Product Differentiation, Multi-Product Firms and Structural Estimation of Productivity*

Jan De Loecker†
Economics Department, K.U. Leuven

Preliminary.

First version: November 2004
Current version: June 2005

Abstract

In this paper I suggest a methodology for getting reliable estimates on (total factor) productivity in an environment of product differentiation and multi-product firms. The model is built on Klette and Griliches (1996) and Olley and Pakes (1996) and I allow for multi-product firms. In addition to the large literature on estimating production functions correcting for the simultaneity bias, I control for pricing behavior and assume imperfect competition in the output market. I aggregate demand and production from product space into firm space, allowing me to use plant-level data to estimate productivity. In this way I get estimates for productivity corrected for demand shocks and as a by-product I get estimates for the elasticity of demand. The suggested methodology is quite flexible and allows for various demand systems. I show that measured productivity increases need not to reflect actual productivity increases, i.e. due to changing demand conditions and the product mix. I apply this methodology on the Belgian Textile Industry using a dataset where I have matched firm-level with product-level information. The resulting production coefficients and productivity estimates change considerably once taking into account the demand variation and the product mix.

Keywords: productivity, demand, multi-product firms

*I would like to thank Peter Davis, David Genesove, Joep Konings, Marc Melitz, Ariel Pakes, Frank Verboven, Patrick Van Cayseele and Hylke Vandenbussche for comments and suggestions. I also thank seminar participants at the ITD Seminar (KU Leuven), IIOC (Atlanta) and CEPR Applied IO Workshop (Munchen). I started this project while I was visiting the Economics Department of Harvard University during the academic year 2003-2004.

†Contact Author: jan.deloecker@econ.kuleuven.be

1 Introduction

In this paper I suggest a methodology for getting reliable estimates on productivity in an environment of imperfect competition in the product market, i.e. product differentiation, and I allow for multi-product firms. The model is related to the original work of Klette and Griliches (1996), where the bias from the use of an industry wide producer price index to deflate firm-level revenue - proxying for output - in estimating a production function is discussed. The focus is mainly on the estimated scale elasticities and there is no explicit discussion on the resulting productivity estimate. Furthermore, they rely on instrumental variables to correct for the simultaneity bias. The latter is a well documented problem when estimating a production function, i.e. the inputs are most likely to be correlated with the unobserved productivity shock and therefore biases the estimates.[1] The recent literature on estimating productivity has focussed mostly on this problem, ignoring or assuming away the omitted price variable bias.

Some recent work has discussed the potential bias of ignoring the demand side when estimating production functions. Katayama et al (2003) start out from a nested logit demand structure and verify the impact of the demand side on the productivity estimates. They need quite some assumptions in order to recover an estimate for productivity and they rely on Bayesian estimation techniques. Melitz and Levinsohn (2002) assume a representative consumer with Dixit-Stiglitz preferences and they feed this through the Levinsohn and Petrin (2003) estimation algorithm. Foster, Haltiwanger and Syverson (2003) discuss the relation between physical output, revenue and firm-level prices. They study this in the context of market selection and they state that productivity based upon physical quantities is negatively correlated with establishment-level prices while productivity based upon deflated revenue is positively correlated with establishment-level prices. The very few papers that explicitly analyze the demand side when estimating productivity or that come up with a strategy to do so, all point to the same direction: higher estimated productivity might just capture pricing power or higher mark-ups in some form.[2]

I combine the underlying structural framework of Olley and Pakes (1996) with a very basic demand side story as in Klette and Griliches (1996).[3] The latter suggests a simple way of identifying a demand parameter and treating the omitted price variable using an industry output variable. However, three main problems remain unchallenged in their

[1] The three main biases when estimating a production function are: i) the transmission or simultaneity bias, ii) the omitted price variable bias and iii) the selection bias.

[2] See Bartelsman and Doms (2000) for a comprehensive review on recent productivity studies using micro data. Concerning the topic of this paper I refer to page 592.

[3] The choice of the simple demand system is made in order to simplify the analysis and to be able to relate my results to the previous literature. Extensions can be made from here, i.e. a constant expenditure CES model can be used and the same results hold

method, which are largely recognized by the authors. Firstly, the industry output might proxy for other omitted variables relevant at the industry level such as industry wide productivity growth and factor utilization. The constant term and the residual in their model should take care of it since time dummies are no longer an option as they would take all the variation of the industry output. Secondly and more importantly, the residual still captures the unobserved productivity shock and biases the estimates on the freely chosen inputs (labor and material inputs). I proxy for this unobserved productivity shock using the method suggested by Olley and Pakes (1996) to overcome the simultaneity bias, i.e. by introducing a polynomial in investment and capital. The third problem is closely related to the solution of the simultaneity problem. Klette and Griliches (1996) end up with a negative capital coefficient partly due to estimating their production function in growth rates.[4] The authors believe that "*a more satisfactory solution to these questions requires an explicit dynamic model of investment behavior, incorporating uncertainty ... we suspect that changes in the capital services are picked up by movements in variable inputs*". The Olley and Pakes (1996) methodology comes straight out of a dynamic optimization problem under uncertainty as suggested in Ericson and Pakes (1995). In the latter the investment decision is an equilibrium outcome as well as the decision to remain active in the market. Essentially, the variation in the variable inputs is subtracted from the output when estimating the capital coefficient. The identifying assumption to estimate the capital coefficient is that capital is predetermined a period before (investment) the productivity shock is drawn, i.e. the news component in the productivity Markov process is uncorrelated with the capital stock.

In addition to addressing the problems discussed above, I allow for multi-product firms and I aggregate the demand from product space into firm-level demand. I present a straightforward estimation strategy and a resulting correction for the presence of multi-product firms and the degree of product differentiation. From this it follows that if one does not correct for these that a measured productivity increase need not to reflect an actual productivity increase.[5] Productivity is thus overestimated essentially due to demand shocks and price effects that are not filtered out. This sheds some light on the numerous studies linking productivity increases with firm characteristics like exporting, direct foreign investment, ... At a more aggregated level, productivity increases are the widely found effect after periods of trade liberalization, deregulation or other major changes in the operating environment.[6] The algorithm results in an estimate for productivity that

[4] I estimate a production function in levels and get sensible capital coefficients of around 0.10.

[5] Measured productivity can still reflect an actual productivity increase, however, to a lesser extent.

[6] See Pavcnik (2002) for the impact of trade liberalization on productivity in the manufacturing sector in Chile, Smarzynska (2004) for the effects of FDI in Lithuania, Van Biesebroeck (2005) for learning by exporting in Sub-Saharan African manufacturing, and Olley and Pakes (1996) for the deregulation of the US telecom industry.

is consistent with multi-product firms and product differentiation. In order for all this to be valid, one has to be willing to make some additional assumptions on the nature of demand and the production of the varieties. I apply this to the Belgian textile industry using product-level information matched to the firm-level data.

This paper is closely related to the Meltiz and Levinsohn (2002) setup. There are, however, some important differences. Firstly, in addition to the plant-level dataset I also have product-level information matched to the plants. This allows me to put more structure on the demand side. They proxy away the number of products per firm by substituting it by the number of firms in an industry. A side from a discussion on the methodology, I compare the differences in estimated productivity using an unique rich firm-level dataset. Secondly, on top of correcting for the omitted price variable I control for the simultaneity bias using the Olley and Pakes (1996) procedure. They follow the LP procedure to do this. This implies that they cannot control for the selection process of lower productivity firms exiting and do not model the dynamics of investment. In addition there has been some discussion on the validity of the LP estimator, i.e. the first stage potentially suffers from multicollinearity and the material input function might not take out all the variation correlated with the inputs (see Ackerberg, Caves and Frazer ; 2004). Finally and on a more conceptual note, by using materials to proxy for the productivity shock they assume that firms drawing a positive productivity shock will always demand more material inputs and thus produce more output and increase revenue. In this way, there is no room for keeping output constant and increase mark-ups, which results in higher revenue as well. This is due to the assumption of perfect competition on the output market where firms all face the same price and hence there is no scope for realizing mark-ups.[7]

This paper is organized as follows: firstly, the production framework and standard approach in estimating production functions is discussed. Secondly, the demand side is introduced starting out with single product firms and later on allowing for multi-product firms. The estimation strategy is discussed as well as the impact on estimated productivity. Thirdly, I present the data used in my application which has matched product-level information in addition to a rich firm-level dataset. The next section applies the suggested methodology on the Belgian textiles, where I compare my estimates with the recent state of the art method (OP) and the OLS benchmark. The last section concludes and discusses further research avenues.

[7]The LP methodology needs perfect competition on the output market in order for the intermediate input to be monotonic increasing in productivity to be able to proceed as in Olley and Pakes (1996). Whereas the latter has no assumption on the degree of competition on the output market.

2 The Production Framework

2.1 The Production Function

Let us start with the production side where a firm i produces (a product) according to the following production function

$$Q_{it} = L_{it}^{\alpha_l} M_{it}^{\alpha_m} K_{it}^{\alpha_k} \exp(\alpha_0 + \omega_{it} + u_{it}^q) \tag{1}$$

where Q stands for the quantity produced, L, M and K are the three inputs labor, materials and capital; and α_l, α_m and α_k are the coefficients, respectively. The parameter ε captures the economies of scale, i.e. $\varepsilon = (\alpha_l + \alpha_m + \alpha_k)$. Productivity is denoted by ω and u^q is an *i.i.d.* component. Below I aggregate over different products within a firm and I assume that the production structure is identical for every product j and therefore no cost synergies or spillovers are modelled here. I refer to Appendix B for a discussion on this.

2.2 The Standard Approach For Identifying the Production Function Parameters

The standard approach in structurally identifying the production function coefficients starts out with a production function as described in equation (1).[8] The physical output Q is then replaced by deflated revenue ($\widetilde{R}$) using an industry price deflator (P_I). Taking logs of equation (1) and relating it to the (log of) observed revenue per firm $r_{it} = q_{it} + p_{it}$, we get the following regression equation (2)

$$r_{it} = x_{it}\alpha + \omega_{it} + u_{it}^q + p_{it} \tag{2}$$

where $x_{it}\alpha = \alpha_0 + \alpha_l l_{it} + \alpha_m m_{it} + \alpha_k k_{it}$. The next step is to use the industry wide price index p_{It} and subtract it from both sides to take care of the unobserved firm-level price

$$\widetilde{r}_{it} = r_{it} - p_{It} = x_{it}\alpha + \omega_{it} + (p_{it} - p_{It}) + u_{it}^q \tag{3}$$

Most of the literature on structural estimation of productivity has worried about the correlation between the chosen inputs x and the unobserved productivity shock ω. The coefficient on the freely chosen variables labor and material inputs will be biased upwards, i.e. a positive productivity shock leads to higher labor and material usage ($E(x_{it}\omega_{it}) > 0$).

Even if this is corrected for, from equation (3) it is clear that if firms have some pricing power the estimates of α will be biased. As mentioned in Klette and Griliches

[8] Dynamic panel data econometrics uses lag structure and IV techniques to identify the production function parameters (Arellano and Bond; 1991).

(1996) inputs are likely to be correlated with the price a firm charges.[9] The error term $(u^q + p_{it} - p_{It})$ still captures any firm-level price deviation from the average (price index) price used to deflate the firm-level revenues. Essentially, any price variation (at the firm level) that is correlated with the inputs biases the coefficients of interest (α), i.e. $E(x_{it}(p_{it} - p_{It})) \neq 0$. The most straightforward correlation goes as follows: the level of inputs are positive correlated with the firm's output, which is negatively correlated with the price. Therefore firm-level inputs (materials and labor) are negatively correlated with the unobserved price and thus underestimates the coefficients on labor and materials. This is referred to as the omitted price variable bias. This bias works in the opposite direction as the simultaneity bias - the correlation between the unobserved productivity shock (ω) and the inputs (x) -, making any prior on the direction of the bias hard.[10]

The same kind of reasoning can be followed with respect to the input markets. However, my explicit treating of the imperfect output market takes - at least partly - care of this. For a more formal treatment I refer to Appendix C. The intuition goes as follows, if material prices are part of the unobservable, an increase in the price of materials will be passed through a higher output price if output markets are imperfect. Therefore controlling for the output price - as suggested in this paper - will also capture the shocks in price of materials. The only case where this reasoning might break down occurs when input markets are imperfect and output markets are perfect competitive, which is not a very likely setup.

Productivity remains an exogenous process and cannot be changed by the firm like e.g. investment (Ericson and Pakes; 1995) or other firm-level decision variables like e.g. export behavior (De Loecker; 2004). In the modelling approach this is translated by the assumption that productivity follows a Markov process.[11]

[9] The interpretation of the correlation is somewhat differerent here since my model is estimated in log levels and not in growth rates as in Klette and Griliches (1996).

[10] In more general terms the bias in the production function coefficients can be shown as follows. There are N firms and K inputs, R is a $Nx1$ matrix capturing firm-level deflated revenues, X is a NxK input matrix and Σ is a $Nx1$ error term, composed of a productivity shock W and noise term U The standard way in estimating a production function $Y = X\beta + \Sigma$, where β is the vector of coefficients, results in the following expression for the vector of coefficients: $plim(\beta) = \beta + plim(X'X)^{-1}X'W + plim(X'X)^{-1}X'U$. The two main biases in estimating β discussed in this paper are given by the last two terms, the simultaneity and omitted price variable bias, respectively.

[11] Muendler (2004) allows productivity to change endogenously and suggest a way to estimate it. Buettner (2004) introduces R&D and models the impact of this controlled process on unobserved productivity. Ackerberg and Pakes (2005) discuss more general extensions to the exogenous Markov assumption of the unobserved productivity shock.

3 Introducing Demand: Product Differentiation and Multi-Product Firms

In this section I introduce the demand system that firms face in the output market. The demand system is based on the standard horizontal product differentiation model, but also captures an unobserved quality component. The demand system needed to be able to identify the parameters of interest is somewhat limited due to missing demand data, i.e. prices and quantities. So one has to be willing to put somewhat more structure on the nature of demand. However, the modeling approach here does not restrict any demand system per se, as long as the inverse demand system can generate a (log-) linear relationship of prices in quantities. What follows then also holds for other demand systems and later on I will turn to some alternative specifications. The inverse demand system is then used to substitute for the unobserved price variable in the revenue generating production function as discussed above. I start out with single product firms and show how this leads to my augmented production function. In a second step I allow for firms to produce multiple products. The focus is on the resulting productivity estimates and in the case of multi-product firms these can be interpreted as average productivity across a firm's products.

3.1 A Simple Demand Structure: Single Product Firms

I follow Klette and Griliches (1996) and later on I extend it by allowing firms to produce multiple products.[12] I start out wit a simple demand system where each firm i produces a single product and faces the following demand (4)

$$Q_{it} = Q_{It} \left(\frac{P_{it}}{P_{It}} \right)^{\eta} \exp(u_{it}^{d}) \tag{4}$$

where Q_{It} is the industry output at time t, (P_{it}/P_{It}) the relative price of firm i with respect to the average price in the industry, u^d is an idiosyncratic shock specific to firm i and η is the substitution elasticity between the differentiated products in the industry. In the empirical analysis I follow Klette and Griliches (1996) and replace the industry output Q_{It} by a weighted average of the deflated revenues, i.e. $q_{It} = (\sum_i ms_{it} r_{it} - p_{It})$ where the weights are the market shares. This comes from the observation that a price index is essentially a weighted average of firm-level prices where weights are market shares (see Appendix A.2).

Later on I allow for the error term to be more structured and decompose the shock into a quality component (ξ_{it}) and an idiosyncratic part (u_{it}^d). The firms are assumed to operate

[12] Or at least I am explicit about it when modelling the demand and supply side.

in an industry characterized by horizontal product differentiation, where η captures the substitution elasticity among the different products and η is finite. As mentioned in Klette and Griliches (1996) similar demand systems have been used extensively under the label of Dixit-Stiglitz demand. The key feature is that monopolistic competition leads to price elasticities which are constant and independent of the number of varieties (see Berry (1994) for a discussion on demand in industries with product differentiation). It is clear that the demand system is quite restrictive and implies one single elasticity of substitution for all products within a product range and hence no differences in cross price elasticities. In the empirical application the elasticity of substitution is allowed to differ among product segments. However, in most productivity studies - up till now - the working assumption is that of perfect competition on the output market, which is clearly hard to defend. The motivation for modeling the demand explicitly here is to include the basic primitive of the market that provide the incentives for firms active in the market in addition to the other primitive - productivity - which is of interest here.

Taking logs of equation (4) and writing the price as a function of the other variables results in the following expression where $x = \ln X$

$$p_{it} = \frac{1}{\eta}(q_{it} - q_{It} - u_{it}^{d}) + p_{It} \tag{5}$$

As discussed extensively in Klette and Girliches (1996) and Melitz and Levinsohn (2003), the typical firm-level dataset has no information on physical output per firm and prices.[13] Commonly, we only observe revenue and we deflate this using an industry-wide deflator. As discussed above, deflating the revenue by an industry-wide price deflator is only valid if firms have no price setting power and all face the same price. The observed revenue r_{it} is then substituted for the true output q_{it} when estimating the production function. Ignoring the price thus leads to an omitted variable bias since it is not unlikely that a firm's price is correlated with the used inputs. I now substitute the price p_{it} by the expression in (5) in equation (2) and from now I pursue with deflated revenue ($\widetilde{r}_{it} = r_{it} - p_{It}$)

$$\widetilde{r}_{it} = r_{it} - p_{It} = \frac{\eta+1}{\eta} q_{it} - \frac{1}{\eta} q_{It} - \frac{1}{\eta} u_{it}^{d} \tag{6}$$

Now I only have to plug in the production technology as expressed in equation (1) and I have a revenue generating production function with both demand and supply variables and parameters.

$$\widetilde{r}_{it} = \left(\frac{\eta+1}{\eta}\right)(\alpha_0 + \alpha_l l_{it} + \alpha_m m_{it} + \alpha_k k_{it}) - \frac{1}{\eta} q_{It} + \left(\frac{\eta+1}{\eta}\right)(\omega_{it} + u_{it}^{q}) - \frac{1}{\eta} u_{it}^{d}$$

[13] Exceptions are Dunne and Roberts (1992), Jaumandreu and Mairesse (2004) and Eslava, Haltiwanger, Kugler and Kugler (2004) where plant-level prices are observed and thus demand and productivity shocks are estimated separately. To my knowledge this is a very rare setup

It is clear that if one does not take into account the degree of imperfect competition on the demand side of the market, i.e. some pricing behavior, that the analysis will be plagued by an omitted price variable bias. On top of this the estimated coefficients are estimates of a reduced form combining the demand and supply side in one equation. Note that I will also treat the transmission/simultaneity bias explicitly using the Olley and Pakes (1996) methodology.

As mentioned above I further decompose the unobservable u^d in equation (4) into a unobserved quality (ξ) and an *i.i.d.* component. This leads to my general estimating equation of the production function

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_\eta q_{It} + (\omega^*_{it} + \xi^*_{it}) + u_{it} \tag{7}$$

where $\beta_h = ((\eta+1)/\eta)\alpha_h$ with $h = l, m, k$, $\beta_\eta = -\eta^{-1}$, $\omega^*_{it} = ((\eta+1)/\eta)\omega_{it}$, $\xi^*_{it} = -\eta^{-1}\xi_{it}$ and $u_{it} = ((\eta+1)/\eta)u^q_{it} - \frac{1}{\eta}u^d_{it}$. When estimating this equation (7) I recover the production function coefficients ($\alpha_l, \alpha_m, \alpha_k$) and returns to scale parameter (ε) controlling for the omitted price variable and the simultaneity bias, as well as an estimate for the elasticity of substitution η. When correcting for the simultaneity bias I follow the Olley and Pakes (1996) procedure and replace the productivity shock ω by a function in capital and investment.

In my empirical analysis I will estimate various versions of the production function described in equation (7). Given the unique dataset where I have product information linked to every firm[14], allowing me to put in some more demand variables (essentially proxying for ξ). Adding those extra (mostly dummy variables) is not going to change the estimated production function coefficients drastically but it will turn out to improve the estimated demand parameter η.

3.2 Multi-Product Firms

Now I explicit model the fact that firms produce several products, i.e. they are all differentiated products and η still captures the substitution elasticity among the varieties. Note that I do not allow for different substitution patterns among products owned by one firm as opposed to the substitution between products owned by different firms. I have no prior on whether consumers are more likely to substitute from a given product to products owned by the same firm versus products owned by other firms. The modelling

[14] A downside is that the product-level information (number of products produced, segments and which products) is time invariant and leaves me with a balanced panel. However, I will run the same regression on an unbalanced dataset where I control for the selection bias as well as suggested in Olley and Pakes (1996). When I do this the results turn out to be very similar as expected since the correction for the omitted price variable is essentially done in the first stage of the estimation algorithm. The variation left in capital is not likely to be correlated with the demand variables and therefore I only find slightly different estimates on the capital coefficient as expected.

approach here - however - does allow for more realistic substitution patterns (like in the spirit of Berry, Levinsohn and Pakes; 1995) among the various products produced.[15]

The demand system is identical to the one expressed in equation (4), only a product subscript j is added. Note that the relevant space is now the product space M. Expressed in terms of the empirical application, I have 308 (N) firm observations and 2,990 firm-product (M) observations, with 563 unique product categories. In the single product case the demand system is relevant for every firm i, whereas in the multiple product case the demand is with respect to product j of firm i.

$$Q_{ijt} = Q_{It}^{p} \left(\frac{P_{ijt}}{P_{It}^{p}} \right)^{\eta_p} \exp(u_{ijt}^{d} + \xi_{ijt}) \tag{8}$$

There are N firms and M products in the industry with each firm producing M_i products, where $M = \sum_i M_i$. The demand for product j of firm i is given by Q_{ij}, Q_I^p is the demand shifter relevant at the product-level, P_I^p is the industry price index relevant at the product level, η_p is the demand elasticity relevant at the product space, ξ_{ijt} is unobserved product quality and u_{ijt}^d is product j specific idiosyncratic shock.[16] In the empirical analysis I only observe the relevant variables at the firm level and I implicitly observe the aggregation of the product level variables. This is the case in any study using firm-level data to estimate a production function. However, as I will discuss later on in detail, I have information on the product market linked to the firm-level data. This allows me to put somewhat more structure on the way the product-level demand and production are aggregated.

Proceeding as outlined above, the revenue of product j of firm i is $r_{ijt} = p_{ijt} + q_{ijt}$ and using the demand system as expressed in equation (8) I get the following expression for the product-firm revenue r_{ijt}

$$r_{ijt} = \left(\frac{\eta_p + 1}{\eta_p} \right) q_{ijt} - \frac{1}{\eta_p} q_{It}^{p} + p_{It}^{p} - \frac{1}{\eta_p} \xi_{ijt} - \frac{1}{\eta_p} u_{it}^{d} \tag{9}$$

As mentioned above I now assume that the production function q_{ij} for every firm i for all its products M_i is given by $q_i = x_i \alpha + \omega_i + u_i^q$ where $x_i = \sum_j^{M_i} x_{ij}$.[17] This implies that the production technology for every product is the same across firms and products, i.e. no cost synergies are allowed on the production side. As before I substitute in the

[15]Note that a large part of the productivity studies are focused on the manufacturing sector making discrete choice models implying a demand of one unit of a good somewhat less appealing.

[16]I make the distinction between the product space and the firm space. In the multi-product model I have to aggregate the revenues per product to the firm's total revenue. The demand shifters are thus depending on the space, therefore I use the supperscript p for the output and price index In the empirical analysis - as in the single product case - I replace the output by the weighted average of deflated revenues. In a more general setup one can allow the demand elasticity to be product or time specific. In the empirical analysis one has to interact all the relevant terms with time dummies.

[17]For ease of notation I write the production function as $x_{ijt}\alpha = \alpha_l l_{ijt} + \alpha_m m_{ijt} + \alpha_k k_{ijt}$. In Appendix B I relax this assumption and allow for synergies among various products.

production technology as given by equation (1) where now a product subscript j is added and I get the following expression

$$r_{ijt} - p_{It}^p = \left(\frac{\eta_p + 1}{\eta_p}\right) x_{ijt}\alpha - \frac{1}{\eta_p} q_{It}^p + \left(\frac{\eta_p + 1}{\eta_p}\right)(\omega_{ijt} + u_{ijt}^q) - \frac{1}{\eta_p}\xi_{ijt} - \frac{1}{\eta_p} u_{ijt}^d \quad (10)$$

For now I assume a constant demand elasticity across products (η). I aggregate the product-firm revenue to the firm revenue by taking the sum over the number of products produced M_i, i.e. $r_{it} = \sum_j^{M_i} r_{ijt}$

$$\widetilde{r}_{it} = r_{it} - p_{It} \quad (11)$$

$$= \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_\eta q_{It} + \sum_{j=1}^{M_i} \left(\left(\frac{\eta + 1}{\eta}\right) \omega_{ijt} - \frac{1}{\eta}\xi_{ijt} \right) + u_{it} \quad (12)$$

where $u_{it} = \sum_j \left(\left(\frac{\eta+1}{\eta}\right) u_{ijt}^q - u_{ijt}^d \right)$. Furthermore p_{It} and q_{It} are the relevant (average) price index and industry output at the industry level, where the latter just captures a demand shifter for the industry as a whole as opposed to a product-level demand shifter.

I rewrite the two unobservables - productivity and quality - in such a way that I can relate it to the information I have in my product data. The terms $\sum_{j=1}^{M_i} \omega_{ijt}$ and $\sum_{j=1}^{M_i} \xi_{ijt}$ capture the sum of all the firm-product productivity shocks and quality shocks, respectively. I rewrite the sum of the productivity (quality) shocks as the product of the firm's average - across the M_i products - productivity (quality) shock $\overline{\omega}_{it}$ ($\overline{\xi}_{it}$) and the number of products (M_i)

$$\sum_{j=1}^{M_i} \omega_{ijt} = M_i \overline{\omega}_{it}$$

$$\sum_{j=1}^{M_i} \xi_{ijt} = M_i \overline{\xi}_{it}$$

where I assume that the number of products per firm are constant over time.[18] By replacing the sum of the productivity (quality) shocks it is clear that for single product firms $M_i \overline{\omega}_{it} = \omega_{it}$. The interpretation of the productivity term for multi-product firms is rewritten here such that it refers to a firm's average productivity across its products. This results in the following equation

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_\eta q_{It} + \left(\frac{\eta + 1}{\eta}\right) M_i \overline{\omega}_{it} + \frac{1}{|\eta|} M_i \overline{\xi}_{it} + u_{it} \quad (13)$$

$$= \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_\eta q_{It} + (\omega_{it}^* + \xi_{it}^*) + u_{it} \quad (14)$$

[18] Since I only have cross sectional information on the number of products produced by firm. In recent work of Bernard, Redding and Schott (2003) a model is presented where firms endogenously choose the product mix and hence their number of products.

where $\omega^*_{it} = \left(\frac{\eta+1}{\eta}\right) M_i \overline{\omega}_{it}$ and $\xi^*_{it} = \frac{1}{|\eta|} M_i \overline{\xi}_{it}$. The last step clearly shows the various components potentially attributed to productivity when not correcting for the degree of product differentiation.

4 Estimating Strategy and Correcting Productivity

I now briefly discuss how to estimate the demand and production function parameters. Secondly, I discuss the resulting productivity estimate and how it should be corrected for in the presence of product differentiation and multi-product firms. Finally I also provide a discussion on the potential miss-measured productivity growth using the standard identification methods.

4.1 Estimation Strategy: Single and Multi-Product Firms

Estimating the regression in (14) is similar to the Olley and Pakes (1996) correction for simultaneity, only now an extra term has to be identified. I have grouped the two 'structural unobservables' productivity and quality into 'productivity'. Introducing the demand side explicitly clearly shows that any estimation of productivity also captures firm/product specific unobservables like e.g. quality.

I assume that the quality and productivity component follow the same stochastic process, i.e. a first order Markov process.[19] Both are known to the firm when making its decision on the level of inputs. The new unobserved state variable in the Olley and Pakes (1996) framework is now $\widetilde{\omega}_{it} = (\omega_{it} + \xi_{it})$ and this is equivalent to Meltiz's (2001) representation. Technically, the equilibrium investment function still has to be a monotonic function with respect to the productivity shock, $\widetilde{\omega}$, in order to allow for the inversion as suggested in equation (15)

$$i_t = i_t(k_t, \widetilde{\omega}_t) \Leftrightarrow \widetilde{\omega}_t = h_t(k_t, i_t) \tag{15}$$

Here I have been more explicit on the nature of this shock (both quality and productivity related); however, it does not change the impact on investment. A firm draws a shock consisting of both productivity and quality and the exact source of the shock is not important as a firm is indifferent to selling more given its inputs due to an increased productivity or increased quality perception of its product(s). We could even interpret

[19] A possible extention to this is to assume that the quality and the productivity shock follow a different Markov process. Therefore one can no longer collapse both variables into one state variable (see Petropoulus for explicit modelling of this). For now I assume a scalar unobservable (productivity and quallity taken together due to the assumption that both following the same Markov process) that follows a first order Markov process. However, I can allow for higher order Markov processes and relax the scalar unobservable assumption as suggested in Ackerberg and Pakes (2005), see later on.

investment in a broader sense, both investment in capital stock and in advertising. I replace the productivity $\widetilde{\omega}_{it}$ component by a polynomial in capital and investment, recovering the estimate on capital in a second stage using non linear least squares. The demand parameters, labor and material are all estimated in a first stage (16)

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_\eta q_{It} + \phi_t(k_{it}, i_{it}) + u_{it} \tag{16}$$

under the identifying assumption that the function in capital and investment proxies for the unobserved product/quality shock. Note that the $\phi(.)$ is a solution to a complicated dynamic programming problem and depends on all the primitives of the model like demand functions, the specification of sunk costs, form of conduct in the industry and others (Ackerberg, Benkard, Berry and Pakes; 2005). My methodology brings one of these primitives - demand - explicitly into the analysis in - however - a crude way and essentially improves the proxy of this complicated function by introducing explicit demand variables in the first stage.[20] The identification of the capital coefficient in a second stage will also improve due to the estimate for $\phi(.)$ out of the first stage.

In a second stage (17) the variation in the variable inputs and the demand variation is subtracted from the deflated revenue to identify the capital coefficient. As in Olley and Pakes (1996) the news component in the productivity/quality process is assumed to be uncorrelated with capital in the same period since capital is predetermined by investments in the previous year.

$$\widetilde{r}_{it+1} - b_l l_{it+1} - b_m m_{it+1} - b_\eta q_{It+1} = c + \beta_k k_{it+1} + \sum_{c=1}^{5} (\widehat{\phi}_t - \beta_k k_{it})^c + e_{it+1} \tag{17}$$

where b is the estimate for β out of the first stage.[21] Note that here I need to assume that quality and productivity follow the exact same Markov process in order to identify the capital coefficient. If the quality term does not follow the same process and is depending on productivity, identification is only possible through an explicit demand estimation as e.g. Berry, Levinsohn and Pakes (1995) in order to produce an estimate for ξ. Another way is to assume that the quality shock is uncorrelated with capital and is has no lag structure, but that would leave us back in the case where quality is essentially ignored when estimating a revenue generating production function.

It is clear that due to setup of my model, that the number of products per firms only come into play when productivity is recovered out of the regression. Productivity (tfp) is then recovered as the residual by replacing the true values by the estimated coefficients, $\widetilde{r}_{it} - b_0 - b_l l_{it} - b_m m_{it} - b_k k_{it} - b_\eta q_{It} = \widehat{tfp}_{it}$.

[20]Here, the model does not explicitly capture the link between the level of investment and future demand expectations. However, it is all captured in the policy function $i = i(k, \omega)$.

[21]Block-bootstrapping (on a firm's year observations) is used to recover the standard error on the estimated capital coefficient.

For now I do not worry about the selection bias since I will use a balanced panel in the first set of regressions. The correction for the selection bias is just as in Olley and Pakes (1996) where a dummy variable denoting survival is regressed on a polynomial in capital and investment. Where now the interpretation on the lower productivity threshold is somewhat different, i.e. it captures both productivity and quality.

4.2 Correcting Productivity

In order to compare with the standard regression with single product firms, it is clear that when estimating equation (18) that the resulting productivity estimate (residual) has to be corrected for i) the number of products a firm produces and ii) the demand parameter, on top of the potentially differently estimated coefficients β_l, β_m, β_k and β_0

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + tfp_{it} + u_{it} \tag{18}$$

For now I assume away the unobserved quality component and focus on the unobserved productivity shock. The resulting productivity tfp relates to the true unobserved productivity $\overline{\omega}_{it}$ (averaged across a firm's products) in the following way

$$\overline{\omega}_{it} = \frac{tfp_{it} - \beta_\eta q_{It}}{M_i}\left(\frac{\eta}{\eta+1}\right) \tag{19}$$

The estimated productivity shock consistent with the product differentiated demand system and multi-product firms is obtained by substituting in the estimates for the true values (β_η and η). This shows that any estimation of productivity - including the recent literature correcting for the simultaneity bias (Olley and Pakes; 1996 and Levinsohn and Petrin; 2003) is biased in the presence of imperfect output markets and multi-product firms. Assuming an underlying product market a simple correction is suggested, i.e. subtract the demand variation, divide the estimated productivity by the number of products and correct for the degree of product differentiation. One can even get the demand parameter out of a separate (and potentially more realistic) demand regression. Note that in the case of single product firms operating in a perfect competitive market the estimated productivity corresponds with the true unobservable (given that the simultaneity and selection bias are addressed as well). Note that the coefficients of the production function change due to the correction for the product differentiation as well in terms of interpretation.

It is clear from equation (19) that the degree of product differentiation (measured by η) only re-scales the productivity estimate. However, when the demand parameter is allowed to vary across product segments, the impact on productivity is not unambiguous. The number of products per firm M_i does change the cross sectional (across firms) variation in productivity and changes the ranking of firms and consequently the impact of changes in the operating environment or firm-level variables on productivity (e.g. trade liberalization/ protection). In more general terms I should include the quality component as well. However, in the empirical analysis I try to get a proxy for this by using various product dummies (see section 5.2).

In a more general framework of time varying number of products per firm (M_{it}) and time and product-space varying demand elasticities (η_{pt}) the bias in measured productivity tfp is given by (20). The traditional measure tfp captures various effects in addition to the actual productivity shock. Equation (20) and (21) show these various components under constant and product varying demand elasticities, respectively.

$$tfp_{it} = \beta_{\eta_t} q_{It} + \left(\frac{\eta_t + 1}{\eta_t} \right) M_{it} \overline{\omega}_{it} + \frac{1}{|\eta_t|} M_{it} \overline{\xi}_{it} \tag{20}$$

$$tfp_{it} = \beta_{\overline{\eta}_{it}} q_{It} + \left(\frac{\overline{\eta}_{it} + 1}{\overline{\eta}_{it}} \right) M_{it} \overline{\omega}_{it} + \frac{1}{|\overline{\eta}_{it}|} M_{it} \overline{\xi}_{it} \tag{21}$$

where in equation (21) a firm specific demand elasticity $\overline{\eta}_i$ captures the average demand elasticity across firm's i products M_i.[22] Measured productivity consists of a pure demand specific term ($\beta_{\eta} q_{It}$) in addition to productivity and quality; and thus demand shocks need to be filtered out from the measured productivity term.

This expression sheds somewhat more light on the discussion whether various competition and trade policies have had an impact on productive efficiency. E.g. Pavcnik (2002) showed that tariff liberalization in Chile led to higher productivity, where essentially time dummies were used to identify the trade liberalization. In terms of my framework, these time dummies might also capture changes in elasticity of demand and/or the product mix of firms. Similar studies have essentially measured productivity in some form as expressed in equation (20). The increased (measured) productivity can be driven by four factors: i) increased average product quality, ii) increased average productivity, iii) more elastic demand and iv) increased number of products. Bernard, Redding and Schott (2003) suggest that an important margin along which firms may adjust to increased globalization and other changes in the competitive structure of markets is through product choice and/or changes in the nature of the production process.

Measuring increased productivity without taking into account the demand side of the output market and the degree of multi-product firms might thus have nothing to

[22] Where $\overline{\eta}_i = \sum_j^{M_i} z_{ij} \eta_p$ with $z = \frac{1}{M_i}$. In the empirical application I consider a weighted average across the various segments a firm is active in. The weights (z) used are the number of products a firm has in a given segment over its total number of products, capturing the importance of each segment in the firm's total demand. The subscript i on the demand elasticity is thus not reflecting a firm specific elasticity, however, it merely reflects that firms operate in different segments and therefore face a different total elasticity.

do with an actual productivity increase.[23] To see this, consider the measured firm-level productivity change over time $\Delta tfp_{it} = (tfp_{it} - tfp_{it-1})$ in equation (22), where for now I abstract from the unobserved quality component.

$$\begin{aligned} \Delta tfp_{it} &= \left(\beta_{\eta_{jt}} q_{It} - \beta_{\eta_{jt-1}} q_{It-1}\right) + \left(\gamma_{it} M_{it} \overline{\omega}_{it} - \gamma_{it-1} M_{it-1} \overline{\omega}_{it-1}\right) && (22) \\ &= \left(\gamma_{it} M_{it} \overline{\omega}_{it} - \gamma_{it-1} M_{it-1} \overline{\omega}_{it-1}\right) && (23) \end{aligned}$$

where $\gamma_{it} = (\overline{\eta}_{it} + 1)/\overline{\eta}_{it}$. In the last step I assume that the basic Klette and Griliches (1996) approach is followed and the residual is reduced to the last term.[24]

Say that firm i experienced no productivity gain at all ($\overline{\omega}_{it} = \overline{\omega}_{it-1}$) and for expositional reasons assume a constant demand elasticity across products. Still we can measure a productivity increase and potentially attributing this to trade or competition policy measures. I distinguish three cases to discuss as given in Table 1: i) constant number of products ($M_{it} = M_{it-1}$), ii) a constant demand elasticity ($\gamma_t = \gamma_{t-1}$) and iii) I allow both the demand elasticity and number of products to change over time.

Table 1: Increased Measured Productivity Without True Productivity Increase

Case	Δtfp_{it}	$\Delta tfp_{it} > 0$ if ...
i) constant number of products	$\Delta\gamma_t M_{it-1}\overline{\omega}_{it-1}$	Demand gets more elastic ($\Delta\gamma_t > 0$)
ii) constant demand elasticity	$\gamma_{t-1}\Delta M_{it}\overline{\omega}_{it-1}$	More products produced ($\Delta M_{it} > 0$)
iii) M_{it} and γ_t change	$(\gamma_t M_{it} - \gamma_{t-1} M_{it-1})\overline{\omega}_{it-1}$	Demand gets more elastic ($\Delta\gamma_t > 0$) More products produced ($\Delta M_{it} > 0$)

.

As indicated in the table above, measured productivity can increase under three different scenario's without having anything to do with an actual increase in productivity. In order for the increased measured productivity to reflect the actual complete productivity increase, the number of products produced and the demand elasticity have to remain constant over time, or one has to control for them.

The numerous of studies that have analyzed the impact of trade liberalization are thus only valid under these assumptions. As discussed in Bernard et al (2003) it is unlikely that the product choice (number and mix) does not change when industries are hit by big

[23] Harrison (1994) builds on the Hall (1988) methodology to verify the impact of trade reform on productivity and concludes that "... *ignoring the impact of trade liberalization on competition leads to biased estimates in the relationship between trade reform and productivity growth*".

[24] The conditions for increased measured productivity without an actual productivity increase when the KG (1996) industry output variable (essentially a demand shifter) is omitted, are somewhat more complicated.

changes in the operating environment like increased globalization. The same is true for the demand elasticity, simple economic reasoning would expect demand to become more elastic as consumers have more products to choose from.

As mentioned above, the cross sectional variation in productivity is miss-measured in the presence of multi-product firms, i.e. it clearly changes the distribution of productivity.[25] An established correlation of measured productivity with other firm specific variables (like export status, ownership structure, trade protection,...) might thus be misleading as well. I will turn to the different biases in my empirical analysis.

Finally, I can express the relative importance of the bias by decomposing the firm-level change in productivity (holding the number of products constant) as follows: a demand related term, the pure productivity change and a covariance term.

$$\Delta tfp_{it} = M_i \left(\underbrace{\Delta\gamma_t \overline{\omega}_{it-1}}_{demand} + \underbrace{\gamma_{t-1}\Delta\overline{\omega}_{it}}_{productivity} + \underbrace{\Delta\gamma_t \Delta\overline{\omega}_{it}}_{covariance} \right) \tag{24}$$

From here one can construct an industry productivity index and verify the different components controlling for the market shares used to construct the index.

4.3 Unobserved Quality and Productivity

So far I have assumed that the unobservable $\widetilde{\omega}$ - including both productivity and quality - can be proxied by a non parametric function in investment and capital. The underlying assumption here is that investment proxies both the shocks in productivity (ω) and product quality (ξ). I now relax this by allowing investment to depend on another unobservable - a demand shock - that varies across firms as suggested in Ackerberg and Pakes (2005). This notion also follows from the discussion throughout the paper that both demand and production related shocks have an impact on observed revenue. Note that quality itself would not enter the production function if we would observe physical output or firm-level prices in the case where quality does not enter the investment policy function. However, when investment is allowed to depend on an unobserved demand shock as well, it enters through the productivity shock even when physical output or firm-level prices are observed. The case discussed here has a demand shock entering both through the productivity shock and through the use of revenue to proxy for output at the firm level. If the unobserved quality shock does not enter the policy function I can just control for it in the first stage of the regression, filtering it out of the non parametric function $\phi(.)$. The latter approach is also discussed in the empirical analysis using firm-product dummies.

[25]Potentially also due to product specific demand elasticities (η_p) if one has the data to identify these (see section 5.2).

Formally, I relax the assumption that investment only depends on the capital stock and the unobserved productivity shock. I now have two unobservables (ω, ξ) and the investment function is now $i_{it} = i_t(k_{it}, \omega_{it}, \xi_{it})$. The demand unobservable ξ is assumed to follow a Markov process that is independent of the productivity process, otherwise I can no longer identify the capital coefficient in a second stage. We now need a second control s_{it} - say advertizement expenditures - to proxy the unobservable in order to control for the productivity shock. Let the bivariate policy function determining (i_{it}, s_{it}) be $\Upsilon(.)$ and assume it is a bijection in (ω_{it}, ξ_{it}) conditional on the capital stock k_{it} (25).

$$\begin{pmatrix} i_{it} \\ s_{it} \end{pmatrix} = \Upsilon_t(k_{it}, \omega_{it}, \xi_{it}) \tag{25}$$

This allows us to invert and rewrite the unobservable productivity as a function of the controls in the following way (26)

$$\widetilde{\omega}_{it} = \Upsilon_t^{-1}(k_{it}, i_{it}, s_{it}) \tag{26}$$

The revenue generating production is as before and the first stage of the estimation algorithm now looks as follows

$$\begin{aligned} \widetilde{r}_{it} &= \beta_0 + \beta_l l_{it} + \beta_m\, m_{it} + \beta_k k_{it} + \beta_\eta q_{It} + \Upsilon_t^{-1}(k_{it}, i_{it}, s_{it}.) + u_{it} && (27) \\ &= \beta_0 + \beta_l l_{it} + \beta_m\, m_{it} + \beta_\eta q_{It} + \widetilde{\phi}_t(k_{it}, i_{it}, s_{it}.) + u_{it} && (28) \end{aligned}$$

where $\widetilde{\phi}_t = \beta_k k_{it} + \Upsilon_t^{-1}(k_{it}, i_{it}, s_{it}.)$. The non parametric function is in three variables, i.e. investment, capital and advertizement expenditures, where the latter controls for the unobserved demand shocks ξ. In addition to the standard Olley and Pakes (1996) methodology I control for both observed and unobserved demand shocks coming from the use of revenue in stead of physical output and from the notion that demand shocks might have an impact on the level of investments.

The second stage hardly changes since the process of the demand shock is assumed to be independent of the productivity shock. Consider the revenue generating production function at time $t+1$

$$\widetilde{r}_{it+1} = \beta_0 + \beta_l l_{it+1} + \beta_m\, m_{it+1} + \beta_k k_{it+1} + \beta_\eta q_{It+1} + E(\widetilde{\omega}_{it+1}|I_t) + v_{it+1} + u_{it+1}$$

where I have used the fact that productivity and the demand shock follow a first-order Markov process, i.e. $\widetilde{\omega}_{it+1} = E(\widetilde{\omega}_{it+1}|\widetilde{\omega}_{it}) + v_{it+1}$, where v is the news term.

$$\begin{aligned} \widetilde{r}_{it+1} - b_l l_{it+1} - b_m\, m_{it+1} - b_\eta q_{It+1} &= \beta_0 + \beta_k k_{it+1} + \widetilde{g}(\widetilde{\omega}_{it}) + v_{it+1} + u_{it+1} \\ &= \beta_0 + \beta_k k_{it+1} + \widetilde{g}(\widetilde{\phi}_{it} - \beta_k k_{it}) + e_{it+1} \end{aligned}$$

where $e_{it+1} = \upsilon_{it+1} + u_{it+1}$. The only difference is that the estimate for $\widetilde{\phi}(.)$ is different to the standard case (16) and leads to more precise estimates for the capital stock. Variation in output (purified from variation in variable inputs and observed demand shock) that is is correlated with the control s is no longer potentially contributed to the variation in capital. The extension discussed here is intuitively closely related to the discussion in section 4.1, where demand variation is introduced in the first stage due to the observation of revenue in stead of physical output. Here the extra control s_{it} comes from the notion that investment might be depending on demand shocks. In order to allow quality to be independent from the productivity shock, i.e. evolve differently over time, in the final stage I would be left with a non parametric function $\widetilde{g}(.)$, consisting of a term depending of investment making identification impossible.[26] In fact the only way out is to assume either that this quality unobservable is uncorrelated with capital and ends up in the error term e. I have assumed that both productivity and quality follow the same Markov process, allowing me to collapse them into one unobservable $\widetilde{\omega}$.[27] This is exactly what I have assumed about the quality component in previous sections. However, now I include variables proxying for the quality unobservable (like advertizement expenditures, product dummies) which take out additional variation related to the demand side (ξ), leading to different estimates for ϕ in the NLLS estimation. When estimating the capital coefficient the identifying assumption is that the demand shocks are independent of the productivity shocks.

5 The Belgian Textiles: Data

I now turn to the dataset I use to apply the methodology suggested above. My data capture the Belgian textiles for the period 1994-2002. The firm-level data are made available by the National Bank of Belgium and are commercialized by BvD BELFIRST. The data contains the entire balance sheet of all Belgian firms that have to report to the tax authorities. In addition to traditional variables such as revenue, value added, employment, various capital stock measures, investments, material inputs; the dataset also has information on entry and exit.

FEBELTEX - the employer's organization of the Belgian Textiles - reports very detailed product-level information on-line (www.febeltex.be). More precisely, they list Belgian firms (311) that produce a certain type of textile product. They split up the textile industry into 5 subsectors: i) interior textiles, ii) clothing textiles, iii) technical textiles,

[26] As Ackerberg and Pakes (2005) show we would be left with a non parametric function $g(\phi_{it} - \beta_k k_{it}, \Upsilon_t^{-1}(k_{it}, i_{it}, s_{it}))$.

[27] For more on the assumption of productivity following a first order Markov process and assuming it to be a scalar unobservable, I refer to Ackerberg and Pakes (2005). Also see De Loecker (2004) for an application of endogenous learning by exporting.

iv) textile finishing and v) spinning. Within each of these subsectors products are listed together with the name of the firm that produces it. From this source I was able to link firms with the number of different products they produce, including other interesting information on the different segments of the textile industry.

I match these firms with BELFIRST and I end up with 308 firms with both firm-level and product-level information. As mentioned in the text, the product-level information is time invariant and this leaves me with a balanced panel of 308 firms for the period 1994 - 2002 (2,772 observations). The average size of the firms in the matched dataset is somewhat higher than the full sample, since mostly bigger firms report the product-level data. Even though I only observe a fraction of the firms in the industry due to the matching, I still cover 70 percent (for the year 2002) of the total employment in the textile industry.

From the BELFIRST dataset I have virtually the entire population of textile producers and this allows me to check for sample selection and sample representativeness. The entry and exit data are detailed in the sense that I know when a firm exits and whether it is a 'real economic exit', i.e. not a merger, acquisition or a scission into other firms.

Another important thing to mention is that by using the information straight from FEBELTEX, it is clear that industrial classification (NACE BELCODE) is sometimes incomplete. If one would merely look at firms producing in the NACE BELCODE 17, there would be left out some important segments of the industry, e.g. the subsector technical textiles also incorporates firms that produce machinery for textile production and these are not in the 17 listings. It is therefore important to take these other segments into the analysis in order to get a complete picture of the industry.

Before I turn to the estimation I report some summary statistics of both the firm-level and product-level data. In Table 2 summary statistics of the variables used in the analysis are given. Since I work with a balanced panel it is not surprising that the average size is increasing over time (11 percent). Value added has gone up over the sample period. In the last column the producer price index (PPI) is presented. It is interesting to notice that since 1996 producer prices fell, only to recover in 2000. Furthermore I also constructed unit prices at a more disaggregated level (3 digit NACEBELCODE) by dividing the production in value by the quantities produced and the drop in prices over the sample period is even more prevalent in specific subcategories of the textile industry and quite different across different subsectors (see Appendix A.2.).

Together with the price decrease, the industry as whole experienced a downward trend in sales. The organization of employers, FEBELTEX, suggests two main reasons for the downward trend in the sales. A first reason is a mere decrease in the production volume, but secondly the downward pressure on the prices due to increased competition has played a very important role. This increased competition stems from both overcapacity in existing segments and from a higher import pressure from low wage countries, Turkey and China more specific.[28] Export still plays an important role, accounting for more than 70% of the total industry's sales in 2002. However, the composition of exports has changed, export towards the EU-15 member states fell back mainly due to the strong position of the euro with respect to the British Pound and the increased competition from low wage countries. This trend has been almost completely offset by the increased export towards Central and Eastern Europe. The increased export is not only due to an increased demand for textile in these countries, but also the lack of local production in the CEECs.[29]

Table 2: Summary Statistics Belgian Textiles (N=308 firms)

Year	Employment	Sales	Value Added	Capital	Material Inputs	PPI
1994	89	18,412	3,940	2,443	13,160	100.00
1995	87	19,792	3,798	2,378	14,853	103.40
1996	83	18,375	3,641	2,177	14,313	99.48
1997	85	21,561	1,365	2,493	16,688	99.17
1998	90	22,869	4,418	2,650	17,266	98.86
1999	88	21,030	4,431	2,574	15,546	98.77
2000	90	23,698	4,617	2,698	17,511	102.98
2001	92	23,961	4,709	2,679	17,523	102.67
2002	99	26,475	5,285	2,805	17,053	102.89
Average	89	21,828	4,367	2,551	16,062	

Note: All monetary variables are expressed in thousands of euro.

To every firm present in this dataset I have matched product-level information. For each firm I know the number of products produced, which products and in which segment(s) the firm is active. There are five segments :1) Interior, 2) Clothing, 3) Technical Textiles, 4) Finishing and 5) Spinning and Preparing.[30] In total there are 563 different products, with 2,990 product-firm observations. On average a firm has about 9 products and 50 percent of the firms have 3 or fewer products. Furthermore, 75 percent of the firms is active in one single segment. Table 3 presents a matrix where each cell denotes the percentage of firms that is active in both segments. For instance, 9.1 percent of the firms is active in both the Interior and Clothing segment. The high percentages in the head diagonal reflect that most firms specialize in one segment.[31] This information is very interesting in itself, since it gives us some information about product mix and product

[28] A clear example of this is the filing of three anti-dumping and anti-subsidy cases against sheets import from India and Pakistan. Legal actions were also undertaken against illegal copying of products by Chinese producers. (Annual Report of Febeltex; 2002)

[29] I refer to other work where the effect of trade liberalization on the Belgian textiles is analyzed (De Loecker; 2005).

[30] I refer to Appendix A for more on the data.

[31] Note that the level of a segment is very disaggregated. There are no conversion tables available to map it into industry classification.

diversification.

Table 3: The Production Structure Across Different Segments
Firms

	Interior	Clothing	Technical	Finishing	Spinning
Interior	*77.2*	9.1	14.4	13.5	4.3
Clothing		*63.6*	10.6	4.5	1.4
Technical			*54.4*	21.3	26.1
Finishing				*53.9*	8.7
Spinning					*59.4*

Note: The cells do not have to sum up to 100 percent by row/column, i.e. a firm can be active in more than 2 segments

The same exercise can be done based on the number of products and as shown in Table 4 the concentration is even more pronounced. The number in each cell denotes the average (across firms) share of a firm's products in a given segment in its total number of products.

Table 4: The Production Structure Across Different Segments
Products

	Interior	Clothing	Technical	Finishing	Spinning
Interior	*83.72*	2.78	8.27	4.41	0.80
Clothing	3.03	*79.28*	15.36	1.86	0.48
Technical	7.01	8.97	*70.16*	9.06	4.79
Finishing	5.75	3.52	15.53	*72.85*	2.35
Spinning	3.72	0.65	27.20	7.40	*61.04*

Note: The cells do sum up to 100 percent by row/column. This table has to be read from the rows only.

The table above has to interpreted in the following way: firms that are active in the Interior segment have (on average) 83.72 percent of all their products in the Interior segment. The analysis based on the product information reveals even more that firms concentrate their activity in one segment. However, it is also the case that firms that are active in the Spinning segment (on average) also have 27.2 percent of their products in the Technical textile segment. Firms active in any of the segments tend to have quite a large fraction of their products in Technical textiles, (8.27 to 27.7 percent).

6 Results

In this section I show how the estimated coefficients of the production function change when a demand side is introduced. In a second step I will discuss the impact on estimated productivity. Here I will verify the extent to which the predicted bias is important for productivity the way it is currently estimated.

6.1 The Estimated Coefficients

I compare my results with a few baseline specifications: [1] a simple OLS estimation of equation (2), the Klette and Griliches (1996) specification in levels [2] and differences [3], $KG\ Level$ and $KG\ Diff$ respectively. Furthermore I compare my results with the Olley and Pakes (1996) estimation technique to correct for the simultaneity bias in specification [4]. I consider three versions of my augmented production function that control for both the omitted price variable bias and the simultaneity bias. In specification [5] I proxy the unobserved productivity shock by a polynomial in investment and capital and the omitted price variable as suggested by Klette and Griliches (1996). In specifications [6] and [7] I include more product level information to proxy for different demand conditions in different segments and segment specific unobservables like quality of products in a given segment. Note that the impact of multi-product firms only comes into play when I analyze the resulting productivity estimates. In specification [6] I include segment dummies and I leave out the Technical Textile segment as a reference segment, although I do not have to since firms can be active in various segments. Finally, specification [7] includes the share of firm's products in a given segment in its total number of products. The latter is to allow for demand conditions and unobserved quality to be different across segments. However, I want to weigh the relevance of this by the share of products a firm sells in any given segment and these are in fact the interaction of the segment dummies with this fraction. The included terms in the last two specifications are essentially capturing the term u^d in equation (4).

Table 5 shows the results for these various specifications. Going from specification [1] to [2] it is clear that the OLS produces reduced form parameters from a demand and a supply structure. As expected, the omitted price variable biases the estimates on the freely chose variables downwards and hence underestimates the scale elasticity. Specification [3] takes care of unobserved heterogeneity by taking first difference as in the original Klette and Griliches (1996) paper and the coefficient on capital goes to zero as expected (see section 1). In specification [4] we see the impact on the estimates of correcting for the simultaneity bias, i.e. the coefficients on the labor decreases somewhat, however, the omitted price variable bias is not addressed in the Olley and Pakes (1996) framework. Both biases are addressed in my three specifications ([5], [6] and [7]). The effect on the estimated coefficients is clear, the correction for the simultaneity and omitted price variable go in opposite direction and therefore by accident making the OLS coefficients look pretty good. The estimate on the capital coefficient does not change much when controlling for the price variation as expected, since capital stock is predetermined. However, it is considerably bigger than in the Klette and Griliches (1996) approach.

The correct scale elasticity estimate is of most concern in KG's work and indeed when correcting for the demand variation, the estimated scale elasticity goes from 0.9477 in the OLS specification to 1.1709 in the KG specification. As mentioned before the latter specification does not take into account of the simultaneity bias which results in upward bias estimates on the freely chosen variables labor and material. This is exactly what I find in specification [5] (also in [6] and [7]), i.e. the implied coefficients on labor drops when correcting for the simultaneity bias (labor from 0.3338 to 0.3075). When estimating the capital coefficient on a balanced panel it is expected to find a lower estimate on capital. Olley and Pakes (1996) find a coefficient on capital of 0.173 and 0.344 on a balanced an unbalanced panel, respectively. To verify this, I estimate a simple OLS production function on an unbalanced dataset capturing the entire textile sector (around 1,000 firms). The capital coefficient obtained in this way is 0.0956 and is very close to my estimate in the balanced panel (0.0879), suggesting that the selection bias is rather small.

The estimated demand elasticity is less negative when correcting for the simultaneity bias as well. The last two columns do not change the estimates a lot, however, it will change the resulting productivity estimates. The message to take out of this table is that both the omitted price variable and the simultaneity bias are important to correct for and that the estimating coefficients change considerably and consequently changes estimated productivity.

An interesting by-product of correcting for the omitted price variable is that I find implied demand elasticities of around -3. I do not place to much weight on the precise point estimate of the demand elasticity.[32] However, I spend some time discussing it as it gives a check on the 'economic' relevance of the demand model I assumed. Before I turn to the impact on the estimated productivity, I allow for a more flexible demand structure by using the product-level information. This allows for the various segments to face different demand elasticities.

[32] Konings, Van Cayseele and Warzynski (2001) use Hall's (1988) method for estimating industry specific mark-ups and they find a Lerner index of 0.26 for the Belgian textile industry, which is well within in the range of my estimates (around 0.32). However, they have to rely on valid instruments to control the for the unobserved productivity shock. A potential solution to overcome this is Roeger's (1995) approach that essentially takes the dual problem of the Hall (1988) approach to overcome the problem of the unobserved productivity shock. This allows to come up with an unbiased estimator for the mark-up, however, no longer able to recover an estimate for productivity.

Table 5: The Estimated Coefficients of the Production Function

Estimates of	OLS	KG Level		KG Diff		OP	Augmented Model					
	[1]	[2]		[3]		[4]	[5]		[6]		[7]	
			Impl Coeff		Impl Coeff			Impl Coeff		Impl Coeff		Impl Coeff
labor	0.2300	0.2319	*0.2967*	0.2451	*0.3338*	0.2113	0.2126	*0.3075*	0.2168	*0.3154*	0.2103	*0.3015*
materials	0.6298	0.6284	*0.8041*	0.5958	*0.8115*	0.6278	0.6265	*0.9063*	0.6365	*0.9260*	0.6360	*0.9118*
capital	0.0879	0.0868	*0.1111*	0.0188	*0.0256*	0.0940	0.1037	*0.1500*	0.0982	*0.1429*	0.1027	*0.1472*
industry output		0.2185		0.2658			0.3087		0.3126		0.3025	
demand elasticity		*-4.58*		*-3.76*			*-3.24*		*-3.20*		*-3.31*	
segment dummies									yes		no	
share segment									no		yes	
Industry dummies						yes	yes		yes		yes	
Nr Obs	1,291	1,291		1,291		985	985		985		985	

[1]: OLS production function
[2]: Klette and Griliches (1996) in levels
[3]: Klette and Griliches (1996) in differences
[4]: Olley and Pakes (1996)
[5]: Correction product differentiation and simultaneity
[6]: [5] + segment dummies
[7]: [5] + segment share dummies
Note: all coefficients are significant at 5 percent level or stricter.

6.2 Segment Specific Demand, Unobserved Product Characteristics and Pricing Strategy

So far, I have assumed that the demand of all the products (and firms) in the textile industry face the same demand elasticity η and I have assumed that the demand shock u_{ijt}^{d} was a pure *i.i.d.* shock. Before I turn to the productivity estimates, I allow for this elasticity to vary across segments and I introduce product dummies. In Appendix A.2 I present the evolution of producer prices in the various subsectors of the textile industry and it is clear that the price evolution is quite heterogenous across the subsectors. The latter suggest that demand conditions were very different across subsectors and now I consider the demand at the 'segment' level.

Firstly, I construct a segment specific demand shifter - segment output deflated- and discuss the resulting estimates and estimated productivity. Secondly, I introduce product dummies to control for product-firm specific shocks, essentially proxying for ξ. In the next section I compare the distribution and present the correlation matrix of the different productivity estimates. Finally I split up my sample according to firms being active in only 1 or more segments. Firms producing in several segments can be expected to have a different pricing strategy since they have to take into account whether their products are complements or substitutes. Note that here the level of analysis is that of a segment, whereas the pricing strategy is made at the individual product level.[33]

6.2.1 Segment Specific Demand

The demand parameter is freed up to be segment-specific by interacting the segment demand shifter (segment output) with the segment share variables.[34] The share variable S_{is} is the fraction of firm i 's products in segment s (M_{is}) in the firm's total number of products ($M_i = \sum_s M_{is}$), where $s = \{1$ (*Interior),* 2 (*Clothing),* 3 (*Technical*), 4 (*Finishing),* 5 (*Spinning and Preparing)*$\}$. Note that now the demand elasticity is allowed even to be firm specific since the share variable is firm specific. As was shown in Tables 3 and 4, using the product information revealed a more correct pattern of activity concentration. Again, my interest lies in the resulting productivity estimate and not as much in the point estimates of the elasticity of demand.

I now turn back to the general setup of the paper with multi-product firms. The demand for every product is given by (8) and q_{It}^{p} captures the product specific demand

[33] In Appendix B I relax the assumption that no synergies exist from producing multiple products by identifying parameters capturing the effect of combining several activities, i.e. the degree of complementarity.

[34] I have also estimated demand relevant parameters one level deeper, see Appendix A.1 for the structure of the segments. However, this leads to a model with 51 different demand elasticities and identification is somewhat harder. In a similar way one can allow for the demand parameter to be time specific and analyze the impact on productivity, see De Loecker (2005) for a more elaborated analysis.

shifter. As in the single product firm case I proxy the demand shifter by output, however, now it is segment output. The segment output I consider is constructed in the following way. I observe firm-level revenue r_{it} and I know the share of the firm's products per segment in its total products produced (S_{is}). I consider the revenue of firm i in segment s to be $r_{ist} = r_{it}S_{is}$. That is, if a firm has 20 percent of its products in segment 1 (Interior Textiles) I assume that 20 percent of its revenue comes from that segment. The relevant weight to construct the segment output is $v_{ist} = \frac{r_{ist}}{\sum_i^{N_s} r_{ist}}$, where N_s is the number of firms active in segment s. The segment output q_{st} is then proxied by $\sum_i v_{ist}\widetilde{r}_{ist}$ as before. I now introduce these terms interacted with the segment share variable in the augmented production function and estimate the following regression (29)

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \sum_{s=1}^{5} \beta_{\eta s} q_{st} S_{is} + \omega_{it}^* + u_{it} \tag{29}$$

Before turning to the estimated productivity ($PDMP2$) I present the estimated coefficients β_{ns} and the distribution of the estimated demand parameter in Table 6. The estimated coefficient are in fact estimates of the Lerner index. One can immediately read of the implied demand parameters for the various segments in the textiles for those firms having all their products in one segment ($S_{is} = 1$).[35]

Table 6: Estimated Demand Parameters and Implied Firm Elasticities

		Interior	Clothing	Technical	Finishing	Spinning
A: Estimated Demand Parameters						
No product dummies	$\beta_{\eta s}$	0.1975*	0.2662*	0.2519*	0.3162*	0.1850*
		(0.0749)	(0.1038)	(0.0916)	(0.1052)	(0.0781)
	η_s $(S_{is}=1)$	*-5.0633*	*-3.7566*	*-3.9698*	*-3.1626*	*-5.4054*
Product dummies	$\beta_{\eta s}$	0.2490*	0.3217*	0.3015*	0.3209*	0.2356*
		(0.0536)	(0.0756)	(0.0649)	(0.0757)	(0.0585)
	η_s $(S_{is}=1)$	*-4.0161*	*-3.1085*	*-4.3168*	*-3.1162*	*-4.2445*
One Segment	$\beta_{\eta s}$	0.2645*	0.3944*	0.3532*	0.4514*	0.2539*
(667 obs)	η_s	*-3.7807*	*-2.5355*	*-2.8315*	*-2.2153*	*-3.9386*
>1 Segments	$\beta_{\eta s}$	0.1573*	0.2263*	0.2219*	0.2214*	0.1452*
(318 obs)	η_s	*-6.3572*	*-4.4189*	*-4.5065*	*-4.5167*	*-6.8871*
B: Firm-Specific Demand Elasticities						
	η_i	mean		*-4.4335*		
		sd		0.7010		
		min		-5.4059		
		max		-3.1627		

Standard errors are given in parantheses and * denotes significance at 1 percent level.

[35]The estimated coefficients of the production function hardly change as expected. Again, the impact of introducing this more flexible demand system has no unambiguous effect on the resulting productivity estimates.

The first row shows the estimated coefficients implying significantly different demand parameters for the various segments. I also include the implied demand parameters relevant for firms having all their products in a given segment. For instance, firms having all their products in the segment Interior face a demand elasticity of -5.1733. In panel B of table 6 I use the firm specific information on the relative concentration (S_{is}) and this results in a firm specific η_i which is in fact a weighted average. I stress that this comes from the fact that firms have multiple products across different segments and therefore the relevant demand condition is different for every firm.[36]

6.2.2 Unobserved Product Characteristics

I now introduce product dummies to capture the product-firm specific demand shocks and time invariant quality unobservables, i.e. $M_i\ \overline{\xi}_i = \sum_j^{M_i} \xi_{ij}$. In terms of section 4.3 the product dummies proxy for the unobserved demand shock - quality - that is firm specific and potentially impacts the investment decision. I assume time invariant unobserved product characteristics. As mentioned above, there are 563 products (K) in total (and a firm produces 9 of these on average) which serve as additional controls in the first stage regression (30). The product dummies are captured by $\sum_{k=1}^{K} \lambda_k PROD_{ik}$ where $PROD_{ik}$ is a dummy variable being 1 if firm i has product k and λ_k are the relevant coefficients. Note that I introduce the product dummies motivating the need to correct for product specific demand shocks and unobserved quality. However, they will also capture variation related to the production side and those two types of variations are not separable.[37] The identifying assumption for recovering an estimate on the capital coefficient is that productivity and the quality shock are independent. However, using the product dummies in the proxy for productivity, the identifying assumption becomes less strong, i.e. I filter out time invariant product unobservables. Note that in the standard approach for identifying the production coefficients, demand variation is not filtered out, both observed and unobserved. Here I allow for product unobservables and demand shocks to impact investment decisions, on top of proxying for the demand shocks proxied by the industry output.

[36] The same is true for the estimated production function coefficients, since they are obtained by correcting for the degree of production differentiation which is firm specific (η_i).

[37] I introduce the product dummies without interactions with the polynomial terms in investment and capital since that would blow up the number of estimated coefficients by K. This then coincides with assuming that the quality unobservable does not enter the investment policy function in the first stage and just correcting for the demand unobservable. However, it matters for the second stage, i.e. this variation is now *not* substracted from deflated sales ($\widetilde{r}$) like the variable inputs. This would imply that the time invariant product dummies would proxy the unobserved demand shock completely. Therefore, the resulting productivity will still capture a time variant quality component.

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m\, m_{it} + \sum_{s=1}^{5} \beta_{\eta s} q_{st} S_{is} + \widetilde{\phi}_t \left(i_{it}, k_{it}, \sum_{k=1}^{K} \lambda_k PROD_{ik} \right) + u_{it} \tag{30}$$

In Table 6 I show that the demand parameters stay almost the same as expected, as well as the production related coefficients. As stressed before, all these extra controls come into play if the interest lies in getting an estimate on productivity (see equation (20)), taking out demand related variation. The estimate of productivity is again obtained by taking the residual out of (30).

In terms of economic interpretation, the table above suggests that firms operating in the Finishing segment (only) face less elastic demand. The high elastic demand segments are Interior and Spinning capturing products - like linen, yarns, wool and cotton - facing high competition from low wage countries.[38] In Appendix A.1 I relate these demand parameters to changes in output prices at more disaggregated level and I find that indeed in those sectors output prices have fallen considerably over the sample period.

6.2.3 Pricing Strategy: Single versus Multi-Product Firms

So far I have assumed that the pricing strategy of firms is the same whether it produces one or more products, or whether it is active in one or more segments. Firms that have products in different segments can be expected to set prices differently since they have to take into account the degree of complementarity between the different goods produced. Remember that the revenue observed at the firm-level is the sum over the different product revenues. Therefore firms producing multiple products will consider the effect of the price of a product on the demand for their other products. I relax this by simply splitting my sample according to the number of segments a firm is active in. The final interest lies in productivity, and again controlling for the demand variation will improve my estimates for productivity. In the third row of Table 6 I present the estimated demand parameters for firms active in only 1 segment and for those active in at least 2. As expected the estimated demand elasticities for the entire sample lies in between. Also, firms producing products in different segments face a more elastic (total) demand since a price increase of one of their product also impacts the demand for their other products in other segments. This is not the case for firms producing only in 1 segment, leading to lower estimated demand elasticities.[39] It is clear that the modeling approach here does allow for various price setting strategies and different demand structures.

[38] See De Loecker; 2005 for a more detailed analysis of increased international competition and Febeltex; 2003 for some suggestive evidence on increased competition in those segments.

[39] Note that now the implied demand elasticities for a firm are given by the weighted sum over the various segments it is active in, where weights are given by the fraction of the number of products in a segment in the total number of products owned by a firm.

6.3 Estimated Productivity: Single versus Multi-product Firms and Perfect Competition versus Product Differentiation

Now I use the residuals from the regressions discussed above to verify the impact on estimated productivity. In turn I verify the impact of correcting for multi-product firms and demand conditions. The estimated productivity (tfp) is obtained by taking the residual of the production functions above and I express them in monetary terms ($tfp = \exp(residual)$), i.e. thousands of euro. In Table 7 I present the mean, standard deviation, minimum and maximum of the estimated productivity distribution for the various specifications. I compare my productivity estimates with the OLS and OP estimates. I consider the estimated productivity corrected for the degree of product differentiation (PD), multi-product firms (MP) and corrected for both using various demand structures ($PDMP1, PDMP2$ and $PDMP3$).

Table 7: Estimated Productivity in euro (1,291 Obs)

Statistic	*OLS*	*OP*	*PD*	*MP*	*PDMP1*	*PDMP2*	*PDMP3*
Mean	10,710	692	639	893	853	578	742
Sd	9,933	629	2,510	160	237	284	231
Min	974	59	16	244	127	57	174
Max	350,982	22,233	90,245	2,812	4,485	2,141	2,210

PD: Product Differentiation, *MP*: Multi-product firms, *PDMP1*:*PD* and *MP*
PDMP2: *PDMP1*+segment demand shocks and *PDMP3*: *PDMP2*+product dummies

In the table above it is very clear that as I introduce a simple demand side, the estimated productivity distribution changes. The correction for the degree of product differentiation scales the estimated productivity by $\left(\frac{\eta}{\eta+1}\right)$ and filters out aggregate demand variation across time proxied by industry output. Remember that the PD estimate corrects (in a simple way) for the degree of product differentiation and thus filters out demand variation from the productivity estimate. I do not rely on this estimate too much since the demand shifter is picking up other industry specific effects (like industry productivity growth) and all these are essentially subtracted from the productivity estimate and later on I allow for segment specific demand shifters.

Assuming a product and time invariant demand elasticity the correction does not change the cross sectional variation of productivity. The correction for the number of products per firm does change this cross sectional variation. When I only correct for the number of products produced per firm, the average productivity is 893 euro and is considerably higher than the 639 euro with the PD estimate. The $PDMP1$ estimate controls for both and the estimated productivity is 853 euro on average. and is around 30 percent higher than the simple model of product differentiation without multi-product firms. Specification $PDMP2$ is based on segment specific demand shifters as expressed in equation (29). The average productivity is 578 euro and is quite different from the standard case ($PDMP1$), i.e. lower. This is mainly due to the strong assumption in the latter that the elasticity of substitution is the same across all different products within the textile industry. Note that it is still much smaller than the OP productivity estimate. Finally, I present the estimate that also controls for product unobservables (proxying time invariant quality) in the last column. It changes the average productivity average considerably, 742 euro, compared to the case without product dummies.

Comparing the average productivity estimates in Table 7 makes the OLS estimate '*look*' pretty good. However, we know that the simultaneity and omitted price variable bias work in the opposite direction. In addition, most research is mostly concern with either productivity growth or the impact of some variable of interest on the productivity, and not the actual productivity measured in monetary units.[40] Therefore I now present the correlation among the different productivity estimates. Re-scaling does not change the results of estimating a regression with the productivity estimate on the left hand side on some variable of interest. However, when the cross sectional variation is changed, this will have an impact. Table 8 shows the correlation coefficients and they are all significant at the 1% significance level.

Table 8: Correlation Coefficients

	OLS	*OP*	*PD*	*MP*	*PDMP1*	*PDMP2*	*PDMP3*
OLS	1.0000						
OP	0.9997	1.0000					
PD	0.9802	0.9801	1.0000				
MP	0.4846	0.4895	0.3920	1.0000			
PDMP1	0.5726	0.5773	0.4861	0.9886	1.0000		
PDMP2	0.1444	0.1459	0.1520	0.6500	0.6054	1.0000	
PDMP3	0.1989	0.2009	0.1883	0.7602	0.7762	0.9395	1.0000

PD: Product Differentiation, *MP*: Multi-product firms, *PDMP1*:*PD* and *MP*
PDMP2: *PDMP1*+segment demand shocks and *PDMP3*: *PDMP2*+product dummies

As expected, the correlation of my productivity estimates (PD, MP, $PDMD1$, $PDMP2$ and $PDMP3$) with the standard OLS and even the OP estimates is very low It is important to note that the correlation between the OP productivity estimate and the corrected estimates is rather low, (0.1459-0.9801), casting some doubt on the wide use of this estimator in various contexts. The fact that the PD estimate is highly correlated with the OP estimate is not surprising given the simple structure in this example, i.e. product and time invariant η. The correction for multi-products seems to really change the distribution of productivity. Note that the correlation between the OP and the PD estimates

[40]In addition, the last specification includes product dummies and changes the interpretation.

is very high although being quite different in levels due to the re-scaling. However, in a regression framework this would not change results. This is the case when multi-product firms and product level demand is considered, the correlation drops drastically with the OLS and the OP estimates. It is also clear that the estimated productivity distribution is fairly robust with respect to my different specifications on the demand side: an industry wide demand shifter versus segment specific demand shifters and introducing product dummies, i.e. the correlation among the last three specifications is around 70 percent.

From the above it is clear that productivity estimates are biased in the presence of imperfect competitive markets and ignoring the underlying product space when considering firm-level variables, i.e. multi-product firms. The number of product per firm correction changes the cross sectional variation in productivity in addition to the re-scaling due to the demand parameter.

Finally, in Table 9 I present the average productivity growth for the various specifications, where I compare my productivity estimate controlling for multi-product firms and product differentiation at the segment level ($PDMP2$) with the OLS and the OP estimates.

Table 9: Average Productivity Growth (percentage)

Segment	OLS	OP	$PDMP2$
Interior	0.12	0.18	-0.76
Clothing	1.15	1.22	0.16
Technical Textiles	1.07	1.12	0.34
Finishing	0.16	0.15	0.41
Spinning and Preparing	0.85	0.89	0.94
Textile Industry	0.52	0.56	0.03
Average	0.67	0.71	0.22

From the table above it is clear that the estimated productivity growth is quite different when correcting for the omitted price variable. Productivity growth is lower in all segments and even switches sign in the Interior segment. Based on expression (22) the difference in measured productivity growth comes from correcting for the product differentiation and the product mix, getting at the actual productivity growth. It is clear that when these variables are then used in a second stage to relate them to variables of interest that results will be different, i.e. smaller effects on productivity.

7 Conclusion

In this paper I suggest a method to correct for the omitted price variable in structural estimation of productivity. I have introduced a very simple demand side and I explicitly allow firms to have multiple products. I introduce a simple aggregation from product space into firm space and derive a straightforward estimation strategy. I show that measured productivity increases need no to reflect actual productivity increase and this sheds some light on papers trying to link trade liberalization and trade protection on firm-level productivity (growth). I illustrate this methodology by analyzing productivity in the Belgian textiles using an unique dataset that in addition to firm-level data has product-level information.

All this is valid if one is willing to make certain assumptions on the demand side and the production of varieties. There is still room for extensions and improvements along those lines. The unobserved quality component and its relation with respect to the productivity shock is not modelled here. Furthermore, productivity is still assumed to follow a Markov process over time and there is no room for a firm to changes this endogenously. However, it is clear that the resulting productivity estimates do change quite drastically if one is no longer ignorant about the product level and the degree of product differentiation in an industry, and how these factors differ over time and firms.

Appendix A: The Belgian Textiles

1. A Quick Overview

The textile industry recently had a lot of attention in the Belgian press because of a severe series of strikes in virtually all firms by the end of 2003. The main reason for this was the unsuccessful wage bargaining between the industry unions and the employers. The Belgian textile and clothing industry production dropped by 7.1% after the year 2000 after a successful period between 1995 and 2000, where total production grew with 13.7%. Besides the decrease in output during the latest years, the industry's employment fell by 30% in 2002 compared to the year 1992.

There are three main applying segments consisting of the Interior, Clothing, and Technical textiles. Besides those there are two supplying segments (Spinning and Finishing) that sell their products to one or more of the applying segments. Following the industry classification of the textile industry into five big subsectors, it is striking to see that the different subsectors have had different experiences. The Interior Textile (40% of industry's value added) increased its production during the last years, mainly by the success of carpets. The Clothing industry (23% of industry's value added) experience is quite different, activity went down by 9% and this is mainly due to the increased competition from low wage countries. The subsector "Technical Textile" captures all products that provide technical solutions to textile problems and is quite heterogeneous. It accounts for 24% of the industry's value added and grew moderately at a rate of 2%. R&D activities are closely connected with other industries. The production in the Spinning sector fell back quite sharp. The same is true for the "Finishing" sector, this sector finishes textile products (e.g. making it waterproof) and clearly follows the general negative trend.[41]

Investments fell back quite drastically over the last few years, with 17,1% and 18,8% in the years 2001 and 2002, respectively. The main reason for this is the low rentability and a decreasing demand forecasts. The decrease in investment's rentability is mainly due to the increased production costs such as the prices of intermediate inputs. These price increases are not being fully transmitted into the prices due to the increased competition (see Table 2 in text).

Finally, I present the structure of the different segments, sub-segments and the products in my dataset in Table A.1. The different levels are important to structure the regressions and serve as additional sources of variation to identify demand parameters.

[41] For a detailed description of what the various segments include I refer to the online database of FEBELTEX (www.febeltex.be).

Table A.1.: Segment Structure: Number of Subsegments and Products per Segment

Interior (9)	Clothing (18)		Technical (9)	Finishing (7)	Spinning (9)
	Clothing fabrics	Knitwear			
Bed linen	Accessories	Accessories	Agrotech	Carpeting	Blended aramid, polyamid or polyacrylic
Carpets	Baby clothes & children's	Babies' wear	Buildtech	Knitted fabrics	Blended artificial yarns
Kitchen linen	Men's wear	Bath	Geotech	Material before spinning	Blended cotton or linen yarns
Matress ticking	Nightwear & underclothing	Children's wear	Indutech	NonWoven	Blended polyester yarns
Table linen	Others	Fabrics for ...	Medtech	Woven fabrics	Blended polypropylene or chlorofibre yarns
Terry toweling articles	Rain-, sportswear & leisure	Nightwear	Mobiltech	Yarns	Blended yarns
Trimming	Women's wear	Outerwear	Packtech	Specialities	Filament Yarns
Upholstery & furnishing fabr.	Workwear & protective suits	Sportswear	Protech		Spun Yarns ($>$ 85% of 1 fibre)
Wallcoverings		Stockings- tights- socks	Sporttech		Synthetic Fibres
		Underwear			
19	61	36	231	132	84

Note: The second row indicates the number of subsegments within a given segment. The last row indicates the number of products within a given segment
I also estimated demand elasticities at the level of aggregation suggested above, i.e. 52 different parameters.

2. Producer Prices and Demand Elasticity

As mentioned in the text a producer price index is obtained by taking a weighted average over a representative number of products within an industry, where weights are based on sales (market shares). In the case of Belgium the National Institute of Statistics (NIS) gathers monthly information of market relevant prices (including discounts if available) of around 2,700 representative products (an 8 digit classification - PRODCOM - where the first 4 are indicating the NACEBELCODE). The index is constructed by using the most recent market share as weights based on sales reported in the official tax filings of the relevant companies. The relevant prices take into account both domestic and foreign markets and for some industries both indices are reported.

I present unit prices at the 3 digit NACEBELCODE (equivalent to 4/5 digit ISIC code). I constructed these by dividing total value of production in a given subcategory by the quantity produced. Table A2 gives the PPI for the various subcategories with 1994 as base year except for the 175 category (Other textile products, mainly carpets). I do not use these to deflate firm-level revenues since I have no information in which category (ies) a firm is active since the product classification cannot be uniquely mapped into the NACEBELCODE and firms are active in various subcategories. The codes have the following description: *171*: Preparation and spinning of textile fibres, *172*: Textile weaving, *173*: Finishing of textiles, *174*:.Manufacture of made-up textile articles, except apparel, *175*: Manufacture of other textiles (carpets, ropes, ...), *176*: Manufacture of knitted and crocheted fabrics and *177*: Manufacture of knitted and crocheted articles.

Table A2: Producer Prices (Unit Prices) at Disaggregated Level

	171	*172*	*173*	*174*	*175*	*176*	*177*
1994	100	100	100	100	-	100	100
1995	99.4	96.7	110.4	111.0	-	100.9	100.7
1996	100.9	94.5	101.1	117.9	100	103.4	94.8
1997	103.7	94.5	101.3	108.5	99.2	93.9	97.5
1998	102.8	96.0	108.0	117.6	101.5	93.3	97.6
1999	95.0	95.8	100.6	118.2	99.6	94.8	92.9
2000	94.3	94.6	119.3	106.2	102.0	84.1	95.5
2001	96.7	93.2	108.4	107.7	104.1	86.9	101.3
2002	97.3	94.2	110.7	103.1	107.2	85.8	106.1
demand elasticity	-5.4054		-3.1626	-3.7566	n.a.	n.a.	-3.7566

Several observations are important to note. Firstly, there is considerable variation across subcategories of the textiles industry in terms of price changes over the period 1994-2002. The sector *Manufacture of knitted and crocheted fabrics* (176) has experienced a severe drop in output prices (14.2 percent) over the sample period, whereas the output prices in the *Finishing of Textiles* (173) has increased with more than 10 percent. Secondly, the evolution in the various subcategories is not smooth, periods of price increases are followed by decreases and the other way around. Thirdly, most of the price decreases occur at the end of the nineties when imports from Central and Eastern Europe were no longer quota restricted as agreed in the Europe Agreements (see De Loecker; 2005 for a detailed analysis of the impact of the trade regime). Finally, in Table 6 estimated demand elasticity were presented for the different segments. It is interesting to note that the segment (Spinning) with the most elastic demand (-5.3135) has indeed experienced a negative price evolution (2.7 percent). The latter segment also captures weaving activities which in turn also experienced a price decrease (5.8 percent). The segment (Finishing) with the least elastic demand (-3.2051) has had a sharp increase in its output prices (10.7 percent). The estimated demand elasticities from Table 6B are given in the last row for those subcategories I could map into segments.

Appendix B: Production Synergies

When aggregating the product-level production function to the firm-level, I have assumed that there are no cost synergies or complementarities in producing several products within one firm. However, we know that the textile sector captures both supplying (*Spinning* and *Finishing*) and applying segments (*Technical textiles*). Firms that produce both type of products can expect to potentially benefit from combining both activities (or more). Therefore, I relax the assumption on the production technology by introducing a parameter σ_{sr} capturing the complementarity in production of combining different products (here segments), where r and s are the different segments. More formally the aggregation from product-level production into firm-level is given by (B.1)

$$q_i = \sum q_{ij} + \sum_{s=1}^{5} \sum_{r=s}^{5} \sigma_{sr} S_{isr} \tag{B.1}$$

where S_{isr} is 1 if a firm i is active in both segment r and s and zero otherwise and σ_{sr} the corresponding coefficients. Proceeding as before, I obtain the following augmented production function (B.2) where $\beta_{\sigma_{sr}} = \left(\frac{\eta_s+1}{\eta_s}\right) \sigma_{sr}$.

$$\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \sum_{s=1}^{5} \beta_{\eta s} q_{st} S_{is} + \sum_{s=1}^{5} \sum_{r=s}^{5} \beta_{\sigma_{sr}} S_{isr} + \omega_{it}^{*} + u_{it} \tag{B.2}$$

The estimated segment demand elasticities are somewhat more negative, however, the same economic interpretations apply, i.e. Interior and Spinning are the most elastic segments (-6.81 and -6.76). I now present the estimated coefficients on the extra term S_{isr} in Table B.1.

Table B.1: Estimated Product Complementarity

$\beta_{\sigma_{sr}}$		s				
		1	2	3	4	5
	1	-0.37*	0.15**	*0.39**	0.04	0.35*
	2		-0.27*	*0.36**	0.08	0.06
r	3			-0.61*	*0.28**	*0.23**
	4				-0.39*	0.22*
	5					-0.41*

Note: * significant at 1% level, **: at 10% level

A positive sign on the coefficients in the table above reflects a (on average) higher output conditional on inputs and demand conditions for a firm active in any two given segments. Firms combining any activity with Technical textiles (3) generate a higher output. To obtain the entire firm relevant effect, we have to add up the relevant terms, e.g for a firm active in segment 1 and 3: $-0.37 + 0.39 = 0.02$, suggesting gains from diversification. The latter is also reflected in the negative coefficients on the head diagonal. In terms of estimated productivity this implies that we potentially overestimate the productivity of firms diversifying by attributing production synergies/complementarities to productivity. As discussed extensively in the text, this might again mislead findings of the impact of changing operating environments like trade liberalization on productivity (growth). Finally, I recover the estimated productivity using the specification described in equation (B.2). Average productivity is 510 euro with a standard deviation of 293 and minimum and maximum productivity are 27 and 1,907, respectively. The correlation with my other specifications discussed in the text is around 0.9 and the correlation with the OLS is only 0.11. As expected, the additional correction for potential synergies does not change the coefficients of the production function, however, it does change the level of estimated productivity and consequently its relation to variables of interest.

Appendix C: Imperfect Input Markets

Throughout the paper the focus has been on the degree of product differentiation and thus imperfect competitiveness of the output market. A methodology has been suggested to correct for this when estimating productivity. However, one can use the same argument on the input market. I now discuss how my methodology implicitly takes partly care of potential imperfect input markets. I illustrate my point assuming an imperfect market for materials, i.e. each firm pays a different price. Denote the cost of material inputs by M and it is the product of the quantity and price, M_q and P^M, respectively. In logs this means that implicitly $m_{it} - p_{it}^m = m_{qi}$.

Estimating the production function as before and ideally the materials used in quantities are obtained by deflating material costs with a firm-level material price

$$
\begin{aligned}
r_{it} &= q_{it} + p_{it} \\
&= \alpha_l l_{it} + \alpha_m (m_{it} - p_{it}^m) + \alpha_k k_{it} + \omega_{it} + p_{it} + u_{it}^q \\
&= \alpha_l l_{it} + \alpha_m m_{it} + \alpha_k k_{it} + \omega_{it} + (p_{it} - \alpha_m p_{it}^m) + u_{it}^q
\end{aligned}
$$

However, this material price variable p_i^m is not observed and becomes an additional unobservable correlated with the material variable m_{it}. A nice feature of my methodology is that p^m is likely to be correlated with output prices p_i. Holding other things constant, a higher price for materials leads to higher output prices in an imperfect competitive output market, the extent of this depends on the relevant mark-up.

Therefore, my control for the unobserved price term (which is in fact substituted by an expression in terms of the production due to a given demand structure, i.e.$((q_{it} - q_{It} - u_{it}^d)\frac{1}{\eta} + p_{It} = p_{it})$ picks up this variation as well. This leads to the following augmented production function

$$
\widetilde{r}_{it} = \beta_0 + \beta_l l_{it} + \beta_m \, m_{it} + \beta_k k_{it} + \left(\beta_\eta q_{It} - \beta_m p_{it}^m\right) + \omega_{it}^* + \xi_{it}^* + u_{it}
$$

Since both correlations with the error terms go in the same direction (negative), the potential correlation problem is reduced to variation in material prices that are not picked up by variation in output prices, i.e. a perfect output market combined with an imperfect input market. Even if I would have an industry wide material deflator P_I^M, the firm-level variation in material price away from the deflator - $(p_{it}^m - p_{It}^m)$ - still biases the estimate on materials and does not solve the problem.

References

[1] Ackerberg, D,.Caves, K and Frazer, G. (2004). Structural Estimation of Production Functions, U.C.L.A. mimeo

[2] Ackerberg, D, Benkard, L., Berry, S. and Pakes, A. (2005). Econometric Tools for Analyzing Market Outcomes, Paper prepared for *Handbook of Econometrics*, J.J. Heckman (ed.)

[3] Ackerberg, D. and Pakes, A. (2005). Notes on Relaxing the Scalar Unobservable Assumption in Olley and Pakes, mimeo, Harvard University.

[4] Arellano M. and Bond, S. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, vol.58, 277-297.

[5] Bartelsman, E. J. and Doms, M. (2000). "Understanding Productivity. Lessons from Longitudinal Micro Data", *Journal of Economic Literature*

[6] Bernard, A.B., Redding, S. and Schott, P.K. (2003). "Product Choice and Product Switching", NBER WP 9789

[7] Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation", *Rand Journal of Economics*, 25(2), 242-262.

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

[9] Buettner, T. (2004). R&D and the Dynamics of Productivity, mimeo, LSE.

[10] De Loecker, J. (2004). "Do Exports Generate Higher Productivity? Evidence from Slovenia", LICOS DP 151.

[11] De Loecker, J. (2005). "Productivity Increase or Different Demand Conditions? The Effect of Trade Liberalization in the Belgian Textiles, K.U. Leuven mimeo.

[12] Dunne, T. and Roberts, M.J. (1992). "Costs, Demand and Imperfect Competition as Determinants of Plant-Level Output Prices", DP CES 92-5, Center for Economic Studies, US Bureau of the Census.

[13] Eslava, M., Haltiwanger, J., Kugler, A. and Kugler, M. (2004). "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia", CEPR DP 4569.

[14] Ericson, R. and Pakes, A. (1995). Markov Perfect Industry Dynamics: A Framework for Empirical Work, *Review of Economic Studies*, Vol 62 (1), 53-82

[15] Febeltex. (2003)."Anual Report 2002-2003 FEBELTEX.

[16] Foster L., Haltiwanger, J. and Syverson, C. (2003). "Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability?", University of Chicago mimeo.

[17] Hall, R.E. (1988)."The Relation between Price and Marginal Cost in US Industry", *Journal of Political Economy*, 96 (5), 921-947.

[18] Harrison, A.E. (1994). "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence, *Journal of International Economics*, 36, 53-73.

[19] Jaumandreu, J. and Mairesse, J. (2004). "Using price and demand information to identify production functions", mimeo.

[20] Katayama, H., Lu, S. and Tybout, J. (2003). "Why Plant-Level Productivity Studies are Often Misleading, and an Alternative Approach to Inference, NBER WP.

[21] Klette, T.J. and Griliches, Z (1996). "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous", *Journal of Applied Econometrics*, vol 11 (4), 343-361.

[22] Konings, J., Van Cayseele P. and Warzynski, F. (2001). "The Dynamics of Industrial Mark-Ups in Two Small Open Economies; Does National Competition Policy Matters?", *International Journal of Industrial Organization*, 19, 841-859.

[23] Levinsohn, J. and Melitz, M. (2002). "Productivity in a Differentiated Products Market Equilibrium", Harvard mimeo.

[24] Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables., *Review of Economic Studies*, Vol. 70, pp. 317-342.

[25] Melitz, M. (2001). "Estimating Firm-Level Productivity in Differentiated Product Industries", Harvard mimeo.

[26] Muendler, M.A. (2004). "Estimating Production Functions When Productivity Change is Endogenous", UCSD WP 05.

[27] Olley, S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry, *Econometrica*, Vol 64 (6), 1263-98.

[28] Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants, *Review of Economic Studies*, 69 (1), 245-76

[29] Petropoulos, W. (2000).."Productivity in an industry with differentiated products", University of Michigan, mimeo.

[30] Roeger, W. (1995). "Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for US Manufacturing", *Journal of Political Economy*, 103 (2), 316-330.

[31] Smarzynska, B. (2004). "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages", *American Economic Review, 94 (3), 605-627.*

[32] Van Biesebroeck, J. (2005). "Exporting Raises Productivity in the Sub-Saharan African Manufacturing", *Journal of International Economics*, forthcoming.