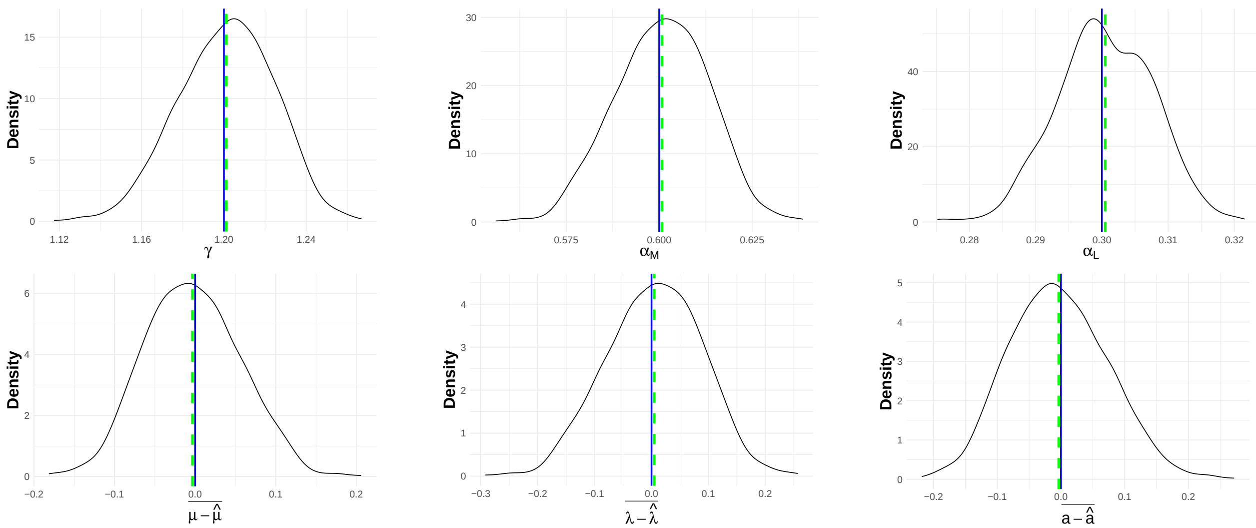
$$L = LAG \left(\frac{\alpha_L \left(Q^{FLEX} \right)^{\frac{1}{\mu}}}{\mu W_L} \Lambda^{\frac{1}{\mu}} \right) \times \exp \nu_L \tag{F-4}$$

$$K = LAG \left(\frac{\alpha_K \left(Q^{FLEX} \right)^{\frac{1}{\mu}}}{\mu W_K} \Lambda^{\frac{1}{\mu}} \right) \times \exp \nu_K \tag{F-5}$$

where $LAG(\cdot)$ is the lag operator; i.e. current levels of capital and labor will be un-correlated with the shocks arising in the current period. ν_L and ν_K are random shocks we draw from then normal distributions $N(0, \sigma_{\nu_L})$ and $N(0, \sigma_{\nu_K})$. Plugging equations (F-4) and (F-5) into (F-1) we can work out Q^*, prices, revenue, material inputs etc. We then can apply the FMMM estimation framework to recover firstly the parameters and secondly, the various shocks; i.e. we run the regressions described in equations (18) and (22).

We draw 1000 samples of 500 firms with 10 time periods of observations. We assume the following parameter values: $\alpha_L = 0.3$, $\alpha_M = 0.6$, $\gamma = 1.2$, $\sigma_{\nu_a} = 0.25$, $\sigma_{\nu_\lambda} = 0.25$, $\sigma_{\nu_\mu} = 0.25$, $\sigma_{\alpha_\mu} = 1$, σ_{ν_K}=1, σ_{ν_L}=1, $\phi_\lambda = 0.5$, $\phi_a = 0.4$. We illustrate the result of this by reporting density plots of the estimated parameters γ, α_M and α_L in the first row of panels in Figure F-1. In all cases we see that this leads to unbiased estimates of the true parameter. Similarly we report density plots of the average difference between estimated and actual markup μ, demand λ and productivity a shocks in the second row of Figure F-1. In each case we see that there are no systematic differences between estimates and actual values. Note that our identification relies on the presence of at least one on quasi fixed production factor that is pre-determined as of time t. It is instructive to conduct a Monte-Carlo exploration of a setting where this assumption goes wrong. Hence we also create artificial data violating this assumption. For that we compute the quasi fixed factors as

Figure F-1: Monte Carlo Results



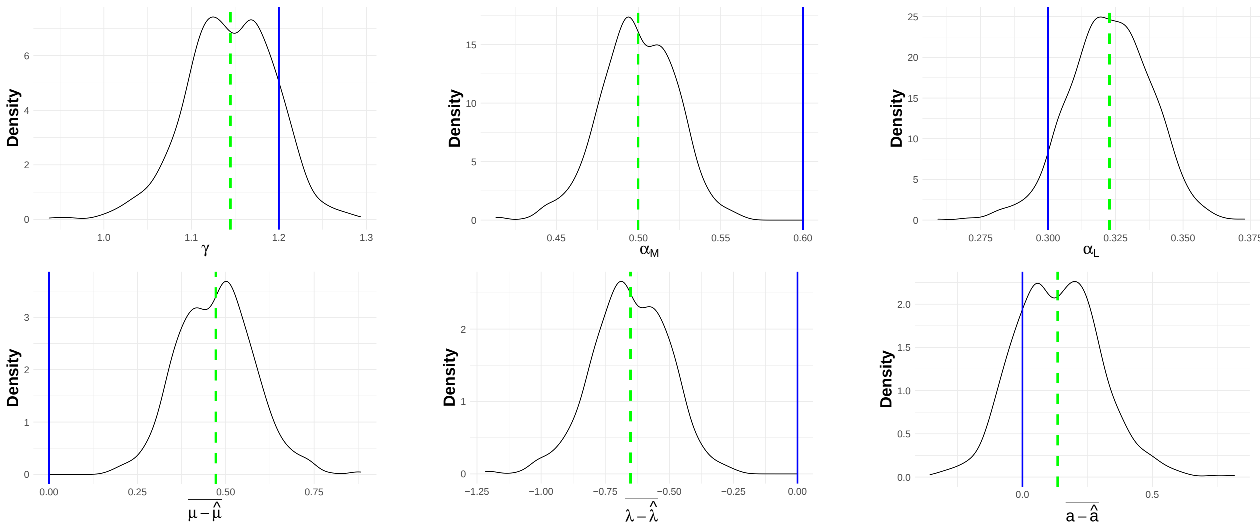
Notes: The figure reports results from a Monte-Carlo analysis of the FMMM estimation framework. We draw 1000 replications of samples of 500 firms over 10 time periods with the following parameter values: $\alpha_L = 0.3n$, $\alpha_M = 0.6$,$\gamma = 1.2$, $\sigma_{\nu_a} = 0.25$, $\sigma_{\nu_\lambda} = 0.25$, $\sigma_{\nu_\mu} = 0.25$, $\sigma_{\alpha_\mu} = 1$, σ_{ν_K}=1, σ_{ν_L}=1, $\phi_\lambda = 0.5$, $\phi_a = 0.4$. Solid vertical lines indicator the true value, dashed vertical lines indicate the average estimated values.

$$L = LAG\left(\frac{\alpha_L\left(Q^{FLEX}\right)^{\frac{1}{\mu}}}{\mu W_L}\Lambda^{\frac{1}{\mu}}\right) \times \exp\nu_L \times (1-b) + b \times \left(\frac{\alpha_L\left(Q^{FLEX}\right)^{\frac{1}{\mu}}}{\mu W_L}\Lambda^{\frac{1}{\mu}}\right)$$

$$K = LAG\left(\frac{\alpha_K\left(Q^{FLEX}\right)^{\frac{1}{\mu}}}{\mu W_K}\Lambda^{\frac{1}{\mu}}\right) \times \exp\nu_K \times (1-b) + b \times \left(\frac{\alpha_K\left(Q^{FLEX}\right)^{\frac{1}{\mu}}}{\mu W_K}\Lambda^{\frac{1}{\mu}}\right)$$

i.e. we allow both to be correlated with contemporaneous shocks to productivity where $b \in [0, 1)$ controls the strength of this correlation. We report results of this in Figure F-2 with $b = 0.4$. This leads to biases in all parameters as well as the markup, demand and productivity shocks. In more detail, note that it leads to a downward bias in α_M. This is because in equation (18) we assign too much of the variation to k leading to an upward bias in b_1. This in turn leads to a downward bias in b_8 in equation (22). As a result we underestimate the scale parameter $\gamma = b_8$. We estimate $\hat{\alpha}_M = \frac{\gamma}{b_1}$. Hence we get a downward bias because both the denominator is too high and the numerator too low. This in turn leads to upward bias in our estimates of μ which is the inverse material share divided by α_M. We estimate $\hat{\lambda} = r\hat{\mu} - q$. As we have upward bias in μ we consequently get downward bias in λ, as we find in Figure F-2. We compute $\hat{\alpha_L} = b_2 \times \alpha_M$. Like for b_1 we have upward bias in b_2 because l will contain short term variation (against model assumptions). This will be compensated

Figure F-2: Monte Carlo Results with Bias



Notes: The figure reports results from a Monte-Carlo analysis of the FMMM estimation framework. We draw 1000 replications of samples of 500 firms over 10 time periods with the following parameter values: $\alpha_L = 0.3n$, $\alpha_M = 0.6$,$\gamma = 1.2$, $\sigma_{\nu_a} = 0.25$, $\sigma_{\nu_\lambda} = 0.25$, $\sigma_{\nu_\mu} = 0.25$, $\sigma_{\alpha_\mu} = 1$, σ_{ν_K}=1, σ_{ν_L}=1, $\phi_\lambda = 0.5$, $\phi_a = 0.4$. We also allow for a bias parameter of $b = 0.4$ Solid vertical lines indicator the true value, dashed vertical lines indicate the average estimated values.

somewhat by the downward bias in α_M leading to only a relatively small upward bias in for α_L in F-2. Finally, we compute $\hat{a}$ as $\hat{a} = q - \alpha_L l - \alpha_M m - (\gamma - \alpha_M - \alpha_L)k$. We over estimate α_L. However, we overestimate α_M and γ. Thus the bias could go either way but with the particular parameter combination in Figure F-2 we end up with an upward bias.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1401	Holger Breinlich	The Effect of Trade Liberalization on Firm-Level Profits: An Event-Study Approach
1400	Richard V. Burkhauser Jan-Emmanuel De Neve Nattavudh Powdthavee	Top Incomes and Human Well-Being Around the World
1399	Rabah Arezki Thiemo Fetzer	On the Comparative Advantage of U.S. Manufacturing: Evidence from the Shale Gas Revolution
1398	Adriana Kocornik-Mina Thomas K.J. McDermott Guy Michaels Ferdinand Rauch	Flooded Cities
1397	Lorenzo Caliendo Giordano Mion Luca David Opromolla Esteban Rossi-Hansberg	Productivity and Organization in Portuguese Firms
1396	Richard Murphy Gill Wyness	Testing Means-Tested Aid
1395	Zack Cooper Stuart Craig Martin Gaynor John Van Reenen	The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured
1394	Hannes Schwandt Amelie Wuppermann	The Youngest Get the Pill: Misdiagnosis and the Production of Education in Germany
1393	Yatang Lin Yu Qin Zhuan Xie	International Technology Transfer and Domestic Innovation: Evidence from the High-Speed Rail Sector in China
1392	Robin Naylor Jeremy Smith Shqiponja Telhaj	Graduate Returns, Degree Class Premia and Higher Education Expansion in the UK

1391	Marc Fleurbaey Hannes Schwandt	Do People Seek To Maximize Their Subjective Well-Being?
1390	Benjamin R. Handel Jonathan T. Kolstad Johannes Spinnewijn	Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets
1389	Kilian Huber	The Persistence of a Banking Crisis
1388	Francesco Caselli Miklos Koren Milan Lisicky Silvana Tenreyro	Diversification through Trade
1387	Hannes Schwandt Janet Currie	Short and Long-Term Effects of Unemployment on Fertility
1386	Marta De Philippis	Multitask Agents and Incentives: The Case of Teaching and Research for University Professors
1385	Ferdinando Monte Stephen J. Redding Esteban Rossi-Hansberg	Commuting, Migration and Local Employment Elasticities
1384	Tito Boeri Juan Francisco Jimeno	The Unbearable Divergence of Unemployment in Europe
1383	Sarah Flèche	Distaste for Centralization: Evidence from a Quasi-Natural Experiment in Switzerland
1382	Johannes Boehm	The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity
1381	Andrew B. Bernard Swati Dhingra	Contracting and the Division of the Gains from Trade

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7404 0612
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk