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

Jordi Jaumandreu[†] Heng Yin[‡]

This version: May 2014

Abstract

We estimate the unobservable demand and cost advantages of firms. We call a product advantage the extent to which a firm sells more units than its competitors at a given price. The most important cost advantage is the unobservable productivity that allows the firm to produce at a different unit cost than its competitors. We do not know the precise source of the unobservable demand advantages as we do not know it for productivity. But we show that using firm-level data on their domestic and export markets these advantages can be identified and estimated. Then we characterize the spectacular growth of China exports and industry re-structuring, with data from 1998 to 2008. Entrant and young firms show cost but no demand advantages, older firms demand advantages, and state owned firms less producticity. Demand advantages are associated to skilled labor and higher pay. R&D performing firms show great demand advantages but at the expense of sharp cost disadvantages. The joint distribution of cost and product advantages shows a strong negative correlation between both advantages. Firms tend to compete either in cost or in product attractiveness. Exports are strongly linked to cost advantages but there is an important and growing fraction linked to technologically developed products.

*We thank Li Shigang, Liu Di and Urbashee Paul for help and M. Arellano, U. Doraszelski, J.C. Farinas, I. Fernandez-Val, M. Manove, Shengyu Li, Shuheng Lin, M. Roberts, M. Rysman and A.Pakes for comments.

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

‡Beijing Nornal University, Dep. of Economics, Business School, 19 Xinjiekouwai Dajie, Haidian, Beijing 100088, PRC. E-mail: yheng@bnu.edu.cn.

1. Introduction

Firms, even for narrowly defined industries, sell products that are different and exhibit different production costs. In addition, they typically set different prices. As a result of their actions some firms keep stable market shares, other just struggle to enlarge theirs after entering, and other exit business after being unable to match other's sales or costs. Only a small part of this heterogeneity can be explained by means of observable factors. Firms show "cost advantages" and "product advantages," from which a big part have an origin that cannot be precisely determined by the econometrician.[1] The most important cost advantage is the unobservable relative productivity that determines the ability of the firm to produce at a different unit cost than its competitors. Similarly, we are going to call unobserved product advantage the degree of product attractiveness and market penetration that drives the possibility for the firm of selling more units than its competitors at a given price (and observed demand advantages). This paper develops a model to estimate the distribution of unobserved cost and demand advantages of firms from revenue and input data.

The most important explanation for persistent unobserved demand advantages of firms is product differentiation (vertical, horizontal or any combination). Other demand advantages are generated by the more observable sales effort and customers exposure to advertising or geographical location. There is also an important dynamic dimension of demand advantages: firms are unlikely to suddenly develop them upon entry, and some firms may exit only after a long period of advantage setbacks. The differences in the level of cost because unobserved productivity among firms is the other key dimension that determines firms'

[1] Trade economists have seen international competition from this point of view at least since Helpman and Krugman (1985, 1989). Recent empirical contributions in this field try to measure aspects as the role of efficiency (Eaton, Kortum and Kramarz, 2011) or the impact of quality (Crozet, Head and Mayer, 2012) in the "anatomy" of trade. Sutton (2001, 2007) has built a model in which industrial development is linked to specific (changing) combinations of cost and quality of products.

shares in domestic as well as export markets. Cost levels of firms are both the result of firm product choices and the pursuit of higher levels of efficiency. Innovative activities of the firms, resulting from R&D investments, shape both products and efficiency.

There are at least three different reasons to care about the separate identification and estimation of unobservable cost and product advantages. First, without this separation we can mistake the impact of the innovation activities of firms. As this paper finds, firms investing in R&D may show important unobservable product advantages and, at the same time, unobservable cost disadvantages. A global assessment based only on cost advantages, for example, had ended by understating the impact of R&D and missing its important role in the growth of Chinese firms. Second, estimating firm-level demands without specifying unobserved persistent product advantages is bounded to produce biased estimates of the demand parameters and of the own production function parameters. This is particularly important because estimating the demand relationship is hardly avoidable when there are no available firm-level prices (see below). The third reason is that the policy instruments to stimulate demand advantages are different from the instruments to promote efficiency. As again this paper shows, unobservable product advantages are linked, for example, to the development of a skilled labor force with higher wages. To the extent that these advantages may be considered endogenous, the knowledge of this relationship is important to inform public policy. Although this paper still doesn't deal with the endogeneity of the advantages under firm action and government policies, we believe that it is the next natural step of this research.

The bulk of economists' effort has been dedicated to the measurement of unobservable cost advantages. The measurement of productivity and its determinants at the firm-level has a long tradition, at least since Griliches (1979). See Syverson (2011) for a review that stresses the amount of heterogeneity in productivity found empirically. Recently, a strain of literature on the estimation of production functions in the presence of unobserved productivity has allowed to recover the distribution of productivity across firms and time. The basic idea is to replace unobserved productivity by a function of observable variables, specifically by inverting the demand for a factor. This is the exercise pioneered by Olley and Pakes

(1996). Contributions to this literature are Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2006) and Wooldridge (2009). This way of estimating productivity constitutes today one of the most applied methods and there is a huge list of applications. Doraszelski and Jaumandreu (2013) have proposed a variant to measure productivity impacted by R&D or "endogenous productivity".

In contrast, only a few papers have tried to estimate unobservable heterogeneous demand advantages outside very specific sectors.[2] The main reason is probably the lack of suitable data and methods. Part of the problem is the multiproduct nature of the firms, which requires complex indices to transform the value of sales into quantity indices.

The first paper that proposed to integrate the estimation of firm-level production and demand functions, Klette and Griliches (1996), wanted to address a frequent mismeasurement problem: the absence of firm-level prices.[3] It rightly pointed out that, without firm-level prices, the researcher is estimating a mix of the production and demand relationships that could bias the production function coefficients. The simultaneous estimation of a demand relationship, in which the production is nested, allows in principle to circumvent the problem of the unobserved prices. See de Loecker (2011) for a recent application. These papers do not allow, however, for more than elementary forms of firm unobserved heterogeneity on the demand side. Unfortunately, this risks to produce so serious biases as it wants to correct for.

A few recent works have addressed more demand heterogeneity measurement in broad samples. Foster, Haltiwanger and Syverson (2008) show how demand and firm's efficiency

[2]Demand estimation since Berry (1994) and Berry, Levinsohn and Pakes (1995) has richly used the discrete choice framework to explain the product shares in specific markets. The price, together with the set of observable characteristics plus the unobservable term, fully describe the observed shares and, in this sense, the demand advantage of each product has been completely described. But this is hardly applicable when we have firms for which we only observe their sales in a broad industry, no product quantities are directly comparable, and only an approximate price index is available in the best of the cases and no price at all in many others.

[3]Katayama, Lu and Tybaut (2009) is a paper stressing the empirical difficulties created by the absence of suitable prices.

have a separate impact on the pricing, quantity sold and market participation of the firms in a sample of quasi-homogeneous good US industries. Jaumandreu and Mairesse (2010) specify firm-level demand and cost functions for a sample of Spanish firms with price data available. Process innovations affect the cost advantages and advertising and product innovations affect the demand advantages. Aw, Roberts and Xu (2011) use a firm-level demand relationship and cost function subject to idiosyncratic shocks, with marginal cost endogenous with respect to R&D, to assess the impact of R&D investments and exporting on the dynamics of productivity.[4] Pozzi and Schivardi (2012) write down a model with productivity and demand unobservable heterogenous terms and compare its theoretical predictions with the results of estimates for a sample of Italian firms. Petrin and Warzynski (2012) use a product level data set for Denmark to apply the discrete choice approach to estimate unobserved demand heterogeneity, that they label as "quality". All these works stress the importance of demand heterogeneity in addition to productivity heterogeneity and the relevance of its estimation.

In this paper we develop a model that specifies and estimates the same level of unobservable heterogeneity in the firm-level demand and production functions by means of first degree Markov processes. Like the papers that only wanted to address the absence of prices, estimation doesn't need the observation of prices at the firm-level, only revenues and input use. Our setup follows Aw, Roberts and Xu (2011) in that each firm sells its product in the domestic and exports markets. With the observation of revenues in two markets with different price elasticities the unobservable advantages can be identified thanks to the different impact of the efficiency of the firm.[5] Having firm-level price data the model could also be estimated market to market. We first show that unobservable cost and demand advantages are characteristics of the production and demand functions of firms that can be nonparametrically identified in the absence of prices. We then apply a simple parametric

[4] Roberts, Xu, Fan and Zhang (2012) have later developed a model for the exports of Chinese footwear producers that draws on the diversity of destinations, prices set and quantities sold in each destination, by each producer, to infer firm demand and cost advantages.

[5] We are working on the conditions to apply our estimates to measure productivity and demand advantages for the firms operating in only one market.

specification that makes the estimation relatively straightforward.

China extended recent period of industry re-structuring and exports growth provides a unique context to apply the model. For recent characterizations of this period using firm-level data see, for example, the works by Brandt, Van Biesebroeck and Zhang (2012) for productivity, and Tang and Zhang (2012) for exports. Our data cover a crucial part of this period, 1998-2008, and we are working to add more recent years. At least three things converge to make this period extraordinary: the spectacular growth of the exports, in particular after China adhered WTO in 2001; the deep re-structuring of the SOE's (state owned firms) and the spread change of the forms of ownership; the intensity of liberalization and competitive pressures, sharply inducing reallocation of activity and entry and exit.

While we still want to perfection some of the paper estimates and inferences, a rich set of facts has emerged from our preliminary application of the model to the Chinese data. They draw a subtle picture of the underlying forces to the enormous increase of exports and the industry restructuring of the 2000's. Our findings contribute evidence, for example, on all the firm-level heterogeneity dimensions enumerated by Melitz and Redding (2012). Had our analysis been based on the pure productive efficiency of firms, the conclusions would be insufficient and with much less policy implications.

The rest of the paper is organized as follows. In the next section we show that the unobservable cost and demand advantages are characteristics nonparametrically identified in the absence of prices, that is, no particular functional form is in principle needed for their estimation. In the third section we set out our particular empirical parametric specification. The fourth section is dedicated to explain how we estimate the econometric model. Section 5 introduces the data, comments the main facts about Chinese manufacturing during the period 1998-2008 according to them, and describes the sample that we use. Section 6 reports the results of estimation and describes the joint distribution and correlations of the estimated cost and demand advantages. Section 7 concludes with some remarks. There are four appendices. Appendix A discusses the parametric specification of demand, Appendix B defines the employed variables, Appendix C details the treatment of the data, and Appendix D details the correspondence of the industries in this paper to the Chinese two-digit breakdown.

2. Model and identification

Te purpose of this section is presenting the basic model and show how the main characteristics of interest, the demand and cost advantages of the firms, are nonparametrically identified from revenue and input data. We start discussing the model assumptions in relation to one market and then we show how having multimarket firms identifies the advantages.

Let us consider a market of firm j. We assume that the market is monopolistically competitive. By monopolistic competition we understand that each firm faces a downward-sloping demand for its product, makes no profit in a situation of long run equilibrium, and a price change by one firm has only a negligible effect on the demand of any other firm (Tirole, 1989). The demand for its product, at moment t, is

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

where P_{jt} is the price set by the firm, Z_{jt} is a vector of observed demand shifters, and δ_{jt} measures unspecified advantages that have a positive impact on the quantity sold by the firm. Some shifters may be endogenously set by the firm (e.g. the level of sales effort).

The firm produces its product with production function

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

where K_{jt}, L_{jt} and M_{jt} stand for capital, labor and materials respectively, and ω_{jt} measures unspecified advantages that have a positive impact on the production level of the firm. The term ω_{jt} is often called productivity.[6] Let us assume that capital is given and that labor and materials are variable inputs in the short run. Cost minimization implies a marginal

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

cost function that can be written as [7]

$$MC_{jt} = MC(K_{jt}, M_{jt,} W_{jt}, P_{Mjt}, \omega_{jt}) = MC(X_{jt}, \omega_{jt}),$$

where W_{jt} and P_{Mjt} are the wage and price of materials, respectively, and $X_{jt} = \{K_{jt}, M_{jt}, W_{jt}, P_{Mjt}\}$.

We assume that the firm knows its demand and marginal cost functions, and sets simultaneously the price and quantity of output which maximize profits, as well as the quantity of variable inputs necessary to carry out the production of this output. From the equation which characterizes optimal price we can write

$$MR_{jt} = P_{jt} + Q(P_{jt}, Z_{jt}, \delta_{jt}) / \frac{\partial Q}{\partial P_{jt}}(P_{jt}, Z_{jt}, \delta_{jt}) = MR(P_{jt}, Z_{jt}, \delta_{jt}) = MC(X_{jt}, \omega_{jt}),$$

and inverting the marginal revenue function for price we have

$$P_{jt} = MR^{-1}(MC(X_{jt}, \omega_{jt}), Z_{jt}, \delta_{jt}). \tag{2}$$

This price equation is often written in terms of a markup on marginal cost which depends on the (modulus of the) elasticity of demand, that is, $P_{jt} = \frac{\eta_{jt}}{\eta_{jt}-1} MC(X_{jt}, \omega_{jt})$. But, at this stage, we want to avoid this expression. The reason is that, without any restriction, we have $\eta_{jt} = \eta(P_{jt}, Z_{jt}, \delta_{jt})$ and the expression would keep price as an argument in the right hand side of the equality.

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

$$R_{jt} = P_{jt} Q(P_{jt}, Z_{jt}, \delta_{jt}). \tag{3}$$

It is important to define its properties. Under monopolistic competition we expect the implicit elasticity of demand to be greater than unity, and hence revenue to be decreasing in price.[8] If we suppose, as it seems natural, that this elasticity is in addition non-increasing

[7] The cost function has the form $C_{jt} = C(K_{jt}, W_{jt}, P_{Mjt}, Q_{jt}, \omega_{jt})$, and the conditional demand for materials $M_{jt} = \frac{\partial C}{\partial P_M}(K_{jt}, W_{jt}, P_{Mjt}, Q_{jt}, \omega_{jt})$. Taking the derivative of the cost function with respect to output and substituting the inverted materials demand for output in the marginal cost function one gets the expression in the text. Alternatively the expression could be developed in terms of the other variable input L_{jt}. Later in this paper we also develop and use other ways to express marginal cost.

[8] $\frac{\partial R_{jt}}{\partial P_{jt}} = Q_{jt} + P_{jt} \frac{\partial Q_{jt}}{\partial P_{jt}} = (1 - \eta_{jt}) Q_{jt} < 0.$

in the demand shifters and the demand advantage, revenue is increasing, as demand, in Z_{jt} and δ_{jt}.[9]

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

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

This equation becomes useful when we cannot observe prices and hence we cannot work with equation (3). It says that revenue depends on the observable factors which determine marginal cost and shifts in demand, and on the two unobservables representing the demand and cost advantages of the firm. Even if we are able to perfectly measure all the observable variables, we can only recover from equation (4) this combination of advantages. Recovering the combination might be interesting on its own, and later we will justify this interest. But here our main objective is to show how ω_{jt} and δ_{jt}, the demand and cost advantages of firms, can be separately nonparametrically identified from revenue data.

What we need is to observe the firm selling the product in (at least) two markets. Suppose, for example, that the firm sells the same product in the exports and domestic market. If this happens we have two revenue functions:

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

The revenue functions, especially the implicit price elasticities, and even the shifters, are presumably different in each equation, but the fact that the product sold is the same in both markets implies that the same unobserved advantages enter to explain both revenues.

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

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

[9] We have, for example, $\frac{\partial R_{jt}}{\partial \delta_{jt}} = \frac{\partial(\frac{\eta_{jt}}{\eta_{jt}-1})}{\partial \delta_{jt}} MC_{jt} Q_{jt} + P_{jt} \frac{\partial Q_{jt}}{\partial \delta_{jt}} = -\frac{1}{(\eta_{jt}-1)^2} \frac{\partial \eta_{jt}}{\partial \delta_{jt}} + P_{jt} \frac{\partial Q_{jt}}{\partial \delta_{jt}} > 0.$

The importance of this inversion is double: it allows us to set an estimable model controlling for persistent unobservables in terms of observables, and gives us the way to back out the advantages from revenue, input quantities and shifters.

What do we need for the system to be invertible? Standard conditions for the invertibility of a nonlinear system of equations (see, for example, Gale and Nikaido, 1965) must hold. Writing equations (5) as the system of equations $R^X(\cdot) - R^X_{jt} = 0$ and $R^D(\cdot) - R^D_{jt} = 0$, and assuming continuous partial derivatives of the revenue functions, invertibility is governed by conditions on the Jacobian

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

For example, a sufficient condition for invertibility (in a broad enough domain) is that no principal minor of the Jacobian vanishes (Theorem 7 of Gale and Nikaido).

To see that invertibility holds in a broad range of conditions it is convenient to think of the corresponding matrix of semi-elasticities

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

We expect all four revenue semielasticities to be positive. In addition, under different elasticities of demand with respect to price, the ratio of revenue semielasticities with respect to ω_{jt} is expected to be different from the ratio of semielasticities with respect to δ_{jt}.[10] Therefore, the scalar minors of this second matrix are positive and the other minor, the determinant of the matrix, does not vanish. Multiplying the first row by R^X_{jt} and the second by R^D_{jt} we get the Jacobian. This kind of transformation preserves the nonvanishing property of the minors and hence the Jacobian meets the invertibility condition.

The intuitive reason by which ω_{jt} and δ_{jt} can be identified is that they have different effects in each market. Cost advantages operate through the price set in each market and,

[10] For a given firm and moment of time we have, for example, $\frac{1}{R^X}\frac{\partial R^X}{\partial \omega} = \frac{\partial R^X}{\partial P^X}\frac{\partial P^X}{\partial MC}\frac{\partial MC}{\partial \omega} = (\eta^X - 1)\left|\frac{1}{MC}\frac{\partial MC}{\partial \omega}\right|$.The ratio of semielasticities with respect to ω can hence be written as $\frac{1}{R^X}\frac{\partial R^X}{\partial \omega} / \frac{1}{R^D}\frac{\partial R^D}{\partial \omega} = (\eta^X - 1)/(\eta^D - 1)$. The ratio of semielasticities with respect to δ is expected to be unity (demand advantages are the same in both markets) or even move in the opposite direction.

as long as the price effects are different because different elasticities, this provides us with a variation in revenues that is enough to identify these advantages. Demand advantages need only to be positive, but they can also have a different impact in each demand.

Cost and demand advantages are likely to be both subject to unexpected shocks and very persistent over time. Joining a practice well established since Olley and Pakes (1996) for modeling unobserved productivity in production functions, we assume that they follow 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 in the level predictable from its value at moment $t-1$ and the unpredictable shocks ξ_{jt} and ε_{jt}. Plugging (6) lagged into (7), and (7) into (5), we have the nonparametric structural econometric model

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

where $g(\cdot)$ and $h(\cdot)$ result from the composition of the unknown functions $q(\cdot)$ and $\omega(\cdot)$, on the one hand, and $s(\cdot)$ and $\delta(\cdot)$ on the other, and vector S_{jt-1} is $S_{jt-1} = \{X_{jt-1}, Z^X_{jt-1}, Z^D_{jt-1}, R^X_{jt-1}, R^D_{jt-1}\}$.

Equations (8) form a system of equations which contain a few variables that are or maybe correlated with the disturbances ξ_{jt} and ε_{jt}, other variables that can be assumed independent, and both disturbances are present in both equations. In the set of variables X_{jt}, only the variable input M_{jt} is correlated with the disturbances ξ_{jt} and ε_{jt} because is chosen when the current values of productivity are known. In the set of shifters Z_{jt}, we will only have a shifter correlated with the disturbances if we assume that its value has been chosen after the innovative productivity shocks have been known. Capital is exogenous because is assumed to be chosen in advance, input prices because are given in competitive markets, and the set of variables in S_{jt-1} because they are determined before the innovative productivity shocks are known.

Nonparametric identification of systems of this type is discussed in Matzkin (2007, 2013). In what follows we content ourselves with specifying and estimating a parametric version of the model. It is clear, however, that the advantages that we want to characterize are identified under much more general specifications.

3. An empirical specification to estimate cost and product advantages

In this section we set out our empirical specification. First we detail the parametric functions that we use and derive the relevant system of equations, leaving unspecified the observable part of marginal cost. Then we briefly discuss the options about how to treat marginal cost and detail the system of estimating equations.

Before starting, however, is useful to summarize the information that is available in our data set. We do not observe the prices and chosen quantities but we do observe the resulting revenues R_{jt}^{X} and R_{jt}^{D}. We also observe K_{jt}. On the other hand, we observe the wage bill and the cost of materials and hence variable cost C_{jt}. In addition we observe the number of workers L_{jt} and a common deflator for the price of materials P_{Mt}. Using L_{jt} and P_{Mt} we can therefore get a firm-level average wage W_{jt} and an index of the quantity of materials M_{jt}. So we observe all the components of variable cost $C_{jt} = W_{jt}L_{jt} + P_{Mt}M_{jt}$.

As in Aw, Roberts and Xu (2011), firms produce a single output (have a unique marginal cost) that sell in a domestic and an export market. Both the domestic (D) and export (X) markets are monopolistically competitive. The demands for the product of firm j are:

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

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

The terms α_0^{X} and α_0^{D} are constants; η_X and η_D are common industry elasticities; $Q_{jt}^{X}, P_{jt}^{X}, Q_{jt}^{D}$ and P_{jt}^{D} the quantities and prices determined by the firm; and the exponents pickup demand shifters. See Appendix A for further discussion on this specification. The expressions in the exponents, $z_{jt}^{X}\alpha_X + \lambda\delta_{jt}$ and $z_{jt}^{D}\alpha_D + \delta_{jt}$, measure the factors, observed and unobserved, by which the product of firm j is valued differently from other products by consumers. These factors tell us how much additional quantity of the product of firm j is bought by consumers when its price is the same that the price of a hypothetical rival for which these demand terms are equal to zero. Notice that we could also write, for example, $P_{jt}^D = \left(Q_{jt}^D/\alpha_0^D\right)^{-\frac{1}{\eta_D}} \exp((z_{jt}^D\alpha_D + \delta_{jt})/\eta_D)$. So, alternatively, these expressions scaled by the corresponding η may be read as describing how much more consumers are willing to pay for the same quantity.

The "demand advantages" of firms hence consist of two components. The first component reflects the impact of a vector of observables z, possibly different in each market. For example, the $z's$ may measure how accessible are global markets from the area of location of the firm or how much amounts its current expenditure in sales promotion.[11] The second component is an idiosyncratic unobservable variable representing the level of "attractiveness" and penetration of the product. This term picks up the advantages (disadvantages) that we are not able to observe and measure (quality, design, level of adequacy to the contractor..., but also being a local or recently introduced good). Notice that the specification allows us to read δ_{jt} ($\times 100$) as approximate percentage points of advantage. Importantly, by means of parameter λ we may allow the unobserved demand advantages of the same product to have a different impact in the export and domestic markets (we normalize the impact to unity in the domestic market).

On the production side, let us assume that the firm has a Cobb-Douglas production function

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

where ω_{jt} represents Hicks neutral productivity.[12] We assume that K_{jt} is given and that

[11] Some shifters may be endogenously determined. Suppose, for example, that the firm sets optimally P_{jt} and z_{jt}^D incurring a cost of one money unit per unit of z. A Dorfman and Steiner (1954) type of condition for z gives

$$z_{jt}^D = \frac{1}{\alpha_D}(\ln \frac{\eta_D}{\alpha_D} + (\eta_D - 1)p_{jt}^D - \delta_{jt}),$$

where p_{jt}^D stands for the log of price. Had we prices, these conditions could be exploited in an Olley and Pakes (1996) type of procedure to estimate demand advantages.

[12] Notice that our demand advantages are also "neutral" with respect to price and shifters, that is, they increase their marginal effects in the same extent.

the firm freely chooses in the short-run L_{jt} and M_{jt}. This implies, as shown in Section 2, a marginal cost function that depends on K_{jt}, the prices of variable inputs, output and ω_{jt}. Marginal cost is directly unobservable, but it can be expressed in many ways in terms of observable variables and productivity ω_{jt}. We discuss later our specific choices, let us write for the moment

$$MC_{jt} = \overline{MC}_{jt} \exp(-\omega_{jt}),$$

where $\overline{MC}_{jt}$ represents the part observable up to a set of parameters.

The firm knows the demands and the marginal cost function and chooses $P_{jt}^X, Q_{jt}^X, P_{jt}^D$ and Q_{jt}^D which maximize short-run profits. Notice that this implies that the firm knows δ_{jt}, ω_{jt}, the value of the shifters and the price of the inputs.[13] Given the price and output choices, variable inputs are determined according to their conditional demand functions and hence are correlated with the current values of δ_{jt} and ω_{jt}.

The first order conditions for prices are:

$$\begin{aligned} P_{jt}^X(1 - \frac{1}{\eta_X}) &= \overline{MC}_{jt} \exp(-\omega_{jt}), \\ P_{jt}^D(1 - \frac{1}{\eta_D}) &= \overline{MC}_{jt} \exp(-\omega_{jt}). \end{aligned} \quad (10)$$

Rewriting the demands (9) in terms of revenues, replacing prices by their optimal choice according to (10) and taking logs (that we will represent by lowercase letters) we have the system

$$\begin{aligned} r_{jt}^X &= \varphi^X - (\eta_X - 1)\overline{mc}_{jt} + z_{jt}^X \alpha_X + (\eta_X - 1)\omega_{jt} + \lambda \delta_{jt} \\ r_{jt}^D &= \varphi^D - (\eta_D - 1)\overline{mc}_{jt} + z_{jt}^D \alpha_D + (\eta_D - 1)\omega_{jt} + \delta_{jt}. \end{aligned} \quad (11)$$

where φ^X and φ^D are constants.[14]

These equations show how revenue in each market depends on the observed part of marginal cost, the corresponding observed demand advantages, the unobserved cost advantage ω_{jt} and the unobserved demand advantage δ_{jt}.

[13] Our model uses only equlibrium conditions which state how optimal prices are related to marginal cost. The determination of price and quantities needs, however, the knowledge of the demand and cost functions.

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

Equations (11) can be solved for ω_{jt} and δ_{jt}. The solution gives

$$
\begin{aligned}
\omega_{jt} &= \gamma^X + (1/d)(r_{jt}^X - z_{jt}^X \alpha_X) - (\lambda/d)(r_{jt}^D - z_{jt}^D \alpha_D) + \overline{mc}_{jt}, \\
\delta_{jt} &= \gamma^D + ((\eta_X - 1)/d)(r_{jt}^D - z_{jt}^D \alpha_D) - ((\eta_D - 1)/d)(r_{jt}^X - z_{jt}^X \alpha_X),
\end{aligned} \tag{12}
$$

where $d = (\eta_X - 1) - \lambda(\eta_D - 1)$.[15] Notice that, if $\lambda = 1$, the inversion requires different demand elasticities in each market.

Let us now assume that the unobservables follow first order exogenous in-homogeneous Markov processes $\omega_{jt} = q_t + q(\omega_{jt-1}) + \xi_{jt}$ and $\delta_{jt} = s_t + s(\delta_{jt-1}) + \varepsilon_{jt}$, where q_t and s_t represent time effects. On the one hand we plug these laws of motion into equations (11) and, using equations (12) lagged, we replace the unobservables in terms of observables dated at time $t-1$. We get[16]

$$
\begin{aligned}
r_{jt}^X = \; & a_t^X - (\eta_X - 1)\overline{mc}_{jt} + z_{jt}^X \alpha_X \\
& + g_1[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - \lambda(r_{jt-1}^D - z_{jt-1}^D \alpha_D) + d\, \overline{mc}_{jt-1}] \\
& + h_1[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + u_{1jt}
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
r_{jt}^D = \; & a_t^D - (\eta_D - 1)\overline{mc}_{jt} + z_{jt}^D \alpha_D \\
& + g_2[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - \lambda(r_{jt-1}^D - z_{jt-1}^D \alpha_D) + d\, \overline{mc}_{jt-1}] \\
& + h_2[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + u_{2jt},
\end{aligned} \tag{14}
$$

where a_t^X and a_t^D are combinations of a constant and time effects, $g_1(\cdot), h_1(\cdot), g_2(\cdot)$ and $h_2(\cdot)$ are unknown functions, $u_{1jt} = (\eta_X - 1)\xi_{jt} + \lambda\varepsilon_{jt}$ and $u_{2jt} = (\eta_D - 1)\xi_{jt} + \varepsilon_{jt}$. Recall that $d = (\eta_X - 1) - \lambda(\eta_D - 1)$. This is the system of equations to estimate, but we still need to specify the observable part of marginal cost.

We assume that the cost minimizing first order condition for materials holds, so we can

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

[16] 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.

get MC_{jt} from

$$MC_{jt}\beta_M \exp(\beta_0) K_{jt}^{\beta_K} L_{jt}^{\beta_L} M_{jt}^{\beta_M - 1} \exp(\omega_{jt}) = P_{Mt}. \tag{15}$$

We are also willing to assume that the cost minimizing first order condition holds for labor, but here we want to allow for some error. We assume that adding up the conditions corresponding to materials and labor, weighted by the quantities of the inputs, we can also get MC_{jt} (up to an unobservable error) from

$$MC_{jt}(\beta_L + \beta_M) \exp(\beta_0) K_{jt}^{\beta_K} L_{jt}^{\beta_L} M_{jt}^{\beta_M} \exp(\omega_{jt}) = C_{jt} \exp(e_{jt}), \tag{16}$$

where e_{jt} is assumed to be a zero mean error, not correlated over time, but not necessarily uncorrelated with the variable inputs.[17] A sensible interpretation for e_{jt} is that it is picking up adjustment costs which determine deviations from the short run optimal choices. We assume, however, that these costs cancel over time and across firms.

For the $\overline{mc}_{jt}$ that enters directly equations (13) and (14) we will use the expression of equation (16), that is

$$\overline{mc}_{jt} = \gamma_1 + c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt} + e_{jt},$$

where $\gamma_1 = -\ln(\beta_L + \beta_M) - \beta_0$. One part of this expression goes to the constants and the errors of the equations are going to be augmented with $-(\eta_X - 1)e_{jt}$ and $-(\eta_D - 1)e_{jt}$. As regards the $\overline{mc}_{jt-1}$ inside the unknown functions $g_1(\cdot)$ and $g_2(\cdot)$ of equations (13) and (14), we will employ equation (15) lagged, so we have

$$\overline{mc}_{jt-1} = \gamma_2 + p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1 - \beta_M) m_{jt-1},$$

where $\gamma_2 = -\ln \beta_M - \beta_0$.

The resulting estimating system of equations can be written as

[17] Some scholars call this the "share" equation beacuse it can be written as $C_{jt}/MC_{jt}Q_{jt} = (\beta_L + \beta_M)\exp(-e_{jt})$.

$$
\begin{aligned}
r_{jt}^{X} &= b_t^X - (\eta_X - 1)(c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt}) + z_{jt}^X \alpha_X \\
&+g_1[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - \lambda(r_{jt-1}^D - z_{jt-1}^D \alpha_D) \\
&\qquad\qquad +d\ (p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1})] \qquad (17) \\
&+h_1[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + v_{1jt}
\end{aligned}
$$

$$
\begin{aligned}
r_{jt}^{D} &= b_t^D - (\eta_D - 1)(c_{jt} - \beta_K k_{jt} - \beta_L l_{jt} - \beta_M m_{jt}) + z_{jt}^D \alpha_D \\
&+g_2[(r_{jt-1}^X - z_{jt-1}^X \alpha_X) - \lambda(r_{jt-1}^D - z_{jt-1}^D \alpha_D) \\
&\qquad\qquad +d\ (p_{Mt-1} - \beta_K k_{jt-1} - \beta_L l_{jt-1} + (1-\beta_M) m_{jt-1})] \qquad (18) \\
&+h_2[(\eta_X - 1)(r_{jt-1}^D - z_{jt-1}^D \alpha_D) - (\eta_D - 1)(r_{jt-1}^X - z_{jt-1}^X \alpha_X)] + v_{2jt},
\end{aligned}
$$

where b_t^X and b_t^D are combinations of a constant and time effects and we slightly abuse of notation by writing $g_1(\cdot)$ and $g_2(\cdot)$ as being the same as in (13) and (14). The disturbances are $v_{1jt} = -(\eta_X - 1)e_{jt} + u_{1jt}$ and $v_{2jt} = -(\eta_D - 1)e_{jt} + u_{2jt}$.

4. Estimation

The model consisting of (17) and (18) is a system of semiparametric equations, that is, equations that have a linear and a nonparametric part (see, for example, 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 loglinear forms. The disturbances are uncorrelated over time and across firms, but can be freely correlated among them. The system is fully nonlinear in parameters, since there are cross-restrictions between the linear part of the equation and the argument on the nonparametric part, and between the two equations. In fact, the restrictions between the parameters in the linear part and the parameters on the arguments of the nonparametric functions determine the scale of the arguments and contribute to identification (we build on the similar uniequational estimation of Doraszelski and Jaumandreu, 2013).

The parameters of interest are the elasticities of the demands, η_X and η_D; the parameters of the marginal cost function, β_K, β_L and β_M; and the vectors of semielasticities of the shifters α_X and α_D.[18] The regressions that we report below use three shifters for each demand, so there are a total of 11 parameters of interest. Using the estimates of these parameters we can recover the unobserved advantages ω_{jt} and δ_{jt} (up to its average for all firms in the industry) by means of equations (12).

We approximate the nonparametric functions by means of polynomials of order three in their arguments. Since for the moment we are modeling the unobserved advantages as exogenous each function is univariate, and requires only the estimation of three coefficients. In addition to the parameters of theoretical interest, we have then to estimate: two constants, eight time dummies in each equation and twelve coefficients of the polynomials. This implies 30 additional parameters.

Our first estimates convinced us that it was very hard the estimation of all parameters at the same time. So we decided to split the problem in two stages. It the first stage we estimate the value of the elasticities η_X and η_D up to the scale parameter, by the procedure that we explain below. In the second stage, we plug these elasticities into (17) and (18) and we proceed to the estimation of the rest of parameters.

We have to estimate the parameters of the marginal cost function, which are the parameters of the production function. In Section 2 we argued that marginal cost is nonparametrically identified together with the rest of of the model. Here identification becomes straightforward thanks to the parametric specification. Our specification of marginal cost has three endogenous variables, c_{jt}, l_{jt} and m_{jt}, in the sense that, because of the moment of its determination, and may be because e_{jt} too, they are correlated with the disturbances v_{1jt} and v_{2jt}. The exogenous components of (contemporaneous or lagged) marginal cost that enter the equations are $k_{jt}, c_{jt-1}, k_{jt-1}, l_{jt-1}$ and m_{jt-1}. As we have only three parameters to estimate (β_K, β_L and β_M) this seems perfectly enough. In practice we will prefer not to treat k_{jt} as exogenous because of presumably errors in its measurement. And we will find that lagged wage, w_{jt-1}, not directly included anywhere, performs as a very reasonable instrument.

[18] Our first estimations adviced us to maintain for the moment the restriction $\lambda = 1$.

In the rest of this section we first detail how we estimate the elasticities and then the GMM procedure that we employ for the rest of parameters.

4.1 Estimating the elasticity of demands.

Here we build on Aw, Roberts and Xu (2011). According to equation (16), variable cost can be written as

$$C_{jt} = \nu MC_{jt} Q_{jt} \exp(-e_{jt}) = \nu (MC_{jt} Q^X_{jt} + MC_{jt} Q^D_{jt}) \exp(-e_{jt})$$

where $\nu = \beta_L + \beta_M$ is the short-run elasticity of scale.

Use the optimal pricing conditions $P^X_{jt} = \frac{\eta_X}{\eta_X - 1} MC_{jt}$ and $P^D_{jt} = \frac{\eta_D}{\eta_D - 1} MC_{jt}$ to substitute for MC_{jt}, divide everything by R_{jt} and denote by S^X_{jt} the share of revenue from exports over total revenue. Manipulating a little the expression and taking logs we have the regression

$$\ln \frac{R_{jt}}{C_{jt}} = \ln \frac{1}{\nu} \frac{\eta_D}{\eta_D - 1} - \ln \left[1 + \left(\frac{\frac{\eta_D}{\eta_D - 1}}{\frac{\eta_X}{\eta_X - 1}} - 1 \right) S^X_{jt} \right] + e_{jt}. \tag{19}$$

This equation allows to estimate the elasticity of each demand by NLS up to the parameter of short run elasticity of scale.[19] Writing the regression as

$$\ln \frac{R_{jt}}{C_{jt}} = \ln a - \ln \left[1 + b S^X_{jt} \right] + e_{jt},$$

we have that $\eta_D - 1 = 1/(a\nu - 1)$ and $\eta_X - 1 = (1 + b)/(a\nu - 1 - b)$. We can estimate a and b and substitute the corresponding expressions for $\eta_D - 1$ and $\eta_X - 1$ in equations (17) and (18), where we are going to search for β_L and β_M and hence the value of $\nu = \beta_L + \beta_M$.

In the preliminary estimations that we report below we have used the simplifying assumption that the firm is producing at the minimum of total average cost. Under this assumption $\frac{C_{jt}}{TC_{jt}} = \frac{W_{jt} L_{jt} + P_{M_t} M_{jt}}{W_{jt} L_{jt} + P_{Mt} M_{jt} + uc_{jt} K_{jt}} = \frac{\nu MC_{jt}}{ATC_{jt}} = \nu$, where uc_{jt} is an estimated user cost for capital. Adding $\ln \frac{C_{jt}}{TC_{jt}}$ to both sides of equation (19), ν disappears from the right hand side. But this assumption is not consistent with the long run equilibrium of monopolistic competition (where firms tend to produce on the left of this minimum), and the employed method is not efficient: we implicitly get two estimates of ν that can be different between them.

[19] Notice that, assuming $\nu = 1$, a linear approximation to this equation can be read as estimating the value of the percentage markup m_D in the domestic market and the difference $m_D - m_X$ between percentage markups in the domestic and exports market: $\ln \frac{R_{jt}}{C_{jt}} \simeq m_D - (m_D - m_X) S^X_{jt} + e_{jt}$.

4.2 GMM estimation.

Write the residuals of (17) and (18), after plugging the elasticity estimates, as a function of the vector θ of parameters that remain to be estimated. The GMM problem is

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

where $A(\cdot)$ is an $L \times T_j$ matrix of functions of the exogenous variables z_j; $\nu_{1j}(\cdot)$ and $v_{2j}(\cdot)$ are the $T_j \times 1$ vectors of residuals, and N is the number of firms. L denotes the number of moments that we are going to use for each equation and T_j the number of observations for firm j. Notice that in principle we are considering the same set of instrument for each equation, although it could be otherwise. We estimate in two steps. In the first step the weighting matrix is

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

while in the second we use the matrix

$$\widehat{W} = \begin{bmatrix} \frac{1}{N}\sum_j A(z_j)v_{1j}(\widehat{\theta})v_{1j}(\widehat{\theta})'A(z_j)' & \frac{1}{N}\sum_j A(z_j)v_{1j}(\widehat{\theta})v_{2j}(\widehat{\theta})'A(z_j)' \\ \frac{1}{N}\sum_j A(z_j)v_{2j}(\widehat{\theta})v_{1j}(\widehat{\theta})'A(z_j)' & \frac{1}{N}\sum_j A(z_j)v_{2j}(\widehat{\theta})v_{2j}(\widehat{\theta})'A(z_j)' \end{bmatrix}^{-1}$$

where $\widehat{\theta}$ represents the estimate of the parameters in the first step. We then compute the asymptotic variance of $\widehat{\theta}$ as

$$Avar\left(\widehat{\theta}\right) = \left\{ \begin{bmatrix} \sum_j A(z_j)\frac{\partial v_{1j}(\widehat{\theta})}{\partial\theta} \\ \sum_j A(z_j)\frac{\partial v_{2j}(\widehat{\theta})}{\partial\theta} \end{bmatrix}' \frac{\widehat{W}}{N} \begin{bmatrix} \sum_j A(z_j)\frac{\partial v_{1j}(\widehat{\theta})}{\partial\theta} \\ \sum_j A(z_j)\frac{\partial v_{2j}(\widehat{\theta})}{\partial\theta} \end{bmatrix} \right\}^{-1}.$$

While this is exactly what we do right now, we are conscious that it should be modified. We are estimating previously the elasticities, so we have at least to correct the asymptotic variance for the variance of these parameters previously estimated. We have still not done because we may also want to take this into account in the weighting matrix of the second step.

The literature on optimal instruments establishes that variance can be minimized by setting instruments of the form

$$A_1(z_j) = E\left[\frac{\partial v_{1j}(\theta_0)}{\partial \theta}|z_j\right] \text{ and } A_2(z_j) = E\left[\frac{\partial v_{2j}(\theta_0)}{\partial \theta}|z_j\right],$$

where θ_0 is the true value of θ. The form of our equations has the advantage that these derivatives are always linear in the endogenous variables and the expectations can be computed. This reintroduces the lagged advantages in the functions $A_1(z_j)$ and $A_2(z_j)$. Most parameters, in addition, are also inside the unknown functions and this implies that derivatives of these functions are going to enter the expectations. This suggests using polynomials on all variables inside the unknown functions and some interactions. These variables are exogenous and this allows for a good prediction of these derivatives.

In the estimation that we report below we use the following instruments in each equation: constant, set of time dummies, a complete polynomial of order three in the key variables k_{jt-1}, l_{jt-1} and m_{jt-1}. We do not include variable p_{Mt} because is common to all firms, and hence perfectly predicted by the dummies. Instead, we add univariate polynomials of order three in two instruments more, lagged variable cost c_{jt-1} and lagged wage w_{jt-1}. We also include polynomials of order three in the lagged shifters.

5. Data

5.1 Source

The source of our data is the Annual Census of Industrial Production, a firm-level survey conducted by the National Bureau of Statistics (NBS) of China. This annual census includes all industrial non-state firms with more than 5 million RMB (about $600,000) in annual sales plus all industrial state-owned firms (SOEs). Our source is then the same used in Brandt, Van Biesebroeck, and Zhang (2012). A comparison with the (complete) Census of 2004 led them to conclude that the aggregates correspond extremely well and that the included firms account for somewhat more than 90% of Chinese industrial output.[20] Our data cover the period 1998-2008.

[20] The same data source has been used, for example, in Hsieh and Klenow (2009) and Lu (2013).

Our sample hence consists basically of large firms and some smaller SOEs. The available information includes firm demographics such as location, industry code, the date of birth and some detail on ownership. We obtain from the data the revenue of the firm (which can be split in domestic sales and exports), physical capital, wage bill, cost of materials, the number of workers and the amount spent in sales promotion and (for a few years) in R&D. In Appendix B we detail the content of these items as well as the construction of variables starting from them.

We want to use the data as a panel of firms, that is, we want to exploit all the observations repeated over time which are available for the same individual. One reason is that our modeling implies persisting productivity and demand advantages evolving over time, whose estimation depends on the sequence of observations on the firm. Another is that we are interested in the impact of variables which depend on the history of the firm, as its experience in the export market or accumulated R&D expenditures. On the other hand, we are particularly interested in detecting the new born firms, as well as the firms that eventually shut down. We want, for example, characterize separately their results. In order to make all this possible we have had to address two important and related questions: the problem of discontinuity of information and the detection of the "economic" entry and exit of firms in the middle of all the additions to and drops from the sample. Appendix C details how we have dealt with these issues and how we have cleaned the resulting data.

5.2 Manufacturing in China during the 2000's

Table 1 provides basic statistical information on the treated data. Columns (2) to (4) report unweighted averages of the firm's levels of revenue, capital and employment, and columns (5) to (7) unweighted averages of their rates of growth. Columns (2) to (4) show that revenue per firm triplicates over the period, while real capital stays at the same level and there is a significant fall in the average number of workers (more than 25%). Columns (5) to (7) show a intense average growth of output, closely followed by capital, and a positive growth for employment after 2002. In column (8) we compute a standard measure of TFP, the growth of deflated revenue minus the weighted growth of capital, labor and materials. We use as weights the average of the cost shares in moment t and $t-1$. Cost is the sum of the wage bill, the cost of materials and a cost of capital calculated using a common user cost. TFP growth is strong, especially after 2001, and averages 2.7%. This estimate matches well the estimates by Brandt, Van Biesebroeck, and Zhang (2012).

It is worthy to dedicate some space to comment on what this data shows about the evolution of the Chinese manufacturing during the 2000s, because the activities of the firms that we analyze in this paper take place in this context. There is implicit in this data an spectacular growth of the industrial output accompanied by a huge growth and reallocation of productive resources. The number of firms is roughly multiplied by a factor of three. This means that, to obtain the growth of the industrial aggregates corresponding to revenue, capital and employment from the reported firm-level means, we should multiply one plus the rate of growth of the corresponding mean by three. This gives the following rough picture: nominal revenue was multiplied during the period by nine, capital by three and employment by two. The increase in output is hence based in an intense increase of productivity of the firms, on the one hand, as the calculation of TFP already made clear. Capital and labor hugely increased as well, but with an important displacement of the leading economic role to firms of smaller size. This is the reason why, despite the increase of the aggregates, capital per firm stays stable and employment per firm diminishes more than one quarter.

Entry and exit play a big role in reallocation, but also the restructuring of the continuing firms. Entry and exit rates are considerable during the entire period (9.6 and 8% on average, respectively) and net entry is positive since 2003 (with an average of 3.3%). Net entry is likely to strongly contribute to the increase of productivity, but further calculations show that it is not a source of net employment growth. Entrants, despite to be more, have a quite smaller average size than exiters. It follows an equally dynamic behavior of many continuing firms, especially of smaller size.

5.3 Sample

Our identification strategy is focused on the firms which sell in the export and domestic markets. We have hence to characterize the importance and role of these companies. Table 2 splits manufacturing in ten sectors which group two-digit industries (see the Appendix for the correspondence). We start by reporting in column (1) TFP growth in these industries, to show that the main characteristics commented for the whole industry are generalized across sectors. Columns (2), (3) and (4) report then the proportion of observations on firms operating in both markets, in the foreign market only, and in the domestic market only, respectively. The proportion of observations on firms selling in the export market and home ranges from 10% (Paper) to 31% (Electronics). The proportion of firms which only sell abroad is only somewhat significant in three (very different) activities (Textile, Timber and Electronics). Column (5) in fact shows that the proportion of firms selling abroad and home sharply fell during the period in Textile while increased in Transport equipment and Electronics, and tended to moderate increases in many other industries.

Columns (6) to (8) report the proportion of firms over the period that sold -at least at some moment- abroad and home, the proportion of industry sales that they accounted for in 2008 and, finally, their average export intensity or proportion of sales which go to the foreign markets. Column (7) shows that these companies can represent between 20% and almost 70% of the sales of the corresponding industry, 40% or more in the most technologically intensive industries. Columns (6) and (7) taken together show that firms that sell both abroad and domestically tend to be smaller in Textile, roughly the same size in other three industries (Food, Paper and Non-metallic minerals), and tend to be significantly bigger in the rest of the industries. There is hence a clear relationship between the technological intensity of the industry and the relative size of the firms that sell in both markets. Column (8) shows that export intensity ranges from 35% to 60%. The proportions for industries Textile and Timber (taken together with the proportions of column (3)) reinforce the idea that, in these activities, tends to be some specialization in the export market.

Summarizing, firms which sell abroad and domestically are a significant part of the sales of any industry, 40% or more in more than half of the industries and particularly important in the most intensive technologically. It seems sensible to base an identification strategy just in these firms, taking however into account that -as the statistics show- they constitute an autoselected sample which responds to some technology and industry characteristics. Our industry panel samples consists of all the firms and (continuous) time sequences that we can get for these firms operating in the two markets.

Table 3 provides descriptive statistics corresponding to this sample. We are going to base our estimations in more than 70,000 firms and almost 280,000 observations. Firms show a high average export intensity, which ranges by industries from 35% to 60%. The average age is between 8 and 14 years, only a small fraction of firms is state owned (between 5% and 20%) but financial capital of the state firms shows a state share above 60%. Firms do not differ too much in its average export experience but they report significant and heterogeneous sales efforts which account from 2% to 5.5% of revenue.

6. Results.

In this section we report and comment on our preliminary results. First, we report the results of the use of equation (19) to estimate the elasticity of demand in the export and domestic markets. Second, we report the results of the estimation of the system of equations (17) and (18), for exports and domestic revenue, after plugging our elasticity estimates. Then we estimate the distributions of cost and product advantages, ω_{jt} and δ_{jt}, and characterize them. In the third subsection we describe the distributions. In the fourth we comment on the correlation between these advantages and the firm-level variables age, state participation, experience in the exports market and worker skills. In the fourth, we describe the average advantages for firms that enter and exit the markets, as well as for firms that perform R&D.

6.1 Estimating elasticities.

Recall that our preliminary results use, as explained in subsection 4.1, an estimate of the short-run scale parameter ν to estimate equation (19). With this caveat in mind, however, it should be recognized that the equation gives very sensible results. They are summarized in Table 4. An advantage of the simplification of the problem is that we can directly read the intercept and the slope as (approximate) margin values and differences between margins (see footnote 16). According to column (1), markups over the short-run marginal cost range from 8% (Textile) to 15% (Electronics) in the domestic market.[21] According to column (2), margins in the export market are systematically lower, with the difference going from 1 to almost 9 percentage points. This implies that export markets are systematically more competitive than domestic markets: firms when sell abroad face more substitutes and higher elasticities. Both the value of the function and of the standard error of the equation (reported in column (5)) point to a very good fit, that makes particularly credible the estimates.

Columns (3) and (4) compute the elasticities corresponding to these margins. Elasticities are sensibly different across industries and mostly show significant differences between the export and domestic markets. Although column (2) reports as significant all proportional markup differences, we drop industries 2 (Textile), 6 (Non-metallic minerals) and 7 (Metals and metal products) from our next preliminary exercises, considering that we cannot get a large enough difference of elasticities.

Our elasticities are similar, for example, to the value obtained for the electronic Taiwanese industry by Aw, Roberts and Xu (2011), although tend to be higher than the elasticities obtained with a different method by Roberts, Xu, Fan and Zhang (2012). Our result contrasts with the idea that markups are systematically higher in the export market (see, for example, de Loecker and Warzynski, 2012). Because of this, it is important to discard potential alternative explanations for our results. Our equation could be seen as saying that marginal costs of selling abroad are higher than the marginal cost of selling domestically. Our model rules this out by assuming that the marginal cost of production is the same and that the costs of exporting are basically sunk and fix. But in the future we want to find ways to test this hypothesis. Another possibility would be that there is a negative correlation between the export share and the shock of the equation. For example, there could be something that explains that exports are higher in bad times for profitability.

We want also to explore the possibility of allowing for different elasticities to different firms in the same industry. The length of time periods that the firms stay in the sample seem

[21] With short-run returns to scale below unity this implies somewhat higher margins on average variable cost.

too short to estimate individual elasticities. But equation (19) provides a nice framework for allowing markups, and hence elasticities, to change with firm characteristics. For the moment, we use the elasticities of Table 5 to estimate the system of equations explaining revenue in all industries but three.

6.2. System for exports and domestic sales.

In this subsection we summarize the results of estimating the system. Table 5 reports the results of the estimation, carried out by nonlinear GMM. The reported coefficients and standard errors are second stage estimates.

Recall that the demand shifters are aimed at controlling for all the demand advantages that can be observed. We have included three shifters in each equation, two are common but we allow for different impacts in each equation. The first common included variable is the level of participation of the state in the financial capital of the firm or *State*. There are reasons for thinking that this may be an important source of advantages both in the exports and domestic markets. One the one hand, government has actively searched to stimulate exports by means of tax exemptions, promotion of geographical spillovers and, sometimes, direct support. On the other, government has traditionally played a leading role in domestic manufacturing production, particularly in some sectors, what may constitute a remaining advantage for the firms under state control. Both kinds of advantages may have been counterbalanced, however, by a lower dynamism in front of the new private firms.

The second common variable is the expenditure in sales promotion, through marketing, advertising, sales forces and so on, or *Sales effort.* Unfortunately we cannot distinguish between expenditures aimed at promotion of exports and domestic sales, so we have to content ourselves with including the total amount in each equation in the hope that the different coefficients will help to pickup the specific effects.

Finally, in the revenue from exports equation, we include the experience of the firm in the export market, measured as the number of years elapsed since the first time that the firm is seen exporting in our sample. We call this variable *Experience.* For the firms staying since the first year in the sample we cannot distinguish if the firm was already exporting before

1998.[22] It is likely, however, that the experience before this year does not affect many firms and was not as important as the experience in later years. In the domestic equation we have found useful to include a shifter characterizing the level of capital intensity, or physical capital per worker. We think that it can be a rough indicator of levels of quality of the product.

Recall from subsection 4.2 that we use the following instruments in each equation: constant, set of time dummies, a complete polynomial of order three in the key variables k_{jt-1}, l_{jt-1} and m_{jt-1}, and univariate polynomials of order three in two instruments more: lagged variable cost c_{jt-1} and lagged wage w_{jt-1}. According to our choice of shifters, we also include in both equations polynomials of order three of the variables *Sales effort* and *Experience* lagged. We include *State* only linearly to avoid exacerbating colinearity.

The main coefficients of interest are the production function parameters, estimated through the specification of marginal cost. Columns (1) to (3) of Table 5 show the point estimates. The results look globally sensible, with plausible values and low standard errors. Returns to scale are very close to unity, and short-run returns to scale are as expected below unity in five of the seven industries. The most difficult coefficient to estimate was capital, that in fact shows somewhat heterogeneous values across industries. It is likely to be a question related to the difficulties of measurement of this variable.[23]

With regards to the shifters, the variable *State* is significantly associated in four sectors to a lower revenue from exports (Food, Furniture, Machinery and Electronics) and to additional revenue in only one (Chemical). The state participation tends to be positively linked to additional revenue in more industries in the case of domestic sales (Paper, Chemical, Transport), but not very significantly, and it has again a negative impact in other two (Machinery and, less significantly, Furniture).

The variable *Sales effort* is positively and very significantly linked to greater sales in the same four sectors in the exports and domestic markets (Paper, Chemical, Transport,

[22] This problem also affects to firms incorporated later that are not recently born.

[23] Our current estimates include the control of some unlikely capital values in a number of industries through the variable capital squared.

Electronics).[24] The variable *Experience,* only included in the exports equation, is significant in two industries (Food and Electronics), and also positive but much more imprecisely estimated in other (Furniture). The capital intensity of the firm attracts in the domestic sales equation a positive significant coefficient (Furniture) and a negative one (Paper).

Our general feeling is that the equations may still improve a little by including, as planned, variables related to the location of the firm and on the subsidies received. It would be unlikely, however, that this changes very much the distribution and properties of the unobserved advantages that we present in the following subsections.

6.3 Distribution of ω_{jt} and δ_{jt}.

We are going to report, for the sake of better comparability, ω_{jt} and δ_{jt}/η (always in differences with respect its global mean). The first variable directly reflects the efficiency-based price differences. The second can be read in terms of the willingness of consumers to pay a price different from the baseline price. Table 6 summarizes the distributions of ω_{jt} and δ_{jt}/η, Figure 1 depicts the marginal densities of ω_{jt} and δ_{jt}/η and their changes over time, and Figure 2 depicts the joint density of ω_{jt} and $\delta_{jt}/$. The distributions turn out to be sensible and very informative.

Columns (1) to (3) of the table report the quartiles of the distributions, and column (4) the standard deviation. Both unobserved advantages tend to have a large dispersion. The typical advantage (an advantage value equal to the typical deviation from the mean of the distribution) may explain differences between 50% and more than 100% of the price, according to the industries. These are either real differences in observed prices, attributable to cost advantages, or in prices that consumers are estimated to be willing to pay given the observed product advantages. This is a notable dispersion, but it simply mirrors the huge dispersion of revenues by firms. In industries 8 and 9 (Machinery and Transport equipment), however, we may be overstating a little the dispersion due to the proximity of the two elasticity estimates. The cost advantages are fairly symmetric, and the product advantages are systematically somewhat skewed to the left, as shown in column (5). That

[24] We are here neglecting the effect of a square term in sales effort, usually negative, that is important to pickup the effect and that may also be controlling for some outliers.

is, the 50% of firms with lower product advantages show some greater numbers (in absolute value) than the numbers of their positive counterpart.

The unobserved advantages ω_{jt} and δ_{jt}/η are strongly negatively correlated, as reported in column (6) and illustrated by Figure 2. As there is not anything in the model that implies such correlation (see, for example, equation (11)), this is an important finding of the exercise. The joint densities of ω_{jt} and δ_{jt}/η, when depicted, show similar traits across industries. The joint densities are single peaked but show important density masses when one advantage is important and the other weak. This is saying that an important bulk of firms either have cost or product advantages, but are not well placed in the other ladder respectively. This strongly suggests that many firms that have important cost advantages sell standard or even low quality products. And that firms that show important product advantages show clear disadvantages in the cost of their products, possibly due to the costs of differentiation (technology, design, quality...)

Columns (7) and (8) of Table 6 show the changes over time in the distributions of ω_{jt} and δ_{jt}/η, and Figure 1 illustrates them. The mean of cost advantages increase for the entire period in all industries but Furniture, from 20 to 100%. The most technological intensive sectors (Chemical, Machinery, Transport equipment, Electronics) show all strong advances above 36%. The standard deviation of the distribution of costs advantages tends to remain stable or decrease a little. Product advantages tend to increase as well, but the increase is much more heterogeneous. Four sectors show increases in the mean advantage in the range 15-20% (Food, Furniture, Paper, Electronics). The mean product advantage is basically stable in another, Chemical. And there are decreases in the mean product advantage in Machinery and, particularly, Transport equipment. Demand advantages also tend to be a little more dispersed over time inside most of the industries.

6.4 Correlations.

Table 7 reports some correlations computed across firms and over all years. Column (1) shows that export intensity is is strongly correlated with cost advantages. Non parametric regressions of export intensity on the cost advantage give continuously increasing relationships in all industries.

The rest of correlations, reported in columns (2) to (9), provide an interesting picture of the links between firm characteristics and market advantages, that seems quite robust across industries. Younger firms tend to present cost advantages and older firms to show demand advantages. State participation is clearly associated to less cost efficiency but also to some demand advantages. The fact that this was not picked up by the regression state dummy suggests that it can be linked to specific firms. Experience in the export market is not particularly relevant in explaining advantages, but it seems clearly more important in providing cost advantages than demand advantages. The index of worker skills is clearly associated to demand advantages.

6.5 Entry, exit, R&D and advantages.

Table 8 completes the picture describing some average cost and demand advantages for entrants and exiters versus continuing firms, as well as for firms performing R&D with respect to the non performing firms.

Column (1) compares the entrants cost advantages with the cost advantages of the incumbents, and column (3) the demand advantages. Entrants tend to show costs advantages but no demand advantages or even negative demand advantages. Columns (2) and (3) perform the same type of comparison for the exiting firms with respect to the continuing. The data tell that exiting firms have more demand than cost disadvantages.

Columns (5) and (6) compare the cost and demand advantages of the firms that perform and not perform R&D in the few years for which we have data on this activity. The results are striking. Firms performing R&D clearly show systematic cost disadvantages that, however, tend to be more than compensated by the demand advantages. This clearly establishes something that could be guessed from the previous correlations and constitutes an important evidence: firms that try to establish some product advantages through technology experience higher costs than the firms that do not. This clearly points out to the coexistence of two types of firms and competition, through product differentiation and technological development, and competition based on low costs.

7. Concluding remarks

The purpose of this paper has been to develop a model to measure the unobservable cost and product advantages of firms and to apply it to characterize the growth of Chinese manufacturing during the 2000's. While unobservable cost advantages of firms have been the object of interest by economists since long ago, unobservable product advantages have only recently drawn attention. There are well established methods to measure the relative firm-level efficiency of firms, that have been the object of continuous refinement by the literature on the estimation of production functions. But almost no models have been advanced in order to assess the symmetric advantages that firms have and develop in their demands.

While we still want to perfection some of the paper estimates and inferences, a rich set of facts has emerged from our preliminary application of the model to the Chinese data. They draw a subtle picture of the underlying forces to the enormous increase of exports and industry restructuring of the 2000's.

Unobservable demand and cost advantages turn out to have a large dispersion, with the cost advantages more symmetrically distributed across firms. They are strongly negatively correlated industry to industry, what suggests that firms have tended to specialize either in the production of low cost standard products or differentiated products, for which they are no longer able to keep the lowest costs. All analyzed industries show an intense increase in the productive efficiency of firms, particularly in the most technologically intensive. More moderate are the increases in the demand or product advantages which, among the most technologically intensive industries, show a sharp advance only in the key Electronics industry. Behind all these changes there is a huge restructuring of the industries, with average entry and exit rates about 9.5 and 8%. Entrants and younger firms show cost (but not demand) advantages, while older firms tend to have more demand advantages. Exits seem much more linked to the absence of demand advantages than to productive inefficiency. State owned firms are clearly less cost efficient. Demand advantages are associated to greater worker skills and pay. Firms that perform R&D show sharp cost disadvantages, more than compensated by the demand advantages. Had we analyzed only productive efficiency we would had likely been puzzled by the low efficiency of the performing R&D firms.

Everything is as firms consciously develop two types of competition that are present industry to industry: competition by means of low cost, competition by means of the (costly) development of higher quality differentiated products. An important part of the Chinese exports during the 2000's has been supported by the first type of firms, but this is by no means the whole picture. An important segment of exports of quality and technologically developed products has been continuously growing in importance.

Appendix A: Demand specification

To discuss demand specification it is useful to start with the often used Dixit and Stiglitz (1977) CES system for a differentiated product industry. A representative consumer chooses the consumption level of the numerarire and the quantities of each variety of a differentiated good, aggregated in the index $Q = \left(\sum Q_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ where we drop subindex t for simplicity and denote the industry value by the absence of subindex. The demand for each variety turns out to be

$$Q_j = \frac{R}{P}\left(\frac{P_j}{P}\right)^{-\sigma} = Q\left(\frac{P_j}{P}\right)^{-\sigma},$$

where $P = \left(\sum P_j^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ is the price of the aggregate quantity and $R = \sum_j P_j Q_j = PQ$ is the income spent in the differentiated good.

Slightly departing from the original formulation, let's use "equivalent" firm-level indices for Q and P, i.e. the equal per-firm value for the N firms in the industry that would give the same value that the aggregate index reaches. These indices are $\overline{Q} = \left(\frac{1}{N}\sum Q_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ and $\overline{P} = \left(\frac{1}{N}\sum P_j^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ and using them we can write the demands as

$$Q_j = \overline{Q}\left(\frac{P_j}{\overline{P}}\right)^{-\sigma},$$

where $\overline{Q}$ and $\overline{P}$ can be conveniently read as (weighted) average values.

We can allow the quantities Q_j be the observed quantities of goods of different degrees of attractiveness to the consumer by adjusting them by means of factors $\exp(\frac{\gamma_j}{\sigma-1})$ which justify the symmetric treatment of the values $(\exp(\frac{\gamma_j}{\sigma-1})Q_j)$ in the quantity index and utility function (see Melitz 2000 for similar specification). When γ_j is high, the consumer reaches the same level of marginal utility with a greater quantity of the good. The non-observed price of the non-observed value can be written as $\exp(-\frac{\gamma_j}{\sigma-1})P_j$. Now demands in terms of the observed prices and quantities become

$$Q_j = \widetilde{Q}\left(\frac{P_j}{\widetilde{P}}\right)^{-\sigma}\exp(\gamma_j),$$

where $\widetilde{Q} = \left(\frac{1}{N}\sum \exp(\frac{\gamma_j}{\sigma})Q_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ and $\widetilde{P} = \left(\frac{1}{N}\sum \exp(\gamma_j)P_j^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ stand for average quantity and price quality-adjusted indices. $\widetilde{Q}$ may also be written as $R/N\widetilde{P}$ (average revenue spent in the goods divided by the quality-adjusted price index). The important thing to notice is that, even with prices and quality fixed, average revenue may change over time both because the resources allocated by the consumer to this industry and the number of firms producing actively in it.

Let's assume, without loss of generality, that $\widetilde{P} = 1$. In fact, when time is relevant, we can usually find an industry price index adjusted by quality to deflate prices by this index. The demand for good j can finally be written as

$$Q_j = \overline{R}\,(P_j)^{-\sigma}\exp(\gamma_j)$$

Introducing time, we will generalize over this model by writing

$$Q_{jt} = \alpha_0 P_{jt}^{-\eta}\exp(z_{jt}\alpha + \delta_{jt}).$$

While retaining the assumption of a common elasticity, here firms are expected to differ over a given average amount of sales by specific advantages evolving over time. Part of these advantages are observed, part come from unobserved sources and are picked up by a presumably heterogeneous and persistent δ_{jt}. Nothing impedes part of the advantages expand or shrink at the same time for all firms in the industry. That is, δ_{jt} may partly reflect gains or losses in the sales of all firms because, for example, a pull of exports affecting all firms in the industry or a decrease on the average expected sales because intense entry.

Appendix B: Variables

Ownership (State)

The share of state in financial capital, computed as the sum of the amounts reported as state and collective capital over total financial capital.

Age

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

Entrant firm

For first time in the sample and born the same year or one of the two previous years.

Revenue and export revenue

Firms report the value of their revenue and the value of industrial export sales at current prices.

Export experience

Current year minus the first year that the exports of the firm are not zero.

Export intensity

Exports divided by firm's total revenue.

Capital stock

Firms report the value of their capital stock at original purchase prices and their capital stock at original purchase prices less accumulated depreciation. To convert these nominal values in an estimate of the real capital stock we need two things: the sequence of real investments and an estimate of the 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 after 1997, it is straightforward to get the starting nominal capital stock and the sequence of nominal investments by difference between the gross capital book values of two years. For those 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 2-digit industry level using 1998-2003 firms' data. We assume that the nominal gross capital observed for the firm comes from the growth at this rate of the capital with which the firm was born. We then estimate the capital stock at birth, deflate it, and compute the real stock in the first year of observation by applying the perpetual inventory method with the series of real investments implied by our calculation.

The investment deflator before 2006 is taken from Brandt, Rawski and Sutton (2008). We we have updated it using the Fixed Asset Investment price index from China Statistical Yearbook.

Wage bill, employment, wage.

Firms report several components of total yearly employees compensation that we add up as wage bill. These components are 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 is the total number of employees, which includes all the full-time production and nonproduction workers, reported by the firm. It excludes part-time and casual workers.

Average wage is obtained by dividing the wage bill by employment.

Skills

We measure skills by the ratio of the firm wage to the unweighted average of wages of the firms in the industry the firm belongs to.

Cost of materials

The NBS definition of intermediate inputs includes direct materials, intermediate inputs used in production, intermediate input in management, intermediate input in business operation (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. To estimate intermediate consumption we have alternatively 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.

Sales effort

All expenditures related to sales (e.g promotion and advertising) as reported by the firm.

Price of output

Output price index at 2-digits from China Statistical Yearbook.

Price of materials

We closely follow Brandt, Van Biesebroeck and Zhang (2012), computing the price of materials for each industry as a weighted average of the output prices of the industries to which the concerned industry purchases its inputs. For the weights we use the Input-Output table corresponding to 2002, which includes 42 sectors. The 2-digit manufacture price indices are from China Statistical Yearbook. The prices of agriculture, construction, transportation, retail and wholesale and some service sectors are calculated by comparing GDP at current prices and constant prices, which are included in the Collection of Statistical Material from 1949 to 2009.

R&D

Firms report the value of R&D expenditure in 2001 and each year of the period 2005-2007.

Appendix C: Data treatment

Discontinuity of information for an existing firm (other than incidental) can happen in the raw data base for two reasons. First, if a firm is non-state owned, falls below the sales threshold of RMB 5 million and it is not surveyed. If the firm re-enters the sample keeping its ID, we only get some missing observations in the time sequence of the firm. But, when the firm doesn't re-enter sample, we unfortunately have strictly no way to distinguish its disappearance from economic shutdown. Second, and more importantly, a firm can have been allocated a different ID (9 digit-code) during the period. Firms occasionally receive a new ID if they are subject to some restructuring (change of name, ownership...), merger or acquisition. This creates a lot of broken sequences and spurious entry and exit.

With regards to the case of the IDs, we have done an intensive work (in the style of Brandt, Van Biesebroeck and Zhang, 2012) to link over time the data of the firms that presumably had the ID changed. This process has used extensively information such the firm's name, corporate representative, 6-digit district code, post code, address, telephone number, industry code, year of birth, and has been implemented in several steps: first checking on neighbor years two by two, then the longer panel sequences with the following/previous years.

The results of treating the sample in this way seem very satisfactory. Focusing on manufacturing, and considering firm time sequences with a minimum of two years, we have a total of 445,397 firms and 2,253,383 firm-year data points with information. So, after our linking, firms stay in the sample by an average of 5 years. We have time sequences of 5 or more years for more than half of the firms and more than 80% of these sequences have no interleaved missings. The degree of response of the sample firms, considered year to year, tends to be higher than 95%.

The linked data details are summarized in Table C1. Column (1) shows that the single observations discarded after the process are a small percentage, except for the starting and final years, at which the process of linking is more difficult. Columns (2) and (3) document the growth of the sample over time, particularly important in the Census year of 2004.

Entry and exit, reported in columns (4) and (5), show very sensible values and explain part of this increase. Entry is defined as the set of firms newly included in the sample and born the same year or any of the two previous years. Its average rate is about 8%. The increase in newly born firms in the Census years of 2004 and 2008 is particularly high, reflecting probably the effort of administrative authorities in being exhaustive. Exit is defined as the set of firms last seen in the sample the previous year. It is hence something indirectly induced by our linking and that can include failures in the linking process as well as mixing some firms in a process of drastic downsizing. But its rate is very sensible, close to the rate of entry, somewhat decreasing over time. This seems a particular good outcome which validates the process. The resulting net entry rate (entry minus exit), reported in column (6), reverses the sign from negative to positive in 2003. Column (7) documents the increases in the sample which are not related to entry and exit. The data seem to denote a quite continuous statistical improvement of the Annual Census too, tending to include more and more firms. Part of this improvement can be related to the increase of the number of firms with a size above the threshold.

We clean the linked data according to the conditions reflected in Table C2. We set to missing value the observation of a year if there are some particularly small values in revenue, capital, wage bill and the cost of materials; some abnormal values in other variables (details in the table); or some consistency problems (revenue less than exports, sales effort or the wage bill plus the cost of materials -variable cost- greater than revenue, or financial capital less than the sum of the reported components). This enlarges the number of data points without information. We then use for each firm the time subsequence (adjacent years) of maximum length provided that is greater than one year. The cleaned sample retains 80% of the firms and 70% of observations.

Appendix D: Industry correspondence

Industry	Two-digit industries (code)
1. Food, drink and tobacco	Agricultural and By-Product Processing (13);
	Food Manufacturing (14);
	Beverage Manufacturing (15);
	Tobacco Products (16).
2. Textile,leather and shoes	Textile (17);
	Apparel, Shoes, and Hat Manufacturing (18);
	Leather, Fur, and Coat Products Manufacturing (19);
3. Timber and furniture	Wood Processing, and Other Wood Products (20);
	Furniture Manufacturing (21).
4. Paper and printing products	Paper Making and Paper Products (22);
	Printing and Recording Media Reproducing (23).
5. Chemical products	Chemical Materials and Products (26);
	Pharmaceutical (27);
	Chemical Fiber (28);
	Rubber Products (29);
	Plastic Products (30).
6. Non-metallic minerals	Nonmetallic Minerals Products (31).
7. Metals and metal products	Ferrous Metal Smelting and Rolling Processing (32);
	Non-Ferrous Metal Rolling Processing (33);
	Metal Products (34).
8. Machinery	General Machinery Manufacturing (35);
	Special Machinery Manufacturing (36).
9. Transport equipment	Transportation Equipment Manufacturing (37).
10. Electronics	Electronic Machinery and Equipment (39);
	Electronic Communication Equipment and Computer (40);
	Instruments, Meter, Stationery and Office Machine (41).

References

Ackerberg, D., K. Caves and G. Frazer (2006), "Structural Identification of Production Functions," Working paper, UCLA.

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

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.

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 University Press, 569-632.

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

Crozet, M., K. Head and T. Mayer (2012), "Quality Sorting and Trade: Firm-level evidence for French Wine," *The Review of Economic Studies*, 79, 609-644.

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

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

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

Eaton, J., S. Kortum and F. Kramarz (2011), "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79, 1453-1499.

Doraszelski, U. and J Jaumandreu (2013), "R&D and Productivity: Estimating Endogenous Productivity," *The Review of Economic Studies,* forthcoming.

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

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

Griliches, Z. (1979), "Issues in Assessing the Contribution of R&D to Productivity Growth," *Bell Journal of Economics*, 10, 92-116

Helpman, E. and P. Krugman (1985), *Market Structure and International Trade,* MIT Press.

Helpman, E. and P. Krugman (1989), *Trade Policy and Market Structure,* MIT Press.

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

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

Katayama, H., S. Lu and J. Tybaut (2009), "Firm-level Productivity Studies: Illusions and a Solution," *International journal of Industrial Organization*, 27, 403-413.

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

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

Lu, D. (2012), "Exceptional exporter performance? Evidence from Chinese manufacturing firms," mimeo, Rochester University.

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.

Melitz, Marc. (2000), "Estimating Firm-level Productivity in Differentiated Product Industries," mimeo, Harvard University.

Melitz, M. and S. Redding (2012), "Heterogeneous Firms and Trade," forthcoming in *Handbook of International Economics*.

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

Petrin, A., and F. Warzynski (2012), "The Impact of Research and Development on Quality, Productivity and Welfare," Working Paper, University of Minnesota and Aarhus University.

Pozzi, A. and F. Schivardi (2012), "Demand or Productivity: What Determines Firm Growth?," Working Paper 9184, CEPR.

Roberts, M., D.Y. Xu, X. Fan and S. Zhang (2012), " A Structural Model of Demand, Cost, and Export Market Selection for Chinese Footwear Producers," NBER Working Paper 17725.

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.

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

Tang, H. and Y. Zhang (2012), "Exchange Rates and the Margins of Trade: Evidence from Chinese Exporters," *CESifo Economic Studies*, 58, 671-702.

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

Wooldridge, J. (2009), "On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables," *Economic Letters*, 104, 112-114.

Table 1: Treated data basic descriptive statistics

Years	Number of firms[a]	Average levels			Average growth rates			TFP[e] growth
		Revenue[b]	Capital[c]	Labor[d]	Revenue	Capital	Labor	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1998	75,518	43.831	27.826	333				
1999	89,866	45.293	27.288	307	0.039	0.020	-0.018	0.023
2000	93,594	53.005	26.057	299	0.080	0.011	-0.006	0.027
2001	102,423	55.874	25.327	275	0.047	0.026	-0.005	0.002
2002	111,653	61.774	24.995	265	0.103	0.059	0.017	0.030
2003	125,422	73.802	24.624	259	0.165	0.093	0.036	0.034
2004	174,371	72.272	20.173	217	0.172	0.105	0.026	0.029
2005	192,097	83.400	21.645	218	0.234	0.166	0.062	0.037
2006	214,517	94.493	22.614	211	0.211	0.128	0.037	0.010
2007	232,664	110.841	23.872	206	0.243	0.127	0.044	0.035
2008	176,268	133.342	28.464	214	0.218	0.181	0.041	0.034
1998-2008	359,135	83.357	24.278	240	0.176	0.110	0.031	0.027

[a] The cleaned sample retains 0.806 of the firms and 0.705 of observations.
[b] Nominal. Millions of RMBs.
[c] Deflated by an investment price index. Millions of RMBs.
[d] Number of workers.
[e] 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.

Table 2: Productivity, type of firms, and firms with export and domestic markets.

	TFP growth[a] average 1999-2008	Proportion of observations with: Export and sales home	Export only	Sales home only	Change prop. 1998-2008	Firms that export and sell home: Prop. of firms[b]	Prop. ind. sales in 2008	Export intensity[c]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	0.006	0.149	0.014	0.837	-0.006	0.212	0.263	0.444
2. Textile,leather and shoes	0.028	0.266	0.082	0.651	-0.122	0.461	0.379	0.603
3. Timber and furniture	0.030	0.203	0.052	0.745	0.020	0.295	0.281	0.569
4. Paper and printing products	0.027	0.096	0.010	0.893	0.016	0.154	0.248	0.350
5. Chemical products	0.029	0.200	0.020	0.780	-0.018	0.270	0.380	0.382
6. Non-metallic minerals	0.038	0.114	0.018	0.868	0.021	0.183	0.205	0.400
7. Metals and metal products	0.020	0.187	0.024	0.790	0.020	0.254	0.426	0.484
8. Machinery	0.032	0.211	0.012	0.778	0.010	0.282	0.445	0.361
9. Transport equipment	0.040	0.200	0.015	0.785	0.079	0.275	0.544	0.364
10. Electronics	0.034	0.313	0.048	0.639	0.048	0.409	0.680	0.482

[a] 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.

[b] Firms that export and sell home at least at some moment over total firms in the industry.

[c] Revenue from exports over total revenue.

Table 3: Sample statistics.

	Number of firms	Number of observations	Export intensity	Age	Proportion of SOEs	Average sate share (only part.)	Export experience	Sales effort
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	5,469	20,684	0.457	10.9	0.151	0.681	3.8	0.050
2. Textile,leather and shoes	17,512	65,469	0.606	9.5	0.084	0.641	3.8	0.022
3. Timber and furniture	2,633	8,891	0.577	7.9	0.058	0.613	3.5	0.039
4. Paper and printing products	1,736	6,563	0.365	10.5	0.092	0.618	3.7	0.033
5. Chemical products	10,877	45,622	0.410	11.2	0.140	0.636	3.9	0.044
6. Non-metallic minerals	3,597	13,168	0.412	11.0	0.133	0.678	3.7	0.055
7. Metals and metal products	6,266	24,513	0.496	11.0	0.113	0.674	3.8	0.029
8. Machinery	8,681	35,414	0.374	13.5	0.158	0.690	3.8	0.039
9. Transport equipment	3,194	13,037	0.379	12.1	0.197	0.689	3.8	0.032
10. Electronics	11,298	46,527	0.508	9.5	0.113	0.590	3.9	0.035

Table 4: Estimating elasticities (NLS).

	$\ln \frac{\eta_D}{\eta_D - 1}$ [a]	$\frac{\frac{\eta_D}{\eta_D - 1}}{\frac{\eta_X}{\eta_X - 1}} - 1$ [a]	$\widehat{\eta}_x$	$\widehat{\eta}_D$	Equation std.err.
	(1)	(2)	(3)	(4)	(5)
1. Food, drink and tobacco	0.162	0.087	13.2	6.7	0.318
	(0.004)	(0.007)			
2. Textile, leather and shoes	0.084	0.014	14.8	12.4	0.203
	(0.002)	(0.002)			
3. Timber and furniture	0.126	0.037	11.7	8.4	0.227
	(0.005)	(0.007)			
4.Paper and printing products	0.114	0.042	14.2	9.3	0.256
	(0.005)	(0.010)			
5. Chemical products	0.159	0.069	11.3	6.8	0.342
	(0.003)	(0.005)			
6. Non-metallic minerals	0.142	0.012	8.2	7.6	0.298
	(0.004)	(0.005)			
7. Metals and metal products	0.101	0.012	11.7	10.4	0.253
	(0.003)	(0.005)			
8. Machinery	0.143	0.036	9.8	7.5	0.294
	(0.002)	(0.005)			
9.Transport equipment	0.123	0.024	10.6	8.6	0.250
	(0.003)	(0.007)			
10. Electronics	0.154	0.069	12.0	7.0	0.285
	(0.002)	(0.004)			

[a] All standard errors are robust to heteroskedasticity and autocorrelation.

Table 5: Estimating the system for exports and domestic sales (Nonlinear GMM)[a].

Industry	Inputs[b]			Exports equation[b]			Domestic sales equation[b]			Function[c]
	k	l	m	*State*	*Sales effort*	*Experience*	*State*	*Sales effort*	k-l	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Food, drink and tobacco	0.042	0.209	0.757	-0.769	-0.455	0.148	0.043	-0.217	0.069	94.666
	(0.018)	(0.029)	(0.036)	(0.298)	(0.213)	(0.029)	(0.288)	(0.089)	(0.051)	
3. Timber and furniture	0.055	0.132	0.933	-1.655	0.143	0.082	-1.127	0.051	0.139	43.381
	(0.027)	(0.054)	(0.077)	(0.888)	(0.133)	(0.060)	(0.804)	(0.097)	(0.082)	
4.Paper and printing products	0.200	0.146	0.675	0.220	1.673	-0.015	0.779	0.690	-0.551	36.796
	(0.035)	(0.037)	(0.058)	(1.105)	(0.324)	(0.104)	(0.514)	(0.204)	(0.149)	
5. Chemical products	0.115	0.065	0.721	1.645	2.131	0.045	1.162	1.348	0.030	70.975
	(0.026)	(0.028)	(0.064)	(0.639)	(0.364)	(0.053)	(0.365)	(0.204)	(0.053)	
8. Machinery	0.040	0.155	0.769	-0.796	-0.058	0.016	-0.774	-0.063	-0.109	148.614
	(0.024)	(0.039)	(0.066)	(0.353)	(0.287)	(0.024)	(0.384)	(0.148)	(0.081)	
9.Transport equipment	0.064	0.126	0.910	-0.095	0.601	-0.133	0.368	0.297	0.017	66.654
	(0.022)	(0.038)	(0.051)	(0.305)	(0.256)	(0.087)	(0.240)	(0.132)	(0.102)	
10. Electronics	0.172	0.298	0.473	-1.104	0.742	0.070	-0.291	0.747		289.061
	(0.019)	(0.032)	(0.050)	(0.284)	(0.200)	(0.042)	(0.203)	(0.200)		

[a] Second stage estimates.
[b] All standard errors are roubust to heteroskedasticity and autocorrelation.
[c] Value of the objective fyunction scaled by N.

Table 6. Distribution of ω and δ

Industry		Quartiles 0.25	0.50	0.75	Standard dev.	Skewness[a]	Correlation between ω and δ	Changes 1998-2008 Mean	Standard dev.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Food, drink and tobacco	ω	-0.288	-0.003	0.276	0.448	0.007		0.316	0.011
	δ	-0.357	0.047	0.404	0.603	-0.078	-0.591	0.188	0.038
3. Timber and furniture	ω	-0.505	-0.027	0.460	0.794	0.034		0.047	-0.069
	δ	-0.496	0.085	0.593	0.897	-0.095	-0.976	0.235	-0.059
4.Paper and printing products	ω	-0.436	-0.090	0.351	0.683	0.132		0.197	-0.003
	δ	-0.0347	0.067	0.439	0.644	-0.104	-0.845	0.198	0.062
5. Chemical products	ω	-0.379	-0.002	0.368	0.620	0.003		0.479	-0.061
	δ	-0.310	0.042	0.383	0.607	-0.069	-0.675	0.070	-0.014
8. Machinery	ω	-0.736	-0.040	0.701	1.114	0.036		0.695	-0.007
	δ	-0.701	0.057	0.757	1.166	-0.049	-0.947	-0.179	0.076
9.Transport equipment	ω	-0.920	0.026	0.901	1.435	0.018		1.030	-0.026
	δ	-0.806	0.053	0.873	1.359	-0.061	-0.987	-0.581	0.017
10. Electronics	ω	-0.502	-0.029	0.451	0.779	0.037		0.359	-0.048
	δ	-0.399	0.046	0.443	0.691	-0.067	-0.641	0.140	0.026

[a] (Mean-Median)/Standard Deviation

Table 7. Correlations.

Industry	ω with export intensity	ω with				δ with			
		Age	*State*	*Experience*	*Skills*	*Age*	*State*	*Experience*	*Skills*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Food	0.694	-0.206	-0.123	0.059	0.075	0.169	0.041	0.090	0.151
3. Furniture	0.874	-0.113	-0.013	0.060	0.041	0.086	0.002	-0.052	-0.011
4.Paper	0.628	-0.051	-0.010	0.097	-0.054	-0.020	-0.077	-0.064	0.149
5. Chemical	0.794	-0.240	-0.242	0.089	-0.082	0.143	0.075	-0.017	0.009
8. Machinery	0.912	-0.311	-0.218	-0.007	-0.163	0.271	0.167	0.109	0.250
9.Transport	0.821	-0.200	-0.170	0.219	-0.057	0.162	0.123	-0.193	0.084
10. Electronics	0.695	-0.264	-0.180	0.048	0.030	0.177	0.078	0.038	0.203

Table 8. Industry re-structuring and investment: Differences of average advantages

Industry	Diff.of means of ω		Diff. of means of δ		Diff. of means of ω	Diff of means of δ
	Entrants vs Incumbents	Exiters vs Continuing.	Entrants vs Incumbents	Exiters vs Continuing.	R&D firms vs no R&D firms	R&D firms vs no R&D firms
	(1)	(2)	(3)	(4)	(5)	(6)
1. Food	0.055	-0.103	-0.142	-0.074	-0.132	0.360
3. Furniture	0.066	0.042	-0.036	-0.116	-0.133	0.163
4. Paper	0.013	-0.005	0.038	-0.093	-0.232	0.252
5. Chemical	0.134	-0.010	-0.066	-0.002	-0.250	0.133
8. Machinery	0.318	0.014	-0.325	-0.151	-0.616	0.779
9. Transport	-0.035	-0.066	0.053	-0.005	-0.260	0.258
10. Electronics	0.182	0.045	-0.109	-0.155	-0.235	0.308

Table C1: Manufacturing linked data[a]

Years	Discarded single obs.[b]	No of firms[c]	Sample growth	Entry rate[d]	Exit rate[e]	Net entry rate[f]	Additions[g]	Response rate[h]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1998	0.153	129,671	-	0.142	-	-	-	1.000
1999	0.026	145,949	0.126	0.044	-	-	-	0.971
2000	0.025	149,371	0.023	0.050	0.093	-0.043	0.066	0.955
2001	0.021	159,471	0.068	0.081	0.110	-0.029	0.097	0.950
2002	0.018	170,979	0.072	0.070	0.075	-0.005	0.077	0.946
2003	0.030	184,537	0.079	0.084	0.080	0.004	0.075	0.943
2004	0.067	247,854	0.343	0.176	0.099	0.077	0.266	0.966
2005	0.009	263,681	0.064	0.069	0.046	0.023	0.041	0.939
2006	0.010	288,433	0.094	0.088	0.055	0.033	0.061	0.953
2007	0.021	315,769	0.095	0.086	0.057	0.029	0.066	0.966
2008[i]	0.195	333,330	0.056	0.170	0.108	-	-	1.000
1998-2008		445,397						0.963

[a] Firms which stay a minimum of two years.
[b] As proportion of the remaining number of firms.
[c] Gives a total 2,253,383 firms-year observations with information.
[d] Newly included firms born in t, $t-1$ or $t-2$, as proportion of number of firms at t.
[e] Firms last seen at $t-1$ as proportion of number of firms at t. Not defined for 1998 and 1999.
[f] Entry rate - exit rate.
[g] Sample growth - net entry.
[h] Proportion of firms in sample at year t which report information.
[i] Entrants of year 2008, 48,369, treated in this line as if they were to stay.

Table C2: Filters used to clean the linked data

Values are set to missing in the following cases,

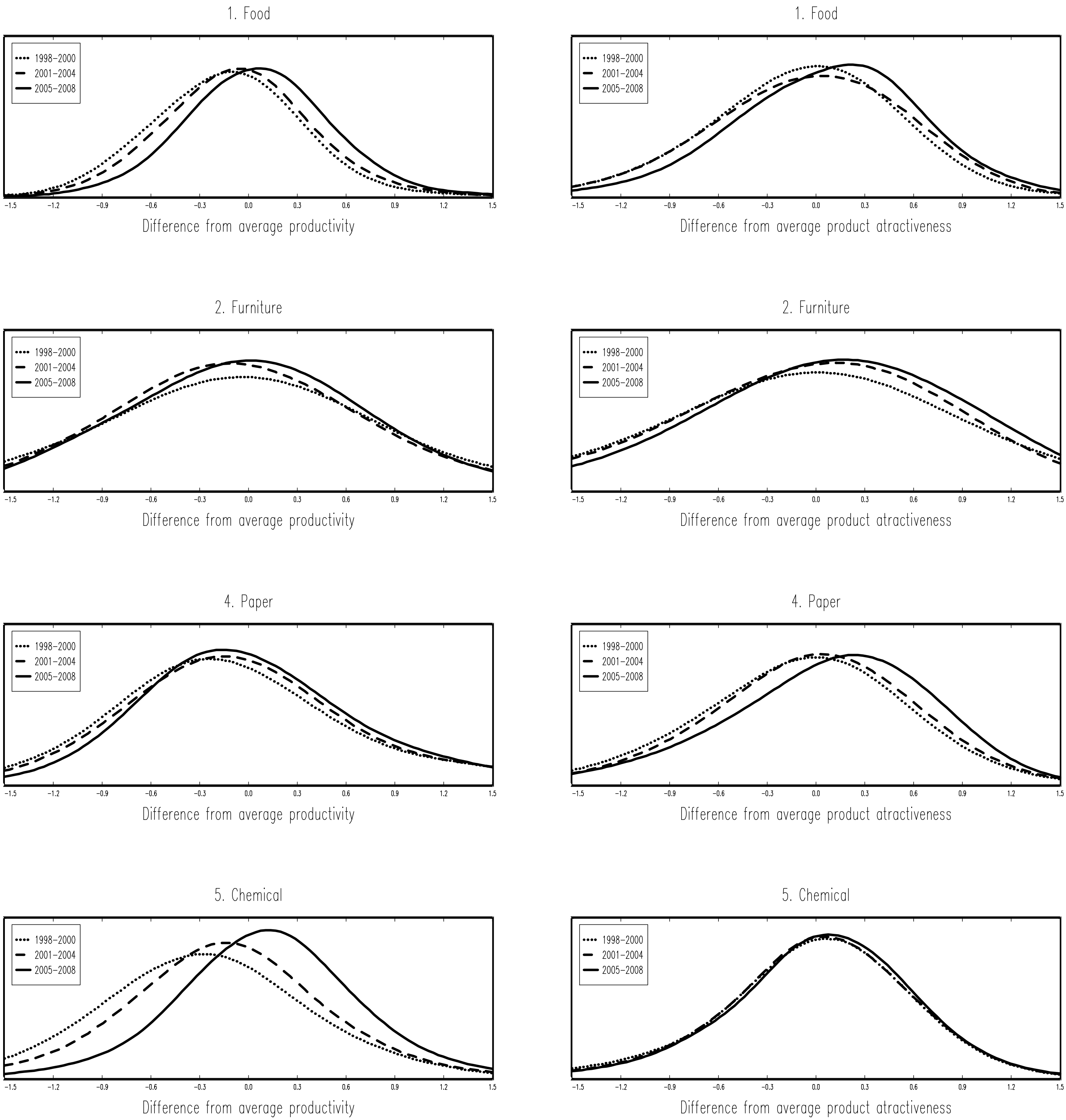
Small values:
- Less than 8 workers or 30,000 RMBs in Revenue, Capital,Wage bill, Cost of materials.

Abnormal values:
- Negative value in Exports or Sales effort.
- Zero or less in Finacial capital or negative finacial components.
- Born before 1949 or after 2008.

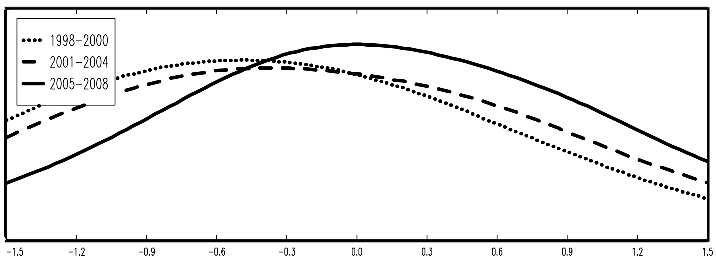
Consistency:
- Revenue less than Exports, Sales effort or Variable cost (Wage bill+ Cost of materials).
- Finacial capital is less that the sum its finacial components.

A missing value determines the interruption of the firm time sequence. We only use the time subsequence of maximum lenght provided that is longer than one year.

Figure 1: Density functions of ω and δ

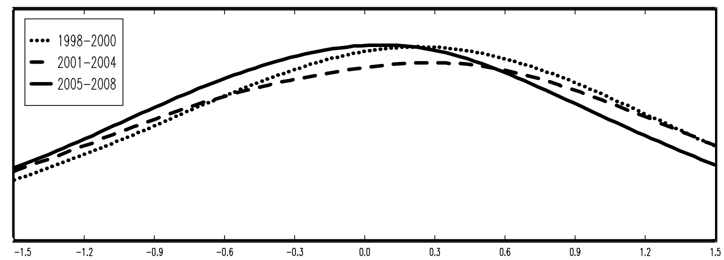


8. Machinery



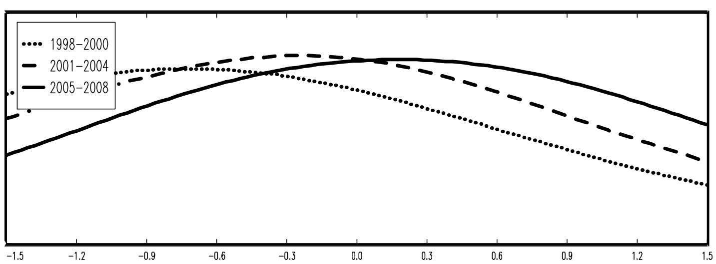
Difference from average productivity

8. Machinery



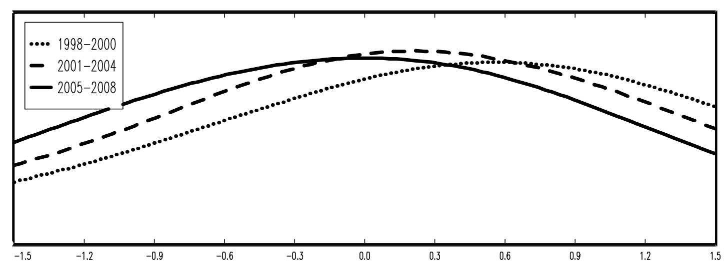
Difference from average product atractiveness

9. Transport



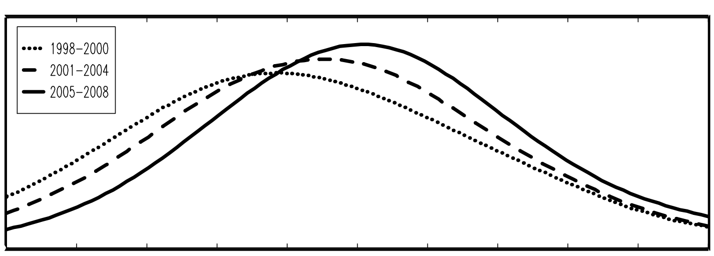
Difference from average productivity

9. Transport



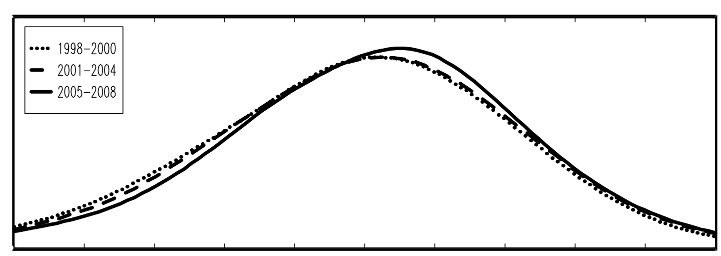
Difference from average product atractiveness

10. Electronics



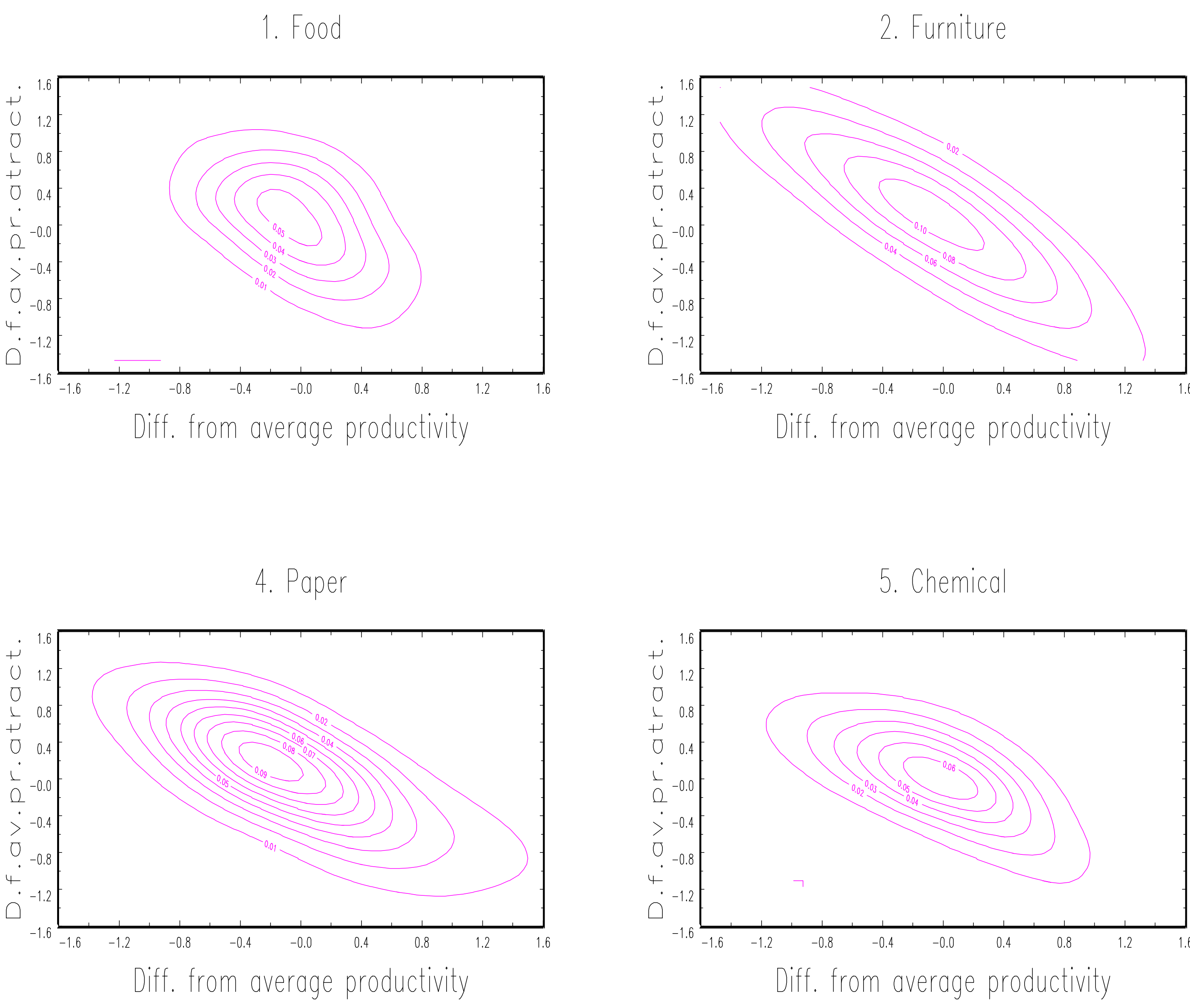
Difference from average productivity

10. Electronics

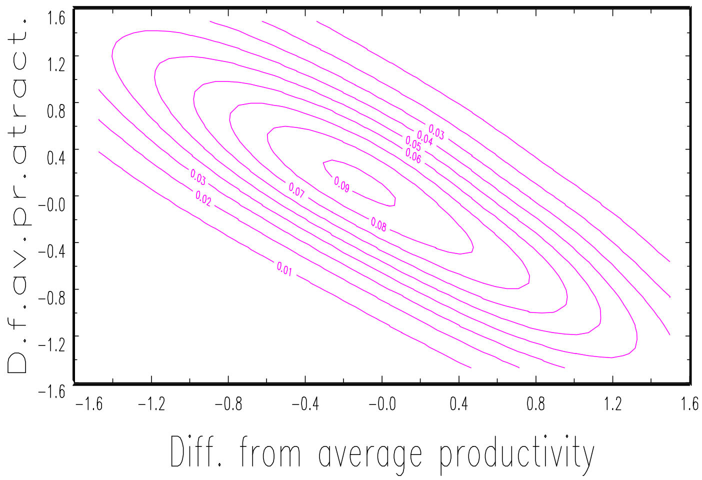


Difference from average product atractiveness

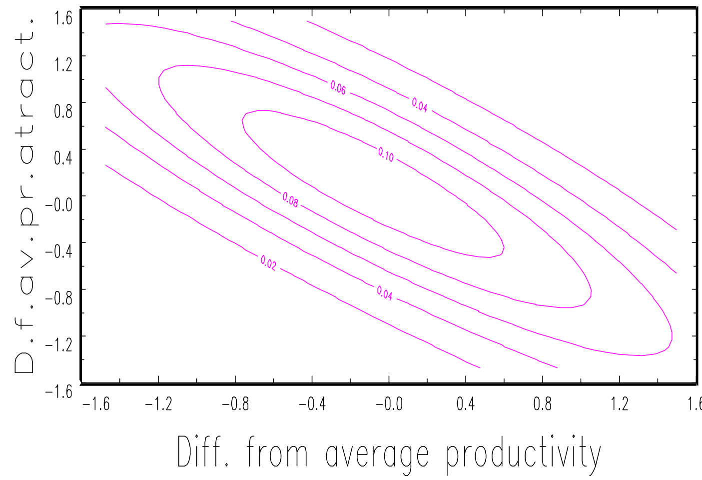
Figure 2: Joint density of ω and δ



8. Machinery



9. Transport



10. Electronics