Are Exporters More Productive than Non-Exporters?

David A. Rivers*

Department of Economics
University of Wisconsin-Madison

November 29, 2009

Job Market Paper

PRELIMINARY
PLEASE DO NOT CITE

Abstract

In an effort to explain the observed heterogeneity in the exporting decisions of firms, the empirical trade literature has concluded that exporting firms are more productive than non-exporting firms. In this paper, I show that the foundation for this conclusion is weak, given that the productivity estimates used in the literature suffer from several sources of potential bias. I apply a new method for estimating production functions to control for these sources of bias. Using data on manufacturing firms in Colombia, I find that, while the measures of productivity used in the literature imply that exporters are more productive, once I correct for the bias, there is no correlation between productivity and export status. This result is inconsistent with productivity being the main determinant of entry into export markets, and suggests the importance of other sources of heterogeneity in explaining firm-level exporting decisions.

*I am grateful to Steven Durlauf, Amit Gandhi, and especially Salvador Navarro for their guidance, comments, and suggestions. Rasmus Lentz, Ben Cowan, and Ignacio Monzón also provided helpful suggestions. I also wish to thank Amil Petrin for sharing with me the data on Colombian manufacturing firms.

1 Introduction

As firm-level data on exports has become available, an interest in understanding the causes of the observed heterogeneity in exporting decisions has developed in the empirical trade literature. One common observation in the data is that, even within narrowly defined industries, only a small percentage of firms export. This suggests the existence of some source of firm-level heterogeneity that causes certain firms to export while other firms do not. The literature has found that exporting firms tend to be "better" in the sense that they sell more output, hire more workers, are more capital intensive, and are more productive. This last finding has become almost a stylized fact in the literature. As a consequence, the literature has focused on productivity as a key determinant of export status, and has worked to develop models of exporting that can replicate this positive correlation between exporting and productivity.

In this paper, I show that the evidence for the conclusion that exporters are more productive is weak, as the underlying estimates of productivity suffer from several sources of potential bias. As opposed to characteristics such as the number of workers or the amount of capital, productivity is not directly observable and has to be recovered from the data. I show that there are several reasons to believe that the measures of productivity commonly used in the literature are biased. First, the use of value-added specifications of production ignores the role of intermediate inputs. Second, a firm's output is typically only measured in revenues as opposed to quantities. Failing to properly control for these unobserved prices leads to biased estimates of productivity. The third source of bias arises due to the common approach of measuring productivity as the residual from a regression of output on inputs. When a firm's input choices are correlated with its productivity shocks, then this causes an endogeneity problem that leads to biased estimates of productivity.

In this paper, I show that the standard methods for controlling for these sources of bias cannot be applied in this setting with exporting firms. As a solution, I develop a new approach for estimating production functions that accounts for exporting firms. This approach is based on the method introduced in Gandhi, Navarro, and Rivers (2009). Using data on Colombian manufacturing firms, I find that the measures of productivity used in the literature over-estimate the productivity advantage of exporting firms relative to non-exporting firms. In fact, after controlling for the sources of bias contained in the commonly used measures, I find no difference in productivity based on export status.

This lack of correlation between productivity and export status is inconsistent with a model in which heterogeneity in productivity is the primary determinant of export status, and is suggestive of the existence of alternative drivers of exporting decisions. Additional justification for this comes from the fact that export intensities—the fraction of output that a firm exports—exhibit high variance in the data and are highly correlated over time. Since this cannot be explained by differences in productivity, this suggests the presence of an additional source of firm-specific heterogeneity that is both persistent and related to exporting decisions. I show that this heterogeneity, which is recoverable from the data, can enter the model through differences in the marginal costs and marginal benefits of exporting.

The rest of the paper is structured as follows. In Section 2, I review the related literature on exporting and productivity. In Section 3, I discuss the potential sources of bias in the measures of productivity used in the literature. In Section 4, I compute the two commonly used measures of productivity using data on Colombian manufacturing firms, and I discuss the effect of using value-added as opposed to gross output specifications of production. In Section 5, I present a method to control for unobserved prices that accounts for the fact that firms sell output both domestically and abroad. In Section 6, I introduce a new method for estimating production functions that both controls for endogeneity due to unobserved productivity and allows for gross output specifications of the production function. Section 7 presents the estimating model that results from assembling all of the components (controlling for exporting, prices, endogeneity of inputs, and gross output specifications). In Section 8, I describe the data, present the empirical results, and discuss the implications for models of exporting. Section 9 concludes.

2 Related Literature on Exporting and Productivity

There are many empirical papers that examine the relationship between firm-level productivity and export status. Most find evidence of a positive correlation between the two. This has prompted a debate as to the cause of this correlation. In a survey of this literature, covering 45 papers and 33 countries, Wagner (2007) notes that the majority of the literature supports the hypothesis of self-selection resulting from the presence of fixed costs of entry into export markets. These fixed costs include identifying and informing potential foreign customers, learning about relevant foreign laws and standards, and forming relationships with distributors. Only the most productive firms find the export market sufficiently profitable to justify paying these fixed costs of exporting, which generates a positive correlation between exporting and productivity. Bernard and Jensen (1999), Clerides, Lach, and Tybout (1998), and Aw, Chung, and Roberts (2000) are the most frequently cited empirical papers in support of this view.

The theoretical foundation for the hypothesis of self-selection is formalized in Melitz (2003). Melitz presents a model with heterogeneous firms and analyzes the effects of intra-industry trade. Each firm produces a different variety of a good in a monopolistically competitive setting. One implication of this model is that, in the presence of fixed costs of exporting, opening the economy to trade induces only the most productive firms to select into exporting. (In the absence of fixed costs of exporting, all firms would export in this model.)

An alternative hypothesis in the literature is that the causation runs in the other direction—that the act of exporting causes increases in firm-level productivity through learning. Exporting firms learn from international trading partners and competitors and use this knowledge to increase their productivity. Although this hypothesis has received less support in the literature, some papers find evidence of learning, particularly in developing countries. Both De Loecker (2007) for Slovenia and Van Biesebroeck (2006) for sub-Saharan Africa find evidence that firm-level productivity increased subsequent to exporting, which supports the learning hypothesis.

Finally, Bernard, Eaton, Jensen, and Kortum (2003) (henceforth BEJK) suggests a third explanation. In BEJK, geographic barriers, which include transportation costs, language barriers, and tariffs, generate heterogeneous marginal costs of exporting. Within each variety of a product, firms compete under Bertrand competition and there are no fixed costs of exporting. Since domestic firms do not face these geographic barriers, in order for a foreign firm to successfully export, it must be able to price more competitively than the domestic firms in the destination country. This implies that, on average, exporting firms are more productive than non-exporting firms.

3 Measures of Productivity

The majority of the empirical trade literature deals with the issue of firm-level productivity being unobservable in one of two ways. First, productivity is approximated by labor productivity, which can usually be measured directly in the data as the ratio of output (measured by either deflated revenues or real value added) to labor input (measured by either real wages, workers, or hours). In addition to being directly observable, the use of labor productivity has the advantage of not requiring data on capital levels, functional form assumptions on the production function, or the estimation of a production function. The main disadvantage of this approach is that labor productivity does not reflect true productivity differences among firms. Labor productivity, by construction, is a function of the other inputs. It is also endogenous as both output and labor input are chosen by the firm. In response to these disadvantages, many papers in the literature also attempt to recover measures of total factor productivity (TFP). Consider a generic production function $F(K, L, M)$, where K denotes capital, L denotes labor, and M denotes intermediate inputs. The amount of output, Q, that a firm produces is given by the following expression:

$$Q = A \times F(K, L, M), \tag{1}$$

where A denotes TFP.[1]

A common approach in the literature is to measure TFP as the residual from a linear regression of the log of real value added on the log of capital and labor. However, there are several reasons to believe that estimates of productivity generated using this method are biased, as I explain in what follows.

A lengthy literature in industrial organization is focused on estimating production functions and TFP. The primary concern in this literature, introduced by Marschak and Andrews (1944), is that if the firm's input decisions respond to productivity shocks, then input levels will be correlated with unobserved productivity, leading to an endogeneity bias. This is often referred to as the "transmission bias" or "simultaneity bias."

A second issue, also first suggested by Marschak and Andrews (1944), but not addressed until more recently by Klette and Griliches (1996), arises because in most datasets the output price and quantity are not observed separately. Rather, only revenues are observed. Since revenues are the only measure of output, unobserved prices are also present in the error term. If a firm's input choices are correlated with its output price, then an additional potential endogeneity problem arises, called the "omitted price bias." Moreover, even if there was not an omitted price bias, the residual would still contain both productivity and price, so the resulting measure of productivity would be confounded with unobserved prices.

A third issue is the use of value-added specifications of the production function as opposed to gross output specifications. This has received much less attention in the literature. Nominal value added, VA, is defined as the difference between the value of output (revenue) and the value of intermediate inputs:

$$\text{VA} = (P \times Q) - \left(P^M \times M\right),$$

where P is the price of output and P^M is the price of intermediate inputs.[2] Value added can then be written as a function of the "primary inputs" of capital and labor,

[1] This measure of productivity captures Hicks-neutral differences in the efficiency of firms. If firm 1 has a TFP that is twice that of firm 2, then with the same amount of inputs, firm 1 can produce twice as much output as firm 2.

[2] Real value added is calculated by deflating nominal value added directly or by deflating revenue and the value of intermediate inputs separately with different deflators (double-deflated value added).

$$\mathrm{VA} = \widetilde{A} \times \widetilde{F}(K, L), \tag{2}$$

where $\widetilde{F}(\bullet)$ is the value-added production function and $\widetilde{A}$ is the value-added approximation to TFP.

The use of value added as a measure of output is popular for a number of reasons. First, value added can be aggregated to measure the total output of an industry or set of industries without double-counting intermediate inputs to production. Second, a value-added production function relates output to labor and capital (but not intermediate inputs), which results in fewer parameters to estimate. Third, the recent structural methods for recovering productivity via the estimation of a production function cannot, in general, handle gross output specifications. The reason is that these methods of Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2006) cannot generally identify coefficients on intermediate inputs. As a result, intermediate inputs have to be subtracted out, and a value-added specification used.

A problem with value added is that it is not the natural measure of the output of the technology of a firm. A firm transforms inputs (including intermediate inputs) into output. Without intermediate inputs, output cannot be produced. In addition, by subtracting out the value of intermediate inputs, value added ignores any potential substitution between intermediate inputs and the "primary inputs" of capital and labor.

In addition to concerns about value added as a concept, the value-added specification is generally not a valid approximation to the gross output specification. Bruno (1978) and Basu and Fernald (1995, 1997) show that there is a limited and restrictive set of conditions under which the parameters of a value-added production function correspond to those of a gross output production function. In particular, Basu and Fernald (1995, 1997) show that when the assumptions of perfect competition and constant returns to scale are violated, the parameters of the value-added specification do not correspond to those of the gross output specification. Since many of the recent theoretical models of trade involve imperfect competition and increasing returns to scale, this seems especially problematic.

An additional reason for preferring a gross output specification derives from the literature that attempts to deal with the "omitted price bias." When prices are unobserved, and output is only measured in revenues, demand has to be modeled in order to control for prices. The problem is that the concept of demand for value added is generally not meaningful. Consumers have demand for final products, and do not care about how much value the firm added to the product.

4 Value-Added versus Gross Output Specifications

In this section, I first replicate the empirical finding that exporters are more productive than non-exporters. Using a dataset on Colombian manufacturing firms, I compute the two most commonly used measures of productivity in the literature. The data are described in more detail in Section 8. For each 3-digit industry, I compute labor productivity as real value added divided by the number of hours worked.[3] Following the empirical trade literature, I also estimate TFP as the residual from an OLS regression of real value added on capital and hours worked:

$$va_{jt} = \alpha l_{jt} + \beta k_{jt} + \nu_{jt}, \tag{3}$$

where va_{jt} is the log of the real value added of firm j in period t, l_{jt} is log labor, k_{jt} is log capital, and the residual ν_{jt} is log TFP (which is equal to the log of $\widetilde{A}$ from equation (2)).[4] I then calculate the productivity premia of exporters as the percentage difference in average productivity between exporting and non-exporting firms. I report these productivity premia in Table 1. The results are consistent with the literature. Exporting firms appear more productive than non-exporting firms in almost all industries, and in many cases the differences are large. In most cases, the premium is higher for labor productivity, which is reflective of the fact that exporters are on average 7 times larger in terms of capital stocks.

As opposed to labor productivity, the concept of TFP captures true efficiency differences across firms. However, estimates of TFP based on value-added may not reflect true productivity differences due to the strong restrictions underlying the approximation. In order to examine the empirical relevance of using value-added, I re-estimate TFP by replacing the value-added specification in equation (3) with the following equation for real gross output:

$$go_{jt} = al_{jt} + bk_{jt} + cm_{jt} + \mu_{jt}, \tag{4}$$

where go_{jt} is the deflated value of log gross output and the residual μ_{jt} is log TFP (which is equal to the log of A from equation (1)). The resulting productivity premia are presented in Table 2. As Table 2 illustrates, the productivity premia based on estimates of TFP from a gross output specification tell a significantly different story than the other two measures. In most cases the premium based on gross output is smaller than its value-added counterpart, and in several cases it is negative. The disparate results yielded by the value-added and gross output specifications provide evidence that the value-added specification is not a good approximation to the gross output specification. Furthermore, the gross output results provide

[3]In the data, firms are classified by industry according to the ISIC Rev. 2 classification.

[4]Throughout this paper I will use lower-case letters to denote logs and upper-case letters to denote levels.

much weaker evidence for a systematic relationship between exporting and productivity, and call into question the conclusion that differences in productivity are the primary driver of exporting decisions.[5]

While using gross output specifications addresses the problem of using meaningful measures of output when recovering productivity, the exercise is far from complete. As described earlier, I still need to control for the biases caused by unobserved prices and unobserved productivity. I address each of these separately in the next two sections. As a consequence of the results described above, I will use gross output specifications for the remainder of my analysis.

5 Controlling for Unobserved Prices

In empirical work on production function and productivity estimation, it is often assumed that output measured in quantities can be observed directly. However, in most datasets we do not observe quantities (or prices) but instead observe only revenues. To see why this introduces further complications, notice that the log of revenue can be expressed as follows:

$$r_{jt} = \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + \omega_{jt} + p_{jt},$$

where p_{jt} is output price and ω_{jt} is log TFP. The typical solution in the literature is to use price deflators to transform revenues into quantities. The problem with this approach is that firm-specific price deflators are typically not available.[6] Therefore, the difference between the firm's price and the price deflator, $p_{jt} - \overline{p_t}$, remains in the error term,

$$r_{jt} - \overline{p}_t = \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + \omega_{jt} + \left(p_{jt} - \overline{p}_t\right).$$

If firms possess any market power, then deflating with an aggregate price index will lead to biased estimates of the production function and, consequently, biased estimates of firm-level productivity. The reason is that $p_{jt} - \overline{p_t}$ reflects a firm's market power, which is likely correlated with input demands, leading to endogeneity. This is known as the "omitted price bias." Moreover, even if endogeneity was not a concern, there is a more obvious source of bias in the productivity estimates themselves: they will be the sum of true productivity (ω_{jt}) and unobserved price deviations $(p_{jt} - \overline{p}_t)$. Consequently, controlling for unobserved prices

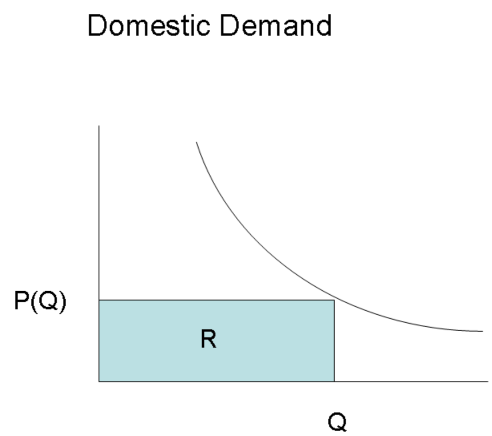
[5] The fact that the gross output specification weakens the apparent relationship between productivity and export status is illustrative of a larger problem. The widespread use of value-added production functions in the broader literature may be resulting in biased conclusions to other questions which rely on estimating a production function.

[6] See Marschak and Andrews (1944) and, for a more recent treatment, Klette and Griliches (1996).

will be important for both reasons.

In the absence of data on prices, solving this problem requires modeling demand (and therefore prices). One option is to use a single constant-elasticity residual demand curve.[7] However, when some firms export, these firms not only sell their goods domestically, but also abroad. As a result, the use of a single demand function to model prices will not be sufficient to measure the "price" of the firm's output, as the average price received by the firm is a weighted average of the domestic and foreign prices. In fact, using a single demand system will *over-estimate* the relative productivity advantage of exporting firms. The intuition for this can be seen in Figure 1.

Figure 1: Inferring Price and Quantity from a Single Demand



Suppose that we observe a firm earning total revenues R. By ignoring the foreign market and using a single domestic demand curve, a quantity Q and price $P(Q)$ are inferred.[8] Given a function for the costs of production, $C(Q)$, the implied profits are $\Pi(Q) = R - C(Q)$. If that firm is exporting some of its output, and there are any costs of exporting, then the firm must be making a higher profit (excluding these costs of exporting) than $\Pi(Q)$. In order for profits to be higher, the costs of production must be lower than $C(Q)$. Under the assumption that the cost function is strictly increasing in quantity, this implies that the total quantity (across both markets) being produced by the exporting firm is $\widetilde{Q} < Q$. As a result, since the model over-estimates the quantity that this exporting firm is producing, it will over-estimate the true productivity of that firm. This is crucial since it directly biases

[7]See Klette and Griliches (1996) and De Loecker (2009) for examples.

[8]The constant-elasticity demand generates a revenue function that is strictly increasing in quantity. Therefore, a given revenue implies a unique quantity.

the relationship I am interested in examining in favor of the common finding that exporters are more productive.

5.1 Model of Prices

In order to control for unobserved prices, I model domestic and foreign demand separately using a standard approach in the trade literature. Within a given industry, firms are assumed to produce horizontally differentiated products and compete in a monopolistically competitive setting. Firms face symmetric constant-elasticity demand curves. There are separate domestic and foreign markets.[9] The demand functions are given by the following equations:

$$P_{jt}^{D} = \overline{P}_{t}^{D}\left(\frac{Q_{jt}^{D}}{\overline{Q}_{t}^{D}}\right)^{\frac{1}{\eta}} \quad ; \quad P_{jt}^{E} = \tau\overline{P}_{t}^{E}\left(\frac{Q_{jt}^{E}}{\overline{Q}_{t}^{E}}\right)^{\frac{1}{\eta}}, \tag{5}$$

where $\overline{Q}_t$ is an aggregate demand shifter related to industry demand at time t,[10] $\overline{P}_t$ is the price deflator, Q_{jt} and P_{jt} are the firm-specific quantity and price, and τ denotes marginal costs of exporting.[11] Superscript D denotes variables related to the domestic market, and superscript E denotes variables related to the export market. The elasticity of demand, η, is assumed to be the same in both markets.[12] Both domestic and foreign prices are measured in nominal units of the domestic currency.[13] In Section 7, I show how these demands are used to control for unobserved prices when estimating productivity.

6 Controlling for Endogeneity Due to Unobserved Productivity

As stated earlier, the concern that inputs are endogenous due to their correlation with unobserved productivity has received a lot of attention in the industrial organization literature. Several methods have been suggested for dealing with this endogeneity.[14] If instruments

[9] This is the setting, for example, in Melitz (2003), with the exception that in that paper there are N symmetric foreign markets. For simplicity, I focus on one foreign market. The results of this paper can be generalized to account for multiple foreign markets.

[10] I form the quantity index as a weighted-average of deflated revenues, where the weights are the market shares.

[11] This is the standard "iceberg" assumption that the marginal costs of exporting are proportional to the value of output that is exported.

[12] This assumption can also be relaxed to allow for the elasticity of demand in the domestic market to differ from the elasticity of demand in the foreign market. It complicates the algebra and adds one more parameter to estimate, but all of the results that I show in the paper still hold.

[13] This implies that the foreign price deflator also captures the exchange rate.

[14] For a summary of these methods and their relative advantages and disadvantages, see Griliches and Mairesse (1998).

are available, then instrumental variables techniques are a natural solution. However, valid instruments for the endogenous inputs, such as input prices, are not typically available. Another approach is to use panel data techniques, but if productivity varies across both firms and time, then it cannot be totally differenced out as fixed effects, and the remaining residual term is still correlated with inputs. In addition, panel data methods remove a lot of the variation in the data, which is needed to identify the parameters of the model.

This leaves the recently-developed proxy variable methods of Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2006).[15] However, for my setting, these methods cannot be employed either. The reason, as I discuss below, is that they cannot generally be used to estimate gross output specifications of the production function. This is due to a problem with the identification of coefficients related to static and competitive inputs that occurs with this approach. This is particularly important given both the empirical evidence in Section 4 regarding the use of value-added versus gross output specifications, and the fact that I need to model demand (for gross output) to control for unobserved prices.

To illustrate, I briefly describe the proxy variable approach. The key insight of these methods is that the firm's demand for a proxy variable λ_{jt} (either investment in physical capital or intermediate input demand) is a monotonic function of the state variables of the firm, productivity ω_{jt} and capital k_{jt}:[16]

$$\lambda_{jt} = g_t\left(\omega_{jt}, k_{jt}\right).$$

Since this relationship is a strictly monotonic function of just one unobservable, it can be inverted to express the unobserved productivity in terms of observable variables,

$$\omega_{jt} = g_t^{-1}\left(\lambda_{jt}, k_{jt}\right).$$

This expression for productivity then replaces productivity in the production function. Temporarily ignoring unobserved prices, log quantity can be expressed as follows:

$$y_{jt} = q_{jt} + \varepsilon_{jt} = \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + g_t^{-1}\left(\lambda_{jt}, k_{jt}\right) + \varepsilon_{jt},$$

where q_{jt} denotes anticipated output in quantities and y_{jt} denotes observed output that is subject to ex-post productivity shocks and/or measurement error, ε_{jt}. This equation no

[15] An alternative to estimating productivity is to recover TFP using an index number approach. These techniques do not suffer from endogeneity concerns as they do not involve running a regression. However, they impose several potentially strong assumptions in order to recover TFP, such as constant returns to scale and perfect competition in the input and output markets.

[16] In Ackerberg, Caves, and Frazer (2006), labor is also assumed to be a state variable.

longer suffers from endogeneity as ε_{jt} is assumed to be uncorrelated with inputs. In addition, it separates ε_{jt} from ω_{jt}, as ω_{jt} does not appear on the right-hand side.

The model also assumes that productivity evolves according to a first-order Markov process,

$$\omega_{jt} = h_t(\omega_{jt-1}) + \xi_{jt}. \tag{6}$$

The innovations to productivity, ξ_{jt}, are assumed to be independent of all inputs that are determined before ξ_{jt} is realized. Then, for a given vector of the parameters of the production function and the inverted proxy equation $\left(g_t^{-1}(\bullet)\right)$, productivity can be formed, and the regression in equation (6) can be computed. The residual of this regression is then interacted with moments to estimate the parameters.

The problem with these methods, as recently pointed out in Bond and Söderbom (2005) and Ackerberg, Caves, and Frazer (2006), is that they cannot generally be used to identify the coefficients on static and competitive inputs. The reason is that there is nothing in the model that varies these inputs (*e.g.,* intermediate inputs) separately from productivity and the other inputs.[17] As a result, a value-added specification, where intermediate inputs are subtracted out and then ignored, has to be used when employing these methods.

6.1 The Share Equation Approach

Since a gross output production function is required for the problem at hand, I need a method that can estimate gross output models. Since none of the methods in the literature can be applied to the problem I study, I apply a new method for estimating production functions from Gandhi, Navarro, and Rivers (2009) (henceforth GNR). This method controls for endogeneity due to unobserved productivity in a gross output setting. I generalize the approach developed in GNR to a setting of imperfect competition in which there are both domestic and foreign markets. I also show how this method can be extended to deal with the complications due to the presence of unobserved prices when some firms are exporting.

The key insight of GNR is that there is important unused information contained in the

[17] Under some specific assumptions on the data generating process, the proxy variable method can be used when investment is the proxy variable. However, using investment as the proxy requires dropping all observations for which there is zero investment in physical capital. This can lead to a significant loss of data, as pointed out by Levinsohn and Petrin (2003). In my dataset, this would result in a loss of 40% of the observations.

In addition to the loss of data, the use of these methods would require the existence of a specific source of variation in demand for intermediate inputs. In particular, this source of variation would have to either have no dynamic effects, or would have to be observable. If the source of variation was unobserved and had dynamic effects, then this would invalidate the necessary assumption that productivity is the only unobservable that affects investment decisions.

firm's input shares. The shares not only provide identifying information for the parameters of the production function but also allow for richer forms of firm-level heterogeneity.[18] To keep the analysis straightforward, I first illustrate the method under the assumption that firms are perfectly competitive in the output market, which is the setting typically analyzed in the literature on production function estimation.

For a generic production technology, consider a firm's maximization problem with respect to a static and competitive input, such as intermediate inputs, M_{jt}:

$$\max_{M_{jt}} P_{jt}Q_{jt} - P_t^M M_{jt} \quad = \quad \max_{M_{jt}} P_{jt}F\left(K_{jt}, L_{jt}, M_{jt}\right)e^{\omega_{jt}} - P_t^M M_{jt}.$$

This results in the following first-order condition:

$$P_{jt}F_M\left(K_{jt}, L_{jt}, M_{jt}\right)e^{\omega_{jt}} - P_t^M = 0,$$

where $F_M(\bullet)$ denotes the partial derivative of $F(\bullet)$ with respect to intermediate inputs, M. Notice that the first-order condition contains both unobserved output prices and unobserved productivity. These are the two sources of endogeneity that are causing problems in the estimation to begin with. However, this expression can be re-arranged in terms of the share of intermediate inputs in total revenue:

$$\frac{P_t^M M_{jt}}{P_{jt}Q_{jt}} = \frac{F_M\left(K_{jt}, L_{jt}, M_{jt}\right) \times M_{jt}}{F\left(K_{jt}, L_{jt}, M_{jt}\right)}. \tag{7}$$

As equation (7) shows, the shares of intermediate inputs can be expressed in terms of the production function, the first derivative of the production function with respect to intermediate inputs, and the level of intermediate inputs.

Taking logs, replacing the product of price and quantity with revenue, and accounting for ex-post productivity shocks and measurement error yields what GNR calls the share equation:

$$s_{jt} \equiv \ln\left(\frac{P_t^M M_{jt}}{R_{jt}}\right) = \ln\left(F_M\left(K_{jt}, L_{jt}, M_{jt}\right)\right) - \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + m_{jt} - \varepsilon_{jt}, \tag{8}$$

where R_{jt} denotes observed revenues. This equation separates ε_{jt} from ω_{jt} (note that ω_{jt}, which appeared in the first-order condition, is now contained in the left-hand side term) and

[18] See Gandhi, Navarro, and Rivers (2009) for a discussion of the benefits of being able to allow for other forms of unobserved heterogeneity (in addition to productivity).

collapses unobserved prices within observed revenues. Furthermore, it provides an additional source of identifying information for the production function directly through the first term: the derivative of the production function with respect to intermediate inputs. The share equation together with the production function can be used to express the two unobservables, ε_{jt} and ω_{jt}, as a function of the parameters of the production function:[19]

$$\begin{aligned} s_{jt} \equiv \ln\left(\frac{P_t^M M_{jt}}{R_{jt}}\right) &= \ln\left(F_M\left(K_{jt}, L_{jt}, M_{jt}\right)\right) - \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + m_{jt} - \varepsilon_{jt} \\ y_{jt} &= \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + \omega_{jt} + \varepsilon_{jt} \end{aligned}$$

The same assumption made in the proxy variable methods—that productivity evolves in a Markovian fashion—can be used to form estimates of the innovation to productivity, ξ_{jt}. Moments with both ε_{jt} and ξ_{jt} can then be formed to recover the production function and hence firm-level productivity (ω_{jt}).

6.1.1 Imperfect Competition

The approach developed in GNR generalizes well to other specifications of the underlying model and can be used under various data restrictions (*e.g.,* observing only revenues as opposed to quantities). In particular, it can handle imperfect competition. When firms charge constant markups, which is the case with the constant elasticity of demand curves I introduced in Section 5.1, the share equation remains the same, with the exception of one additional term, which is the log of the markup:[20]

$$s_{jt} \equiv \ln\left(\frac{P_t^M M_{jt}}{R_{jt}}\right) = \ln\left(F_M\left(K_{jt}, L_{jt}, M_{jt}\right)\right) - \ln\left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + m_{jt} - \ln\left(\frac{\eta}{\eta+1}\right) - \varepsilon_{jt}. \tag{9}$$

Moreover, the share equation is the same for both exporting and non-exporting firms.[21] These results highlight an appealing feature of the approach in GNR: relaxing assumptions within this framework results in equations that are still easy to use.

[19] Note that equation (8) could be estimated directly and non-parametrically without specifying a functional form for the production function. The right-hand side of the equation is just a function of the inputs to production and measurement error, which is uncorrelated with inputs.

[20] With the constant-elasticity demand system, the markup that firms optimally charge is $\left(\frac{\eta}{\eta+1}\right)$, where η is the elasticity of demand.

[21] The fact that the share equation does not depend on export status is a result of the assumption that the domestic and foreign demand elasticities are the same. When this assumption is relaxed, the share equation for exporters becomes a slightly more complicated function that includes an extra parameter—the foreign demand elasticity—but the intuition and results of the model still hold.

Thus far I have shown how the share equation can be used to control for the endogeneity due to unobserved productivity. I have also shown how the share equation can be modified to account for both imperfect competition and exporting. The remaining step is to use the model of demand from Section 5.1 to control for unobserved prices in the production function.

7 The Empirical Model: Revenues and Exporting

In this section, I assemble all of the pieces: the demand systems to control for unobserved prices, the share equation to control for endogeneity of inputs due to unobserved productivity, and the gross output specification. I begin by showing that the "revenue production function" needs to be adjusted to account for the different sources of revenues. I then show how to solve the problem that this creates—by pairing the demand systems introduced in Section 5.1 with the firm's optimal allocation of output across the domestic and foreign markets. This leads to what I call the "domestic revenue production function," which paired with the share equation, constitutes my empirical model.

As I discussed in Section 5, when some firms export, their revenues come from both domestic and foreign markets, each with their own demands. Total revenues for exporting firms are then the sum of domestic and foreign revenues. By replacing the prices using the demand systems in Section 5.1, total revenues can be expressed as follows:

$$\begin{aligned} R_{jt} &= R_{jt}^{D} + R_{jt}^{E} = P_{jt}^{D} Q_{jt}^{D} + P_{jt}^{E} Q_{jt}^{E} \\ &= \frac{\overline{P}_{t}^{D}}{\left(\overline{Q}_{t}^{D}\right)^{\frac{1}{\eta}}} \left(Q_{jt}^{D}\right)^{1+\frac{1}{\eta}} + \tau \frac{\overline{P}_{t}^{E}}{\left(\overline{Q}_{t}^{E}\right)^{\frac{1}{\eta}}} \left(Q_{jt}^{E}\right)^{1+\frac{1}{\eta}} . \end{aligned}$$

Two main problems with this expression prevent it from being used directly. First, marginal costs of exporting, τ, are not observed. Second, there is no model for the quantity of output sold on the domestic market, Q_{jt}^{D}, separate from the quantity of output sold on the foreign market, Q_{jt}^{E}. There is only a model for total quantity: the production function. In order to address these challenges, I derive a model for the *domestic* revenues of the firms separately from the *foreign* revenues. In doing so, I am able to estimate the elasticity of demand and all of the parameters of the production function, including unobserved productivity. I accomplish this by taking advantage of another static first-order condition of the firm. Specifically, I look at the firm's maximization problem with respect to its allocation of output between the domestic and foreign markets.

I define the fraction of a firm's physical output that is sold on the domestic market as

$$\theta_{jt} = \frac{Q_{jt}^D}{Q_{jt}^D + Q_{jt}^E}.$$

As I show in Appendix A, when the elasticities of demand are the same in both markets, firms choose to allocate output such that the prices received by the firm in both markets are equal. Since the prices are equal, this implies that the division of quantities across markets, which *is not observed* in the data, is equal to the division of revenues across markets, which *is observed* in the data.[22] As a result, I can derive an expression for domestic quantity as a function of the division of revenues and the production function,

$$Q_{jt}^D = \theta_{jt} \times Q_{jt} = \left(\frac{R_{jt}^D}{R_{jt}^D + R_{jt}^E} \right) \times F\left(K_{jt}, L_{jt}, M_{jt}\right) \times e^{\omega_{jt}}.$$

By using the domestic demand curve, I can obtain a model for the observed measure of domestic output (*i.e.*, domestic revenues). The log of deflated domestic revenues, $\widetilde{r}_{jt}^D$, is given by the following equation:

$$\widetilde{r}_{jt}^D = \left(1 + \frac{1}{\eta}\right) \ln \left(\frac{R_{jt}^D}{R_{jt}^D + R_{jt}^E} \right) - \frac{1}{\eta} \overline{q}_t^D + \left(1 + \frac{1}{\eta}\right) \left[\ln \left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + \omega_{jt} \right] + \varepsilon_{jt}.$$

This equation controls for unobserved prices when some firms are exporting. Additionally, when combined with the share equation, it controls for endogeneity due to unobserved productivity, all within a gross output setting.

Putting the two equations together yields the set of estimating equations:

$$\begin{aligned} s_{jt} &= \ln \left(F_M\left(K_{jt}, L_{jt}, M_{jt}\right)\right) - \ln \left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + m_{jt} - \ln \left(\tfrac{\eta}{\eta+1}\right) - \varepsilon_{jt} \\ \widetilde{r}_{jt}^D &= \left(1 + \tfrac{1}{\eta}\right) \ln \left(\tfrac{R_{jt}^D}{R_{jt}^D + R_{jt}^E}\right) - \tfrac{1}{\eta} \overline{q}_t^D + \left(1 + \tfrac{1}{\eta}\right) \left[\ln \left(F\left(K_{jt}, L_{jt}, M_{jt}\right)\right) + \omega_{jt}\right] + \varepsilon_{jt}. \end{aligned} \tag{10}$$

Note that the share in the first equation is a function of nominal total revenues, and the dependent variable of the second equation is deflated domestic revenues.

[22] When the elasticities of demand in the domestic and foreign markets are not the same, the ratio of prices is equal to the ratio of the markups.

7.1 Learning-by-Exporting

As discussed in Section 2, some papers in the trade literature suggest that exporting leads to an increase in a firm's productivity. If this is the case, then the process governing the evolution of productivity is now a controlled Markov process, as a firm's decision about whether or not to export has an effect on its realization of productivity. This needs to be taken into account in the estimation strategy. Under the timing assumption that lagged export status, D_{jt-1}, affects the realization of productivity, the process for productivity can be written as follows:[23]

$$\omega_{jt} = \tilde{h}\left(\omega_{jt-1}, D_{jt-1}\right) + \xi_{jt}. \tag{11}$$

Testing for the presence of learning-by-exporting is therefore embedded directly into the estimation procedure.

8 Data and Estimation Results

8.1 Data

My data come from an annual census of Colombian manufacturing plants over the period 1981-1991. The data cover all firms with 10 or more employees. This dataset has been used previously in several studies (for example, Roberts and Tybout (1997) and Clerides, Lach, and Tybout (1998)) and contains information about each plant's capital stocks, investment flows, expenditures on labor and intermediate inputs, number of workers, wages, value of total output, and value of output that is exported.

For the structural estimates I focus on the largest 3-digit industry, which is Apparel (industry 322). Since I model firms as being monopolistically competitive, I need an industry that contains a large number of firms for this assumption to be valid. Additionally, choosing the largest industry yields the most observations for the estimation. After dropping observations with missing values, a total of 4,490 observations remain for 732 firms, of which 18% exported in at least one year. In Table 3, I report some summary statistics for the data.

8.2 Parametrization

So far, my discussion of the method of GNR that I employ has not relied on any functional form assumptions on the production function. In my estimation, I use a CES specification

[23] In the estimation procedure I assume that $\tilde{h}\left(\omega_{jt-1}, D_{jt-1}\right)$ is linear and additively separable in its arguments. I can allow for more general functions of ω and D. The only key assumption here is separability.

of the production function. The parametric versions of the share equation and the domestic revenue production function for CES are:

$$\begin{aligned} s_{jt} \equiv \ln\left(\frac{P_t^M M_{jt}}{R_{jt}}\right) &= -\ln\left(\alpha_l L_{jt}^\rho + \alpha_k K_{jt}^\rho + \alpha_m M_{jt}^\rho\right) + \ln\left(\alpha_m r\right) - \ln\left(\frac{\eta}{\eta+1}\right) + \rho m_{jt} - \varepsilon_{jt} \\ \tilde{r}_{jt}^D &= \left(1+\frac{1}{\eta}\right)\ln\left(\frac{R_{jt}^D}{R_{jt}^D+R_{jt}^E}\right) + \left(1+\frac{1}{\eta}\right)\frac{r}{\rho}\ln\left(\alpha_l L_{jt}^\rho + \alpha_k K_{jt}^\rho + \alpha_m M_{jt}^\rho\right) \\ &\quad + \left(1+\frac{1}{\eta}\right)\omega_{jt} - \frac{1}{\eta}\overline{q}_t^D + \varepsilon_{jt}, \end{aligned} \tag{12}$$

where α_l, α_k and α_m are share parameters, ρ is the CES parameter ($\left(\frac{1}{1-\rho}\right)$ is the elasticity of substitution), and r is returns to scale.

8.3 Parameters of the Revenue Production Function

As a first step, I estimate a baseline version of the model under the assumption that firms are perfectly competitive. In this setting, all price variation is captured by the time-varying price index, so unobserved prices are perfectly controlled for by the price index. As a result, only endogeneity due to unobserved productivity needs to be controlled for. To do this, I implement a version of the share equation method under perfect competition. The estimating equations in this setting are:

$$\begin{aligned} s_{jt} \equiv \ln\left(\frac{P_t^M M_{jt}}{R_{jt}}\right) &= -\ln\left(\alpha_l L_{jt}^\rho + \alpha_k K_{jt}^\rho + \alpha_m M_{jt}^\rho\right) + \ln\left(\alpha_m r\right) + \rho m_{jt} - \varepsilon_{jt} \\ y_{jt} &= \frac{r}{\rho}\ln\left(\alpha_l L_{jt}^\rho + \alpha_k K_{jt}^\rho + \alpha_m M_{jt}^\rho\right) + \left(1+\frac{1}{\eta}\right)\omega_{jt} + \varepsilon_{jt}, \end{aligned} \tag{13}$$

where y_{jt} is total deflated revenues. The results for the baseline model are presented in Table 4.

The coefficients on each of the inputs in the production function $(\alpha_l, \alpha_k, \alpha_m)$ are not robust to differences in the scaling of inputs. As a result, I report mean input elasticities with respect to each of the inputs, as opposed to the parameter estimates themselves.[24] Standard errors are reported in parentheses below the parameter estimates.[25]

The results for the baseline model are reasonable. Output is most elastic with respect to intermediate inputs, and the labor elasticity is about twice that of capital. The CES

[24] In fact, the sum of these three parameters is not separately identified from the mean of productivity. Because of this, I normalize the sum of the three parameters to be one, and define $\alpha_m = 1 - \alpha_l - \alpha_k$.

[25] Standard errors for the share parameters (not shown), returns to scale, and the CES parameter are based on the asymptotic distribution. For the input elasticities and the elasticity of demand, standard errors are computed by sampling 5,000 sets of parameters from their asymptotic distribution, constructing these elasticities, and computing the standard deviation of the resulting elasticities.

parameter suggests an elasticity of substitution that is larger than for Cobb-Douglas. The results also suggest that the technology exhibits constant returns to scale. There is no estimate for the elasticity of demand or markups since firms are assumed to be perfectly competitive. These results are in line with typical production function estimates in the literature, which find roughly constant returns to scale and labor elasticities that are about twice as large as capital elasticities.

In Table 4, I present the estimates of the full model (equation (12)) where I account for all of the potential sources of bias. The estimates for the full model do not differ much in terms of the input elasticities or the CES parameter. The biggest difference is in the estimates of returns to scale and market power. I find moderate increasing returns to scale as well as markups of about 11%.[26] These results are consistent with other papers in the literature that address the omitted price bias (*e.g.,* Klette and Griliches (1996), Gandhi, Navarro, and Rivers (2009), and De Loecker (2009)). In addition these results are consistent with the recent trade theory in which firms face downward-sloping demand curves and exhibit increasing returns to scale.

8.4 Learning-by-Exporting

As I discussed in Section 7.1, I can directly test for evidence of learning-by-exporting by explicitly allowing the evolution of productivity to depend on lagged export status. I use the following specification for the process on productivity in the estimation:

$$\omega_{jt} = \delta_0 + \delta_1 \omega_{jt-1} + \delta_2 D_{jt-1} + \xi_{jt},$$

where the estimate of δ_2 denotes the percentage increase in productivity due to exporting in the previous period. The point estimate for δ_2 is -0.4% and is not statistically different from zero, which implies a lack of evidence for learning in the data. This result is not surprising given, as I show next, that I find no correlation between exporting and productivity.

8.5 Productivity Comparison

Given the parameter estimates from the full model, which corrects for the "transmission bias," the "omitted price bias" with exporting, and the bias from using a value-added rather than a gross output specification, I can recover an unbiased estimate of total factor productivity

[26]It is not surprising that I find increasing returns to scale given that I find evidence of market power. The baseline model, which ignores unobserved prices, finds constant returns to scale. That means that if a firm doubles it inputs, it doubles its output, which is measured as deflated revenues. However, if that firm has market power, then as it increases its output, its price decreases. Therefore, in order for a doubling of inputs to lead to a doubling of revenues, there must be increasing returns to scale in production.

for each firm. These estimates identify differences in efficiency across firms separately from differences in input levels or size. Using these estimates, I compute the productivity premia for exporters. I define a firm to be an exporter in each year in which the firm exports a positive amount, and a non-exporter in all other years. I find that exporters are not more productive than non-exporters. I present this result in Table 5. For comparison, I also report the productivity premia for the Apparel industry (322) that were derived from the two common measures of productivity in the empirical trade literature: labor productivity and TFP measured as the residual from a OLS regression of log value added on log inputs. (These results were originally reported in Table 1.) Standard errors for the estimated premia are reported in parentheses below the point estimates.[27]

The point estimate from the full model suggests that exporters are 1% less productive than non-exporters, although the estimate is not statistically different from zero. This result differs sharply from the productivity premia obtained by measuring productivity as labor productivity or as the residual from an OLS regression of value added on inputs, which imply large statistically significant productivity advantages for exporters.

Since one might be concerned that these results, which compare mean productivity across export status, are driven by outliers and are not representative of the entire distribution, I also report the distributions of productivity by export status. These are reported in Figure 2. Although there is more variation in productivity for exporters, the distributions look very similar, which suggests that there are no systematic differences between exporting firms and non-exporting firms in terms of productivity.

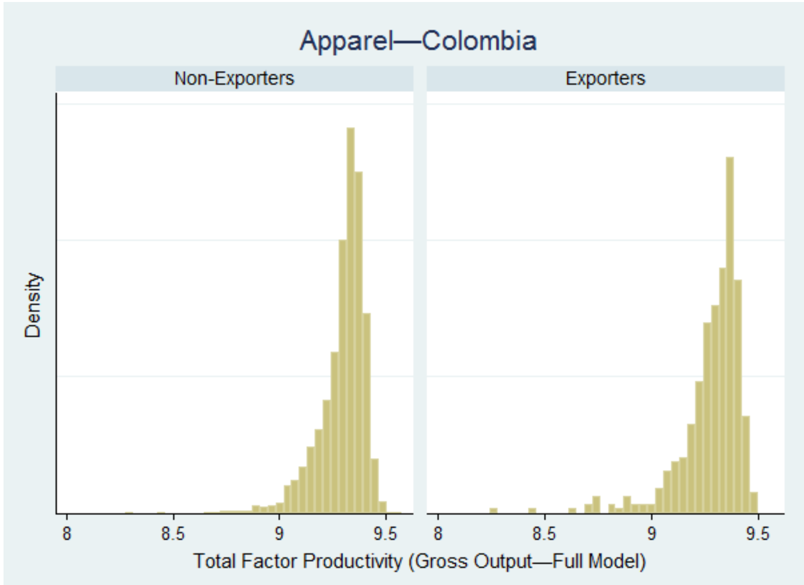
8.6 Implications for Models of Exporting

Once unbiased estimates of total factor productivity are obtained, exporters no longer appear to be more productive than non-exporters. This suggests that differences in technological efficiency are not the primary determinant of export status. This begs the question that, if productivity differences do not explain heterogeneous exporting decisions, then what does? One possibility is differences in capital stock. As shown in Figure 3, exporters are on average much larger, in terms of capital, than non-exporters.

This is consistent with the basic mechanism in Melitz's model as he does not include capital in the model. If capital is not perfectly flexible, then persistent differences in capital

[27]For the productivity premia based on labor productivity and TFP from a value-added OLS specification, standard errors are computed via the non-parametric bootstrap. For the productivity premium based on TFP estimates from the full model using gross output, standard errors are computed by sampling 5,000 sets of parameters from their asymptotic distribution, computing firm-level productivity, computing the productivity premium for exporters, and then calculating the standard deviation of the resulting premia over the 5,000 samples.

Figure 2: Distribution of TFP by Export Status



stocks will operate like differences in productivity and generate the same type of selection into exporting, with the selection being on capital, rather than productivity. However, as Figure 3 illustrates, differences in capital are not sufficient to explain exporting decisions, as there are many large firms that do not export, and some small firms that do. This suggests that some other form of firm-level heterogeneity must be important in determining exporting decisions.

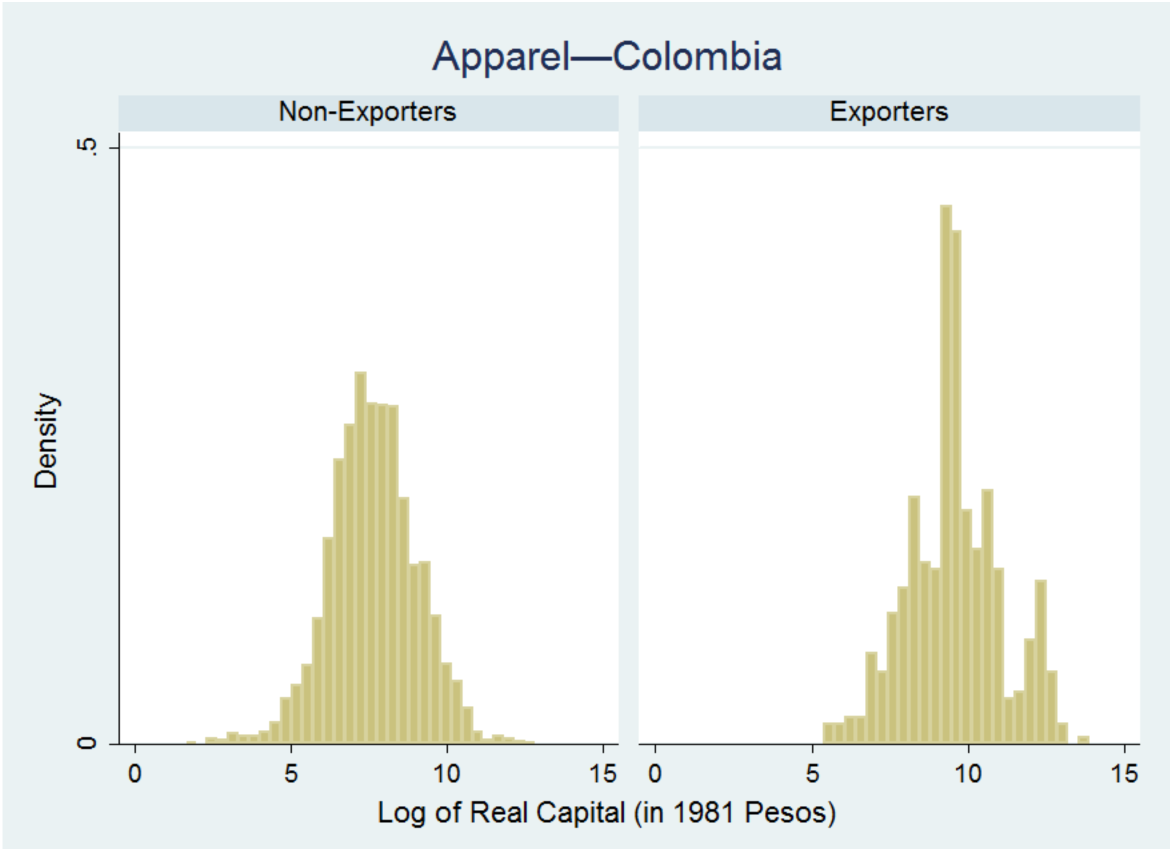
One benefit of the model that I have derived is that it contains additional information that suggests the potential sources of this heterogeneity. In Section 7, I showed how the model can be used to compute the share of total output that is sold on the domestic market, θ_{jt}. In Figure 4, I present a histogram of export intensity (the percentage of output sold on the foreign market), $1-\theta_{jt}$, conditional on a firm being an exporter. The figure is for the Apparel industry, but similar patterns hold in other industries.

As Figure 4 illustrates, there is a lot of heterogeneity in export intensity. This heterogeneity cannot be explained by differences in either productivity or capital stocks. Export intensity is only a function of the relative marginal revenues of the markets. Recall the demand equations from Section 5.1:

$$P_{jt}^{D} = \overline{P}_{t}^{D}\left(\frac{Q_{jt}^{D}}{\overline{Q}_{t}^{D}}\right)^{\frac{1}{\eta}} \quad ; \quad P_{jt}^{E} = \tau\overline{P}_{t}^{E}\left(\frac{Q_{jt}^{E}}{\overline{Q}_{t}^{E}}\right)^{\frac{1}{\eta}}.$$

As I show in Appendix A, profit maximization implies setting the marginal revenues equal to each other. This in turn implies that the optimal fraction of quantity sold on the domestic market, θ_{jt}^{*}, is given by the following equation:

Figure 3: Capital Stock by Export Status



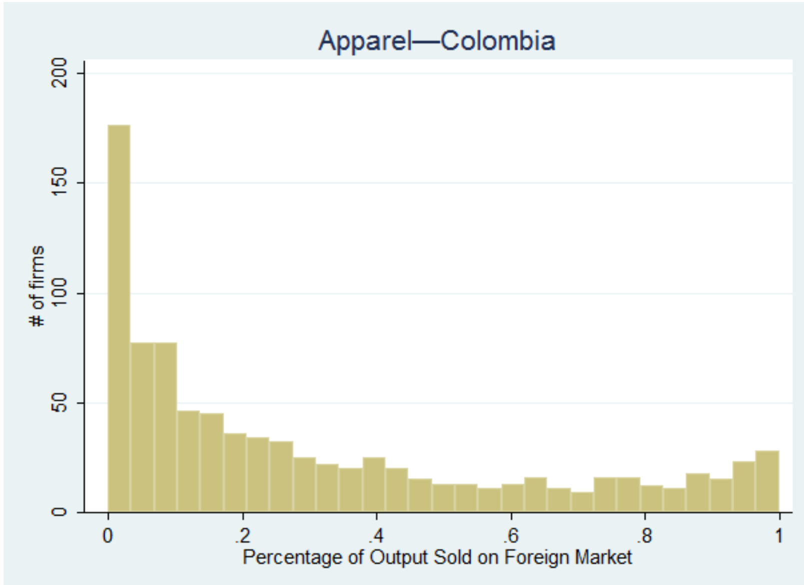
$$\theta^*_{jt} = \frac{\tau \frac{\overline{P}_t^{E\eta}}{\overline{Q}_t^E}}{\tau \frac{\overline{P}_t^{E\eta}}{\overline{Q}_t^E} + \frac{\overline{P}_t^{D\eta}}{\overline{Q}_t^D}}, \tag{14}$$

which is just a function of the relative price and quantity indices and the marginal exporting costs.

So far I have presented the marginal costs of exporting and the foreign price and quantity indices as varying across time but not across firm. However, none of my results have relied on this assumption. Rather, this was for clarity in the exposition. If different firms export to different countries and face different transportation costs, then these terms capturing the marginal costs and marginal benefits of exporting will be heterogeneous across firms as well. All of this heterogeneity can be expressed by an index,

$$\pi_{jt} \equiv \tau_{jt} \times \frac{\overline{P}_{jt}^E}{\left(\overline{Q}_{jt}^E\right)^{\frac{1}{\eta}}},$$

Figure 4: Export Intensity



where the terms on the right hand side are allowed to vary by firm.

By looking at the counterpart to the domestic revenue production function, we can see how this heterogeneity enters the model. Foreign revenues are given by the following equation:

$$\begin{aligned} r_{jt}^{E} = & \ln\left(\tau_{jt} \times \frac{\overline{P}_{jt}^{E}}{\left(\overline{Q}_{jt}^{E}\right)^{\frac{1}{\eta}}}\right) + \left(1+\frac{1}{\eta}\right)\ln\left(1-\theta_{jt}\right) \\ & + \left(1+\frac{1}{\eta}\right)\left[\frac{r}{\rho}\ln\left(\alpha_l L_{jt}^{\rho} + \alpha_k K_{jt}^{\rho} + \alpha_m M_{jt}^{\rho}\right) + \omega_{jt}\right] + \varepsilon_{jt}. \end{aligned} \tag{15}$$

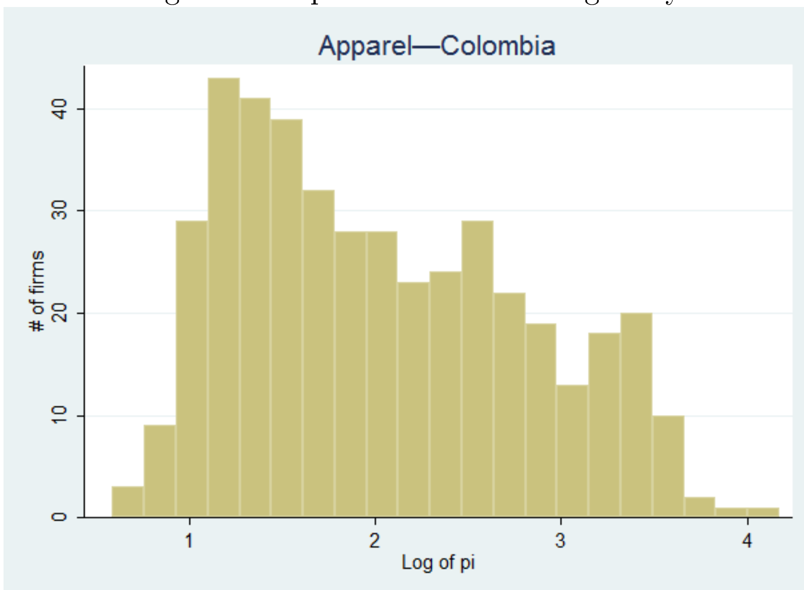
By substituting in π_{jt}, equation (15) can be written as a function of this heterogeneity,

$$\begin{aligned} r_{jt}^{E} = & \ln\left(\pi_{jt}\right) + \left(1+\frac{1}{\eta}\right)\ln\left(1-\theta_{jt}\right) \\ & + \left(1+\frac{1}{\eta}\right)\left[\frac{r}{\rho}\ln\left(\alpha_l L_{jt}^{\rho} + \alpha_k K_{jt}^{\rho} + \alpha_m M_{jt}^{\rho}\right) + \omega_{jt}\right] + \varepsilon_{jt}. \end{aligned} \tag{16}$$

I can now obtain an estimate of π_{jt} directly from equation (16). Note that I already have estimates of all of the other parameters of equation (16) from the estimation of the domestic revenue production function. I also have an estimate of TFP (ω_{jt}) and of the ex-post productivity/measurement error term (ε_{jt}). In Figure 5, I plot this measure of the underlying export-related heterogeneity, π_{jt}.

Not surprisingly, there is a lot of heterogeneity in π_{jt}. In addition, I find that π_{jt} is highly persistent over time, with an average correlation coefficient of 0.85. This persistence suggests that this underlying heterogeneity is not random noise but rather evidence of persistent differences across firms that affect exporting decisions. In particular, this suggests that

Figure 5: Export-Related Heterogeneity



differences in the geography of trade (through differences in destination and marginal costs of exporting) are important determinants of firm-level exporting decisions.

9 Conclusion

In this paper, I examine the empirical finding that exporting firms are more productive than non-exporting firms. I replicate this stylized fact on data for Colombian manufacturing firms using two commonly used measures of productivity from the trade literature. However, I show that these commonly used productivity measures are potentially biased due to both the use of value-added specifications of the production function and endogeneity caused by productivity and prices being unobserved and unaccounted for. I find evidence that the measures of productivity used in the literature over-estimate the productivity advantage of exporting firms. By extending a new strategy for estimating production functions, I am able to control for unobserved productivity and unobserved prices within a gross output setting.

There are two key findings in this paper. First, I find that once unbiased productivity estimates are obtained, exporting firms no longer appear more productive than non-exporting firms. In fact, the distributions of productivity across export status are very similar. This suggests that productivity is not the main determinant of exporting decisions. Consequently, some other form of firm-level heterogeneity is driving exporting decisions. Second, using data on export intensity, I show that heterogeneity associated with differences in the geographic barriers faced by firms can explain exporting decisions. While the first finding alone does not refute that fixed costs of exporting are a determinant of export status, it does show that they may not be as important as commonly believed. Furthermore, the second finding suggests that a model of exporting that ignores other sources of heterogeneity will miss key facts in the data. Together, these two results suggest that future research should emphasize the role that geography plays in determining patterns in firm-level exporting decisions.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2006): "Structural Identification of Production Functions," Unpublished Manuscript, UCLA Economics Department.

Aw, B., S. Chung, and M. Roberts (2000): "Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)," *World Bank Economic Review*, 14, 65–90.

Basu, S. and J. G. Fernald (1995): "Are Apparent Productive Spillovers a Figment of Specification Error?" *Journal of Monetary Economics*, 36, 165 – 188.

——— (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105, 249 – 283.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.

Bernard, A. B. and J. B. Jensen (1999): "Exceptional Exporter Performance: Cause, Effect or Both?" *Journal of International Economics*, 47, 1–25.

Bond, S. and M. Söderbom (2005): "Adjustment Costs and the Identification of Cobb-Douglas Production Functions," Unpublished Manuscript, The Institute for Fiscal Studies, Working Paper Series No. 05/4.

Bruno, M. (1978): "Duality, Intermediate Inputs, and Value-Added," in *Production Economics: A Dual Approach to Theory and Practice*, ed. by M. Fuss and M. D., Amsterdam: North-Holland, vol. 2.

Clerides, S. K., S. Lach, and J. R. Tybout (1998): "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *The Quarterly Journal of Economics*, 113, 903–947.

De Loecker, J. (2007): "Do Exports Generate Higher Productivity? Evidence from Slovenia," *Journal of International Economics*, 73, 69–98.

——— (2009): "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity," NYU Stern, Unpublished Manuscript.

Gandhi, A., S. Navarro, and D. Rivers (2009): "Identifying Production Functions Using Restrictions from Economic Theory," Unpublished manuscript, University of Wisconsin-Madison.

Griliches, Z. and J. Mairesse (1998): "Production Functions: The Search for Identification," in *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, New York: Cambridge University Press, 169–203.

Klette, T. J. and Z. Griliches (1996): "The Inconsistency of 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, 317–342.

Marschak, J. and W. Andrews (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205.

Melitz, M. (2003): "The Impact of Trade on Intra-Industry Reallocation and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Roberts, M. and J. Tybout (1997): "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87, 545–564.

Van Biesebroeck, J. (2006): "Exporting Raises Productivity in sub-Saharan African Manufacturing Firms," *Journal of International Economics*, 67, 373–391.

Wagner, J. (2007): "Exports and Productivity: A Survey of the Evidence from Firm-level Data," *World Economy*, 30, 60–82.

Table 1: Export Productivity Premia–Colombia

Measure \ Industry	311	312	313	314	321	322	323	324	331	332
Labor Productivity–VA	655%	118%	183%	151%	46%	43%	20%	27%	57%	11%
TFP–VA	368%	15%	114%	-17%	-3%	17%	-5%	24%	36%	25%

Measure \ Industry	341	342	351	352	353	354	355	356	361	362
Labor Productivity–VA	89%	96%	46%	88%	-37%	379%	80%	71%	300%	124%
TFP–VA	-3%	43%	20%	27%	1%	127%	-5%	14%	4%	6%

Measure \ Industry	369	371	372	381	382	383	384	385	390	Mean
Labor Productivity–VA	95%	196%	113%	85%	38%	88%	77%	73%	46%	116%
TFP–VA	20%	6%	49%	16%	10%	14%	21%	5%	20%	33%

Firms are classified by industry according to the ISIC Rev. 2 classification.

"Labor Productivity–VA" is based on estimates of labor productivity computed as real value-added per worker. "TFP–VA" is based on estimates of TFP computed as the residual from a regression of log real value-added on log capital and log labor.

Table 2: Export Productivity Premia–Colombia

Measure \ Industry	311	312	313	314	321	322	323	324	331	332
Labor Productivity–VA	655%	118%	183%	151%	46%	43%	20%	27%	57%	11%
TFP–VA	368%	15%	114%	-17%	-3%	17%	-5%	24%	36%	25%
TFP–GO	23%	0%	52%	0%	1%	3%	-4%	2%	14%	4%

Measure \ Industry	341	342	351	352	353	354	355	356	361	362
Labor Productivity–VA	89%	96%	46%	88%	-37%	379%	80%	71%	300%	124%
TFP–VA	-3%	43%	20%	27%	1%	127%	-5%	14%	4%	6%
TFP–GO	2%	9%	-3%	2%	10%	4%	3%	4%	1%	-1%

Measure \ Industry	369	371	372	381	382	383	384	385	390	Mean
Labor Productivity–VA	95%	196%	113%	85%	38%	88%	77%	73%	46%	116%
TFP–VA	20%	6%	49%	16%	10%	14%	21%	5%	20%	33%
TFP–GO	5%	1%	9%	4%	-1%	4%	8%	1%	2%	5%

Firms are classified by industry according to the ISIC Rev. 2 classification.

"Labor Productivity–VA" is based on estimates of labor productivity computed as real value-added per worker. "TFP–VA" is based on estimates of TFP computed as the residual from a regression of log real value-added on log capital and log labor. "TFP–GO" is based on estimates of TFP computed as the residual from a regression of log real gross output on log capital, log labor, and log intermediate inputs.

Table 3: Summary Statistics for Colombian Apparel Firms

	Mean	Standard Deviation
Gross Output	49.2	148.7
Value Added	21.3	63.4
Capital Stock	10.3	35.7
Number of Workers	82.3	163.4
Wages	10.7	26.3
Value of Intermediate Input	27.9	88.1
Percentage of Output Exported	9.8%	28.8%
Percentage of Output Exported (Exporters Only)	41.7%	37.8%

All figures are reported in thousands of 1981 pesos, with the exception of percentage of output exported and number of workers.

Table 4: Production Function Parameter Estimates–Apparel

	Mean L Elas.	Mean K Elas.	Mean M Elas.	r	ρ	η	Markup
Baseline Model	0.30	0.15	0.53	0.99	0.79	–	–
	(0.007)	*(0.006)*	*(0.001)*	*(0.001)*	*(0.005)*	–	–
Full Model	0.29	0.20	0.59	1.08	0.82	-10.40	1.11
	(0.08)	*(0.09)*	*(0.18)*	*(0.27)*	*(0.07)*	*(27.34)*	*(0.33)*

r : returns to scale

ρ : CES parameter (elasticity of substitution $= \frac{1}{1-\rho}$)

η : elasticity of demand

Standard errors are reported in parentheses below the point estimates.

Table 5: Export Productivity Premia–Apparel

Measure	Productivity Difference
Labor Productivity–Value-Added	43%
	(0.08)
TFP–Value-Added–OLS	17%
	(0.06)
TFP–Gross Output–Full Model	-1%
	(0.13)

"Labor Productivity–Value-Added" is based on estimates of labor productivity computed as real value-added per worker. "TFP–Value-Added–OLS" is based on estimates of TFP computed as the residual from a regression of log real value-added on log capital and log labor. "TFP–Gross Output–Full Model" is based on estimates of TFP computed as the residual from a regression of log real gross output on log capital, log labor, and log intermediate inputs.

Standard errors are reported in parentheses below the point estimates.

Appendix A: Optimal Allocation of Output

Conditional on a firm deciding to export, it must determine how to optimally allocate quantity to the domestic and foreign markets. Since goods sold on the domestic and foreign markets share the same production technology, their marginal costs are the same. Therefore, profit maximization implies that the firm wants to set the marginal revenue in the domestic market equal to the marginal revenue in the foreign market. Given the demands in equation (5), domestic marginal revenue can be derived as follows:

$$
\begin{aligned}
R_{jt}^{D} &= P_{jt}^{D} Q_{jt}^{D} \\
&= \overline{P}_t^{D} \left(\tfrac{1}{\overline{Q}_t^{D}}\right)^{\frac{1}{\eta}} \left(Q_{jt}^{D}\right)^{1+\frac{1}{\eta}} \\
\Rightarrow & \\
MR_{jt}^{D} &= \left(1+\tfrac{1}{\eta}\right) \overline{P}_t^{D} \left(\tfrac{1}{\overline{Q}_t^{D}}\right)^{\frac{1}{\eta}} \left(Q_{jt}^{D}\right)^{\frac{1}{\eta}} \\
&= \left(1+\tfrac{1}{\eta}\right) P_{jt}^{D}.
\end{aligned}
$$

Similarly, foreign marginal revenue is equal to:

$$MR_{jt}^{E} = \left(1+\frac{1}{\eta}\right) P_{jt}^{E}.$$

Consequently, setting the marginal revenues equal to each other implies setting the prices equal to each other.

$$
\begin{aligned}
MR_{jt}^{D} = MR_{jt}^{E} \ \Rightarrow \qquad P_{jt}^{D} &= P_{jt}^{E} \\
\Rightarrow \ \overline{P}_t^{D} \left(\tfrac{1}{\overline{Q}_t^{D}}\right)^{\frac{1}{\eta}} \left(Q_{jt}^{D}\right)^{\frac{1}{\eta}} &= \tau \overline{P}_t^{E} \left(\tfrac{1}{\overline{Q}_t^{E}}\right)^{\frac{1}{\eta}} \left(Q_{jt}^{E}\right)^{\frac{1}{\eta}}
\end{aligned}
$$

By replacing Q_{jt}^{D} with $\theta_{jt} \times Q_{jt}$ and Q_{jt}^{E} with $(1-\theta_{jt}) \times Q_{jt}$, I obtain the following expression:

$$\overline{P}_t^{D} \left(\tfrac{1}{\overline{Q}_t^{D}}\right)^{\frac{1}{\eta}} \left(\theta_{jt} \times Q_{jt}\right)^{\frac{1}{\eta}} = \tau \overline{P}_t^{E} \left(\tfrac{1}{\overline{Q}_t^{E}}\right)^{\frac{1}{\eta}} \left((1-\theta_{jt}) \times Q_{jt}\right)^{\frac{1}{\eta}}$$

Solving this expression for θ_{jt} yields the expression for the optimal fraction of quantity sold on the domestic market, θ_{jt}^{*}:

$$\theta_{jt}^{*} = \frac{\tau \frac{\overline{P}_t^{E\eta}}{\overline{Q}_t^{E}}}{\tau \frac{\overline{P}_t^{E\eta}}{\overline{Q}_t^{E}} + \frac{\overline{P}_t^{D\eta}}{\overline{Q}_t^{D}}},$$

which is just a function of the transportation costs and the price and quantity indices in both markets.