Production Function Estimation with Multi-Destination Firms

Geoffrey Barrows*, Hélène Ollivier†, Ariell Reshef‡

April 10, 2024

Abstract

We develop a production function estimator for the case when firms endogenously select into multiple destination markets where they compete imperfectly, and when researchers observe output denominated only in value. We show that ignoring the multi-destination dimension (i.e., exporting) yields biased and inconsistent inference, leading to unrealistic inference in the data. In contrast, our estimator is consistent and performs well in finite samples. In French manufacturing data, our estimator recovers increasing total returns to scale, decreasing returns to flexible inputs, elasticities of demand between -21.5 and -3.4, and learning-by-exporting effects between 0 and 4% per year.

Keywords: production function, learning by exporting, trade, productivity
JEL Classification: F12, F63, D24

*CNRS, CREST, Ecole Polytechnique. Address: Ecole polytechnique, Departement d'Economie, Route de Saclay, 91128 Palaiseau, France. Email: geoffrey-masters.barrows@polytechnique.edu.

†Paris School of Economics - CNRS, Address: Paris School of Economics, 48 Boulevard Jourdan 74014 Paris, France. Email: helene.ollivier@psemail.eu.

‡Paris School of Economics - CNRS, Address: Paris School of Economics, 48 Boulevard Jourdan 74014 Paris, France. Email: ariell.reshef@psemail.eu.

1 Introduction

Production function estimation is a central component of many economic analyses.[1] While early work relied on restrictive assumptions with respect to the evolution of unobserved shocks to supply, remarkable progress has been made in the last 30 years towards specifying flexible conditions under which structural elements of supply can be identified from firm or plant-level data (Olley & Pakes, 1996; Blundell & Bond, 2000; Levinsohn & Petrin, 2003; Wooldridge, 2009; Ackerberg et al., 2015; Gandhi et al., 2020). Nevertheless, significant gaps remain between the conditions assumed by the literature and the real-world datasets confronted by practitioners (De Loecker & Goldberg, 2014; De Loecker & Syverson, 2021).

One fundamental problem emphasized in the literature is that since output is usually observed in monetary terms—i.e., sales, not physical quantities—the pricing decisions of firms directly influence the outcome variable (sales). In this case, unobserved shocks to demand bias the estimation of output elasticities, even if unobserved shocks to supply are adequately controlled for (Klette & Griliches, 1996; Foster et al., 2008; De Loecker & Goldberg, 2014; De Loecker & Syverson, 2021). The standard practice of deflating firm-level revenues by industry-wide price indices only addresses the problem if all firms within an industry sell at the same price;[2] and the oft-cited workaround from Klette & Griliches (1996) (KG hereafter), where the inverse demand function is substituted for missing price data, delivers consistent estimates only in the case that all firms sell to a *single market*.[3] In reality, many firms serve multiple destinations and charge firm-destination specific prices,

[1] Researchers estimate production functions for a wide array of purposes. The structural coefficients can be themselves of interest, as in studies of returns to scale (Caballero & Lyons, 1992) or can be used to estimate markups (Hall, 1986) or wedges (Hsieh & Klenow, 2009). Alternatively, researchers may aim to control for estimated productivity in order to address omitted variable bias (Almunia et al., 2021), or may estimate productivity as the residual of a production function and then regress this residual on other explanatory variables in order to study the determinants of productivity (Harrigan et al., 2023). For a recent review of the literature, see De Loecker & Syverson (2021).

[2] A sufficient condition for a unique price is homogeneous output. However, this approach is often used to analyze markets for goods that are clearly not homogeneous. This is of particular concern when researchers use estimated output elasticities to compute firm-specific markups. In a market with homogeneous output and free entry (or perfect competition), one would expect marginal cost pricing, i.e., no markups. In this case, the object of inquiry (variable markups) is inconsistent with the assumptions of the model, as noted by Bond et al. (2021).

[3] Klette & Griliches (1996) is frequently cited as the work-around to unobserved firm-specific output prices; for example, see Melitz & Levinsohn (2006), De Loecker (2011) and Grieco et al. (2016). Melitz & Levinsohn (2006) and Grieco et al. (2016) assume that firms sell to a single market. De Loecker (2011) allows that firms sell to different "segments" of the market, where a segment is defined as a set of products, but De Loecker (2011) implicitly assumes that the representative consumer for each segment exists in a single destination. We refer to this model as a "single-market" model due to the single destination, even though there are different product segments.

which implies that the moment conditions employed by existing "revenue production function" estimators are unlikely to hold.[4] As we show, these export market features have important implications for estimation, even when researchers are not explicitly interested in exporting or the effect of exporting *per se*. To date, the literature has overlooked this problem.

In this paper, we develop a procedure to estimate structural elements of supply and demand from firm-level data when firms may serve multiple destination markets, wherein they face firm-market-time specific demand conditions, and when outputs are only observed in monetary terms. Exploiting nonparametric identification results from Gandhi et al. (2020), we estimate the output elasticities of flexible and quasi-fixed inputs without imposing any assumptions on the production function beyond multiplicative separability between the (unobserved) productivity shifter and the rest of the production function. The flexible form of the production function allows for non-constant marginal cost functions, which has important implications for optimization and market entry. On the demand side, we allow for firm-destination-year taste shocks, but impose a common industry-wide elasticity of substitution between varieties.[5] Importantly, the common elasticity does not imply constant markups, as we explain below. The elasticity of substitution is identified from variation in firm-year export revenue shares, information that is available in most firm-level balance sheets. Finally, we identify the elasticity of productivity to observable characteristics, allowing these determinants to endogenously affect the evolution of productivity.

We start by specifying a data generating process in which heterogeneous firms endogenously select destination markets to serve, destination-specific quantities and prices, and quantities of flexible inputs to hire, each period. We make the standard assumptions that firms must pay both fixed and variable trade costs to ship goods across borders, and that firms neglect the effect of their own pricing decision on market-wide price indices (i.e., monopolistic competition). Non-constant marginal cost functions together with fixed market entry costs imply that optimization requires firms to solve a combinatorial discrete choice problem in which market entry decisions and destination-specific quantities and

[4] Simply restricting the sample to non-exporters is insufficient to address the problem for two reasons. First, restricting the sample introduces sample selection bias. Second, in a sample of non-exporters, identification of the demand elasticity is based purely on time series variation in the domestic aggregate demand shifter, which is insufficient in most panels to deliver precise estimates. We return to these points in Sections 3 and 5.

[5] Most estimates of "trade elasticities" vary only by product or industry, regardless of the destination. See, for example, Broda & Weinstein (2006), Kee et al. (2008), Fontagné et al. (2022), Shapiro (2016) and Boehm et al. (2023).

prices are chosen jointly and simultaneously. Given the combinatorial structure of the optimization, demand conditions in any given market can affect the sales in all other markets both through the extensive margin of market entry, and through the intensive margin (i.e., conditional on the same set of destinations), as in Almunia et al. (2021).[6]

On the demand side, we assume that a representative consumer in each market aggregates quantities of individual varieties with a constant elasticity of substitution (CES) utility function, but we augment this standard demand system with two forms of firm-destination-year demand heterogeneity. First, representative agents have *ex ante* taste shocks that are revealed to the firm prior to making production plans, and hence affect input choices. Second, representative agents have *ex post* taste shocks that are realized at the point of sale, and hence are unknown to firms at the time that they choose flexible inputs. If not controlled for, the *ex ante* demand shocks generate omitted variable bias—or *transmission bias*, as it is called in the literature—because they influence flexible input choices and directly affect revenues. The *ex post* shocks rationalize variation in flexible input shares in revenues across firms that use the same quantities of inputs. *Ex post* firm-destination-year shocks also generate variation in prices and markups both across firms and within firms across destinations. Hence, while we impose some structure on demand in order to circumvent missing price data, our model is flexible enough to allow for heterogeneous pricing behavior that is potentially important in real-world markets.[7]

We base our estimation on the two-step procedure from Gandhi et al. (2020). First, we project the expenditure share of flexible inputs (e.g., materials) in revenues on all inputs. This "factor share" regression nonparametrically identifies what is sometimes called the "revenue elasticity" of flexible inputs—i.e., composites of output elasticities and the industry-specific demand elasticity. While this first-step regression exploits the first order conditions of the firm, identification does not rely on any particular functional form of the production function. According to our model, the residual of this regression corresponds to a weighted average of unobserved *ex post* firm-destination-year demand shocks. Combining this residual with firm-level export revenue shares allows us to build a proxy for the weighted average of *ex ante* firm-specific demand shocks. Variation in this demand proxy allows us to estimate the curvature of demand without building aggregate demand shifters

[6]Under certain conditions on the profit function, this combinatorial discrete choice problem belongs to the class of problems studied by Arkolakis & Eckert (2017), and hence can be solved with their algorithm. Though our estimation strategy does not require that these conditions hold.

[7]For example, we allow for variable markups, pricing to market (Blum et al., 2023), and output reallocation across markets (Almunia et al., 2021). It is important to note that while our model allows for these features, we cannot estimate markups or study strategic pricing behavior, given the data constraints.

in each market from industry-wide price indices—as in the KG single-market approach—and without even knowing the set of destinations served by firms.

Next, we use the flexible input elasticities from the first step regression to compute the contribution of flexible inputs to revenues. In our nonparametric setting, the flexible input elasticities from the first step regression define a system of partial differential equations that can be integrated to compute the total contribution of flexible inputs, without ever specifying a functional form for the production function. We compute this contribution and subtract it off from revenues, along with the residual from the first step and, in the second step, regress the resulting values on predetermined inputs (e.g., capital) and the firm-specific demand proxy. We estimate the second step regression by generalized method of moments (GMM), allowing the unobserved productivity shock to depend on participation in the export market, as in De Loecker (2013).[8] The demand elasticity is identified by both cross-firm and time series variation in the share of export sales in total sales. This variation is driven by firm-specific time and market-specific demand shocks, time series variation in destination market aggregate demand shifters, and time and market-specific variation in fixed costs of entry. The demand elasticity is then used to recover output elasticities and retrieve the production function itself.

The model could also be estimated via the more popular control function methods of Blundell & Bond (2000), Wooldridge (2009) or Ackerberg et al. (2015).[9] However, recent work indicates two practical issues with the implementation of such methods. First, Gandhi et al. (2020) demonstrate that the control function method relies on time series variation in material input prices for identification. Given levels of input price variation observed in practice, Gandhi et al. (2020) argue that the control function approach is likely biased and inconsistent due to weak instruments. Second, Ackerberg et al. (2020) show that there are, in fact, multiple solutions to the GMM optimization in the standard control function framework, even when the sample size goes to infinity. Hence, results may be sensitive to the initial conditions given to the numerical search procedure, and there may not be obvious ways to choose among multiple solution vectors. In contrast, the factor shares approach does not rely on material input price variation for identification, and the GMM

[8]We focus on learning by exporting, but of course, one could control for other factors that are believed to affect firm productivity.

[9]These are the models most frequently cited when researchers use the control function method. But in fact, these models are all written for value-added production functions. As far as we are aware, Gandhi et al. (2020) is the only paper that establishes conditions under which the gross output production function is identified via the control function method. The control function approach we consider does include materials and follows the version developed in Gandhi et al. (2020).

used in the second step does not admit multiple solutions.[10]

We perform Monte Carlo simulations in order to compare the statistical properties of our multi-market estimator to: (1) the standard practice of deflating firm-level revenues by industry-wide price indices, (2) the factor shares approach that includes a correction for demand in a single market as in KG, and (3) the control function approach with single-market demand correction. We demonstrate that if the data generating process coincides with our multi-destination trade model with demand heterogeneity, then our estimator is consistent, while the considered alternatives are not. In addition, we find that our estimator has appealing finite-sample properties: the bias is small and confidence intervals have good "coverage ratios" (i.e., 95% confidence intervals contain the true parameters in about 95% of the simulated samples). Other estimators are inconsistent, strongly biased (compared to ours), and have confidence intervals with poor coverage ratios.

Finally, we use our procedure to study returns to scale, the elasticity of demand, and the effect of exporting on productivity in a panel of French manufacturing firms in 1994–2016. Using our estimator, we find price elasticities of demand ranging between -21.5 and -3.4, depending on industries, which is a range that is consistent with estimates from the gravity literature (for example, Shapiro 2016; Fontagné et al. 2022). On the supply side, we estimate that returns to flexible inputs are less than 1, on average. Decreasing returns to flexible inputs imply negative cross-market cost complementarities as in Almunia et al. (2021). We find overall increasing returns to scale, on average around 1.15. We also find evidence of learning by exporting (LBE) between zero and 4 percent year-on-year. These estimates imply cross-section differences in productivity between exporters and non-exporters of up to 40 percent. The model with no demand correction yields lower returns to scale, consistent with unaddressed transmission bias. The model with a single-market correction delivers unrealistic elasticities of demand (despite the relatively long period) and unrealistic returns to scale. Overall, our estimator outperforms existing methods in terms of obtaining sensible estimates of supply and demand elasticities in the data.

Our main contribution is to develop a production function estimator that exploits moment conditions that are consistent with a model in which firms potentially serve multiple destination markets and face heterogeneous demand shocks, without relying on quantity data. When quantity data are observed, output elasticities can be estimated without im-

[10] In Monte Carlo simulations, we find that indeed the control function approach yields a biased estimator, unless input prices vary significantly over time and only if the GMM parameter search starts from the true underlying structural parameter values of the model—the latter being an untenable condition in practice. Additionally, Doraszelski & Jaumandreu (2023) show that the first step in Ackerberg et al. (2015) leaves the demand heterogeneity partially uncontrolled for, which results in large biases when estimating markups.

posing structure on pricing behavior (as in, for example, Aw et al. 2011; Roberts et al. 2018; Blum et al. 2023). But in this case, structural assumptions are required on the supply side with respect to how firms apportion inputs across multiple production lines or products. To balance high demands on the data stemming from the existence of multiples production lines, quantity-based multi-product productivity estimators often rely on restrictive functional form assumption for the production function (Cobb-Douglas), e.g., Blum et al. 2023; de Roux et al. 2021. Additionally, it may be difficult to compare quantities across firms and products in a meaningful way, as noted by De Loecker & Goldberg (2014).[11] Moreover, physical quantity data are only available in rare datasets, and even then only for particular industries, limiting the applicability of quantity-based multi-product estimation techniques.[12] In contrast, we offer an approach that can be implemented in a wide range of differentiated-product markets with information that is widely available.

Beyond the production function literature mentioned above, our paper is related to a small literature on cross-market cost complementarities. Berman et al. (2015), Aghion et al. (2022), Barrows & Ollivier (2021) and Almunia et al. (2021) all estimate the effect of demand shocks in a given market on sales in a different market. If the returns to flexible inputs are decreasing, then more supply to one market increases the cost of serving other markets. Hence, positive *ex ante* demand shocks in one market should lower sales in another market. Neither Berman et al. (2015) nor Barrows & Ollivier (2021) estimate production function parameters, but rather focus on the reduced form connection between demand shocks in one market and sales in another market. Almunia et al. (2021) specifies a similar model to the one we develop. However, the procedure used by Almunia et al. (2021) to estimate the production function and the elasticity of demand is not consistent with the conditions necessary for cross-market cost complementarities nor with multi-destination markets, which we demonstrate in Appendix G.

Finally, our application is related to the literature on productivity-enhancing effects of exporting (Van Biesebroeck, 2005; De Loecker, 2007; Wagner, 2007, 2012; Garcia-Marin & Voigtländer, 2019; Atkin et al., 2017; Buus et al., 2022). It is often stated in the literature

[11] Quoting directly from De Loecker & Goldberg (2014), "the introduction of additional data creates its own challenges; although more data may help alleviate some of the problems discussed above, they are not a panacea" (page 206).

[12] For example, Blum et al. (2023) focus on only 10 3-digit ISIC industries for which there is a standard and uniform measure of physical quantities of output across firms, which yields only 2749 firms for Chilean manufacturing. Dhyne et al. (2022) analyze just 6 2-product environment in which firms tend to produce the same two 6-digit goods, thereby excluding roughly 90% of firm-year observations. de Roux et al. (2021) focus only on firms producing rubber and plastic products in the Columbian manufacturing survey (covering 362 firms).

that learning by exporting (LBE) is a phenomenon that affects firms from the developing world, if it affects any firms at all.[13] We follow De Loecker (2013) in estimating the effect of LBE by including an export dummy in the controlled Markov process, but study a set of developed-world firms for which outputs are denominated in value. When quantities are observed, then one can study LBE without explicitly modeling the export decision and export market. In contrast, if outputs are denominated in value, then it is critical to employ a multi-destination model that includes a correction for demand from multiple markets. In fact, we do not know a paper on LBE—including those cited just above—that addresses this issue.[14]

The rest of the paper is organized as follows. In Section 2 we describe the model that gives rise to an estimation equation for the production function. In Section 3 we explain different ways of estimating the production function, including our new estimator. In Section 4 we report results from Monte Carlo simulations that demonstrate the consistency and small sample properties of our estimator, and compares to those of other estimators. In Section 5 we describe the estimation results from French manufacturing data. Section 6 concludes.

2 Model

We specify a model in which heterogeneous single-product firms engage in monopolistic competition across horizontally differentiated varieties on multiple destination markets. The model is in partial equilibrium, as we seek only to link firm-level output to firm-level inputs and demand shifters. Closing the model would not alter estimation in any way, and hence we take the demand side of the model as exogenous. The model delivers an estimation equation for the production function and demand curvature, as well as a data generating process for Monte Carlo simulations.

[13]For example, in their review of industrial policy, Harrison & Rodríguez-Clare (2010) say "learning from exporting is most likely in technologically backward countries and among less productive firms". Or from Berman & Rebeyrol (2010) "One interesting regularity in the empirical papers aiming at testing the relevance of the learning by exporting hypothesis is that most studies finding a positive effect of the export status use data from developing countries".

[14]Van Biesebroeck (2005), De Loecker (2007) and De Loecker (2013) all deflate sales or value added by industry price indices—invariably, domestic price indices—in order to approximate quantities. This leaves firm-level variation in demand shocks a source of transmission bias. In addition, using domestic price indices implies that price conditions faced by exporters are identical in the domestic and foreign markets, which is at odds with the existence of variable trade barriers and different market conditions.

2.1 Demand

There are a fixed number of destination markets indexed by $d \in \{1, ..., \mathscr{D}\}$, origin markets indexed by $o \in \{1, ..., \mathscr{D}\}$, and industries indexed by $i \in \{1, ..., \mathscr{I}\}$. In each destination market a representative consumer aggregates consumption in two tiers. In the top tier, the consumer aggregates over industry-level consumption bundles with a flexible utility function:

$$U_t^d = U_t^d\big(B_{1t}^d, B_{2t}^d, ..., B_{\mathscr{I}t}^d\big), \tag{1}$$

where t indexes time. Within a generic industry i, consumers aggregate over varieties f produced in country of origin o with a CES structure:

$$B_{it}^d = \left[\sum_o \sum_{f \in \Theta_{it}^{od}} \big(X_{ft}^{od}\big)^{\rho_i} \exp(\epsilon_{ft}^{od} + u_{ft}^{od})\right]^{1/\rho_i}, \tag{2}$$

where X_{ft}^{od} is the quantity consumed of variety f in destination d sourced from o in time t, ϵ_{ft}^{od} is an *ex ante* variety-specific demand shock (realized prior to production), u_{ft}^{od} is an *ex post* variety-specific demand shock (realized at the point of sales), Θ_{it}^{od} is the set of varieties in industry i shipped from origin o to destination d in year t, and $\rho_i < 1$ is a parameter that governs the substitutability of varieties within the industry, with constant price elasticity of demand $\eta_i = 1/(\rho_i - 1) < -1$. The CES price index at the industry level is defined in the usual way:

$$\Upsilon_{it}^d = \left[\sum_o \sum_{f \in \Theta_{it}^{od}} \big(P_{ft}^{od}\big)^{\frac{\rho_i}{\rho_i - 1}} \exp\left(\frac{1}{1-\rho_i}\left(\epsilon_{ft}^{od} + u_{ft}^{od}\right)\right)\right]^{\frac{\rho_i - 1}{\rho_i}} \tag{3}$$

where P_{ft}^{od} is the price of variety f sourced from o that is paid by consumers in destination market d at time t.

The representative consumer's objective is to maximize her utility (1) given her budget constraint. The CES structure yields an expression for expenditures R_{ft}^{od} on each variety f in destination d:

$$R_{ft}^{od} = \big(X_{ft}^{od}\big)^{\rho_i} \frac{\Upsilon_{it}^d}{\big(B_{it}^d\big)^{\rho_i - 1}} \exp(\epsilon_{ft}^{od} + u_{ft}^{od}). \tag{4}$$

Given the empirical applications we consider, we make two notational simplifications. First, as we perform our analysis industry-by-industry, we drop the industry index i. Second, we assume that researchers only observe varieties and firms coming from a single origin country, which we refer to as $o = 1$. Hence, we drop the o index from now on.

2.2 Production

Firms produce a single differentiated variety which they may ship to many destination markets. To serve a given market, firms must pay a firm-destination-year specific fixed cost C^d_{ft} and a destination-specific ad valorem "iceberg" cost $\tau^d_t \geq 1$. For simplicity, we assume that there are no domestic fixed costs, so that $C^1_{ft} = 0$. This ensures that all firms sell on the domestic market. We also normalize the iceberg cost to sell on the domestic market to 1. To sell X^d_{ft} units to destination market d, firm f must produce $Q^d_{ft} = \tau^d_t X^d_{ft}$ units. At each period t, the sum of all units sold to all destination markets must equal total output: $\sum_d Q^d_{ft} = Q_{ft}$.

Firms produce outputs using flexible inputs (written in logs) $\boldsymbol{v_{ft}} = (v^1_{ft},...,v^{\mathscr{V}}_{ft})$ and quasi-fixed inputs (written in logs) $\boldsymbol{\kappa_{ft}} = (\kappa^1_{ft},...,\kappa^{\mathscr{K}}_{ft})$. Flexible inputs are chosen optimally each period given input prices (in levels) $\boldsymbol{W_t} = (W^1_t, ..., W^{\mathscr{V}}_t)$. Quasi-fixed inputs (such as capital) evolve each period according to the depreciation rate and an endogenous investment choice. For each quasi-fixed input κ^j, we have

$$\exp(\kappa^j_{ft}) = (1 - \varrho^j) \exp(\kappa^j_{ft-1}) + \iota^j_{ft-1}, \tag{5}$$

where ϱ^j denotes the rate of depreciation and ι^j_{ft-1} is the investment choice in period $t-1$.[15]

Quantity produced is a deterministic function of a Hicks-neutral productivity shock ω_{ft} and a twice continuously differentiable transformation of variable and quasi-fixed inputs $F(\cdot)$:

$$Q_{ft} = \exp(\omega_{ft}) F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) \iff q_{ft} = \omega_{ft} + f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}), \tag{6}$$

where lower case indicates logs.

2.3 Optimization

The firm solves a combinatorial discrete choice problem each period in which it chooses a vector that indicates which markets to serve $\boldsymbol{I_{ft}} = (I^1_{ft}, ..., I^{\mathscr{D}}_{ft})$—where I^d_{ft} as an indicator that equals 1 if firm f serves market d in year t and equals 0 otherwise—a vector of destination-specific output shares $\boldsymbol{\chi_{ft}} = (\chi^1_{ft}, \chi^2_{ft}, ..., \chi^{\mathscr{D}}_{ft})$, and a vector of flexible inputs

[15] We later entertain the case of inputs that evolve in a more general manner (partial adjustment).

$\boldsymbol{v_{ft}}$ to maximize expected profits, given flexible input prices, quasi-fixed inputs, fixed and iceberg trade costs, and market-specific demand conditions. The firm takes expectations over *ex post* demand shocks u_{ft}^d, which are assumed to be i.i.d. with a constant mean u and variance σ_u^2, that are both known to the firm.[16]

Using (4), we write the optimization problem as

$$\begin{aligned} \max_{\boldsymbol{I_{ft}}} \max_{\boldsymbol{\chi_{ft}}, \boldsymbol{v_{ft}}} \quad \mathscr{L} &= E\left[\exp(\rho\omega_{ft}) F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})^\rho \sum_d \left(\chi_{ft}^d\right)^\rho D_t^d \exp(\epsilon_{ft}^d + u_{ft}^d) \right] \\ &- \sum_j \exp(v_{ft}^j) W_t^j + \lambda_{ft}\left(1 - \sum_d \chi_{ft}^d\right) - \sum_d I_{ft}^d C_{ft}^d \,, \end{aligned} \tag{7}$$

where $D_t^d \equiv \Upsilon_t^d \left(B_t^d\right)^{1-\rho} \left(\tau_t^d\right)^{-\rho}$ is a destination-industry-specific demand shifter, and λ_{ft} is the Lagrangian associated to the constraint $\sum_d \chi_{ft}^d = 1$. We ignore for simplicity the additional constraints $\chi_{ft}^d \geq 0$ for all d, with the understanding that $\chi_{ft}^d > 0$ whenever $I_{ft}^d = 1$.

We first solve for the optimal $\boldsymbol{\chi_{ft}}$ and $\boldsymbol{v_{ft}}$, given a set of destinations, Ω_{ft}, that are served with strictly positive quantities. Assuming monopolistic competition implies that firms take price indices as given. First order conditions yield, for each destination $d \in \Omega_{ft}$,

$$E\left[\exp(u)\right] (Q_{ft})^\rho \rho \left(\chi_{ft}^d\right)^{\rho-1} D_t^d \exp(\epsilon_{ft}^d) = \lambda_{ft} \tag{8}$$

and for each flexible input v^j

$$\rho \exp(\rho\omega_{ft}) E\left[\exp(u)\right] \left[\sum_{d\in\Omega_{ft}} \left(\chi_{ft}^d\right)^\rho D_t^d \exp(\epsilon_{ft}^d) \right] \left(F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})\right)^{\rho-1} \frac{\partial F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})}{\partial \exp(v_{ft}^j)} = W_t^j, \tag{9}$$

given $E\big[\exp(u_{ft}^d)\big] = E\big[\exp(u)\big]$, a constant, for all firms and destinations.

For any two markets d and d' served by firm f, we have from (8)

$$\chi_{ft}^d = \chi_{ft}^{d'} \left[\frac{D_t^{d'} \exp(\epsilon_{ft}^{d'})}{D_t^d \exp(\epsilon_{ft}^d)} \right]^{\frac{1}{\rho-1}}. \tag{10}$$

Summing over destinations and rearranging yields the optimal quantity share for any des-

[16] As in Gandhi et al. (2020), *ex post* shocks are necessary to rationalize variation in input expenditure shares across firms. Whereas Gandhi et al. (2020) assume that these shocks are i.i.d. draws from the same distribution function with constant mean within a market, we extend this assumption to the multi-market context.

tination market d served by firm f

$$\chi_{ft}^{d} = \frac{\left(D_t^d \exp(\epsilon_{ft}^d)\right)^{\frac{1}{1-\rho}}}{\sum_{z\in\Omega_{ft}} \left(D_t^z \exp(\epsilon_{ft}^z)\right)^{\frac{1}{1-\rho}}}. \tag{11}$$

Plugging the last equation into (9) we get

$$\rho \exp(\rho\omega_{ft}) E\left[\exp(u)\right] \left[\sum_{d\in\Omega_{ft}} \left(D_t^d \exp(\epsilon_{ft}^d)\right)^{\frac{1}{1-\rho}}\right]^{1-\rho} (F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}))^{\rho-1} \frac{\partial F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})}{\partial \exp(v_{ft}^j)} = W_t^j. \tag{12}$$

This is a system of $\mathscr{V}$ equations with $\mathscr{V}$ unknowns. A sufficient condition for a unique interior solution is that $F(\cdot)$ is concave in each flexible input.

Next, firms choose the set of destinations that maximizes total expected profits over all possible sets. Firms can either solve the combinatorial problem by computing the expected profits for each possible set, or, under certain conditions, by employing an efficient algorithm as in Arkolakis & Eckert (2017).[17]

We can write the optimal vector of indicators for market entry by firm as

$$\boldsymbol{I_{ft}} = I\left(\omega_{ft}, \epsilon_{\mathbf{ft}}, \boldsymbol{\kappa_{ft}}, \mathbf{C_{ft}}, \mathbf{D_t}, \mathbf{W_t}\right) \tag{13}$$

to stress that each indicator depends on the firm-specific productivity ω_{ft}, quasi-fixed inputs $\boldsymbol{\kappa_{ft}}$, the entire vector of destination-industry market potentials $\mathbf{D_t}$, the vector of input prices $\mathbf{W_t}$, and each firm's entire vector of fixed costs $\mathbf{C_{ft}}$ and vector of *ex ante* demand shocks $\epsilon_{\mathbf{ft}}$. The optimal input demand and the quantities sold on each market are implicit in the solution for the optimal $\boldsymbol{I_{ft}}$.

Given the optimal Q_{ft}, destination-specific prices can be found using (4) and (11)

$$P_{ft}^d = \tau_t^d \left(Q_{ft}\right)^{\rho-1} \left(\sum_{z\in\Omega_{ft}} \left(D_t^z \exp(\epsilon_{ft}^z)\right)^{\frac{1}{1-\rho}}\right)^{1-\rho} \exp(u_{ft}^d). \tag{14}$$

Hence, prices vary across destinations within the firm both due to variable trade barri-

[17] The algorithm of Arkolakis & Eckert (2017) requires either decreasing or increasing differences in the extensive margin of exports. If the production function output elasticities and overall return to scale are "relatively stable", or Cobb-Douglas type, then with a constant demand elasticity we can satisfy the necessary conditions for the algorithm (either for complementarity or substitutability across destinations). With a general production function, we cannot ensure that these conditions hold, though we do not require these conditions for our estimation.

ers (τ_t^d) and because of firm-destination demand heterogeneity (u_{ft}^d). Firm-destination markups can be computed as the ratio of prices to marginal cost, the latter being equal to expected marginal revenues on any given destination. This yields

$$\mu_{ft}^d = \frac{\exp(u_{ft}^d)}{\rho E\left[\exp(u)\right]}. \tag{15}$$

Hence, like prices, markups vary across firms and across destinations within the firm because of firm-destination demand heterogeneity.

2.4 Estimation equation

Plugging in the solution (11) for χ_{ft}^d into total revenues and taking logs, we have

$$r_{ft} = \rho\omega_{ft} + \rho f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) + (1-\rho)\ln\left[\sum_{d\in\Omega_{ft}} \left(D_t^d \exp(\epsilon_{ft}^d)\right)^{\frac{1}{1-\rho}}\right] + \ln\psi_{ft}, \tag{16}$$

where the set Ω_{ft} represents the optimal choice of $\boldsymbol{I_{ft}}$ from (7) and where the last term of (16) is the log of a weighted average of *ex post* demand shocks, with $\psi_{ft} \equiv \sum_{d\in\Omega_{ft}} \chi_{ft}^d \exp(u_{ft}^d)$.

The second to last term of (16) is the firm-specific demand shifter. It depends on the set Ω_{ft}, aggregate industry-destination demand shifters D_t^d and firm-destination specific demand shocks ϵ_{ft}^d, which are both observed by the firm before making production decisions. Hence, the term affects input decisions and must be controlled for to avoid transmission bias.

Rather than building a proxy for this demand shifter—which would be quite demanding from a data perspective—we exploit the first order conditions. Since all firms serve the domestic market $d = 1$ (given $C_{ft}^1 = 0$), first order conditions (8) imply, for any destination d served by a firm f,

$$\frac{R_{ft}^d}{R_{ft}^1} = \left(\frac{D_t^d \exp(\epsilon_{ft}^d)}{D_t^1 \exp(\epsilon_{ft}^1)}\right)^{\frac{1}{1-\rho}} \frac{\exp(u_{ft}^d)}{\exp(u_{ft}^1)}. \tag{17}$$

Rearranging and summing over destinations yields

$$\sum_{d\in\Omega_{ft}} \left(D_t^d \exp(\epsilon_{ft}^d)\right)^{\frac{1}{1-\rho}} \exp(u_{ft}^d) = \left(D_t^1\right)^{\frac{1}{1-\rho}} \exp\left(\frac{\epsilon_{ft}^1}{1-\rho} + u_{ft}^1\right) \sum_{d\in\Omega_{ft}} \frac{R_{ft}^d}{R_{ft}^1}. \tag{18}$$

Using the definition of ψ_{ft} and rearranging yields

$$\sum_{d\in\Omega_{ft}} \left(D_t^d \exp(\epsilon_{ft}^d)\right)^{\frac{1}{1-\rho}} = \left(D_t^1\right)^{\frac{1}{1-\rho}} \exp\left(\frac{\epsilon_{ft}^1}{1-\rho} + u_{ft}^1\right) \frac{R_{ft}}{R_{ft}^1} \frac{1}{\psi_{ft}}. \tag{19}$$

Plugging this expression back into (16), we obtain our estimation equation

$$r_{ft} = \ln D_t^1 + \rho f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) + (1-\rho) \ln D_{ft} + \epsilon_{ft}^1 + (1-\rho)\, u_{ft}^1 + \rho\omega_{ft} + \ln \psi_{ft}. \tag{20}$$

with

$$D_{ft} \equiv \frac{R_{ft}}{R_{ft}^1} \frac{1}{\psi_{ft}}.$$

With this substitution, we split the endogenous demand shifter into an aggregate component D_t^1 that can be absorbed into time fixed effects, and a firm-specific component D_{ft} that depends only on observable data (the export share) and the weighted average of *ex post* demand shocks—which can be estimated from the data, as we demonstrate below. The last four terms of (20) collect unobserved shocks to productivity, *ex ante* domestic demand shocks (ϵ_{ft}^1), and *ex post* demand shocks.

3 Empirical strategy

We base our estimator on the two-step factor shares approach of Gandhi et al. (2020). In their main text, Gandhi et al. (2020) treat the case in which outputs are denominated in quantity; but in Appendix O6-4, they consider the "revenue production function", i.e. the case in which outputs are denominated in value. The primary difference between our estimation and theirs is that we allow firms to serve multiple destination markets on which they face heterogeneous demand conditions. We also present alternative estimation strategies based on the more popular control function approach, although we prefer the factor shares approach for reasons we discuss below.

3.1 Multi-market estimator: first step

Following Gandhi et al. (2020), estimation proceeds in two steps. In the first step, output elasticities with respect to flexible inputs are identified from projecting factor expenditure shares on logs of input levels. The estimation equation is derived from the first order conditions for flexible inputs. When outputs are denominated in value, these elasticities are inclusive of the demand-side parameter ρ.[18]

In the multi-destination market case, we combine (16) with (12) and obtain the cost share in revenue of flexible input v^j

$$\ln s^j_{ft} = \ln\left[\exp(-E\left[\ln(\psi)\right])E\left[\exp(u)\right]\beta^j_{ft}(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})\right] + \varphi_{ft} \tag{21}$$

where we define $s^j_{ft} \equiv \frac{W^j_t \exp(v^j_{ft})}{R_{ft}}$, where $\beta^j_{ft}(\cdot) \equiv \rho\sigma^j_{ft}(\cdot) \equiv \rho\frac{\partial F(\boldsymbol{v_{ft}},\boldsymbol{\kappa_{ft}})}{\partial \exp(v^j_{ft})}\frac{\exp(v^j_{ft})}{F_{ft}}$ denotes the output elasticity of flexible input v^j multiplied by ρ, or the "revenue elasticity" of input v^j, and where $\varphi_{ft} \equiv E\left[\ln(\psi)\right] - \ln\left[\psi_{ft}\right]$. We add and subtract the constant $E\left[\ln(\psi)\right]$ because $E\left[\ln\psi_{ft}\right] \neq 0$ due to Jensen's inequality.[19]

To operationalize the estimator, we follow Gandhi et al. (2020) in approximating $\beta^j_{ft}(\cdot)$ with a complete polynomial function of degree 2 in all inputs.[20] We estimate $\beta^j_{ft}(\cdot)$ by NLLS for each flexible input v^j:

$$\min_{\boldsymbol{g^j}} \quad \sum_f\sum_t\Bigg\{\ln s^j_{ft} - \ln\Bigg(g^j_0 + \sum_{z\in\{v^1,\ldots,v^{\mathscr{V}},k^1,\ldots,k^{\mathscr{K}}\}} g^j_z z_{ft}$$
$$+ \sum_{\ell\in\{1,\ldots,\mathscr{V}\}}\ \sum_{z\in\{v^\ell,\ldots,v^{\mathscr{V}},k^1,\ldots,k^{\mathscr{K}}\}} g^j_{v^\ell z} v^\ell_{ft} z_{ft} + \sum_{\ell\in\{1,\ldots,\mathscr{K}\}}\ \sum_{z\in\{v^1,\ldots,v^{\mathscr{V}},\kappa^\ell,\ldots,\kappa^{\mathscr{K}}\}} g^j_{\kappa^\ell z}\kappa^\ell_{ft} z_{ft}\Bigg)\Bigg\}^2 \tag{22}$$

where all the g^j coefficients include the constant $\exp(-E\left[\ln(\psi)\right])E\left[\exp(u)\right]$. To purge this constant from the g^j coefficients we compute

$$\widehat{\exp(-E\left[\ln(\psi)\right])}E\left[\exp(u)\right] = \frac{1}{N}\sum_f\sum_t \exp(-\widehat{\varphi}_{ft}) \tag{23}$$

where N is the number of firm-year observations and $\widehat{\varphi}_{ft}$ is the residual from (22).[21] We

[18]In general, the "revenue elasticities" of flexible inputs are identified from the factor share regressions as long as the markup does not depend on input levels (and as long as the orthogonality condition discussed below is met). But this does not require that markups are fixed. Markups may vary over time and across firms in our model because of *ex post* demand shocks.

[19]In the single-market case, the residual φ_{ft} is simply the single *ex post* demand shock $-u_{ft}$, which is mean zero and exogenous by assumption.

[20]This sieve series estimator is one way of operationalizing a nonparametric regression. Another possibility would be a local linear regression. One advantage of the sieve estimator is that the integral of $\beta^j_{ft}(\cdot)$ will have a closed-form solution. But the use of polynomials does not entail any assumptions about the functional form of $F(\cdot)$.

[21]Since (23) calls for the use of the exponential function, the estimate of $\widehat{\exp(-E\left[\ln(\psi)\right])}E\left[\exp(u)\right]$ may be sensitive to the presence of extreme outliers. In the French data, we exclude any firm that ever has a

then divide all g^j coefficients by this constant and compute

$$\begin{aligned}\widehat{\beta}^j_{ft}(\boldsymbol{v_{ft}},\boldsymbol{\kappa_{ft}}) &= \widehat{g}^j_0 + \sum_{z\in\{v^1,...,v^{\mathscr{V}},k^1,...,k^{\mathscr{K}}\}} \widehat{g}^j_z z_{ft} + \sum_{\ell\in\{1,...,\mathscr{V}\}} \sum_{z\in\{v^\ell,...,v^{\mathscr{V}},k^1,...,k^{\mathscr{K}}\}} \widehat{g}^j_{v^\ell z} v^\ell_{ft} z_{ft} \\ &+ \sum_{\ell\in\{1,...,\mathscr{K}\}} \sum_{z\in\{v^1,...,v^{\mathscr{V}},\kappa^\ell,...,\kappa^{\mathscr{K}}\}} \widehat{g}^j_{\kappa^\ell z} \kappa^\ell_{ft} z_{ft} \end{aligned} \quad (24)$$

Identification of equation (22) requires orthogonality between φ_{ft} and all variable and quasi-fixed inputs. In the multi-market case, it is not *a priori* obvious that this condition holds, since the weights χ^d_{ft} are potentially endogenous to input choices. Nevertheless, we have

Proposition 1. $E[\varphi_{ft}|\boldsymbol{v_{ft}},\boldsymbol{\kappa_{ft}}]=0$*, hence, the share regression (22) identifies the revenue elasticity of flexible input* v^j*,* β^j_{ft}*, and the residual* φ_{ft}*.*

Proof: see Appendix A

The key to the proof is that even though the destination weights χ^d_{ft} are endogenous to input choices, the weights are orthogonal to the realized demand shocks u^d_{ft}. Hence, by the law of iterated expectations, $E\Big[\sum_{d\in\Omega_{ft}}\chi^d_{ft}\exp(u^d_{ft})\Big]=E[\exp(u)]$, a constant, and φ_{ft} is orthogonal to input choices.

3.2 Multi-market estimator: second step

The second step of the procedure is to use the information from the first step to recover the rest of the production function. The basic insight from Gandhi et al. (2020) is that the flexible input elasticity defines a partial differential equation that can be integrated to compute the part of the production function related to each flexible input j.

By the fundamental theorem of calculus, for each flexible input v^j,

$$\int_{v^j_0}^{v^j_{ft}} \beta^j_{ft}(\boldsymbol{v_{ft}},\boldsymbol{\kappa_{ft}})dv^j_{ft} = \rho f(\boldsymbol{v_{ft}},\boldsymbol{\kappa_{ft}}) + \rho\mathscr{C}^j\left(v^1_{ft},...,v^{j-1}_{ft},v^{j+1}_{ft},...,v^{\mathscr{V}}_{ft},\kappa^1_{ft},...,\kappa^{\mathscr{K}}_{ft}\right) \quad (25)$$

where v^j_0 is the minimum possible value of flexible input v^j and $\mathscr{C}^j(\cdot)$ is a constant of integration that depends on all quasi-fixed inputs and all flexible inputs except for input v^j. As noted in Appendix O6-3 of Gandhi et al. (2020), these differential equations can

material expenditure share or a labor expenditure share greater than 20 or less than 0.001. This restriction excludes less than 0.1% of the data.

be combined to construct the production function up to a constant that depends only on predetermined inputs (also see Varian 1992, pages 483-484).[22]

Substituting this expression of the production function into (20), we compute revenues net of the contribution of flexible inputs and $\widehat{\varphi}_{ft}$:

$$\begin{aligned}\widetilde{r}_{ft} \equiv r_{ft} \quad &- \quad \int_{v_0^1}^{v_{ft}^1} \beta_{ft}^1(z^1, v_0^2, ..., v_0^{\mathscr{V}}, \kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}) dz^1 \\ &- \quad ... - \int_{v_0^{\mathscr{V}}}^{v_{ft}^{\mathscr{V}}} \beta_{ft}^{\mathscr{V}}(v_{ft}^1, v_{ft}^2, ..., z^{\mathscr{V}}, \kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}) dz^{\mathscr{V}} + \widehat{\varphi}_{ft}. \end{aligned} \tag{26}$$

We then transform (20) into

$$\begin{aligned}\widetilde{r}_{ft} \quad &= \quad \alpha_t + \beta^D \ln \widehat{D}_{ft} + \sum_{j\in\{1,...,\mathscr{K}\}} b_{\kappa^j} \kappa_{ft}^j + \sum_{j\in\{1,...,\mathscr{K}\}} \sum_{z\in\{\kappa^j,...,\kappa^{\mathscr{K}}\}} b_{\kappa^j z} \kappa_{ft}^j z_{ft} \\ &+ \quad \epsilon_{ft}^1 + (1-\rho)\, u_{ft}^1 + \rho \omega_{ft}, \end{aligned} \tag{27}$$

where $\alpha_t \equiv \ln D_t^1 + \rho E[\ln(\psi)]$ collects industry-period terms, with $\rho E[\ln(\psi)]$ carrying over from the first-step estimation of $\widehat{\varphi}_{ft}$, $\widehat{D}_{ft} \equiv (R_{ft}/R_{ft}^1)\exp(\widehat{\varphi}_{ft})$ proxies for the firm-specific demand shock and identifies the demand-side parameter $\beta^D \equiv 1-\rho$, and the term $\rho\mathscr{C}\left(\kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}\right)$ is approximated by a complete polynomial function of degree 2 in quasi-fixed factors (last two terms in the first line of (27)).

In equation (27), ω_{ft}, ϵ_{ft}^1, and u_{ft}^1 are all endogenous to $\widehat{D}_{ft}$ both through the endogenous choice of destinations and through realized sales in the domestic market. Additionally, if ω_{ft}, ϵ_{ft}^1, and u_{ft}^1 are persistent, then they correlate with all quasi-fixed inputs through the investment rule. We assume that ω_{ft} and ϵ_{ft}^1 evolve according to first order Markov processes and exploit timing for identification. In particular, we assume productivity ω_{ft} follows an AR(1) process and depends on lagged export participation indicator $e_{f,t-1}$, as in De Loecker (2013):

$$\omega_{ft} = h\omega_{f,t-1} + \mu e_{f,t-1} + \widetilde{\omega}_{ft} \tag{28}$$

[22]Combining differential equations for each flexible input, we have

$$\begin{aligned}\rho f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) \quad &= \quad \int_{v_0^1}^{v_{ft}^1} \beta_{ft}^1(z^1, v_0^2, ..., v_0^{\mathscr{V}}, \kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}) dz^1 + \int_{v_0^2}^{v_{ft}^2} \beta_{ft}^2(v_{ft}^1, z^2, v_0^3, ..., v_0^{\mathscr{V}}, \kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}) dz^2 \\ &+ \quad \ldots + \int_{v_0^{\mathscr{V}}}^{v_{ft}^{\mathscr{V}}} \beta_{ft}^{\mathscr{V}}(v_{ft}^1, v_{ft}^2, ..., z^{\mathscr{V}}, \kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}) dz^{\mathscr{V}} - \rho\mathscr{C}\left(\kappa_{ft}^1, ..., \kappa_{ft}^{\mathscr{K}}\right). \end{aligned}$$

where h is a scalar and $\widetilde{\omega}_{ft}$ represents an i.i.d. shock to productivity.[23] The parameter μ indicates the effect of lagged export participation on current productivity—the effect of "leaning by exporting" (LBE). We further assume that the domestic *ex ante* demand shock follows an AR(1) process with the same persistence parameter h,[24]

$$\epsilon^1_{ft} \quad = \quad h\epsilon^1_{f,t-1} + \widetilde{\epsilon}^1_{ft}, \tag{29}$$

where $\widetilde{\epsilon}^1_{ft}$ represents i.i.d. shocks to domestic demand. The assumption that productivity shocks and demand shocks share the same persistence parameter h allows us to combine them into a composite shock, in a similar fashion to De Loecker (2011) and Melitz & Levinsohn (2006): $\nu_{ft} \equiv \epsilon^1_{ft} + \rho\omega_{ft}$, which by assumptions (28) and (29) gives

$$\nu_{ft} = h\nu_{f,t-1} + \rho\mu e_{f,t-1} + \xi_{ft}, \tag{30}$$

where $\xi_{ft} \equiv \widetilde{\epsilon}^1_{ft} + \rho\widetilde{\omega}_{ft} + (1-\rho)u^1_{ft} + h(1-\rho)u^1_{ft-1}$ is an MA(1) error term.

Substituting ν_{ft} into (27) yields

$$\widetilde{r}_{ft} \quad = \quad \alpha_t + \beta^D \ln \widehat{D}_{ft} + \sum_{j\in\{1,\dots,\mathscr{K}\}} b_{\kappa^j}\kappa^j_{ft} + \sum_{j\in\{1,\dots,\mathscr{K}\}}\sum_{z\in\{\kappa^j,\dots,\kappa^{\mathscr{K}}\}} b_{\kappa^j z}\kappa^j_{ft}z_{ft} + \nu_{ft} \tag{31}$$

For any candidate vector $\left(\beta^{D^*}, b^*_{\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}}, b^*_{\kappa^1\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}\right)$, we can compute

$$\widehat{\nu_{ft} + \alpha_t} \quad = \quad \widetilde{r}_{ft} - \beta^{D^*} \ln \widehat{D}_{ft} - \sum_{j\in\{1,\dots,\mathscr{K}\}} b_{\kappa^j}\kappa^j_{ft} - \sum_{j\in\{1,\dots,\mathscr{K}\}}\sum_{z\in\{\kappa^j,\dots,\kappa^{\mathscr{K}}\}} b_{\kappa^j z}\kappa^j_{ft}z_{ft}\kappa^{j'}_{ft}, \tag{32}$$

and then regress $\widehat{\nu_{ft} + \alpha_t}$ on $\nu_{f,t}\widehat{_{-1} + \alpha}_{t-1}$, the past exporting decision $e_{f,t-1}$ and time fixed effects, and compute the residual $\widehat{\xi}_{ft}\left(\beta^{D^*}, b^*_{\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}}, b^*_{\kappa^1\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}\right)$. We then build the following moment conditions:

$$E\left\{\widehat{\xi}_{ft}\left(\beta^{D^*}, b^*_{\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}}, b^*_{\kappa^1\kappa^1}, \dots, b^*_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}\right)\begin{pmatrix}\ln \widehat{D}_{f,t-2} \\ \kappa^1_{ft} \\ \vdots \\ \left(\kappa^{\mathscr{K}}_{ft}\right)^2\end{pmatrix}\right\} = 0 \tag{33}$$

[23]De Loecker (2013) assumes a flexible first order Markov process, but with endogenous firm-specific demand shocks we must impose linearity for estimation.

[24]Note that at this stage we make no assumptions on the evolution of ϵ^d_{ft} for $d \neq 1$, i.e., on any other market that is not the domestic one; we discuss below assumptions that may be required—but not necessarily—for identification.

and minimize deviations from these moments by GMM.

At the true parameter values, $\widehat{\xi}_{ft}$ is orthogonal to all quasi-fixed inputs in period t. This is because $\widehat{\xi}_{ft}$ contains only period t innovations to productivity $\tilde{\omega}_{ft}$ and domestic demand $\tilde{\epsilon}^1_{ft}$, and *ex post* domestic demand shocks u^1_{ft} and $u^1_{f,t-1}$, none of which influence the investment decision in period $t-1$. However, even at the true parameter values $\widehat{\xi}_{ft}$ correlates with $\widehat{D}_{ft}$ and $\widehat{D}_{f,t-1}$ through the endogenous set of destinations and through sales on the domestic market.[25] Thus, to build the objective function, we use $\ln \widehat{D}_{f,t-2}$, which is orthogonal to $\widehat{\xi}_{ft}$.

Finally, we compute $\widehat{\rho} = 1 - \widehat{\beta}^D$ and the output elasticity for each quasi-fixed input κ^k:[26]

$$\widehat{\sigma^k_{ft}} \quad = \quad \frac{1}{\widehat{\rho}}\left(\frac{\partial \widetilde{r}_{ft}}{\partial \kappa^k_{ft}} \quad + \sum_{j\in\{1,\ldots,\mathscr{V}\}} \frac{\partial}{\partial \kappa^k_{ft}}\left[\int \beta^j_{ft}(\cdot) dv^j_{ft}\right]\right) \tag{34}$$

and for flexible inputs

$$\widehat{\sigma^j_{ft}} \quad = \quad \widehat{\beta^j_{ft}}/\widehat{\rho}. \tag{35}$$

We compute the LBE effect as the point estimate on the export lag from the regression estimates of the Markov process, deflated by $\widehat{\rho}$.

Since the second step uses estimated objects from the first step, we bootstrap the entire two-step procedure to compute standard errors. The bootstrap procedure samples firms rather than individual observations, which is akin to clustering standard errors by firm.

Before proceeding to alternative estimators, we discuss the source of identification of ρ in (33). As mentioned earlier, the assumptions of the model imply that $\ln \widehat{D}_{f,t-2}$ is orthogonal to $\widehat{\xi}_{ft}$, so the moment condition should hold. But what about relevance? Conditional on quasi-fixed inputs, time fixed effects, and $\nu_{f,t-1}\widehat{+\alpha}_{t-1}$, there are at least two explanations for the correlation between $\ln \widehat{D}_{f,t-2}$ and $\ln \widehat{D}_{f,t}$.

---

[25] Recall that in order to build $\widehat{D}_{f,t-1}$ we use realized domestic sales, which are directly affected by $u^1_{f,t-1}$.

[26] Assuming a second degree polynomial for both the first step and the second step yields

$$\begin{aligned}\widehat{\sigma^k_{ft}} \quad = \quad & \frac{1}{\widehat{\rho}}\Bigg(\widehat{b}_{\kappa^k} + 2\widehat{b}_{\kappa^k\kappa^k}\kappa^k_{ft} + \sum_{j\in\{1,..k-1,k+1,..,\mathscr{K}\}} \widehat{b_{\kappa^j\kappa^k}}\kappa^j_{ft} + \sum_{j\in\{1,\ldots,\mathscr{V}\}} \widehat{g}^j_{\kappa^k} v^j_{ft} \\ & + \quad 2\sum_{j\in\{1,\ldots,\mathscr{V}\}} \widehat{g}^j_{\kappa^k\kappa^k} v^j_{ft}\kappa^k_{ft} + \sum_{j\in\{1,\ldots,\mathscr{V}\}}\sum_{z\in\{k^1,\ldots,k^{k-1},k^{k+1}..,k^{\mathscr{K}}\}} \widehat{g}^j_{z\kappa^k} z_{ft} v^j_{ft} + \frac{1}{2}\sum_{j\in\{1,\ldots,\mathscr{V}\}} \widehat{g}^j_{v^j\kappa^k} v^j_{ft} v^j_{ft}\Bigg)\end{aligned}$$

First, persistence in the *ex ante* foreign demand shocks ϵ^d_{ft}, for $d \neq 1$, yields correlation between $\ln \widehat{D}_{f,t-2}$ and $\ln \widehat{D}_{f,t}$. To see this, consider two firms with the same evolution of ν_{ft} (which includes only domestic demand shocks ϵ^1_{ft}) and quasi-fixed inputs, serving the same set of destinations Ω_{ft}. Suppose that the first firm has persistently higher draws for ϵ^d_{ft} for some $d > 1$ than the second firm. The former firm will tend to earn a higher share of revenue from the export market than the latter, and hence will tend to have higher $\ln \widehat{D}_{ft}$ in all periods.

The second mechanism that generates correlation between $\ln \widehat{D}_{f,t-2}$ and $\ln \widehat{D}_{ft}$ relies on persistent firm-specific fixed costs of market entry. To see this, consider two firms with the same evolution of ν_{ft} and quasi-fixed inputs, but different fixed costs of reaching different markets. In this case, the two firms will likely serve different markets. If these fixed costs are persistent, then the two firms will be exposed to different aggregate shocks. Suppose that the first firm has lower fixed costs compared to the second firm for serving a particular large foreign market. Then the former firm will tend to earn more from exporting than the latter, all else equal, and thus will tend to have a higher $\ln \widehat{D}_{ft}$ in all periods.[27]

3.3 Factor share method with no demand correction

When estimating production function parameters with data denominated in value, the vast majority of researchers simply deflate firm-level revenues by the domestic price deflator and treat the resulting series as if they were quantities. Given our data generating process, only under the assumption that $\rho = 1$ would deflating by the domestic price index convert firm-level revenues into firm-level quantities. Rationalizing this common practice therefore implies that firms produce homogeneous goods. In this case, the conditions from the main text of Gandhi et al. (2020) would be met, and thus their factor shares estimator could be applied.

However, in the case that goods are not perfect substitutes ($\rho < 1$), then deflating revenues by the domestic price index and implementing the estimation procedure from Gandhi et al. (2020) will lead to biased estimates of output elasticities and LBE effects. To see this, we write the second step estimation equation of Gandhi et al. (2020) in our

[27] Alternatively, we could exploit a shift-share instrument instead of $\ln \widehat{D}_{f,t-2}$ in (33), where the weights would be pre-period market shares for each firm and the shocks would reflect industry-destination-period demand B_t. However, this would require knowledge of the entire destination network of each firm and measures of aggregate demand. We prefer to use $\ln \widehat{D}_{f,t-2}$ as the instrument because it requires only knowledge of the domestic share in revenues and it allows persistence in the ϵ^d_{ft} draws to contribute to the relevance of the instrument.

notation, assuming that there are in fact multiple destination markets:

$$\begin{aligned}\widetilde{r}_{ft}^{NC} &= \alpha + \sum_{j\in\{1,\ldots,\mathscr{K}\}} b_{\kappa^j}\kappa_{ft}^j + \sum_{j\in\{1,\ldots,\mathscr{K}\}}\sum_{z\in\{\kappa^j,\ldots,\kappa^{\mathscr{K}}\}} b_{\kappa^j z}\kappa_{ft}^j z_{ft} \\ &+ \epsilon_{ft}^1 + (1-\rho)\,u_{ft}^1 + \rho\omega_{ft} + \beta^D\left(\ln D_{ft} + \ln B_t^1\right). \end{aligned} \tag{36}$$

with $\widetilde{r}_{ft}^{NC} \equiv \widetilde{r}_{ft} - \ln\Lambda_t^1$, where Λ_t^1 is the empirical analogue to the true CES price index in the domestic market, and α is a constant that absorbs the industry-specific normalization of the price index.[28] Inspecting (36), we see that both $\ln D_{ft}$ and $\ln B_t^1$ influence the residual ξ_{ft} as constructed in the second step from Gandhi et al. (2020). The aggregate term $\ln B_t^1$ can be controlled for by time fixed effects, but the firm-specific demand shifter $\ln D_{ft}$ can not. Since $\ln D_{ft}$ depends on quasi-fixed input levels, failure to control for $\ln D_{ft}$ implies a violation of the second step moment conditions.[29]

The violation causes biases in ways that are hard to determine and likely depend on parameter values. For example, $\ln D_{ft}$ depends positively on quasi-fixed inputs, since higher quasi-fixed input levels lead to lower marginal costs, higher marginal revenues, higher likelihood of exporting to any given destination, and hence higher export share. Leaving $\ln D_{ft}$ for the error term will thus tend to generate upward bias in the b_j coefficients (ignoring the bias stemming from $\ln B_t^1$).[30] But since the true σ_{ft}^k depends on b_j terms and ρ (see equation (34)), the overall effect on $\widehat{\sigma_{ft}^k}$ is not clear, because implicitly setting $\rho = 1$ will tend to bias *downward* $\widehat{\sigma_{ft}^k}$. The two sources of bias work in opposite directions, and we cannot in general determine which force dominates.[31]

[28] In fact, even if there are multiple destination markets and only revenues are observed, the moment condition for the factor shares first step NLLS from Gandhi et al. (2020) holds. Hence, the bias enters only in the second step. This is because the empirical steps outlined in section 3.1 are exactly the same steps outlined in Gandhi et al. (2020), though the interpretation of the estimated objects differs. The point of section 3.1 was to prove that the moment condition for the NLLS holds even if there are multiple destination markets.

[29] Additionally, it is not possible for $\ln D_{ft}$ to follow the same AR(1) as ϵ_{ft}^1, since $\ln D_{ft}$ depends inversely on ϵ_{ft}^1. So $\ln D_{ft}$ cannot simply be absorbed into ν_{ft} either.

[30] If there is only one quasi-fixed input that enters linearly in (36), then the bias is clearly positive. With multiple quasi-fixed inputs and higher order terms and interactions, it is not clear that omitting $\ln D_{ft}$ leads to upward bias in all estimated b_j terms.

[31] Since the moment conditions for the first-step NLLS holds regardless of the number of markets, the no demand correction estimator should lead to a downward bias in the estimated elasticities for flexible inputs, simply because—given the model—the revenue elasticity is inclusive of $\rho < 1$ (see equation (35))

3.4 Factor share method with a single-market correction

The few papers that explicitly address the value versus quantity distinction in the context of production function estimation implement some version of the Klette & Griliches (1996) procedure, which calls for including a proxy for the CES quantity index in the second step GMM (De Loecker, 2011; Grieco et al., 2016). We present a version of this model here and later estimate it in both simulated data and in the French data for comparison.

In Appendix O6-4, Gandhi et al. (2020) present the Klette & Griliches (1996) approach adapted to the factor shares method.[32] The first step NLLS estimation is exactly the same as in section 3.1, though the interpretation of the estimated objects differs. Moreover, whether or not there are multiple destination markets, the moment condition for this NLLS estimation holds (see footnote 28).

In the second step, Gandhi et al. (2020) introduce a proxy for the CES quantity index, which in their model is unique, because they posit only a single market. We call this aggregate quantity B_t^{proxy}. Defining $\widetilde{r}_{ft}^{KG} \equiv \widetilde{r}_{ft} - \ln \Lambda_t$, where Λ_t is the empirical price index, the second step estimation equation can be written as

$$\widetilde{r}_{ft}^{KG} = \alpha + \beta^D \ln B_t^{proxy} + \sum_{j\in\{1,...,\mathscr{K}\}} b_{\kappa^j} \kappa_{ft}^j + \sum_{j\in\{1,...,\mathscr{K}\}} \sum_{z\in\{\kappa^j,...,\kappa^{\mathscr{K}}\}} b_{\kappa^j z} \kappa_{ft}^j z_{ft} + \nu_{ft}^{KG} \tag{37}$$

where α absorbs the industry-specific price index normalization, and $\nu_{ft}^{KG} \equiv \rho\omega_{ft} + \epsilon_{ft}^1 = h\nu_{f,t-1} + \rho\mu e_{f,t-1} + \xi_{ft}$. For any candidate vector $\left(\beta^{D^*}, b_{\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}}^*, b_{\kappa^1\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}^*\right)$, we can compute

$$\widehat{\nu_{ft}^{KG} + \alpha} = \widetilde{r}_{ft}^{KG} - \beta^{D^*} \ln B_t^{proxy} - \sum_{j\in\{1,...,\mathscr{K}\}} b_{\kappa^j} \kappa_{ft}^j - \sum_{j\in\{1,...,\mathscr{K}\}} \sum_{z\in\{\kappa^j,...,\kappa^{\mathscr{K}}\}} b_{\kappa^j z} \kappa_{ft}^j z_{ft} \tag{38}$$

regress on $\widehat{\nu_{f,t-1}^{KG} + \alpha}$ and $e_{f,t-1}$, compute the residual $\widehat{\xi}_{ft}\left(\beta^{D^*}, b_{\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}}^*, b_{\kappa^1\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}^*\right)$ and build the moment conditions

$$E\left\{\widehat{\xi}_{ft}\left(\beta^{D^*}, b_{\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}}^*, b_{\kappa^1\kappa^1}^*, ..., b_{\kappa^{\mathscr{K}}\kappa^{\mathscr{K}}}^*\right) \begin{pmatrix} \ln B_t^{proxy} \\ \kappa_{ft}^1 \\ \vdots \\ \left(\kappa_{ft}^{\mathscr{K}}\right)^2 \end{pmatrix}\right\} = 0. \tag{39}$$

In the case that there is actually only one destination market (and the empirical price index

[32] De Loecker (2011) andGrieco et al. (2016) use control function methods.

and B_t^{proxy} are computed in a theory-consistent way, see appendix B) we have $\Lambda_t = \Upsilon_t/\Upsilon_0$ and $B_t^{proxy} = B_t/\Upsilon_0$, where Υ_0 captures the price index normalization. In this case, $\widehat{\xi}_{ft}$ is orthogonal to quasi-fixed inputs in period t because at the true parameter values $\xi_{ft} \equiv \widetilde{\epsilon}_{ft} + \rho\widetilde{\omega}_{ft}$.[33] Moreover, the aggregate demand shifter $\ln B_t^{proxy}$ is orthogonal to $\widehat{\xi}_{ft}$ by assumption. Hence, the parameter β^D is identified by time series variation in industry-wide demand aggregates. Thus, in the case that there is only one output market, this estimation procedure identifies the demand parameter $\rho = 1 - \beta^D$, as well as all output elasticities.

However, in the case that there are, in fact, multiple destination markets into which firms select endogenously, then the moment conditions (39) do not hold. To see this, we re-write (27), moving the price index for the domestic market to the left hand side,

$$\begin{aligned}\widetilde{r}_{ft}^{KG} &= \rho E[\ln\psi] + \beta^D \ln B_t^{proxy} + \sum_{j\in\{1,\ldots,\mathscr{K}\}} b_{\kappa^j}\kappa_{ft}^j + \sum_{j\in\{1,\ldots,\mathscr{K}\}}\sum_{z\in\{\kappa^j,\ldots,\kappa^{\mathscr{K}}\}} b_{\kappa^j z}\kappa_{ft}^j z_{ft} \\ &+ \epsilon_{ft}^1 + (1-\rho)\, u_{ft}^1 + \rho\omega_{ft} + \beta^D \ln D_{ft} + \beta^D\Big(\ln B_t^1 - \ln B_t^{proxy}\Big). \qquad (40)\end{aligned}$$

Multiple sources of bias arise in (40). First, unless $\ln D_{ft}$ follows exactly the same AR(1) process as ϵ_{ft}^1 and ω_{ft} (which we argued above is not possible, given the model), then $\widehat{\xi_{ft}}$—as constructed via the KG approach—includes $\ln D_{ft}$. Since $\ln D_{ft}$ depends on quasi-fixed inputs, this implies a violation of the moment conditions in (39). If $B_t^{proxy} \propto B_t^1$, i.e. measured without error, then the omission of $\ln D_{ft}$ from the estimation equation will tend to bias the estimator for all b_j coefficients upward and bias the estimator for β^D downward, as $\ln D_{ft}$ correlates positively with all quasi-fixed inputs, and negatively with $\ln B_t^1$.[34] However, even in this case, the effect on $\widehat{\sigma_{ft}^k}$ is not clear, because $\widehat{\sigma_{ft}^k}$ depends directly on b_j coefficients and inversely on $\widehat{\rho}$. Given the preceding argument, we expect all these coefficients to be biased upwards, which has thus an ambiguous effect on $\widehat{\sigma_{ft}^k}$.[35]

Second, if there are multiple markets, then the demand shifter B_t^{proxy} is likely measured with error. Gandhi et al. (2020) do not explain exactly how to construct B_t^{proxy} from the data, but they cite De Loecker (2011), who proposes to set B_t^{proxy} equal to the weighted sum of deflated *total* revenues of domestic firms. We show in Appendix B that if the price

[33] When all firms serve a single market, the *ex post* demand shock u_{ft} is identified by the factor share regression in the first step, and hence does not appear in the second step.

[34] As in the case of the no demand correction estimator, this holds if there is only one quasi-fixed input that enters (40) linearly. If there are multiple quasi-fixed input that enter (40) along with higher order terms and cross terms, then it is not clear in which direction the bias goes.

[35] Since the moment conditions for the first-step NLLS holds regardless of the number of markets, the KG estimator should lead to a downward bias in the estimated elasticities for flexible inputs, simply because – given the model – the revenue elasticity is inclusive of $\rho < 1$ (see equation (35))

deflator is constructed in a theory-consistent way, then the domestic quantity index B_t^1 can be constructed up to a normalization from price deflators and *total domestic absorption*, i.e. total domestic sales of domestic firms plus total imports from foreign firms. If, instead, B_t^{proxy} is constructed from total revenues of domestic firms (either using weights or not), then B_t^{proxy} will not be equivalent to B_t^1, even up to a normalization. The difference between the two is the trade deficit, which is relegated to the error term, multiplied by β^D. The trade deficit may be positively or negatively correlated with B_t^{proxy}, depending on whether local demand shocks or foreign supply shocks dominate, which means that measurement error in B_t^{proxy} can lead to violations of (39), and it may be difficult to predict in which direction the measurement error biases estimates.

In light of these concerns, a tempting strategy would be to estimate the factor shares method with a single-market correction for a set of non-exporters. For all non-exporting firms, $D_{ft} = \exp(-u_{ft}^1)$ because all sales are domestic ($R_{ft} = R_{ft}^1$), and there are only domestic shocks ($\varphi_{ft} = -u_{ft}^1$). In this case, although the error term in (27) follows an AR(1) process and $\ln D_{ft}$ drops out, bias persists for two reasons. First, B_t^{proxy} is still likely measured with error. Second, sample selection bias violates the orthogonality conditions in (39): the residual from the AR(1) does not have a zero mean conditional on quasi-fixed inputs. If higher levels of quasi-fixed inputs are associated with a greater probability to export (e.g., due to increasing returns to scale), then the conditional mean of the residual will be negatively correlated with them because the sample never admits exporters, and this innovation may induce a firm to export due to cross-market complementarities. The direction of the bias may be different under different cross-market complementarities and overall returns to scale.

Whether estimating the KG model in the full sample or in a sub-sample of non-exporters, two additional problems arise. First, identification in this model relies on time-series variation in aggregate demand, which may not be sufficient in short panels. Second, given our application—the productivity effects of learning by exporting—using the factor share method with a single-market correction entails an additional conceptual issue: there is no exporting in a single market model. Of course, the model can be estimated in the data, because in fact firms do export. But there is a logical inconsistency in positing a single destination and then studying the effect of serving different markets.

3.5 Control function method

The multi-destination model could also be estimated using an amended control function method, which we specify in Appendix C. There are two reasons why we prefer the factor shares method.

First, the control function method relies on time series variation in material input prices for identification. Gandhi et al. (2020) show that when input price variation is low, the control function method suffers from finite sample bias. This is because with low material price variation, the lag of materials is a weak instrument for contemporaneous materials, conditional on productivity and capital. We replicate this finding in Section 4.1 in Monte Carlo simulations for a single-market version of the model.

Second, even when the sample size goes to infinity, the GMM objective function admits multiple solutions in the standard control function framework, as demonstrated recently by Ackerberg et al. (2020). Ackerberg et al. (2020) argue that choosing among these candidate solutions is not as simple as just choosing the parameter combination that yields the lowest objective function value. This is because there are, in fact, multiple parameter vectors for which the moment conditions are satisfied and the objective equals zero. Hence, the optimization problem is under identified.[36] Ackerberg et al. (2020) argue that additional moment restrictions are necessary for identification in the control function method, and they propose a set of such moments, but this procedure is still work in progress.

Our amended multi-market version of the control function method may be less prone to the weak instrument critique of Gandhi et al. (2020) because it introduces cross-firm variation in addition to the time series variation. This motivates us to estimate this amended version of the control function method, in addition to the factor shares procedure that we develop. However, cross-firm variation does not address the "weak moments" problem highlighted by Ackerberg et al. (2020).

4 Monte Carlo simulations

In this section, we study the consistency and finite sample properties of the different estimators presented in Section 3 using Monte Carlo simulations.

In the first set of experiments, we simulate data as described in Section 2 assuming

[36] It is well known that nonlinear estimation like GMM can be sensitive to initial values as well as searching algorithms (Knittel et al., 2014). As shown by Ackerberg et al. (2020), the problem with the control function method is more severe than mere numerical challenges.

there is just a single destination market. With these simulated data, we estimate output elasticities and the curvature of the demand function using the factor shares single-market estimator from Section 3.4 and a single-market version of the control function method that is described in Appendix C. These simulations extend the Monte Carlo experiments from Gandhi et al. (2020) and Ackerberg et al. (2020) to the case of heterogeneous products with missing output price data, and highlight the advantages of the factor share method over the control function approach.

In the second set of experiments, we simulate the multi-destination model from Section 2. With these simulated data, we estimate output elasticities, the curvature of the demand function, and LBE using the factor shares multi-market method from Sections 3.1-3.2, the factor shares approach with no demand correction from Section 3.3, and the factor shares approach with a single-market correction from Section 3.4. With multiple destination markets, only the factor shares multi-market model should be consistent.

4.1 Single market simulations

We report results from Monte Carlo experiments in which the data generating process is described in Section 2 for just a single destination market. We simulate 100 samples of a single industry with 500 firms over 50 periods.[37]

For the data generating process, we impose that firms produce with a Cobb-Douglas production function with one flexible input, materials (M), and one quasi-fixed input, capital (K):

$$Q_{ft} = \exp\left(\omega_{ft}\right) M^{\gamma^M} K^{\gamma^K} \tag{41}$$

with $\gamma^M = 0.8$ and $\gamma^K = 0.3$. Capital updates each period according to the law of motion: $K_{ft} = 0.9K_{f,t-1} + \iota_{f,t-1}$, where $\iota_{ft} = \exp(0.8\rho\omega_{ft} + 0.8\epsilon_{ft})\left(K_{ft}\right)^{0.2}$. We fix $\rho = 0.8$. While we build the data according to these restrictions for simplicity, we obviously need not impose any functional form in the estimation.

Within each replication we draw total expenditures and quantity series, and homogeneous (across firms) material input prices. At the firm level, we draw initial capital stocks $K_{f,1} \sim U(1, 201)$, initial productivity shocks $\omega_{f,1} \sim N(0, 0.01)$, and initial *ex ante* demand

[37] In the single-market case, we follow closely the experiments presented in Gandhi et al. (2020). Gandhi et al. (2020) posit a long panel (50 periods) in order to give the control function a reasonable chance to identify the structural parameters. Since the output elasticity for materials is identified purely from time-series variation in the material input price, there is little chance that the control function identifies structural parameters in panels of only 10-15 years (the type of duration one usually observes in balance sheet datasets). In the multi-market simulations below, we can entertain much shorter panels.

shocks $\epsilon_{f,1} \sim N(0, 0.0009)$. We let ω and ϵ update according to the same AR(1) process described in (28) and (29), with $h = 0.8$ and where $\widetilde{\omega}_{ft} \sim N(0, 0.01)$ and $\widetilde{\epsilon_{ft}} \sim N(0, 0.0009)$. We draw *ex post* demand shocks $u_{ft} \sim N(0, 0.0009)$. Firm-period quantities, revenues and inputs M_{ft} are determined given productivity, capital, materials prices and aggregate demand.

We estimate in each sample of simulated data the factor shares approach and the control function approach assuming researchers observe R_{ft}, M_{ft}, K_{ft}, W_t^M and B_t. For the factor shares model, we set initial conditions for the first-step NLLS estimation for M based on an OLS estimation of the regression

$$\ln\left[\frac{W^m_{ft}M_{ft}}{R_{ft}}\right] = g^m_0 + g^m_m m_{ft} + g^m_k k_{ft} + g^m_{mm} m_{ft} m_{ft} + g^m_{kk} k_{ft} k_{ft} + g^m_{mk} m_{ft} k_{ft} + \vartheta_{ft},$$

where ϑ_{ft} is a regression residual. For the second step GMM, we set initial conditions based on an OLS estimation of the regression

$$\widetilde{r}_{ft} = g_k k_{ft} + g_{kk} kk_{ft} + g_D B_t + \vartheta'_{ft},$$

where ϑ'_{ft} is a regression residual and B_t is the true CES quantity index.

For the control function approach, we set initial conditions for the second-step GMM based on an OLS estimation of the regression

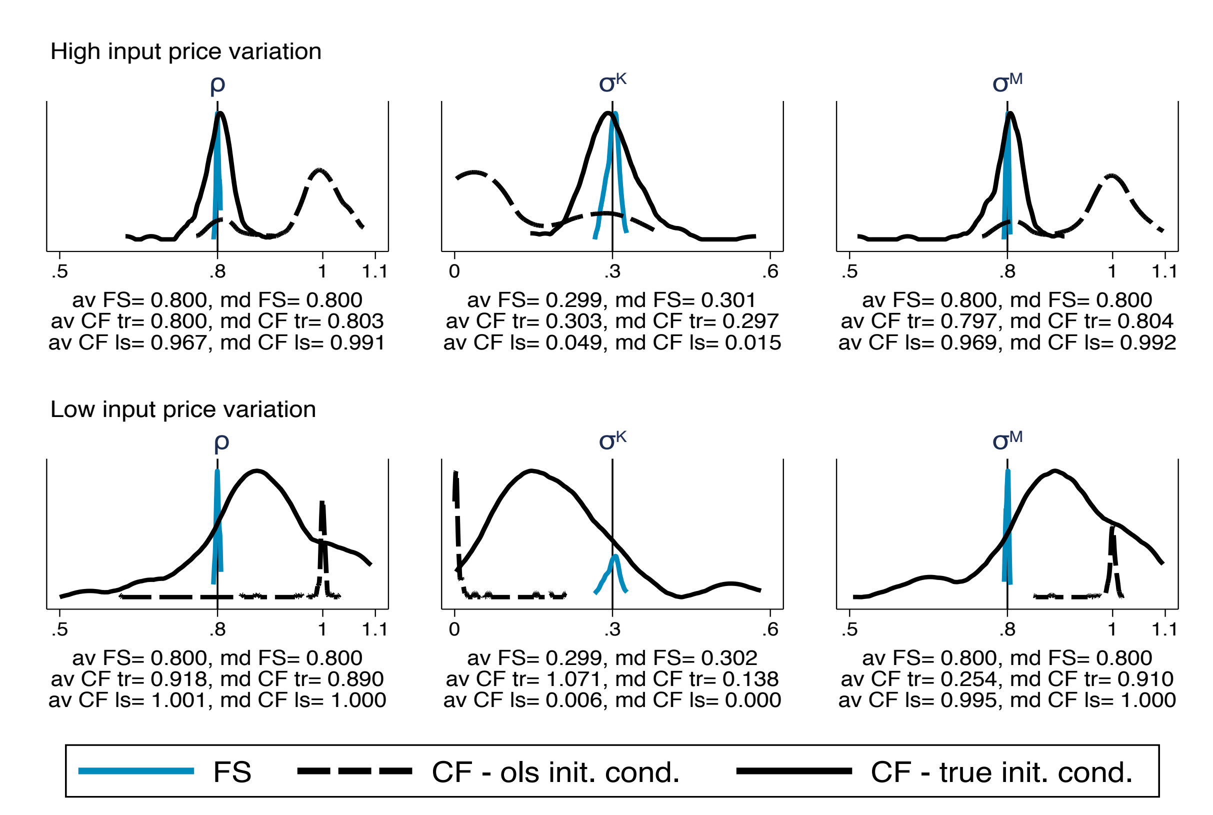
$$\widetilde{r}^{CF}_{ft} = g_0 + g_D B_t + g^m_m m_{ft} + g^m_k k_{ft} + g^m_{mm} m_{ft} m_{ft} + g^m_{kk} k_{ft} k_{ft} + g^m_{mk} m_{ft} k_{ft} + \vartheta''_{ft},$$

where $\widetilde{r}^{CF}_{ft}$ represents log revenues net of the residual from the control function first step, and ϑ''_{ft} is a regression residual. We denote the resulting control function estimates as "CF ls", since the GMM starts from the OLS point estimates. We also start the control function estimation from the true parameter values and refer to the resulting estimates as "CF tr".

In Figure 1 we present the distribution of estimates of ρ and the average material and capital output elasticities across the 100 samples by estimator, along with the average ("av") and median ("md") of the distributions. In the top row we present the case of high input price variation. True material and capital output elasticities are constant across firms and over time ($\sigma^M = \gamma^M = 0.8$ and $\sigma^K = \gamma^K = 0.3$), and are depicted with vertical black lines.[38]

[38] Estimated material and capital output elasticities vary both due to sampling error and with the level of capital because we allow for higher order terms in capital and interactions between inputs in the estimation

Figure 1: Parameter Estimates in the Single-Market Simulations



Notes. The figure reports the distribution of averages of estimates across 100 Monte Carlo samples. Top (bottom) row presents results for high (low) input price variation. True parameter values are depicted as vertical lines. Averages ("av") and medians ("md") of distributions for each estimator are reported below each subfigure. "FS" indicates the factor shares method. "CF - ls" indicates control function method where the GMM optimization starts from the OLS values. "CF - tr" indicates control function method where the GMM optimization starts from the true model parameters.

The distribution of the estimates from the factor shares approach is depicted in solid blue. For each empirical object ρ, σ^K , σ^M, the distribution of estimates appears to be centered on the true values. Averages and medians of the distributions are identical with the truth out to at least two decimal places. Similarly, the distribution of the control function estimates *taking true values as the initial conditions* (black solid line) also appears to be centered on the truth, with averages and medians of distributions identical to the truth out to at least two decimal places. The distribution of the solid blue line is clearly narrower than the distribution of the solid black line, indicating that the factor shares approach is more efficient.

process (see equations (34) and (35)).

In contrast, when the second step of the control function method starts the optimization algorithm from the OLS values (black dashed line), the distributions of estimates are clearly biased. Estimates of ρ and σ^M tend to center around 1, and estimates of σ^K center around 0. These values coincide with the results from a naïve OLS estimate of the production function, where transmission bias tends to bias upward the estimator for output elasticities of flexible inputs, and tends to bias downward output elasticities of quasi-fixed inputs.[39]

Comparing the solid black line to the dashed black line, it is clear that the control function is sensitive to initial starting conditions. We explore this sensitivity further by plotting the distribution of parameter estimates from the control function method when varying systematically the initial values of the second step of the GMM procedure (Appendix D). The results clearly demonstrate that the control function method is highly sensitive to initial parameter guesses. Moreover, there appears to be multiple modes of the distributions. This multi-modal pattern results from the existence of multiple solutions to the GMM optimization (Ackerberg et al., 2020).[40]

Also in Figure 1, we present in the bottom row the distribution of estimates for the case of low input price variation. The factor shares method still recovers unbiased estimates of structural parameters. However, the estimates from the control function method are biased even when the GMM optimization starts from the true parameters. As explained by Nelson & Startz (1990), instrumental variables estimators are biased towards OLS in finite samples with weak instruments. We can see this pattern from the medians of the distributions in solid black. The distributions are wide, and outliers severely distort the

[39]The OLS result can be demonstrated analytically for the case of Cobb-Douglas production. To see this, consider equation (21) and impose Cobb-Douglas production with a single flexible input (materials) and a single quasi-fixed input (capital). We have then

$$\ln[(M_{ft}W_t^M)/R_{ft}] = \ln\Big[\exp(-E[\ln(\psi)])E[\exp(u)]\,\rho\gamma^M\Big] + \varphi_{ft}$$

and rearranging, we get

$$\ln R_{ft} = \theta_t + \ln M_{ft} - \varphi_{ft}$$

where the θ_t absorbs all constant terms and the homogeneous material price. We see from the last equation that $\ln M_{ft}$ enters with a coefficient of 1, and $\ln K_{ft}$ and $\ln D_{ft}$ enter with a coefficient of 0 (they drop out of the equation). Hence, if we regresses $\ln R_{ft}$ on $\ln M_{ft}$, $\ln K_{ft}$, and $\ln D_{ft}$ along with time fixed effects, we would expect to find regression coefficients of 1, 0, and 0, respectively, which imply returns to scale of 1. Ackerberg et al. (2015) made a similar point with respect to the non-identification of the elasticity of output with respect to labor in the first step of the Levinsohn & Petrin (2003) control function estimator.

[40]The estimates using the factor share method are invariant to starting conditions out to several decimal places. Both steps of the factor share method rely on non-linear optimization algorithms, so numerical error can produce some tiny variation in the results. But the variation is so small that it is not worth presenting in a figure.

means, but the medians indicate that the estimates of σ^M are biased up and the estimates of σ^K are biased down, as they are in OLS. This is the same pattern found in Monte Carlo simulations by Gandhi et al. (2020) for the single-market case in which quantities are observed.

4.2 Multi market simulations

In this section we present results from Monte Carlo simulations in which firms serve multiple destination markets. For these experiments, in order to keep the computational burden manageable, we posit 4 destination markets. We simulate 100 samples of a single industry with 2,000 firms over 6 periods. Industry-wide expenditures and quantity indices are destination specific and drawn in the same fashion as in the single-market case. Capital, productivity, and demand shocks are all simulated as in the single-market case, though in this case demand shocks are firm-destination-period specific.

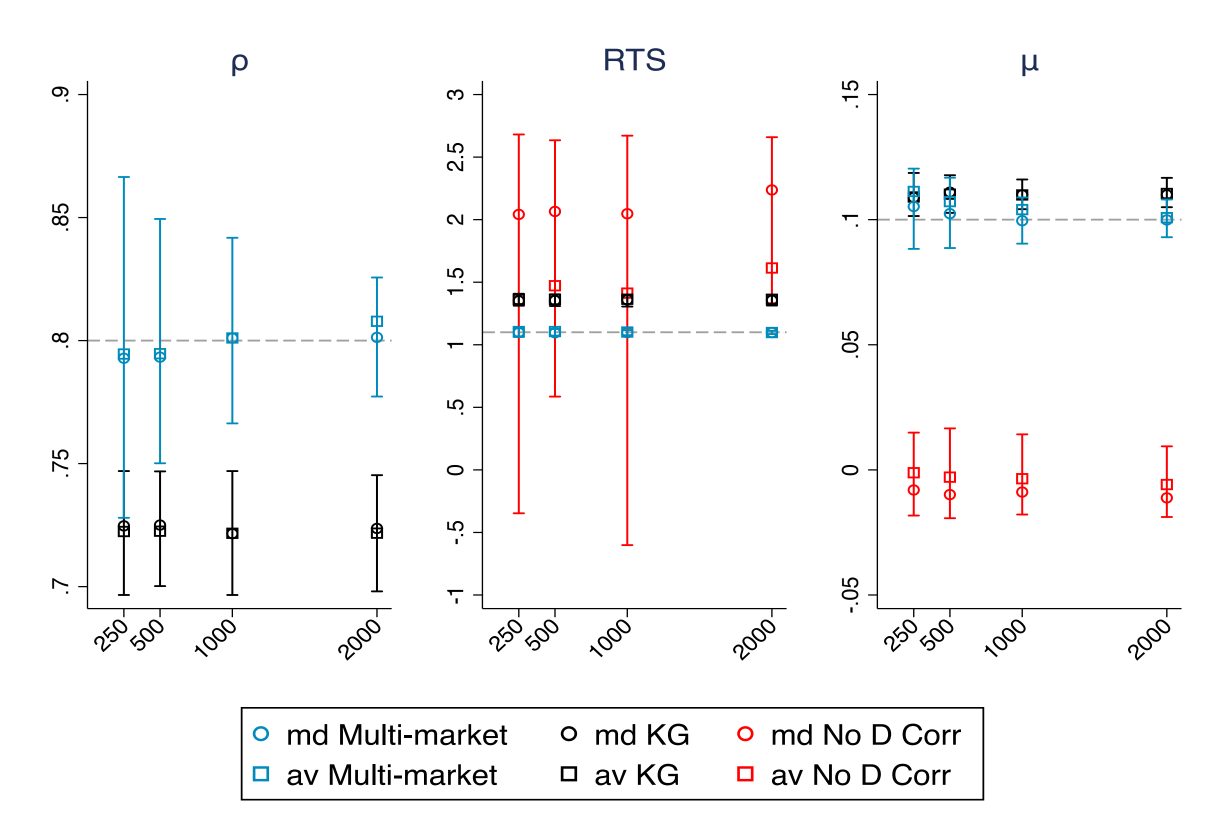
Fixed costs of reaching the foreign markets rationalizes heterogeneous participation in the export market. There are no fixed costs of serving the domestic market ($d = 1$), whereas fixed costs of entry to foreign markets are drawn from a log normal distribution with mean 6 and standard deviation 0.6. Taking expectations over the *ex post* demand shocks u^d_{ft}, firms choose the combination of destinations that yields the highest expected profits.

We simulate the model period by period. In the first period, we solve for the set of destinations that maximizes expected profits for each firm. From these values, we determine who is active on the export market. We then update firm productivity for period 2 (which includes the LBE effect), setting the learning-by-exporting coefficient $\mu = 0.05$. Given $\omega_{f,2}$ and $K_{f,2}$, we then solve the combinatorial problem for each firm in period 2. We again determine which firms are active on the export market in period 2, and update firm productivity accordingly. We continue in this fashion until the final period.

With these simulated data, we estimate output elasticities, the curvature of demand and the LBE parameter using our factor shares multi-market method, the factor shares method with a single-market demand correction, and the factor shares method with no demand correction. We study the consistency of these estimators in Figure 2. The Figure reports for each estimator the mean, median and inter-quartile range of estimates of μ, ρ and returns to scale $(\widehat{\sigma}^K_{ft} + \widehat{\sigma}^M_{ft})$ for increasingly larger sample sizes. The true parameter values are indicated by dashed horizontal lines. We can see that our multi-market estimator, depicted in blue, is consistent, while the estimator with no demand correction ("No D

Corr", in red) and the estimator with a single-market correction ("KG", in black) are not.

Figure 2: Consistency Properties in the Multi-Market Simulations

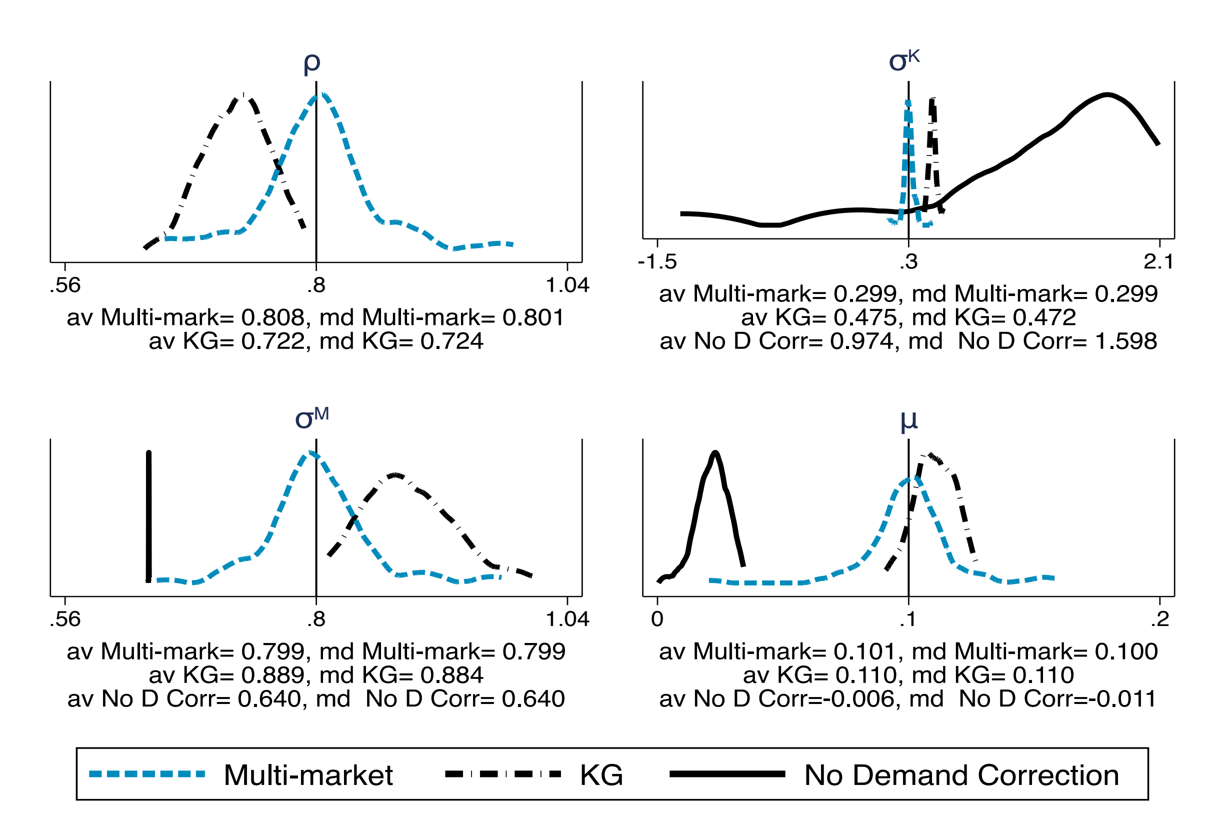


Notes. The figure reports the mean, median and inter-quartile range of three estimators across 100 simulations of the multi-market model with 250, 500, 1,000 and 2,000 firms each. The true parameter values are depicted as horizontal dashed lines. The first estimator is our multi-market estimator (blue); the second includes a correction for demand in a single market("KG", black); the third makes no correction for demand ("No D Corr", red). Means are denoted by squares, medians are denoted by circles, and the inter-quartile ranges are denoted by the bars.

Figure 3 presents finite sample properties of the estimators. We find that the distribution of estimates from our multi-market factor shares method (dashed blue line) is centered on the true parameters, with means and medians quite close to the true values. For the single-market factor shares method (dashed black line labeled "KG"), the distribution of the estimates of ρ is biased downwards and the distributions of estimates of σ^K and σ^M are biased upwards. The estimator that assumes no correction for demand yields results that center far away from the truth.

Finally, we investigate inference with our estimator compared to the estimator that

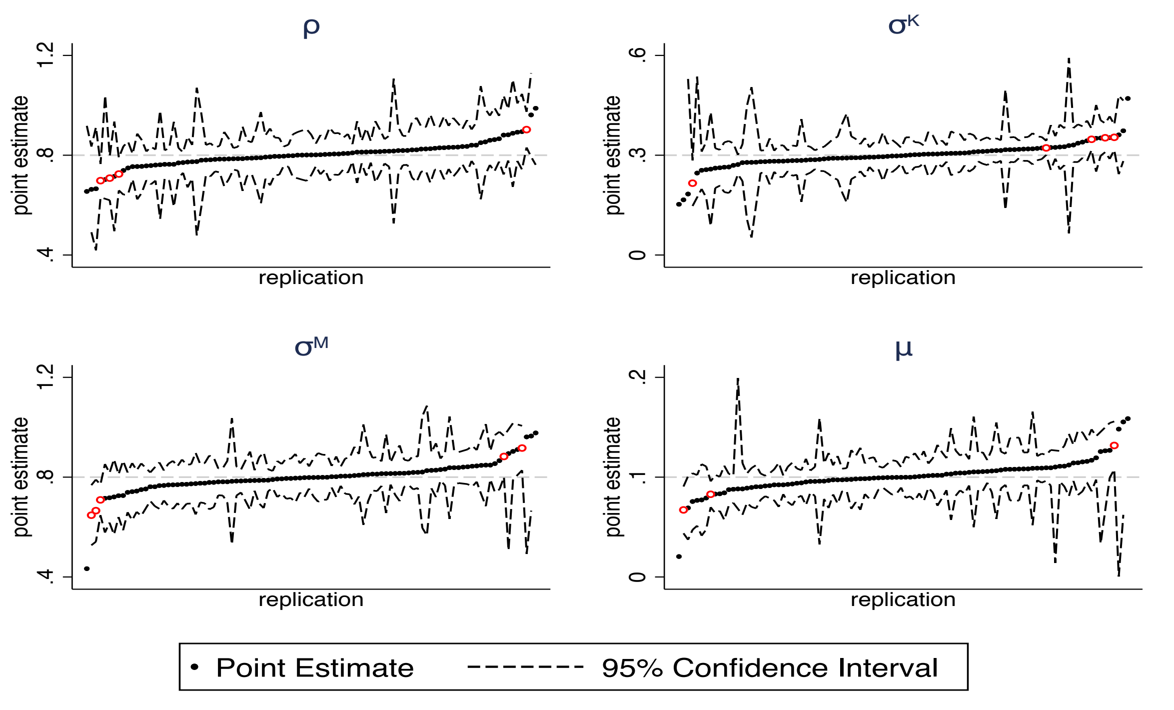
Figure 3: Finite Sample Properties in the Multi-Market Simulations



Notes. The figure reports the distribution of three estimators across 100 simulations of the multi-market model with 2,000 firms each. True parameter values are depicted as vertical lines. The first is our multi-market estimator ("Multi-mark") ; the second includes a correction for demand in a single market as in Klette & Griliches (1996) using the factor shares approach ("KG"); the third estimator makes no correction for demand ("No D Corr"). Averages ("av") and medians ("md") of distributions for each estimator are reported below each sub-figure.

corrects only for demand in a single market. We compute for each replication the bootstrapped 95% confidence interval and investigate the proportion of replications in which the confidence interval contains the true parameter value. In Figure 4 , we plot the estimated value and the 95% confidence interval by replication and parameter. Black dots indicate replications for which the 95% confidence interval includes the true value, and red open dots indicate replications for which the 95% confidence interval excludes the true value. As expected, our estimator's 95% confidence intervals include the true parameter value about 95% of the time. In contrast, Figure 5 illustrates that the 95% confidence intervals for the estimator with a single-market correction rarely include the true parameter.[41]

Figure 4: Coverage Ratios for the Multi-Market Estimator



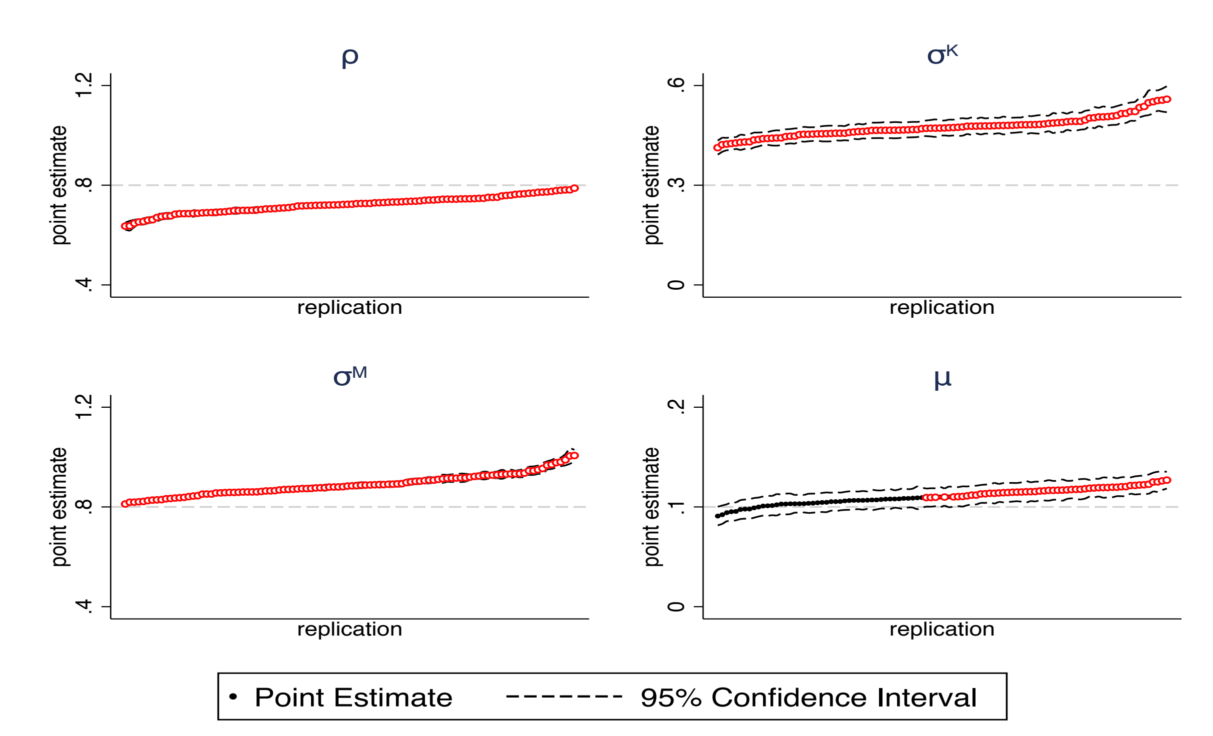
Notes. The figure reports point estimates and 95% confidence intervals of our multi-market estimator across 100 simulations of the multi-market model with 2,000 firms each. Solid dots mark point estimates for which the true parameter value lies within the 95% confidence interval. Red circles mark point estimates for which the true parameter value lies outside of the 95% confidence interval. All confidence intervals are computed using bootstrapped standard errors.

5 Application to French manufacturing

In this section we describe our data sources, report essential descriptive statistics, and then report estimates of returns to scale, the elasticity of demand, the output elasticities of inputs, and learning-by-exporting effects for French manufacturing firms. We report

[41]We note that standard errors of the estimates from the single-market demand correction estimator, especially for ρ are quite small. This is because we estimate the model using the true, exogenous process for B_t, which is common to all firms and across all replications (only productivity draws across firms differ within and across replications).

Figure 5: Coverage Ratios for the Estimator with Single-Market Correction



Notes. The figure reports point estimates and 95% confidence intervals of the *single*-market estimator across 100 simulations of the *multi*-market model with 2,000 firms each. Solid dots mark point estimates for which the true parameter value lies within the 95% confidence interval. Red circles mark point estimates for which the true parameter value lies outside of the 95% confidence interval. All confidence intervals are computed using bootstrapped standard errors.

results for our multi-market estimator, the single-market correction estimator, and the standard estimator that makes no correction for demand. For the main results we assume that labor is predetermined within the period. In the end we discuss results that assume a partial adjustment process for labor and results using the control function method, while relegating detailed results to Appendix F.2 and Appendix F.3, respectively.

5.1 Data and descriptive statistics

We use administrative data sources to build a quasi-exhaustive panel of the universe of French manufacturing firms in 1994–2016. Most of the data comes from firm balance sheets from the FICUS and FARE datasets, which originate in firms' tax declarations. We use total revenues, material expenditures, employment, and book-value of capital stocks. We obtain information on firms' exports from the French Customs. It is straightforward to merge the customs data to FICUS/FARE because they use the same firm-level SIREN identifier. We deflate expenditures on materials by industry-level input price indices that we obtain from the EU KLEMS dataset. We build firm-level capital stocks using the methodology of Bonleu et al. (2013) and Cette et al. (2015). Appendix E provides more information on the source data and explains how we construct firm-level capital stocks starting from book values.

We report descriptive statistics in Table 1. The skewed firm size distribution is apparent from the difference between means and medians, for example, in revenue and employment. This feature is common in many manufacturing datasets. The high percentages of exporting firms is typical of European economies, who trade intensively within Europe. On average, 25% of firms in our data export at least once, but this varies considerably across industries, with a low of 6.6% in "Food, beverage, tobacco", and a high of 71% in "Chemical products".

Table 1: Descriptive Statistics

No.	Industry		Revenue (mn euros)	Labor (employment)	Materials (mn euros)	Capital (mn euros)	Obs.	No. firms	No. exporters	Percent exporters
1	Autos and transport equipment	Mean	44.33	144.55	27.78	18.37	50403	5507	2774	50.4
		Median	1.01	10.00	0.39	0.21				
2	Chemical products	Mean	52.54	97.35	29.95	27.82	52047	4943	3510	71.0
		Median	2.36	14.00	0.94	0.55				
3	Computer, electronics	Mean	11.04	59.99	5.04	4.80	52845	5736	3158	55.1
		Median	0.79	8.00	0.26	0.12				
4	Electrical equipment	Mean	13.25	70.62	7.12	5.14	42476	4584	2321	50.6
		Median	0.98	9.00	0.35	0.13				
5	Food, beverage, tobacco	Mean	3.11	12.74	1.81	1.21	884753	113119	7498	6.6
		Median	0.24	3.50	0.08	0.10				
6	Machinery and equipment	Mean	3.78	22.65	1.73	1.01	323815	34802	11297	32.5
		Median	0.55	5.00	0.17	0.09				
7	Basic metal and fabricated metal	Mean	4.46	27.28	1.93	2.23	352083	33769	12975	38.4
		Median	0.82	9.00	0.16	0.24				
8	Other manufacturing	Mean	1.56	12.15	0.62	0.55	250297	30933	6584	21.3
		Median	0.22	3.00	0.05	0.06				
9	Rubber and plastic	Mean	6.95	39.42	3.08	4.16	163847	16121	7006	43.5
		Median	0.88	8.00	0.30	0.26				
10	Textiles, wearing apparel	Mean	3.31	24.42	1.41	0.98	149369	21384	9139	42.7
		Median	0.49	6.00	0.14	0.08				
11	Wood, paper products	Mean	2.75	17.41	1.21	1.62	295484	32356	9789	30.3
		Median	0.47	5.00	0.11	0.14				
	Total	Mean	5.54	24.84	2.90	2.47	2617419	303254	76051	25.1
		Median	0.38	5.00	0.11	0.11				

Notes. The table reports descriptive statistics for the estimation sample, where capital is the book value reported by the firm and materials are expenditures. Exporters are defined as firms that exported at least once during the sample. Source: FICUS/FARE datasets and French Customs.

In Table 2 we report descriptive statistics for the export intensity among firms that export. Within exporting firms, the export share also varies considerably, both across industries and across firms within industries. While the median exporter obtains 4.2% of revenue from exporting, the 90^{th} percentile firm obtains almost 40% of revenue from foreign markets.

Table 2: Percent exports in revenue for exporters

No.	Industry	Mean	p5	p10	p50	p90	p95
1	Autos and transport equipment	14.6	0.2	0.4	5.8	43.1	55.9
2	Chemical products	22.5	0.2	0.6	11.4	64.6	77.5
3	Computer, electronics	19.5	0.2	0.5	8.3	58.2	74.5
4	Electrical equipment	15.4	0.2	0.5	6.1	46.3	60.1
5	Food, beverage, tobacco	10.1	0.1	0.2	2.8	31.1	48.3
6	Machinery and equipment	12.4	0.1	0.3	4.1	38.8	57.1
7	Basic metal and fabricated metal	10.6	0.1	0.3	3.5	31.7	47.9
8	Other manufacturing	12.8	0.3	0.5	4.8	38.3	52.7
9	Rubber and plastic	11.4	0.1	0.2	3.7	35.7	52.2
10	Textiles, wearing apparel	17.8	0.4	0.8	9.4	48.5	62.2
11	Wood, paper products	7.8	0.1	0.2	1.6	23.8	41.9
	Total	12.8	0.1	0.3	4.2	39.7	56.7

Notes. The table reports the distribution of the percent of exports in revenue for exporters in the estimation sample. Percent exports in revenue for exporters is computed for firms and years in which exports are positive. Source: FICUS/FARE datasets and French Customs.

Tables 1 and 2 make two important points. First, the fact that many firms sell on more than their domestic market (at least once) implies that estimation methods that assume that all sales are on the domestic market ignore important information. In particular, building theory-consistent demand aggregates for B_t^{proxy} in (37) is not feasible using only information from the domestic market. Second, variation in the extensive exporting margin and the high variation in export intensity among exporting firms jointly indicate that there is sufficient variation to identify ρ in our setting, coming from, *inter alia*, the cross section of firms. This is in contrast to methods that assume only one market, where only time series variation identifies ρ.

5.2 Main results

In this section we report results for the factor shares method using our multi-market estimator, the single-market estimator, and the standard estimator that makes no correction for demand.

An important decision when taking the factor shares method to the data is whether to classify inputs as flexible or predetermined. It is quite standard in production function applications to treat capital as a quasi-fixed input and to treat materials as a flexible input. The treatment of labor varies by application. Many applications in the developing world (e.g., Colombia, Chile, Mexico) treat labor like a flexible input. Applications to developed-world data sometimes treat labor as a flexible input, and sometimes treat it as a quasi-fixed input. Presumably, developed economies have stricter labor market regulations, which makes it harder to adjust labor stocks to contemporaneous shocks. With French data, researchers tend to treat labor as a quasi-fixed input (Harrigan et al., 2023). This is the assumption that we adopt for our main specification.[42]

Total Returns to Scale. We present estimates of total returns to scale (RTS) in the top left panel of Figure 6. Detailed estimates are reported in Tables F.1, F.2, and F.3, where we also report the persistence parameter h, the demand curvature parameter ρ, and the long run effect of exporting $\mu/(1-h)$, as well as bootstrapped standard errors for all estimates.

We start with the estimator that makes no demand correction (red triangles) in order to assess the bias in this commonly-used estimator, and to compare to our preferred specification. Measured on the left axis of the upper left panel in Figure 6, we find estimated total returns to scale slightly below 1 for most industries. The mean and median of the industry-specific average returns to scale are both equal to 0.96. These estimates are close to what researchers tend to find with this approach. For example, in their factor shares approach (deflating revenues by industry-wide price indices and interpreting as quantities) Gandhi et al. (2020) find average returns to scale in Colombia between 0.99 and 1.06, and between 1.04 and 1.15 for Chile.

Interestingly, the factor share estimates of total returns to scale from Gandhi et al. (2020) do not differ much from naïve OLS estimates of returns to scale (see Gandhi et al. 2020, Table 2). We find a similar result; OLS estimates of total returns to scale, which we report in appendix Figure F.4 (top left panel) are remarkably similar to a more elaborate estimator that neglects any correction for demand in Figure 6.[43] This reinforces Klette &

[42]We later investigate the sensitivity of the results to the possibility that firms partially adjust labor to contemporaneous shocks; we find that this does not materially alter the main results.

[43]Recall that the naïve OLS is expected to yield estimated total returns to scale of around 1. See footnote 39 for more details.

Griliches (1996) and our argument, that controlling for price variation is essential, even if adequately controlling for supply shocks.

The upper left panel in Figure 6 also reports our estimates of returns to scale using the multi-market estimator (depicted by blue circles, left axis). Average total returns to scale range from 1.05 (Electrical equipment) to 1.22 (Wood, paper products), and for one industry up to 1.37 (Food, beverage, tobacco). The mean (median) estimate across the 11 industries is 1.15 (1.13), which in line with estimates in Antweiler & Trefler (2002), and is substantially higher than corresponding OLS estimates (Figure F.4). As hypothesized by Klette & Griliches (1996), the constant returns to scale estimated by the no demand correction model mask returns that are actually increasing (in our notation, $\sigma_{ft}^M + \sigma_{ft}^L + \sigma_{ft}^k > 1$). Klette & Griliches (1996) argue that ignoring unobserved firm-specific prices would tend to lead to a downward bias in estimated returns to scale, *ceteris paribus*. As we discuss in section 3.3, there are, in fact, several forces that bias the estimator with no demand correction, and the overall sign cannot be determined in general. Nevertheless, the evidence in the upper left panel of Figure 6 is consistent with the central hypothesis from Klette & Griliches (1996).

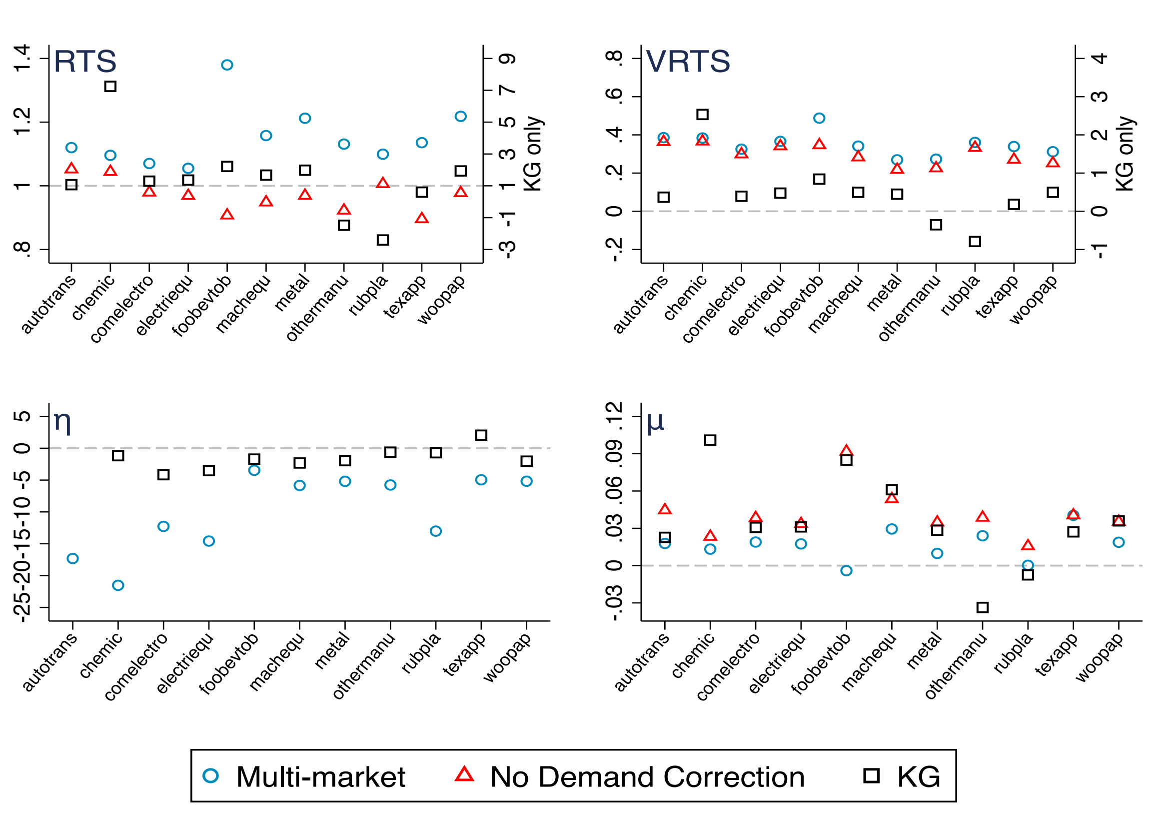
If the data generating process coincides with the multi-destination model from section 2, then the single market estimator from Klette & Griliches (1996) (and more recently, the appendix of Gandhi et al. 2020) does not entirely address the transmission bias stemming from missing output prices. But it remains to be seen how well our version of the single market estimator from Section 3.4 performs in practice. In the top left panel of Figure 6, it is clear that the answer is: not very well. Estimates (depicted by black squares) vary wildly across industries—so much so that we cannot comfortably fit the estimates on the same scale with the other two estimators. Measured on the right axis (KG only), we find that the average returns to scale range from -2.4 (Rubbers and plastics) to 7.3 (Chemicals). For the KG estimator we find only three industries with plausible estimates for returns to scale: Auto and transportation (1.07), Communication electronics (1.29), and Electrical equipment (1.36). Estimated average returns to scale are implausibly high or implausibly low for all other industries.[44]

The large range of estimates for returns to scale is largely due to the range of estimates of ρ. Recall that the KG estimator uses the estimate of ρ to "deflate" the revenue elasticities. With the KG estimator, we estimate a range for ρ from -0.64 to 1.48 across industries. When ρ is estimated to be close to zero, then the returns to scale become very large in

[44] The high and low estimates of average returns to scale is not just a matter of outlier observations either. Medians within the industry are very close to the means.

absolute value, as they do for Chemicals. When the estimate of ρ is negative, this leads to negative estimates of returns to scale (Rubbers and plastics and Other manufacturing). These extreme estimates of ρ are not merely due to estimation uncertainty; the estimates are quite precise (see Table F.3 for bootstrapped standard errors).

Figure 6: Estimates by Industry



Notes. The figure reports factor share estimates of average returns to scale, returns to materials, the demand curvature $\eta = 1/(\rho - 1)$, and the LBE learning by exporting (LBE) parameter μ by industry and estimator. For total and variable returns to scale, the KG estimator is reported on the right axis, while the multi-market and no demand correction estimators are reported on the left axis. Detailed estimates are reported in Tables F.1, F.2, and F.3

In Section 3.4, we show that both transmission bias and measurement error in B_t^{proxy} could bias the KG factor shares estimator when the true data generating process features multiple destinations. Even when the sign of the bias on estimated coefficients is clear, the sign of the bias on estimated returns to scale is ambiguous, since estimated returns to scale is a nonlinear transformation of estimated coefficients with biases of potentially different signs. Consistent with this, there is no discernible pattern in the comparison of the estimated returns to scale with the KG estimator versus the multi-market estimator.

The wide range of KG estimates of returns to scale is notable because we find much smaller range of estimated returns to scale in the Monte Carlo experiments when using the KG estimator. A likely explanation for the difference is that in the Monte Carlo simulations, we assume the researcher observes the true domestic quantity index B_t^1. In the application we follow the common practice to approximate B_t^1 using all of the firms' revenues (not distinguishing exports and domestic sales, nor adding imports) and use price indices built from producer prices in the domestic market. These discrepancies highlight an additional, practical advantage for our multi-market estimator: it does not require making such data compromises, as it does not require building B_t^1, nor does it require to deflate firm revenues.

Given the theoretical drawbacks of applying the KG estimator to a sample of multi-destination firms, we can alternatively apply the KG estimator to a sample of never exporters, as discussed in Section 3.4. In this case measurement error in B_t^{proxy} and selection bias could still lead to biased estimates. Table F.4 reports results for the sample of non exporters, where we find very similar results to the main KG estimates when we do not drop exporters (Table F.3). This suggests that measurement error in B_t^{proxy} is likely the key driver of the wide range of estimates in the KG estimator. This is not surprising, since identification in the KG estimator relies on time series variation in aggregate consumption, for which there are at most 21 observation per industry in our sample. If this key variable is measured poorly—or does not vary much over time—estimates should indeed vary substantially.[45]

In our preferred specification (Figure 6, blue circles) average returns to scale are greater than 1 for all industries. This indicates that there are efficiency gains from size embedded in the technology used by firms, regardless of how total factor productivity evolves. Our estimates imply that returns to scale are increasing for virtually all firm-year observations, not just on average (see F.3).[46] From a welfare perspective, increasing returns imply a cost

[45]Recall that the measurement error is not classical, so the direction of the bias is not necessarily towards zero.

[46]Several mechanisms could explain this phenomenon. The simplest explanation for increasing returns to scale is the presence of fixed costs of operation. Alternatively, complementarities within the firm could generate increasing returns at any point along the firm-size distribution. For example, externalities across workers could lead to increasing returns (e.g., learning by doing), as in for example Kellogg (2011); Hjort (2014), among others.

to diversification that weighs against love of variety, as hypothesized by Krugman (1979). In addition, increasing returns imply larger business cycle fluctuations, and may provide a rationale for targeted interventions during downturns.

Returns to flexible inputs. We turn to estimates of returns to scale for flexible inputs (VRTS), which, in the case that labor is pre-determined each period, are just the output elasticity with respect to materials. The top right panel of Figure 6 reports averages across industries by estimator. The mean (median) of the average estimates across industries is 0.34 (0.35) with the multi-market estimator and 0.30 (0.30) with the no demand correction estimator. Notice that the first step for both estimators is identical—but their interpretation differs. In the no demand correction case the first step identifies directly the output elasticity, whereas in the multi market estimator the first step identifies the revenue elasticity with respect to materials, and must be divided by ρ in order to obtain the output elasticity. Since ρ is estimated to be less than 1 when using the multi-market estimator, the estimated output elasticity of materials is larger.

Returns to flexible inputs well below 1 imply negative cross-market complementarities in the short-to-medium run. For example, a positive demand shock in one market leads to more sales to that market, an increase in marginal costs, lower sales to other markets and a reduction in the likelihood of selling to other markets. This is consistent with findings in Almunia et al. (2021), who argue that the massive negative demand shock in Spain during the financial crisis caused an increase in exporting, presumably due to a reduction in scale and in marginal costs.[47]

Variable returns to scale estimated via the single-market correction method (measured on the right axis, KG only) yield implausibly large or even negative estimates, ranging from -0.79 (Rubber and plastic) to 2.53 (Chemical products), with an average and median around 0.45. As noted above, this variation is mostly due to variation in the estimates of ρ.

Elasticity of Demand. In the bottom left panel of Figure 6 we report estimates of the price elasticity of demand $\eta = 1/(\rho - 1)$. With the single-market correction estimator, most of the demand elasticities fall within the range -4.14 to -0.6, which is a range that is mostly lower (in magnitude) than what people tend to estimate in the literature. For example,

[47] Almunia et al. (2021) perform production function and productivity estimation, but when developing their estimator they ignore cross-market complementarities. We explain how their estimator differs fundamentally from ours in Appendix G.

using data on trade flows and trade costs, Shapiro (2016) estimates an average trade elasticity across industries of -8.16 (see Shapiro 2016, Table 2), which translates into a price elasticity of demand of -9.16.[48] Additionally, with the KG method, we estimate a positive demand elasticity for Textile and apparel ($\eta = 2.05$), and a very high (in magnitude) demand elasticity for Autos and transport equipment (-58.7).[49]

With the multi-market estimator, we estimate a range of demand elasticities from -21.5 (Chemicals) to -3.4 (Food, beverage and tobacco), with no implausible outlier estimates. The mean (median) estimate across industries is -9.9 (-5.8), which is much closer to the mean and median estimates that are typically estimated in gravity regressions (e.g., Shapiro 2016). Bootstrapped standard errors are reported in Table F.1. Standard errors for η become extremely large as ρ approaches 1 (as in the case of electrical equipment). It is more instructive to look at the standard errors for ρ. Here, the bootstrapped confidence interval is very small. For instance, we can easily reject $\rho = 1$ for all industries, which indicates—as expected—that we can reject a homogeneous good hypothesis.

Learning by Exporting. Finally, in the bottom right panel, we report estimates of LBE effects by industry and estimator. With the multi-market estimator, we estimate LBE effects in the range of -0.004 (Food, beverage and tobacco) to 0.040 (Textile and apparel). For the two industries with very low estimated LBE effects (Food, beverage and tobacco and Rubbers and plastics), we cannot reject zero effect; for all other industries, we do reject 0. Across all 11 industries, the mean (median) estimate of LBE is 0.017 (0.018). These estimates are lower than the estimates with the no demand correction estimator (mean = 0.04 and median = 0.038) and the single-market correction estimator (mean = 0.035 and median = 0.030). The comparison suggests that estimated LBE effects are biased upward in the other estimators, possibly because these estimators mistakenly attribute the effect of foreign demand shocks to LBE.

It is often stated in the literature that LBE effects are only found in developing-world firms. With French manufacturing data, we find robust evidence of significant LBE effects, contrary to this perception. With our multi-market procedure, our estimates translate into as much as 40% long run cross-sectional differences in productivity between exporters and non-exporters (Table F.1, last column).[50] These estimates are quite precisely estimated,

[48] In Shapiro (2016), the trade elasticity is equal to 1 minus the elasticity of substitution across varieties, which is equal to $1 + \eta$ in our notation.

[49] We leave the estimate for Autos and transportation equipment out of Figure 6 for ease of reading.

[50] The long run, cross-sectional difference is computed as $\mu/(1 - h)$, where μ is the effect of exporting today on productivity tomorrow and h is the persistence parameter in the AR(1) process for productivity.

which is not surprising given the high number of firm-year observations per industry. Compared to previous work, these effects are smaller than those estimated via RCT with Egyptian firms (Atkin et al., 2017) and via structural approaches with Chilean, Colombian, and Mexican firms, e.g., (Garcia-Marin & Voigtländer, 2019), but larger than estimates from Danish firms using a quasi-natural experiment (Buus et al., 2022).[51]

5.3 Results assuming partial labor adjustment or using the control function method

In this section, we report two additional sets of results. The first allows for dynamic, partial adjustment for labor and serves as a robustness check for the main results. The second set of results applies the control function method, which is just for comparison, as we expect finite sample bias and weak moments problems.

We start with results that allow labor to partially adjust to contemporaneous productivity and demand shocks. This permits entertaining the possibility that firms have the ability to flexibly adjust part of employment, while another part is pre-determined within the period. This is particularly interesting in the context of the French dual labor market, which features both short-term fixed employment contracts and long term indefinite duration contracts. Even though the French dual labor market is known for its rigidity, it is certainly possible that French firms adjust the current labor stock to contemporaneous supply and demand shocks, even if not completely (Saint-Paul, 1996; Reshef et al., 2022). To allow for this possibility, we need only adjust the factor shares second step moment condition (33) to replace all contemporaneous labor measures with lagged measures.

We report detailed results for the four models estimated above (multi-market, no demand correction, single market, and single market with no exporters) in Appendix F.2. The results are quite similar to our main specification, in which we treat labor as quasi-fixed. We estimate slightly higher returns to scale and lower elasticities of substitution for the multi-market estimator, and a slightly larger range of values for LBE (-0.014 to 0.045). Estimates based on the single-market correction estimator still vary wildly by industry, and the estimates of LBE still appear biased up in the two misspecified estimators.

We now turn to report results using the control function approach across the four models

[51]The other notable comparisons in the literature is De Loecker (2013), who estimates LBE effects in the range of 0.017 to 0.066 across Slovenian manufacturing industries. Note that De Loecker (2013) does not correct for demand, so these estimates are best compared to the results from our no-demand-correction estimator, with which we find in a similar range of 0.016 to 0.092.

estimated above in Appendix F.3. The three models that apply some correction for demand (multi-market, single market, and single-market while excluding exporters) yield implausible returns to scale and output elasticities, or implausible demand curvatures, or both. For example, with the multi-market estimator, we estimate *positive* demand elasticities for 5 out of 11 industries (though these estimates are quite imprecise). Estimated returns to scale with the multi-market control function model are also quite close to 1, which is similar to results applying the naïve OLS approach.[52] The single market correction model yields erratic and implausible estimates of both returns to scale and demand elasticities, regardless of whether we exclude exporters. The model with no correction for demand yields mostly plausible estimates of returns to scale (except for Wood and paper products), but with quite low estimates of returns to capital and high returns to materials—a telltale sign of transmission bias. We conclude that the control function approach delivers, in practice, a poor estimator of the production function and demand parameters, which is consistent with the results from the Monte Carlo simulations.

6 Conclusion

Production function estimation is key to many economic analyses, but the conditions assumed in theory rarely match those faced by applied researchers (De Loecker & Goldberg, 2014; De Loecker & Syverson, 2021). In particular, most datasets report output only in values, not in quantities. In addition, many firms serve multiple destination markets, wherein they face heterogeneous demand conditions, which is inconsistent with models that control for a single market-wide demand shifter. This inconsistency is important for estimation, even when researches are not interested in exporting or the effect of exporting *per se*.

In this paper, we show how to estimate output elasticities, the price elasticity of demand, the elasticity of productivity to observable determinants, and productivity itself when firms serve multiple destination markets and when outputs are denominated only in monetary terms. We show that existing production function estimators that use revenue to identify output yield biased and inconsistent inference in this case. Our estimator is no harder to implement than existing methods and requires only one additional piece of information: firms' export shares. Our estimator does not rely on functional form assumptions for the production function, and although it relies on a common industry-wide elasticity of

[52]Since we use initial conditions from naïve OLS estimates of the production functions and demand curvature, we are not surprised to find similar results after performing the GMM search for the non-linear estimator (Ackerberg et al., 2020).

demand, it allows for firm-destination-year prices and markups.

In addition to our main contribution, we confirm results from Gandhi et al. (2020) and Ackerberg et al. (2020): the control function approach requires substantive time-series variation in flexible input prices for identification, and is sensitive to initial conditions. In contrast, our estimator does not have these drawbacks.

We demonstrate the practical advantages of our estimator relative to existing approaches using balance sheet information for the universe of French manufacturing firms. In the French data, we estimate demand elasticities between -21.5 and -3.4, which are in a range that is consistent with much of the literature. We estimate average returns to scale ranging from 1.05 to 1.22 with one industry at 1.37, and average returns to flexible inputs uniformly below 1. The latter result implies cross-market complementarities: additional production for a given market raises the cost of serving all other markets in the short run. Alternative approaches yield implausible estimates of returns to scale or demand curvature, or both. We also estimate learning-by-exporting effects ranging from 0 to 4% per year, which imply cross-sectional differences in productivity between exporters and non-exporters up to 40%.

Overall, the tools that we develop in this paper deliver more credible estimates of production functions in contexts in which price and quantity data are unavailable, and in settings where firms endogenously select into potentially multiple destination markets. It also allows us to study a determinant of productivity such as learning by exporting using a productivity estimation framework that is consistent with heterogeneous firms' export decisions and pricing to market.

References

Ackerberg, D., Frazer, G., Luo, Y., & Su, Y. (2020). Under-Identification of Structural Models Based on Timing and Information Set Assumptions. *Working Paper Available at SSRN: http://dx.doi.org/10.2139/ssrn.3717757*.

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411–2451.

Aghion, P., Bergeaud, A., Lequien, M., & Melitz, M. J. (2022). The heterogeneous impact of market size on innovation: Evidence from French firm-level exports. *The Review of Economics and Statistics*, (pp. 1–56).

Almunia, M., Antràs, P., Lopez-Rodriguez, D., & Morales, E. (2021). Venting Out: Exports during a Domestic Slump. *American Economic Review*, 111(11), 3611–62.

Antweiler, W. & Trefler, D. (2002). Increasing returns and all that: a view from trade. *American Economic Review*, 92(1), 93–119.

Arkolakis, C. & Eckert, F. (2017). Combinatorial discrete choice. *Available at SSRN 3455353*.

Atkin, D., Khandelwal, A. K., & Osman, A. (2017). Exporting and firm performance: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 132(2), 551–615.

Aw, B. Y., Roberts, M. J., & Xu, D. Y. (2011). R&D investment, exporting, and productivity dynamics. *American Economic Review*, 101(4), 1312–44.

Barrows, G. & Ollivier, H. (2021). Foreign demand, developing country exports, and CO2 emissions: Firm-level evidence from India. *Journal of Development Economics*, 149, 102587.

Berman, N., Berthou, A., & Héricourt, J. (2015). Export dynamics and sales at home. *Journal of International Economics*, 96(2), 298–310.

Berman, N. & Rebeyrol, V. (2010). Exporter dynamics and productivity growth. *Working Paper, EUI MWP, 2010/26*.

Blum, B. S., Claro, S., Horstmann, I., & Rivers, D. A. (2023). The ABCs of firm heterogeneity when firms sort into markets: The case of exporters. *Journal of Political Economy*, Just Accepted.

Blundell, R. & Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric reviews*, 19(3), 321–340.

Boehm, C. E., Levchenko, A. A., & Pandalai-Nayar, N. (2023). The long and short (run) of trade elasticities. *American Economic Review*, 113(4), 861–905.

Bond, S., Hashemi, A., Kaplan, G., & Zoch, P. (2021). Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 121, 1–14.

Bonleu, A., Cette, G., & Horny, G. (2013). Capital utilization and retirement. *Applied Economics*, 45(24), 3483–3494.

Broda, C. & Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly Journal of Economics*, 121(2), 541–585.

Buus, M. T., Munch, J. R., Rodrigue, J., & Schaur, G. (2022). Do export support programs affect prices, quality, markups and marginal costs? Evidence from a natural policy experiment. *Review of Economics and Statistics*, (pp. 1–45).

Caballero, R. J. & Lyons, R. K. (1992). External effects in US procyclical productivity. *Journal of Monetary Economics*, 29(2), 209–225.

Cette, G., Dromel, N., Lecat, R., & Paret, A.-C. (2015). Production factor returns: the role of factor utilization. *Review of Economics and Statistics*, 97(1), 134–143.

De Loecker, J. (2007). Do exports generate higher productivity? evidence from slovenia. *Journal of International Economics*, 73(1), 69–98.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5), 1407–1451.

De Loecker, J. (2013). Detecting learning by exporting. *American Economic Journal: Microeconomics*, 5(3), 1–21.

De Loecker, J. & Goldberg, P. K. (2014). Firm performance in a global market. *Annu. Rev. Econ.*, 6(1), 201–227.

De Loecker, J. & Syverson, C. (2021). An Industrial Organization Perspective on Productivity. *Becker Friedman Institute Working Paper No. 2021-109*.

de Roux, N., Eslava, M., Franco, S., & Verhoogen, E. (2021). Estimating Production Functions in Differentiated-Product Industries with Quantity Information and External Instruments. *National Bureau of Economic Research Working Paper*, 28323.

Dhyne, E., Petrin, A., Smeets, V., & Warzynski, F. (2022). Theory for extending single-product production function estimation to multi-product settings. *National Bureau of Economic Research Working Paper*, 30784.

Doraszelski, U. & Jaumandreu, J. (2023). Reexamining the De Loecker & Warzynski (2012) method for estimating markups. *Working Paper*, Available at https://faculty.wharton.upenn.edu/wp-content/uploads/2016/11/Doraszelski-Reexamining-the-De-Loecker-Paper-10.30.23.pdf.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *American Economic Review*, (pp. 157–177).

Fontagné, L., Guimbard, H., & Orefice, G. (2022). Tariff-based product-level trade elasticities. *Journal of International Economics*, 137, 103593.

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1), 394–425.

Gandhi, A., Navarro, S., & Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8), 2973–3016.

Garcia-Marin, A. & Voigtländer, N. (2019). Exporting and plant-level efficiency gains: It's in the measure. *Journal of Political Economy*, 127(4), 1777–1825.

Grieco, P. L., Li, S., & Zhang, H. (2016). Production function estimation with unobserved input price dispersion. *International Economic Review*, 57(2), 665–690.

Hall, R. E. (1986). Market structure and macroeconomic fluctuations. *Brookings papers on economic activity*, 1986(2), 285–338.

Harrigan, J., Reshef, A., & Toubal, F. (2023). Techies, Trade and Skill-Biased Productivity. NBER Working Paper No. 31341.

Harrison, A. & Rodríguez-Clare, A. (2010). Trade, foreign investment, and industrial policy for developing countries. *Handbook of Development Economics*, 5, 4039–4214.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics*, 129(4), 1899–1946.

Hsieh, C.-T. & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4), 1403–1448.

Kee, H. L., Nicita, A., & Olarreaga, M. (2008). Import demand elasticities and trade distortions. *The Review of Economics and Statistics*, 90(4), 666–682.

Kellogg, R. (2011). Learning by drilling: Interfirm learning and relationship persistence in the Texas oilpatch. *The Quarterly Journal of Economics*, 126(4), 1961–2004.

Klette, T. J. & Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4), 343–361.

Knittel, C. R., Metaxoglou, K., et al. (2014). Estimation of Random-Coefficient Demand Models: Two Empiricists' Perspective. *The Review of Economics and Statistics*, 96(1), 34–59.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9(4), 469–479.

Levinsohn, J. & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2), 317–341.

Melitz, M. & Levinsohn, J. A. (2006). Productivity in a differentiated products market equilibrium. *Mimeo, Princeton University.*

Nelson, C. R. & Startz, R. (1990). Some further results on the exact small sample properties of the instrumental variable estimator. *Econometrica*, 58(4), 967.

Olley, S. & Pakes, A. (1996). Market share, market value and innovation in a panel of British manufacturing firms. *Econometrica*, 64(6), 1263–1297.

Redding, S. J. & Weinstein, D. E. (2020). Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences. *The Quarterly Journal of Economics*, 135(1), 503–560.

Reshef, A., Saint-Paul, G., & Toubal, F. (2022). Labor Market Flexibility and Firms' Resilience. *Mimeo, Paris School of Economics.*

Roberts, M. J., Yi Xu, D., Fan, X., & Zhang, S. (2018). The role of firm factors in demand, cost, and export market selection for Chinese footwear producers. *The Review of Economic Studies*, 85(4), 2429–2461.

Saint-Paul, G. (1996). *Dual labor markets*. MIT Press Cambridge.

Sato, K. (1976). The ideal log-change index number. *The Review of Economics and Statistics*, (pp. 223–228).

Shapiro, J. S. (2016). Trade costs, CO2, and the environment. *American Economic Journal: Economic Policy*, 8(4), 220–254.

Van Biesebroeck, J. (2005). Exporting raises productivity in sub-saharan african manufacturing firms. *Journal of International economics*, 67(2), 373–391.

Varian, H. R. (1992). *Microeconomic analysis*, volume 3. Norton Press New York.

Vartia, Y. O. (1976). Ideal log-change index numbers. *Scandinavian Journal of Statistics*, (pp. 121–126).

Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm-level data. *World economy*, 30(1), 60–82.

Wagner, J. (2012). International trade and firm performance: A survey of empirical studies since 2006. *Review of World Economics*, 148, 235–267.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics letters*, 104(3), 112–114.

Appendix

A Proof of proposition 1

Identification of equation (22) requires orthogonality between φ_{ft} and all variable and quasi-fixed inputs. The term φ_{ft} depends on output shares χ_{ft}, which depend on the *levels* of $\boldsymbol{D}$ and $\boldsymbol{\epsilon}$, as do all variable and quasi-fixed inputs.

We must show that $E\left[\varphi_{ft}|v^1_{ft},...,v^{\mathscr{V}}_{ft},\kappa^1_{ft},...,\kappa^{\mathscr{K}}_{ft}\right]=0$. Suffice to show that $E\left(\ln\psi_{ft}|\boldsymbol{D},\boldsymbol{\epsilon}\right)=$ constant (it is not equal to zero; see above) and does not depend on $\boldsymbol{D}$ and $\boldsymbol{\epsilon}$.

First, we develop the Taylor expansion of $\ln\sum_j\chi_j e^{u_j}$ around $\mathbf{u}=0$ (mean value for u's). The base term is

$$\sum_j \chi_j e^{u_j}\bigg|_{\boldsymbol{u}=0} = \ln\sum_j \chi_j e^0 = \ln\sum_j \chi_j = \ln 1 = 0\ . \tag{A.1}$$

The first order expansion term is:

$$\frac{1}{1!}d\ln\sum_j \chi_j e^{u_j}\bigg|_{\boldsymbol{u}=0} = \frac{1}{1!}\frac{1}{\sum_j\chi_j e^{u_j}}\sum_j \chi_j e^{u_j}du_j\bigg|_{\boldsymbol{u}=0} = \sum_j \chi_j du_j\ . \tag{A.2}$$

The second order expansion term is:

$$\frac{1}{2!}d^2\ln\sum_j \chi_j e^{u_j}\bigg|_{\boldsymbol{u}=0} = \frac{1}{2}\frac{1}{\sum_j\chi_j e^{u_j}}\sum_j \chi_j e^{u_j}d^2u_j\bigg|_{\boldsymbol{u}=0} = \frac{1}{2}\sum_j \chi_j d^2u_j \tag{A.3}$$

because $(e^u)^n=e^u$ for any n. And so on. The Taylor expansion around $\boldsymbol{u}=0$ is thus

$$\ln\sum_j \chi_j e^{u_j} = \sum_j \chi_j u_j + \frac{1}{2}\sum_j \chi_j u_j^2 + \frac{1}{3!}\sum_j \chi_j u_j^3 ... \tag{A.4}$$

The structure is linear, and thus amenable to the expectation operator. Substituting into

$E\left(\ln\psi_{ft}|\boldsymbol{D},\boldsymbol{\epsilon}\right)$ we have

$$
\begin{aligned}
& E\left(\ln\psi_{ft}|\boldsymbol{D},\boldsymbol{\epsilon}\right) \\
= \;& E\left[\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)u_{ft}^{d}+\frac{1}{2}\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)(u_{ft}^{d})^{2}+\frac{1}{3!}\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)(u_{ft}^{d})^{3}+...\right|\boldsymbol{D},\boldsymbol{\epsilon}\Bigg] \quad \text{(A.5)} \\
= \;& \sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)E\left[u_{ft}^{d}\middle|\boldsymbol{D},\boldsymbol{\epsilon}\right]+\tfrac{1}{2}\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)E\left[(u_{ft}^{d})^{2}\middle|\boldsymbol{D},\boldsymbol{\epsilon}\right] \\
& +\tfrac{1}{3!}\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)E\left[(u_{ft}^{d})^{3}\middle|(\boldsymbol{D},\boldsymbol{\epsilon})\right]+... \quad \text{(A.6)} \\
= \;& E\left[u_{ft}^{d}\right]\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)+\tfrac{1}{2}E\left[(u_{ft}^{d})^{2}\right]\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right) \\
& +\tfrac{1}{3!}E\left[(u_{ft}^{d})^{3}\right]\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\left(\boldsymbol{D},\boldsymbol{\epsilon}\right)+... \quad \text{(A.7)} \\
= \;& E\left[u_{ft}^{d}\right]+\frac{1}{2}E\left[(u_{ft}^{d})^{2}\right]+\frac{1}{3!}E\left[(u_{ft}^{d})^{3}\right]+\frac{1}{4!}E\left[(u_{ft}^{d})^{4}\right]... \quad \text{(A.8)} \\
= \;& 0+\frac{1}{2}E\left[(u_{ft}^{d})^{2}\right]+\frac{1}{3!}E\left[(u_{ft}^{d})^{3}\right]+\frac{1}{4!}E\left[(u_{ft}^{d})^{4}\right]...\;, \quad \text{(A.9)}
\end{aligned}
$$

which is a constant that does not depend on $\boldsymbol{\epsilon}$. *QED.*

As a by-product, we now know what $E\left(\ln\psi_{ft}\right)$ is equal to:

$$E\left(\ln\psi_{ft}\right)=E_{\boldsymbol{D},\boldsymbol{\epsilon}}\left[E\left(\ln\psi_{ft}|\boldsymbol{D},\boldsymbol{\epsilon}\right)\right]=E\left(\ln\psi_{ft}|\boldsymbol{D},\boldsymbol{\epsilon}\right)\;. \quad \text{(A.10)}$$

where we apply the law of iterated expectations.

B Building market quantity proxy from price indices

In a single-market estimation model, the demand-side parameter is identified from time series variation in the industry-wide CES demand index. In this section, we discuss how to construct this index.

Essentially, the quantity index can be recovered from expenditure data and industry-wide price deflators. Assuming just a single market (hence dropping the d superscript), and using (2) and (4), we can write

$$B_{t}^{\rho}\;=\;\sum_{f\in\Theta_{t}}\exp\left(\epsilon_{ft}+u_{ft}\right)X_{ft}^{\rho}=\sum_{f\in\Theta_{t}}\frac{R_{ft}}{\Upsilon_{t}}B_{t}^{\rho-1} \quad \text{(B.11)}$$

This implies

$$B_t \;=\; \sum_{f\in\Theta_t} \frac{R_{ft}}{\Upsilon_t} \tag{B.12}$$

Hence, if we observe the true CES price index in levels, we can construct the CES quantity index from aggregate deflated revenues.[53]

But of course, the true CES price index is not observed in levels. First, price indices are almost always reported relative to some base year normalization. This implies

$$B_t = \sum_{f\in\Theta_t} \frac{R_{ft}}{\Upsilon_t} = \Upsilon_0 \underbrace{\sum_{f\in\Theta_t} \frac{R_{ft}}{\Lambda_t}}_{\equiv B_t^{proxy}} \tag{B.13}$$

where Λ_t is the empirical analogue to the true CES price index normalized to base-year $t = 0$, and Υ_0 is the unobserved base-year normalization.

Second, the CES price index is a theoretical construct that depends on structural parameters. How does this theoretical object correspond to Λ_t? Sato (1976) and Vartia (1976) prove for a symmetric CES with no entry and exit, there exists a set of weights wt_{ft} such that

$$\ln \frac{\Upsilon_t}{\Upsilon_0} \;=\; \sum_{f\in\Theta_t} wt_{ft} \ln\left(\frac{p_{ft}}{p_{fi0}}\right) \tag{B.14}$$

I.e., the log change in the true CES price index is a weighted average of the log change in the prices of individual firms. Sato (1976) and Vartia (1976) give the analytical expression for these weights, which ends up being very close to a simple chain weight. Feenstra (1994) extends to the case of entry and exit. Redding & Weinstein (2020) extends to the asymmetric CES (which corresponds to our demand system (2)). If we assume that Λ_t is computed using Weinstein-Redding weights, then (B.13) holds.

[53] In the only work we are aware of that explains how to construct the CES quantity index, De Loecker (2011) computes the weighted average of deflated revenues (see De Loecker (2011) equation B.1.9 in the appendix), though – as we show in (B.12) – theory indicates the gross sum is called for.

C Control function method

In this section, we describe the control function approach to estimating our multi-destination model. The procedure is based on the control function estimation of the gross output production function described by Gandhi et al. (2020).[54]

The control function approach proceeds in two steps. In the first step, *ex post* shocks (possibly inclusive of measurement error) are computed as the residual of a non-parametric regression of revenues on all input levels. Identification relies on substituting for the endogenous unobservable with the material demand function. In the second step, the *ex post* shock is subtracted off from revenues and all structural parameters are identified via GMM.

In the first step, Gandhi et al. (2020) invert the material demand function to substitute for the unobserved shock ($\nu_{ft} = \rho\omega_{ft} + \epsilon^1_{ft}$, in our case). Since we assume cost minimizing behavior, we can use the first order conditions instead to accomplish this substitution. Labeling material demands v^1_{ft}, we substitute (12) into (20) and write

$$r_{ft} = f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) - \ln\left[\frac{\partial F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}})}{\partial e^{v^1_{ft}}}\right] + \ln W^{v^1}_t - \ln \rho E\left[\exp(u)\right] + \ln\psi_{ft} \quad \text{(C.15)}$$

Collecting the first two terms into an unknown function, we get

$$r_{ft} = \Phi(v^1_{ft}, ..., v^{\mathscr{V}}_{ft}, \kappa^1_{ft}, ..., \kappa^{\mathscr{K}}_{ft}) + \ln W^{v^1}_t - \ln\rho E\left[\exp(u)\right] + \ln\psi_{ft} \quad \text{(C.16)}$$

We estimate this model approximating $\Phi(\cdot)$ with polynomials, including time fixed effects to control for input prices, and label the residual $\widehat{\varphi_{ft}}$.[55]

[54] The procedure from Gandhi et al. (2020) is virtually the same as the procedure proposed by Ackerberg et al. (2015), except that Ackerberg et al. (2015) consider the value-added production function. Hence, Ackerberg et al. (2015) do not identify the material input elasticity.

[55] We could alternatively substitute for ν_{ft} using the inverse material demand. In this case, the demand shifter $\ln\left[\frac{R_{ft}}{R^1_{ft}}\right]$ does not cancel. With this method, we could write

$$r_{ft} = \Phi(v^1_{ft}, ..., v^{\mathscr{V}}_{ft}, \kappa^1_{ft}, ..., \kappa^{\mathscr{K}}_{ft}, \ln\left[\frac{R_{ft}}{R^1_{ft}}\right]) + \delta_t + \rho\ln\psi_{ft} \quad \text{(C.17)}$$

In this formulation, we identify $\rho\ln\psi_{ft}$ in the first step, not $\ln\psi_{ft}$. If we identify $\rho\ln\psi_{ft}$, then we can subtract it off from both sides of (20) to write

$$\widetilde{r}_{ft} = \alpha_t + \rho F(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) + (1-\rho)\ln\left[\frac{R_{ft}}{R^1_{ft}}\right] + \nu_{ft} \quad \text{(C.18)}$$

The difference between this approach to the control function first step and the approach using the first

In the second step, we subtract off $\widehat{\varphi_{ft}}$ from the both sides of (20) to write

$$\widetilde{r}_{ft} = \alpha_t + \rho f(\boldsymbol{v_{ft}}, \boldsymbol{\kappa_{ft}}) + (1-\rho)\ln \widehat{D}_{ft} + \nu_{ft} \tag{C.19}$$

with $\widetilde{r}_{ft} \equiv r_{ft} - \widehat{\varphi_{ft}}$, $\alpha_t = \ln D_t^1 + E\left[\ln \psi\right]$ and $\widehat{D}_{ft} = \left[\frac{R_{ft}\exp\left(\widehat{-\varphi_{ft}}\right)}{R_{ft}^1}\right]$. By assumptions (28) - (29), we have that $\nu_{ft} = h\nu_{f,t-1} + \mu e_{f,t-1} + \xi_{ft}$ and $\xi_{ft} \equiv \widetilde{\epsilon}_{ft}^1 + (1-\rho)u_{ft}^1 + h(1-\rho)u_{f,t-1}^1 + \rho\widetilde{\omega}_{ft}$.

We adopt a complete polynomial of degree 2 and write

$$\begin{aligned}\widetilde{r}_{ft} &= \alpha_t + \beta^D \ln \widehat{D}_{ft} + \sum_{j\in\{1,\dots,\mathscr{V}\}} g_{v^j} v_{ft}^j + \sum_{j\in\{1,\dots,\mathscr{K}\}} b_{\kappa^j}\kappa_{ft}^j + \sum_{j\in\{1,\dots,\mathscr{V}\}}\ \sum_{z\in\{v^j,\dots,v^{\mathscr{V}},\kappa^1,\dots,\kappa^{\mathscr{K}}\}} g_{v^j z} v_{ft}^j z_{ft} \\ &+ \sum_{j\in\{1,\dots,\mathscr{K}\}}\ \sum_{z\in\{\kappa^j,\dots,\kappa^{\mathscr{K}},v^1,\dots,v^{\mathscr{V}}\}} b_{\kappa^j z}\kappa_{ft}^j z_{ft} + \nu_{ft}\end{aligned} \tag{C.20}$$

For any candidate vector, we can compute $\widehat{\nu_{ft} + \alpha}_t$, regress $\widehat{\nu_{ft} + \alpha}_t$ on $\widehat{\nu_{f,t-1} + \alpha}_{t-1}$, $e_{f,t-1}$, and time fixed effects, and compute the residual $\widehat{\xi}_{ft}(\cdot)$. We then build moment conditions by multiplying $\widehat{\xi}_{ft}(\cdot)$ with the levels of all quasi-fixed inputs and $\ln \widehat{D}_{f,t-2}$ and all flexible inputs, along with the appropriate interaction and square terms. At the true parameter values, $\widehat{\xi}_{ft}$ correlates with $\ln \widehat{D}_{ft}$ and all flexible inputs in period t. But given the timing assumptions, $\widehat{\xi}_{ft}$ is orthogonal to the lags of all flexible inputs and $\ln \widehat{D}_{f,t-2}$.

Finally, we compute $\widehat{\rho} = 1 - \widehat{\beta^D}$, deflate all the $\widehat{g}$, $\widehat{b}$ coefficients, and compute factor output elasticities.

---

order condition is that here, we condition on $\ln\left[\frac{R_{ft}}{R_{ft}^1}\right]$ in the first step and then leave φ_{ft} out of the construction of the firm-specific demand shifter in the second step. Since we already assume monopolistic competition and cost minimizing behavior to solve the model, there is no reason not to use the first order condition in the control function first step. But we certainly could adopt this alternative method.

D Sensitivity to initial guesses in the control function approach

In order to investigate the sensitivity to initial values discussed by Ackerberg et al. (2020), we return to the single-market data generating process from section 4.1. In the main text we present Monte Carlo results for the control function approach when the second-step GMM procedure starts from the true parameter vector and when the GMM starts from the OLS estimates. We document in Figure 1 that the solution to the GMM varies with the starting value.

Here, we confirm the analysis from Ackerberg et al. (2020) by systematically varying the starting conditions. We run the same Monte Carlo simulations as in the main text. We vary initial conditions for $\beta_0^D \in \{0, 0.05, 0.1, 0.15, 0.2, 0.25\}$, $\beta_0^M \in \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$ and $\beta_0^K \in \{0, 0.1, 0.2, 0.3, 0.4\}$. We compute the GMM solution starting from every combination defined by these three sets. Recall that the true parameter values are $\beta^D = 1 - \rho = 0.2$, $\beta^M = \rho\sigma^M = 0.64$ and $\beta^K = \rho\sigma^K = 0.24$.

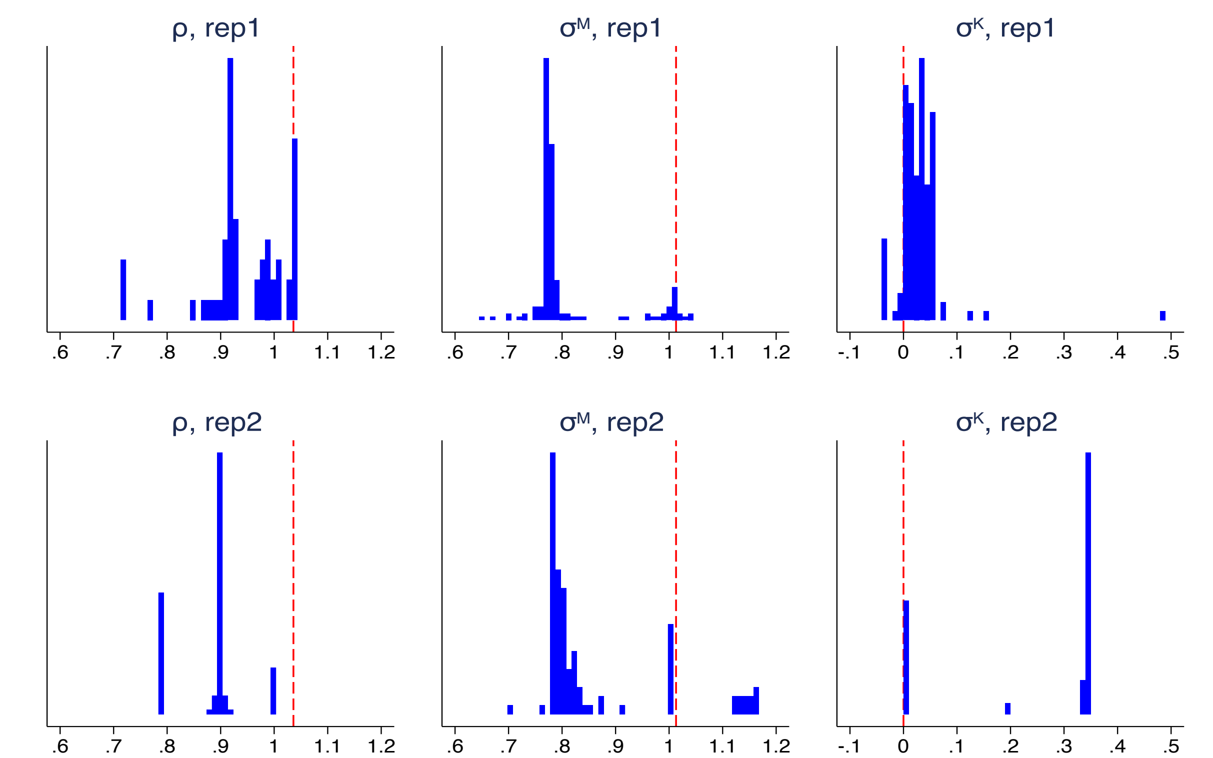
In Figure D.1 (D.2), we present results when the data is simulated with a high (low) degree of time series variation in material input prices. We display the distribution of GMM solutions for each parameter for two different Monte Carlo samples, one in each row. Results starting from the OLS estimates are depicted with a vertical red dashed line. Results from all other starting values are depicted in solid bars in blue.

In Figure D.1, we see that the estimates based on the OLS initial values coincide with the results in Figure 1: starting from the OLS values, the GMM solution tends towards $\rho = 1$, $\sigma^M = 1$, and $\sigma^K = 0$ (red dashed line). In Figure D.1, we also see that when the GMM starts from other initial conditions there is a mass point of convergence around the same values, although we also see other mass points. This bunching pattern is consistent with Ackerberg et al. (2020).

In Figure D.2, we see that the estimates converges more often towards the OLS results. That is, the distribution is bunched more tightly around the OLS estimates than in Figure D.1. This is what we would expect, as with low input price variation, the GMM suffers from weak instruments, which tends to bias the esitmates towards the OLS result.

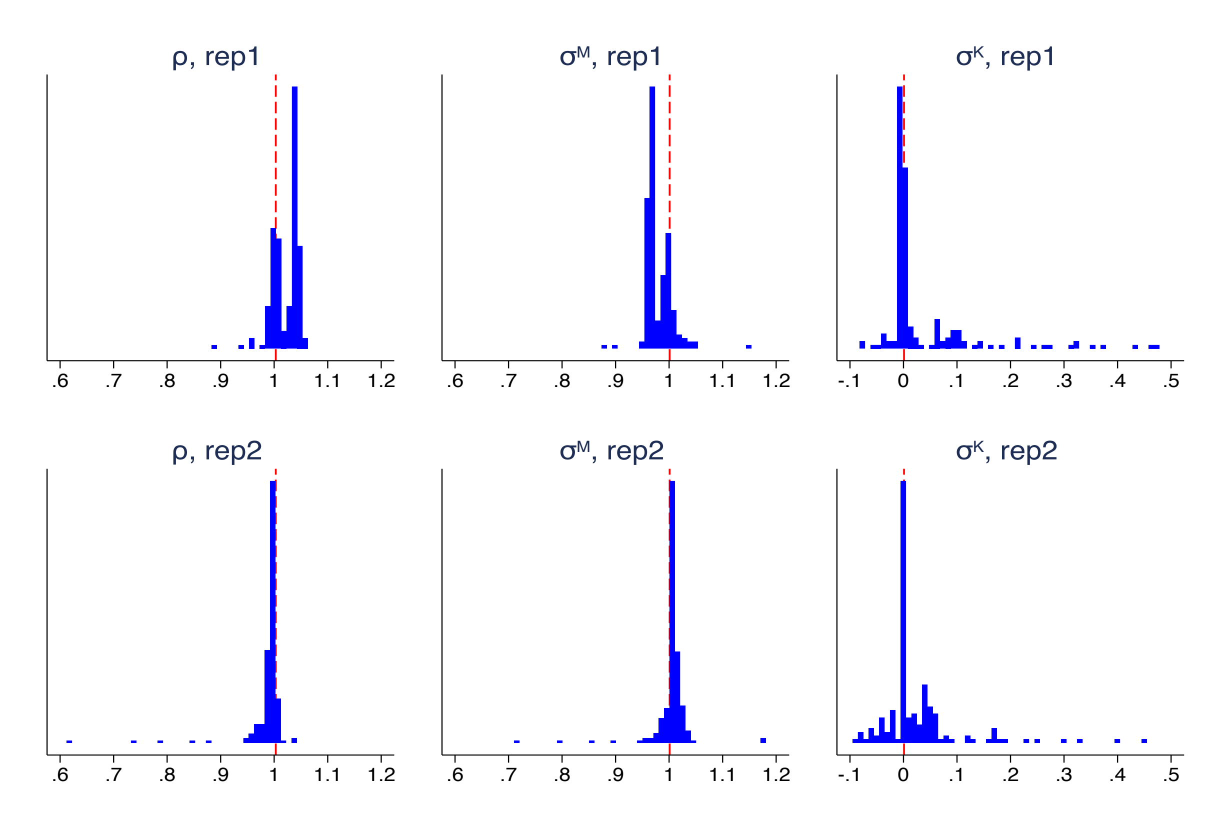
Overall, we see that given the exact same data, the GMM solution can vary wildly depending on the initial guess for the parameters.

Figure D.1: Different Starting Values for Control Function Estimation, High Material Price Variation



Notes. The figure reports the distribution of estimates of ρ and averages of estimates of factor output elasticities resulting from the control function method taking different parameter vectors as starting values for the second-step GMM procedure. Each row reports results for a single Monte Carlo sample. Data is generated based on the single-market scenario described in Section 4.1, with high input price variation. The true parameter values are $\rho = 0.8$, $\sigma^M = 0.8$ and $\sigma^K = 0.3$.

Figure D.2: Different Starting Values for Control Function Estimation, Low Material Price Variation



Notes. The figure reports the distribution of estimates of ρ and averages of estimates of factor output elasticities resulting from the control function method taking different parameter vectors as starting values for the second-step GMM procedure. Each row reports results for a single Monte Carlo sample. Data is generated based on the single-market scenario described in Section 4.1, with low input price variation. The true parameter values are $\rho = 0.8$, $\sigma^M = 0.8$ and $\sigma^K = 0.3$.

E Data

Firm-level balance sheet information is reported in the FICUS (*Fichier complet unifié de SUSE*) and FARE (*Fichier Approché des Résultats ESANE*) datasets, which cover the periods 1994–2007 and 2008–2016, respectively. These data originate in tax declarations of all firms in France, and are collected by the French National Institute of Statistics and Economic Studies, INSEE. We use total revenue, expenditure on materials, employment and the book value of capital.

We construct capital stocks following the methodology proposed by Bonleu et al. (2013) and Cette et al. (2015). We start with the book value of capital. Since the stocks are recorded at historical cost, i.e., the value at the time of entry into the firm i's balance sheet, an adjustment has to be made to move from stocks valued at historic cost ($K^{BV}_{i,s,t}$) to stocks valued at current prices ($K_{i,s,t}$). We deflate K^{BV} by an industry-specific price index (sourced from INSEE) that assumes that the price of capital is equal to the sectoral price of investment T years before the date when the first book value was available, where T is the corrected average age of capital, hence $p^K_{s,t+1} = p^I_{s,t-T}$. The average age of capital is computed using the share of depreciated capital, $DK^{BV}_{i,s,t}$ in the capital stock at historical cost:

$$T = \frac{DK^{BV}_{i,s,t}}{K^{BV}_{i,s,t}} \times \widetilde{A}$$

where

$$\widetilde{A} = median_{i \in S} \left(\frac{K^{BV}_{i,s,t}}{\Delta DK^{BV}_{i,s,t}} \right)$$

where S the set of firms in a sector. We use the median value $\widetilde{A}$ to reduce the volatility in the data, as investments within firms are discrete events.

Data on firms' exports are from the French Customs. For each observation, we know the value of exports of the firm. We use the firm-level SIREN identifier to match the trade data to FICUS/FARE. This match is not perfect. The imperfect match is because there are SIRENs in the trade data for which there is no corresponding SIREN in FICUS/FARE. This may lead to measurement error: for some firms, we will observe zero exports even when true exports are positive. This is not a big concern because most of the missing values are in the oil refining industry, which we drop from our sample.

F Additional Results from French Manufacturing

F.1 Results with pre-determined labor

Table F.1: Estimates using Multi-Market Estimator, Predetermined Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.488	0.546	0.346	1.380	0.710	-3.453	-0.004	0.792	-0.019
	(0.009)	(0.010)	(0.007)	(0.025)	(0.013)	(0.163)	(0.003)	(0.005)	(0.013)
Textiles, wearing apparel	0.339	0.560	0.237	1.136	0.798	-4.956	0.040	0.898	0.396
	(0.024)	(0.037)	(0.017)	(0.076)	(0.051)	(1.486)	(0.003)	(0.004)	(0.023)
Wood, paper products	0.312	0.719	0.187	1.218	0.807	-5.178	0.019	0.830	0.111
	(0.006)	(0.013)	(0.005)	(0.022)	(0.014)	(0.357)	(0.001)	(0.006)	(0.006)
Chemical products	0.384	0.535	0.177	1.096	0.954	-21.523	0.013	0.870	0.102
	(0.007)	(0.015)	(0.009)	(0.015)	(0.012)	(6.614)	(0.002)	(0.015)	(0.016)
Rubber and plastic	0.360	0.536	0.203	1.099	0.923	-13.015	0.000	0.870	0.002
	(0.010)	(0.016)	(0.012)	(0.036)	(0.027)	(5.364)	(0.001)	(0.007)	(0.010)
Basic metal and fabricated metal	0.269	0.734	0.209	1.212	0.808	-5.195	0.010	0.836	0.060
	(0.006)	(0.017)	(0.006)	(0.028)	(0.018)	(0.500)	(0.001)	(0.005)	(0.007)
Computer, electronics	0.325	0.595	0.150	1.070	0.918	-12.266	0.019	0.838	0.117
	(0.007)	(0.015)	(0.007)	(0.021)	(0.017)	(2.952)	(0.002)	(0.008)	(0.014)
Electrical equipment	0.366	0.533	0.156	1.055	0.931	-14.575	0.017	0.833	0.105
	(0.011)	(0.019)	(0.009)	(0.029)	(0.025)	(27.360)	(0.003)	(0.010)	(0.014)
Machinery and equipment	0.341	0.676	0.141	1.158	0.829	-5.838	0.029	0.786	0.137
	(0.014)	(0.028)	(0.012)	(0.052)	(0.033)	(0.653)	(0.002)	(0.008)	(0.007)
Autos and transport equipment	0.386	0.562	0.172	1.120	0.942	-17.316	0.018	0.792	0.086
	(0.006)	(0.014)	(0.010)	(0.018)	(0.014)	(3.865)	(0.003)	(0.018)	(0.012)
Other manufacturing	0.273	0.631	0.227	1.131	0.827	-5.767	0.024	0.856	0.167
	(0.006)	(0.012)	(0.007)	(0.021)	(0.017)	(0.631)	(0.001)	(0.006)	(0.008)

Notes. The table reports estimates based on the multi-market estimator, treating labor as predetermined (like capital): average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.2: Estimates using no Demand Correction, Predetermined Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.346	0.401	0.160	0.907	-	-	0.092	0.923	1.185
	(0.001)	(0.003)	(0.002)	(0.003)			(0.002)	(0.002)	(0.027)
Textiles, wearing apparel	0.271	0.455	0.169	0.895	-	-	0.041	0.951	0.822
	(0.002)	(0.006)	(0.004)	(0.005)			(0.002)	(0.002)	(0.029)
Wood, paper products	0.252	0.581	0.145	0.977	-	-	0.035	0.949	0.695
	(0.001)	(0.004)	(0.003)	(0.003)			(0.002)	(0.002)	(0.019)
Chemical products	0.366	0.511	0.167	1.044	-	-	0.023	0.969	0.751
	(0.004)	(0.013)	(0.009)	(0.006)			(0.003)	(0.005)	(0.084)
Rubber and plastic	0.333	0.493	0.179	1.006	-	-	0.016	0.971	0.546
	(0.002)	(0.005)	(0.005)	(0.003)			(0.001)	(0.002)	(0.039)
Basic metal and fabricated metal	0.218	0.596	0.156	0.969	-	-	0.035	0.931	0.505
	(0.001)	(0.004)	(0.003)	(0.003)			(0.001)	(0.002)	(0.014)
Computer, electronics	0.299	0.536	0.144	0.979	-	-	0.038	0.936	0.597
	(0.004)	(0.010)	(0.006)	(0.009)			(0.003)	(0.007)	(0.044)
Electrical equipment	0.341	0.484	0.144	0.969	-	-	0.034	0.953	0.710
	(0.004)	(0.011)	(0.008)	(0.008)			(0.003)	(0.004)	(0.051)
Machinery and equipment	0.283	0.556	0.109	0.948	-	-	0.053	0.880	0.446
	(0.002)	(0.007)	(0.003)	(0.007)			(0.002)	(0.008)	(0.020)
Autos and transport equipment	0.363	0.541	0.147	1.052	-	-	0.045	0.924	0.586
	(0.004)	(0.010)	(0.008)	(0.008)			(0.005)	(0.014)	(0.069)
Other manufacturing	0.226	0.531	0.166	0.923	-	-	0.039	0.943	0.681
	(0.001)	(0.005)	(0.003)	(0.004)			(0.002)	(0.002)	(0.021)

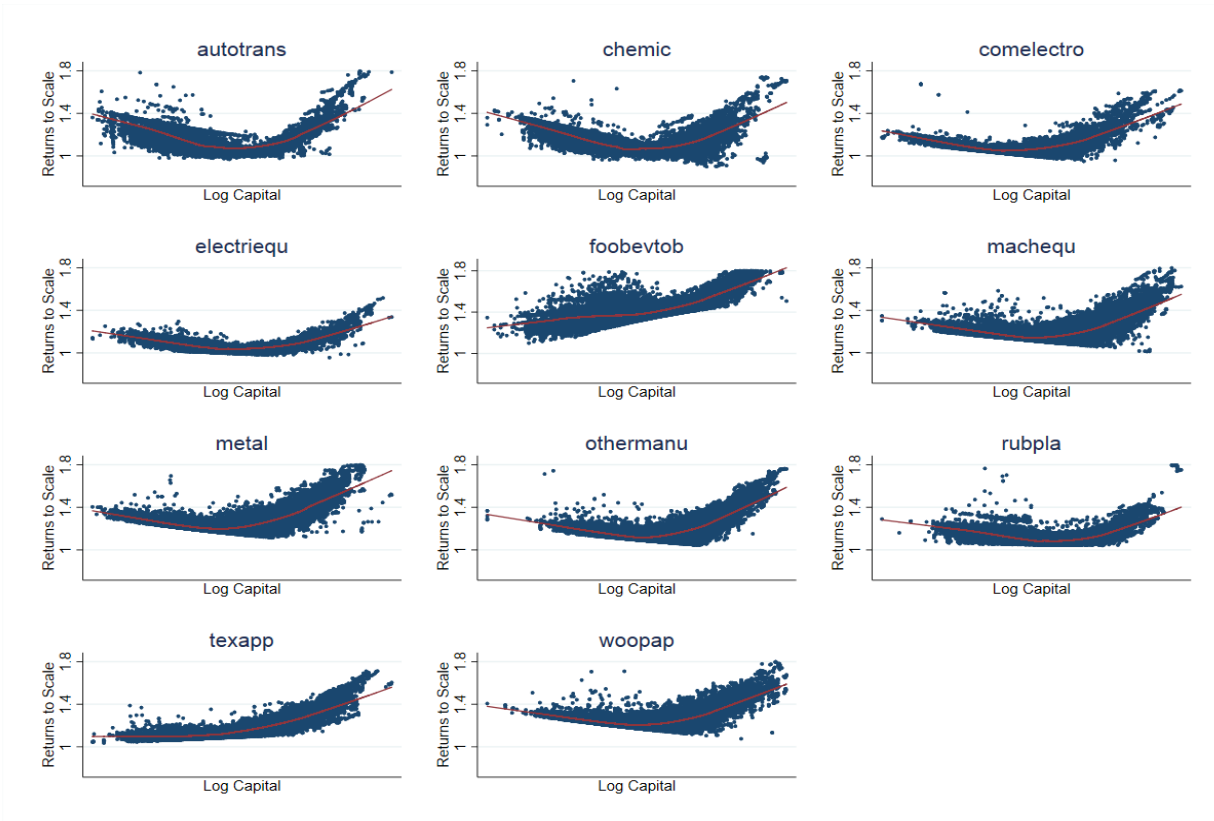
Notes. The table reports estimates without correcting for demand at all, using the factor share approach, treating labor as predetermined (like capital): average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.3: Estimates using Single-Market Estimator (KG), Predetermined Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.843	0.964	0.411	2.218	0.411	-1.698	0.085	0.849	0.563
	(0.018)	(0.024)	(0.010)	(0.050)	(0.009)	(0.025)	(0.004)	(0.005)	(0.019)
Textiles, wearing apparel	0.182	0.310	0.113	0.605	1.487	2.054	0.027	0.894	0.257
	(0.003)	(0.006)	(0.003)	(0.010)	(0.022)	(0.094)	(0.001)	(0.004)	(0.008)
Wood, paper products	0.497	1.141	0.296	1.935	0.506	-2.025	0.036	0.848	0.236
	(0.029)	(0.064)	(0.020)	(0.112)	(0.030)	(0.126)	(0.002)	(0.007)	(0.016)
Chemical products	2.534	3.532	1.183	7.249	0.144	-1.169	0.101	0.873	0.797
	(0.537)	(0.786)	(0.243)	(1.553)	(0.025)	(0.034)	(0.026)	(0.015)	(0.220)
Rubber and plastic	-0.790	-1.158	-0.452	-2.399	-0.421	-0.704	-0.007	0.875	-0.059
	(0.083)	(0.125)	(0.041)	(0.247)	(0.051)	(0.024)	(0.002)	(0.008)	(0.015)
Basic metal and fabricated metal	0.448	1.161	0.373	1.982	0.486	-1.944	0.029	0.851	0.192
	(0.014)	(0.033)	(0.014)	(0.058)	(0.014)	(0.054)	(0.002)	(0.006)	(0.010)
Computer, electronics	0.393	0.719	0.181	1.293	0.759	-4.150	0.031	0.837	0.189
	(0.028)	(0.052)	(0.016)	(0.093)	(0.055)	(1.559)	(0.003)	(0.009)	(0.019)
Electrical equipment	0.477	0.698	0.188	1.362	0.715	-3.511	0.031	0.834	0.188
	(0.018)	(0.030)	(0.011)	(0.049)	(0.025)	(0.338)	(0.003)	(0.011)	(0.017)
Machinery and equipment	0.498	0.982	0.194	1.674	0.568	-2.316	0.061	0.794	0.296
	(0.057)	(0.106)	(0.031)	(0.193)	(0.042)	(0.153)	(0.005)	(0.014)	(0.065)
Autos and transport equipment	0.370	0.545	0.159	1.073	0.983	-58.695	0.023	0.795	0.111
	(0.005)	(0.010)	(0.009)	(0.009)	(0.004)	(14.148)	(0.002)	(0.018)	(0.010)
Other manufacturing	-0.352	-0.818	-0.304	-1.475	-0.640	-0.610	-0.034	0.876	-0.271
	(0.074)	(0.175)	(0.060)	(0.308)	(0.108)	(0.042)	(0.009)	(0.007)	(0.054)

Notes. The table reports estimates based on the single-market estimator *à la* Klette & Griliches (1996), using the factor share approach, treating labor as predetermined (like capital): average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

Figure F.3: Returns to Scale by Industry



Notes. This figure presents estimated returns to scale via the multi-market factor shares estimator by firm-year against log of capital. Labor is assumed to be pre-determined.

Table F.4: Estimates using Single-Market Estimator (KG) on Sample of Non-Exporters, Predetermined Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.824	0.950	0.421	2.196	0.420	-1.725	-	0.857	-
	(0.018)	(0.024)	(0.010)	(0.050)	(0.009)	(0.026)		(0.005)	
Textiles, wearing apparel	0.179	0.318	0.127	0.624	1.515	1.942	-	0.913	-
	(0.003)	(0.007)	(0.003)	(0.011)	(0.024)	(0.092)		(0.003)	
Wood, paper products	0.521	1.210	0.325	2.056	0.483	-1.935	-	0.851	-
	(0.032)	(0.071)	(0.023)	(0.124)	(0.030)	(0.115)		(0.006)	
Chemical products	2.838	4.035	1.361	8.235	0.129	-1.148	-	0.874	-
	(0.729)	(1.078)	(0.343)	(2.136)	(0.026)	(0.034)		(0.014)	
Rubber and plastic	-0.789	-1.162	-0.454	-2.405	-0.422	-0.703	-	0.874	-
	(0.082)	(0.125)	(0.041)	(0.247)	(0.051)	(0.024)		(0.008)	
Basic metal and fabricated metal	0.456	1.193	0.395	2.044	0.477	-1.913	-	0.854	-
	(0.014)	(0.034)	(0.014)	(0.059)	(0.014)	(0.051)		(0.005)	
Computer, electronics	0.458	0.852	0.229	1.539	0.652	-2.874	-	0.843	-
	(0.039)	(0.073)	(0.024)	(0.133)	(0.058)	(0.618)		(0.009)	
Electrical equipment	0.489	0.727	0.212	1.428	0.697	-3.304	-	0.838	-
	(0.019)	(0.031)	(0.012)	(0.053)	(0.026)	(0.299)		(0.011)	
Machinery and equipment	0.510	1.025	0.219	1.753	0.555	-2.245	-	0.804	-
	(0.073)	(0.140)	(0.043)	(0.255)	(0.046)	(0.155)		(0.013)	
Autos and transport equipment	0.370	0.555	0.168	1.093	0.982	-55.556	-	0.801	-
	(0.005)	(0.011)	(0.009)	(0.010)	(0.004)	(12.718)		(0.018)	
Other manufacturing	-0.309	-0.738	-0.276	-1.323	-0.729	-0.578	-	0.881	-
	(0.055)	(0.132)	(0.045)	(0.231)	(0.108)	(0.037)		(0.007)	

Notes. The table reports estimates based on the single-market estimator *à la* Klette & Griliches (1996), using the factor share approach, treating labor as predetermined (like capital), where we restrict the sample to non-exporting firms: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

F.2 Results allowing partial adjustment for labor

Table F.5: Estimates using Multi-Market Estimator, Partially Adjusting Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.486	0.846	0.199	1.532	0.712	-3.475	0.006	0.723	0.021
	(0.008)	(0.013)	(0.004)	(0.024)	(0.012)	(0.142)	(0.003)	(0.004)	(0.010)
Textiles, wearing apparel	0.373	0.779	0.185	1.337	0.725	-3.641	0.045	0.849	0.296
	(0.021)	(0.072)	(0.013)	(0.089)	(0.039)	(0.541)	(0.003)	(0.005)	(0.020)
Wood, paper products	0.323	0.941	0.102	1.365	0.780	-4.545	0.011	0.776	0.051
	(0.005)	(0.019)	(0.004)	(0.024)	(0.012)	(0.246)	(0.002)	(0.005)	(0.007)
Chemical products	0.390	0.703	0.105	1.198	0.937	-15.926	-0.000	0.821	-0.002
	(0.006)	(0.023)	(0.012)	(0.017)	(0.010)	(2.817)	(0.003)	(0.012)	(0.016)
Rubber and plastic	0.374	0.736	0.124	1.233	0.890	-9.120	-0.014	0.818	-0.079
	(0.009)	(0.032)	(0.005)	(0.041)	(0.022)	(1.453)	(0.003)	(0.006)	(0.015)
Basic metal and fabricated metal	0.277	0.898	0.126	1.301	0.786	-4.683	0.005	0.802	0.024
	(0.006)	(0.024)	(0.004)	(0.030)	(0.017)	(0.364)	(0.001)	(0.004)	(0.007)
Computer, electronics	0.328	0.711	0.095	1.133	0.911	-11.225	0.009	0.810	0.046
	(0.007)	(0.024)	(0.008)	(0.027)	(0.018)	(2.619)	(0.003)	(0.007)	(0.016)
Electrical equipment	0.368	0.644	0.099	1.112	0.926	-13.430	0.012	0.799	0.062
	(0.010)	(0.031)	(0.012)	(0.032)	(0.024)	(5.545)	(0.003)	(0.008)	(0.016)
Machinery and equipment	0.348	0.811	0.092	1.251	0.813	-5.360	0.019	0.758	0.080
	(0.015)	(0.055)	(0.004)	(0.071)	(0.034)	(0.583)	(0.003)	(0.005)	(0.010)
Autos and transport equipment	0.382	0.637	0.129	1.148	0.951	-20.317	0.009	0.767	0.037
	(0.006)	(0.016)	(0.010)	(0.018)	(0.013)	(4.929)	(0.003)	(0.015)	(0.013)
Other manufacturing	0.292	0.962	0.138	1.392	0.771	-4.374	0.008	0.775	0.035
	(0.005)	(0.022)	(0.005)	(0.027)	(0.014)	(0.277)	(0.002)	(0.004)	(0.010)

Notes. The table reports estimates based on the multi-market estimator, allowing labor to partially adjust to current period shocks: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1 - h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.6: Estimates using Single-Market Estimator (KG), Partially Adjusting Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.469	0.914	0.084	1.466	0.739	-3.826	0.054	0.775	0.238
	(0.012)	(0.018)	(0.007)	(0.033)	(0.018)	(0.262)	(0.003)	(0.003)	(0.013)
Textiles, wearing apparel	0.171	0.287	0.115	0.573	1.586	1.706	0.029	0.852	0.195
	(0.003)	(0.012)	(0.007)	(0.010)	(0.029)	(0.085)	(0.002)	(0.005)	(0.012)
Wood, paper products	0.508	1.429	0.193	2.130	0.496	-1.982	0.026	0.810	0.137
	(0.032)	(0.095)	(0.016)	(0.136)	(0.030)	(0.120)	(0.003)	(0.004)	(0.013)
Chemical products	23.606	44.825	5.790	74.221	0.015	-1.016	-0.310	0.836	-1.890
	(85.397)	(163.150)	(19.452)	(267.802)	(0.151)	(0.471)	(0.804)	(0.015)	(5.118)
Rubber and plastic	-0.803	-1.554	-0.293	-2.650	-0.415	-0.707	0.020	0.830	0.116
	(0.087)	(0.154)	(0.034)	(0.272)	(0.054)	(0.025)	(0.003)	(0.007)	(0.016)
Basic metal and fabricated metal	0.352	0.900	0.274	1.526	0.618	-2.615	0.024	0.810	0.127
	(0.008)	(0.034)	(0.012)	(0.041)	(0.014)	(0.096)	(0.002)	(0.005)	(0.010)
Computer, electronics	0.428	0.893	0.150	1.471	0.697	-3.300	0.020	0.812	0.107
	(0.030)	(0.069)	(0.016)	(0.107)	(0.049)	(0.649)	(0.003)	(0.008)	(0.017)
Electrical equipment	0.485	0.828	0.134	1.448	0.702	-3.358	0.023	0.799	0.115
	(0.021)	(0.062)	(0.018)	(0.071)	(0.030)	(0.372)	(0.004)	(0.012)	(0.020)
Machinery and equipment	0.554	1.218	0.172	1.944	0.510	-2.042	0.056	0.776	0.251
	(0.078)	(0.176)	(0.033)	(0.286)	(0.044)	(0.131)	(0.007)	(0.010)	(0.053)
Autos and transport equipment	0.370	0.615	0.127	1.112	0.982	-54.193	0.011	0.770	0.050
	(0.004)	(0.016)	(0.011)	(0.011)	(0.004)	(14.038)	(0.003)	(0.015)	(0.010)
Other manufacturing	2.353	7.098	1.496	10.948	0.096	-1.106	0.252	0.803	1.274
	(9.305)	(28.465)	(5.755)	(43.521)	(0.060)	(0.075)	(0.928)	(0.005)	(4.851)

Notes. The table reports estimates based on the single-market estimator *à la* Klette & Griliches (1996), using the factor share approach, allowing labor to partially adjust to current period shocks: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1 - h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.7: Estimates using Single-Market Estimator (KG) on Sample of Non-Exporters, Partially Adjusting Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.455	0.905	0.084	1.444	0.761	-4.179	-	0.777	-
	(0.012)	(0.018)	(0.008)	(0.033)	(0.019)	(0.331)		(0.003)	
Textiles, wearing apparel	0.162	0.305	0.114	0.581	1.673	1.487	-	0.850	-
	(1.479)	(1.440)	(1.419)	(4.338)	(0.248)	(0.361)		(0.012)	
Wood, paper products	0.522	1.491	0.200	2.214	0.482	-1.930	-	0.811	-
	(0.034)	(0.101)	(0.017)	(0.145)	(0.030)	(0.114)		(0.004)	
Chemical products	18.729	35.354	4.628	58.711	0.020	-1.020	-	0.836	-
	(36.482)	(69.525)	(8.114)	(114.013)	(0.098)	(0.079)		(0.015)	
Rubber and plastic	-0.812	-1.548	-0.299	-2.659	-0.410	-0.709	-	0.832	-
	(0.088)	(0.157)	(0.035)	(0.277)	(0.054)	(0.025)		(0.007)	
Basic metal and fabricated metal	0.357	0.933	0.280	1.569	0.610	-2.563	-	0.811	-
	(0.008)	(0.034)	(0.012)	(0.041)	(0.013)	(0.089)		(0.005)	
Computer, electronics	0.465	0.997	0.163	1.625	0.642	-2.792	-	0.815	-
	(0.035)	(0.083)	(0.019)	(0.129)	(0.050)	(0.456)		(0.008)	
Electrical equipment	0.494	0.870	0.140	1.504	0.690	-3.222	-	0.801	-
	(0.022)	(0.063)	(0.019)	(0.073)	(0.030)	(0.338)		(0.011)	
Machinery and equipment	0.578	1.312	0.186	2.076	0.489	-1.958	-	0.783	-
	(0.094)	(0.221)	(0.041)	(0.355)	(0.045)	(0.125)		(0.009)	
Autos and transport equipment	0.370	0.627	0.128	1.124	0.981	-53.237	-	0.773	-
	(0.004)	(0.016)	(0.011)	(0.011)	(0.004)	(13.594)		(0.015)	
Other manufacturing	3.197	10.031	1.991	15.219	0.071	-1.076	-	0.804	-
	(406.503)	(1280.868)	(252.535)	(1939.906)	(0.059)	(0.069)		(0.005)	

Notes. The table reports estimates based on the single-market estimator *à la* Klette & Griliches (1996), using the factor share approach, allowing labor to partially adjust to current period shocks, where we restrict the sample to non-exporting firms: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1 - h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.8: Estimates using no Demand Correction, Partially Adjusting Labor Input

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.346	0.635	0.070	1.052	-	-	0.142	0.861	1.024
	(0.001)	(0.008)	(0.004)	(0.006)			(0.003)	(0.004)	(0.041)
Textiles, wearing apparel	0.271	0.433	0.192	0.896	-	-	0.046	0.948	0.869
	(0.002)	(0.021)	(0.009)	(0.013)			(0.003)	(0.003)	(0.038)
Wood, paper products	0.252	0.708	0.092	1.052	-	-	0.046	0.908	0.507
	(0.001)	(0.010)	(0.005)	(0.006)			(0.001)	(0.005)	(0.025)
Chemical products	0.366	0.830	0.025	1.221	-	-	0.011	0.869	0.088
	(0.004)	(0.339)	(0.153)	(0.186)			(0.091)	(0.027)	(0.902)
Rubber and plastic	0.333	0.689	0.097	1.118	-	-	0.026	0.944	0.462
	(0.002)	(0.022)	(0.010)	(0.013)			(0.002)	(0.005)	(0.041)
Basic metal and fabricated metal	0.218	0.629	0.129	0.976	-	-	0.041	0.898	0.400
	(0.001)	(0.010)	(0.005)	(0.006)			(0.002)	(0.007)	(0.021)
Computer, electronics	0.299	0.611	0.115	1.025	-	-	0.043	0.903	0.443
	(0.004)	(0.023)	(0.011)	(0.014)			(0.003)	(0.016)	(0.062)
Electrical equipment	0.341	0.569	0.105	1.015	-	-	0.054	0.920	0.675
	(0.004)	(0.040)	(0.022)	(0.021)			(0.008)	(0.037)	(0.174)
Machinery and equipment	0.283	0.624	0.085	0.992	-	-	0.055	0.858	0.390
	(0.002)	(0.006)	(0.003)	(0.005)			(0.002)	(0.007)	(0.021)
Autos and transport equipment	0.363	0.602	0.120	1.085	-	-	0.046	0.937	0.730
	(0.004)	(0.014)	(0.010)	(0.010)			(0.004)	(0.009)	(0.084)
Other manufacturing	0.226	0.680	0.129	1.035	-	-	0.053	0.922	0.686
	(0.001)	(0.012)	(0.005)	(0.009)			(0.002)	(0.003)	(0.033)

Notes. The table reports estimates without correcting for demand at all, using the factor share approach, allowing labor to partially adjust to current period shocks: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

F.3 Results using the control function method in French manufacturing data

We report here estimates of production functions, demand parameters and controlled Markov processes across different models—multi-market model, single market model with and without exporters, and without correction for demand—using the control function method. In doing so we cannot apply a quasi-non-parametric approach as we did above when using the factor shares approach. Instead, we must make a slightly stronger assumption on the structure of the production function. Since the data clearly reject a Cobb-Douglas production function, we apply a translog, which is a second order approximation.

We use OLS estimates for setting the initial values for the GMM search, which is a common practice. This procedure is prone to the Ackerberg et al. (2020) critique, whereby the GMM search tends not to move away from the OLS point estimates. However, this does not restrict the results to be similar across different models. Indeed, the results are distinct across models.

Table F.9: Estimates using Multi-Market Model, Control Function Estimator

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.685	0.246	0.107	1.038	0.909	-11.021	-0.002	0.847	-0.014
	(0.086)	(0.041)	(0.024)	(0.087)	(0.116)	(23.559)	(0.002)	(0.020)	(0.013)
Textiles, wearing apparel	0.468	0.416	0.093	0.977	1.020	50.659	0.002	0.860	0.013
	(0.055)	(0.029)	(0.014)	(0.041)	(0.047)	(280.992)	(0.003)	(0.014)	(0.019)
Wood, paper products	0.398	0.464	0.073	0.935	1.035	28.470	0.001	0.718	0.003
	(0.333)	(0.097)	(0.038)	(0.370)	(0.208)	(32.896)	(0.016)	(0.042)	(0.088)
Chemical products	0.592	0.356	0.100	1.049	0.971	-34.610	-0.002	0.871	-0.016
	(1.354)	(0.443)	(0.046)	(1.832)	(0.185)	(706.536)	(0.178)	(0.018)	(1.176)
Rubber and plastic	0.541	0.420	0.085	1.045	0.941	-16.945	-0.003	0.806	-0.018
	(0.412)	(0.023)	(0.065)	(0.488)	(0.149)	(103.942)	(0.028)	(0.014)	(0.160)
Basic metal and fabricated metal	0.324	0.483	0.112	0.919	1.058	17.159	0.003	0.807	0.014
	(0.051)	(0.015)	(0.017)	(0.068)	(0.076)	(134.295)	(0.003)	(0.016)	(0.017)
Computer, electronics	0.527	0.431	0.088	1.045	0.957	-23.480	-0.004	0.820	-0.023
	(0.092)	(0.058)	(0.015)	(0.103)	(0.094)	(228.626)	(0.010)	(0.023)	(0.053)
Electrical equipment	0.520	0.373	0.079	0.972	1.020	49.988	0.001	0.818	0.006
	(8.582)	(3.284)	(0.292)	(12.149)	(0.462)	(279.117)	(0.502)	(0.017)	(2.980)
Machinery and equipment	0.441	0.488	0.069	0.998	0.993	-136.461	-0.002	0.767	-0.008
	(0.032)	(0.039)	(0.004)	(0.067)	(0.076)	(617.599)	(0.005)	(0.009)	(0.019)
Autos and transport equipment	0.623	0.379	0.078	1.081	0.931	-14.571	-0.003	0.831	-0.017
	(0.310)	(0.686)	(0.059)	(0.774)	(0.504)	(139.466)	(0.026)	(0.022)	(0.138)
Other manufacturing	0.372	0.444	0.115	0.931	1.018	55.878	0.001	0.836	0.007
	(0.013)	(0.008)	(0.004)	(0.019)	(0.024)	(3157.226)	(0.001)	(0.012)	(0.006)

Notes. The table reports estimates based on the multi-market model, using the control function estimator: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho-1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.10: Estimates using Single-Market Model (KG), Control Function Estimator

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	6.272	1.400	1.386	9.059	0.101	-1.112	-0.006	0.824	-0.032
	(2.624)	(0.351)	(0.681)	(3.627)	(0.285)	(0.557)	(0.012)	(0.046)	(0.072)
Textiles, wearing apparel	0.242	0.341	0.040	0.624	1.734	1.362	-0.009	0.816	-0.048
	(0.136)	(0.147)	(0.048)	(0.054)	(0.370)	(11.992)	(0.018)	(0.020)	(0.097)
Wood, paper products	1.052	-4.057	-0.015	-3.019	-0.525	-0.656	0.017	0.824	0.095
	(2.211)	(21.919)	(0.577)	(20.339)	(0.849)	(19.361)	(0.155)	(0.029)	(0.896)
Chemical products	1.288	0.259	0.344	1.891	0.501	-2.006	0.036	0.853	0.246
	(0.480)	(0.501)	(0.076)	(0.961)	(0.226)	(2.695)	(0.025)	(0.052)	(0.164)
Rubber and plastic	-2.141	-1.671	-0.391	-4.203	-0.233	-0.811	-0.001	0.805	-0.004
	(5.523)	(1.390)	(1.112)	(7.943)	(0.263)	(0.437)	(0.264)	(0.030)	(2.012)
Basic metal and fabricated metal	0.671	0.041	0.215	0.927	0.920	-12.504	-0.004	0.931	-0.058
	(0.180)	(1.772)	(0.110)	(1.708)	(0.453)	(5.492)	(0.008)	(0.010)	(0.141)
Computer, electronics	0.491	0.459	0.082	1.032	0.968	-30.831	0.001	0.767	0.003
	(0.016)	(0.384)	(0.043)	(0.351)	(0.199)	(46.306)	(0.003)	(0.030)	(0.016)
Electrical equipment	0.541	0.332	0.080	0.953	1.034	29.313	-0.001	0.815	-0.004
	(0.017)	(0.092)	(0.020)	(0.091)	(0.070)	(237.863)	(0.006)	(0.022)	(0.045)
Machinery and equipment	1.207	0.973	0.181	2.361	0.420	-1.723	-0.014	0.805	-0.073
	(0.322)	(4.167)	(0.119)	(4.345)	(0.906)	(117.510)	(0.211)	(0.021)	(0.821)
Autos and transport equipment	0.776	0.105	0.104	0.985	0.995	-198.898	-0.006	0.828	-0.034
	(0.411)	(0.538)	(0.052)	(0.087)	(0.021)	(558.209)	(0.011)	(0.026)	(0.064)
Other manufacturing	0.985	0.624	0.365	1.974	0.464	-1.865	-0.013	0.786	-0.061
	(1.134)	(1.643)	(0.376)	(2.942)	(1.066)	(0.796)	(0.018)	(0.011)	(0.086)

Notes. The table reports estimates based on the single-market model *à la* Klette & Griliches (1996), using the control function estimator: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1 - h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.11: Estimates using Single-Market Model (KG), on Sample of Non-Exporters, Control Function Estimator

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	6.288	1.403	1.390	9.081	0.101	-1.112	-	0.824	-
	(2.568)	(0.355)	(0.666)	(3.555)	(0.234)	(14.683)		(0.050)	
Textiles, wearing apparel	0.243	0.350	0.044	0.637	1.672	1.489	-	0.817	-
	(0.425)	(0.432)	(0.127)	(0.125)	(0.312)	(1.979)		(0.020)	
Wood, paper products	0.831	-2.156	0.048	-1.277	-1.568	-0.389	-	0.821	-
	(1.818)	(17.279)	(0.628)	(16.420)	(0.869)	(42.193)		(0.023)	
Chemical products	7.214	6.731	0.830	14.776	0.071	-1.077	-	0.849	-
	(0.403)	(0.683)	(0.078)	(0.983)	(0.252)	(1.442)		(0.051)	
Rubber and plastic	-2.143	-1.671	-0.391	-4.205	-0.233	-0.811	-	0.805	-
	(7.894)	(2.003)	(1.711)	(11.467)	(0.289)	(0.338)		(0.030)	
Basic metal and fabricated metal	0.669	0.029	0.211	0.908	0.929	-14.178	-	0.931	-
	(0.407)	(2.157)	(0.286)	(2.056)	(0.537)	(4.951)		(0.012)	
Computer, electronics	0.483	0.550	0.065	1.097	0.913	-11.504	-	0.793	-
	(0.025)	(0.103)	(0.026)	(0.080)	(0.076)	(32.094)		(0.030)	
Electrical equipment	0.599	0.361	0.127	1.087	0.924	-13.227	-	0.844	-
	(0.022)	(0.084)	(0.020)	(0.076)	(0.060)	(3113.756)		(0.022)	
Machinery and equipment	1.241	1.002	0.190	2.433	0.405	-1.681	-	0.803	-
	(0.339)	(3.677)	(0.107)	(3.900)	(0.962)	(5.230)		(0.019)	
Autos and transport equipment	0.775	0.101	0.104	0.980	0.995	-195.595	-	0.829	-
	(0.274)	(0.375)	(0.038)	(0.071)	(0.016)	(1405.936)		(0.024)	
Other manufacturing	-0.015	-1.166	-0.002	-1.183	-1.041	-0.490	-	0.801	-
	(0.736)	(1.293)	(0.241)	(2.135)	(1.278)	(1.202)		(0.009)	

Notes. The table reports estimates based on the single-market model *à la* Klette & Griliches (1996), using the control function estimator, where we restrict the sample to non-exporting firms: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho - 1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

Table F.12: Estimates using No Demand Correction, Control Function Estimator

Industry	σ^M	σ^L	σ^K	RTS	ρ	$\eta = \frac{1}{(\rho-1)}$	μ	h	$\mu/(1-h)$
Food, beverage, tobacco	0.613	0.264	0.094	0.971	-	-	0.026	0.948	0.505
	(0.004)	(0.010)	(0.005)	(0.008)			(0.004)	(0.002)	(0.063)
Textiles, wearing apparel	0.463	0.427	0.114	1.005	-	-	0.018	0.955	0.401
	(0.038)	(0.067)	(0.010)	(0.031)			(0.010)	(0.015)	(0.177)
Wood, paper products	-1.383	3.542	-0.090	2.070	-	-	0.161	0.828	0.934
	(0.531)	(0.906)	(0.047)	(0.329)			(0.052)	(0.048)	(0.301)
Chemical products	0.581	0.309	0.117	1.007	-	-	0.021	0.916	0.253
	(0.019)	(0.049)	(0.020)	(0.018)			(0.007)	(0.043)	(0.263)
Rubber and plastic	0.360	0.878	0.045	1.283	-	-	0.016	0.879	0.134
	(0.249)	(0.589)	(0.055)	(0.311)			(0.030)	(0.026)	(0.254)
Basic metal and fabricated metal	0.388	0.458	0.126	0.972	-	-	0.011	0.958	0.271
	(0.018)	(0.026)	(0.006)	(0.004)			(0.002)	(0.014)	(0.049)
Computer, electronics	0.711	0.061	0.209	0.982	-	-	-0.108	0.834	-0.647
	(0.277)	(0.389)	(0.102)	(0.020)			(0.106)	(0.048)	(0.693)
Electrical equipment	0.532	0.394	0.058	0.984	-	-	0.071	0.889	0.634
	(0.245)	(0.360)	(0.064)	(0.075)			(0.078)	(0.041)	(0.649)
Machinery and equipment	0.483	0.433	0.063	0.979	-	-	0.002	0.910	0.020
	(0.340)	(0.385)	(0.029)	(0.021)			(0.038)	(0.055)	(0.242)
Autos and transport equipment	0.632	0.271	0.092	0.994	-	-	0.022	0.879	0.178
	(0.084)	(0.110)	(0.015)	(0.017)			(0.023)	(0.037)	(0.272)
Other manufacturing	0.397	0.443	0.118	0.958	-	-	0.012	0.957	0.273
	(0.045)	(0.093)	(0.019)	(0.031)			(0.011)	(0.033)	(0.081)

Notes. The table reports estimates without correcting for demand at all, using the control function estimator: average output elasticities σ^j for materials input ($j = M$), labor ($j = L$) and capital ($j = M$), overall returns to scale (RTS), the demand elasticity $\eta = 1/(\rho-1)$, the coefficient to learning by exporting μ, the persistence parameter in the controlled Markov h, and the long run effect of exporting $\mu/(1-h)$. Bootstrap standard errors clustered by firm in parentheses.

F.4 OLS Results

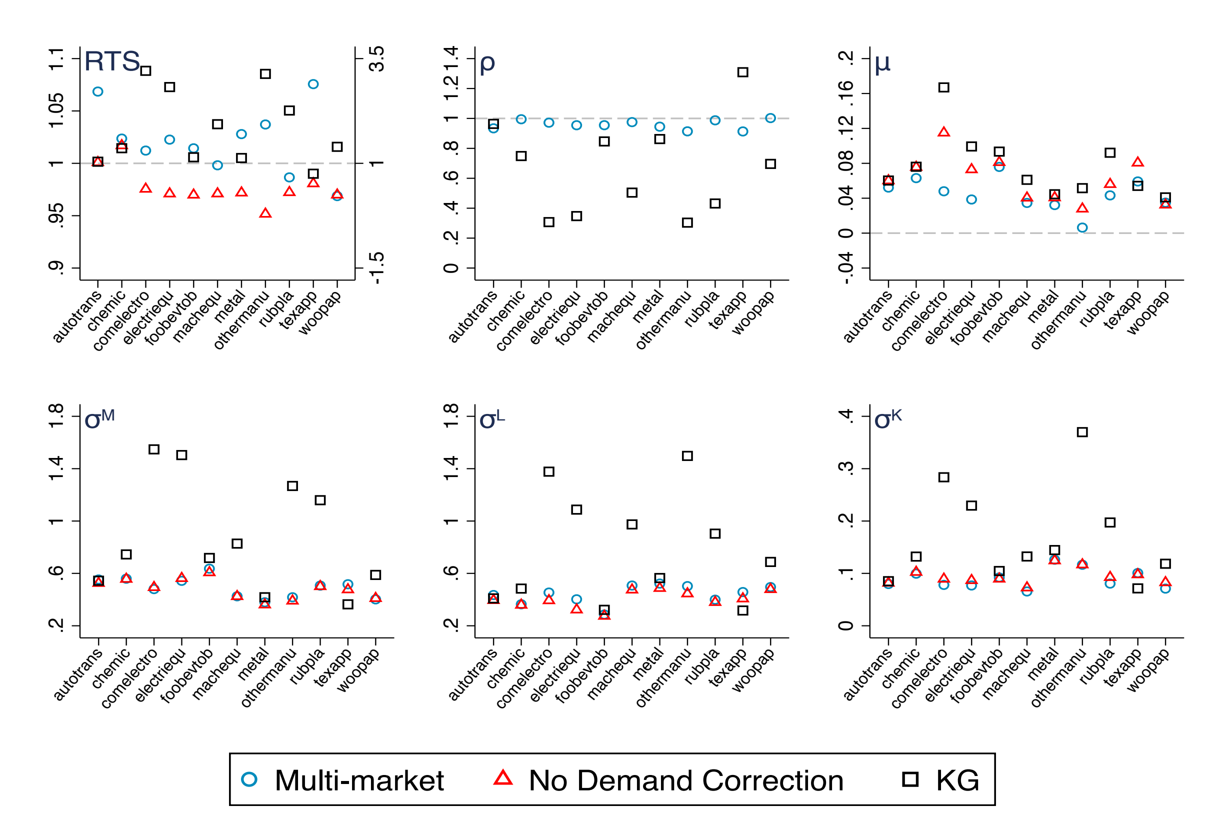
In this section, we present results from OLS estimates of (20), assuming translog form for $f(\cdot)$.

For the multi-market estimator (blue circles) we proxy D_{ft} with just $\frac{R_{ft}}{R^1_{ft}}$, as with OLS there is no first step to identify ψ_{ft}. In this case, the outcome variable is simply log revenues. We include time fixed effects to absorb the D^1_t term.

For the KG method (black squares), we deflate revenues and control for B^{proxy}_t. In this case, we cannot control flexibly for time effects, as the indicator variables would absorb the B^{proxy}_t term.

The no demand correction model (red triangles) includes time fixed effects.

Figure F.4: Estimates using näive OLS



Notes. The figure reports OLS estimates of average total returns to scale, ρ, μ, and average factor output elasticities by industry and estimator. For total returns to scale, the KG estimator is reported on the right axis, while the multi-market and no demand correction estimators are reported on the left axis.

G Comparison to Almunia et al. (2021)

In recent work, Almunia et al. (2021) also specify a model with multiple destinations, monopolistic competition, and quasi-fixed capital. They also estimate production function parameters and the elasticity of demand, as we do. There are three key differences between the production function estimation in Almunia et al. (2021) and ours. First, Almunia et al. (2021) do not control for firm-specific demand shifters when estimating the capital coefficients. Implicitly, Almunia et al. (2021) assume that all firms within a sector sell to the same unique market, despite having written a model with multiple destinations. Second, Almunia et al. (2021) assume that firms are myopic with respect to *ex post* destination specific demand shocks ($E\left[e^{u_{ft}}\right] = 1$). Third, Almunia et al. (2021) implicitly assume constant returns to flexible inputs when estimating the elasticity of demand, which is inconsistent with the thrust of their main findings.

Almunia et al. (2021) claim (see Appendix F.1 of Almunia et al. (2021)) that the assumption of monopolistic competition implies (in their notation) that

$$R_{it} - C_{it}^{v} = \frac{1}{\sigma} R_{it}, \tag{G.1}$$

where R_{it} is total revenues of firm i at time t, C_{it}^{v} denotes the total variable cost of firm i at time t, and σ is the elasticity of substitution. Equivalently, we can define in our notation the total variable cost of firm f in time t,

$$Cost_{ft}\left(Q_{ft}\right) = \sum_{j} e^{v_{ft}^{j}} W_{t}^{j}, \tag{G.2}$$

or assuming just two flexible inputs—materials and labor—(G.2) becomes

$$Cost_{ft}\left(Q_{ft}\right) = e^{m_{ft}} W_{t}^{m} + e^{l_{ft}} W_{t}^{l}. \tag{G.3}$$

We can then re-write (G.1) in our notation

$$R_{ft} - Cost_{ft}\left(Q_{ft}\right) = \left(1 - \rho\right) R_{ft}. \tag{G.4}$$

Based on assumption (G.1), and substituting for C_{t}^{v} with expenditures on flexible inputs, Almunia et al. (2021) derive the following moment condition (in their notation)

$$E\left[\ln\left(\frac{\sigma-1}{\sigma}\right)+r_{it}^{obs}-\ln\left(P_{it}^{M}M_{it}+w_{it}L_{it}\right)\right]=0, \tag{G.5}$$

or in our notation

$$E\left[\ln\rho+r_{ft}-\ln\left(W_{ft}^{m}e^{m_{ft}}+W_{ft}^{l}e^{l_{ft}}\right)\right]=0. \tag{G.6}$$

If the assumption in (G.1) were to hold, then indeed the moment condition (G.6) could be exploited to identify the curvature of demand (ρ, in our notation). But as we show below, (G.1) requires both (in our notation) $E\left[e^{u_{ft}}\right]=1$ and constant marginal costs. This implies that firms are myopic with respect to *ex post* destination specific demand shocks, and that variable returns to scale are unitary.

We can rewrite the optimization problem of a firm from section 2 for a fixed set of destinations using the variable cost function (G.3):

$$\max_{\boldsymbol{\chi_{ft}},Q_{ft}} \quad \mathscr{L} = E\left[Q_{ft}^{\rho}\sum_{d\in\Omega_{ft}}\left(\chi_{ft}^{d}\right)^{\rho}D_{t}^{d}e^{\epsilon_{ft}^{d}+u_{ft}^{d}}\right]-Cost_{ft}\left(Q_{ft}\right)+\lambda_{ft}\left(1-\sum_{d\in\Omega_{ft}}\chi_{ft}^{d}\right) \tag{G.7}$$

which leads to first order condition for Q_{ft}

$$\rho\left(Q_{ft}\right)^{\rho-1}\left[\sum_{d\in\Omega_{ft}}\left(D_{t}^{d}e^{\epsilon_{ft}^{d}}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}E\left[e^{u}\right]=\frac{\partial Cost_{ft}\left(Q_{ft}\right)}{\partial Q_{ft}}, \tag{G.8}$$

Multiplying both sides by Q_{ft}, we have

$$\rho\left(Q_{ft}\right)^{\rho}\left[\sum_{d\in\Omega_{ft}}\left(D_{t}^{d}e^{\epsilon_{ft}^{d}}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}E\left[e^{u}\right]=\frac{\partial Cost_{ft}\left(Q_{ft}\right)}{\partial Q_{ft}}Q_{ft}, \tag{G.9}$$

and substituting with total revenues

$$\rho E\left[e^{u}\right]R_{ft}\psi_{ft}^{-1}=\frac{\partial Cost_{ft}\left(Q_{ft}\right)}{\partial Q_{ft}}Q_{ft}, \tag{G.10}$$

Now if we set $E\left[e^{u}\right]=1$ and we assume $\frac{\partial Cost_{ft}\left(Q_{ft}\right)}{\partial Q_{ft}}Q_{ft}=Cost_{ft}\left(Q_{ft}\right)$, we get the moment condition (G.6). But with non-constant marginal cost, $\frac{\partial Cost_{ft}(Q_{ft})}{\partial Q_{ft}} Q_{ft} \neq Cost_{ft}(Q_{ft})$. In particular, with decreasing returns to flexible inputs – the necessary condition for cross-market complementarities – $\frac{\partial Cost_{ft}(Q_{ft})}{\partial Q_{ft}} Q_{ft} \neq Cost_{ft}(Q_{ft})$.

Hence, the assumptions necessary for identification of the curvature in demand in (G.1) are inconsistent with the mechanism under study in (G.1). Moreover, if one wants to estimate returns to flexible inputs, as we do, the assumptions embedded in (G.6) entail that returns to flexible inputs are unitary, so there is no need to estimate them.