Theory for Extending Single-Product Production Function Estimation to Multi-Product Settings *

Emmanuel Dhyne
National Bank of Belgium

Amil Petrin
University of Minnesota & NBER,

Valerie Smeets
Aarhus University

Frederic Warzynski
Aarhus University[†]

September 29, 2020: Preliminary

Abstract

We introduce a new methodology for estimating multi-product production functions. It embeds the seminal contributions of Diewert (1973) and Lau (1976) in the semi-parametric econometric framework of Olley and Pakes (1996). We address the simultaneity of inputs and outputs while allowing for the possibility of one unobserved productivity term for each product, all of which may be freely correlated with inputs and outputs. We recover estimates of firm-product marginal costs using the input and output elasticities by extending Hall's (1988) single-product result to our multi-product setting using McFadden (1978). We focus on six 6-digit Belgian "industries" that produce two products, finding all but five of the forty-eight input coefficients are positive and thirty eight are strongly significant. We find outputs are substitutes as the coefficient on "other good output" is always negative and highly significant. 100% of marginal cost estimates are positive and close to 80% of markups are estimated to be greater than 1. We find very similar results when we move to 4-digit industries, when we use similar multi-product data from France, and when we use the trans-log approximation We develop two tests for the appropriateness of approximating multi-product production with a collection of single-product production functions.

---

*We thank several very helpful referees, Andrew Bernard, Jan De Loecker, Swati Dhingra, Penni Goldberg, John Haltiwanger, Marc Muendler, Ariel Pakes, Peter Schott, Chad Syverson, Dan Trefler, and seminar participants at the 2014 NBB Conference, Riksbank, the 11th International Conference, Monash University, Mannheim, EITI2015, IIOC2015, ATW2015, Rice University, Texas A&M, DIEW2015, NOITS2015, Uppsala, the final COMPNET conference, ESWC2015, LSE, Hong Kong University, UCD, Lund, Toronto, UCLouvain, Chicago Fed, RIDGE, AEA2017, Bundesbank, SED2017, Oxford and the NBER IO Winter meeting 2020. Petrin, Smeets and Warzynski thank the National Bank of Belgium for its financial support. The authors are also extremely grateful for the data support provided by the NBB Statistical department. All presented results respect the all confidentiality restrictions associated with the data sources. The views expressed are those of the authors and do not necessarily reflect the views of the NBB. All errors are ours.

[†]Corresponding author - email : fwa@asb.dk

1 Introduction

A well-known anecdotal fact is that most firms that produce one good also produce at least one or more other similar goods. This fact has been confirmed in recent micro-level production data across several countries.[1] Prior to these micro-level data being available, most micro-level data sets on production only included total revenue generated by all of a firm's products, in addition to measurements on categories of input expenditures like the wage bill, spending on capital or intermediate inputs. In these new data - in addition to the same input measurements - product-level quantities and revenues are reported separately for each product. In this paper we show how this kind of data can be used to improve upon previous estimation of production functions and the implied estimates of marginal costs and markups.

We introduce a new methodology for estimating multi-product production functions. It embeds the seminal contributions of Diewert (1973) and Lau (1976) in the semi-parametric econometric framework following Olley and Pakes (1996) and the ensuing literature. The intuition behind the approach is straightforward. The standard single product production function gives the maximal output for any combination of inputs (e.g. labor, capital, and intermediate inputs). A multi-product production function extends the single product setting by giving the maximal output achievable of any one good holding inputs levels *and* the levels of other output goods produced constant. This approach allows for the possibility that any one output can be a substitute or a complement with any other output conditional on inputs.

We (re-)derive several testable conditions that must hold for the estimated input and output elasticities to be consistent with multi-product production. We also extend results from Petropoulos (2001) and Ackerberg, Benkard, Berry, and Pakes (2007) to address the simultaneity of inputs and outputs. Our approach allows for the possibility of one unobserved productivity term for each firm-product, each one of which may be freely correlated with inputs and outputs.

We focus on six 6-digit Belgian "industries" that produce two products. Consistent with multi-product production theory, in the Belgian data all but five of the forty-eight input coefficients are positive, and thirty eight of these forty-three positive input coefficient estimates are strongly statistically significant. The coefficient on "other good output" is always negative and highly significant suggesting quantities are substitutes for one another holding inputs constant. Returns to scale range from 0.93 to 1.12 for 11 of 12 estimated specifications with several constant returns technologies. 100% of marginal cost estimates are positive and close to 80% of markups are estimated to be greater than 1. We find

[1] See e.g. Bernard, Redding and Schott (2010, 2011) and Mayer, Melitz and Ottaviano (2014, 2018).

very similar results when we move to 4-digit industries, when we use similar multi-product data from France, and when we use the trans-log approximation.

Hall (1986, 1988) shows in the case of single-product production, minimization of the variable cost function yields a relationship between the markup and the elasticities of output with respect to an input and observed input expenditures. We derive the multivariate analog where we express marginal costs as a function of output-input elasticities, individual input expenditures, and output quantities. Deloecker and Warzynski (2012) assume single-product production and recover firm-specific markups by applying Hall's insights to standard micro-level production data where only total revenue is recorded. Using this new micro-level data on individual output quantities and revenues we show how one can allow for and estimate one unobserved marginal cost term for each firm-product using the multi-product variable cost minimization problem.

Using Indian manufacturing data, De Loecker et al (2016) is the first paper to tackle these questions since the availability of this kind of data. They impose two key assumptions in their representation of multi-product production. *Assumption A1* maintains that all production is single-product production. They estimate production function parameters using only the observations on single-product firms. *Assumption A4* maintains that it is possible to partition inputs across these different single-product production functions. They then show that cost minimization - under *A1 and A4* - provides for a rule for partitioning inputs across the different single-product production technologies, and they use this rule for identifying firm-product estimates of marginal costs. We develop two tests for the appropriateness of approximating multi-product production with a collection of single-product production functions.

Several follow-up papers using multi-product data extend the De Loecker et al (2016) methodology on important economic dimensions. Valmari (2016) extends the cost minimization conditions to profit maximization conditions by adding a demand side to the model. Gong and Sickles (2018) show how to allow for different production functions for multi-product firms versus single-product firms in a setting where stochastic frontier analysis is the maintained production model. Orr (2019) provides alternatives to assumptions *A1 and A4* and Itoga (2019) follows Orr's extension. De Loecker et al (2016) and all of the papers that have followed are complementary to our approach in the sense that they provide several restrictions that may be useful for improving the precision of our parameter estimates.[2]

The rest of the paper is structured as follows. In Sections 2 and 3 we cover the

[2]There is also an interest among practitioners in extending the Directional Distance Function to allow for multi-product production (see e.g. Fare, Martins-Filho, and Vardanyan (2009) or Kuosmanen and Johnson (2019).)

theory and the data. Section 4 discusses estimation of marginal costs. Section 5 discusses estimation. Section 6 has the results and Section 7 concludes.

2 Multi-Product Production

Using Diewert (1973) and Lau (1976) we review the theoretical conditions under which single- and multi-product production functions exist and their testable implications.

2.1 Single Product Firms

The primitive of production analysis is the firm's production possibilities set T. In the single-product setting T lives in the non-negative orthant of R^{1+N} and contains all values of the single output q that can be produced by using N inputs $x = (x_1, x_2, \ldots, x_N)$, so if $(\tilde{q}_1, \tilde{x}) \in T$, then $\tilde{q}_1$ is producible given $\tilde{x}$, The single-product production function $F(x)$ - the production frontier - is defined as:

$$q^* = F(x) \equiv max\{q \,|\, (q, x) \in T\}.$$

$F(x)$ admits some well-known testable properties. If inputs are *freely disposable* then an output level achieved with the vector of inputs x' can always be achieved with a vector of inputs x'' where $x'' \geq x'$. This implies the production function is weakly increasing in inputs (Diewert (1973)). The production function $F(x)$ should also be concave in the freely variable inputs holding fixed inputs constant and it should be quasi-concave in the fixed inputs holding the freely variable inputs constant (Lau (1976)).

2.2 Multi-Product Firms

With M outputs and N inputs the firm's production possibilities set T lives on the non-negative orthant of R^{M+N}. It contains all of the combinations of M non-negative outputs $q = (q_1, q_2, \ldots, q_M)$ that can be produced by using N non-negative inputs $x = (x_1, x_2, \ldots, x_N)$ so if $(\tilde{q}, \tilde{x}) \in T$ then $\tilde{q} = (\tilde{q_1}, \ldots, \tilde{q_J})$ is achievable using $\tilde{x} = (\tilde{x_1}, \ldots, \tilde{x_N})$. For good j produced by the firm let the output production of other goods be denoted by q_{-j}. For any (q_{-j}, x) Diewert (1973) defines the *transformation function* as

$$q_j^* = F_j(q_{-j}, x) \equiv max\{q_j \,|\, (q_j, q_{-j}, x) \in T\}.$$

If no positive output of q_j is possible given (q_{-j}, x) then he assigns

$$F_j(q_{-j}, x) = -\infty.$$

We develop the properties of $F_j(q_{-j}, x)$ under a mix of assumptions from Diewert (1973) and Lau (1976).

We follow Lau (1976) and divide outputs and inputs (q_{-j}, x) into those that are variable v in the short-run and those that are not, denoted by K. Alternatively, we could do all of our analysis conditional on q_{-j}, with (v, K) partitioning only the variable from the fixed inputs. We sometimes abuse notation by expressing (q_{-j}, x) as (v, K) and by writing $F_j(v, K)$.

We assume the production possibilities set T satisfies the following four *Conditions P*:

(i) P.1 T is a non-empty subset of the non-negative orthant of R^{M+N}

(ii) P.2 T is closed and bounded,

(iii) P.3 If $(q, x_k, x_{-k}) \in T$ then $(q, x'_k, x_{-k}) \in T \ \forall x'_k \geq x_k$.

(iv) P.4 The sets $T^K = \{v \,|\, (v, K) \in T\}$ are convex for every K; the sets $T^v = \{K \,|\, (v, K) \in T\}$ are convex in K for every v.

Conditions P.1 and P.2 are weak regularity conditions on T that require the production set to be non-empty, closed, and bounded. Condition P.3 is a free disposal condition on inputs; if you can produce q_j given (q_{-j}, x), then you can produce q_j with any $x^{'} \geq x$. Diewert (1973) uses these free disposal conditions to prove that output is weakly increasing in any input holding all other inputs and outputs constant. Diewert (1973) then shows if we add the condition that T is convex, there exists a well-defined production function that is concave in the inputs, ensuring decreasing marginal rates of substitution among inputs. The standard concavity tests in a single product setting can be directly extended to test whether the multi-product theory holds.

Convexity on T rules out the possibility of increasing returns to scale. Condition P.4 extends the convexity on T assumption from Diewert (1973) to the disjoint biconvexity assumption of Lau (1976). Under disjoint biconvexity we have convexity of the freely variable inputs and outputs v holding fixed inputs and outputs K constant, and convexity in the fixed variables K holding the freely variable v constant. This setup allows for the possibility of overall increasing returns to scale - non-convexities in T - while maintaining decreasing marginal rates of substitution between elements in v and similarly for the elements in K. The analysis can be done unconditionally or conditional on outputs q_{-j}.

The following theorem formalizes the above claims.

Theorem 2.1 (Transformation Function) *Under P.1-P.4 the function $F_j(q_{-j}, x)$ is an extended real-valued function defined for each $(q_{-j}, x) \geq (0_{M-1}, 0_N)$ and is non-negative* *on the set where it is finite.* $F_j(q_{-j}, x)$ *is non-decreasing in* x *holding* q_{-j} *constant,* $F_j(v, K)$ *is concave in* v *for all* K*, and* $F_j(v, K)$ *and quasi-concave in* K *for all* v*.*

Proof: See Appendix A.

The empirical implication of disjoint convexity is as follows. Convexity in the elements of v (conditional on any K) results in a production function that is concave in v holding K constant. For the elements in K convexity in K given v results in the production function being quasi-concave in K given v.

2.3 "Unobserved" Inputs

Historically, in the single-product production literature it is common to allow for a component of the error to affect output and be observed by the firm when it is making its input decisions (Griliches and Mairesse (1995)). This factor is an "unobserved" technical efficiency term that is unobserved to the researcher and is allowed to be freely correlated with input choices. In our setting with multiple products we want to allow for one possible "unobserved" technical efficiency term for each output produced, with the entire vector of these unobserved shocks denoted

$$\omega = (\omega_1, \omega_2, \ldots, \omega_M).$$

In this section we briefly outline how to incorporate these factors into our theory framework. The main result is all of the components of the theorem continue to hold with the caveat now that everything is conditional on ω.

We extend the production possibilities set to the case where - in addition to containing observed M outputs q and observed N inputs x - we now add the "unobserved" M inputs ω, so $(q, x, \omega) \in R^{M+N+M}$. $(q, x, \omega) \in T$ if (e.g.) the vector of outputs q can be produced with observed and unobserved inputs x and ω respectively. Define

$$q*_j = F_j(q_{-j}, x, , \omega) = max(q_j | q_j, q_{-j}, x, \omega) \in T$$

and let it equal $-\infty$ if there is no non-negative q_j such that $(q_j, q_{-j}, x, \omega) \in T$.

Now we restate what the production possibilities set T must satisfy *Conditions P'*:

(i) P'.1 T is a non-empty subset of the non-negative orthant of R^{M+N}

(ii) P'.2 T is closed and bounded,

(iii) P'.3 If $(q, x_k, x_{-k}, \omega) \in T$ then $(q, x'_k, x_{-k}, \omega) \in T$ $\forall x'_k \geq x_k$.

(iv) P'.4 The sets $T^K = \{v \,|\, (v, K, \omega) \in T\}$ are convex for every K given ω; the sets $T^v = \{K \,|\, (v, K, \omega) \in T\}$ are convex in K for every v given ω.

All of the results for the transformation function hold but now they are conditional on ω.

Theorem 2.2 (Transformation Function with Unobserved Inputs) *Under P'.1-P'.4 the function $F_j(q_{-j}, x, \omega)$ is an extended real-valued function defined for each $(q_{-j}, x) \geq (0_{M-1}, 0_N)$ and is non-negative on the set where it is finite. $F_j(q_{-j}, x, \omega)$ is non-decreasing in x holding q_{-j} and ω constant. Given ω, $F_j(v, K, \omega)$ is concave in v for all K quasi-concave in K for all v.*

Proof: See Appendix A.

Convexity in the elements of v conditional on any K and ω results in a production function that is concave in v holding K and ω constant. For the elements in K convexity in K given v and ω results in the production function being quasi-concave in K given v and ω. All previously discussed tests are available in this setting after conditioning on ω.

2.4 Returns to Scale

In the multi-product case where we normalize to output j and use the log-linear functional form we write:

$$lnq_{jt} = \beta_0^j + \beta_l^j lnl_t + \beta_k^j lnk_t + \beta_m^j lnm_t + \gamma_{-j}^j \prime lnq_{-jt}. \tag{1}$$

We follow Caves, Christensen, and Diewert (1982) and define returns to scale as

$$RTS(q, x, K, \omega) = \frac{\partial ln\Delta}{\partial ln\theta} \tag{2}$$

evaluated at $\theta = \Delta = 1$ such that

$$f_j(\Delta lnq_{-j}, \theta x, \theta K, \omega) - \Delta lnq_j = 0.$$

It is straightforward to show returns to scale is equal to

$$\frac{1 + \gamma_1^j + \ldots + \gamma_{M-1}^j}{\beta_l^j + \beta_k^j + \beta_m^j}.$$

We note that there are an infinite set of output tuples that can be achieved for any given increase in the input tuple. This definition of RTS is the total derivative evaluated in the current direction of outputs and inputs. The transformation function is a scalar field and thus can be evaluated in any direction of outputs and inputs. For Cobb-Douglas, trans-log, and quadratic approximations to production the total derivative exists so there exists a Taylor-series approximation in any direction of inputs and outputs at any level of input use and output production. We now turn to tests of single product production approximations to multi-product production.

3 Data

3.1 The Belgian PRODCOM survey

Statistical offices around the world are running production surveys through which they collect precise information about the products made by firms that are intended for use in industrial statistics. These datasets cover a large subset of mostly manufacturing firms and typically contain both values and quantities for each good produced by firms.

In this paper, we use the firm-product level production data based on a production survey (PRODCOM) collected by Statistics Belgium.[3] The survey is designed to cover at least 90% of production value in each NACE 4-digit industry by including all Belgium firms with a minimum of 10 employees or total revenue above 2.5 million Euros.[4] The sampled firms are required to disclose monthly product-specific revenues and quantities sold of all products at the PRODCOM 8 digit level (e.g. 11.05.10.00 for "Beer made from malt", 23.51.11.00 for "Cement clinkers" in the PRODCOM 2008-2017 classification).

Our analysis covers the entire period through which the data is available, 1996-2017. This creates two difficulties: in 2008, PRODCOM both significantly reduced its sample size to administrative costs and changed its classification system (the first 4 digits of a PRODCOM code refer to a NACE 4 digit sector and the NACE classification has been revised in 2008 implying a complete redefinition of the PRODCOM codes). In addition to that major revision, PRODCOM codes are marginally revised on a yearly basis. We therefore use annual concordance tables provided by Eurostat to follow the specific products over our sample period and use only those products with no confusion regarding the concordance.

In our empirical analysis, we perform several cleaning procedures to avoid outliers. First, we only keep firms that have their principal business activities in manufacturing as classified by NACE. First, for each 4-digit industry we compute the median ratios of total revenue over employment, capital over employment, total revenue over materials and wage bill over labor (average wage), and we exclude those observations more than five times the interquartile range below or above the median. Second, we only keep firm-product observations where the share of the product's revenue in the firm's total revenue is at least 5%. Third, we use the Value Added Tax revenue data that provides us with a separate check against the revenue numbers firms report to PRODCOM. Comparing

[3] See http://statbel.fgov.be/fr/statistiques/collecte_donnees/enquetes/prodcom/ and http://statbel.fgov.be/nl/statistieken/gegevensinzameling/enquetes/prodcom/ for more details in French and Dutch, or Eurostat in English (http://ec.europa.eu/eurostat/web/prodcom). This dataset was previously used in Bernard et al., 2019 and Amiti et al., forthcoming.

[4] NACE is a French acronym for the European Statistical Classification of Economic Activities.

the tax administrative data revenue numbers with the revenue numbers reported in the PRODCOM data, we find that between 85% and 90% of firms report similar values for both. We exclude firms if they do not report a total value of production to PRODCOM that is at least 90% of the revenue they report to the tax authorities.

Table 1: Average share of a firm's revenue derived by its individual products, 1996 to 2017

Product ranking within a firm determined by its share of the firm's total revenue.

	Number of products produced by the firm at the Prodcom 8-digit level						
	1	2	3	4	5	More than 5	N
Product rank							
1	100	78.1	69.9	64.8	60.0	50.5	
2		21.9	22.9	23.2	22.5	21.7	
3			7.2	9.0	10.6	11.6	
4				3.0	5.0	6.5	
5					1.9	3.8	
6+						5.9	
Share of manufacturing output	24.7	17.8	11.4	9.4	3.7	33.0	100
# observations	37,284	34,068	22,875	18,324	12,380	79,199	204,130

Note: For any product rank i each column j reports the average share (in %) of the i-th product in total output for firms producing j products.

We aggregate monthly revenues and quantities to the quarterly level and calculate the associated quarterly unit price. Table 1 shows the average revenue share of products in firms' portfolios when they are producing a different number of products at two levels of aggregation (8-digit and 2-digit PRODCOM). We observe 204,130 firm-product yearly observations between 1996 and 2017. As has been noted in other product-level data sets, the majority of firms produce multiple products.[5] At the 8-digit level of disaggregation, multi-product firms are responsible for 75.3% of total value of manufacturing output. Most firms produce between one and five products and these firms account for 67% of the value of manufacturing output. For firms producing two goods the core good accounts for 78.1% of revenue. Similarly for firms producing three goods 69.9% of revenue comes from the core product. Even for firms producing six or more goods the core good is responsible for 50.5% of revenue.

To test the pure Diewert-Lau framework, our analysis requires the identification of firms producing the same subset of products. For this purpose, we identified a few specific industries where firms producing two goods were the most commonly observed form

[5] See e.g. Bernard et. al (2010) or Goldberg et. al (2010).

of production. We identified six 2-product environment (combos) that fit to our requirements[6]: bread and cake; marble and other building stones; doors of plastic and doors of metal; structures of iron, steel and aluminium, and doors of metal; windows of wood, and joinery and carpentry of wood; and bricks, and prefabricated structures of cement.

Table 2 shows the product portfolio description for those 6 environments. The main message that this table conveys is that, for these 6 economic environments that we identified, the most observed form of production is when firms produce these two exact products associated to a combination, or at least this type of production pattern is a common form of production. The more obvious example is bread and cake: out of 9,621 observations, firms producing bread produce 2-products in 8,064 cases; out of these 8,064 two-product firms, 7,855 also produce cake. As we go down the list, the number of observations becomes lower and the share of single product firms also goes up.[7]

3.2 Firm Input Measurements

Quarterly measurements of firms inputs from 1997 to 2016 are obtained from the VAT fiscal declarations of firm revenue, the National Social Security database, and the Central Balance Sheet Office database. For tax liability purposes, Belgian firms have to report in their VAT fiscal declarations both their sales revenues and their purchases. Purchases are reported into three separate categories: material inputs and services directly used for production, other inputs and services used for supporting activities, and acquisition of capital goods. Using this information, we construct quarterly measures for both types of intermediate inputs and for investments. For measures of firm employment, we use data from the National Social Security declarations, where firms report on a quarterly basis their level of employment and their total wage bill. To construct a quarterly measure of capital we start with data from the Central Balance Sheet Office, which records annual measures of firm assets for all Belgian firms. For the first year a firm is in our data, we take the total fixed assets as reported in the annual account as their starting capital stock. We then use standard perpetual inventory methods to build out a capital stock for each firm-quarter.[8]

[6]See Appendix B for the full product description

[7]For one of our combinations, we realize that firms producing PRODCOM products within a 6-digit code were providing different unit of measurements. We therefore used the most common unit provided for code 222314.

[8]In order to build the capital stock, we assume a constant depreciation rate of 8% per year for all firms. Real capital stock is computed using the quarterly deflator of fixed capital gross accumulation. The initial capital stock in $t = t_0$, where period t_0 represents the 4th quarter of the first year of observation of the firm, is given by

$$K_{t_0} = \frac{Total\ fixed\ assets_{first\ year\ of\ observation}}{P_{K;t_0}}$$

Table 2: Number of observations by product and product scope, selected combinations of products

	# obs.	# obs. w/ 1 product	# obs. w/ 2 products	# obs. w/ 3 products	More than 3 products		# obs.	# obs. (w/ same unit)
107111	9,621	721	8,064	542	294	combo	7,855	7,855
107112	10,020	761	8,155	578	526	107111-107112		
237011	2,050	386	1,178	417	69	combo	1,132	1,132
237012	2,922	953	1,394	542	33	237011-237012		
222314	4,539	1,345	2,421	614	159	combo	1,766	1,065
251210	6,780	2,757	3,000	831	192	222314-251210		
251123	13,892	8,895	3,292	1,094	611	combo	892	892
251210	6,780	2,757	3,000	831	192	251123-251210		
162311	4,602	1,875	1,169	904	654	combo	595	595
162319	3,580	1,181	1,045	422	932	162311-162319		
236111	3,048	1,413	752	593	290	combo	464	464
236112	5,163	3,008	1,298	678	179	236111-236112		

4 Identification of Marginal Costs

Hall (1986, 1988) shows in the case of single-product production cost minimization identifies the markup as a function of the observed elasticities of revenue with respect to an input and the observed firm expenditures on that input. Using McFadden (1978) we derive the multivariate analog using the variable cost function to show how to express marginal costs as a function of output-input elasticities, individual input expenditures, and output quantities. We show how/when we can identify one unobserved marginal cost term for each firm-product. In this section we use x to exclusively denote the N_1 freely variable inputs and K to exclusively denote the remaining $N - N_1$ fixed inputs. For the purposes of illustration we use the Cobb-Douglas log-linear production function described earlier:

$$lnq_{jt} = \beta_0^j + \beta_l^j lnl_t + \beta_k^j lnk_t + \beta_m^j lnm_t + \gamma_{-j}^j \prime lnq_{-jt}. \tag{3}$$

Minimization of the variable cost function given the desired output vector of $q^* = (q_1^*, q_2^*, \dots, q_M^*)$ is given by

$$Min_x \;\; P * x \;\; s.t. \;\; f_j(q_{-j}^*, x, K, \omega) - q_j^* \geq 0$$

where $P = [P_1 \cdots P_{N_1}]'$ denotes the input prices for the variable inputs. The Lagrangian is

$$L = \;\; P * x - \lambda_j(f_j(q_{-j}^*, x, K, \omega) - q_j^*),$$

which yields the first-order conditions of which optimal input choice x^* is the solution:

$$P_i = \lambda_j \frac{\partial f_j(q_{-j}^*, x^*, K, \omega)}{\partial x_i} \quad i = 1, \dots, N_1,$$

The marginal cost is given by λ_j, the derivative of the total cost function with respect to q_j evaluated at the desired output level q^*. If we let β_i^j denote the elasticity of the output of good j with respect to input i and we solve for λ_j we have

$$\lambda_j = \frac{P_i}{\frac{\partial f_j(q_{-j}^*, x^*, K, \omega)}{\partial x_i}} = \frac{P_i x_i^*}{\beta_i^j * q_j^*} \quad i = 1, \dots, N_1.$$

so marginal cost is equal to the expenditure on input i divided by the output elasticity of input i times the quantity of output j that is produced. Given λ_j we can use j's

The capital stock in the subsequent periods is given by

$$K_t = (1 - 0.0194)\, K_{t-1} + \frac{I_t}{P_{K;t}}$$

We assume that the new investment is not readily available for production and that it takes one year from the time of investment for a new unit of capital to be fully operational.

transformation function estimates to recover an estimate of marginal cost for all other goods $l \neq j$ given by

$$\lambda_l = -\lambda_j * \frac{\partial f_j(q^*_{-j}, x^*, K, \omega)}{\partial q_l} \quad l \neq j$$

or

$$\lambda_l = -\lambda_j * \gamma_l \frac{q^*_l}{q^*_j}, \;\; l \neq j$$

for the Cobb-Douglas log-linear approximation. Thus we have N_1 estimates for the marginal cost for each output good, one for each freely variable input.

In the single-product case it simplifies down to

$$\lambda = \frac{P_i x^*_i}{\beta_i * q^*},$$

which is the result from Hall (1986, 1988). Letting p_{q^*} denote the price of output and multiplying this formula through by $\frac{1}{p_{q^*}}$ and inverting we have the markup given as

$$\mu = \frac{p_{q^*} q^*}{P_i x^*_i} * \beta_i$$

where $\mu = \frac{p_{q^*}}{\lambda}$. This is the approach proposed in De Loecker and Warzynski (2012) to invert out markups in standard plant-level data where only input expenditures and revenue are observed, and where one has an estimate of the elasticity of output with respect to input i. The difference between these last two expressions illustrates the value of observing quantities of outputs (or, alternatively, individual prices of outputs); without them, one can estimate markups using observed revenue shares and estimated elasticities, but one cannot separate price from marginal cost.

5 Estimation

We review the standard proxy approach in the single-product production setup and then turn to our multi-product extension.

5.1 Single-product production setting

We have for q_t:

$$q_t = \beta_l l_t + \beta_k k_t + \beta_m m_t + \omega_t + \epsilon_t \tag{4}$$

where the error has two components. ϵ_t is assumed to be i.i.d. error upon which the firm does not act (like measurement error or specification error). ω_t is the technical efficiency shock, a state variable observed by the firm but unobserved to the econometrician. ω_t

is assumed to be first-order Markov and is the source of the simultaneity problem as firm observe their shock before choosing their freely variable inputs l_t and m_t. k_t also responds to ω_t but with a lag as investments made in period $t-1$ come online in period t. This assumption allows k_t to be correlated with expected value of ω_t given ω_{t-1} - denoted $E[\omega_t|\omega_{t-1}]$ - but maintains that the innovation in the productivity shock, $\xi_t = \omega_t - E[\omega_t|\omega_{t-1}]$, is unknown at the time the investment decision was made in $t-1$ and is therefore uncorrelated with current k_t.

The control function approaches of Olley and Pakes (1996) and Levinsohn and Petrin (2003) both provide conditions uner which there exists a proxy variable $h(k_t, \omega_t)$ that is a function of both state variables and that is monotonic in ω_t given k_t. The variables may include either investment (OP) or materials, fuels, electricity, or services (LP) (e.g.). Given the monotonicity there exists some function $g(\cdot)$,

$$\omega_t = g(k_t, h_t)$$

allowing ω_t to be written as a function of k_t and h_t. Kim, Petrin, and Song (2016) extend Hu and Schennach (2008) to allow for measurement error in all of the variables in the proxy function.

For estimation Wooldridge (2009) uses a single index restriction to approximate unobserved productivity, writing

$$\omega_t = g(k_t, h_t) = \mathbf{c}(k_t, h_t)'\beta_\omega$$

where $\mathbf{c}(k_t, h_t)$ is a known vector function of (k_t, h_t) chosen by researchers with parameter vector β_ω to be estimated. The conditional expectation $E[\omega_t|\omega_{t-1}]$ can then be written as

$$E[\omega_t|\omega_{t-1}] = f(\mathbf{c}(k_{t-1}, h_{t-1})'\beta_\omega)$$

for some unknown function $f(\cdot)$, which Wooldridge (2009) approximates using a polynomial.

Replacing ω_t with its expectation and innovation, the estimating equation becomes

$$q_t = \beta_l l_t + \beta_k k_t + \beta_m m_t + E[\omega_t|\omega_{,t-1}] + \xi_t + \epsilon_t \tag{5}$$

For expositional transparency we use only the first-order approximation term for $f(\cdot)$, which yields our error term

$$[\xi_t + \epsilon_t](\theta) = q_t - \beta_l l_t - \beta_k k_t - \beta_m m_t - \mathbf{c}(h_{t-1}, k_{t-1})'\beta_\omega \tag{6}$$

with the parameters to $\beta = (\beta_l, \beta_k, \beta_m, \beta_\omega)$.

Let θ_0 denote the true parameter value. Wooldridge shows that the conditional moment restriction

$$s(x_t;\theta) \;\; \equiv \;\; E[[\xi_t + \epsilon_t](\theta)|x_t] \text{ and } s(x_t;\theta_0) = 0$$

is sufficient for identification of β in the single product case (up to a rank condition on the instruments).In equation (4) a function of m_{t-1} and k_{t-1} conditions out $E[\omega_t|\omega_{t-1}]$. ξ_t is not correlated with k_t, so k_t can serve as an instrument for itself. Lagged labor l_{t-1} and twice lagged materials m_{t-2} serve as instruments for l_t and m_t.

5.2 Multivariate Control Functions

We denote the vector of technical efficiency shocks as $\omega_t = (\omega_{1t}, \omega_{2t}, \ldots, \omega_{M_t})$ and assume $E[\omega_t|\omega_{t-1}] = \omega_{t-1}$. Choices of inputs will now generally be based not only on ω_{jt} but also on all of the other technical efficiency shocks ω_{-jt}. This frustrates the "inverting out" of ω_t that allows one to express ω_t as a function of k_t and a single proxy h_t as is done in the single product case.

We extend suggestions from Petropoulos (2001) and Ackerberg, Benkard, Berry, and Pakes (2007) to allow for these multiple unobserved technical efficiency shocks. Suppose we observe (at least) one proxy variable for every technical efficiency shock. Let $\mathbf{h}_t = (h_{1t}, \ldots, h_{Lt})$ denote the $1XL$ vector of available proxies. Each of these variables will generally be a function of k_t and $(\omega_{1t}, \omega_{2t}, \ldots, \omega_{M_t})$ and we write the vector of proxies as $\mathbf{h}_t(k_t, \omega_t)$. Conditional on k_t, if $\mathbf{h}_t(k_t, \omega_t)$ is one-to-one with ω_t then we can invert the proxy variables to get the $1XL$ vector of functions $\omega_t = \mathbf{g}(k_t, \mathbf{h}_t)$. Included in this vector of functions is

$$\omega_{jt} = g_j(k_t, \mathbf{h}_t), \qquad j = 1 \cdots M$$

which then motivates including a function of (k_t, h_t) in the estimation to control for ω_{jt}.

Using the first-order conditions from cost minimization we have the following expression when evaluated at the optimal x:

$$\frac{\partial^2 q_j(x^*, K, \omega)}{\partial x\,\partial x'} * \frac{\partial x^*}{\partial \omega'} + \frac{\partial^2 q_j(x^*, K, \omega)}{\partial x\;\partial \omega'} = 0.$$

Using the implicit function theorem we see invertibility of $\frac{\partial^2 q_j(x^*,K,\omega)}{\partial x\,\partial x'}$ and full rank of $\frac{\partial^2 q_j(x^*,K,\omega)}{\partial x\;\partial \omega'}$ are the conditions required to solve for $\frac{\partial x^*}{\partial \omega'}$ as

$$\frac{\partial x^*}{\partial \omega'} = \Big(\frac{\partial^2 q_j(x^*, K, \omega)}{\partial x\,\partial x'}\Big)^{-1} * \frac{\partial^2 q_j(x^*, K, \omega)}{\partial x\;\partial \omega'}.$$

5.3 Multi-product production setting

The estimation routine proceeds in a manner similar to the single-product case. We use the same single index restriction to control for the unobserved productivity terms, with one possible for each product, so we have

$$\omega_{jt} = g_j(k_t, \mathbf{h}_t) = \mathbf{c}_j(k_t, \mathbf{h}_t)'\beta_{\omega_j}$$

where $\mathbf{c}_j(k_t, \mathbf{h}_t)$ is a known vector function of $(k_t, \mathbf{h}_t)$ chosen by researchers. $E[\omega_{jt}|\omega_{t-1}]$ is now given as

$$E[\omega_{jt}|\omega_{t-1}] = f_j(\mathbf{c}_j(k_{t-1}, \mathbf{h}_{t-1})'\beta_{\omega_j})$$

for some unknown function $f_j(\cdot)$. Again we use only the first-order approximation term for $f_j(\cdot)$ to keep exposition to a minimum.

Re-expressing in terms of firm's expectations we have

$$q_{jt} = \beta_l^j l_t + \beta_k^j k_t + \beta_m^j m_t + \gamma_{-j}^j{}' q_{-jt} + E[\omega_{jt}|\omega_{t-1}] + \xi_{jt} + \epsilon_{jt} \qquad (7)$$

with $\xi_{jt} = \omega_{jt} - E[\omega_{jt}|\omega_{t-1}]$. The error is

$$[\xi_{jt} + \epsilon_{jt}](\theta) = q_{jt} - \beta_l^j l_t - \beta_k^j k_t - \beta_m^j m_t - \gamma_{-j}^j{}' q_{-jt} - \mathbf{c}_j(k_{t-1}, \mathbf{h}_{t-1})'\beta_{\omega_j}$$

with the new parameters γ_{-j}^j added to $\beta^j = (\beta_l^j, \beta_k^j, \beta_m^j, \gamma_{-j}^j, \beta_{\omega_j})$.

An additional key difference from the single product case is the need for instruments for q_{-jt}, which might either be lagged values of q_{-jt} or inputs lagged even further back. Let θ_0 denote the true parameter value. The conditional moment restriction

$$s(x_{jt};\theta) \equiv E[[\xi_{jt} + \epsilon_{jt}](\theta)|x_{jt}] \text{ and } s(x_{jt};\theta_0) = 0$$

continues to be sufficient for identification of β as long as a rank condition holds.

6 Results

6.1 6-digit analysis

Table 3 shows the coefficients of our generalized transformation function for the 6 selected combinations of two goods. The top panel shows the results when the log of quantity of the first good (bread in the first example) is considered as left hand side variable and regressed on aggregate firm-level inputs and the log of quantity of the second good (cake in the first column). The bottom panel shows a similar regressions when log of quantity of the second good (cake in column 1') is regressed on inputs and the log of quantity of the first good (bread in column 1').[9]

The first row of each panel shows the coefficient of the log of production of the other good conditional on input use. The coefficient is always negative and highly significant suggesting quantities are substitutes for one another holding inputs constant.

Consistent with multi-product production theory, all but five of the forty-eight input coefficients are positive, and thirty eight of these forty-three positive input coefficient estimates are strongly statistically significant. These findings are based on relatively small sample sizes for each industry which, except for Bread and Cakes, ranges from between 255 to 996 observations.

Table 4 reports the estimated returns to scale using estimates from Table 3. We calculate returns to scale for both quantity normalizations leading to 12 estimates of returns to scale. 11 of these 12 estimate are between 0.93 and 1.11. All four estimates from the two industries with the most observations are almost exactly equal to one and do not vary across the different quantity normalizations.

6.2 4-digit analysis

We next try to adopt several aggregation strategies to consider different product markets and possibly increase our sample size. Our first approach aggregates physical output within a 4-digit PRODCOM code for firms operating in two 4-digit environments. We estimate this framework for a subset of firms in the furniture industry. Table 5 shows the results are largely consistent with the findings from the 6-digit industries with quantities being substitutes and input coefficients being positive. Columns 1 and 3 and columns 2 and 4 are the different normalizations of the same transformation function and they yield similar returns to scale.

6.3 Robustness with French data

We replicate the analysis using a sample of French firms (see Smeets and Warzynski (2019) for more information about the dataset). Data are collected annually for the period 2009-2017. They record only one aggregate measure of materials but otherwise the estimation is similar to the Belgian case. Table 6 shows the results at the 6-digit level and Table 7 at the 4-digit level. The number of observations is a bit lower because of the relatively shorter panel and the annual dimension of the data, but results again are largely consistent with our findings from the Belgian data.

[9] See Dhyne, Petrin and Warzynski (2016) for a joint estimation of production function, demand function and cost function for the Belgian bread and cake industry.

Table 3: Multi-product production function estimates at 6-digit Prodcom level, Belgian data

Dependent variable q_{ijt} is log of the quantity sold in physical units at the 6-digit product level of good j by firm i at time t

	(1)	(2)	(3)	(4)	(5)	(6)
	107111-107112	237011-237012	222314-251210	251123-251210	162311-162319	236111-236112
$q_{(-j)}$	-0.374***	-0.629***	-0.769***	-0.417***	-0.303***	-0.334***
	(0.010)	(0.048)	(0.032)	(0.058)	(0.041)	(0.038)
l	0.405***	0.442***	0.364***	0.267***	0.650***	-0.071
	(0.017)	(0.066)	(0.068)	(0.074)	(0.043)	(0.091)
k	0.101***	0.163***	0.070***	0.286***	0.606***	0.722***
	(0.009)	(0.050)	(0.022)	(0.032)	(0.075)	(0.064)
$m1$	0.602***	0.979***	1.216***	0.571***	0.404***	0.655***
	(0.016)	(0.060)	(0.062)	(0.085)	(0.058)	(0.049)
$m2$	0.305***	0.042	0.350***	0.219***	-0.240***	0.108
	(0.012)	(0.066)	(0.036)	(0.041)	(0.054)	(0.084)
	(1')	(2')	(3')	(4')	(5')	(6')
	107112-107111	237012-237011	251210-222314	251210-251123	162319-162311	236112-236111
$q_{(-j)}$	-0.555***	-0.390***	-0.720***	-0.687***	-1.307***	-0.963***
	(0.015)	(0.033)	(0.030)	(0.102)	(0.161)	(0.108)
l	0.547***	0.276***	0.365***	-0.037	0.741***	0.342**
	(0.019)	(0.055)	(0.066)	(0.098)	(0.146)	(0.141)
k	0.145***	0.399***	0.005	0.068	0.841***	1.027***
	(0.010)	(0.037)	(0.022)	(0.054)	(0.185)	(0.108)
$m1$	0.721***	0.955***	1.187***	1.185***	1.077***	0.682***
	(0.019)	(0.040)	(0.060)	(0.061)	(0.095)	(0.104)
$m2$	0.163***	-0.234***	0.276***	0.108*	-0.459***	0.142
	(0.015)	(0.052)	(0.037)	(0.057)	(0.110)	(0.131)
			Correlation between ω_1 and ω_2			
	0.81	0.89	0.92	0.84	0.78	0.95
# obs.	7,262	996	895	596	255	360

Note: Each column reports the estimated coefficients using a modified variant of the GMM Wooldrige estimator. Explanatory variables are in logs and include firm-level labor, the standard real indices for materials and for capital - i.e. the dollar value of each - and the physical quantity of the other good produced by the firm. We include the product's price as an additional control. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Multi-product production function estimates at 6-digit Prodcom level, Belgian data

Implied Returns to Scale
6-digit Prodcom level, Belgian data

	(1) 107111-107112	(2) 237011-237012	(3) 222314-251210	(4) 251123-251210	(5) 162311-162319	(6) 236111-236112
ReturnsToScale	1.022	0.994	1.126	0.936	1.089	1.059
	(1') 107112-107111	(2') 237012-237011	(3') 251210-222314	(4') 251210-251123	(5') 162319-162311	(6') 236112-236111
ReturnsToScale	1.007	0.993	1.052	0.780	0.956	1.112
# obs.	7,262	996	895	596	255	360

6.4 Marginal costs and markups

Using the theory from section 4, we can derive estimates of marginal costs. We apply the formulas on our subsample of bread and cake producers where the number of observations is the largest. All marginal cost estimates are positive.

Table 5: Multi-product production function estimates at 4-digit Prodcom level, Belgian data

Dependent variable q_{ijt} is log of the quantity sold in physical units at the 4-digit product level of good j by firm i at time t

	(1) 3100-3109	(2) 3102-3109	(1') 3109-3100	(2') 3109-3102
$q_{(-j)}$	-0.261***	-0.262***	-0.732***	-0.797***
	(0.021)	(0.027)	(0.061)	(0.075)
l	0.416***	0.652***	0.638***	1.693**
	(0.049)	(0.117)	(0.088)	(0.167)
k	0.156	0.152	0.860***	-0.134
	(0.200)	(0.262)	(0.352)	(0.407)
m	0.910***	0.447***	0.631***	0.125
	(0.048)	(0.090)	(0.103)	(0.144)
RTS	1.166	0.984	1.225	0.938
# obs.	1,205	885	1,205	885

Note: Each column reports the estimated coefficients using a modified variant of the GMM Wooldrige estimator. Explanatory variables are in logs and include firm-level labor, the standard real indices for materials and for capital - i.e. the dollar value of each - and the physical quantity of the other good produced by the firm We include the product's price as an additional control. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Multi-product production function estimates at 6-digit Prodcom level, French data

Dependent variable q_{ijt} is log of the quantity sold in physical units at the 6-digit product level of good j by firm i at time t

	(1)	(2)	(3)	(4)	(5)	(6)
	181212-181219	162311-162319	310912-310913	237011-237012	310210-310913	251123-251210
$q_{(-j)}$	-0.652***	-0.301***	-0.074	-0.561**	-0.212***	-0.464***
	(0.110)	(0.049)	(0.048)	(0.241)	(0.047)	(0.079)
l	0.340***	0.714***	0.311***	0.633***	0.235***	0.312***
	(0.077)	(0.117)	(0.106)	(0.190)	(0.084)	(0.154)
k	0.073**	0.142**	0.045	0.052	0.298***	0.334***
	(0.036)	(0.052)	(0.052)	(0.062)	(0.056)	(0.084)
m	1.200***	0.464***	0.652***	0.856***	0.696***	0.838***
	(0.094)	(0.114)	(0.093)	(0.231)	(0.071)	(0.115)
	(1')	(2')	(3')	(4')	(5')	(6')
	181219-181212	162319-162311	310913-310912	237012-237011	310913-310210	251210-251123
$q_{(-j)}$	-0.541***	-0.601***	-0.224*	-0.410***	-0.532***	-0.399***
	(0.081)	(0.087)	(0.086)	(0.087)	(0.130)	(0.071)
l	0.334***	0.585***	0.011	0.547***	0.076	0.066
	(0.074)	(0.172)	(0.169)	(0.191)	(0.161)	(0.147)
k	-0.009	0.186**	0.118	0.113	0.145**	0.230***
	(0.034)	(0.074)	(0.077)	(0.083)	(0.074)	(0.070)
m	1.092***	0.513***	0.986***	0.818***	1.148***	1.098***
	(0.075)	(0.150)	(0.141)	(0.170)	(0.122)	(0.123)
# obs.	569	349	380	359	334	312

Note: Each column reports the estimated coefficients using a modified variant of the GMM Wooldrige estimator. Explanatory variables are in logs and include firm-level labor, the standard real indices for materials and for capital - i.e. the dollar value of each - and the physical quantity of the other good produced by the firm We include the product's price as an additional control. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Multi-product production function estimates at 4-digit Prodcom level, French data

Dependent variable q_{ijt} is log of the quantity sold in physical units at the 4-digit product level of good j by firm i at time t

	(1)	(2)	(1')	(2')
	3102-3109	3101-3109	3109-3102	3109-3101
$q_{(-j)}$	-0.153***	-0.297***	-0.239***	-0.492***
	(0.024)	(0.043)	(0.061)	(0.094)
l	0.244***	0.285**	0.350***	0.393**
	(0.059)	(0.129)	(0.076)	(0.157)
k	0.192***	0.160**	0.112*	0.322***
	(0.033)	(0.063)	(0.045)	(0.101)
m	0.750***	0.726***	0.702***	1.021***
	(0.049)	(0.146)	(0.084)	(0.183)
# obs.	1,078	342	1,079	342

Note: Each column reports the estimated coefficients using a modified variant of the GMM Wooldrige estimator. Explanatory variables are in logs and include firm-level labor, the standard real indices for materials and for capital - i.e. the dollar value of each - and the physical quantity of the other good produced by the firm We include the product's price as an additional control. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Summary statistics on marginal costs estimates and prices. Bread and cake producers, Belgium

	Marginal cost		**Price**		**Markup**	
	Bread	Cake	Bread	Cake	Bread	Cake
mean	1.616	3.004	1.474	5.032	1.099	2.128
10%	0.755	1.064	1.103	3.15	0.518	0.808
25%	0.968	1.459	1.149	3.431	0.731	1.255
50%	1.377	2.295	1.325	3.802	1.051	1.976
75%	1.985	3.791	1.715	5.127	1.411	2.825
90%	2.817	6.156	1.971	8.894	1.746	3.653
std dev	0.90	2.153	0.422	3.109	0.471	1.104

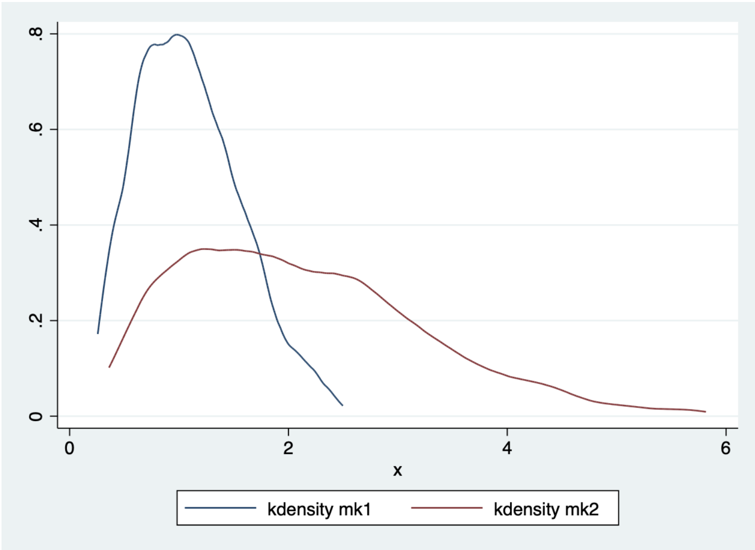
We then use our marginal cost estimates to compute markups. We find an average markup of 1.09 for bread and 2.12 for cake. The distribution of markups for bread and cake is shown in figure 1. We observe for both products a concentration around one, and a fat tail on the right of 1.

We then correlate our marginal costs and markup measures to the productivity for both bread and cake. Price and marginal cost are negatively correlated with technical efficiency and markups are positively correlated with technical efficiency. These results are in line with previous research using product level information and estimation markups and productivity (see e.g. Foster, Haltiwanger and Syverson, 2008 and De Loecker et al., 2016).

Table 9: Relationship between prices, marginal costs, markups and TFPQ

		Bread	
	logp	*logMC*	Markup
TFPQ	-0.303***	-0.886***	0.217***
	(0.004)	(0.014)	(0.008)
R^2	0.36	0.34	0.08
# obs.	7,726	7,726	7,726
		Cake	
	logp	*logMC*	Markup
TFPQ	-0.378***	-1.182***	0.198***
	(0.005)	(0.020)	(0.015)
Adj. R^2	0.38	0.31	0.02
# obs.	7,536	7,536	7,536

Figure 1: Distribution of markups for bread (mk_1) and cake (mk_2)



6.5 Translog specification

In this subsection, we show the results when we instead adopt a trans-log specification.

Table 10 shows the average and median elasticities for the bread and cake producers. These figures are comparable to those with the Cobb Douglas but allow the elasticities to be firm-product specific.

Table 11 shows the computed marginal costs and markups using the estimates from the trans-log specification. Average and median marginal costs are similar to those of the CD specification.

Table 10: Multi-product production function estimation with translog specification: median and mean elasticities

	mean	median
$q_{(-j)}$	-0.399	-0.404
l	0.375	0.358
k	0.112	0.112
$m1$	0.738	0.773
$m2$	0.238	0.219
# obs.	6,930	

Note: xxx

Table 11: Summary statistics on marginal costs and markups using translog estimates and elasticities. Bread and cake producers, Belgium

	Marginal cost		**Markup**	
	Bread	Cake	Bread	Cake
mean	1.362	2.665	1.452	2.543
10%	0.570	0.970	0.580	0.837
25%	0.743	1.312	0.912	1.351
50%	1.087	2.014	1.344	2.281
75%	1.644	3.424	1.920	3.422
90%	2.568	5.412	2.437	4.566
std dev	0.894	1.899	0.713	1.526

7 Conclusion

We introduce a new methodology for estimating multi-product production functions. It embeds the seminal contributions of Diewert (1973) and Lau (1976) in the semi-parametric econometric framework following Olley and Pakes (1996). We address the simultaneity of inputs and outputs while allowing for one unobserved technical efficiency term for each firm-product, each one of which may be freely correlated with inputs and outputs. We show how to use these estimates to recover estimates of firm-product marginal costs by extending the Hall (1988) single-product result to our multi-product setting. The main advantage of our framework is that it does not require multi-product production to be a collection of single-product production functions. Our empirical results using panel multi-production production data are largely consistent with our derived necessary conditions from the theory.

References

[1] Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. Econometric Tools for Analyzing Market Outcomes. Chapter 63 in Handbook of Econometrics: Volume 6A, Edited by J. Heckman and E. Leamer, North Holland 2007.

[2] Bernard, A., E. Blanchard, I. Van Beveren, and H. Vandenbussche, 2012. Multi-Product Exporters, Carry-Along Trade and the Margins of Trade. NBER Working Paper #18246.

[3] Bernard, A., S. Redding, and P. Schott, 2010. Multiple-Product Firms and Product Switching. American Economic Review 100, 70-97.

[4] Bernard, A., S. Redding, and P. Schott, 2011. Multi-product Firms and Trade Liberalization. Quarterly Journal of Economics 126, 1271â€“1318.

[5] De Loecker, J, and Warzynski, F. 2012. "Markups and Firm-Level Export Status." American Economic Review, 102 (6): 2437-71.

[6] De Loecker, J., 2011. Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity. Econometrica 79, 1407-1451.

[7] De Loecker, J., 2013. Detecting Learning by Exporting. American Economic Journal: Microeconomics 5, 1-21.

[8] De Loecker, J., P. Goldberg, A. Khandelwal, and N. Pavcnik, 2016. Prices, markups and Trade Reform. Econometrica 84, 445-510.

[9] Dhyne, E., A. Petrin, and F. Warzynski, 2016. Deregulation and Spillovers in Multi-Product Production Settings. Mimeo, Aarhus University.

[10] Diewert, W. 1973. Functional Forms for Profit and Transformation Functions. Journal of Economic Theory 6, pp. 284-316.

[11] Duprez, C. 2014. Value Creation in Exports: A Diagnostic for Belgium. NBB Economic Review, September 2014, pp. 25-39.

[12] Fare, Martins-Filho, and Vardanyan, 2009. On functional form representation of multi-output production technologies. Journal of Productivity Analysis.

[13] Foster, L., J. Haltiwanger, and C. Syverson, 2008. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? American Economic Review 98, 394-425.

[14] Garcia Marin, A. and Voigtlander, N., (2013). Exporting and plant-level efficiency gains: It's in the measure. NBER Working Paper #19033.

[15] Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova, 2010a. Imported Intermediate Input and Domestic Product Growth: Evidence from India. Quarterly Journal of Economics 126, 1727-1767.

[16] Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova, 2010b. Multiproduct Firms and Product Turnover in the Developing World: Evidence from India. Review of Economics and Statistics 92, 1042-1049.

[17] Gong and Sickles, 2018. Resource Allocation in the Multi-divisional Multi-product Firms. Rice University Working Paper.

[18] Griliches and Mairesse, 1995. Production Functions: The Search for Identification. NBER working paper # 5067.

[19] Hall, R.E., 1986. Market Structure and Macroeconomic Fluctuations. Brookings Papers on Economic Activity 2, 285-322.

[20] Hall, R. E., 1988. The Relation between Price and Marginal Cost in U.S. Industry. Journal of Political Economy 96, 921-947.

[21] Hu and Schennach (2008). Instrumental Variable Treatment of Nonclassical Measurement Error Models. Econometrica. Vol. 76, Issue 1, pp. 195-216.

[22] Itoga, 2019. Within-Firm Reallocation and the Impacts of Trade under Factor Market Imperfection. Working Paper.

[23] Kim, Petrin, and Song, 2016. Estimating Production Functions with Control Functions When Capital Is Measured With Error. Journal of Econometrics, 190 (2), February 2016, pp. 267-279.

[24] Klette, T. J. and Z. Griliches, 1996. The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous. Journal of Applied Econometrics 11, 343-361.

[25] Kuosmanen, T. and Johnson, A., 2019. Modeling joint production of multiple outputs in StoNED: Directional distance function approach. Aalto University Working Paper.

[26] Lau, L., 1976. A Characterization of the Normalized Restricted Profit Function. Journal of Economic Theory 12, 131-163.

[27] Levinsohn, J., and A. Petrin, 2003. Estimating Production Functions Using Inputs to Control for Unobservables. Review of Economic Studies 70, 317-341.

[28] Mayer, T., M. Melitz, and G. Ottaviano, 2014. Market Size, Competition, and the Product Mix of Exporters. American Economic Review 104, 495-536.

[29] Mayer, T., M. Melitz, and G. Ottaviano, 2016. Product Mix and Firm Productivity Responses to Trade Competition. Work in progress.

[30] McFadden, D., "Cost, Revenue, and Profit Functions," in Duality of Production, Cost, and Profit Functions, 1978,

[31] Olley, S., and A. Pakes, 1996. The Dynamics of Productivity in the Telecommunications Industry. Econometrica 64, 1263-1297.

[32] Orr, S., 2019. Within-Firm Productivity Dispersion: Estimates and Implications. University of British Columbia Working Paper.

[33] Pavcnik, N., 2002. Estimating Production Functions Using Inputs to Control for Unobservables. Review of Economic Studies 70, 317-341.

[34] Petropoulos, W, 2000. Industry Productivity Dynamics with Capacity Utilization and Differentiated Products, University of Michigan Ph.D.

[35] Roberts, M. and D. Supina, 2000. Output price and markup dispersion in micro data: the roles of producer heterogeneity and noise. Industrial Organization 9, 1-36.

[36] Smeets, V. and Warzynski, F., 2019. Multi-Product Firms in French Manufacturing and the Rise of Subcontracting. Mimeo, Aarhus University.

[37] Valmari, N., 2016. Estimating Production Functions of Multi-product Firms. ETLA Working Paper No. 37, March, 2016.

[38] Wooldridge, J. M., 2009. On estimating firm-level production functions using proxy variables to control for unobservables. Economics Letters 104, 112-114.

Appendix

The first result is from Diewert (1973) and the last two results are from Lau (1976). The results below when we classify outputs and inputs into flexible (a ν) ones and fixed ones (a K).

Proof of Theorem 2.1

Under P.1-P.4 $F_j(q_{-j}, x)$ is

(1) non-decreasing in x

Let $(q_{-g}, x) \geq (0_{M-1}, 0_N)$ and suppose $q_g^* = F(q_{-g}, x)$ is finite. Then $(q_g^*, q_{-g}, x') \in T \ \forall x' \geq x$ by free disposal. But $F(q_{-g}, x') \geq q_q^* = F(q_{-g}, x)$.

Under P.1-P.4 $F_j(v, K)$ is

(2a) concave in v ∀ K

Suppose $q_j^* = F(v_j, K) \quad j = 1, 2$. Then $(q_j^*, v_j, K) \in T \quad j = 1, 2$.
By convexity $(\lambda\, q_1^* + (1-\lambda)q_2^*, \lambda\, v_1 + (1-\lambda)v_2, K) \in T \ \ 0 < \lambda < 1$.

$$\begin{aligned} Then\ q^* &= F(\lambda\, v_1 + (1-\lambda)v_2, K) \\ &= max(q \,|\, (q, \lambda\, v_1 + (1-\lambda)v_2, K) \in T) \\ &\geq \lambda\, q_1^* + (1-\lambda)q_2^* \\ &= \lambda\, F(v_1, K) + (1-\lambda)F(v_2, K) \end{aligned}$$

(2b) quasi-concave in K ∀ v

Suppose $q_j^* \equiv F(v, K_j)\ for\, j = 1, 2$. Then $(q_j^*, v, K_j) \in T\ for\, j = 1, 2$. Let $\tilde{q} = min(q_1^*, q_2^*)$. Then $(\tilde{q}, v, K_j) \in T\ for\, j = 1, 2$. Then convexity of T in K $\forall v$ implies $(\tilde{q}, v, \lambda K_1 + (1-\lambda)K_2) \in T\ for 0 < \lambda < 1$ With $K_\lambda \equiv \lambda K_1 + (1-\lambda)K_2$ we have

$$q_\lambda = F(v, K_\lambda) \geq \tilde{q} = min(F(v, K_1)\,,\, F(v, K_2).$$

Theorem 2.2 follows immediately after conditioning on ω.

Appendix B: choice of 6-digit combinations

Specification 1

10.71.11 Fresh bread

10.71.12 Fresh pastry goods and cakes

Specification 2

23.70.11 Marble, travertine, alabaster, worked, and articles thereof (except setts, curbstones, flagstones, tiles, cubes and similar articles); artificially coloured granules, chippings and powder of marble, travertine and alabaster

23.70.12 Other worked ornamental or building stone and articles thereof; other artificially coloured granules and powder of natural stone; articles of agglomerated slate

Specification 3

22.23.14 Doors, windows and frames and thresholds for doors; shutters, blinds and similar articles and parts thereof, of plastics

25.12.10 Doors, windows and their frames and thresholds for doors, of metal

Specification 4

25.11.23 Other structures and parts of structures, plates, rods, angles, shapes and the like, of iron, steel or aluminium

25.12.10 Doors, windows and their frames and thresholds for doors, of metal

Specification 5

16.23.11 Windows, French windows and their frames, doors and their frames and thresholds, of wood

16.23.19 Builders' joinery and carpentry, of wood, n.e.c.

Specification 6

23.61.11 Tiles, flagstones, bricks and similar articles, of cement, concrete or artificial stone

23.61.12 Prefabricated structural components for building or civil engineering, of cement, concrete or artificial stone