Cost and Product Advantages: A Firm-level Model for the Chinese Exports and Industry Growth*

Jordi Jaumandreu† Heng Yin‡

This version: October 31, 2016
First draft: January 2013

Abstract

We use data from 70,000 Chinese manufacturing firms, which are both domestic sellers and exporters, to estimate the joint distribution of unobserved productivity (cost advantages) and unobserved demand heterogeneity (product advantages) from 1998 to 2008. Product advantages are negatively correlated with cost advantages (positively correlated with marginal cost). We characterize growth and sketch examples to show that splitting the advantages produces useful analytical insights. The state is not good at developing product advantages. A fraction of firms specialize in low-cost-low-quality exports. Many marginal cost differences across firms come from heterogeneous output-embodied levels of quality and technology, not "price distortions."

Keywords: productivity, cost advantages, demand heterogeneity, product advantages.

JEL: C33,D24,L11,O30,O47.

*We thank Alexander Callis, Shigang Li, Di Liu, Urbashee Paul, Xiaoyu Tang and Xiaoxi Zhao for help; U. Doraszelski and Shuheng Lin for many useful discussions; V. Aguirregabiria, M. Arellano, J. Van Biesebroeck, J.C. Farinas, I. Fernandez-Val, M. Manove, M. Melese, Shengyu Li, B. Peters, M. Roberts, M. Rysman, A.Pakes and numerous seminar audiences for comments. Jaumandreu acknowledges partial funding from Fundacion Rafael del Pino.

†Boston University, Dep. of Econ., 270 Bay State Road, Boston, MA02215, USA. E-mail: jordij@bu.edu.

‡Renmin University of China, National Academy of Development and Strategy, Haidian District, Beijing 100872, PRChina. E-mail: yheng@ruc.edu.cn.

1. Introduction

Imagine we observe that two firms sell two substitute products at the same location, with the same price, and incur the same promotional expenses. However, the first firm sells X% more units than the second. The first product may meet the tastes of a greater proportion of consumers with its combination of horizontal and/or vertical (quality, technology or design) characteristics. The product may have been around for longer, and over time it has entered the awareness or trust set of more consumers.[1] The product may have a better distribution network or is sold under long-term contracts to large buyers. Whatever the reason, we say that the first firm has an unobserved product advantage of magnitude X% over the second. Unobserved advantages in production are traditionally called productivity, unobserved product advantages are customarily refered to as demand heterogeneity. Formally, we call unobserved productivity or TFP the proportion by which one firm produces a larger (smaller) quantity of output with the same inputs than a hypothetical average firm in the same industry would. Similarly, we define an unobserved product advantage as the proportion by which the firm sells a larger (smaller) quantity of output at the same price than a hypothetical average firm in the industry would, once the observed explanatory factors have been controlled for (in this paper: location, age, state support, and sales effort).

Productivity generates, by duality, unobserved cost advantages. In imperfectly competitive markets, the profits and growth of firms are as crucially dependent on unobserved product advantages as they are on cost advantages.[2] However, while there is a huge literature analyzing productivity distributions (see Bartelsman and Doms, 2000, and Syverson, 2011, for surveys), demand heterogeneity distributions have only been examined recently (see below).[3] One leading reason is that

[1] This effect can be partially observed through the impact of variables measuring age and market experience of the firm.

[2] Product advantages are more "rival" than cost advantages. The increase of the level of productivity of one firm does not affect the level of productivity reached by another firm (although it may end in stealing of demand through price competition). However, the development of a product advantage by a firm can be predatory to the advantage of another firm (although this is not necessarily the case).

[3] Demand estimation since Berry (1994) and Berry, Levinsohn and Pakes (1995) has richly used the discrete-choice framework to explain product shares in specific markets, but the focus is consumer tastes. Market shares are a function of the observable product characteristics, price, and an unobserved linear utility effects of omitted characteristics usually denoted as ξ_j for product j. Some authors model ξ_j as an AR(1) process (see Lee, 2013, and Sweeting, 2013). Our product advantages are basically a combination of the ξ_j term and the nonlinearities of the

unobserved product advantages have been taken as inseparable from productivity without firm-level information on output prices. When the data at hand contains no firm-specific output price to deflate revenue, demand heterogeneity is unavoidably brought into the productivity relationship that has to be estimated. Klette and Griliches (1996) started this analysis, and De Loecker (2011) blends this idea with an Olley and Pakes (1996) procedure of estimation. Both papers treat residual demand heterogeneity as iid, but recent papers admit more persistent demand shocks and give up in separating the two unobservables. Some examples are Hsieh and Klenow (2009), Gandhi, Rivers and Navarro (2013), Asker, Collard-Wesler and De Loecker (2014), Boler, Moxnes and Ullveit-Moe (2015), Peters, Roberts, Van Ahn and Fryges (2016) and Bilir and Morales (2016). Many assume that there is an unobservable composite of productivity and demand heterogeneity that follows a Markov process.

Using a sample of roughly 70,000 Chinese manufacturing firms, that are both domestic sellers and exporters, we estimate the joint distribution of unobserved cost advantages and unobserved product advantages, and how it changed from 1998 to 2008. Then we use the distribution to characterize the growth of Chinese manufacturing and examine its weaknesses. Additionally, we sketch a few examples to show that the split of advantages into cost and product generates useful analytical insights. Our data is particularly suitable for this exercise. In the eleven years our data covers the average domestic output of the firms in our sample increases by a factor of 2.3 and exports by 3.3. As Foster, Haltiwanger and Syverson (2008) stress, firms grow by increasing productivity and setting lower prices (a movement along their demand curves) or by expanding the sales of their products by means of observed and unobserved demad-expanding actions and investment (a shift in their demand curves). The advantage of analyzing a sample from a period of extremely fast growth is that firms are both rapidly improving productivity and trying to build product advantages.[4]

expression for s_j. However, between the usual industry-specific BLP exercise and the exercise here there are two important additional differences. First, observed product characteristics typically reach an important level of detail that is not available for an interindustry study. Second, with interindustry data, the usual firm-level observation of a multi-product firm refers to the composite of product-specific demands (that are likely to belong to different markets).

[4] Here are two specific examples. *Konka* is a big TV and electronic producer that, in 2008, was getting 20% of its sales from exports. From 2001 to 2008, its total sales increased by a factor of 1.8 with exactly the same employment (17,000), while the price of a 21 inch TV fell from 3000 RMB to 1000 RMB. *Tingyi* is a producer of instant noodles and beverages whose "Master Kong" brand had about 38% of the domestic market in 2008. Sales from 1998 to 2008 incresed by a fator of 7.3, while employment more than doubled (up to 50,000). However, the price of standard noodles increased from 1.5 RMB in 1998 to 2.8 RMB in 2008. Sales soared as the firm triplicated its centers of distribution and switched to a tighter relationship with retailers instead of relying on wholesalers.

Separating the unobservable advantages into cost and product has several benefits. First, it allows the analyst to attribute the observed facts to productivity or demand factors (for example, she can weigh up the extent to which firms self-select into the export market because of higher productivity or because of superior products). Second, it permits separate assessment of dispersion, persistence and trends. Third, both advantages should be considered endogenous, in the sense of being impacted by the investments of firms (in knowledge, human capital, organization). Investments will typically have different productivity and feasibility limits, and will confront a likely trade-off among advantages. Therefore, only through separation we can diagnose market-specific situations and devise suitable strategies for sustainable growth (for example, should firms focus on competing either in prices or in the development of higher quality products? with which tools?). Finally, the analysis of the allocation of resources faces questions that cannot be answered without the separation of advantages (for example, how much should a particular product be manufactured if it increases sales and profits while at the same time it diminishes productivity and increases marginal cost?).

To separate cost and product advantages, we start by specifying two demands for the product of the firm (exports and domestic, as in Das, Roberts and Tybout, 2007, and Aw, Roberts and Xu, 2011) which depend on the firm prices for each market, the observable shifters, and a persistent time evolving unobservable reflecting the product advantages of the firm. As we do not observe output prices, we transform the demands into revenue equations and replace the explanatory price by its optimal level in terms of the firm specific marginal cost. This is similar to De Loecker (2011), who writes the inverse demand and replaces output by the production function. Marginal cost has an observable part but depends on unobservable productivity too. The transformation gives us two equations in which the sales of the firm depend, in addition to the observed cost and demand shifters, on productivity and product advantages. We specify the unobservables as Markov processes and estimate them using an Olley and Pakes (1996) approach: inverting the system to get the unobservables in terms of observables. Intuitively, the system can be inverted because the elasticity of demand in the export market is higher and cost advantages have a greater impact on the export market.[5] We estimate the system by nonlinear GMM semiparametric methods drawing on Doraszelski and Jaumandreu (2013). We estimate the elasticity of demands, the parameters of the cost (production) function, the impact of observed shifters and the value of the unobservables.

[5] We show that a sufficient condition to nonparametrically recover the unobservables is that the demand elasticities of the two markets are different. We recover the unobservables by inverting our parametric system of equations, but we are sure that we are picking up something more generally identified.

We are not the first researchers to deal with the separation of the unobserved advantages. Foster, Haltiwanger and Syverson (2008) disentangle productivity and demand heterogeneity using a sample of US quasi-homogeneous good industries for which they can use unit values as prices. This was pioneering in assessing demand heterogeneity. They estimated static residual demand effects. Foster, Haltiwanger and Syverson (2016) use the same type of data to estimate a dynamic model of demand accumulation that underlines the reality of slowly building market shares. Using data from more than 7,000 Italian firms in three industries (textiles, metals and machinery), Pozzi and Schivardi (2016) build an analysis in terms of time differences. They know price changes and have a subjective assessment of demand elasticity from managers. They compute TFP growth and demand shocks, then explore their role in the growth of firms. Roberts, Xu, Fan and Zhang (2016) observe product exports across world destinations for a sample of 738 Chinese footwear producers and take unit values as prices. They assess the relative importance of a firm idiosyncratic demand effect and firm specific marginal cost, considering in addition fixed cost effects.

Some other papers have computed TFP and product advantages to measure their relative role in an empirical relationship of interest. Eslava, Haltiwanger, Kugler and Kugler (2004) check the impact on reallocation of output among Colombian firms, Aw and Lee (2014) on foreign investments of Taiwanese firms, and Gervais (2015) on the export decisions and export intensity of US firms. These papers share a static residual demand approach to the measurement of product advantages, and use TFP or TFP of the rest of the firms to instrument price or output (the last paper uses labor productivity). Hottman, Redding and Weinstein (2014) is a tightly parametrized CES model for supermarket goods, allowing for the measurement of productivity and demand heterogeneity.[6]

In comparison with these works, our paper has three specific differences. First, it utilizes a larger sample that includes all kinds of differentiated products. We use firms from China's entire manufacturing spectrum (split into ten broad industries). Second, we disentangle cost and product advantages without observing output prices. Third, our focus is on robustness: we estimate the unobservables simultaneously, as non-funcionally-dependent and freely correlated Markov processes, addressing endogeneity by means of an Olley and Pakes (1996)/ Levinsohn and Petrin (2003) method of estimation.[7]

Our results abound in stylized facts and new insights on traditional estimates. Firm demands are

[6] Jaumandreu and Mairesse (2010, 2016) explore exogenous and endogenous determinants of the shifts of product advantages and productivity.

[7] We model them as *exogenous* Markov processes but we think that the model should be generalized. See Section 8, Concluding remarks.

estimated to be very elastic, especially in the exports market. Therefore, export markets emerge as more competitive than domestic markets, and firms set lower prices and get smaller margins. Productivity has a big dispersion (comparable to other estimates), and mean productivity experiences a huge change during this period specific to China. Product advantages, compared in a proper scale, are even more dispersed than productivity, but they change quite slowly and very heterogeneously across products. Although with Chinese specific traits, our results here match the findings of Foster, Haltiwanger and Syverson (2016).

A crucial novel result is that product advantages turn out to be negatively correlated with cost advantages (positively correlated with marginal cost). Foster, Haltiwanger and Syverson (2008, 2016) do not share this finding because their identifying assumption is the absence of correlation between TFP and demand heterogeneity. Pozzi and Schivardi (2016) cannot assess this correlation because they look at the change in the unobservables over time. However, our results perfectly match the findings of Roberts, Xu, Fan and Zhang (2016) that "demand differences are costly to produce." They find a positive correlation between the firm demand fixed effect across makets and the firm marginal cost. We elaborate later on the implications of our result.

Chinese manufacturing experienced a big change in the allocation of production. Only a small proportion of starting firms survived and an overwhelming majority of production is ultimately controlled by firms born during the period, smaller in employment or capital. This allows us to check whether product advantages show important roles in selection into the market and survival (as in Foster, Haltiwanger and Syverson, 2008 and 2016, and Roberts, Xu, Fan and Zhang, 2016). The split of the advantages adds here insights to the results of Brandt, Van Biesebroeck and Zhang (2012). Entrants contribute product advantages, but these entrants tend to be higher-cost producers.[8] Reallocation among survivors is also important: the most productive firms tend to become bigger, but product advantages become more concentrated in the smaller firms.

We further work with three examples to show that the split of advantages in cost and demand can produce useful analytical insights. We find that product advantages are primarily experienced by private firms, particularly the newly born ones. Firms in ownership transition were more sluggish in the development of product advantages but realized faster growth with regard to productivity. We uncover that some firms choose to completely orient their activity towards the export market (specialization) based on cost advantages combined with very low product advantages. We check that both R&D activities and the quality of the workforce increase a firm's product advantages

[8]See page 35 for the definition of survivors, entrants and exitors in the context of our sample.

with the tradeoff of having higher production costs. This challenges the view that marginal cost differences are mainly "price distortions."

What are the implications of the negative correlation finding? Profits are checked to be positively correlated with both productivity and product advantages. However, product advantages and productivity are negatively correlated. This means that many firms that show important product advantages exhibit comparatively modest productivity. Conversely, many firms that reveal strong cost advantages have no product advantages. In fact, product advantages turn out to have a strong positive correlation with total marginal cost. Everything suggests that developing both productivity and product advantages is costly and that there is a trade-off between them. Combining quality, design or technology with low costs, has technological and firm knowledge/ability limits which impact the growth paths. The picture that we obtain for the whole period is that Chinese firms relied heavily on cost competition to grow, and relied much more modestly on product advantages (although these product advantages are sharply developed in electronics and machinery). This is a weakness that can hurt Chinese exports, particularly as other developing countries engage more intensely in the race (see Sutton, 2001 and 2007, for insights on a development model based on the mix of cost and product advantages). This matches the diagnosis of the policy-makers who designed "Made in China 2025."[9]

Our results also suggest some methodological issues. The magnitude and persistence of demand heterogeneity, and its correlation with productivity, have implications. First, IV is not a suitable technique to estimate firm-level demand relationships of firms when product advantages are uncontrolled and, in particular, TFP and input prices are not legitimate instruments. The use of this technique is likely to induce a downward bias in the estimated elasticity of demand. However, a specification of product advantages in a way analogous to the Markov specification of productivity may allow the use of moments based on variables uncorrelated with the unpredictable part of demand shocks (e.g., lagged prices of the inputs). Second, the consistency of OP/LP procedures of separated estimation of production functions in imperfectly competitive markets requires very strong assumptions. Generally, the input demand that is inverted depends on unobserved demand heterogeneity via marginal revenue, even if the researcher has output prices.[10] There is still no

[9] "Made in China 2025" is an ambitious plan for manufacturing to become more innovation-driven, higher quality, greener and based on greater human capital.

[10] Call K a fixed input, X a vector of variable inputs with prices W_X, and P the ouput price. Solving the system of FOCs, the unconditional demand for X is $X = X(K, \frac{W_X}{MR}, \omega)$, where MR is marginal revenue and $MR = MC$. Notice that equal input and output prices across firms is, in addition to being unrealistic, not necessary nor sufficient

method to control for the unobserved variability that this introduces.[11] This makes progress in modeling demand heterogeneity an interesting avenue of research for improving the estimation of production functions.

The rest of the paper is organized as follows. In Section 2, we show that the unobservable cost and product advantages are characteristics nonparametrically identified in the absence of prices. In principle, no particular functional form is needed for their estimation. In Section 3, we set out our particular empirical parametric specification. Section 4 explains how we estimate the econometric model. Section 5 introduces the data and describes the sample that we use. Section 6 reports the results of estimation, describes the joint distribution and the correlations of the estimated cost and product advantages, and performs some descriptive exercises. Section 7 develops three examples in which separating cost and product advantages is useful for analysis. Section 8 concludes. There are five appendices and an Online Appendix.

2. Model and identification

In this section, we present the model and show that the main characteristics of interest, the demand and cost advantages of the firms, are nonparametrically identified from revenue, input prices, input quantities and demand shifters.

2.1 Revenue as a function of cost and product advantages.

Firm j produces a product that sells in two or more monopolistically competitive markets.[12] Let us consider market I of firm j. The demand for the product at moment t is

$$Q_{jt}^{I} = Q^{I}(P_{jt}^{I}, Z_{jt}^{I}, \delta_{jt}), \tag{1}$$

to write $X = X_t(K, \omega)$ as many researchers do. In general, $MR = MR(P, Z, \delta)$ where Z and δ represent observed and unobserved demand heterogeneity respectively (see later in this paper). Additionally, $MC = MC(K, W_X, Q, \omega)$, where substituting for $Q = Q(P, Z, \delta)$ reintroduces Z and δ. Conditional demands depend, in turn, on Q. Therefore, only quite restrictive assumptions can control for demand heterogeneity in the relationship that has to be inverted.

[11] Doraszelski and Jaumandreu (2013) model the elasticity of demand as a nonparametric function of the price and a demand intercept; De Loecker, Goldberg, Khandelval and Pavnik (2016) include many observable variables in the inverted demand (page 466). Ackerberg, Caves and Frazier (2015) mention the problem.

[12] From monopolistic competition, we use the properties that each firm faces a downward-sloping demand for its product and that a price change by one firm has a negligible effect on the demand of any other firm (Tirole, 1989).

where P_{jt}^I is the price set by the firm, Z_{jt}^I is a vector of observed market and firm specific demand shifters, and δ_{jt} is a scalar unobservable that measures unspecified advantages linked to the firm's product. We assume that $Q^I(\cdot)$ is monotonic in δ_{jt} and that the impact of δ_{jt} is positive without a loss of generality. Some demand shifters may be set by the firm (e.g. the level of sales effort).

The firm has production function

$$Q_{jt} = F(K_{jt}, L_{jt}, M_{jt}, \omega_{jt}),$$

where $Q_{jt} = \sum_I Q_{jt}^I$ is total firm output, variables K_{jt}, L_{jt} and M_{jt} stand for capital, labor and materials respectively, and ω_{jt} is a scalar unobservable that measures unspecified advantages with a positive impact on the production level of the firm. We assume that $F(\cdot)$ is monotonic in ω_{jt}. The term ω_{jt} is usually called productivity.[13] Let us write the dual marginal cost as $MC_{jt} = MC(X_{jt}, \omega_{jt})$ where X_{jt} is a vector of observable prices and quantities of the inputs.[14]

Multiplying both sides of equation (1) by P_{jt}^I we get the revenue expression

$$R_{jt}^I = P_{jt}^I Q^I(P_{jt}^I, Z_{jt}^I, \delta_{jt}) \tag{2}$$

and, inverting the profit maximization condition $MR(P_{jt}^I, Z_{jt}^I, \delta_{jt}) = MC(X_{jt}, \omega_{jt})$,[15] we can write

$$P_{jt}^I = MR^{-1}(MC(X_{jt}, \omega_{jt}), Z_{jt}^I, \delta_{jt}). \tag{3}$$

Combining equations (2) and (3) we finally have

$$R_{jt}^I = R^I(MC(X_{jt}, \omega_{jt}), Z_{jt}^I, \delta_{jt}). \tag{4}$$

This equation[16] is useful when prices are not observed and we cannot work with equation (1). Equation (4) says that revenue is a function of both the observable factors which determine marginal

[13] Productivity is almost universally specified as Hicks neutral. Therefore, the production function is written as $Q_{jt} = F(K_{jt}, L_{jt}, M_{jt}) \exp(\omega_{jt})$. We keep, for the moment, a more general specification that is symmetric with the specification of the demand advantages δ_{jt}.

[14] Consider the following example. Given K_{jt}, and calling wage W_{jt} and the price of materials P_{Mjt}, the variable cost function is $C_{jt} = C(K_{jt}, W_{jt}, P_{Mjt}, Q_{jt}, \omega_{jt})$ and $MC_{jt} = \frac{\partial C}{\partial Q_{jt}}(\cdot)$. The conditional demand for materials is $M_{jt} = \frac{\partial C}{\partial P_M}(K_{jt}, W_{jt}, P_{Mjt}, Q_{jt}, \omega_{jt})$. Solving this demand for output, and replacing output in the marginal cost function, one gets the expression of the text with $X_{jt} = \{K_{jt}, M_{jt}, W_{jt}, P_{Mjt}\}$.

[15] We assume that MR is monotonic in price. A suffient condition is that the absolute value of elasticity is not decreasing in price.

[16] It is easy to show that $\frac{\partial R_{jt}^I}{\partial MC} < 0$. If the demand elasticity is non-increasing for the demand shifters and δ_{jt}, we have $\frac{\partial R_{jt}^I}{\partial Z_{jt}} > 0$ and $\frac{\partial R_{jt}^I}{\partial \delta_{jt}} > 0$.

cost and the demand shifters, and of the two unobservables representing the demand and cost advantages of the firm. Even if we were able to perfectly measure all the observable variables, we cannot separately recover ω_{jt} and δ_{jt} from equation (4). Recovering a combination might be interesting on its own, but our main objective is to show how ω_{jt} and δ_{jt} can be separately nonparametrically identified.

2.2 Recovering ω_{jt} and δ_{jt}.

What we need is to observe the firm selling the product in (at least) two markets. Suppose, for example, the firm sells the product in the exports (X) and domestic (D) market. We have two revenue functions

$$
\begin{aligned}
R_{jt}^X &= R^X(MC(X_{jt},\omega_{jt}), Z_{jt}^X, \delta_{jt}),\\
R_{jt}^D &= R^D(MC(X_{jt},\omega_{jt}), Z_{jt}^D, \delta_{jt}).
\end{aligned}
\tag{5}
$$

If this system can be solved, we can get ω_{jt} and δ_{jt} expressed in terms of observables

$$
\begin{aligned}
\omega_{jt} &= \omega(X_{jt}, Z_{jt}^X, Z_{jt}^D, R_{jt}^X, R_{jt}^D),\\
\delta_{jt} &= \delta(X_{jt}, Z_{jt}^X, Z_{jt}^D, R_{jt}^X, R_{jt}^D).
\end{aligned}
\tag{6}
$$

This inversion allows us to set an estimable model controlling for persistent unobservables in terms of observables and gives us a way to back out the advantages from revenue, input prices, input quantities and shifters.

Let us discuss when the system be inverted. Call λ_{jt} the ratio of semielasticities of revenue with respect to the product advantages, i.e. $\lambda_{jt} = \frac{1}{R_{jt}^D}\frac{\partial R^D}{\partial \delta_{jt}} / \frac{1}{R_{jt}^X}\frac{\partial R^X}{\partial \delta_{jt}}$. Let η_{Xjt} and η_{Djt} be the absolute value of the elasciticity of demand in the export and domestic market. Then we can establish

Proposition. If the ratio of elasticities $(\eta_{Xjt}-1)/(\eta_{Djt}-1)$ is different from λ_{jt} system (5) can be inverted.

Proof: See Appendix A.

The intuitive reason by which ω_{jt} and δ_{jt} can be identified is that their effects are different in each market. Cost advantages operate through the price set in each market. As long as the price effects are different, the variation in revenues identifies the advantages.[17] One particular case happens when product advantages have the same impact in each market, $\lambda_{jt} = 1$. In this case, it is sufficient for identification that the demand elasticities are different in the two markets.

[17] Except when the ratio of these effects exactly matches the relative effects of the product advantages.

2.3 An estimable model.

Cost and product advantages are likely to be both persistent over time and subject to unexpected shocks. We use the modeling for unobserved productivity in production functions introduced by Olley and Pakes (1996). We assume that the cost and product advantages follow the first order Markov processes

$$\begin{aligned} \omega_{jt} &= q(\omega_{jt-1}) + \xi_{jt} \\ \delta_{jt} &= s(\delta_{jt-1}) + \varepsilon_{jt} \end{aligned} \tag{7}$$

where $q(\cdot)$ and $s(\cdot)$ are unknown functions. Advantages at moment t are decomposed into the level predictable from its value at moment $t-1$ and the mean independent shocks ξ_{jt} and ε_{jt}. Unobservables ω_{jt-1} and δ_{jt-1} can be recovered using (6) lagged and plugged into (7). Then (7) can be inserted into (5), so that we have the nonparametric structural econometric model

$$\begin{aligned} R^X_{jt} &= R^X(MC(X_{jt}, g(S_{jt-1}) + \xi_{jt}), Z^X_{jt}, h(S_{jt-1}) + \varepsilon_{jt}) \\ R^D_{jt} &= R^D(MC(X_{jt}, g(S_{jt-1}) + \xi_{jt}), Z^D_{jt}, h(S_{jt-1}) + \varepsilon_{jt}), \end{aligned} \tag{8}$$

where $g(\cdot) = q(\omega(\cdot))$, $h(\cdot) = s(\delta(\cdot))$ and $S_{jt-1} = \{X_{jt-1}, Z^X_{jt-1}, Z^D_{jt-1}, R^X_{jt-1}, R^D_{jt-1}\}$.

Equations (8) form a system which contains a few variables that maybe correlated with ξ_{jt} and ε_{jt}. Other variables are assumed independent, and both disturbances are present in both equations. Matzkin (2007, 2013) discusses nonparametric identification of systems of this type. In what follows, we specify and estimate a parametric version of the model. However, the advantages that we want to characterize are identified under much more general specifications.

3. An empirical specification to estimate cost and product advantages

3.1 Demand.

Firms produce a single product, in practice a set of products that we treat as one, that sell in the domestic (D) and export (X) markets. Both markets are monopolistically competitive. The demands for the product of firm j are

$$Q^X_{jt} = \alpha^X_0 \left(\frac{P^X_{jt}}{P^X_t}\right)^{-\eta_X} \exp(z^X_{jt}\alpha_X + \delta_{jt}),$$

$$Q^D_{jt} = \alpha^D_0 \left(\frac{P^D_{jt}}{P^D_t}\right)^{-\eta_D} \exp(z^D_{jt}\alpha_D + \delta_{jt}). \tag{9}$$

The terms α^X_0 and α^D_0 are constants, η_X and η_D are common industry elasticities, and P^X_t, and P^D_t industry price indices.[18]

The firm's demand is shifted by two components in each market. The first component is the impact of a vector of observables z^I_{jt}.[19] The second component is the idiosyncratic unobservable δ_{jt} representing the unexplained level of advantages of the product.[20] We model δ_{jt} as firm specific, persistent over time and embodying unexpected shocks (see below). Two firms with a similar products, prices, and the same number of years in the market (and/or other relevant similar observable advantages), can still show a different level of market penetration given by the level of their unobserved product advantages. By its definition, δ_{jt} also includes demand improvements (deteriorations) common to all firms in the market. For example, a pull of industry exports affecting all firms or a decrease in the level of demand available to each firm due to the entry of new firms in the market.

A restriction of our empirical modeling is the assumption that the impacts of the unobserved advantages δ_{jt} are the same in both markets (unit semielasticities).[21] This seems natural for many advantages, but not for others. This limitation stems from the lack of firm-level prices. We need two equations to disentangle ω_{jt} from δ_{jt}. The estimation of a different δ_{jt} in each market would require a third equation. In the empirical part we check for the robustness of our assumption by allowing the impact of product advantages to differ across markets. The same product characteristics are

[18] We further discuss this specification in section A3 of the Online Appendix.

[19] Some shifters may be endogenously determined by the firm in the short run. This is likely to happen with sales effort. Let z^I_{jt} represent the log of expenditures on advertising and promotion in market I and suppose that the firm optimally sets P^I_{jt} and $\exp(z^I_{jt})$. The Dorfman and Steiner (1954) condition for optimal determination of $\exp(z^I_{jt})$ gives $\frac{\exp(z^I_{jt})}{R_{jt}} = \frac{\alpha_I}{\eta_I}$, which can be also written as $z^I_{jt} = \frac{1}{1-\alpha_I}(\ln \alpha^I_0 + \ln \frac{\alpha_I}{\eta_I} - (\eta_I - 1)p^I_{jt} + \delta_{jt})$, where p^I_{jt} stands for the log of price. Note that if we had prices, this latest equation could be exploited in an Olley and Pakes (1996) type of procedure to estimate demand advantages.

[20] The terms $z^X_{jt}\alpha_X + \delta_{jt}$ and $z^D_{jt}\alpha_D + \delta_{jt}$ tell us the additional quantity of the product of firm j that is bought by consumers when its price is equal to the price of a rival for whom these demand terms are equal to zero. We could also write $P^D_{jt} = P^D_t \left(Q^D_{jt}/\alpha^D_0\right)^{-\frac{1}{\eta_D}} \exp((z^D_{jt}\alpha_D + \delta_{jt})/\eta_D)$. The same terms scaled by the corresponding η can be read as describing how much more the consumers are willing to pay for the same quantity of the good with respect to the price of a product with zero advantages.

[21] Roberts, Xu, Fan and Zhang (2016) specify a common firm effect across destination markets that turns out to be the dominant effect of their model.

supposed to have an impact $\lambda\delta_{jt}$ in the domestic market and δ_{jt} on exports.

3.2 Production and cost.

Firm j produces its product (set of products) with Cobb-Douglas production function

$$Q_{jt} = \exp(\beta_0) K_{jt}^{\beta_K} L_{jt}^{\beta_L} M_{jt}^{\beta_M} \exp(\omega_{jt}), \tag{10}$$

where ω_{jt} represents Hicks neutral productivity.[22] We assume that K_{jt} is given and that the firm freely chooses L_{jt} and M_{jt} in the short-run. We denote the short-run elasticity of scale by $\nu = \beta_L + \beta_M$. We call the corresponding variable cost C_{jt} and marginal cost MC_{jt}. A consequence of Hicks neutrality is that MC_{jt} can be separated into observed variables and unobserved ω_{jt}, so we write $MC_{jt} = \overline{MC}_{jt} \exp(-\omega_{jt})$.

The marginal cost of domestic and export sales is the same. However, this assumption may be restrictive: firms produce multiple products and marginal costs may differ across products. For example, firms may choose to export a product (or range of products) different from the product (range of products) that they sell domestically implying different marginal costs. Even if they sell the same products, the composition of sales may lead to a different cost. Theoretical trade literature has just started to deal with these possibilities,[23] but there is still no empirical evidence.[24] It is easy to generalize our model to the presence of different marginal costs due to the capital used, input prices or both (see below). However, we have nothing in our data that indicates varying product choices or allows to test for them. We leave this extension to future research.

3.3 Firm equilibrium.

According to demands (9) and the cost implied by (10), the firm sets the prices and quantities $P_{jt}^X, Q_{jt}^X, P_{jt}^D$ and Q_{jt}^D to maximize short-run profits. To do so the firm takes into account K_{jt}, the current values of the shifters, and the values of δ_{jt} and ω_{jt} (unobservable for the econometrician but

[22] Notice that our unobserved demand advantages δ_{jt} are also "neutral" with respect to other shifters.

[23] In Mayer, Melitz and Ottaviano (2014, 2016), firms export only a subset of their product range, the products that are expected to perform best according to the increased toughness of competition in foreign markets.

[24] For example, Manova and Zhang (2012) show with detail the multiproduct character of the exports of Chinese firms and the wide price discrimination practiced across destinations. However, the character of the data (customs data) impedes the comparison with the domestic sales.

observable for the firm). The first order conditions can be written as

$$\begin{aligned} P_{jt}^X(1-\frac{1}{\eta_X}) &= MC_{jt}, \\ P_{jt}^D(1-\frac{1}{\eta_D}) &= MC_{jt}. \end{aligned} \tag{11}$$

Simultaneous to deciding the price and output choices, the firm determines the variable input quantities M_{jt} and L_{jt} according to the cost minimizing conditions:

$$\begin{aligned} MC_{jt}\beta_M \exp(\beta_0)K_{jt}^{\beta_K}L_{jt}^{\beta_L}M_{jt}^{\beta_M-1}\exp(\omega_{jt}) &= P_{Mt}, \\ MC_{jt}\beta_L \exp(\beta_0)K_{jt}^{\beta_K}L_{jt}^{\beta_L-1}M_{jt}^{\beta_M}\exp(\omega_{jt}) &= W_{jt}(1+\Delta_{jt}), \end{aligned} \tag{12}$$

where Δ_{jt} is a shock to the price of labor reflecting the impact of adjustment costs in the short-run equilibrium.[25]

Importantly, equations (9), (10), (11) and (12) together imply that variable inputs are correlated with the unobservables δ_{jt} and ω_{jt}. Since both unobservables are persistent, capital K_{jt} is correlated too (because past investment choices of the firm are correlated with past values of the unobservables). The firm-level wage W_{jt} is likely to reflect the productivity level of the firm and possibly the product advantages, so it is likely to be correlated as well. In estimations, it is very important that we control for the predictable part of the unobservables δ_{jt} and ω_{jt}. This will limit endogeneity to the variables that are chosen after the realization of the unpredictable part of δ_{jt} and ω_{jt} (we discuss which ones in subsection 4.3).

3.4 Estimating equations.

Multiplying conditions (12) by M_{jt} and L_{jt} respectively and adding them we get $\nu MC_{jt}Q_{jt} = C_{jt}(1+\frac{W_{jt}L_{jt}}{W_{jt}L_{jt}+P_{Mt}M_{jt}}\Delta_{jt})$. Using conditions (11) to replace MC_{jt} in $MC_{jt}Q_{jt} = MC_{jt}Q_{jt}^D + MC_{jt}Q_{jt}^X$, dividing everything by total revenue $P_{jt}^X Q_{jt}^X + P_{jt}^D Q_{jt}^D = R_{jt}$, inverting the ratio and taking logs (that we represent henceforth by lowercase letters) we arrive at the equation

$$\ln\frac{R_{jt}}{C_{jt}} \equiv r_{jt}-c_{jt} = \ln\frac{1}{\nu}\frac{\eta_D}{\eta_D-1} - \ln\left[1+\left(\frac{\frac{\eta_D}{\eta_D-1}}{\frac{\eta_X}{\eta_X-1}}-1\right)S_{jt}^X\right]+e_{jt}, \tag{13}$$

where S_{jt}^X represents the share of revenue from exports in total revenue and the disturbance stands for the shock $e_{jt} = -\ln(1+\frac{W_{jt}L_{jt}}{W_{jt}L_{jt}+P_{Mt}M_{jt}}\Delta_{jt})$.

[25] For the dynamic framework that justifies this shadow price of labor see Doraszelski and Jaumandreu (2013, 2016).

This equation describes the log of revenue over variable cost, or price-average cost margin of the firm (PACM),[26] as the result of the domestic markup multplied by $\frac{1}{\nu}$ and the effect of the possible difference of markups between the foreign and domestic markets. It generalizes Das, Roberts and Tybout (2007) and Aw, Roberts and Xu (2011). Under the assumption $E(S_{jt}^X e_{jt}) = 0$, it identifies the elasticities of demand up to parameter ν.[27] This equation can be easily extended to relax the assumption of common elasticities in the whole industry by estimating different elasticities for specific groups of firms. In the empirical part, we take advantage of this feature to check the robustness of our basic estimate.

On the other hand, multiplying demands (9) by output prices, replacing prices on the right hand side by their optimal choice according to (11), splitting marginal cost and taking logs we have the revenue system

$$\begin{aligned} r_{jt}^X &= \varphi^X + \eta_X p_t^X - (\eta_X - 1)\overline{mc}_{jt} + z_{jt}^X \alpha_X + (\eta_X - 1)\omega_{jt} + \delta_{jt} \\ r_{jt}^D &= \varphi^D + \eta_D p_t^D - (\eta_D - 1)\overline{mc}_{jt} + z_{jt}^D \alpha_D + (\eta_D - 1)\omega_{jt} + \delta_{jt}. \end{aligned} \tag{14}$$

where φ^X and φ^D are constants.[28,29]

These equations show how revenue in each market depends on the $\overline{mc}_{jt}$ component of marginal cost, observed product advantages, the unobserved cost advantage ω_{jt} and the unobserved demand advantage δ_{jt}. They generalize Aw, Roberts and Xu (2011). Appendix B develops the corresponding equations if marginal cost differs across the two markets.

Equations (14) can be solved for ω_{jt} and δ_{jt}. The solution gives

$$\begin{aligned} \omega_{jt} &= \gamma^X - p_t + (1/d)((r_{jt}^X - z_{jt}^X \alpha_X) - (r_{jt}^D - z_{jt}^D \alpha_D)) + \overline{mc}_{jt}, \\ \delta_{jt} &= \gamma^D - p_t + ((\eta_X - 1)/d)(r_{jt}^D - z_{jt}^D \alpha_D) - ((\eta_D - 1)/d)(r_{jt}^X - z_{jt}^X \alpha_X), \end{aligned} \tag{15}$$

[26] $\ln \frac{R_{jt}}{C_{jt}} = \ln(1 + \frac{R_{jt} - C_{jt}}{C_{jt}}) \simeq \frac{P_{jt}Q_{jt} - AC_{jt}Q_{jt}}{AC_{jt}Q_{jt}} = \frac{P_{jt} - AC_{jt}}{AC_{jt}} = PACM_{jt}$ is a profitability measure that we call price-average cost margin (PACM). Alternatively, profitability can be measured with the rate of short-run economic profitability $\pi_{jt} = \frac{R_{jt} - C_{jt}}{R_{jt}}$. Notice that $\pi_{jt} = \frac{PACM_{jt}}{1 + PACM_{jt}}$.

[27] This way to estimate elasticities can be related to De Loecker and Warzynski (2012) estimation of firm's markups. Let's suppose only one market and call the markup $\mu_{jt} = \frac{P_{jt}}{MC_{jt}}$. De Loecker and Warzynski (2012) propose to estimate markups as $\mu_{jt} = \nu/(C_{jt}/R_{jt})\exp(-\epsilon_{jt})$ using previous estimates of ν and the disturbance ϵ_{jt}. Our equation reorders this expression as $R_{jt}/C_{jt} = \frac{1}{\nu}\mu_{jt}\exp(\epsilon_{jt})$ and estimates in one stage, giving a different interpretation to the disturbance. They stress the heterogeneity of μ_{jt}, we are mainly interested in splitting it as the outcome of operating in two different markets.

[28] The marginal cost component $\overline{mc}_{jt}$ can take different forms. We discuss later our specific choices.

[29] $\varphi^X = \ln \alpha_0^X - (\eta_X - 1)\ln \frac{\eta_X}{\eta_X - 1}$ and $\varphi^D = \ln \alpha_0^D - (\eta_D - 1)\ln \frac{\eta_D}{\eta_D - 1}$.

where $d = (\eta_X - 1) - (\eta_D - 1)$ and $p_t = (1/d)(\eta_X p_t^X - \eta_D p_t^D)$.[30]

We specify equations (14) as follows. First, we replace the unobservables by first order exogenous Markov processes with ω_{jt-1} and δ_{jt-1} replaced by their expressions according to (15). We use in-homogeneous Markov processes which include time effects because the equations (15) contain a common price term that we cannot strictly observe, $\omega_{jt} = q_t + q(\cdot) + \xi_{jt}$ and $\delta_{jt} = s_t + s(\cdot) + \varepsilon_{jt}$, where q_t and s_t represent time effects. These time effects collapse with the other time effects present in the equations (the time effects representing $\eta_X p_t^X$ and $\eta_D p_t^D$).

Second, we specify $\overline{mc}_{jt}$ and $\overline{mc}_{jt-1}$ using two different expressions. Inside the unknown $q(\cdot)$ function we use the lagged first order condition for materials solved for $\overline{mc}_{jt-1}$, so we have $\overline{mc}_{jt-1} = -\ln \beta_M - \beta_0 + p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1}$. For $\overline{mc}_{jt}$, we use the expression that results from aggregating equations (12), $\overline{mc}_{jt} = -\ln(\beta_L + \beta_M) - \beta_0 + c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt} - e_{jt}$. Part of this expression goes to the constants and another to the disturbances.

The resulting system of revenue equations can be written as[31]

$$\begin{aligned} r_{jt}^X = \; & a_t^X - (\eta_X - 1)(c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt}) + z_{jt}^X \alpha_X \\ & + g_1[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - (r_{jt-1}^D - z_{jt-1}^D \alpha_D) \\ & \qquad + d\ (p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1})] \\ & + h_1[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + v_{1jt} \end{aligned} \tag{16}$$

$$\begin{aligned} r_{jt}^D = \; & a_t^D - (\eta_D - 1)(c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt}) + z_{jt}^D \alpha_D \\ & + g_2[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - (r_{jt-1}^D - z_{jt-1}^D \alpha_D) \\ & \qquad + d\ (p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1})] \\ & + h_2[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + v_{2jt}, \end{aligned} \tag{17}$$

where a_t^X and a_t^D are combinations of a constant and time effects. The terms $g_1(\cdot), h_1(\cdot), g_2(\cdot)$ and $h_2(\cdot)$ are unknown functions and the disturbances are $v_{1jt} = (\eta_X - 1)e_{jt} + (\eta_X - 1)\xi_{jt} + \varepsilon_{jt}$ and $v_{2jt} = (\eta_D - 1)e_{jt} + (\eta_D - 1)\xi_{jt} + \varepsilon_{jt}$.

[30] $\gamma^X = (\varphi^D - \varphi^X)/d$ and $\gamma^D = -((\eta_X - 1)\varphi^D - (\eta_D - 1)\varphi^X)/d$.

[31] We use the fact that an unknown function $\widetilde{q}(d + x)$, where d is a constant, can be written as $c + q(x)$, where c is another constant. We also collapse in the coefficients of the function any parameters that multiply the unknown function or its argument.

3.5 Identification and back up of ω_{jt} and δ_{jt}.

Identification hinges on equations (13), (16) and (17). We need to estimate parameters β_K, β_L and β_M of the production function (marginal cost), demand elasticities η_X and η_D and shift parameters α_X and α_D to be able to backup ω_{jt} and δ_{jt} using equation (15). In principle, all parameters can be estimated from equations (16) and (17), but identification of the elasticities and the parameter of scale ($\nu = \beta_L + \beta_M$) using only these equations seems quite weak. Equation (13) is a very robust relationship that cannot identify the elasticities and the parameter of scale by itself. We apply NLS to equation (13) to estimate the functions $a = \ln \frac{1}{\nu} \frac{\eta_D}{\eta_D - 1}$ and $b = \frac{\eta_D}{\eta_D - 1} / \frac{\eta_X}{\eta_X - 1} - 1$. We then plug these estimates as restrictions in the system formed by (16) and (17) and estimate all parameters of the revenue system by nonlinear GMM.

To back up ω_{jt} and δ_{jt}, we employ equations (15) implemented using a rough estimate of the common time index p_t.[32]

4. Estimation

4.1 A system of semiparametric equations.

The model consisting of (16) and (17) is a system of semiparametric equations, the equations have a linear and a nonparametric part (see Robinson 1988). Each equation has two nonparametric functions, the pairs (g_1, h_1) and $(g_2,\ h_2)$. The arguments of the nonparametric functions are log-linear expressions of observed variables. The disturbances are uncorrelated over time and across firms, but can be freely correlated among them.

The system is fully nonlinear in parameters for three reasons: we impose the restrictions implied by equation (13), for each equation there are cross-restrictions between the parameters of the linear part and the nonparametric part, and there are cross-restrictions between the two equations. In fact, the restrictions involving the linear part and the nonparametric functions contribute to identification (we build on the similar uniequational estimation by Doraszelski and Jaumandreu, 2013).

We approximate the nonparametric functions by means of third order polynomials. We are mod-

[32] We have no separate observations on P_t^X and P_t^D. In estimating (14), the price movements are absorbed in the time dummies. To recover ω_{jt} and δ_{jt}, we will approximate the changes in p_t from the changes in the only industry index that is available. The approximation works well because both price indices are likely to move during the period in parallel. In fact, more refined alternatives have produced the same results.

eling the unobserved advantages as exogenous processes, therefore each function is univariate and requires only the estimation of three coefficients. In the empirical part, we use four demand shifters in each equation (*Location*, *Age/Experience*, *Subsidy*, *Sales effort*, see subsection 6.2). This implies a total of 13 parameters of theoretical interest ($\beta_K, \beta_L, \beta_M$,$\eta_D$, η_X, and the four-dimension vectors α_X and α_D). However, we have to estimate 30 more: two constants, eight time dummies in each equation and twelve coefficients of the polynomials. To avoid a nonlinear search on 43 parameters, we proceed concentrating-out the 32 parameters that enter linearly.

4.2 Selection.

Our sample consists of time sequences of observations for firms that are observed both exporting and selling in the domestic market. It might be that the disturbances of equations (13), (16) and (17) are correlated with the decision to export, therefore creating a sample selection bias. We wish to draw inferences which are valid for all firms. For example, we are interested in demand elasticities or production function coefficients that can be attributed to all firms, not only the exporting subset. We proceed checking whether there are biases and, if this is the case, addressing them.

Let us discuss the subtle case of why selection may happen with the shocks of equations (16) and (17). Theoretical models provide reasons by which firms self-select into the export market according to their productivity levels. Empirical papers have found that this is the case (see, for example, Melitz and Redding, 2012). In our case, self-selection may happen for both productivity and product advantages. The predictable part of productivity and product advantages has been replaced in our equations by observables, so they do not constitute a problem. However, the current (unpredictable at $t-1$) shocks are still present.

Suppose, generalizing on Olley and Pakes (1996), that the firm's rule for exporting is that the combination of productivity and product advantages should be above some threshold that is a function of the state variables, namely capital and the value of the demand shifters. This is what happens if we consider the firm taking dynamically profit maximizing decisions in the presence of fixed cost of exporting. If the firm makes the decision to export in the same period of the shock, then the state variables and the shock will be correlated conditionally in the continuation in the export market (only firms with more capital and stronger demand shifters value will accommodate the most negative shocks). This is indeed possible. However, it is not particularly likely. As Ackerberg, Benkard, Berry and Pakes (2007) remark, this case depends on the anticipation of the shock and the immediate reaction of the firm entering or withdrawing (from the export market).

We estimate equation (13) including the inverse Mills ratio based on a probit estimation for the decision to export in the universe of exporters and non-exporters. In the system, we check for possible sample selection by extending Olley and Pakes (1996) procedure of inverting the probability of exporting to control for the unobservable threshold in our two-dimensional setting. We find a slight but significant effect of selection in margins and elasticities, which we correct accordingly, but we do not find any effect in the system.

4.3 Endogeneity.

Once ω_{jt-1} and δ_{jt-1} have been replaced by observables, the problems of endogeneity are limited to the possible correlation between any of the included variables and the composite disturbances v_{1jt} and v_{2jt} through the components e_{jt}, ξ_{jt} and ε_{jt}.

Our specification of marginal cost brings three endogenous variables: c_{jt}, l_{jt} and m_{jt}. They are endogenous because they are determined at a moment of time at which ξ_{jt} and ε_{jt} are known. Variables l_{jt} and c_{jt} are also correlated with e_{jt}. The exogenous marginal cost determinants are $k_{jt}, c_{jt-1}, k_{jt-1}, l_{jt-1}$ and m_{jt-1}. In practice, k_{jt} and k_{jt-1} are strongly correlated and we focus on k_{jt-1}. We have to estimate only three marginal cost related parameters (β_K, β_L and β_M) so the four remaining variables are enough to identify them. Notice that the lagged values help to estimate the coefficients of the endogenous variables because the coefficients in different parts of the equations are the same (for example, variables l_{jt} and l_{jt-1} share the same coefficient β_L).

Some of the demand shifters are potentially endogenous, they might be correlated with the disturbances v_{1jt} and v_{2jt} through their components ξ_{jt} and ε_{jt}. This is not the case with the location of the firm or its age/experience in the market, because location and entry were probably decided time before the realization of the disturbances. It is more likely that the reception of a subsidy is related to a shock suffered contemporaneously by the firm. The choice in sales effort likely occurs after the disturbances are realized (implying the same timing that we assume for the variable inputs l_{jt} and m_{jt}). To be safe, we only use moments dated at time $t-1$ for all demand shifters.

4.4 Instruments.

Let $\widehat{\gamma} = (\widehat{a}, \widehat{b})$ be the parameter estimate from equation (13). After plugging in this estimate, write the residuals of (16) and (17) as a function of variables x_j and vector θ of parameters that remain to be estimated. Let them be the $T_j \times 1$ vectors $v_{1j}(x_j, \theta, \widehat{\gamma})$ and $v_{2j}(x_j, \theta, \widehat{\gamma})$. We base estimation on the moments

$$E\begin{bmatrix} A(z_j)v_{1j}(x_j, \theta, \widehat{\gamma}) \\ A(z_j)v_{2j}(x_j, \theta, \widehat{\gamma}) \end{bmatrix} = 0,$$

where $A(\cdot)$ is a matrix of functions of the exogenous variables z_j, with dimensions $L \times T_j$, with L denoting the number of moments (we employ the same set of instruments for each equation). The literature on optimal instruments (Amemiya, 1974; Newey, 1990, 1993) establishes that variance can be minimized by functions of the form

$$A(z_j) = E\left[\frac{\partial v_{\cdot j}(x_j, \theta_0, \widehat{\gamma})}{\partial \theta} | z_j\right],$$

where the dot indicates 1 or 2 and θ_0 is the true value of θ.

In our equations, the derivatives inside the expectation turn out to be linear in the endogenous variables, and these variables can be expressed in terms of the lagged observables. In addition, the derivatives of the unknown functions enter the expectation because parameters show up inside these functions. All this suggests that a good approximation to the expectations can be obtained using polynomials on all variables inside the unknown functions and some interactions.

We use the following instruments for each equation: a constant, a set of time dummies, the dummy for location; a complete third order polynomial in the key variables k_{jt-1}, l_{jt-1}, and m_{jt-1}; a third order polynomial in c_{jt-1}; variable p_{Mt}. We add univariate third order polynomials in the lagged shifters Age_{jt-1}, $Experience_{jt-1}$, $Subsidy_{jt-1}$, $Saleseffort_{jt-1}$ that we enlarge with a polynomial in the variable $State\ participation_{jt-1}$. Additionally, we found the interactions between $Subsidy_{jt-1}$ and m_{jt-1}, as well as $Sales\ effort_{jt-1}$ and m_{jt-1} to be important. We use 50 instruments in each equation to identify 43 parameters. We get reasonable estimates in the 10 industries using exactly the same set of instruments.

4.5 Estimation procedure and consistent standard errors.

We set the GMM problem as

$$\min_{\theta} \begin{bmatrix} \frac{1}{N}\sum_j A(z_j)v_{1j}(x_j, \theta, \widehat{\gamma}) \\ \frac{1}{N}\sum_j A(z_j)v_{2j}(x_j, \theta, \widehat{\gamma}) \end{bmatrix}' \widehat{W} \begin{bmatrix} \frac{1}{N}\sum_j A(z_j)v_{1j}(x_j, \theta, \widehat{\gamma}) \\ \frac{1}{N}\sum_j A(z_j)v_{2j}(x_j, \theta, \widehat{\gamma}) \end{bmatrix}$$

where N is the number of firms and we use the consistent weighting matrix

$$\widehat{W} = \begin{bmatrix} (\frac{1}{N}\sum_j A(z_j)A(z_j)')^{-1} & 0 \\ 0 & (\frac{1}{N}\sum_j A(z_j)A(z_j)')^{-1} \end{bmatrix}.$$

Our two-stage procedure implies that we have to estimate consistent standard errors (see Wooldridge, 2010). Stacking all moments in the vector $g(w_j, \theta, \widehat{\gamma}) = \begin{bmatrix} A(z_j)v_{1j}(x_j, \theta, \widehat{\gamma}) \\ A(z_j)v_{2j}(x_j, \theta, \widehat{\gamma}) \end{bmatrix}$, where w_j is the union of vectors x_j and z_j, the GMM problem can be more compactly written as

$$\min_{\theta}[\frac{1}{N}\sum_j g(w_j, \theta, \widehat{\gamma})]'\widehat{W}[\frac{1}{N}\sum g(w_j, \theta, \widehat{\gamma})],$$

and the assymptotic variance of $\widehat{\theta}$ expressed as

$$Avar(\widehat{\theta}) = \frac{(G'WG)^{-1}G'WDWG(G'WG)^{-1}}{N}.$$

where $G = E[\nabla_\theta g(w_i, \theta_0, \gamma_0)]$ and W is the probability limit of $\widehat{W}$. The derivation of matrix D is in Appendix C. It reflects both the variance of the moments in the GMM estimation and the previous NLS estimation. The assymptotic variance $Avar(\widehat{\theta})$ is estimated by replacing the probability limits by estimates and computing matrix D as shown in the appendix.

5. Data.

5.1 Source and treatment.

Our data comes from the Annual Census of Industrial Production, a firm-level survey conducted by the National Bureau of Statistics (NBS) of China. The target of the census is all industrial non-state firms with more than 5 million RMB in annual sales plus all industrial state-owned firms (SOEs).[33] The source is the same as in Brandt, Van Biesebroeck, and Zhang (2012); we draw intensively on their work at the time of treating the data.[34] Our data was collected from 1998 to 2008. In what follows, we describe how we build a panel data set, construct variables and clean the data.[35]

[33] After 2006, SOEs with less than 5 million RMB are excluded from the survey. This affects only a few firms; we count 22 firms in 2006 that did not answer the survey the following year.

[34] Other recent studies which use this source are Roberts, Xu, Fan and Zhang (2016); Lu and Yu (2015); Ma, Tang and Zhang (2014); Aghion, Cai, Dewatripont, Du, Harrison and Legros (2013); Brandt, Van Biesebroeck, Wang and Zhang (2012); Song, Storesletten and Zilibotti (2011); Lu (2010) and Hsieh and Klenow (2009).

[35] We want to use the data as a panel of firms. We want to exploit all the observations repeated over time which are available for the same individual. One reason is that our modeling implies persistent productivity and product advantages that evolve over time. Therefore, their estimation depends on the sequence of observations for the firm.

In the raw data, the same firm can show up at different moments with different identifiers. It is very important to link these separate observations for two reasons: to get the right time sequences of observations for each firm, and to determine if the firm shutdown during the period. We describe the linking process and analyze its results in Section A5.1 of the Online Appendix.[36] After linking the dat,a we find reasonable rates of economic entry, expansion and exit, which average 9.4%, 7.8% and 7.9% respectively. Many additions come from firms growing large enough to be included in the survey. Additionally, there are improvements over time in statistical coverage that need to be accounted for to get the right interpretation of the numbers.[37]

The survey information includes location, industry code, the date of creation, details on ownership and some financial information. We obtain or construct revenue (split into domestic sales and exports), an estimate of physical capital, wage bill, cost of materials, subsidies, the number of workers and the amount spent on sales promotion and (for a few years) on R&D investment. In Appendix D, we detail the content of these items as well as the definition of other variables. Using the industry codes, we allocate firms into ten industries. In Appendix E, we describe the correspondence with the two-digit codes breakdown and list the number of four-digit codes included in each industry.

We check for consistency of the variables and clean the data by dropping abnormal observations.[38] We then use the firm's longest time subsequence of complete data, provided that is longer than one year. The cleaned data set retains 84% of the firms and 74% of the raw observations.

5.2 Growth and reallocation.

The treated data shows that Chinese firms underwent important gowth during this period. Aditionally, there was a large reallocation of manufacturing activity.[39] Only 25% of the firms in the starting year reach the final year (survivors). The rest shutdown before the final year (exitors). However, due to entry and additions, the total number of firms in the data nearly triples. About 77% of the firms in the final year are born during the time period (entrants). The other 23% consists of surviving firms plus additions to the database. Survivors in the final year only represent about

[36] Our manufacturing linked data is very similar to Brandt, Van Biesebroeck and Zhang (2012), but the focus of our analysis is individual firms dynamics. We exclude firms with a single observation from this analysis and we systematically and separately identify entrants from additions to the survey.

[37] Additions often sell more than 5 M RMB and the ratio of data aggregates to industry GDP estimates in the China Statistical Yearbook is increasing. See Section A5.1 of the Online Appendix for details.

[38] As described in Section A5.2 of the Online Appendix.

[39] Details can be checked in Table 0c of the Online Appendix.

9% of firms and the additions to the survey constitute 14%.

Survivors experience significant growth over time, but entrants and additions to the survey are significantly smaller.[40] The exitors, although smaller than survivors, are in turn larger than entrants and additions. The result is that output and productivity increase sharply at the same time that production becomes dominated by newer smaller firms. The production of the average firm roughly triples. However, average capital tends to decrease and average employment decreases by a third from its starting level.

5.3 Sample.

We draw our sample by selecting all the available (continuous) time sequences of firms operating in the domestic and foreign markets. The sample shares all the previously discussed characteristics with two important distinctions. First, turnover now includes firms that start to export and stop exporting. Some of the firms present initially leave the sample because, although still alive they stop exporting. A consequence is that the proportion of firms in the first year of the sample which stay until the final year is somewhat smaller (20%). Firms that leave the sample now include the firms that stop exporting in addition to the shutdowns. Firms that join the sample now include existing firms that start to export together with the new born entrants. Second, the average sizes of all categories of firms are roughly twice global averages.

We compute a standard TFP measure: the growth of deflated revenue minus the growth of capital, labor and deflated materials, weighted by the average cost shares between t and $t-1$ computed using a common cost of capital. TFP growth is strong, especially after 2001, and averages 2.8%, for all of the treated data and our sample. This estimate exactly matches the main estimate by Brandt, Van Biesebroeck, and Zhang (2012).

5.4 Descriptive statistics.

Tables 1 and 2 provide descriptive statistics of the sample by industry. In total, there are more than 73,000 firms and 290,000 observations.

For firms in the sample, column (3) of Table 1 reports the share of industry sales in 2008 and column (4) states the firms' average export intensity (the proportion of sales in foreign markets). Firms in the sample represent between 20% to 70% of corresponding industry sales, 40% or more

[40] From here on, we measure firm size by employment. Results are similar if we use capital.

in the most technologically intensive industries. The average export intensity ranges from 35% to 60%, depending on industry. In each industry, less than 25% of firms are located in Middle or Western China (column 5), and the percentage tends to be significantly lower in most industries. The average age is between 8 and 14 years (column 6), but firms differ little in their average export experience (column 7). To summarize, firms in the sample explain an important fraction of sales in each industry. They tend to represent a greater portion the more technologically intensive the industry is, and export a large part of their sales. However, on average, firms are younger with limited experience in the export market.

The 2000s witnessed a massive change of ownership of Chinese firms. In the sample, columns (8) to (11) document this fact. For simplicity, we do not pay attention to the exact level of participation of the state[41] and we categorize firms with participation versus firms without participation. Then, adding a dynamic dimension referred to the whole sample period, we classify them as "always state (participated)," "always private" and firms that experience a change in their participation (mostly from state participated to private).

At the beginning of the period, in 1998, each type of firm represents roughly a third of all firms. At the end of the period, in 2008, "always state" firms represent 4.5% while "always private" firms account for at least 80% in all industries. This radical change in composition has two sources: the shutdown of many state participated firms and the overwhelming proportion of entrants that are private (at least 90% in all industries). The rest are firms that experience a change in their status, mostly from state participated to private. The absolute number of this type of firms is roughly stable over the time period. However, its relative share has shrunk due to entry of new firms.

Table 2 contains more descriptive statistics. Between 15% to 30% of firms receive state subsidies (columns 1 and 2), but their average value is very small (less than 1% of revenue). Subsidies are orthogonal to the participation of state in the financial capital of the firm. Foreign participation (columns 3 and 4) occurs in 20% to 35% of the firms and, when it is present, it represents, on average, a solid majority of financial capital.

Economic and investment data show significant heterogeneity across industries. The average size of firms, measured by employment in column (5), is between 290 to 760 workers. Across industries, margins (PACMs) range from 10% to 20%. Virtually all firms show some sales effort (column 7). The average intensity is between 2.5% to 6% of sales (column 8). The proportion of firms that invest

[41] State participation, when positive, is on average very high (60% to 70% of capital), but 30% to 40% of firms have state participations under 50% of finacial capital.

in R&D (column 9) ranges from 8.5% to 32%. The average intensity of these investments (column 10) goes from a little less than 1% to 2%, indicating varying degrees of technological sophistication. Finally, we report the log of the firm-level average wage divided by the industry-wide average wage (column 11). Under the assumption of a competitive labor market, where firms pay the value of marginal productivity to their workers, this ratio provides a measure of the degree by which the firm-level average marginal productivity of labor exceeds (falls short of) the industry level average. This average is likely to be closely related to the degree of workforce skills. It is an indicator of the relative quality of the firm-level labor input. This index of labor quality tends to show moderately negative average values and great intraindustry dispersion.

6. Results.

6.1 Estimating functions a and b.

Table 3 reports the estimates of functions a and b carried out in order to constrain the estimation of the demand elasticities η_D and η_X and the parameter of scale ν (see subsection 3.5). Columns (1) and (2) report the result of regressing the dependent variable, the log of revenue over variable cost or margin (PACM), on a constant and the nonlinear effect of the share of exports in sales (*Export intensity),* according to specification (13). Column (3) reports the root mean square error of the equations and reveals a reasonable fit. In fact, the estimated equation is linear enough for the R-squares to be meaningful. They range from 0.54 to 0.66.

The first result that emerges from the estimation is that the more a firm exports the lower its total margin. Taken as a simple (quasi) linear predictor, the equation says that the domestic margin is the largest margin and that the total margin decreases with the intensity of the exports. The value of the domestic margin by industries is given by the intercept in column (1). The slope of the predictor can be read as the difference in percentage points between the typical domestic and export margins. The difference ranges from the 3 percentage points in *Metals* to the 11 percentage points in *Food.* The average across all industries is 6 percentage points.

Under our structural interpretation, we are estimating the functions $a = \ln \frac{1}{\nu} \frac{\eta_D}{\eta_D - 1}$ and $b = \frac{\eta_D}{\eta_D - 1} / \frac{\eta_X}{\eta_X - 1} - 1$. At this stage, we cannot disentangle the value of demand elasticity from ν in the domestic market and we can only get a relative assessment of the elasticities by looking at the estimated b function. However, the estimates for function b anticipate an important difference in the elasticities. The data confirms what we expect from theory: the demand for exports has a larger elasticity given competition with a higher number of susbtitute goods.[42] This has an important pricing implication: given marginal cost the firm is expected to set a lower price in the exports market to equate marginal revenue in both markets.[43]

Robustness checks and selection.

Because our estimates of the functions a and b play an important role in the estimation of the system, we want to check their robustness with respect to potential problems of mismeasurement. We check the possible effects of adjustment costs of labor, subsidy distorsions (Hsieh and Klenow, 2010), shocks to demand, transportation costs and time varying margins.[44] Later, we also report the result of relaxing the assumption of common demand elasticicities for the firms in a given industry. We find the estimation quite robust for all of these misspecifications and leave the equation unchanged. However, sample selection is relevant and we correct for it as follows.

We estimate the probability of exporting, using observations from all firms in each industry, by means of a nonparametric specification.[45] The inverse Mills ratio corrects for a slight downward bias when estimating the domestic margins, and it leaves the difference between domestic and exports margin almost unchanged (see columns 4, 5 and 6). To estimate the system, we use the a and b in columns (4) and (5). Columns (7) and (8) report for reference (and to help with intuition) the levels of the domestic and exports price-average cost margins.

6.2. System for exports and domestic sales.

Table 4 summarizes the results of estimating equations (16) and (17) subject to the restrictions implied by functions a and b estimated in the first step.

[42] This coincides with what classical structural IO literature tended to find: less market power and smaller margins in export markets. See, for example, Bernstein and Mohnen (1991), Bughin (1996) and Moreno and Rodriguez (2004). The likely higher toughness of competition in foreign markets has also been recently underlined by the theoretical trade literature (see, for example, Mayer, Melitz and Ottaviano, 2014 and 2016).

[43] One may wonder what is the average difference of PACMs between exporters and non-exporters. This question is not relevant here but is related to the difference in markups between these two kind of firms addressed, for example, in De Loecker and Warzysnski (2012). We give an answer to the question of the difference of PACMs in Section A6 of the Online Appendix, where we show that exporters tend to have slighly smaller *global* margins than non-exporters but also slightly greater *domestic* margins. We cannot strictly give an answer to the question of different markups because it cannot be excluded that exporters and non-exporters diverge systematically in the ratios average cost/marginal cost. However, see the comments in the appendix.

[44] Table 3c in the Online Appendix reports the robustness checks, which are commented on there.

[45] We consider a complete second order polynomial in the following lagged variables: capital, wage, materials, age, subsidy and sales effort.

The production function parameters are key coefficients. Columns (1) to (3) show their point estimates and standard errors. The elasticities of the inputs are reasonable, as are the global returns to scale. In *Chemical*, *Transport* and *Electronics* the returns to scale are not distinguishible from unity. In *Food*, *Textile*, *Furniture*, *Paper*, *Metals* and *Machinery* they range between 0.92 and 0.96. Only *Non-metals* are 0.90.

The elasticities of demand, estimated simultaneously, are reported in columns (4) and (5). Their relative values make full sense. In *Electronics,* firms have the greatest market power, both domestically and abroad, whereas *Textile* is (almost) perfectly competitive. Elasticities in the world market are systematically greater than in the domestic market, sometimes by a significant extent. This supports the identification approach of this paper.

Our elasticities lie on the right tail of the distribution of elasticities estimated in the IO and trade literature. This reflects the specificity of the Chinese economy and its exports. However, it is also a byproduct of the way we estimate them. The elasticities of demand are identified by the observed margins and the simultaneously estimated parameter ν. Our method of identification, in contrast with other methods, requires the mutual consistency of three measurements: short-run profits, parameter of scale and elasticities.[46] The three estimates are sensible, which is unusual for exercises of this type.[47] One important characteristic is that the estimated elasticities are robust with respect to the presence of product advantages correlated with prices.

Recall that demand shifters are used to control for all observable product advantages. We have included four shifters in each equation: *Location,*[48] *Age* of the firm (sometimes replaced by *Expe-*

[46]The rate of short-run economic profits is a simple transformation of the price-average cost margins: $\pi_{jt} = \frac{R_{jt}-C_{jt}}{R_{jt}} = 1-\nu(\frac{\eta_D-1}{\eta_D}S_{jt}^D + \frac{\eta_X-1}{\eta_X}S_{jt}^X) = 1-\nu+\frac{\nu}{\eta_{jt}}$ where $\frac{1}{\eta_{jt}} = \frac{1}{\eta_D}S_{jt}^D + \frac{1}{\eta_X}S_{jt}^X$ is a weighted average of the inverse of the elasticities (or aggregate markup). This generalizes a similar expression which holds when there is only one market and therefore one elasticity. Short-run economic profits, scale parameter and elasticities are linked by this expresion.

[47]To estimate markups, the literature has followed two broad approaches. In one, elasticities are estimated from the specification of a demand system (see for example Hottman, Redding and Weinstein, 2014, for a recent application). In the other, markups are derived from the first order of one factor or several factors together. This is the traditional Solow-based Hall (1990) method, recently applied by De Loecker and Warzynski (2012). Wathever the approach, the estimates have implications for profits, but often they are not developed or tested against observations. Short-run profits are equal to markups only if $\nu = 1$, and profits are greater (smaller) than markups if ν is below (above) unity. When the first method is employed there is usually no available estimate of ν and profits remain ignored. When the second method is applied, ν is often left implicit although it offers (through profits) a nice test about the likelihood of the estimates (Gordinchenko, 2012, makes also this point). For example, an elasticity of 3 with a parameter $\nu = 0.9$ implies short-run profits of 40% , which is hard to believe.

[48]In the face of the difficulties for treating the dummy variable *Location* as an argument of the nonparametric

rience),[49] *Subsidy* and *Sales effort*. The shifters are common to both equations but we allow for different impacts in each market.

The variable *Sales effort* is the most important shifter (see columns 9 and 13). We enter the expenditure in logs, so that the coefficients can be read as elasticities of revenue with respect to the value of these expenditures. Elasticities are positive and, in nine out of the ten industries, significant both in the domestic and exports equation. The average elasticities of the significant values are 1.5 and 1 respectively, but they range from 0.3 to 3.8. Promotion tends to be more effective in the export market.[50]

The shifter *Subsidy* (columns 8 and 12) is often non-significant.[51] In general, subsidies are associated with more sales (domestic and exports) in *Paper, Machinery* and *Transport,* and with less sales in *Food* and *Chemical.* The variable *Age,* columns (7) and (11), explains significant positive differences in sales in both markets for *Chemical, Metals, Machinery, Transport and Electronics* (in the export market for *Machinery* and *Electronics* the variable used is *Experience*). Firms located in the *Middle-West* area tend to have less sales, particularly in the export market.

Robustness checks on the estimation of the system.

We carry out robustness checks for our assumption of equal elasticities, the presence of arbitrary forms of heterogeneity across four-digit subindustries, the effect of selection, and products having market specific impacts.[52]

We reestimate equation (13) with elasticities that change with the size of firms, quality of products (measured through workforce skills), and foreign participation, and then we reestimate the system.[53] Neither the coefficients of the system nor the estimated productivity and product advantages (com-

functions we finally gave up and included it in the linear part of the equations. Its coefficients shoud be consequently read as reduced form impacts.

[49] We never include *Age* and *Experience* together because are highly correlated variables.

[50] These elasticities give us an interesting check of the internal consistency of the estimates and of the price elasticities. Recall that, by the Dorfman and Steiner (1954) condition, the optimal value of sales effort expenditures over revenue should equal the ratio of elasticities with respect to sales effort and price. Dividing column (9) by column (4) and column (13) by column (5) one gets the optimal values of sales effort according to our estimates. The domestic values range from 5% to 18% , with a mean of 9.7%, and the exports values range from 3% to 10% with a mean of 5.6%. The ratios are all quite sensible and hence a reason to trust the estimates.

[51] The shifter *Subsidy* is in per unit terms and coefficients are therefore semielasticities. For example, in the *Food* industry, a subsidy of 1 percent of sales is associated to 2 percent less sales for a given price (and rest of shifters) in the domestic market, and 6 percent less sales in the exports market. Recall from Table 2 that the average subsidy is 1% of revenue.

[52] The results are reported in detail in Section A6 of the Online Appendix.

[53] We find elasticity effects of these variables but none for the location and age of the firm.

pared with the estimates that we are going to report for the main specification) change significantly. Our final assessment is that the specification of heterogeneous elasticities along these lines is a feasible refinement that does not modify the basic results. To consider arbitrary forms of heterogeneity, we reestimate the system including the corresponding subindustry dummies at the four-digit level. This uses 392 dummies (see Appendix E). The new specification induces very small changes in the estimates of the coefficients, productivity and product advantages.

Our sample only considers firms that simultaneously sell in the domestic and export markets. As explained in Section 4, this raises the possibility that the system should be corrected for sample selection. Recall that if there is sample selection the expectation of the Markov processes becomes a function of an unobserved threshold. We test for possible sample selection bias and we conclude that there is no need for correction.

The same product attributes can have a different impact in the domestic and export market. We allow for this possibility by estimating an additional parameter λ as coefficient of δ_{jt} in the domestic market. We conclude that the generalization to models with market specific product advantages is highly desirable. However, the current model is not too restrictive when imposing constraints, at least for an important part of the industries.

6.3 The estimated ω_{jt} and δ_{jt}.

Once we estimate the parameters of the system, we can back up ω_{jt} and δ_{jt} using equations (15). We back up both unobservables in differences with respect to the mean in each industry and, abusing notation, we keep the symbols unchanged. As a result, we can read the values of ω_{jt} and δ_{jt} as reflecting percentage differences with respect to the advantages of a hypothetical firm with average advantages in this particular industry and period. An important outcome of this transformation is that we can compare the values of ω_{jt} and δ_{jt} across industries.

We report, for better comparability, ω_{jt} and $\delta_{jt}/(\eta_D - 1)$. Variable ω_{jt} reflects productivity differences. With markup pricing with common elasticities, these differences become also efficiency-driven price differences between firms. Recall from the model that δ_{jt} reflects percentage quantity advantages given price, so it has a different scale than ω_{jt}. Inverting revenue, it is easy to see that $\delta_{jt}/(\eta_D - 1)$ can be read as the implicit willingness of consumers to pay a different price from the baseline price. We choose to divide δ_{jt} by $\eta_D - 1$ as a matter of convenience, but the results could also be presented in terms of $\delta_{jt}/(\eta_X - 1)$ (recall that the ratio η_D/η_X is constant for each industry). At some point, we are going to multiply δ_{jt} by the firm-level weighted average of the inverse elasticities $\frac{1}{\eta-1} = S_{Djt}\frac{1}{\eta_D-1} + S_{Xjt}\frac{1}{\eta_X-1}$, where S_{Djt} and S_{Xjt} are the firm-level revenue shares of domestic sales and exports. However, we prefer not to abuse this expression because shares are endogenous.

Distribution of ω_{jt} and δ_{jt}

Columns (1) through (9) of Table 5 summarize the marginal distributions of ω_{jt} and $\delta_{jt}/(\eta_D-1)$. Figure 1 depicts the level sets of the joint density of ω_{jt} and $\delta_{jt}/(\eta_D-1)$ at the starting and final time intervals (1998-2000 and 2005-2008).[54]

Columns (1) to (6) of Table 5 report the quartiles of ω_{jt} and $\delta_{jt}/(\eta_D-1)$ in the initial and final year of the sample, 1998 and 2008. Columns (7) and (8) report the standard deviations of ω_{jt} and $\delta_{jt}/(\eta_D-1)$ for the same years. Column (9) provides a measure of the skewness for the whole distribution of values ω_{jt} and $\delta_{jt}/(\eta_D-1)$.

Both the interquartile ranges and the standard deviations describe a significant dispersion that tends to be somewhat greater in the product advantages (with the exception of industries *Timber* and *Non-metals,* the industries with less product differentiation). The interquartile ranges of ω_{jt} show differences between 40% and 60% (in *Electronics* the interquartile ranges of ω_{jt} are larger in both years). The corresponding differences in willigness to pay range between 35% and 85% (in *Machinery* and *Electronics,* there are larger interquartile ranges). Given the values of the elasticities, the ranges of $\delta_{jt}/(\eta_D-1)$ imply huge differences in sales for the same prices. This is a notable dispersion, but it simply mirrors the real dispersion of revenues for firms with similar costs and productivity. As shown in column (9), cost advantages are fairly symmetric, but product advantages are systematically skewed to the left (except in *Metals*).

Our ability to compare our estimates to previous measurements is limited because of differences in methodology and samples. Our results are not "revenue" measurements (our two-markets strategy allows for the recovery of the "quantities-consistent" ω_{jt} and $\delta_{jt}/(\eta_D-1)$), we specify both unobservables as general Markov processes, and we estimate them in broad industries of differentiated products.

Foster, Haltiwanger and Syverson (2008) carry out "physical" TFP measurements with firm-level quantities of selected quasi-homogeneous products. They report a standard deviation of TFP of 0.26. We get standard deviations that tend to be slightly less than twice this value. This makes sense if we take into account the high heterogeneity of products included in our industries. Hsieh

[54] In the Online Appendix, Figure A1, we depict the marginal densities and their changes over three moments of time (1998-2000, 2001-2004 and 2005-2008).

and Klenow (2009) use the same data source for the years 1998-2005. They estimate a standard deviation of their "quantity" TFP of 0.95 and an interquartile range of 1.28. However, they do not try to separate the impact of demand heterogeneity. As they remark, their measure "is a composite of process effiency and idiosyncratic demand terms coming from quality and variety".[55]

Foster, Haltiwanger and Syverson (2016), working with a sample similar to the one of the 2008 article, use the within product-year residual of their demand estimate to asses demand heterogeneity. Their standard deviation is 1.47. We can transform this into a deviation that is roughly comparable with ours dividing by their highest demand elasticity estimate of 3: $\frac{1.47}{3-1} = 0.735$. In 2008, our average standard deviation for $\delta_{jt}/(\eta_D - 1)$ is 0.548. This difference may occur because, when we measure $\delta_{jt}/(\eta_D - 1)$, we have already subtracted a lot of variation in demand through our included observed shifters.[56]

In summary, the distributions turn out to be sensible and very informative. We get sensible measurements of persistent productivity and product advantages for broad product differentiated industries, which compare favorably with other measurements in more homogeneous settings. Our results underline two important things that were first shown by Foster, Haltiwanger and Syverson (2008) for their quasi-homogeneos goods sample. Productivity and product advantages show a significant heterogeneity, which tends to be greater for product advantages. In addition, we are able to characterize with detail the joint distribution of cost and product advantages as follows.[57]

Negative correlation, other correlations and implication.

The unobserved advantages ω_{jt} and $\delta_{jt}/(\eta_D - 1)$ have a strong negative correlation, as reported in column (10) and illustrated by the level sets of the joint densities in Figure 1. There is nothing in the model that implies such correlation, so this is a very important finding of our exercise. This says that firms that possess unobserved cost advantages tend to have weak unobserved product

[55] Their VA measurement for TFP in the presence of observed and unobserved demand shifters would be, in terms of our notation and with total revenue R_j

$$\exp(\omega_j) = cons \times \frac{R_j^{\frac{\eta}{\eta-1}} \exp(-\frac{z_j\alpha+\delta_j}{\eta-1})}{K_j^{\beta} L_j^{1-\beta}}.$$

Omitting the shifters on the right hand side implies to add their value to ω_j.

[56] In fact, the availability of new observed demand shifters has significantly narrowed the dispersion of the estimated unobservable advantages in different versions of this paper.

[57] The Foster, Haltiwanger and Syverson (2008 and 2016) identification strategy is based on the assumption that TFP and demand advantages are unocrrelated (TFP is used as instrument for price). In what follows, we show that demand advantages are highly correlated with TFP. As there is no reason to think that this correlation is absent in quasi-homogeneous good industries, this introduces an important doubt on the consistency of their specific estimates.

advantages, and firms that have unobserved product advantages tend to show less unobserved cost advantages.

Recall that $mc_{jt} = \overline{mc}_{jt} - \omega_{jt}$, therefore unobserved ω_{jt} is only a part of marginal cost. Hence, it is relevant to characterize the correlation between the observed and unobserved parts. Column (11) shows that $Corr(\overline{mc}_{jt}, \omega_{jt})$ is strongly positive. This simply indicates that productivity is positively associated to higher observed costs.[58] At the same time, it is also crucial to characterize the correlation of product advantages $\delta_{jt}/(\eta_D - 1)$ with total marginal cost mc_{jt}. Column (12) shows that $Corr(mc_{jt}, \delta_{jt}/(\eta_D - 1))$ is also strongly positive.[59] More product advantages are afforded with the trade off of higher costs, both observed and unobserved (less productivity).

All of this strongly suggests one of the main conclusions of this paper: many firms that have important cost advantages sell standard or even low quality products that are cheaper to produce. Additionally, firms which show important product advantages acquire them at the expense of clear disadvantages in the cost of their products, presumably due to the higher costs of producing the goods which embody these advantages (technology, design, quality,...).[60] Columns (13) and (14) show that despite the cost of advantages, real profits and estimated advantages (both taken separately and jointly) are mostly positively related. Therefore, firms have incentives to strive for both kind of advantages.

Consider a plane of $(\omega_{jt}, \delta_{jt}/(\eta_D - 1))$ pairs. If firms are equal for all observed factors (costs, shifters) and able to freely choose their combination of advantages from a balanced concave frontier given as endowment, it is easy to show that profit maximizing firms would show positively related amounts of ω_{jt} and $\delta_{jt}/(\eta_D - 1)$.[61] With $\eta_X > \eta_D$ firms would simply tend to prefer more cost than product advantages, but observed advantages would lie on a positively sloped line with slope less than unity. However, the observations of the real $(\omega_{jt}, \delta_{jt}/(\eta_D - 1))$ pairs tend to be spread along the negatively sloped isoprofit curves. Firms tend to reach similar levels of profitablity with very different combinations of advantages. This suggests two characteristics. First, advantages have an important uncertain component which escapes the direct control of firms. Second, even if firms are able to invest to impact the advantages and their relative importance, the abilities of firms to influence each kind of advantage (the real transfomation curves) are likely to be very heterogeneous

[58] For example, higher wages and greater cost of high quality materials.

[59] Roberts, Xu, Fan and Zhang (2016) find a positive correlation of 0.795 between their firm effect, formally comparable to our product advantages, and their marginal cost specification.

[60] For more specific insights on this relationship see subsection 7.3.

[61] See the explanation in section A6.3 of the Online Appendix.

(according to technological knowledge and past investments, say). The exogenous Markov processes that we have used in our modeling seem to be perfectly able to detect in practice this heterogeneity. However, this result points to a completely new aim of research that is particularly policy relevant: the determinants of the different advantages, including the possibilities, incentives and limits of firms' investment in the development of each advantage.

Changes in the means.

The change in the means of ω_{jt} and $\delta_{jt}/(\eta-1)$ over time provides an estimate of the growth of average productivity and product advantages.[62] Columns (1) and (2) of Table 6 show this growth. To report the growth of product advantages, we multiply δ_{jt} by the weighted average of the inverse elasticities. This facilitates the decomposition of the total growth of the product advantages into a gross component and an effect of entry (see below).

The increase in the means of ω_{jt} is huge and relatively even, ranging across industries from 24% to 60%. The growth of the means of the product advantages is, on the contrary, extremely heterogeneous. In two industries, *Electronics* and *Machinery,* product advantages grow at the same large rate as cost advantages. However, the other industries show very modest gains or none at all. Aditionally, in *Transport* the average product advantages decresase. Figure 2 shows the evolution of mean cost and product advantages over time.

As markets have been subject to significant net entry, it is possible that the greater demand available to firms has been counterbalanced by the increase in the number of firms competing for this demand. To check this conjecture, we estimate the net entry into markets and decompose the mean of $\delta_{jt}/(\eta-1)$ into two components: gross growth and the effect of entry[63]. Columns (3) and (4) show that entry tends to have a negative impact on the individual product advantages but it is small.

Aggregate changes.

To assess the sources of global changes, we weight productivity and product advantages by revenue shares. Then we aggregate and decompose the change of these aggregates over time in terms of the "dynamic Olley and Pakes decomposition" proposed by Melitz and Polanec (2015).[64] Columns (5) to (12) report the results. In this decomposition, entrants and exitors contribute to aggregate growth

[62] Subtracting the mean of ω_{jt} for 1998 from the mean for 2006 cancels the global mean that we have previously substracted to get ω_{jt}.

[63] See section A6.3 of the Online Appendix for details

[64] We could have presented the decompositions for each one of the markets (firms' sales vary) but patterns across markets turn out to be quite similar.

if their productivity or product advantages diverge from the ones of survivors.[65]

To simplify language, we are going to use the standard names of "survivors", "entrants" and exitors." However, it its important to recognize the specific content of these concepts. Our decomposition refers to the aggregates of firms active both in the domestic and export markets. Our sample has also additions over time. We consider "survivors" the firms that are in the sample for the starting year, 1998, and remain until 2008. In addition, we consider "survivors" firms that are present only in 2008 but were already born in 1996.[66,67] Entrants are the firms present in the sample in 2008 that are born during the period or existing non-exporting firms that start exporting. Exitors are all firms present in 1998 that either shut down or stop exporting durring the period. Newly created and shutting down firms dominate the two sets respectively.

The contributions of entry and exit to productivity growth tend to be unimportant. This means that the productivity of entrants and exitors compared with the productivity of the survivors shows small differences. At the end of the period, entrants turn out to be slightly less productive than survivors. This points to two main facts: entrants tend to enter with less productivity or cost advantages and the process to acquire them is slow. Exitors tend to be firms that show less productivity than the firms that survive. Therefore, their disappearance tends to contribute positively to the growth of aggegate productivity. However, none of the differences are dramatic.

The contributions of survivors to the growth of aggregate product advantages is negative in seven industries and virtually zero in another. The group of survivors loses product advantages over the period. The contribution of entrants is instead positive in seven industries. In five industries, the growth explained by entrants is larger than the negative growth induced by survivors. The disappearance of the exitors also makes some significant positive contributions. This implies that product advantages possesed by exitors were significantly lower than survivors' advantages.

Additionally, the growth of productivity and product advantages of survivors can be split into the shift of the mean of their distribution and the change of the covariance between the involved variable and the survivors' shares ("reallocation" component). Reallocation among survivors makes an important positive contribution to the growth of productivity and is responsible for the negative growth of product advantages.[68] At the end of the period, product advantages are displaced towards

[65] Details can be found in Section A6.3 of the Online Appendix along with additional results in Table 6b.

[66] Therefore, the number of survivors is different in 1998 and 2008.

[67] There is a minor ambiguity here: when a firm is an addition that is exporting we are not sure when it started to export. We take it notwithstanding as survivor.

[68] See Table 6b of the Online Appendix.

the smallest survivors.

Summarizing, there are no dramatic differences in productivity and productivity growth between survivors, entrants and exitors. However, product advantages are developed by the entrants, possibly at the expense of exitors and some survivors. Selection into the market is determined more by product advantages than productivity. Reallocation is also important: productivity becomes linked more to the largest market shares and product advantages to the smallest survivors.[69]

7. Privatization, trade specialization, technological investment.

In this section, we briefly sketch three examples of economic questions in which the distinction and quantification of cost and product advantages are relevant. In the first example, the distinction helps give a richer description of the process of privatization and shows that the state is particularly bad at the development of product advantages. In the second, the relative degree of cost advantages is closely associated with the degree of specialization in exports by Chinese firms. Specialization in exports is a puzzling trait of firms' heterogeneity in trade. In the third, the technological investment of firms and the use of a highly skilled workforce are shown to build product advantages for products which have higher cost of production than their substitutes. It is an important relationship that puts forward an idea very relevant for theoretical and empirical studies on reallocation: "cost differences" are not equivalent to "cost distortions".

7.1 Privatization and firms efficiency.

Colums (1) to (4) of Table 7 show the evolution of ω_{jt} and $\delta_{jt}/(\eta_D - 1)$ for the groups of "always private" firms and firms that experience a change of status (see Section 5 for the details of this taxonomy).[70] Columns (1) and (2) show that average productivity growth is systematically higher

[69] Brandt, Van Biesebroeck and Zhang (2012), with very different methods and starting from classical measurements of productivity, conclude with an optimistic assessment of the contribution of market entry to productivity and a pessimistic evaluation of the effiency-enhancing input reallocations among active firms. Notice that one possible reason of this mismatch with our conclusions is simply the lack of split of their measurement into productivity and product advantages. In our analysis, entrants do not have larger productivity but tend to show greater product advantages. Reallocation is made of productivity-enhancing shift of shares combined with relative loss of product advantages for the biggest firms. The aggregation of these trends may produce conclusions closer to theirs.

[70] We exclude the group "always state participated" firms because the state tends to retain only a small amount of very well performing firms. The behavior of productivity and product advantages of this group of firms is almost entirely determined by the selection operated over the years.

during the period for firms in the process of privatization. In contrast, columns (3) and (4) show that product advantages grow at the same pace for privatized and private firms in four industries, and evolve better for private firms in other five [71]

To further explore these numbers, for each industry we form a panel subsample of status-changing firms subject to the condition that firms start as state participated and end as private (although we admit back and forth changes in participation in between). For this subsample, we explain the evolution of measured productivity and product advantages by means of the following regression:

$$y_{jt} = \alpha_j + \alpha_t + \beta private_{jt} + v_{jt},$$

where $y_{jt} = \omega_{jt}$ or $\delta_{jt}/(\eta - 1)$, α_t represents a time effect common to all firms and the variable $private_{jt}$ is a step dummy that takes the value one the first time that a firm is observed to be without state participation and in all subsequent periods.[72] We estimate two versions of the model: replacing the fixed effects α_j by a constant and keeping the fixed effects as a form to allow for firm specific levels of productivity and product advantages.[73]

Columns (5) to (8) of Table 7 report the estimation of β, and show a set of stylized facts. Privatization during the period first affected the firms with relatively high productivity (column 5) and, in half of the industries, firms with relatively low product advantages (column 7). The high growth of productivity, which charaterizes the firms in transition, is weakly related to privatization itself (column 6). The growth of product advantages is not influenced by privatization (column 8).

To summarize, the privatization of firms somewhat helped increase productivity but did nothing to develop product advantages, despite this being a motive for privatization. It follows that the product advantages are contributed by the "always private firms," particularly the newly born private firms (recall the analysis of entry in Section 6). Firms coming from the intervention of the state seem more sluggish in the development of product advantages.

[71] Brandt, Van Biesebroeck and Zhang (2012) find a somewaht higher productivity growth of the firms in transition (page 349).

[72] Small variants in the construction of $private_{jt}$ do not significantly change the results.

[73] Using a constant, if privatization was earlier in more efficient firms we expect a positive bias in β. The reason is that we have more observations with one in the indicator coming from relatively efficient firms. Conversely, we expect a negative bias if privatization was earlier in the relatively less efficient firms. The introduction of fixed effects offers a different perspective: the estimate of β is exclusively based in comparing the residual efficency of each firm under privatization with its efficiency before privatization.

7.2 Specialization in exports.

In all industries, Chinese exporters show a bimodal distribution of export intensitiy with a pronounced "U-shaped" form. Columns (1) and (2) of Table 8 show two extreme intervals of the distribution (exporting less than 20% of sales, exporting more than 80%). These intervals concentrate between 50% to 60% of exporters. We show that the degree of specialization in exports is highly associated with the firm's relative intensity in cost advantages. This suggests that developing cost advantages and becoming a manufacturer highly specialized in exports is an optimal decision for many firms. Developing a full model for this choice is beyond of the scope of this paper. Our aim is to simply show that the distinction between cost and product advantages is a relevant component to explain the trade heterogeneity of firms.[74]

First, using our estimates of ω_{jt} and $\delta_{jt}/(\eta-1)$, we contruct the index of relative cost advantages ica_{jt}.[75] Then, calling export intensity ei_{jt}, years of experience in the export market $xper_{jt}$, the effect of other unobserved factors u_{jt}, and using a logit transformation, we estimate the OLS model

$$\ln \frac{ei_{jt}}{1-ei_{jt}} = \alpha_0 + \alpha_1 ica_{jt} + \alpha_2 ica_{jt}^2 + \alpha_3 ica_{jt}^3 + \alpha_4 xper_{jt} + u_{jt}.$$

This model could be further improved with interactions between the included variables and with the addition of other explanatory factors, but we feel that the basic form is sufficient for our current purposes. Columns (3) and (4) of Table 8 report the marginal effects of ica_{jt} and $xper_{jt}$ on ei_{jt}. Column (5) reports the R^2 of the regression. The R^2 is not high in *Textile* and *Metals*, but this very simple model explains one third of the variance for *Paper* and *Machinery* and more than half of the variance for the other industries.

The index of relative cost advantages has a uniform impact. Except for *Textile*, an additional percentage point of cost advantages implies an increase between 1.3 and 2.3 percentage points of export specialization. Interestingly, in most industries the youngest firms in the market are the most specialized.

[74] Lu (2010) stresses the "U-shaped" form of the distribution of export intensity of Chinese manufacturers. Puzzled by the appearance that exporters have lower productivity, this paper tries an explanation in which domestic markets select the most efficient firms. Our distinction between cost and demand advantages allows for another look at unexplained facts put forward in her paper: many heavy exporters are very cost efficent but deprived of the product advantages that characterize firms with greater domestic sales.

[75] To do so, we drop the values of ω_{jt} and $\delta_{jt}/(\eta-1)$ below the first decile of each distribution ($\omega_{0.1}$ and $\delta_{0.1}/(\eta-1)$) and we compute the index of cost advantages as $ica_{jt} = (\omega_{jt} - \omega_{0.1})/[(\omega_{jt} - \omega_{0.1}) + (\delta_{jt}/(\eta-1) - \delta_{0.1}/(\eta-1))]$. Of course, this a somewhat arbitrary construction that could be modified in many ways.

Columns (6) and (7) show the "U shaped" pattern of the distribution of predicted export intensities. The model fails to explain the observed pattern in *Textile* and *Metals*, however, it does a good job in the other industries. Figure 3 illustrates reproduces the complete distributions for all industries excluding *Textile* and *Metals*.

We conclude that intensity in cost advantages is strongly associated to the specialization of some firms in the export market. A natural way to interpret these results is to think of firms that choose to produce standard products (in technology, design, quality...) but are able to reach significantly lower costs in producing them.

7.3 Technological investments and workforce skills.

We use the data on R&D and workforce skills (see Section 5) to investigate the relationship between technological investments/quality of labor and the estimated cost and product advantages.

Columns (1) and (2) of Table 9 show that firms that perform R&D activities have, in 7 industries, some cost disadvatage (relatively lower ω_{jt}). Because ω_{jt} reflects the efficiency with which production inputs are used, this indicates that firms which undertake technological activities require a larger quantity of factors to produce a given quantity of their products. Conversely, columns (4) and (5) show that in the same 7 industries, firms with R&D expenditures have higher product advantages. This implies that the relatively higher cost of the products of the firms undertaking R&D results in superior characteristics that enhance demand. Columns (3) and (6) illustrate that the relationship described above generalizes to R&D intensity. All industries show a negative relationship between R&D intensity of firms and cost advantages.[76] However, R&D intensity has a positive correlation with product advantages for firms in 7 industries. The upper graphs of Figure 4 depict the nonparametric regressions of ω_{jt} and $\delta_{jt}/(\eta-1)$ on R&D intensity.

Columns (7) and (8) report the correlation between ω_{jt} and $\delta_{jt}/(\eta-1)$ and the quality of labor. Quality of labor input is positively associated with cost advantages in 8 industries. Notice the apparent paradox: firms with higher wages show more cost advantages. Firms experience greater reductions of their marginal cost because the impact of productivity associated with the quality of labor. On the other hand, the quality of labor is positively associated to product advantages in 7

[76] Note that our ω_{jt} purely reflects cost advantages of firms and is not comparable with other measurements of productivity. Some studies explicitly mix productive efficiency with demand shifters, other studies may implicitly mix them by an imperfect deflation of firm revenues. In fact, the simple addition in our sample of cost and demand effects tends to reverse the disadvantage of R&D firms.

industries. Both relationships taken together say that firms that have relatively high quality workers are firms with both relative unobserved cost and product advantages. The bottom graphs of Figure 7 depict the nonparametric regressions of ω_{jt} and $\delta_{jt}/(\eta-1)$ on the quality of labor.

To summarize, both R&D activities and the quality of the workforce push forward the product advantages of firms while at the same time are associated to increasing production costs. This points out at a missing piece of the current theoretical and empirical studies on reallocation of resources. These studies typically interpret all cost differences as coming from marginal productivities of inputs that are not equalized due to frictions or intervenned input prices.[77] Without denying such distortions, the above analysis shows a fundamental heterogeneity of firms that implies observed and unobserved costs with a counterpart in product advantages. This has two important implications. First, measurements of cost distortions that ignore the presence of product advantages based on higher production costs may overstate distortions, particularly in developing countries where quality heterogeneity across firms may be larger. Second, the right allocation policy differs from the policy against pure distorsions. Both profit and total welfare objectives should explicitly consider the positive effects of allocating resources to activities with higher costs but important positive demand effects (see subsection 6.3).

8. Concluding remarks.

Using a sample of Chinese manufacturing firms, which operate domestically and in the export market, we estimate the joint distribution of cost advantages (unobserved productivity) and product advantages (unobserved demand heterogeneity), and how it changed from 1998 to 2008. Using the multimarket character of the firms, we disentangle cost and product advantages, without observing output prices, and estimate the unobservables simultaneously as non funcionally-dependent and freely correlated Markov processes. Particularly significant is that product advantages turn out to be negatively correlated with cost advantages (positively correlated with marginal cost). Using the distribution, we have characterized the growth of Chinese manufacturing and described its

[77] Consider, for example, the starting point of the influential paper by Hsieh and Klenow (2009): the value of marginal productivity of capital and labor (they consider a VA production function) diverges exclusively by the idiosyncratic price distortions experienced by firms. Other papers have stressed the need for considering the shadow prices of the dynamic inputs (Bartelsman, Haltiwanger and Scarpetta, 2013; Asker, Collard-Wexler and De Loecker, 2014) or financial restrictions (Midrigan and Xu, 2014). However, there are no papers integrating into the framework the heterogeneity of input prices because of product differentiation.

weaknesses. Chinese firms relied heavily on cost competition to grow and much more modestly on product advantages.

Let us mention a few aspects that we leave for further research. We model δ_{jt} and ω_{jt} as exogenous Markov processes. However, the natural extension would be to consider how product and cost advantages shift across firms and over time with the technological and human capital investments of firms.[78] Another important extension would be to apply the results to analyze the distribution of the non-exporters that have been excluded from our sample.

There are two more ways in which the results could be extended. The first is the assumption that the product advantages have a similar impact in both the domestic and export markets. We check that inferences are relatively immune to this assumption. However, it is possible to further relax the assumption at the cost of giving more structure to the differences between markets, and this seems something worth trying. Another is the assumption of common industry elasticities. We show that to relax this assumption in our framework is not difficult. A systematic exploration of elasticity variation across firms in the industry would provide insights on markup heterogeneity. Such investigation should address the difficult question of the separate identification of product advantages and elasticity variation.

[78] That is, to consider "endogenous" Markov processes as in Doraszelski and Jaumandreu (2013, 2016).

Appendix A: Proof of the proposition.

Proof : Let's consider the matrix

$$\begin{bmatrix} \eta_{Xjt}-1 & 1 \\ \eta_{Djt}-1 & \lambda_{jt} \end{bmatrix}.$$

By assumption all principal minors of this matrix do not vanish. Multiplying the first column by $\left|\frac{1}{MC_{jt}}\frac{\partial MC_{jt}}{\partial \omega_{jt}}\right|$ and the second column by $\frac{1}{R^X_{jt}}\frac{\partial R^X}{\partial \delta_{jt}}$ we get the matrix of semielasticities

$$\begin{bmatrix} \frac{1}{R^X_{jt}}\frac{\partial R^X}{\partial \omega_{jt}} & \frac{1}{R^X_{jt}}\frac{\partial R^X}{\partial \delta_{jt}} \\ \frac{1}{R^D_{jt}}\frac{\partial R^D}{\partial \omega_{jt}} & \frac{1}{R^D_{jt}}\frac{\partial R^D}{\partial \delta_{jt}} \end{bmatrix}.$$

We have multiplied each column by a positive value, and therefore we preserve the property of non vanishing principal minors. Now multiply the first row by R^X_{jt} and the second by R^D_{jt}. For the same reason as before, we get a matrix of derivatives with non vanishing principal minors

$$\begin{bmatrix} \frac{\partial R^X}{\partial \omega_{jt}} & \frac{\partial R^X}{\partial \delta_{jt}} \\ \frac{\partial R^D}{\partial \omega_{jt}} & \frac{\partial R^D}{\partial \delta_{jt}} \end{bmatrix}.$$

Writing equations (5) in the text as the system of equations $R^X(\cdot) - R^X_{jt} = 0$ and $R^D(\cdot) - R^D_{jt} = 0$, we observe that the above matrix is the Jacobian of the system. A system is invertible if no principal minor of its Jacobian vanishes (Theorem 7 of Gale and Nikaido, 1965)■

Appendix B: A model with different marginal costs.

Let's assume that

$$\begin{aligned} MC^D_{jt} &= \frac{1}{v}\exp(\beta_C)h^D_{jt}(Q^D_{jt})^{(1-\nu)/\nu}\exp(-\omega_{jt}/\nu), \\ MC^X_{jt} &= \frac{1}{v}\exp(\beta_C)h^X_{jt}(Q^X_{jt})^{(1-\nu)/\nu}\exp(-\omega_{jt}/\nu), \end{aligned}$$

where $\beta_C = \frac{\beta_0}{\nu}+\ln v+\frac{1}{\nu}\ln \beta_L^{-\beta_L}\beta_M^{-\beta_M}$. The terms h^D_{jt} and h^X_{jt} are $h^D_{jt} = (K^D_{jt})^{-\beta_K/\nu}(W^D_{jt})^{\beta_L/\nu}(P^D_{Mjt})^{\beta_M/\nu}$ and $h^X_{jt} = (K^X_{jt})^{-\beta_K/\nu}(W^X_{jt})^{\beta_L/\nu}(P^X_{Mjt})^{\beta_M/\nu}$. Therefore we are admitting that the capital used and/or the prices of the inputs are different between the good produced for export and domestic sale.

Using optimal pricing expressions, marginal costs can be rewritten in terms of revenue as

$$\begin{aligned} MC_{jt}^D &= \left(\frac{\eta_D}{\eta_D - 1}\right)^{-(1-\nu)} \left(\frac{h_{jt}^D}{\nu}\right)^{\nu} (R_{jt}^D)^{1-\nu} \exp(-\omega_{jt}), \\ MC_{jt}^X &= \left(\frac{\eta_X}{\eta_X - 1}\right)^{-(1-\nu)} \left(\frac{h_{jt}^X}{\nu}\right)^{\nu} (R_{jt}^X)^{1-\nu} \exp(-\omega_{jt}). \end{aligned}$$

Rewrite optimal prices in terms of these expressions. Take equations (9) and plug in the price expressions. Rearranging and taking logs we get

$$\begin{aligned} r_{jt}^X &= \phi^X - \frac{\nu(\eta_X - 1)}{\zeta_X} \ln h_{jt}^X + \frac{1}{\zeta_X}(z_{jt}^X \alpha_X + (\eta_X - 1)\omega_{jt} + \delta_{jt}), \\ r_{jt}^D &= \phi^D - \frac{\nu(\eta_D - 1)}{\zeta_D} \ln h_{jt}^D + \frac{1}{\zeta_D}(z_{jt}^D \alpha_X + (\eta_D - 1)\omega_{jt} + \delta_{jt}), \end{aligned}$$

where $\phi^X = \frac{1}{\zeta_X} \ln \alpha_0^X - \frac{\nu(\eta_X - 1)}{\zeta_X} \ln \frac{1}{\nu} \frac{\eta_X}{\eta_X - 1}$ and $\phi^D = \frac{1}{\zeta_D} \ln \alpha_0^D - \frac{\nu(\eta_D - 1)}{\zeta_D} \ln \frac{1}{\nu} \frac{\eta_D}{\eta_D - 1}$ and where $\zeta_X = 1 + (1-\nu)(\eta_X - 1)$ and $\zeta_D = 1 + (1-\nu)(\eta_D - 1)$.

Appendix C: Correcting the standard errors for two-stage estimation.

Our NLS estimator solves the problem

$$\min_{\gamma} \frac{1}{N} \sum_j [y_j - m(w_j, \gamma)]^2,$$

which has first order condition

$$\sum_j \bigtriangledown_{\gamma} m(w_j, \widehat{\gamma})' [y_j - m(w_j, \widehat{\gamma})] = 0.$$

To estimate the parameters θ of the system we use the GMM estimator that solves

$$\min_{\theta} [\frac{1}{N} \sum_j g(w_j, \theta, \widehat{\gamma})]' \widehat{W} [\frac{1}{N} \sum g(w_j, \theta, \widehat{\gamma})].$$

Because we expect $E[\bigtriangledown_{\gamma} g(w_j, \theta_0, \gamma_0)] \neq 0$ we have to correct the standard errors of $\widehat{\theta}$ to ensure their consistency (Newey and McFadden, 1994).

The first order condition for $\widehat{\theta}$ is

$$[\sum_j \bigtriangledown_{\theta} g(w_j, \widehat{\theta}, \widehat{\gamma})]' \widehat{W} [\sum_j g(w_j, \widehat{\theta}, \widehat{\gamma})] = 0.$$

Expanding $\sum_j g(w_j, \widehat{\theta}, \widehat{\gamma})$ around θ_0 and substituting it back into the first-order condition we have

$$0 = [\sum_j \bigtriangledown_{\theta} g(w_j, \widehat{\theta}, \widehat{\gamma})]' \widehat{W} [\sum_j g(w_j, \theta_0, \widehat{\gamma})] + [\sum_j \bigtriangledown_{\theta} g(w_j, \widehat{\theta}, \widehat{\gamma})]' \widehat{W} [\sum_j \bigtriangledown_{\theta} g(w_j, \overline{\theta}, \widehat{\gamma})](\widehat{\theta} - \theta_0),$$

where $\overline{\theta}$ is the value that makes the expansion exact according to the mean value theorem. Dividing the sums of $g(\cdot)$ and its derivatives by N, replacing the result in the case of the derivatives by the probability limit $G = E[\bigtriangledown_\theta g(w_i, \theta_0, \gamma_0)]$, replacing $\widehat{W}$ by its probability limit W, and solving for $\sqrt{N}(\widehat{\theta} - \theta_0)$ yields

$$\sqrt{N}(\widehat{\theta} - \theta_0) = -(G'WG)^{-1}G'W \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \widehat{\gamma}) + o_p(1).$$

This expression allows us to deduce the variance of $\widehat{\theta}$.

Given the presence of $\widehat{\gamma}$, we have to expand $\sum_j g(w_j, \theta_0, \widehat{\gamma})$ around γ_0,

$$\begin{aligned} \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \widehat{\gamma}) &= \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \gamma_0) + [\frac{1}{N} \sum_j \bigtriangledown_\gamma g(w_j, \theta_0, \overline{\gamma})] \sqrt{N}(\widehat{\gamma} - \gamma_0) \\ &= \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \gamma_0) + G_\gamma \sqrt{N}(\widehat{\gamma} - \gamma_0) + o_p(1), \end{aligned}$$

where $G_\gamma = E[\bigtriangledown_\gamma g(w_j, \theta_0, \gamma_0)]$. Similarly to $\widehat{\theta}$, an expansion and subsequent rearrangement of the first order condition for $\widehat{\gamma}$ gives the expression for $\sqrt{N}(\widehat{\gamma} - \gamma_0)$

$$\sqrt{N}(\widehat{\gamma} - \gamma_0) = E[\bigtriangledown_\gamma m(w_j, \gamma_0)' \bigtriangledown_\gamma m(w_j, \gamma_0)]^{-1} \frac{1}{\sqrt{N}} \sum_j \bigtriangledown_\gamma m(w_j, \gamma_0)'(y_j - m(w_j, \gamma_0)).$$

Plugging this representation into the expansion of $\frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \widehat{\gamma})$, we have

$$\begin{aligned} & \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \widehat{\gamma}) \\ = & \frac{1}{\sqrt{N}} \sum_j g(w_j, \theta_0, \gamma_0) \\ & + G_\gamma E[\bigtriangledown_\gamma m(w_j, \gamma_0)' \bigtriangledown_\gamma m(w_j, \gamma_0)]^{-1} \frac{1}{\sqrt{N}} \sum_j \bigtriangledown_\gamma m(w_j, \gamma_0)'(y_j - m(w_j, \gamma_0)) + o_p(1). \end{aligned}$$

Defining

$$\widetilde{g}(w_j, \theta_0, \gamma_0) = g(w_j, \theta_0, \gamma_0) + G_\gamma E[\bigtriangledown_\gamma m(w_j, \gamma_0)' \bigtriangledown_\gamma m(w_j, \gamma_0)]^{-1} \bigtriangledown_\gamma m(w_j, \gamma_0)'(y_j - m(w_j, \gamma_0)),$$

the new expression to derive the variance of $\widehat{\theta}$ turns out to be

$$\sqrt{N}(\widehat{\theta} - \theta_0) = -(G'WG)^{-1}G'W \frac{1}{\sqrt{N}} \sum_j \widetilde{g}(w_j, \theta_0, \gamma_0) + o_p(1).$$

Defining

$$D = E[\widetilde{g}(w_j, \theta_0, \gamma_0)\widetilde{g}(w_j, \theta_0, \gamma_0)'],$$

we have

$$Avar(\widehat{\theta}) = \frac{(G'WG)^{-1}G'WDWG(G'WG)^{-1}}{N}.$$

The asymptotic variance can be estimated by replacing the probability limits with estimates and matrix D using an estimate based on $g(w_j, \widehat{\theta}, \widehat{\gamma})$, $\widehat{G}\gamma$, $\bigtriangledown_\gamma m(w_j, \widehat{\gamma})$ and $y_j - m(w_j, \widehat{\gamma})$.

Appendix D: Variables

Middle-West location. Dummy that takes the value one for firms located in the *Middle* and *Western* parts of China.

Year of birth. Year the firm was born.

Age. Current year minus the year in which the firm was born.

Entry. We consider that the firm is an "economic" entrant if when it is included in the sample for the first time it was born that year or one of the two previous years.

Exit. We consider all dissapearances from the sample as "shutdowns".

Experience. Current year minus the first year that the exports of the firm are non-zero (years after "entering" the export market).

Subsidy. State aid received by the firm as proportion of sales.

State participation. We compute the share of the state in financial capital as the sum of the reported state and collective capital over total financial capital. The "always state" are firms that, while in sample, are state participated. The "always private" are the firms that, while in sample, never have state participation. The remaining category are the firms that go over a change.

Foreign participation. The amount of capital owned by foreign firms over total financial capital.

Revenue. Revenue after taxes, at current prices, as reported by the firm.

Exports. Value of industrial export sales after taxes, at current prices, as reported by the firm.

Export intensity. Exports divided by revenue.

Price of output. Output price index of the two-digit industry the firm belongs to, taken from China Statistical Yearbook (CSY).

Capital. Real stock constructed as follows. Firms report the value of their capital stock at original purchase prices and their capital stock at original purchase prices less accumulated depreciation. From these nominal values, we estimate a sequence of real investments and real capital stock at the starting year. Capital is then constructed by applying the perpetual inventory method assuming a yearly depreciation of 9%. For firms founded before 1998, we apply a method similar to Brandt, Van Biesebroeck, and Zhang (2012). We first estimate a yearly nominal rate of investment in fixed assets at the two-digit industry level using 1998-2003 firms' data. We assume that capital accumulates constantly at this rate from when the firm was created. We then estimate the capital stock at birth, deflate it, and compute the real stock in the first year. The investment deflator is taken from Brandt, Rawski and Sutton (2008), updated using the Fixed Asset Investment price index from CSY.

Cost of materials. Estimate of the intermediate consumption in production as follows. The survey definition of intermediate inputs includes direct materials, intermediate inputs used in production, intermediate input in management, intermediate input in business operations (sales cost) and financial expenses. As we want to use a measure of variable cost, the inclusion of general management expenses, sales cost and financial costs is problematic. Alternatively we started by the manufacturing costs (which include materials), labor cost and depreciation of capital during the process of production. From these manufacturing costs, we have then deduced the imputted wage bill and imputted depreciation of capital. From 2004 to 2007, we can do this using the detailed information on the structure of intermediate inputs. For the rest of years we assume the same proportions.

Price of materials. Estimate of a price index for the intermediate consumption of the industry the firm belongs to. As Brandt, Van Biesebroeck and Zhang (2012) we did compute a weighted average of the output prices for the industries from which the industry of interest purchases its inputs. For the weights, we use the Input-Output table from 2002, that includes 42 sectors. The two-digit manufacturing price indices are from CSY. The prices of agriculture, construction, transportation, retail, wholesale and some service sectors are calculated by comparing GDP at current prices and constant prices of the Collection of Statistical Material from 1949 to 2009.

Materials. Cost of materials divided by the price of materials.

Wage bill. We add up wages, unemployment insurance premium, pension and medical insurance premium, housing mutual fund and total welfare fees. It should be taken into account that firms only began to report retirement and health insurance in 2003, and housing benefits in 2004.

Employment. Total number of employees, which includes all the full-time production and nonproduction workers, as reported by the firm. It excludes part-time and casual workers.

Wage. Wage bill divided by employment.

Variable Cost. Sum of the cost of materials and wage bill.

Revenue over Variable Cost. Revenue divided by variable cost.

Sales effort. (Log of) All expenditures related to sales (e.g salesforce wages and advertising expenditures) as reported by the firm.

Sales effort intensity. Sales effort divided by revenue.

R&D. Expenditures in R&D activities as reported by the firm. There is only data for the year 2001 and the period 2005-2007.

R&D intensity. R&D expenditure over revenue.

Workforce s*kills.* Ratio of the firm wage to the average of wages of all the firms in the industry.

Appendix E: Industry correspondence and number of subindustries

Industry	Two-digit industries	Four-digit ind. (No.)
1. Food,	13. Agricultural and by-product proc.	49
drink and tobacco.	14. Food manufacturing	
	15. Beverage manufacturing	
	16. Tobacco products	
2. Textile,	17. Textile	33
leather and shoes.	18. Apparel, shoes, and hat manuf.	
	19. Leather, fur, and coat prod. manuf.	
3. Timber	20. Wood proc., and other wood prod.	13
and furniture.	21. Furniture manufacturing	
4. Paper and	22. Paper making and paper products	10
printing products.	23. Printing and recording media reprod.	
5. Chemical products.	26. Chemical materials and products	61
	27. Pharmaceutical	
	28. Chemical fiber	
	29. Rubber products	
	30. Plastic products	
6. Non-metallic minerals.	31. Non-metallic minerals products	30
7. Metals	32. Ferr. metal smelting and rolling proc.	37
and metal products.	33. Non-ferrous metal rolling processing	
	34. Metal products	
8. Machinery.	35. General machinery manufacturing	73
	36. Special machinery manufacturing	
9. Transport equipment.	37. Transportation equipment manuf.	23
10. Electronics.	39. Electronic machinery and equipment	63
	40. Elec. commun. equip. and computer	
	41. Instr., meter, stat. and office machine	

References

Ackerberg, D., L. Benkard, S. Berry and A. Pakes (2007), Econometric Tools for analyzing Market Outcomes, in Heckman, J. and E. Leamer eds., *Handbook of Econometrics,* vol. 6A, 4171-4276.

Ackerberg, D., K. Caves and G. Frazer (2015), "Structural Identification of Production Functions," *Econometrica*, 6, 2411-2451.

Aghion, P., J. Cai, M. Dewatripont, L. Du, A. Harrison and P. Legros (2015), "Industrial Policy and Competition," *American Economic Journal: Macroeconomics,* 7, 1-32.

Amemiya, T. (1974), "The Nonlinear Two-stage Least-squares Estimator," *Journal of Economic Literature*, 2, 105-110.

Asker, J., A. Collard-Wexler and J. DeLoecker (2014), "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 22, 1013-1063.

Aw, B.Y., and Y. Lee (2014), "A Model of Demand, Productivity and Foreign Location Decision among Taiwanese Firms," *Journal of International Economics*, 92, 304-316.

Aw, B.Y., M. Roberts and D.Y. Xu (2011), "R&D Investment, Exporting and Productivity dynamics," *American Economic Review*, 101, 1312-1344.

Bartelsman, E. and M. Doms (2000), "Understanding Productivity: Lessons from Longitudinal Data," *Journal of Economic Literature,* 38, 569-594.

Bartelsman, E., J. Haltiwanger and S. Scarpetta (2013), "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review,* 103, 305-334.

Bernstein, J. and P. Mohnen (1991), "Price-Cost Margins, Exports and Productivity Growth: with an Application to Canadian Industries," *Canadian Journal of Economics*, 24, 638-659.

Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242-262.

Berry, S., J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.

Bilir, L.K. and E. Morales (2016), "The Impact of Innovation in the Multinational Firm," mimeo, University of Wisconsin-Madison and Princeton University.

Boler, E., A. Moxnes and K. Ullveit-Moe (2015), "R&D, International Sourcing, and the Joint Impact on Firm Performance," *American Economic Review,* 105, 3704-3739.

Brandt, L., T.G. Rawski and J. Sutton (2008), "China's Industrial Development," in L. Brandt and T. Rawski (eds.), *China's Great Economic Transformation*, Cambridge U.Press, 569-632.

Brandt, L., J. Van Biesebroeck and Y. Zhang (2012), "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 97, 339-351.

Brandt, L., J. Van Biesebroeck, L. Wang and Y. Zhang (2012), "WTO Accession and Performance of Chinese Manufacturing Firms," Discussion Paper 9166, CEPR.

Bughin, J. (1996), "Capacity Constraints and Export Performance: Theory and Evidence from Belgian Manufacturing," *Journal of Industrial Economics*, 44, 187-204.

Das, S., M. Roberts, and J.Tybout (2007), "Market Entry Costs, Producer Heterogeneity, and Export Dynamics," *Econometrica*, 75, 837-873.

De Loecker, J. (2011), "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79, 1407-1451.

De Loecker, J., P. Goldberg, A. Khandelval and N. Pavnik (2016), "Prices, Markups and Trade Reform," *Econometrica,* 84, 445-510.

De Loecker, J. and F. Warzynski (2012), "Markups and Firm-level Export Status," *American Economic Review*, 102, 2437-2471.

Doraszelski, U. and J Jaumandreu (2013), "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies,* 80, 1338-1383.

Doraszelski, U. and J Jaumandreu (2016), "Measuring the Bias of Technological Change," mimeo, University of Pennsylvania and Boston University.

Dorfman, R. and P.Steiner (1954), "Optimal Advertisisng and Optimal Quality," *American Economic Review*, 44, 826-836.

Eslava, M., J. Haltiwanger, A. Kugler and M. Kugler (2004), "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia," *Journal of Development Economics,* 75, 333-371.

Foster, L., J. Haltiwanger and C. Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98, 394-495.

Foster, L., J. Haltiwanger and C. Syverson (2016), "The Slow Growth of New Plants: Learning About Demand?," *Economica,* 83, 91-129.

Gale, D. and H. Nikaido (1965), "The Jacobian Matrix and Global Univalence of Mappings," *Matematische Annalen*, 159, 81-93.

Gandhi, A., S. Navarro and D. Rivers (2013), "On the Identification of Production Functions: How Heterogeneous is Productivity?," mimeo, University of Wisconsin-Madison.

Gervais, A. (2015), "Product Quality and firm Heterogeneity in International Trade," *Canadian Journal of Economics,* 48,1152-1174.

Gordinchenko, Y. (2012), "Using Firm Optimization to Evaluate and Estimate Productivity and Returns to Scale," mimeo, University of California Berkeley.

Hall, R.E. (1990), "Invariance Properties of Solow's Productivity Measure." In

Growth/Prod./Unemp.: Essays to Celebrate Bob Solow's Birthday, Diamond, P. ed., MIT Press.

Hottman, C., S. Redding and D. Weinstein (2014), "What is *Firm Heterogeneity* in Trade Models? The Role of Quality, Scope, Markups and Cost," NBER Working Paper 20436.

Hsieh, C.T. and P. Klenow (2009), "Misallocation and manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403-1448.

Jaumandreu, J. and J. Mairesse (2010), "Innovation and Welfare: Results from Joint Estimation of Production and Demand Functions," NBER Working Paper 16221.

Jaumandreu, J. and J. Mairesse (2016), "Disentangling the Effects of Process and Product Innovation on Cost and Demand," *Economics of Innovation and New Technology.*

Klette, T. and Z. Griliches (1996), " The Inconsistency of the Common Scale Estimators when Output Prices are Unobserved and Endogenous," *Journal of App.Econometrics*, 11, 343-361.

Lee, R. (2013), "Vertical Integration and Exclusivity in Platforms and Two-sided Markets," *American Economic Review,* 103, 2960-3000.

Levinsohn, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317-341.

Lu, D. (2010), "Exceptional exporter performance? Evidence from Chinese manufacturing firms," mimeo, Rochester University.

Lu, Y. and L. Yu (2015), "Trade Liberalization and Markup Dispersion: Evidence from China's WTO Accession," *American Economic Journal: Applied Economics*, 7, 221-253.

Ma, Y., H. Tang and Y. Zhang (2014), "Factor Intensity, Product Switching, and Productivity: Evidence from Chinese Exporters," *Journal of International Economics,* 92, 349-362.

Manova, K. and Z. Zhang (2012), "Export Prices Across Firms and Destinations," *The Quarterly Journal of Economics,* 1-59.

Matzkin, R. (2007), "Nonparametric Identification," in Heckman, J.J. and E.E. Leamer (eds.), *Handbook of Econometrics*, vol. 6, Elsevier, 5307-68.

Matzkin, R. (2013), "Nonparametric Identification in Structural Economic Models," *Annual Reviews of Economics*, 5, 457-486.

Mayer, T., M. Melitz and G. Ottaviano (2014), "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 104, 495-536.

Mayer, T., M. Melitz and G. Ottaviano (2016), "Product Mix and Firm Productivity Responses to Trade Competition," CEPR Working Paper.

Melitz, M. and S. Redding (2014), "Heterogeneous Firms and Trade," in Helpman, E., K. Rogoff and G. Gopinath (eds.), *Handbook of International Economics,* Volume 4, 1-54.

Melitz, M. and S. Polanec (2015), "Dynamic Olley-Pakes decomposition with Entry and Exit," *Rand Journal of Economics,* 46, 362-375.

Midrigan, V. and D.Y. Xu (2014), "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104, 422-458.

Moreno, L. and Rodriguez, D. (2004), "Domestic and Foreign Price/Marginal-cost Margins: An Application to Spanish Manufacturing Firms," *Review of International Economics*, 12, 60-80.

Newey, W. (1990), "Efficient Instrumental Variables Estimation of Nonlinear Models," *Econometrica*, 58, 809-837.

Newey, W. (1993), "Efficient Estimation of Models with Conditional Moment restrictions," in Maddala, G., C. Rao and H. Vinod (eds.), *Handbook of Statistics*, Elsevier, vol. 11, 419-454.

Olley, S. and A. Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263-1298.

Peters, B., M. Roberts, V.A. Vuong and H. Fryges (2016), "Estimating Dynamic R&D Choice: An Analysis of Cost and Long-Run Benefits," *Rand journal of Economics.*

Pozzi, A. and F. Schivardi (2016), "Demand or Productivity: What Determines Firm Growth?," *Rand Journal of Economics*, .

Roberts, M., D.Y. Xu, X. Fan and S. Zhang (2016), "The Role of Firm Factors in Demand, Cost, and Export Market Selection for Chinese Footwear Producers," mimeo.

Robinson, P. (1988), "Root-n-consistent Semiparametric Regression," *Econometrica*, 56, 931-954.

Song, Z., K. Storesletten and F. Zilibotti (2011), "Growing like China," *American Economic Review, 101,* 196-233.

Sutton, J. (2001), "Rich Trades, Scarce Capabilities: Industrial Development Re-visited," Keynes Lecture, Proceedings of the British Academy.

Sutton, J. (2007), " Quality, Trade and the Moving Window: The globalization Process," *The Economic Journal*, 117, 469-498.

Sweeting, A. (2013), "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *Econometrica,* 81, 1763-1803.

Syverson, C. (2011), "What Determines Productivity?," *Journal of Econ. Literature*, 49, 326-365.

Tirole, J. (1989), *The Theory of Industrial Organization,* MIT Press.

Wooldridge, J. (2010), *Econometric Analysis of Cross Section and Panel Data,* MIT Press.

Table 1: Descriptive statistics.[a]

	Number of firms	Number of obs.	Prop. of ind. sales in 2008	Export intensity[b]	Middle-West location prop.[b]	Age[b,c]	Exper.[b,c]	State participation: proportion of firms — 1998: Always state part.	1998: Always private	2008: Always state part.	2008: Always private
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Food, drink and tobacco	5,548	21,048	0.263	0.444	0.221	10.4	3.8	0.426	0.220	0.045	0.795
2. Textile,leather and shoes	18,108	68,191	0.379	0.603	0.089	9.6	3.9	0.331	0.318	0.018	0.883
3. Timber and furniture	2,747	9,343	0.281	0.569	0.181	8.0	3.5	0.356	0.365	0.014	0.921
4. Paper and printing products	1,791	6,797	0.248	0.350	0.105	10.5	3.7	0.379	0.330	0.033	0.849
5. Chemical products	11,184	47,318	0.380	0.382	0.136	11.2	4.0	0.326	0.296	0.037	0.808
6. Non-metallic minerals	3,652	13,481	0.205	0.400	0.244	11.0	3.7	0.333	0.263	0.03	0.825
7. Metals and metal products	6,499	25,521	0.426	0.484	0.125	11.0	3.9	0.351	0.310	0.026	0.861
8. Machinery	9,008	36,944	0.445	0.361	0.118	13.3	3.8	0.375	0.257	0.029	0.843
9. Transport equipment	3,308	13,638	0.544	0.364	0176	12.0	3.8	0.362	0.271	0.035	0.848
10. Electronics	11,691	48,367	0.680	0.482	0.063	9.5	4.0	0.264	0.368	0.020	0.860

[a] Years 1998-2008.
[b] Average 1998-2008.
[c] Number of years.

Table 2: Descriptive statistics (cont'd).[a]

	Subsidy[b]		Foreign partic.[b]				Sales effort[b]		R&D[b]		Workforce
	Prop. of obs.	Mean subs.[c]	Prop. of obs.	Mean partic.[c]	Employ-ment[b]	$\ln \frac{R}{C}^b$	Prop. of obs.	Mean intensity[c]	Prop. of obs.	Mean intensity[c]	skills[b] (s. d.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Food, drink and tobacco	0.192	0.011	0.287	0.691	383	0.186	0.953	0.052	0.151	0.007	-0.068 (0.668)
2. Textile,leather and shoes	0.190	0.006	0.229	0.676	460	0.122	0.893	0.025	0.085	0.008	0.006 (0.482)
3. Timber and furniture	0.186	0.008	0.219	0.700	287	0.150	0.954	0.041	0.096	0.007	-0.020 (0.523)
4. Paper and printing products	0.153	0.009	0.234	0.800	368	0.173	0.924	0.035	0.105	0.009	0.004 (0.626)
5. Chemical products	0.222	0.009	0.258	0.762	404	0.196	0.955	0.046	0.243	0.013	-0.015 (0.630)
6. Non-metallic minerals	0.179	0.013	0.245	0.695	453	0.215	0.960	0.058	0.197	0.011	0.069 (0.619)
7. Metals and metal products	0.206	0.006	0.226	0.755	626	0.147	0.929	0.031	0.150	0.010	-0.101 (0.575)
8. Machinery	0.243	0.010	0.267	0.758	448	0.193	0.946	0.041	0.270	0.019	-0.028 (0.592)
9. Transport equipment	0.289	0.010	0.294	0.746	764	0.178	0.951	0.034	0.321	0.016	-0.267 (0.590)
10. Electronics	0.226	0.008	0.347	0.810	599	0.173	0.945	0.037	0.303	0.020	-0.0150 (0.608)

[a] Years 1998-2008.
[b] Average(s) 1998-2008.
[c] Mean of non-zero values.

Table 3: Estimating the a and b functions.[a] Dependent variable: $\ln \frac{R_{jt}}{C_{jt}}$.

	NLS			With sample selection correction			Dom. margin	Export margin
	a (s. e.)[c]	b (s. e.)[c]	Standard error of equ.	a (s. e.)[c]	b (s. e.)[c]	Coeff. on Mills r. (s. e.)[c]	$exp(a)-1^b$ (s. e.)[d]	$\frac{\exp(a)}{1+b}-1^b$ (s. e.)[d]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	0.234	0.108	0.174	0.255	0.108	-0.018	0.290	0.164
	(0.005)	(0.008)		(0.008)	(0.003)	(0.003)	(0.010)	(0.007)
2. Textile, leather and shoes	0.144	0.035	0.108	0.147	0.036	-0.005	0.159	0.119
	(0.002)	(0.002)		(0.002)	(0.000)	(0.001)	(0.002)	(0.002)
3. Timber and furniture	0.175	0.044	0.126	0.189	0.045	-0.014	0.207	0.155
	(0.004)	(0.006)		(0.006)	(0.001)	(0.003)	(0.007)	(0.006)
4.Paper and printing products	0.193	0.055	0.141	0.215	0.055	-0.018	0.240	0.175
	(0.005)	(0.008)		(0.007)	(0.002)	(0.003)	(0.009)	(0.007)
5. Chemical products	0.232	0.089	0.176	0.253	0.089	-0.022	0.288	0.183
	(0.003)	(0.005)		(0.004)	(0.002)	(0.002)	(0.006)	(0.004)
6. Non-metallic minerals	0.229	0.034	0.158	0.264	0.036	-0.028	0.302	0.256
	(0.004)	(0.007)		(0.005)	(0.002)	(0.003)	(0.007)	(0.006)
7. Metals and metal products	0.163	0.033	0.136	0.188	0.031	-0.026	0.207	0.171
	(0.003)	(0.004)		(0.004)	(0.001)	(0.002)	(0.004)	(0.004)
8. Machinery	0.218	0.068	0.147	0.243	0.065	-0.028	0.275	0.198
	(0.002)	(0.004)		(0.003)	(0.001)	(0.002)	(0.004)	(0.003)
9.Transport equipment	0.198	0.053	0.128	0.217	0.051	-0.020	0.242	0.182
	(0.003)	(0.006)		(0.004)	(0.001)	(0.003)	(0.005)	(0.004)
10. Electronics	0.214	0.083	0.144	0.220	0.084	-0.008	0.246	0.149
	(0.002)	(0.004)		(0.003)	(0.001)	(0.002)	(0.003)	(0.002)

[a] $a = \ln \frac{1}{\nu}\frac{\eta_D}{\eta_D - 1}$, $b = \frac{\frac{\eta_D}{\eta_D-1}}{\frac{\eta_X}{\eta_X-1}} - 1$.

[b] $exp(a) - 1 = \frac{1}{\nu}\frac{\eta_D}{\eta_D-1} - 1$, $\frac{\exp(a)}{1+b} - 1 = \frac{1}{\nu}\frac{\eta_X}{\eta_X-1} - 1$.

[c] Standard errors are robust to heteroskedasticity and autocorrelation.

[d] Standard errors computed using the delta method.

Table 4: Estimating the system for exports and domestic sales. Nonlinear GMM.

Industry	Input elasticity			Demand elasticity		Shifters domestic sales equation				Shifters exports equation			
	k (s. e.)[b]	l (s. e.)[b]	m (s. e.)[b]	η_D	η_X	*Middle-West* (s. e.)[b]	*Age* (s. e.)[b]	*Subsidy* (s. e.)[b]	*S. effort* (s. e.)[b]	*Middle-West* (s. e.)[b]	*Age*[a] (s. e.)[b]	*Subsidy* (s. e.)[b]	*S. effort* (s. e.)[b]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Food, drink and tobacco	0.040	0.274	0.648	6.3	14.6	-0.227	0.083	-1.986	0.087	-0.617	0.015	-5.867	0.048
	(0.014)	(0.042)	(0.043)			(0.054)	(0.016)	(1.009)	(0.074)	(0.131)	(0.019)	(3.614)	(0.164)
2.Textile, leather and shoes	0.030	0.391	0.517	20.3	66.4	-0.586	0.504	-0.014	2.816	-1.984	-0.596	-7.512	3.767
	(0.006)	(0.017)	(0.017)			(0.103)	(0.075)	(4.478)	(0.270)	(0.554)	(0.323)	(17.349)	(0.804)
3. Timber and furniture	0.020	0.220	0.676	13.2	29.3	-0.345	-0.048	11.916	0.771	-0.988	-0.070	24.095	1.605
	(0.018)	(0.037)	(0.037)			(0.155)	(0.103)	(39.458)	(0.214)	(0.422)	(0.127)	(87.113)	(0.560)
4.Paper and printing products	0.060	0.273	0.623	10.0	19.9	-0.564	0.232	35.415	1.547	-1.140	0.054	113.205	1.571
	(0.013)	(0.042)	(0.039)			(0.156)	(0.148)	(23.909)	(0.403)	(0.321)	(0.096)	(51.136)	(0.567)
5. Chemical products	0.066	0.055	0.867	6.3	12.0	0.046	0.039	-1.929	0.625	0.013	0.025	-7.032	0.477
	(0.009)	(0.017)	(0.015)			(0.066)	(0.015)	(1.437)	(0.073)	(0.140)	(0.013)	(1.835)	(0.083)
6. Non-metallic minerals	0.078	0.300	0.524	14.9	29.9	-0.501	-0.141	8.073	0.791	-1.196	-0.194	-0.170	1.755
	(0.014)	(0.027)	(0.030)			(0.146)	(0.089)	(10.801)	(0.180)	(0.301)	(0.118)	(20.753)	(0.427)
7. Metals and metal products	0.059	0.223	0.665	14.9	26.1	-0.431	0.184	3.637	2.652	-0.696	0.098	-4.300	2.670
	(0.009)	(0.020)	(0.021)			(0.101)	(0.064)	(5.590)	(0.364)	(0.168)	(0.049)	(10.822)	(0.513)
8. Machinery	0.074	0.202	0.685	8.7	17.2	-0.343	0.179	13.334	0.336	-0.766	0.245	20.508	0.536
	(0.006)	(0.020)	(0.021)			(0.059)	(0.020)	(9.442)	(0.066)	(0.118)	(0.028)	(15.765)	(0.108)
9.Transport equipment	0.093	0.118	0.777	10.0	18.3	-0.065	0.051	23.039	1.028	-0.306	0.025	22.116	0.793
	(0.011)	(0.018)	(0.021)			(0.093)	(0.013)	(6.004)	(0.125)	(0.168)	(0.010)	(7.006)	(0.191)
10. Electronics	0.077	0.505	0.454	6.1	10.7	-0.482	0.206	1.101	0.303	-0.918	0.277	-0.096	0.436
	(0.009)	(0.035)	(0.033)			(0.064)	(0.017)	(1.215)	(0.066)	(0.121)	(0.023)	(2.820)	(0.109)

[a] In industries 8 and 10 the variable is *Experience*.

[b] Standard errors robust to heteroskedasticity and autocorrelation and corrected for two-step estimation.

Table 5. Distribution of ω and $\delta/(\eta_D - 1)$.

Industry	Quartiles 1998[a]			Quartiles 2008[a]			Standard dev.[a]		Skewness[a,b]	Correl. between	Correl. ω with	Correl. $\frac{\delta}{\eta_D-1}$	Correlation of profits with	
	0.25	0.50	0.75	0.25	0.50	0.75	1998	2008	1998-08	ω and δ	$\overline{mc}$	with mc	ω $\frac{\delta}{\eta_D-1}$	$\omega + \frac{\delta}{\eta_D-1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Food, drink and tobacco	-0.476	-0.197	0.119	-0.201	0.067	0.329	0.438	0.431	0.010	-0.400	0.669	0.587	0.010	0.147
	-0.468	-0.016	0.396	-0.376	0.063	0.459	0.676	0.665	-0.131				0.134	
2.Textile, leather and shoes	-0.518	-0.224	0.069	-0.110	0.178	0.481	0.461	0.458	0.029	-0.285	0.883	0.691	0.177	0.002
	-0.330	0.042	0.321	-0.272	-0.001	0.223	0.649	0.480	-0.158				-0.151	
3. Timber and furniture	-0.469	-0.221	-0.009	-0.095	0.101	0.301	0.358	0.303	0.016	-0.317	0.781	0.828	0.102	0.261
	-0.169	0.059	0.190	-0.147	0.061	0.242	0.329	0.322	-0.175				0.203	
4.Paper and printing products	-0.508	-0.251	0.008	-0.072	0.158	0.398	0.377	0.381	-0.022	-0.449	0.713	0.733	0.199	0.013
	-0.415	0.023	0.344	-0.380	-0.051	0.278	0.649	0.587	-0.033				-0.120	
5. Chemical products	-0.499	-0.213	0.060	-0.151	0.126	0.391	0.449	0.441	0.019	-0.839	0.271	0.899	-0.049	0.109
	-0.371	0.031	0.360	-0.331	0.056	0.403	0.684	0.661	-0.127				0.095	
6. Non-metallic minerals	-0.725	-0.407	-0.089	-0.095	0.197	0.476	0.487	0.447	-0.020	-0.037	0.903	0.774	0.141	0.283
	-0.176	0.017	0.189	-0.118	0.047	0.203	0.301	0.274	-0.148				0.305	
7. Metals and metal products	-0.500	-0.276	-0.048	-0.091	0.138	0.370	0.366	0.350	-0.001	-0.395	0.673	0.427	0.084	-0.057
	-0.389	-0.057	0.296	-0.337	-0.058	0.230	0.534	0.517	0.042				-0.113	
8. Machinery	-0.467	-0.193	0.052	-0.138	0.105	0.344	0.415	0.381	0.028	-0.525	0.584	0.938	0.087	0.204
	-0.869	-0.282	0.228	-0.162	0.193	0.477	0.695	0.556	-0.220				0.120	
9.Transport equipment	-0.673	-0.400	-0.116	-0.024	0.193	0.429	0.402	0.346	-0.003	-0.658	0.651	0.886	0.020	0.098
	-0.246	0.056	0.322	-0.295	-0.002	0.243	0.454	0.487	-0.044				0.059	
10. Electronics	-0.813	-0.291	0.216	-0.330	0.108	0.568	0.789	0.732	0.034	-0.507	0.595	0.890	0.110	0.225
	-1.129	-0.300	-0.383	-0.358	0.226	0.734	1.170	0.935	-0.160				0.118	

[a] First row reports ω, second row $\delta/(\eta_D - 1)$.
[b] (Mean-Median)/Standard Deviation

Table 6: Growth of ω and $\frac{\Delta\delta}{(\eta-1)}$, weighted growth, and contributions to weighted growth 1998-2008.

Industry	$\Delta\omega$	$\frac{\Delta\delta}{(\eta-1)}$ [a]	Comp. of $\frac{\Delta\delta}{(\eta-1)}$ [a]		Weighted growth of ω and contributions[b,c]				Weighted growth of $\frac{\delta}{(\eta-1)}$ and contributions[b,d]			
			G.growth	Entry	Total	Survivors[e]	Entrants[f]	Exitors[g]	Total	Survivors[e]	Entrants[f]	Exitors[g]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. Food, drink and tobacco	0.244	0.073	0.060	0.013	0.150	0.127	-0.070	0.093	0.089	-0.124	-0.002	0.215
2.Textile, leather and shoes	0.415	0.026	0.032	-0.006	0.404	0.506	-0.142	0.040	0.035	-0.118	0.278	-0.125
3. Timber and furniture	0.321	0.014	0.036	-0.022	0.279	0.255	-0.016	0.041	0.036	0.092	-0.076	0.020
4.Paper and printing products	0.411	0.020	0.005	0.015	0.596	0.718	-0.080	-0.042	-0.209	-0.606	0.183	0.214
5. Chemical products	0.336	0.027	0.046	-0.019	0.421	0.440	-0.064	0.045	-0.016	-0.050	0.108	-0.074
6. Non-metallic minerals	0.601	0.028	0.033	-0.005	0.693	0.632	-0.004	0.065	0.010	0.006	-0.051	0.055
7. Metals and metal products	0.414	0.000	0.010	-0.010	0.479	0.705	-0.114	-0.112	-0.006	-0.266	0.333	-0.072
8. Machinery	0.314	0.375	0.386	-0.011	0.377	0.485	-0.081	-0.027	0.391	0.182	0.115	0.094
9.Transport equipment	0.600	-0.038	-0.033	-0.005	0.669	0.743	-0.115	0.041	-0.117	-0.332	0.116	0.099
10. Electronics	0.430	0.448	0.476	-0.028	0.617	0.621	-0.173	0.169	0.185	-0.132	0.182	0.132

[a] $\frac{1}{\eta-1} = S_D \frac{1}{\eta_D - 1} + S_X \frac{1}{\eta_X - 1}$, where S_D, S_X are firm level revenue shares of domestic sales and exports.
[b] 1% of observations at each tail of the distribution of have been trimmed for this exercise.
[c] $\sum w_{j08}\omega_{j08} - \sum w_{j98}\omega_{j98}$.
[d] $\sum w_{j08}\frac{\delta_{j08}}{(\eta-1)} - \sum w_{j98}\frac{\delta_{j98}}{(\eta-1)}$.
[e] Includes additions that were already born in 1996.
[f] Includes starts in the export market.
[g] Includes firms that stop exporting.

Table 7: Cost and demand advantages in the process of privatizacion.

	Growth 1998-2008				Impact of ownership change			
	$\Delta\omega$		$\Delta\delta/(\eta-1)$		ω		$\delta/(\eta-1)$	
Industry	With change	Always private	With change	Always private	No-FE (s.e.)	FE (s.e.)	No-FE (s.e.)	FE (s.e.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	0.260	0.187	0.110	0.004	0.106	0.021	-0.112	0.014
					(0.024)	(0.015)	(0.033)	(0.011)
2.Textile, leather and shoes	0.592	0.418	-0.244	-0.044	0.041	0.037	-0.013	-0.020
					(0.019)	(0.008)	(0.018)	(0.005)
3. Timber and furniture	0.276	0.258	0.047	0.046	0.063	0.023	0.020	-0.005
					(0.040)	(0.022)	(0.032)	(0.014)
4.Paper and printing products	0.434	0.371	-0.164	-0.079	0.079	0.036	-0.102	0.005
					(0.045)	(0.025)	(0.087)	(0.023)
5. Chemical products	0.301	0.228	0.046	0.110	0.064	0.009	-0.069	0.011
					(0.017)	(0.009)	(0.018)	(0.010)
6. Non-metallic minerals	0.620	0.436	0.102	0.100	0.060	0.018	0.009	0.007
					(0.033)	(0.016)	(0.015)	(0.007)
7. Metals and metal products	0.493	0.331	-0.210	-0.026	0.084	0.033	-0.020	-0.005
					(0.024)	(0.011)	(0.029)	(0.012)
8. Machinery	0.384	0.287	0.181	0.168	0.079	0.012	-0.069	0.004
					(0.021)	(0.009)	(0.035)	(0.007)
9.Transport equipment	0.670	0.415	-0.063	0.006	0.022	0.002	-0.053	0.018
					(0.037)	(0.015)	(0.032)	(0.020)
10. Electronics	0.460	0.327	0.351	0.352	0.027	0.011	0.034	0.012
					(0.042)	(0.018)	(0.051)	(0.015)

Table 8: Cost and demand advantages and export specialization.

Industry	Observed distribution of export intensity (ei)		Marginal effects in the regression of export intensity			Predicted distribution of export intensity (ei)	
	$P(ei \leq 0.2)$	$P(ei \geq 0.8)$	Relative cost advantage (s. d.)	Experience (s. d)	R^2	$P(ei \leq 0.2)$	$P(ei \geq 0.8)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Food, drink and tobacco	0.308	0.278	2.246	-0.017	0.599	0.303	0.227
			(0.023)	(0.002)			
2.Textile, leather and shoes	0.167	0.406	0.160	0.009	0.133	0.000	0.051
			(0.017)	(0.002)			
3. Timber and furniture	0.187	0.390	2.175	0.011	0.667	0.114	0.328
			(0.033)	(0.003)			
4.Paper and printing products	0.398	0.168	1.333	-0.026	0.348	0.333	0.042
			(0.036)	(0.005)			
5. Chemical products	0.307	0.187	1.936	-0.021	0.860	0.311	0.173
			(0.006)	(0.001)			
6. Non-metallic minerals	0.369	0.171	2.081	0.023	0.518	0.378	0.102
			(0.041)	(0.002)			
7. Metals and metal products	0.253	0.315	1.428	-0.030	0.268	0.097	0.198
			(0.021)	(0.002)			
8. Machinery	0.356	0.184	1.466	-0.014	0.354	0.294	0.068
			(0.014)	(0.002)			
9.Transport equipment	0.380	0.182	1.888	-0.035	0.671	0.391	0.141
			(0.022)	(0.002)			
10. Electronics	0.216	0.326	1.840	-0.001	0.523	0.173	0.290
			(0.015)	(0.001)			

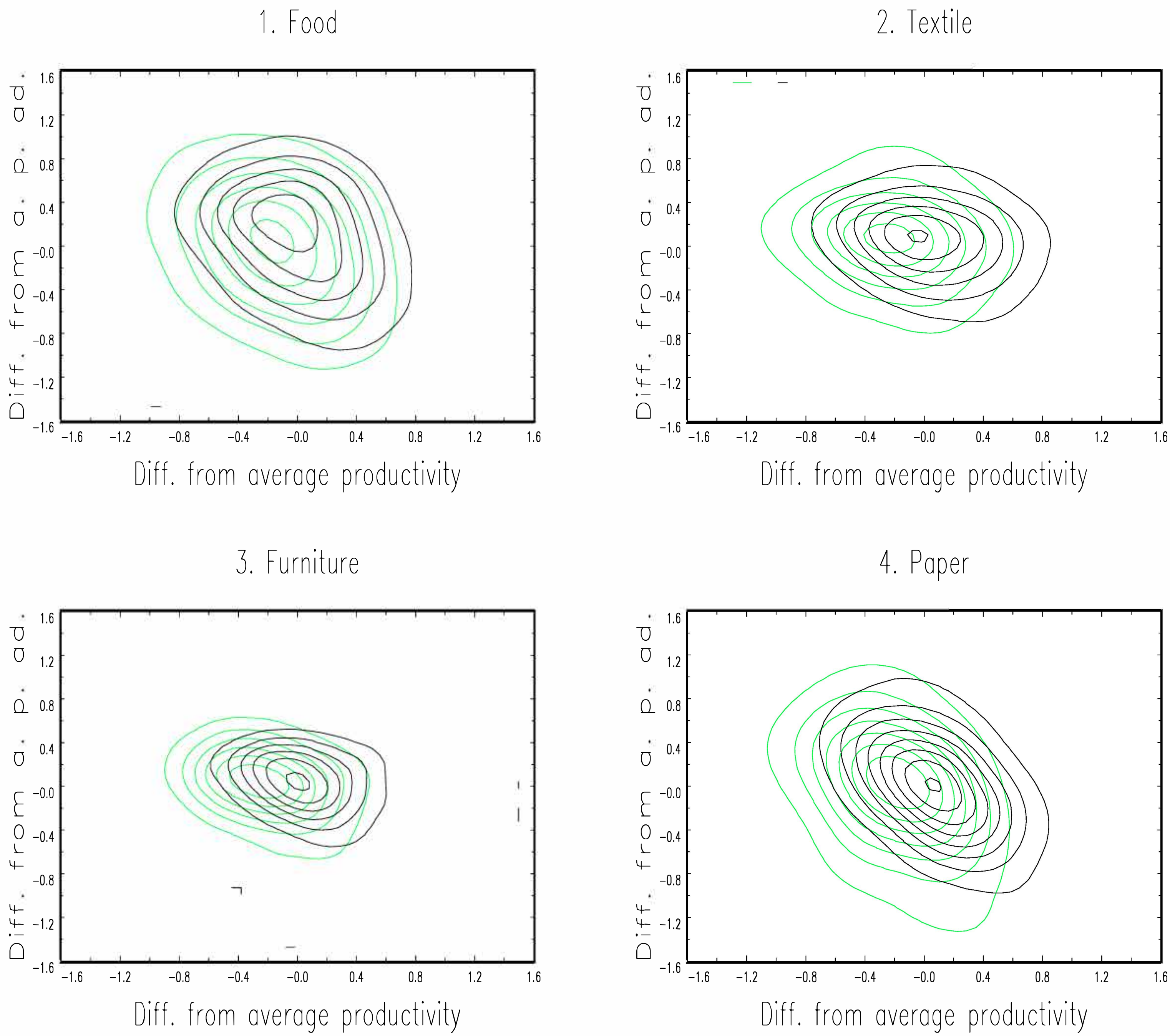
Table 9: R&D investment, workforce skills and cost and demand advantages.

Industry	$R\&D$ investment[a]						Workforce skills	
	ω			δ				
	No R&D	R&D	$Corr(\omega, \frac{R\&D}{R})^b$	No R&D	R&D	$Corr(\delta, \frac{R\&D}{R})^b$	$Corr(\omega, qual)$	$Corr(\delta, qual)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	0.025	-0.050	-0.049	-0.001	0.298	-0.078	0.159	0.128
2.Textile, leather and shoes	0.028	0.211	-0.013	0.045	-0.221	0.022	0.215	-0.051
3. Timber and furniture	0.026	0.010	-0.137	-0.002	0.114	0.070	0.185	0.058
4.Paper and printing products	0.034	0.099	-0.111	0.043	-0.119	-0.006	0.233	-0.039
5. Chemical products	0.062	-0.038	-0.040	-0.028	0.129	0.007	-0.033	0.049
6. Non-metallic minerals	0.102	0.053	-0.072	0.001	0.070	0.068	0.170	0.036
7. Metals and metal products	0.037	0.049	-0.119	0.051	-0.203	0.079	0.193	-0.104
8. Machinery	0.038	0.015	-0.138	0.061	0.071	0.065	0.017	0.189
9.Transport equipment	0.088	0.023	-0.048	-0.016	0.007	-0.014	-0.004	0.062
10. Electronics	0.066	0.007	-0.140	-0.003	0.214	0.094	0.137	0.157

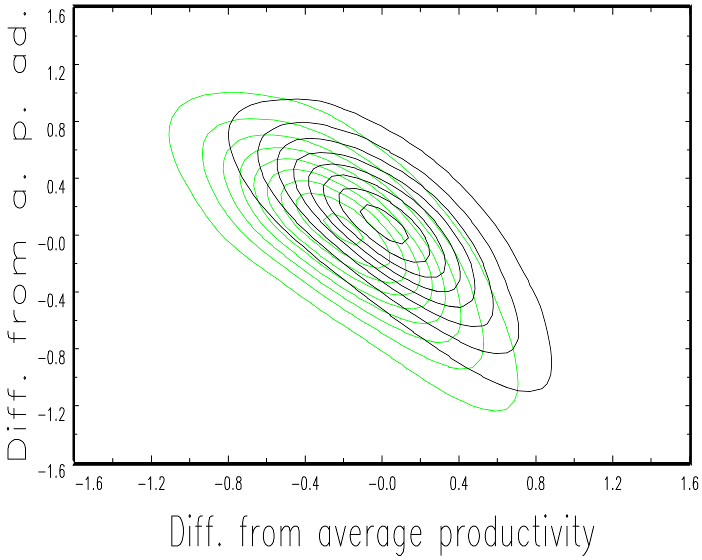
[a] Statistics compued over the years 2001 and 2005 to 2007.
[b] Computed for firms with R&D expenditure.

Figure 1: Joint density of ω and $\delta/(\eta_D\text{-}1)$

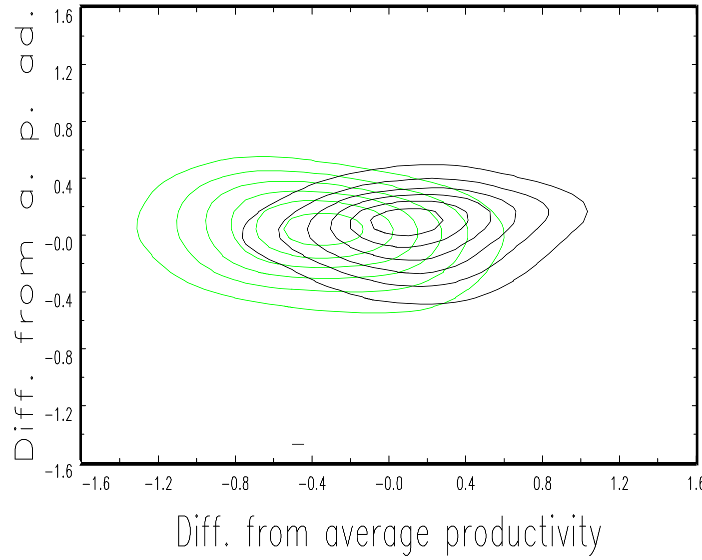
Change from 1998–2000 (light) to 2005–2008



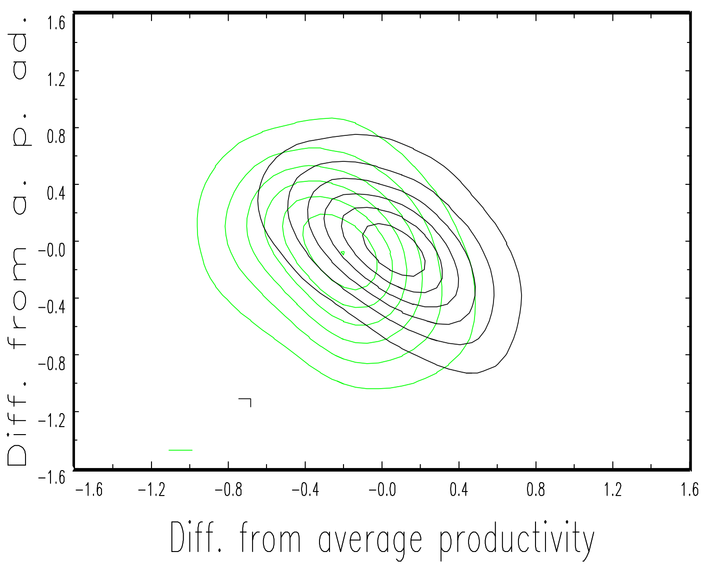
5. Chemical



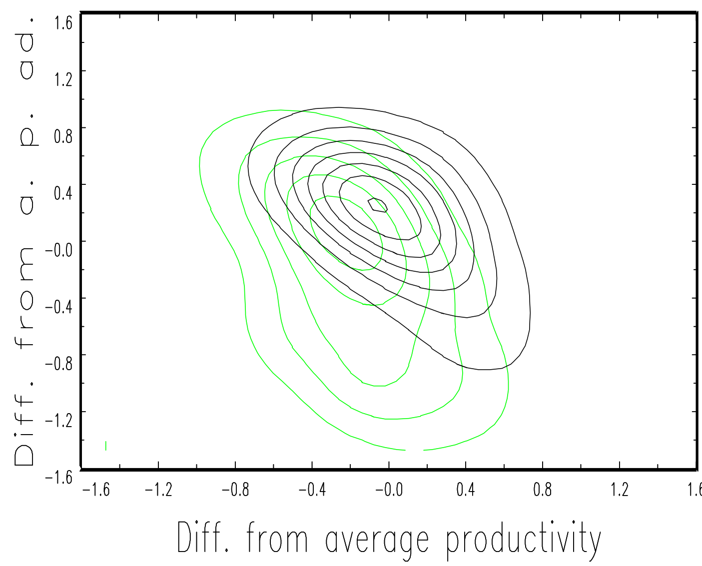
6. Non-met. minerals



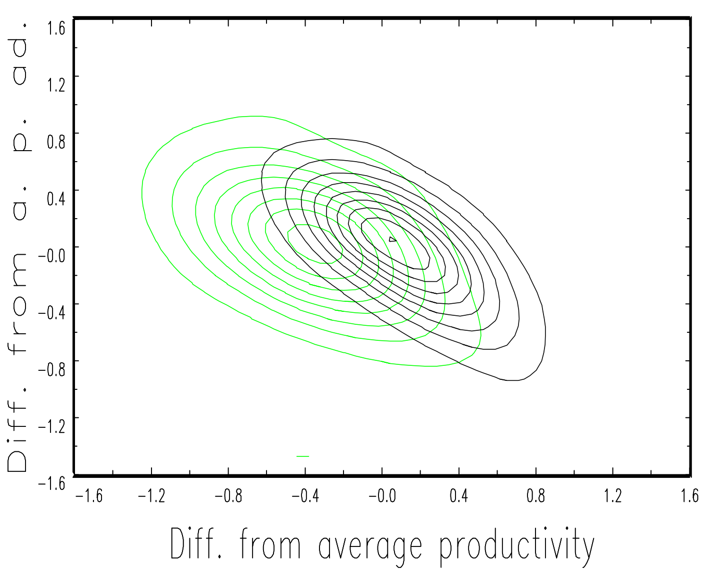
7. Metal products



8. Machinery



9. Transport



10. Electronics

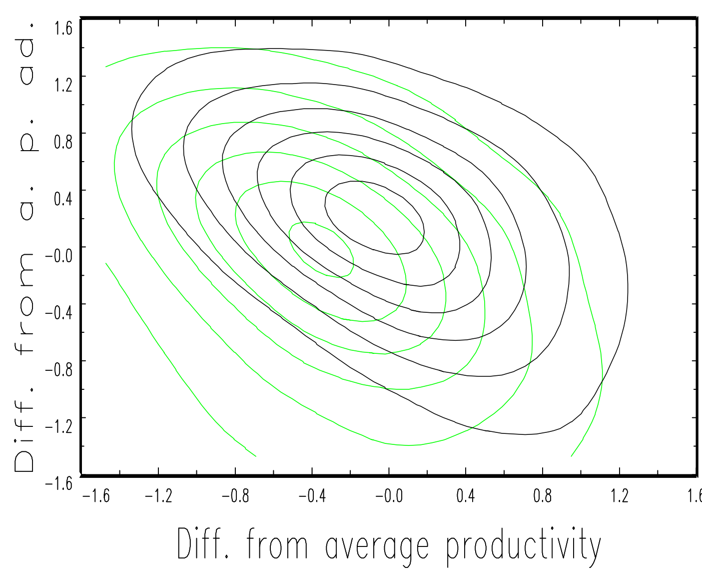


Figure 2: Changes in the mean of ω and $\delta/(\eta_D\text{-}1)$

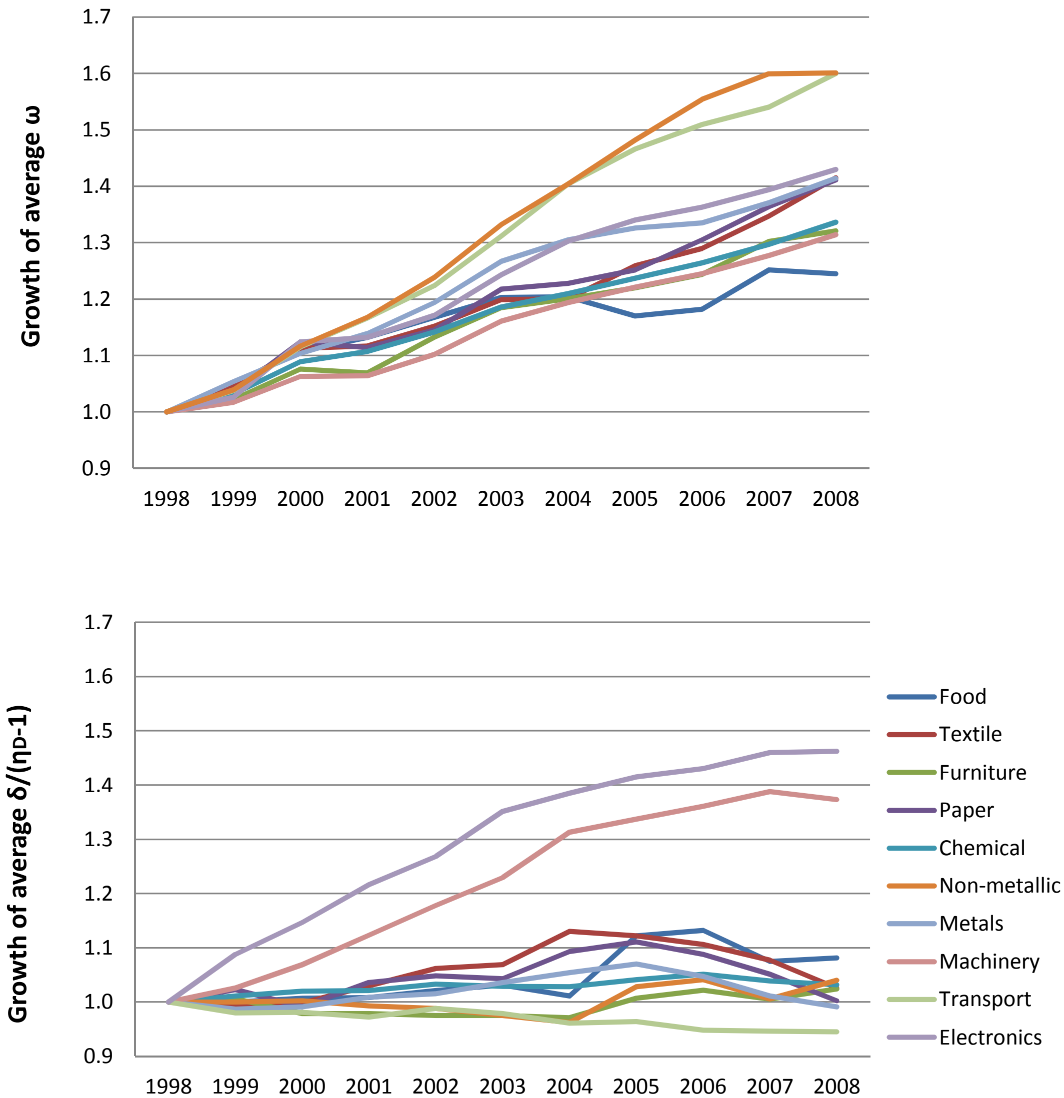


Figure 3: Predicting export intensities

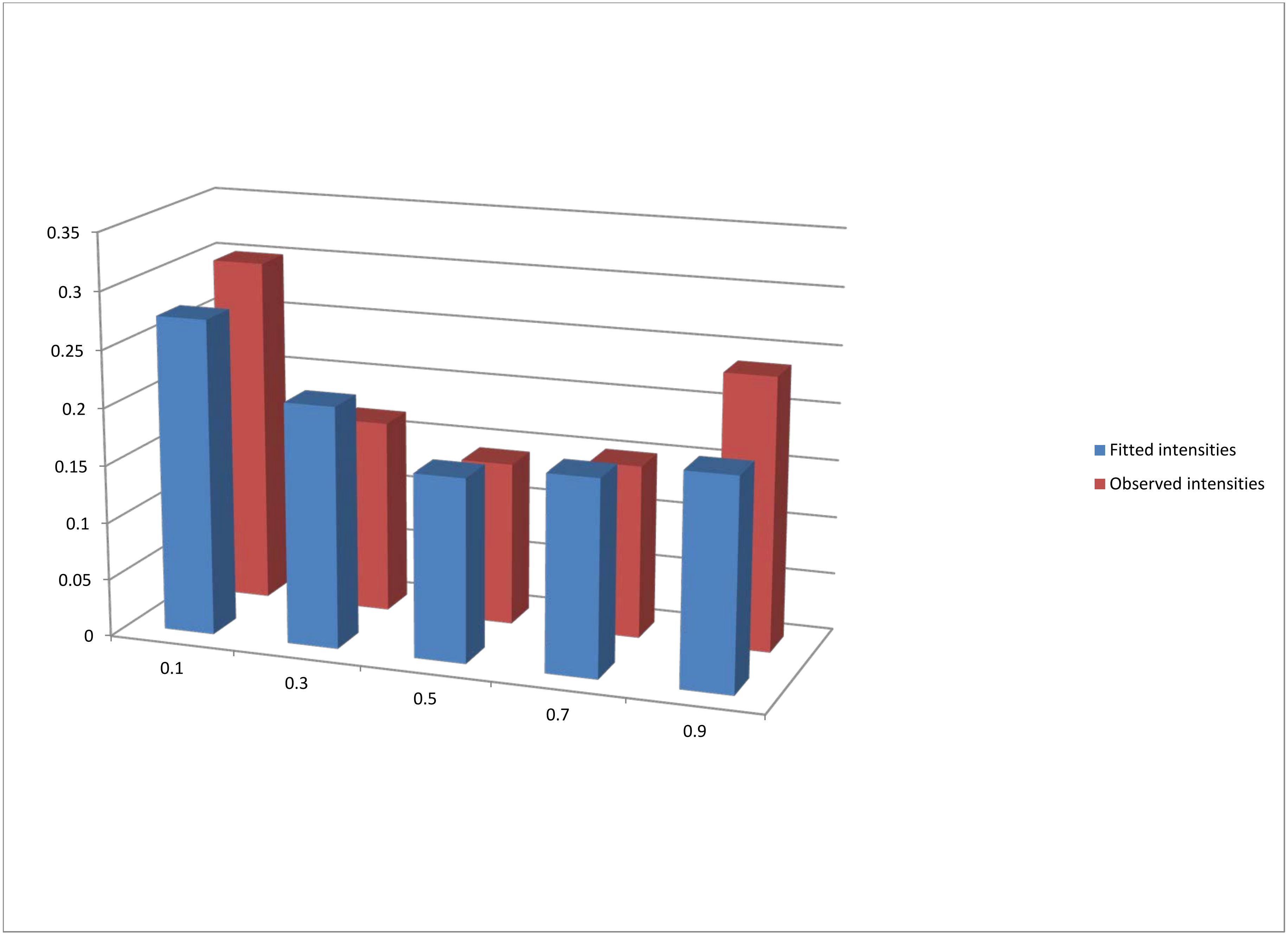


Figure 4: Advantages, R&D and labor quality.