Cost and Product Advantages: Evidence from Chinese Manufacturing Firms*

Jordi Jaumandreu[†] Heng Yin[‡]

This version: May 2018
First draft: January 2013

Abstract

We use data on 70,000 Chinese manufacturing firms that sell domestically and export to robustly estimate the joint distribution of unobserved productivity (cost advantages) and unobserved demand heterogeneity (product advantages) from 1998 to 2008. Product advantages show a trade off with cost advantages and are positively related to observed costs. Using the advantages we characterize Chinese manufacturing, that grew competing more on costs than in product advantages (which account for a significant but small 24% of growth). Our estimation highlights important biases affecting the estimates of the coefficients of the production function, demand elasticities and markups, when heterogeneity of demand or its correlation with productivity are ignored. With the separation of cost and product advantages, we revisit and reinterpret recent studies to find new results which change their policy consequences.

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

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

*This paper previously circulated under the title "Cost and Product Advantages: A Firm-level Model for the Chinese Exports and Industry Growth." 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, F. De Carolis, J. DeLoecker, J.C. Farinas, I. Fernandez-Val, M. Gebressilase, P. Goldberg, M. Manove, Shengyu Li, G. Ottaviano, B. Peters, M. Roberts, M. Rysman, A.Pakes and numerous seminar audiences for comments.Yin acknowledges the support of a research grant from the National Natural Science Foundation of China (project code 71673305). 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 that we observe two firms that 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. 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.[1] While unobserved advantages in production are called productivity, unobserved product advantages are referred 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).[2] Importantly, we adopt a residual and nonparametric definition of demand heterogeneity, symmetric to the prevalent definition of productivity.

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 sales by means of observed and unobserved demad-expanding actions (a shift in their demand curves).[3] This explains why firms

[1] The degree of observability of product advantages depends on the availability of appropriate variables to control for them. For example, the awareness effect can be partially observed through the impact of variables measuring the 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 result in the 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] Here are two specific examples referred to our database. *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 employees), 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

in imperfectly competitive markets make at least three different types of investment: physical capital to carry out production, knowledge and technology to compete in costs, and knowledge and demand enhancement to compete in products.[4] We adopt the dynamic model of endogenous productivity of Doraszelski and Jaumandreu (2013), where firms invest in R&D to enhance productivity, and we extend it to include the investments that determine shifts of the product advantages.

While there is an extensive literature analyzing productivity distributions (see Bartelsman and Doms, 2000, and Syverson, 2011, for surveys), demand heterogeneity and its distribution have only been examined recently. One likely reason is that demand heterogeneity has often shown up as a nuisance. Take the example of the estimation of the production function and markups. When a researcher has access to firm-level output prices but competition is imperfect, the most frequently used methods become inconsistent if one admits the presence of demand heterogeneity.[5,6] The researcher has incentives to ignore the problem. However, since the significance of demand heterogeneity and its correlation with included variables has become increasingly documented, it is pertinent to worry about the robustness of the empirical results based on ignoring it.

As time goes by demand heterogeneity has become an object of interest and research. When firm-level output prices are not available, then a researcher has no other solution that specifying how is the heterogeneity of demand. Klette and Griliches (1996) first showed this, a point that was

50,000 employees). However, the price of standard noodles increased from 1.5 RMB in 1998 to 2.8 RMB in 2008. Sales soared as the firm tripled its centers of distribution and engaged in closer relationships with retailers instead of relying on wholesalers.

[4] This is, for example, recognized by the official statistical approaches to R&D where breakthroughs are categorized as "process" or "product" innovations. Recent papers have started to discuss the relative incentives for a firm's investment in process and product innovation: Dhingra (2013), Flach and Irlacher (2018).

[5] This is recognized by Ackerberg, Caves and Frazer (2015) when they write: "Alternatively (to price takers in the input and output markets), consider a situation where firms face downward sloping output demand curves (and/or upward sloping input supply curves). In this case, one will generally need to assume that firms are all facing identical demand/supply curves; otherwise, firms will have different intermediate input (or investment) demand functions (i.e. the scalar unobservable assumption 4, 4b or 4c will be violated)."

[6] Under imperfect competition, the first order conditions for input choice include firm marginal revenue (if profit maximization is assumed) or marginal cost (if cost minimization is assumed). Hence, an Olley and Pakes (1996)/Levisohn and Petrin (2003) method to estimate the production function must replace unobservable productivity with an inverted input demand that includes unobserved marginal revenue or marginal cost. Non-observable marginal cost can in principle be replaced by its determinants, but then the relevant firm's output or its determinants should be included. When researchers specify ouput determinants, they implicitly drop the unobservable heterogeneity of demand that would make estimation inconsistent. This is what papers as De Loecker and Warzynski (2012), De Loecker, Goldberg, Khandelwal and Pavcnik (2016), or Brandt, Van Biesebroek, Wang and Zhang (2017), for example, do.

recently restated in De Loecker (2011).[7] Both papers solved the involved problems by assuming that unobserved heterogeneity of demand is iid.

However, in practice, heterogeneity of demand has been found to be persistent and correlated with included variables. Many researchers have then switched to estimating a "composite" of productivity and demand heterogeneity (that happens to show up with weights of a third unobservable, the elasticity of demand), calling it loosely "productivity".[8] This is problematic at least for two reasons. First, estimation requires the assumption that this composite follows a Markov process. This assumption is incompatible with each unobservable following a different Markov process. Second, nothing guarantees that the estimate is going to behave as productivity, particularly in the presence of correlation among its components. In a recent paper, Guillard, Jaumandreu and Olivari (2018) find, among other problems, that the composite of the firms that perform R&D does not maintain the stochastic dominance properties over the composite of non R&D firms that productivity does.

Fortunately, there is already a significant literature that tries to separately estimate demand heterogeneity, and thus also establishes a methodology for the consistent estimation of the production function and productivity in the presence of imperfect competition and demand heterogeneity. Foster, Haltiwanger and Syverson (2008), in a pioneering work, disentangled productivity and demand heterogeneity using a sample of US quasi-homogeneous good industries, for which they used unit values as prices. 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. Pozzi and Schivardi (2016) develop, in a broader sample, a similar analysis in terms of time differences using price changes and a subjective assessment of demand elasticity from managers. Roberts, Xu, Fan and Zhang (2017) observe the destination of exports at the product-level for a sample of 738 Chinese footwear producers and take unit values as prices. They assess the relative importance of the firm idiosyncratic demand effect across markets and firm specific marginal costs.[9]

[7] This problem is also stressed in Katayama, Lu and Tybout (2009).

[8] Some examples are Hsieh and Klenow (2009), Gandhi, Navarro and Rivers (2017), 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).

[9] 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 the TFP of the rest of the firms to instrument price or output

An alternative literature has used tightly parametrized frameworks to do the same job. Hottman, Redding and Weinstein (2016) model a sample of US supermarket goods with available prices, allowing for the measurement of productivity and demand heterogeneity. Firm products "appeal", for example, is derived from the consumers' valuation weight of each good quantity in a CES function symmetric over variety valuations. Eslava and Haltiwanger (2017) use the output and input prices available for Colombian manufacturing to apply a similar CES model to assess the sources of long-run growth. Forlani, Martin, Mion and Muuls (2017) also uses a tight functional specification of demand heterogeneity in an exercise aimed at measuring productivity, demand heterogeneity and markups with Belgian data.[10]

We depart from these works in three main aspects. First, we extend the exercise of separating productivity from demand advantages to a sample that includes all kinds of differentiated products. We use firms from China's entire manufacturing (split into ten broad industries). Second, we construct a robust framework. Our product advantages do not depend on the functional specification and we allow productivity and demand heterogeneity to be freely correlated. Foster, Haltiwanger and Syverson (2008, 2016), Hottman, Redding and Weinstein (2016), and Eslava and Haltiwanger (2017) base identification on the orthogonality of productivity and demand heterogeneity, but the papers that free this correlation consistently show a negative relationship (Roberts, Xu, Fan and Zhang, 2017; Forlani, Martin, Mion and Muuls, 2017) that we strongly confirm in this paper. This correlation denies the legitimacy of productivity as instrument for price in the presence of unobserved demand heterogeneity. Third, we estimate quantity productivity and quantity demand advantages without using prices. This makes the technique to separate productivity and demand advantages widely usable because firm-level output prices are rarely available. A corollary is that this makes the estimates not dependent on the specific assumptions used to construct price indices as happens with many papers. Of course, the advantages of having a more comprehensive exercise, nonparametrically oriented and without using prices, come at the cost of some restrictive assumptions. We assess how restrictive after explaining with some detail how we identify the model.

Here is our identification strategy. We select firms that sell in both the exports and domestic markets. We assume that each firm sells the same product, group of products, in each market and verify that the elasticity of demand in the export market is higher than in the domestic market. In

(the last paper uses labor productivity). Taking advantages as a fact, Jaumandreu and Mairesse (2017) explore exogenous and endogenous determinants of their shifts.

[10] Grieco, Li and Zhang (2017) is also a simultaneous estimation of unobserved productivity and "quality" imposing strong functional restrictions.

this setting, cost advantages have a greater impact on the export market. This allows us to invert the system of demands for exports and domestic sales, once that prices have been substituted for, backing up productivity and demand heterogeneity from observed data.[11]

To implement this strategy, we specify the two demands for the product of the firm.[12] Each demand depends on the firm price for the specific market, observable shifters, and a persistent unobservable to reflect the product advantages of the firm, which we allow to evolve over time. As we do not observe output prices, we transform the demands into revenue equations and replace the explanatory price variable by its optimal level in terms of the firm-specific marginal cost.[13] Marginal cost has an observable part but depends on unobservable productivity too. This transformation gives us two equations in which the firm's sales depend on the unobservables productivity and demand heterogeneity in addition to the observed cost and demand shifters. We specify each unobservable as a Markov process.

The identification of the parameters of the production function, embedded in the marginal cost expression, requires the simultaneous estimation of the elasticity of demand for the firm's products in both markets. This cannot be done relying solely on the system of demands. Assuming static pricing[14], we specify a third equation with the ratio revenue-variable cost (the inverse of the share of variable cost in revenue) as dependent variable. This ratio gives a measure of the price average-cost ratio, which is a function of the elasticities entangled with the short-run returns to scale parameter of the firm. Neither this equation nor the system of demands are able to identify the elasticiticies and the parameter of scale by themselves, but their simultaneous estimation can.[15] This new methodology can potentially be applied in other contexts.

Exit and entry are quantitatively important in our sample. There is no particular reason, however, to think that this turnover creates a sample selection problem and we follow the current practice of not trying to account for it.[16] However, our selection of firms that sell domestically and abroad

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

[12] Exports and domestic, as in Das, Roberts and Tybout, 2007, and Aw, Roberts and Xu, 2011

[13] This can be compared to De Loecker (2011), who writes the inverse demand and replaces output by the production function.

[14] Like virtually all the empirical works that estimate production function and markups. See, for example, De Loecker and Warzynski (2012). See Jaumandreu and Lin (2017) for a departure from this assumption.

[15] Technically, we use a two stage procedure to first estimate a function of the elasticities and scale parameter from the margin equation, which we use in a second step as restriction in the estimation of the system of demands.

[16] We follow Levinsohn and Petrin (2003), Doraszelski and Jaumandreu (2013) and Ackerberg, Caves and Frazer

may be more persuasively argued to create sample selection biases in the estimated coefficients. We check for these biases adapting the procedure of Olley and Pakes (1996) to the presence of two unobservables (to continue exporting, we require that the combination of productivity and demand advantages exceeds a threshold).

We estimate the elasticity of demands, the parameters of the cost (production) function, the impact of observed shifters and the unobservables. To simultaneously identify productivity and demand heterogeneity, we use the same basic assumptions used for the "structural" estimation of production functions: the Markovian character of productivity (extended to demand heterogeneity) and the "timing" assumption that capital is chosen one period before variable inputs.[17] However, identification also relies on the additional explicit assumptions that the firm has the same marginal cost whether the product (group of products) is sold domestically or abroad, and that product advantages are similar in both markets.

Let us make a brief assessment of these additional assumptions. First, it is important to notice that they are less restrictive than the assumptions embodied in the exercises that implicitly assume a unique firm demand (and hence a unique and invariant elasticity and markup). The existence of an aggregate of exports and domestic sales, responding to an aggregate price index, requires very restrictive assumptions on the composition of sales. Our framework frees these restrictions, allowing for different elasticities, intercepts, and endogenous prices. Second, some theoretical models restrict the products that are sold in each market through the introduction of product-specific fixed costs[18], but the empirical relevance of this exercise is still unclear when domestic versus global export sales are involved.[19] Third, it is very easy (but data demanding) to generalize our model to the presence of different marginal costs, and it is not difficult to model the impact of products in each market as differing in terms of observables. We leave these extensions for future research. Finally, we check that the effect of product advantages can be sensibly taken as the same in both markets by allowing the unobservable to impact differently exports and the domestic market. The results demonstrate that the assumption of similar impact is not unreasonable for most of the markets.

(2015). The likelihood of immediate reaction to the negative (positive) shocks which can trigger biases is low.

[17] See the Ackerberg, Caves and Frazer (2015) summary of assumptions on Olley and Pakes (1996), Levinsohn and Petrin (2003), and Doraszelski and Jaumandreu (2013).

[18] 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.

[19] 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.

The estimation of the model produces three kinds of results. First, it generates an unrestricted estimate of the joint distribution of productivity and demand heterogeneity, as well as its change over time. Second, it highlights biases affecting other estimates of the coefficients of the production function, demand elasticities and markups, attributable to ignoring the heterogeneity of demand. Third, with the analytical scalpel that separates cost and product advantages, we revisit and reinterpret recent studies on China manufacturing, often reversing their policy consequences. In the next section, we summarize these results comparing them with the relevant literature.

The rest of the paper is as follows. Section 2 summarizes results and related literature. In Section 3, we show that the unobservable cost and product advantages are characteristics that are nonparametrically identified in the absence of prices. Section 4 introduces the data and describes the sample. In Section 5, we set out our particular empirical parametric specification. Section 6 explains how we estimate the econometric model. Section 7 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 8 concludes. There are five appendices and an online appendix.

2. Results and related literature overview

We get a reasonable marginal distribution of productivity that has a moderate dispersion, which remains stable over time, and a huge change in mean during the period.[20] Product advantages, compared in a proper scale, are slightly more dispersed and change very slowly and heterogeneously (only *Machinery* and *Electronics* display a change of magnitude comparable to productivity).[21] Despite its heterogeneity, unobservable product advantages explain about 24% of the revenue growth based on productivity and demand heterogeneity. The joint distribution shows a strong negative correlation of the unobserved cost and demand advantages. This negative correlation of the unobservables is also found by Roberts, Xu, Fan and Zhang (2017) and Forlani, Martin, Mion and Muuls (2017).[22] Our interpretation is that many firms that show unobserved product advantages (quality, technology, design, distribution...) tend to show additional costs not captured by the observed

[20] The dispersion of our distributions is not far from the result obtained by Foster, Haltiwanger and Syverson (2008) with a quasi-homogeneous goods sample, but it is far below the dispersion obtained by Hsieh and Klenow (2009) who use the same kind of sample that we use. This is likely due to our separation of productivity and demand heterogeneity.

[21] Our demand heterogeneity estimates have less dispersion than those in Foster, Haltiwanger and Syverson (2016).

[22] Grieco and MacDevitt (2016) also find a negative relationship between productivity and quality of the product in their industry-specific analysis of a health industry.

wage and materials bill (non-wage costs of skills, costs of management of superior materials and costs of organization, etc.). As the observed part of marginal cost is positively correlated with the product advantages (something that it is also found by Hottman, Redding and Weinstein, 2016), the observed and unobserved parts of marginal cost generate a strong positive correlation between marginal cost and product advantages. Developing product advantages is affected by a trade-off. Combining quality, design, technology or better distribution with low costs, has technological and firm knowledge/ability limits which impact the growth paths of firms.[23]

We get elasticities of demand that are larger than the elasticities estimated in many studies, although perfectly consistent with our estimated short-run production elasticity of scale and short-run profitabilities. Our average elasticity is 10 in domestic markets and elasticities roughly double in export markets (similar to Das, Roberts and Tybout, 2007). Export markets hence emerge as more competitive, and firms set lower prices and get smaller margins. We also get reasonable estimates of the marginal cost (production function) coefficients for all industries. The average elasticity of materials is 0.66, 0.24 for labor, and 0.90 together. The implicit average markup in the domestic market is 0.12 and 0.06 in the export market.[24]

Our estimates suggest that uncontrolled heterogeneity of demand is likely to bias significantly downward the estimates of elasticity of demand through two different channels. When the elasticity of demand is directly estimated by regressing sales on prices, the noncontrolled positive correlation between price (through marginal cost) and product advantages induces small elasticities. An example of this happens when productivity is considered a valid instrument for price (Foster, Haltiwanger and Syverson, 2008 and 2016; Hottman, Redding and Weinstein, 2016, and Eslava and Haltiwanger, 2017). On the other hand, when markups are computed by dividing a production elasticity by the (corrected) corresponding input share, a method popularized by De Loecker and Warzynski (2012), the absence of control for the heterogeneity of demand in the estimation of the production elasticity induces large markups (small implicit demand elasticities). For example, Brandt, Van Biesebroek, Wang and Zhang (2017), with identical data to ours, get an average coefficient on materials of 0.913 that is by itself greater than our elasticity of scale and an average markup as large as 0.23

[23] This starts to be a well-established fact for differentiated products, which suggests that we are applying the simple production function conceived for homogeneous outputs to activities that are quality heterogeneous. Jaumandreu and Yin (2018) suggests that a next step of research should be estimating productivity gross of quality, or productivity keeping its potential to become quality.

[24] Markups can be computed using the price average-cost margins in Table 3 combined with the elasticity of scale reported in Table 4.

(the positive margins of Lu and Yu 2015, estimated by the same method, show the same average value).[25] These biases are likely to impact also the correlations on which inferences are based. This urgently calls for the adoption of methods that are robust to the heterogeneity of demand.

The separated estimation of productivity and product advantages allows us to reinterpret many recent results on China manufacturing. Our data show again how important turnover is in the transformation of China manufacturing, as already pointed out by Brandt, Van Biesebroeck and Zhang (2012) and Hsieh and Song (2015). However, entrants are initially less cost efficient. This is contrary to the conclusions of Brandt, Van Biesebroeck and Zhang (2012) who are mislead by the fact that entrants contribute important product advantages. Reallocation among incumbents according to productivity growth happens to be important, but again this can be missed if one cannot separate productivity growth from the decline of product advantages of big firms (as happens to Brandt, Van Biesebroeck and Zhang, 2012). Controlling by selection, our results give a very modest role to privatization in the increase of intramural productivity, which is very similar to what Hsieh and Song (2017) found, and still smaller in the development of product advantages.[26] Our splitting of advantages supplements a nice description of what happened. Productivity grows strongly under the reform at firms that remain under state control, so the relative increases associated with privatization are small. However, all firms that either were or stay under the control of the state seem to be prevented from achieving the large increases in product advantages that the new private entrants show. Private activity appears much more skilled at detecting new commercial opportunities than the state, while the state is quite able to reinforce productivity of established activities.

We also find the key to explaining the apparent "low efficiency" that seems to characterize big exporters and has puzzled researchers who have dived into the question (see, for example, Lu 2010). A significant set of firms with big cost advantages and no product advantages (low wages plus low technology) tend to specialize in the export market, selling most of their production abroad. Their productivity is large, but this is not apparent in an analysis without separation because of their lack of product advantages.

The whole picture that we obtain for the period is that Chinese firms relied heavily on cost competition to grow, and more modestly on product advantages (although these product advantages sharply developed in *Electronics* and *Machinery*). This mix fits well with the findings in Fan, Li

[25] For Brandt, Van Biesebroeck, Wang and Zhang (2017), see Table A2 and Table A3 of the Online Appendix. For Lu and Yu (2015), see Table 5 of the Appendix.

[26] Chen, Igami, Sawada and Xiao (2017) is a paper that tries to dig on the role of privatization on the increase of productivity with structural methods. The challenge is to deal simultaneously with selection.

and Yeaple (2015), who reveal improved products and prices over time, and with Kee and Tang (2016), who detect an increasing ratio of domestic value added on total exports. However, we detect a slowness in acquiring product advantages that could hurt Chinese exports in the long-run, particularly as other countries engage more intensely in the race (see Sutton, 2001 and 2007, for insights on a development model based on a mix of cost and product advantages). Interestingly, this matches the implicit diagnosis of the policy-makers who designed "Made in China 2025."[27]

Our findings, however, do not support the popular idea advanced by Hsieh and Klenow (2009) of a productivity gap in China economy sustained by "price distortions". The methodology and measurement that support this idea are based on the (until now) untested assumption that the value of the so-called "revenue productivity" should be equal for all firms in an industry, and that differences in this value exclusively reflect exogenous (policy induced) distortions of prices.[28] The implication is that the reallocation of resources that would increase productivity is the switching of output to the lowest cost firms. What we find with the same data is a huge endogenous heterogeneity of product quality and technology associated with a corresponding endogenous heterogeneity of input prices, efficiency and marginal costs. Optimal reallocation policy should consist of stimulating the welfare maximizing combination of product and cost advantages, what would include to boost the production of some of the highest cost firms in the industry. This difference of policy implications suggests that the literature on reallocation of resources needs to accommodate the separation between cost and product advantages.[29]

3. 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.

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

[28] A recent paper by Haltiwanger, Kulick and Syverson (2017) criticizes the restrictions implicit in the functional forms used by Hsieh and Klenow (2009) as responsible for the invariance of "revenue productivity", a property that they do not find to hold in their data. Our stress here is that endogenously heterogeneous input prices warrant, even if the functional form restrictions were right, that "revenue productivity" is going to change across firms with a motive that clearly differs from "price distortions".

[29] Dhingra and Morrow (2018) is a paper that analyzes how the productivity heterogeneity of firms impacts the welfare analysis of reallocation. However, nothing has been developed, to our knowledge, about how demand heterogeneity impacts this analysis.

3.1 Revenue as a function of cost and product advantages.

Firm j produces a product that sells in two or more monopolistically competitive markets.[30] Let us consider market I of firm j. The demand for the product at moment t is

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

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.[31] We assume without loss of generality that $Q^{I}(\cdot)$ is monotonic in δ_{jt} and that the impact of δ_{jt} is positive. 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.[32] 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.[33]

[30] 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).

[31] 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, with focus in 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 expression for s_j. However, between the usual industry-specific BLP exercise and the exercise here there are two important 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).

[32] 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}.

[33] 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 $X_{jt} = \{K_{jt}, M_{jt}, W_{jt}, P_{Mjt}\}$.

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

$$R^I_{jt} = P^I_{jt} Q^I(P^I_{jt}, Z^I_{jt}, \delta_{jt}) \tag{2}$$

and, inverting the profit maximization condition $MR(P^I_{jt}, Z^I_{jt}, \delta_{jt}) = MC(X_{jt}, \omega_{jt})$,[34] we can write

$$P^I_{jt} = MR^{-1}(MC(X_{jt}, \omega_{jt}), Z^I_{jt}, \delta_{jt}). \tag{3}$$

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

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

This equation[35] 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 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.

3.2 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

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

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

$$\begin{aligned} \omega_{jt} &= \omega(X_{jt}, Z^X_{jt}, Z^D_{jt}, R^X_{jt}, R^D_{jt}), \\ \delta_{jt} &= \delta(X_{jt}, Z^X_{jt}, Z^D_{jt}, R^X_{jt}, R^D_{jt}). \end{aligned} \tag{6}$$

[34] We assume that MR is monotonic in price. Firm equilibrium is unique if the profit function is strictly quasiconcave, that is a standard condition assumed in monopolistic competition models (see, for example, Zhelobodko, Kokovin, Parenti and Thisse, 2012). Quasiconcavity of the profit function when marginal costs is constant, for example, is sufficient to ensure the invertibility of marginal revenue.

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

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 can 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 elasticity 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.[36] 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.

3.3 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

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

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

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.

4. Data

In what follows, we firstly describe how we build a panel data set, construct variables and clean the data.[37] Then we assess the dynamics of the data and sample that we are going to use and comment the descriptive statistics.

4.1 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).[38] 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.[39] Our data was collected from 1998 to 2008.

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.[40] After linking

[37] 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.

[38] 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.

[39] Other recent studies which use this source are Brandt, Van Biesebroek, Wang and Zhang (2017); Roberts, Xu, Fan and Zhang (2016); Lu and Yu (2015); Lu (2010) and Hsieh and Klenow (2009).

[40] 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 identify entrants separately from additions to the survey.

the data we find reasonable rates of economic entry, additions and exit, which average 9.4%, 7.8% and 7.9% respectively. Many additions are likely to come from firms growing large enough to be included in the survey. But this is not everything, additionally there are improvements over time in statistical coverage that need to be accounted for to get the right interpretation of the numbers.[41]

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 B, 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 C, 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.[42] 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.

4.2 Growth and reallocation.

The treated data shows that Chinese firms underwent important growth during this period. Additionally, there was a large reallocation of manufacturing activity.[43] 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 9% of firms and the additions to the survey constitute 14%.

Survivors grow over time, entrants and additions to the survey are significantly smaller.[44] The exitors, although smaller than survivors, are in turn larger than entrants and additions. The result is that output and productivity increase sharply while 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.

[41] 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.

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

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

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

4.3 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 or entered later 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 join the sample now include existing firms that start to export together with the new born entrants. Second, firms are bigger. The average sizes of all categories of firms are roughly twice the 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%, both for all the treated data and our sample. This estimate exactly matches the main estimate by Brandt, Van Biesebroeck, and Zhang (2012).

4.4 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 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. On average, firms are young and 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[45] 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 (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 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, 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. It tends to show moderately negative average values and great intraindustry dispersion.

[45] 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.

5. An empirical specification to estimate cost and product advantages

5.1 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

$$Q_{jt}^{X} = \alpha_0^X \left(\frac{P_{jt}^X}{P_t^X} \right)^{-\eta_X} \exp(z_{jt}^X \alpha_X + \delta_{jt}),$$

$$Q_{jt}^{D} = \alpha_0^D \left(\frac{P_{jt}^D}{P_t^D} \right)^{-\eta_D} \exp(z_{jt}^D \alpha_D + \delta_{jt}). \tag{9}$$

The terms α_0^X and α_0^D are constants, η_X and η_D are common industry elasticities, and P_t^X, and P_t^D industry price indices.[46]

The firm's demand is shifted by two components in each market. The first component is the impact of a vector of observables z_{jt}^I.[47] The second component is the idiosyncratic unobservable δ_{jt} representing the unexplained level of advantages of the product.[48] We model δ_{jt} as firm specific, persistent over time and embodying unexpected shocks (see below). Two firms with similar prices and relevant 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

[46] Expressions (9) in logs coincide with first order approximations to any demand. We further discuss this specification in section A3 of the Online Appendix.

[47] Some shifters may be endogenously determined by the firm in the short run. This is likely to happen with sales effort. Let z_{jt}^I represent the log of expenditures on advertising and promotion in market I and suppose that the firm optimally sets P_{jt}^I and $\exp(z_{jt}^I)$. The Dorfman and Steiner (1954) condition for optimal determination of $\exp(z_{jt}^I)$ gives $\frac{\exp(z_{jt}^I)}{R_{jt}} = \frac{\alpha_I}{\eta_I}$, which can be also written as $z_{jt}^I = \frac{1}{1-\alpha_I}(\ln \alpha_0^I + \ln \frac{\alpha_I}{\eta_I} - (\eta_I - 1)p_{jt}^I + \delta_{jt})$, where p_{jt}^I 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.

[48] The terms $z_{jt}^X \alpha_X + \delta_{jt}$ and $z_{jt}^D \alpha_D + \delta_{jt}$ 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 $P_{jt}^D = P_t^D \left(Q_{jt}^D / \alpha_0^D \right)^{-\frac{1}{\eta_D}} \exp((z_{jt}^D \alpha_D + \delta_{jt})/\eta_D)$. 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.

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

5.2 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.[50] 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 $MC_{jt} = \overline{MC}_{jt} \exp(-\omega_{jt})$.

5.3 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 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 the price and output choices, the firm determines the variable input quantities

[49] 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.

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

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.[51]

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. But, in estimations, we control for the predictable part of the unobservables δ_{jt} and ω_{jt}. This limits endogeneity to the variables that are chosen after the realization of the unpredictable part of δ_{jt} and ω_{jt} (we discuss later which ones).

5.4 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 $e_{jt} = \ln(1 + \frac{W_{jt}L_{jt}}{W_{jt}L_{jt}+P_{Mt}M_{jt}}\Delta_{jt})$.

This equation describes the log of revenue over variable cost, or price-average cost margin of the firm (PACM),[52] as the result of the domestic markup multiplied 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 $E(S_{jt}^X e_{jt}) = 0$, the elasticities of demand up to parameter ν can be estimated by NLS[53] If S_{jt}^X and e_{jt} are correlated, identification

[51] For the dynamic framework that justifies this shadow price of labor see Doraszelski and Jaumandreu (2013, 2018).

[52] $\ln \frac{R_{jt}}{C_{jt}} = \ln(1 + \frac{R_{jt}-C_{jt}}{C_{jt}}) \simeq \frac{P_{jt}Q_{jt}-AC_{jt}Q_{jt}}{AC_{jt}Q_{jt}} = \frac{P_{jt}-AC_{jt}}{AC_{jt}} = PACM_{jt}$ 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 $\pi_{jt} = \frac{PACM_{jt}}{1+PACM_{jt}}$.

[53] This way to estimate elasticities can be related to De Loecker and Warzynski (2012) estimation of firm's markups.

is still possible by GMM (we use IV as robustness check). The 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

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

a revenue system where φ^X and φ^D are constants.[54,55]

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 D 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

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

where $d = (\eta_X - 1) - (\eta_D - 1)$ and $p_t = (1/d)(\eta_X p_t^X - \eta_D p_t^D)$.[56]

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. 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 $\overline{mc}_{jt-1} =$

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.

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

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

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

$-\ln \beta_M - \beta_0 + p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1}$. 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[57]

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

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

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 $v_{2jt} = (\eta_D - 1)e_{jt} + (\eta_D - 1)\xi_{jt} + \varepsilon_{jt}$.

5.5 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

[57] 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.

using a rough estimate of the common time index p_t.[58]

6. Estimation

6.1 A system of semiparametric equations.

The model (16) and (17) is a system of semiparametric equations (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 single equation estimation by Doraszelski and Jaumandreu, 2013).

We approximate the nonparametric functions by means of third order polynomials. We are modeling 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*). 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 (we finally enter linearly two alphas).

6.2 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

[58] 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.

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 with some detail 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, 2014). 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. To anticipate the results, 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.

6.3 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.

6.4 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.

6.5 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 asymptotic variance of $\widehat{\theta}$ expressed as

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

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

7. Results

7.1 Estimation.

Estimating functions a and b.

Table 3 reports the estimates of functions a and b carried out to constrain the estimation of the demand elasticities η_D and η_X and the parameter of scale ν (see subsection 5.5). 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 function a 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 substitute goods.[59] 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.[60]

[59] 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).

[60] 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

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 distortions (Hsieh and Klenow, 2009), shocks to demand, transportation costs and time varying margins.[61] Later, we also report the result of relaxing the assumption of common demand elasticities 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.[62] 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.

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.

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 distinguishable 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.

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.

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

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

Our elasticities lie on the right tail of the distribution of elasticities estimated in the IO and trade literature. This can partly reflect the specificity of the Chinese economy and its exports. However, it is also the result 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.[63] The three estimates are sensible, which is unusual for exercises of this type.[64] The key difference that can explain our results 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,*[65] *Age* of the firm (sometimes replaced by *Experience*),[66] *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

[63] The rate of short-run economic profits is a simple transformation of the price-average cost margins: $\pi_{jt} = \frac{R_{jt}-C_{jt}}{R_{jt}} = 1-\nu(\frac{\eta_D-1}{\eta_D}S^D_{jt} + \frac{\eta_X-1}{\eta_X}S^X_{jt}) = 1-\nu+\frac{\nu}{\eta_{jt}}$ where $\frac{1}{\eta_{jt}} = \frac{1}{\eta_D}S^D_{jt} + \frac{1}{\eta_X}S^X_{jt}$ 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.

[64] 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 and 2016, for a recent application). In the other, markups are derived from the first order condition of one factor or several factors together. This is the traditional Solow-based Hall (1990) method, recently revisited 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.

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

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

market. The estimated elasticities give us in fact an interesting check of the internal consistency of the whole estimates and of the price elasticities.[67]

The shifter *Subsidy* (columns 8 and 12) is often non-significant.[68] 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 our assumption of an equal impact across markets of product advantages.[69]

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.[70] Neither the coefficients of the system nor the estimated productivity and product advantages (compared 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 C). 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 6, 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

[67] 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.

[68] 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.

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

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

a function of an unobserved threshold. We test for possible sample selection bias and we conclude that there is no such bias.

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 the constraint $\lambda = 1$, at least for an important part of the industries.

7.2 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, 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 also 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).[71]

Columns (1) to (6) of Table 5 report the quartiles of ω_{jt} and $\delta_{jt}/(\eta_D - 1)$ in the initial and final

[71] 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).

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 willingness 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*).

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. And it is far below of Hsieh and Klenow (2009), that use the same data source that ours for the years 1998-2005. They estimate a standard deviation of their "quantity" TFP of 0.95 and an interquartile range of 1.28. The reason is that they do not try to separate the impact of demand heterogeneity. As they remark, their measure "is a composite of process efficiency and idiosyncratic demand terms coming from quality and variety".[72]

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. The difference may occur because, when we

[72] 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.

measure $\delta_{jt}/(\eta_D - 1)$, we have already subtracted a lot of variation in demand through our included observed shifters.[73]

Forlani, Martin, Mion and Muuls (2017) get standard deviations of their productivity and "product appeal" measures which are greater than ours and quite similar to each other. The model of Hottman, Redding and Weinstein (2016) distinguishes four sources of "demand size" variation: cost advantages, product "appeal", markups and product scope (a role for the number of products associated to the CES assumption). It concludes that "appeal" and product scope, two effects that we implicitly collapse in our product advantages, account for four fifths of the size effects, something quite extreme that can be an effect of the low estimated elasticities (and hence small price effects).

In summary, our estimated distributions turn out to be sensible and very informative. We get reasonable measurements of persistent productivity and product advantages for broad product differentiated industries, which compare favorably with other measurements in more homogeneous settings or other trials to separate the advantages with heavy parametrization. Our results generalize 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. We are able, in addition, to characterize with detail the unrestricted joint distribution of cost and product advantages as follows.[74]

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 an important finding of our exercise. This says that firms that possess unobserved cost advantages tend to have weak unobserved product 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 indicates that productivity is positively

[73] 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.

[74] 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. This is also an identification assumption in Hottman, Redding and Weinstein (2014 and 2016).

associated to higher observed costs. For example, to higher wages of more skilled workers and greater cost of high quality materials. 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.[75] The conclusion is that 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 is because they 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,...). 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.

Can we say something on the economics of the advantages? Think of the plane of all possible pairs $(\omega_{jt}, \delta_{jt}/(\eta_D - 1))$. If firms are equal in all observed factors (costs, shifters), and able to freely choose their combination of advantages from its idiosyncratic endowment in a family of "parallel" balanced concave frontiers, it is easy to see that profit maximizing firms would show positively related amounts of ω_{jt} and $\delta_{jt}/(\eta_D - 1)$.[76] With $\eta_X > \eta_D$ firms tend to prefer more cost than product advantages, but the locus of profit maximization pairs would be a positively sloped line. 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 reach similar levels of profitability with very different combinations of advantages. This suggests two facts. 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 transformation curves) are very heterogeneous (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 outcome points to a completely new aim of research that is particularly policy relevant: the possibilities, incentives and limits of firms' investment in the development of each advantage.

[75] 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.

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

7.3 Changes over time.

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. 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. Additionally, in *Transport* the average product advantages decrease. 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[77]. Columns (3) and (4) show that entry tends to have a negative impact on the individual product advantages but it is small.

Contributions to growth.

Adding the two terms in the unobservables of each one of equations (14) it is possible to compute the total effect of the growth of advantages on revenue growth in each market. Doing this calculation with the estimates of the cost and demand advantages, and for the whole period, one can determine the proportion of growth attributable to each unobservable. We then compute a sales weighted average of the growth contributions across industry revenues, domestic and exports, excluding the industry of transport equipment whose product advantages evolve quite negatively. Despite its heterogeneity, unobservable product advantages explain about 24% of the revenue growth based on productivity and demand heterogeneity.

Decomposition of aggregate changes.

To assess the sources of global changes, we weight productivity and product advantages by revenue

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

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). Columns (5) to (12) report the results.[78] In this decomposition, entrants and exitors contribute to aggregate growth if their productivity or product advantages diverge from the ones of survivors.[79]

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 aggregate 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 possessed 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 but it is also responsible for the negative growth of product advantages.[80] At the end of the period, product advantages are displaced towards the smallest survivors.

[78] We could have presented the decompositions for each one of the markets but patterns across markets turn out to be quite similar.

[79] We consider survivors the firms that are in the sample for the starting year, 1998, and remain until 2008, but also the firms that are present only in 2008 but were already born in 1996. Therefore, the number of survivors is different in 1998 and 2008. There is also 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. 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 during the period. Newly created and shutting down firms dominate the sets of entrants and exitors respectively. Details of the decomposition can be found in Section A6.3 of the Online Appendix along with additional results in Table 6b.

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

Summarizing, there are no dramatic differences in productivity and productivity growth between survivors, entrants and exitors, only the traditional cost disadvantage of entrants which resumes over time. 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 more linked to the largest market shares and product advantages to the smallest survivors.

7.4 Three examples.

In this subsection, 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 separation 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".

P**rivatization and firms efficiency.**

Columns (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 4 for the details of this taxonomy).[81] Columns (1) and (2) show that average productivity growth is systematically higher 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 [82]

To further explore these numbers, we form for each industry a panel subsample of status-changing firms subject to the condition that firms start as state participated and end as private (although

[81] 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.

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

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.[83] We estimate two versions of the model: replacing α_j by a constant and keeping the fixed effects α_j as a form to allow for firm specific levels of productivity and product advantages.

Columns (5) to (8) of Table 7 report the estimates of β. The comparison of the different estimates allows to establish several facts.[84] Privatization during the period first affected the firms with relatively high productivity and, in half of the industries, firms with relatively low product advantages. The high growth of productivity, which characterizes the firms in transition, is however weakly related to privatization itself, as the control by fixed effects shows. The growth of product advantages is in fact even not influenced by privatization.

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 subsection 7.3). Firms coming from the intervention of the state seem more sluggish in the development of product advantages.

Specialization in exports.

In all industries, Chinese exporters show a bimodal distribution of export intensity 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. In what follows 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

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

[84] 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.

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.[85]

Using our estimates of ω_{jt} and $\delta_{jt}/(\eta-1)$, we construct the index of relative cost advantages ica_{jt}.[86] 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. In most industries the youngest firms in the market are the most specialized.

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 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. The 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.

Technological investments and workforce skills.

We use the data on R&D and workforce skills (see Section 4) to investigate the relationship

[85] 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, her 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.

[86] 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.

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 disadvantage. 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.[87] 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 of ω_{jt} and $\delta_{jt}/(\eta-1)$ with the quality of labor. Quality of labor is positively associated with cost advantages in 8 industries. Notice the apparent paradox: firms with higher wages show more cost advantages. What happens is that 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 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 intervened input prices. Without denying such distortions, the above analysis shows a fundamental heterogeneity of firms that implies observed and unobserved

[87] While a negative relationship between ω_{jt} and performing R&D is consistent with the rest of the findings here, it is in partial contradiction with the usual finding that R&D investment stimulates productivity (see Doraszelski and Jaumandreu, 2013). R&D would be more associated with heterogeneous quality than with the effiency in the production of similar varieties, something China and period specific that matches well the small proportion of performing firms. Some necessary caveats are: the incomplete character of the data and the two steps treatment of the relationship.

costs with a counterpart in product advantages.

8. Concluding remarks.

With a sample of Chinese manufacturing firms, which operate domestically and in the export market, we have succeeded in estimating separately 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 have disentangled cost and product advantages without observing output prices, and estimated the unobservables simultaneously as non funcionally-dependent and freely correlated Markov processes. Using its distribution, we have characterized the growth of Chinese manufacturing and described its weaknesses. But the distinction of advantages is of more general interest: it has turned out useful to develop new insights and policy implications in many traditional topics.

This paper also has methodological consequences. Dealing explicitly with demand heterogeneity has turned out to be important for at least two estimation aims. First, to assess properly productivity heterogeneity and to uncover an inverse link between demand heterogeneity and classical (quantity) productivity measures. Second, to avoid biases induced by heterogeneity of demand in the estimation of the production function coefficients and to compute realistic demand elasticities and markups, consistent with sensible short-run returns to scale and profits.

The results can be extended in several ways. First, a model implying different marginal costs in each market can be tested with some more data. Second, the assumption that product advantages have a similar impact in both the domestic and export markets can be relaxed at the cost of giving more structure to the differences between markets, something which seems worth trying. Finally, our model has shown how easy is to relax the assumption of common industry elasticities. A systematic exploration of elasticity variation across firms in the industry would provide additional insights on markup heterogeneity.

There are also two more general pending tasks, that here have been excluded only for simplicity. On the one hand, applying the results to analyze the distribution of the non-exporters that have been excluded from our sample. On the other, to explicitly include R&D and the human capital investments of firms in the processes δ_{jt} and ω_{jt}.[88] This is the way to assess the ultimate determinants of growth and start an exploration of the economics of the advantages.

[88] As Doraszelski and Jaumandreu (2013, 2017) do with R&D.

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: 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 disappearances 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 imputed wage bill and imputed 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 skills. Ratio of the firm wage to the average of wages of all the firms in the industry.

Appendix C: 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.	

Industry	Two-digit industries	Four-digit ind. (No.)
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	

Appendix D: A model with different marginal costs.

Let's assume that

$$\begin{aligned} MC_{jt}^{D} &= \frac{1}{\nu}\exp(\beta_C)h_{jt}^{D}(Q_{jt}^{D})^{(1-\nu)/\nu}\exp(-\omega_{jt}/\nu), \\ MC_{jt}^{X} &= \frac{1}{\nu}\exp(\beta_C)h_{jt}^{X}(Q_{jt}^{X})^{(1-\nu)/\nu}\exp(-\omega_{jt}/\nu), \end{aligned}$$

where $\beta_C = \frac{\beta_0}{\nu} + \ln v + \frac{1}{\nu}\ln \beta_L^{-\beta_L}\beta_M^{-\beta_M}$. The terms h_{jt}^{D} and h_{jt}^{X} are $h_{jt}^{D} = (K_{jt}^{D})^{-\beta_K/\nu}(W_{jt}^{D})^{\beta_L/\nu}(P_{Mjt}^{D})^{\beta_M/\nu}$ and $h_{jt}^{X} = (K_{jt}^{X})^{-\beta_K/\nu}(W_{jt}^{X})^{\beta_L/\nu}(P_{Mjt}^{X})^{\beta_M/\nu}$. 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

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

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

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

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 E: 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 $y_j - m(w_j, \widehat{\gamma})$.

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.

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.

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), "Innovation in the Global Firm," NBER Working Paper 22160.

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, L. Wang and Y. Zhang (2017), "WTO Accession and Performance of Chinese Manufacturing Firms," *American Economic Review,* 107, 2784-2820.

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.

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

Chen, Y., M. Igami, M. Sawada and M.Xiao (2017), "Privatization and Productivity in China," mimeo, University of Arizona.

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.

Dhingra, S. (2013), "Trading Away Wide Brands for Cheap Brands," *American Economic Review*, 103, 2554-2584.

Dhingra, S. and J. Morrow (2018), "Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity," forthcoming *Journal of Political Economy.*

Dhyne, E., A.Petrin and F. Warzynski (2012), "Prices, Markups and Quality at the Firm-Product Level," mimeo, Aarhus University.

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 (2018), "Measuring the Bias of Technological Change," forthcoming *Journal of Political Economy.*

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.

Eslava, M. and J. Haltiwanger (2017), "The Drivers of Life-Cycle Growth of Manufacturing Plants," mimeo, University of Maryland.

Fan, H., Y. Li and S.R. Yeaple (2015), "Trade Liberalization, Quality, and Export Prices," *Review of Economics and Statistics,* 97, 1033-1051.

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.

Flach, L. and M. Irlacher (2018), "Product versus Process: Innovation Strategies of Multiproduct Firms," *American Economic Journal: Microeconomics*, 236-277.

Forlani, E., R. Martin, G. Mion and M. Muuls (2017), "Unraveling Firms: Demands, Productivity and Markups Heterogeneity," Discussion Paper 11058, CEPR.

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 (2017), "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.

Grieco, P. and R. McDevitt (2017), "Productivity and Quality in Health Care: Evidence from the Dyalisis Industry," *Review of Economic Studies,* 84, 1071-1105.

Grieco, P., S. Li and H. Zhang (2017), "Input Prices, Productivity and Trade Dynamics: Long-run Effects of Liberalization on Chinese Paint Manufacturers," mimeo, Pennsylvania State University.

Guillard, C., J. Jaumandreu and J. Olivari (2018), "Endogenous Productivity and Unobserved Prices," mimeo, Boston University.

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.

Haltiwanger, J., R. Kulick and C. Syverson (2017), "MIsallocation Measures: The Distortion that Ate the Residual," mimeo, University of Maryland.

Hottman, C., S. Redding and D. Weinstein (2016), "Quantifying the Sources of Firm Heterogeneity," *Quarterly Journal of Economics*, 1291-1364.

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

Hsieh, C.T. and Z. Song (2015), "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, Spring, 295-366.

Jaumandreu, J. and S. Lin (2017), "Pricing Under Innovation: Evidence from Spanish Manufacturing Firms," mimeo, Boston University and Xiamen University.

Jaumandreu, J. and J. Mairesse (2017), "Disentangling the Effects of Process and Product Innovation on Cost and Demand," *Economics of Innovation and New Technology*, 26, 150-167.

Jaumandreu, J. and H. Yin (2018), "Comparing Productivity when Products and Inputs Differ in Quality: China Manufacturing Growth 1998-2013," mimeo, Boston University.

Katayama, H., S. Lu and J. Tybout (2009), "Why Plant-Level Productivity Studies Are Often Misleading, and an Alternative Approach to Inference," *International Journal of Industrial Organization*, 27, 403-413.

Kee, H. and H. Tang (2016), "Domestic Value Added in Exports: Theory and Firm Evidence from China," *American Economic Review,* 106, 1402-1436.

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.

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 11389.

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.

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 (2017), "Estimating Dynamic R&D Choice: An Analysis of Cost and Long-Run Benefits," *Rand journal of Economics,* 48, 409-437.

Pozzi, A. and F. Schivardi (2016), "Demand or Productivity: What Determines Firm Growth?," *Rand Journal of Economics*, 47, 608-630.

Roberts, M., D.Y. Xu, X. Fan and S. Zhang (2018), "The Role of Firm Factors in Demand, Cost, and Export Market Selection for Chinese Footwear Producers," forthcoming *Review of Economic Studies*.

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

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.

Zhelobodko, E., S. Kokovin, M. Parenti and J-F. Thisse (2012), "Monopolistic Competition: Beyond the Constant Elasticity of Substitution," *Econometrica,* 2765-2784.

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 ω and δ	Correl. ω with $\overline{mc}$	Correl. $\frac{\delta}{\eta_D-1}$ with mc	Correlation of profits with	
	0.25	0.50	0.75	0.25	0.50	0.75	1998	2008	1998-08				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.

			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]			
Industry	$\Delta\omega$	$\frac{\Delta\delta}{(\eta-1)}$ [a]	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.

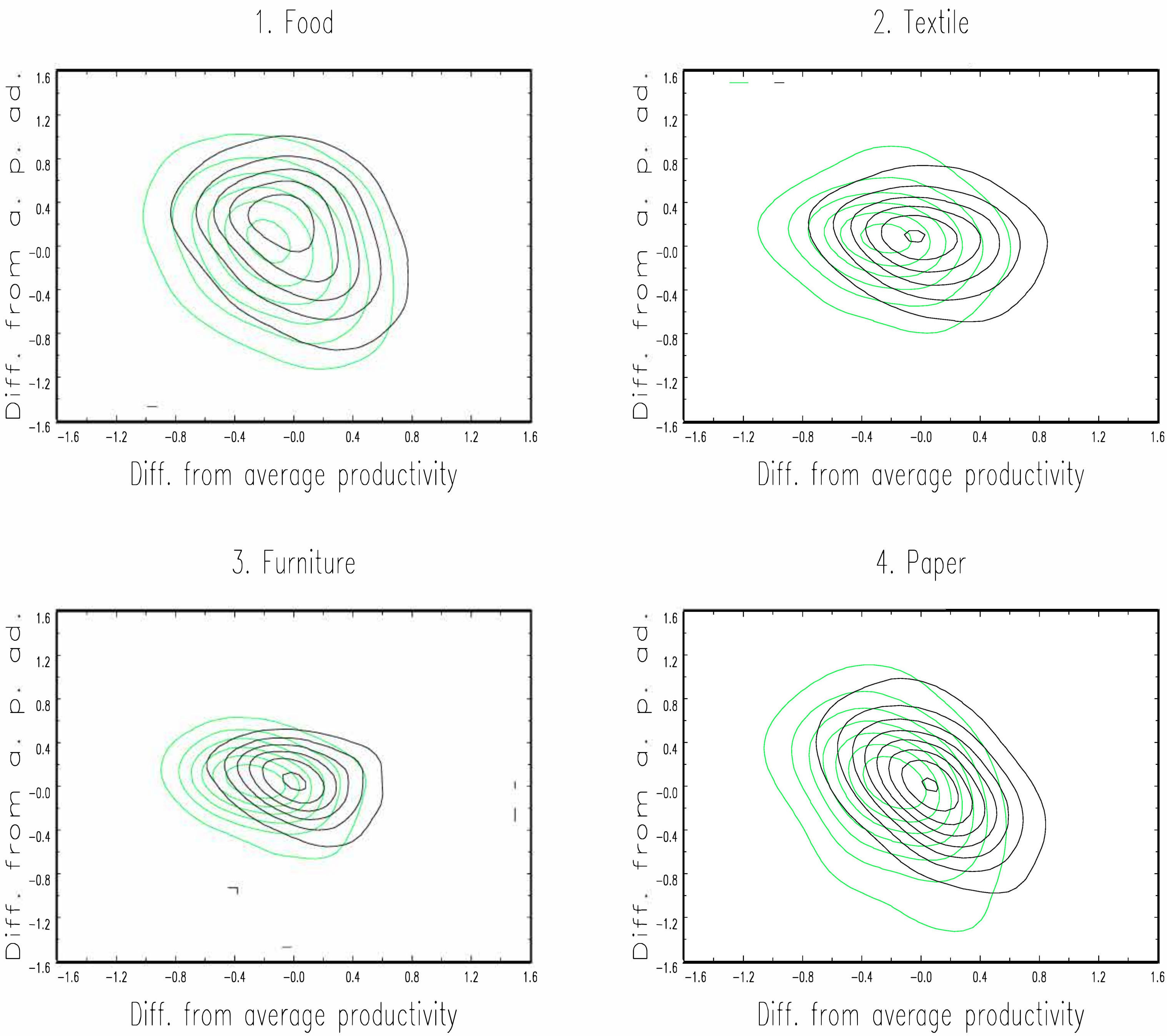
Industry	Observed distribution of export intensity (ei)		Marginal effects in the regression of export intensity			Predicted distribution of export intensity (ei)	
	$P(ei \leq 0.2)$	$P(ei \geq 0.8)$	Relative cost advantage (s. d.)	Experience (s. d)	R^2	$P(ei \leq 0.2)$	$P(ei \geq 0.8)$
	(1)	(2)	(3)	(4)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

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

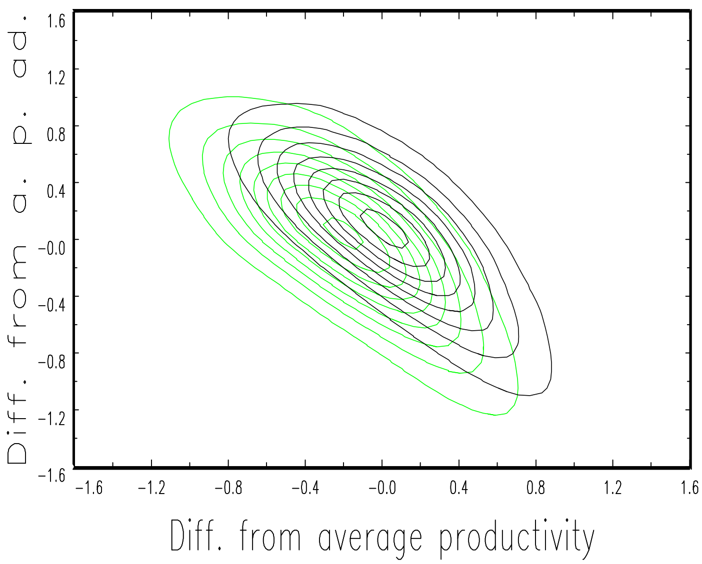
Industry	$R\&D$ investment[a]						Workforce skills	
	ω			δ				
	No R&D	R&D	$Corr(\omega, \frac{R\&D}{R})^b$	No R&D	R&D	$Corr(\delta, \frac{R\&D}{R})^b$	$Corr(\omega, qual)$	$Corr(\delta, qual)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	0.025	-0.050	-0.049	-0.001	0.298	-0.078	0.159	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.

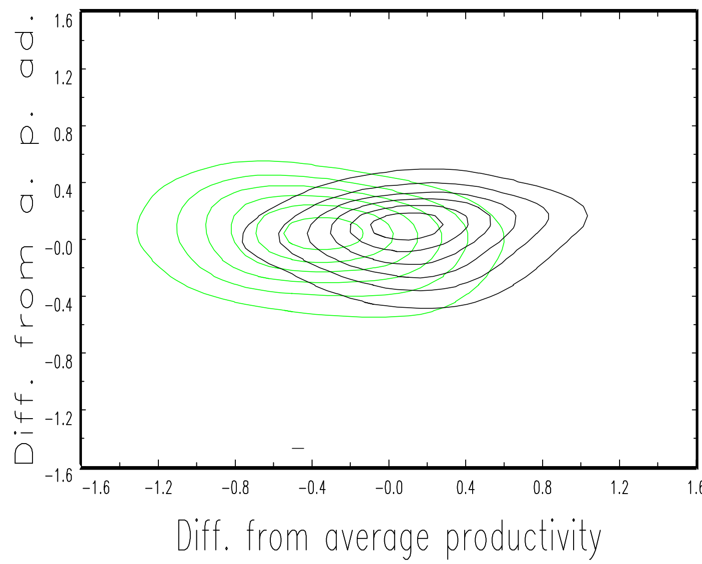
Figure 1: Joint density of ω and $\delta/(\eta_D\text{-}1)$
Change from 1998–2000 (light) to 2005–2008



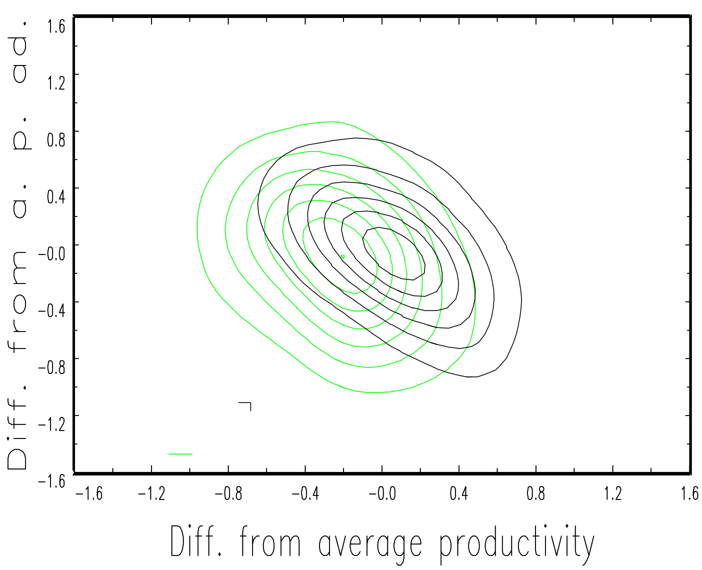
5. Chemical



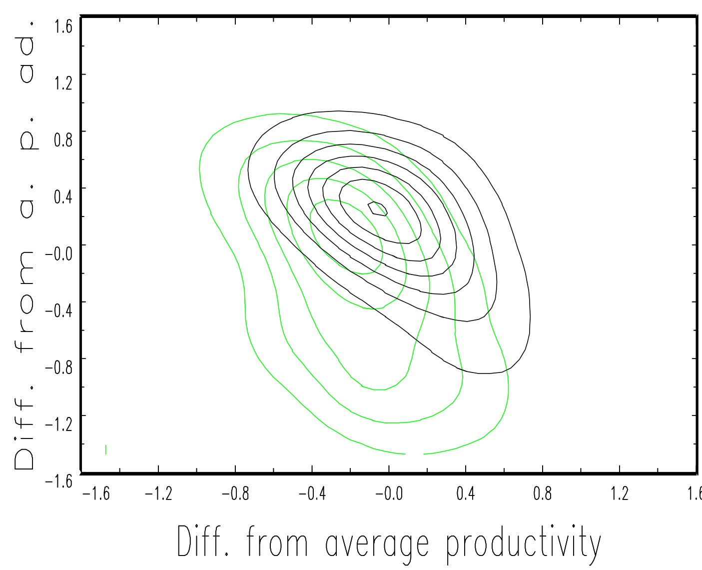
6. Non-met. minerals



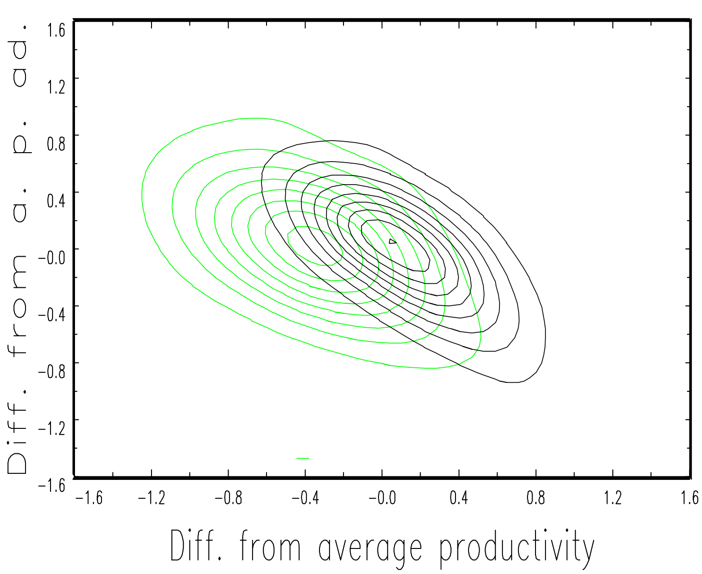
7. Metal products



8. Machinery



9. Transport



10. Electronics

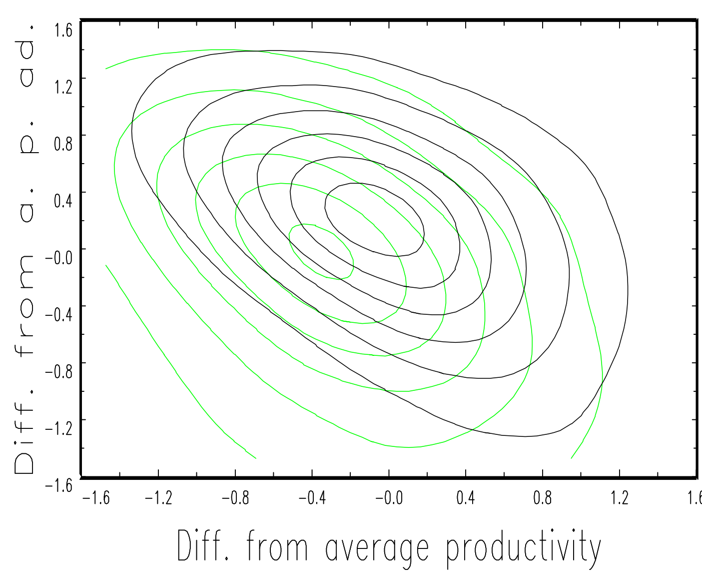


Figure 2: Changes in the mean of ω and $\delta/(\eta_D-1)$

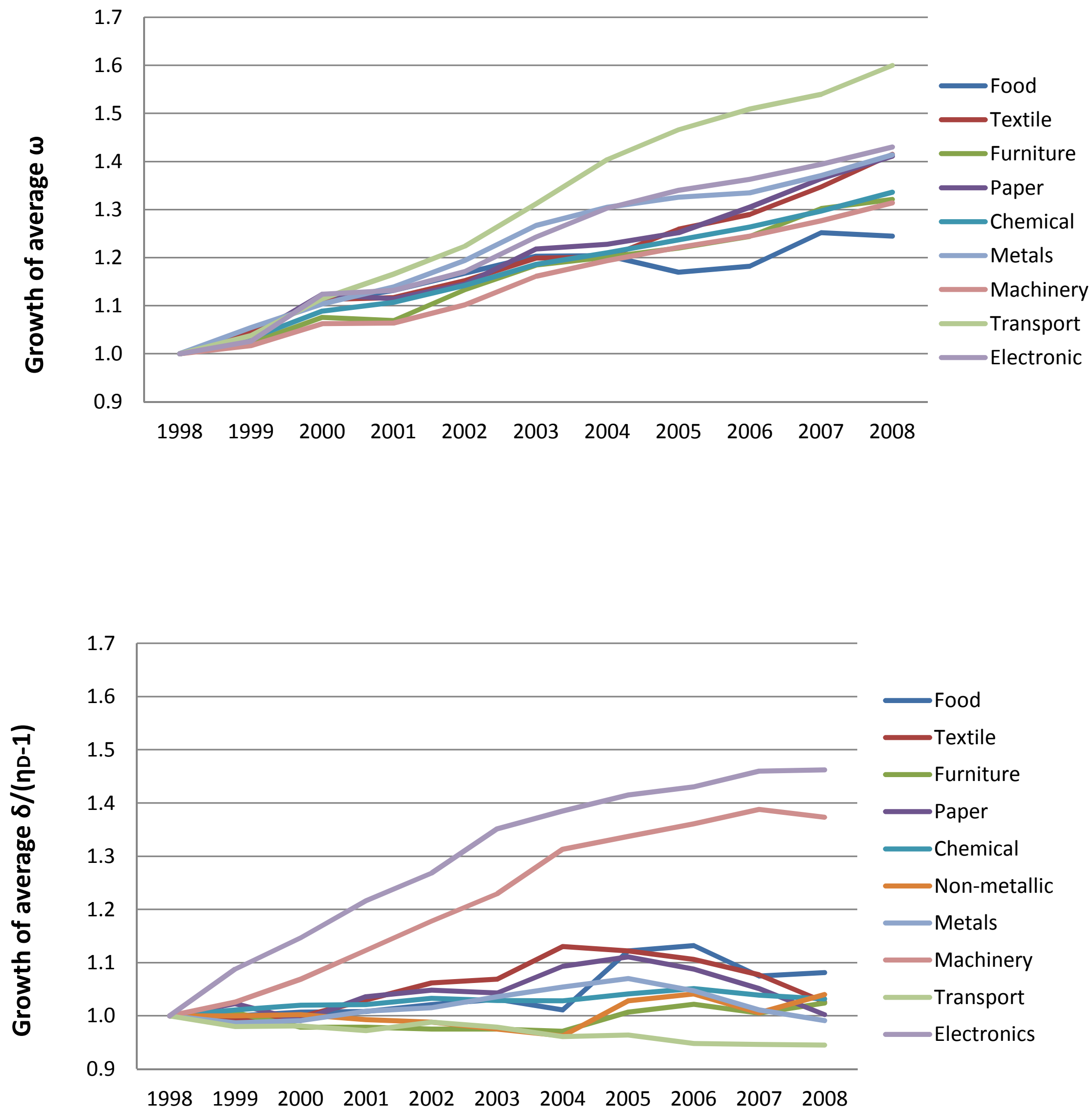


Figure 3: Predicting export intensities

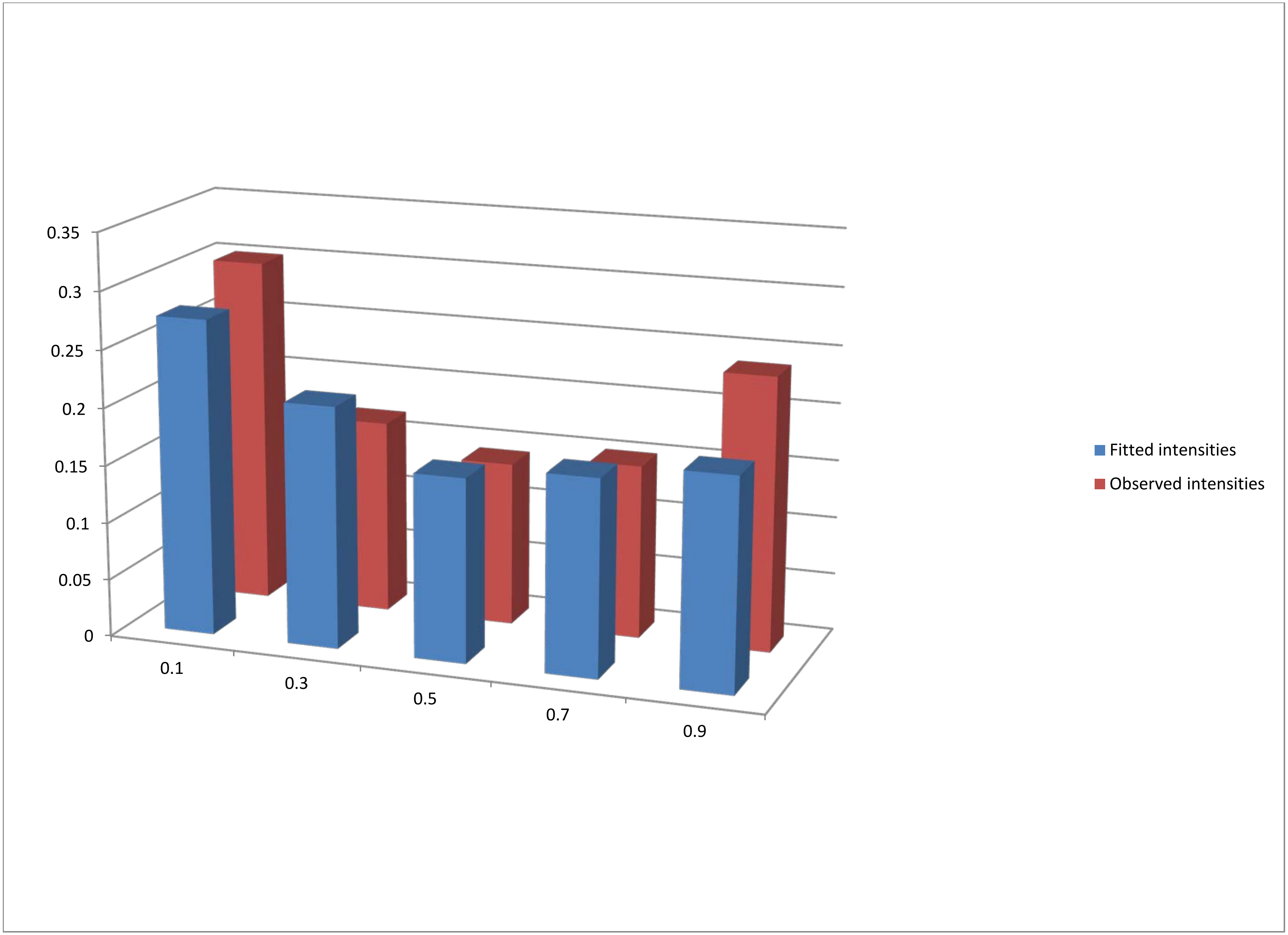


Figure 4: Advantages, R&D and labor quality.