Firm Productivity, Heterogeneity and Exports*

Emir Malikov

University of Nevada, Las Vegas

February 10, 2020

Abstract

Exporters tend to be more productive than non-exporting firms, and uncovering the extent of these productivity differentials as well as roots thereof has long been one of the core topics in the literature. This paper focuses on the issues pertaining to robust empirical modeling of the intricate endogenous relationship between the firm's export behavior and its productivity. I provide a new synthesized methodology for the structural control-function-based identification of unobserved firm productivity in the presence of *(i)* endogenous self-selection into export markets, *(ii)* learning-by-exporting effects on firm productivity and *(iii)* potential heterogeneity in the production of for-export and for-domestic-sale outputs. The model offers a useful addition to the toolkit of practitioners interested in studying the nexus between firm productivity and export behavior. I showcase the proposed methodology by applying it to the firm-level data from China's toy industry, one of the country's most export-oriented manufacturing sectors, during the 1999–2006 period.

Keywords: control function, export, production function, productivity, proxy, selectivity

JEL Classification: C23, D24, F10, L10

* *Email*: emir.malikov@unlv.edu. I would like to thank Subal Kumbhakar and participants of the 2018 North American Productivity Workshop at the University of Miami for many insightful comments on an earlier draft of the paper.

1 Introduction

Exporting firms are generally more productive than their domestically oriented counterparts. Not only are the exporters usually more productive *a priori* (Clerides, Lach & Tybout, 1998; Delgado, Farinas & Ruano, 2002) but they are also more likely to enjoy productivity gains due to learning by exporting, quality and variety effects, exposure to tougher competition as well as the absorption of new technologies from abroad (Pavcnik, 2002; Amiti & Konings, 2007; De Loecker, 2007, 2013). In this paper, I focus on the issues pertaining to robust empirical modeling of this intricate endogenous relationship between the firm's export behavior and its productivity.

I contribute to the literature by offering a new synthesized nonparametric methodology for the structural control-function-based identification of unobserved firm productivity in the presence of *(i)* endogenous self-selection into export markets, *(ii)* dependence of productivity evolution on the firm's past export experiences and, importantly, *(iii)* potential heterogeneity in the production of for-export and for-domestic-sale outputs.

The firm's self-selection into foreign markets is likely to be correlated with its quasi-fixed inputs and productivity which, if unaccounted for, will produce inconsistent and biased estimates of both the production function and productivity thereby making it imperative to model the firm's endogenous export experience. Similar problems may arise when one fails to explicitly model the role of exports in shaping the firm's future productivity (De Loecker, 2013). More importantly, it is reasonable to expect that exporting and non-exporting firms operate with different technologies; after all, exporters tend to be more capital intensive, invest more frequently and adopt new technologies more often (Van Biesebroeck, 2005). Such a heterogeneity may arise, for instance, due to different packaging or labeling requirements for products to be exported abroad. It may also reflect the underlying differentiation in products by their intended markets of sale based on design differences, different capital-labor requirements, the use of somewhat different machinery or materials, etc. This implies a trichotomous heterogeneity in production across non-exporters, complete exporters (both types targeting single markets) and partial exporters selling their products in two markets. To my knowledge, the distinction between partial and complete exporters of this kind has not been previously studied in the literature. In such a flexible framework, the firm's self-selection into exporting also entails a subsequent reconfiguration of its production. Taking this (endogenously determined) heterogeneity for granted would therefore amount to the bias-inducing model misspecification capable of producing misleading conclusions about the interplay between the firm's productivity and its export behavior. I handle the latter by adapting De Loecker, Goldberg, Khandelwal & Pavcnik's (2016) ideas for modeling multi-product production to the case of production in the presence of export opportunities.

To identify heterogeneous production technologies, I assume that complete non-exporters and complete exporters use the same technology to produce outputs intended for sale in a given market as do partial exporters. I therefore can use data on such firms to estimate heterogeneous intended-market-specific production functions, with the knowledge of which I then can recover unobservable cross-product input allocations within partial exporter firms. Among other things, this differential treatment of non-, partial and complete exporters (while controlling for self-selection) also permits the identification of endogenously export-dependent latent firm productivity which, in turn, enables me to consistently estimate the exporter productivity premium. Unobserved firm-level productivity is proxied via a control-function approach à la Olley & Pakes (1996) and Levinsohn & Petrin (2003). The multi-step model is implemented using log-polynomial sieves with the selection process estimated via Khan's (2013) semiparametric distribution-free sieve nonlinear least-squares estimator. The proposed methodology constitutes a useful addition to the toolkit of practitioners interested in studying the nexus between firm productivity and export behavior, the work on which has proliferated greatly since Bernard & Jensen's (1995) seminal paper on the subject matter (e.g., Aw & Hwang, 1995; Aw, Chung & Roberts, 2000; Baldwin & Gu, 2003; Girma, Greenaway & Kneller, 2004; Wagner, 2007, in addition to those already cited).

My methodology complements other available approaches to tackling export behavior in the structural production-function estimation. While De Loecker (2013) also pays close attention to the export dependence of firm productivity, my approach to endogenous exports goes a step farther by also explicitly modeling the firm's export status self-selection in addition to accommodating potential heterogeneity in the production of products intended for sale in domestic versus foreign export markets. Given the latter two aspects, this paper's framework is also more general than that by Van Biesebroeck (2005), De Loecker (2007) and Malikov, Zhao & Kumbhakar (2020) who model endogenous exports in the input-allocation paradigm whereby changes in firm's export status are conceptualized as an "investment" in a dynamic state variable as opposed to specifying an explicit export decision rule. My methodology also presents a significant improvement over few earlier attempts in the literature to explicitly model self-selection into exporting such as Girma et al. (2004), who estimate the productivity-dependent selection process *after* already having recovered the estimates of productivity under the assumption of its independence from the firm's export behavior. In contrast, consistent with theoretical underpinnings (e.g., Melitz, 2003; Melitz & Ottaviano, 2008), I model the interdependence of the two simultaneously.[1]

I showcase the proposed methodology by applying it to the firm-level data from China's "toy, stationery, sports and cultural goods" industry during the 1999–2006 period. China is the major global toy producer and exporter, with the estimated 80% of toys in the world made in China (HKTDC, 2019). The toy industry is one of China's most export-oriented manufacturing sectors (75% firms in our sample engage in exports) which makes it an especially suitable application to illustrate merits of my model. I find overwhelming evidence in support of the statistically and economically significant exporter productivity premium in the industry, both at fixed marginal moments as well as across the entire distribution of firms. Importantly, this conclusion would have been completely different had I not explicitly accommodated the heterogeneity across exporters and non-exporters subject to their endogenous self-selection during the estimation of latent firm productivity. Namely,

[1]Incidentally, so do Clerides et al. (1998) who are a notable exception. However, unlike me, they do it parametrically under the distributional assumption and also, when doing so, they do not differentiate between partial and complete exporters as well as do not allow for heterogeneous production technologies.

when using conventional proxy-based estimators popularly used in the empirical literature on the export–productivity nexus, I find that the estimated productivity hardly differs across exporters and non-exporters and one misleadingly cannot reject their equality. The latter speaks in support of a richer modeling framework capable of allowing for endogenously determined cross- and within-firm heterogeneity dependent on the firms' export behavior like the one proposed here.

The rest of the paper proceeds as follows. Section 2 describes the conceptual framework of heterogeneous production in the presence of export opportunities. I discuss the identification and estimation strategy in Section 3. Section 4 briefly describes the data. The results are discussed in Section 5 followed by concluding remarks in Section 6.

2 Conceptual Framework

Consider the production process of a firm $i = 1, \ldots, N$ at time period $t = 1, \ldots, T$ where its physical capital K_{it}, labor L_{it} and an intermediate input such as materials M_{it} are transformed into (possibly two) outputs $Y_{s,it}$ differentiated on the basis of their intended markets of sale—domestic or export—respectively indexed by $s \in \{\mathrm{d}, \mathrm{e}\}$. The outputs are produced via time-varying (to capture temporal technological changes) intended-market-specific production technologies:

$$Y_{s,it} = F_{s,t}(K_{s,it}, L_{s,it}, M_{s,it}) \exp\{\omega_{it} + \eta_{s,it}\} \qquad \forall\ s \in \{\mathrm{d}, \mathrm{e}\}, \tag{2.1}$$

where $\exp(\omega_{it} + \eta_{s,it})$ is the log-additive "composite" Hicks-neutral productivity term consisting of *(i)* the firm's persistent productivity ω_{it} and *(ii)* a random transitory productivity shock $\eta_{s,it}$.[2] The error $\eta_{s,it}$ is allowed to be output-specific and is said to be *i.i.d.* with a normalized zero mean: $\mathbb{E}[\eta_{s,it}|\ \mathcal{I}_{it}] = \mathbb{E}[\eta_{s,it}] = 0\ \forall s$, where $\mathcal{I}_{it}$ denotes the information available to the firm i for making period t decisions. This shock is observable by firms in period t only *ex post* after all production decisions take place.

By allowing the production function $F_{s,t}(\cdot)$ to vary with s, this paper's framework explicitly accommodates potential heterogeneity in the production of outputs intended for sale in the domestic versus foreign export markets. Such a heterogeneity may arise, for instance, due to different packaging or labeling requirements for products to be exported abroad; it may also reflect the underlying differentiation (horizontal or vertical) in products themselves, say, involving design differences, etc. Econometrically, by modeling heterogeneous production functions, I am able to allow for potential discontinuities in the production function across exporting and non-exporting firms. Function $F_{s,t}(\cdot)$ is still assumed to satisfy the standard neoclassical assumptions including differentiability, positive monotonicity and concavity in inputs.

Motivated by De Loecker et al. (2016), all inputs are assumed to be attributable to the either type of outputs so that each output-specific input can be written as a share of the firm's total input. Specifically, defining $\rho_{it} \in [0, 1]$ as a share of inputs used in the production of $Y_{\mathrm{d},it}$, I have

[2] Sometimes the latter shock is alternatively interpreted as a measurement error in the output.

that $V_{\mathrm{d},it} = \rho_{it} V_{it}$ and $V_{\mathrm{e},it} = (1-\rho_{it})V_{it}$ for $V \in \{K, L, M\}$. Note that ρ_{it}, reflective of the input allocation across product lines within the firm, is *un*observed. Further, following Gandhi, Navarro & Rivers (2017) and Malikov et al. (2020), I assume that both K_{it} and L_{it} are subject to adjustment frictions (e.g., time-to-install, hiring costs) and thus are quasi-fixed, whereas M_{it} is freely varying. That is, M_{it} is determined by the firm in period t, while K_{it} and L_{it} are determined in period $t-1$. Being dynamic inputs, K_{it} and L_{it} are the state variables evolving according to deterministic laws of motion whereby $K_{it} = I_{it-1} + (1-\delta)K_{it-1}$ and $L_{it} = H_{it-1} + L_{it-1}$ with I_{it}, H_{it} and δ being the gross investment, net hiring and the depreciation rate, respectively.

Following popular practice in modeling multi-product firms, I assume that persistent productivity ω_{it} is *firm*-specific and applies to the production of both the domestic and export products (should the firm engage in both). The latter seems appropriate especially given that ω_{it} is normally interpreted as capturing unobserved firm characteristics such as management quality and tacit knowledge that are important for the production of outputs *regardless* of their intended market of sale. At the same time, I allow the evolution of ω_{it} to be impacted by the firm's past export behavior to capture the so-called "learning by exporting" by letting the firm productivity follow a *controlled* first-order Markov process of the following form (De Loecker, 2013):

$$\omega_{it} = h(\omega_{it-1}, \mathbf{X}_{it-1}) + \zeta_{it} \quad \text{with} \quad \mathbb{E}[\zeta_{it} | \, \mathcal{I}_{it-1}] = 0, \tag{2.2}$$

where $h(\cdot)$ is an unknown function, and $\mathbf{X}_{it}$ includes some measures of the firm's endogenously determined export experience. This export dependence of productivity is usually rationalized by the argument that exporters are exposed to tougher competition, which induces them to raise their productivity, and are more likely to enjoy productivity gains due to learning, quality/variety effects and new technologies imported from abroad. In this paper, I set $\mathbf{X}_{it}$ to be a (scalar) measure of the firm's export intensity (x_{it}) conventionally defined as the nominal share of its total output produced for the sale in export markets. Such a measure facilitates modeling the firm's productivity-enhancing export experiences in a richer, continuous framework. Lastly, it is convenient to interpret ζ_{it} in (2.2) as a mean-zero innovation in persistent productivity, unobservable to firms in period $t-1$. Firms do however observe ω_{it}, consisting of the conditional mean $h(\cdot)$ and innovation ζ_{it}, in period t when decisions concerning freely varying inputs are being made.

Based on a (continuous) export intensity measure $x_{it} \in [0,1]$, I can characterize the firm's (discrete) export status in a trichotomous fashion by differentiating between *(i)* domestic-market-only oriented non-exporters for whom $Y_{\mathrm{e},it} = 0$ and thus $x_{it} = 0$, *(ii)* export-market-only oriented "complete" exporters for whom $Y_{\mathrm{d},it} = 0$ and thus $x_{it} = 1$, and *(iii)* "partial" exporters selling their products both domestically and abroad for whom $Y_{\mathrm{d},it} > 0$ and $Y_{\mathrm{e},it} > 0$ implying that $x_{it} \in (0,1)$. As it is to become apparent in Section 3, such a ternary categorization of firms enables me to identify heterogeneous intended-market-specific production functions as well as the (endogenous) exporter productivity premium. More concretely, expanding on Olley & Pakes's (1996) methodology, I model the firm's endogenous choice of its export status in the form of a productivity-driven self-selection rule whereby, consistent with economic theory (Melitz, 2003) and empirical evidence (Delgado et al., 2002), engagement in the export market requires higher productivity. In addition, it is reasonable to presuppose that the degree of firm's export orientation (partial vs. complete) is an increasing function of its productivity, which can be rationalized by the increasing marginal cost of exporting. With this, the firm's export status selection rule can be stylized as

$$\Xi_{it} = \begin{cases} 1 & \text{if } \begin{cases} \omega_{it} \leq \underline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}) & \text{for a non-exporter with } Y_{\mathrm{d},it} > 0, Y_{\mathrm{e},it} = 0 \\ \underline{\omega}_{it}(\cdot) < \omega_{it} < \overline{\omega}_{it}(\cdot) & \text{for a partial exporter with } Y_{\mathrm{d},it} > 0, Y_{\mathrm{e},it} > 0 \\ \omega_{it} \geq \overline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}) & \text{for a complete exporter with } Y_{\mathrm{d},it} = 0, Y_{\mathrm{e},it} > 0 \end{cases} \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

where Ξ_{it} is a binary indicator such that it equals one if, in period t, the firm decides to leave its export status *unchanged* for the period $t+1$ (i.e., $\Delta x_{it+1} = 0$) and zero otherwise.

Effectively, equation (2.3) models the firm's export status (non-exporter, partial exporter and complete exporter) as an increasing step-function of productivity as captured by (cut-off) thresholds $\underline{\omega}_{it}(\cdot)$ and $\overline{\omega}_{it}(\cdot)$ which, respectively, represent the level of productivity below which the firm is better-off selling all of its products domestically and the level of productivity above which the firm finds itself better-off selling all products abroad. Firms with productivity levels between the thresholds sell outputs in both markets. These productivity thresholds are functions of the production-related state variables K_{it}, L_{it} and $\mathbf{X}_{it}$ as well as some contextual variables $\mathbf{S}_{it}$ capturing relevant exogenous firm characteristics [more on the latter below].

From (2.3), it is evident that the decisions pertaining to the firm's export behavior in period t are assumed to be made in period $t-1$. That is, I assume that changes in the firm's export status are costly and thereby subject to delay, which is meant to capture irreversible adjustment costs that may include time for and cost of finding new intermediaries/buyers, contract (re)negotiations, obtaining new permits and, more importantly, reconfiguring production if products intended for sale abroad are distinct from those sold domestically.[3] Such sunk costs associated with exporting are among the key features of modern trade models with heterogeneous firms. For instance, Melitz & Ottaviano (2008) show that the existence of such costs produces a productivity threshold around which the firm's (productivity-based) export status rule is defined. Similar selection rules with productivity cut-offs also appear in the models of heterogeneous multi-product firms in which changes in the mix of produced products entail irreversible adjustment costs (Mayer, Melitz & Ottaviano, 2014). My framework shares some commonalities with the latter models as well since, in light of the differentiation of outputs based on their intended markets of sale, partial exporters essentially can be conceptualized as multi-product firms.

Overall, modeling export decisions via (2.3) enables me to control for the firm's export self-selection although it is worthy of emphasis that, in the estimation, I do not *a priori* impose the

[3]A quasi-fixed treatment of the firm's export status is along the lines of that implicitly assumed by De Loecker (2013).

productivity ranking across the three types of firms implied by the above selection rule. Hence, the estimation results are not driven by this assumptions (as to become more evident from the self-selection correction procedure in Section 3). To this end, note that the cut-offs in (2.3) are firm-specific and, thus, the implied productivity ranking across zero, partial and complete exporting is *within*-firm and need not uniformly hold *across* firms. Therefore, the instances when, say, owing to the differences in contextual state variables not all firms with the same productivity level are deemed "productive enough" to engage in exports (partially or completely) have not been not ruled out. Furthermore, the formulation in (2.3) whereby the firm's complete export orientation is predicated on it being more productive than a partial exporter is also not of consequence here and can be switched, if desired. The latter would then be more aligned with the conceptualization of partial exporters as multi-product (i.e., two-product) firms which are oftentimes argued to be more productive than the less diversified firms (here, one-product non- and complete exporters).

Given the dynamic state-variable-like implications for $\mathbf{X}_{it}$ that follow from the firm's export status selection rule in (2.3) and its underlying assumption that adjustments in export behavior are costly and delayed by a period, I can derive an implicit law of motion for $\mathbf{X}_{it}$ whereby $\mathbf{X}_{it} = \mathbf{X}_{it-1} + \mathcal{X}_{it-1}$, where $\mathcal{X}_{it}$ is the net adjustment in the firm's export behavior. As implied by (2.3), the optimal policy function for such adjustments can be represented by $\mathcal{X}_{it} = \mathbb{X}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$, where I omit cut-off thresholds because they too depend on $(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it})$.

To allow the exporting and non-exporting firms to allocate their inputs differently, I also let the firm's optimal policy functions for total input quantities depend on its predetermined export behavior à la Van Biesebroeck (2005) and Malikov et al. (2020): the gross investment function $I_{it} = \mathbb{I}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$, the net hiring function $H_{it} = \mathbb{H}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$ and the conditional demand for materials $M_{it} = \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$.

Lastly but not least importantly, in the discussion above, I have allowed the firm's production decisions to be affected by contextual variables $\mathbf{S}_{it}$ capturing relevant firm characteristics. My choice of such variables is dictated by the context of empirical application. First, one of the salient features of the Chinese production data is that the government has a stake in many manufacturing firms. Some firms are also the recipients of direct government subsidies. Both aspects proxy well for the firms' political connections which are of non-trivial importance given China's institutions as well as the government's active public policies aimed at promoting exports. Another feature of China's manufacturing is that it is a large recipient of foreign direct investment. Just like with the state interest in firm equity, foreign ownership (be it partial or complete) of firms is also likely to influence both its production and exporting decisions. Thus, it is imperative to condition the firm's decisions on these firm characteristics.

I respectively measure the government's and foreign investors' participation in the firm's ownership using the ratio of government-owned and foreign-owned equity over total equity. Let these government and foreign equity shares be denoted by $G_{it} \in [0, 1]$ and $D_{it} \in [0, 1]$, respectively. In addition, I use a binary "subsidy recipient" indicator $B_{it} \in \{0, 1\}$ if the firm receives any subsidy from the government. Together, $\mathbf{S}_{it} = (G_{it}, D_{it}, B_{it})'$.

Obviously, the $\mathbf{S}_{it}$ characteristics are the result of both the firm's behavior (e.g., submitting applications for subsidies, soliciting foreign investments) as well as the decisions by external parties (e.g., the government adjusting its share in the firm equity or approving/denying subsidies, foreign companies choosing the target of their investments). Given these complexities, I model the determination of $\mathbf{S}_{it}$ and its evolution over time without a full-fledged structural conceptualization. Instead, I adopt a more reduced-form approach and model $\mathbf{S}_{it}$ in the "dynamic input allocation" paradigm whereby the elements in $\mathbf{S}_{it}$ are treated as dynamic state variables with the changes therein being subject to adjustment frictions. In such a setup, the government equity share G_{it}, the foreign equity share D_{it} and the subsidy recipient indicator B_{it} are determined at time $t-1$ akin to dynamic inputs and evolve according to their respective laws of motion: $G_{it} = G_{it-1} + \Delta G_{it-1}$, $D_{it} = D_{it-1} + \Delta D_{it-1}$ and $B_{it} = B_{it-1} + \Delta B_{it-1}$, where "$\Delta$" denotes the adjustments. The adjustments $\Delta\mathbf{S}_{it} = (\Delta G_{it}, \Delta D_{it}, \Delta B_{it})'$ are said to depend on all state variables including firm productivity $(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$ which describe the firm's *internal* characteristics, as well as on an exogenous variable ξ_{it} representative of *external* factors that also affect the determination of $\mathbf{S}_{it}$ and are attributable to the government and/or foreign investors. That is, $\Delta\mathbf{S}_{it} = \mathbb{S}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it}, \xi_{it})$.

Two points are worthy of emphasis here. First, by conditioning $\Delta\mathbf{S}_{it}$ on ω_{it}, I account for the important role that firm productivity plays in the government's and foreign investors' considerations of investments into firms. However, via costliness of adjustments in $\mathbf{S}_{it}$, I effectively assume that firms do not experience *immediate* changes in their government/foreign equity share and subsidy status in light of productivity shocks. This structural timing assumption renders $\mathbf{S}_{it}$ weakly exogenous with respect to both ζ_{it} and $\eta_{s,it}$ because $\mathbf{S}_{it}$ (via $\Delta\mathbf{S}_{it-1}$) depends on ω_{it-1}. Second, the presence of an exogenous determinant ξ_{it} from *outside* the firm's production environment is important for identification of the production function because it enables me to circumvent the Gandhi et al. (2017) critique. Crucially, such a variable need not be observed for the estimation purposes. I discuss this in more detail in the next section.

3 Identification and Estimation Strategy

The identification of *heterogeneous* production functions relies on the fact that the fully specialized firms (non-exporters and complete exporters) use the same technology to produce outputs intended for sale in a given market as do partial exporters. I therefore can use data on non-exporters and complete exporters to estimate intended-market-specific production functions $F_{s,t}(\cdot)\ \forall\ s \in \{\mathrm{d}, \mathrm{e}\}$. Further, separate treatment of the three types of firms (while controlling for self-selection) permits the identification of endogenously export-dependent latent productivity which, in turn, enables me to measure the exporter productivity premium (if any).

3.1 Non- and complete exporters

Non-exporters ($s = \mathrm{d}$) and complete exporters ($s = \mathrm{e}$) sell their products in one market only thereby rendering the differentiation between intended-market-*specific* and *total* outputs/inputs redundant. Hence, for them, (2.1) simplifies to

$$Y_{it} = F_{s,t}(K_{it}, L_{it}, M_{it}) \exp\{\omega_{it} + \eta_{it}\} \tag{3.1}$$

where I intentionally continue to index the production function with subscript s to emphasize that the two types of firms may use different technologies. Importantly, the estimation of (3.1) requires *no* information on unobserved within-firm input allocations across product lines. In what follows, the discussion of identification applies to non-exporters and complete exporters one at a time, implying *separate* estimations for the two types of firms.

Taking logs of both sides of (3.1) and using the Markovian nature of ω_{it} from (2.2) yields

$$y_{it} = f_{s,t}(K_{it}, L_{it}, M_{it}) + h\left(\omega_{it-1}, \mathbf{X}_{it-1}\right) + \zeta_{it} + \eta_{it}, \tag{3.2}$$

where the lower-case variables/functions denote those in logs. I proxy for unobservable productivity in the spirit of Levinsohn & Petrin (2003) by making use of the conditional material demand function. Assuming that $\mathbb{M}(\cdot)|M_{it} > 0$ is strictly monotonic in a scalar ω_{it} for any given $(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it})$, the demand function can then be inverted to control for productivity via $\omega_{it} = \mathbb{M}^{-1}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it})$. Substituting for ω_{it-1}, from (3.2) I get

$$y_{it} = f_{s,t}(K_{it}, L_{it}, M_{it}) + \varphi\left(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}\right) + \zeta_{it} + \eta_{it}, \tag{3.3}$$

where $\varphi(\cdot) \equiv h\left(\mathbb{M}^{-1}(\cdot), \cdot\right)$ is another unknown function.

Since the data are realized for the self-selected sample of non-exporter or complete exporter firms, ignoring it would lead to the selection (omitted variable) bias. More formally, the expectation of (3.3) conditional on $\mathcal{I}_{it-1}$ and the fully specialized firm's decision to continue its current export behavior ($\Xi_{it} = 1$) is given by

$$\begin{aligned}\mathbb{E}[y_{it}|\, \mathcal{I}_{it-1}, \Xi_{it} = 1] &= \mathbb{E}[f_{s,t}(K_{it}, L_{it}, M_{it})|\, \mathcal{I}_{it-1}] + \varphi(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \mathbb{E}[\zeta_{it}|\, \mathcal{I}_{it-1}, \Xi_{it} = 1] \\ &\equiv \mathbb{E}[f_{s,t}(K_{it}, L_{it}, M_{it})|\, \mathcal{I}_{it-1}] + \begin{cases} \psi(\underline{\omega}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) & \text{for non-exporter} \\ \psi(\overline{\omega}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) & \text{for complete exporter,} \end{cases}\end{aligned} \tag{3.4}$$

where unknown function $\psi(\cdot)$ is the sum of separably unidentifiable $\varphi(\cdot)$ and the unknown selection bias term. Both unobservable productivity thresholds $\underline{\omega}_{it}$ and $\overline{\omega}_{it}$ can be proxied using the firm's export status self-selection probability (propensity score) given by

$$\mathbb{P}\mathrm{r}[\Xi_{it}=1|\ \mathcal{I}_{it}] = \begin{cases} \mathcal{P}(\underline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it}),K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for non-exporter} \\ \mathcal{P}(\overline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it}),K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for complete exporter} \end{cases}$$
$$= \mathcal{P}(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) \equiv \varrho_{it}. \tag{3.5}$$

See Appendix A for the detailed derivation of (3.4) and (3.5).

Then, inverting for thresholds via $\mathcal{P}^{-1}(\varrho_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})$ using the first equality in (3.5), from (3.4) I get

$$\mathbb{E}[y_{it}|\ \mathcal{I}_{it-1},\Xi_{it}=1] = \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\ \mathcal{I}_{it-1}] + \psi(\varrho_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}), \tag{3.6}$$

where $\psi\left(\mathcal{P}^{-1}(\varrho_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}),K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}\right) \equiv \psi(\varrho_{it},K_{it-1},$ $L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})$. Thus, the probability ϱ_{it} is effectively a control function for the endogenous selection bias. Note that ϱ_{it} depends on the same covariates that also directly appear inside the proxy function $\psi(\cdot)$. To ensure that the rank condition holds, some structure on the relationship between ϱ_{it} and its determinants needs be imposed, which can be done in a number of flexible semiparametric ways. In doing so, $\psi(\cdot)$ is identified by (partially parametric) "functional form."[4]

Thus, the selection-bias-corrected analogue of (3.3) is given by

$$y_{it} = f_{s,t}(K_{it},L_{it},M_{it}) + \psi(\varrho_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) + \zeta^*_{it} + \eta_{it}, \tag{3.7}$$

where the new innovation $\zeta^*_{it} \equiv \zeta_{it} - \lambda(\cdot)$ is such that $\mathbb{E}[\zeta^*_{it}|\ \mathcal{I}_{it-1},\Xi_{it}=1]=0$ by construction, with $\lambda(\cdot)$ being the selection bias term defined in (A.1) of Appendix A.

To complete identification of the selection-bias-corrected intended-market-specific production function $f_{s,t}(\cdot)$ in (3.7), the remaining challenge that needs tackling is the endogeneity of M_{it} with respect to ζ^*_{it} (through ζ_{it}) which the firm observes at time t before choosing freely varying materials. The remaining right-hand-side covariates are weakly exogenous under the structural assumptions (predetermined at time $t-1$) and therefore self-instrument.

Gandhi et al. (2017) have recently shown that the proxy-based production function estimators suffer from inherent under-identification due to the presence of freely varying inputs in the production function that lack *relevant* excluded instruments thereby failing to satisfy the rank condition despite the abundance of exogenous, albeit irrelevant, lagged inputs. Thus, the identification of (3.7) can only be achieved if the additional predetermined variables provide some excluded *relevant* (exogenous) variation for M_{it} after conditioning on the already included self-instrumenting variables. Below I show that the Gandhi et al. (2017) critique does not apply to my model.

Under the adopted structural assumptions, the freely varying input M_{it} appearing inside the production function $f_{s,t}(\cdot)$ in (3.7) has a valid instrument, namely $\mathbf{S}_{it}$, which provides *excluded relevant*

[4] Doraszelski & Jaumandreu (2013, p.1349) use a similar identification scheme. So do Olley & Pakes (1996, p.1276).

exogenous variation, conditional on the already included self-instrumenting $(K_{it}, L_{it}, \varrho_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$. I am therefore able to identify the model.

The (structural) validity of $\mathbf{S}_{it}$ as an instrument rests on *(i)* its exogeneity whereby $\mathbb{E}[\zeta^*_{it} + \eta_{it}|\, \mathbf{S}_{it}] = 0$, *(ii)* the containment of independent variation from *outside* the production model, and *(iii)* its relevancy for the firm's choice of M_{it}. The first condition holds owing to the quasi-fixed, predetermined nature of firm's contextual characteristics due to the costly adjustment therein. To see how $\mathbf{S}_{it}$ meets the remaining two conditions, note that the freely varying input M_{it}

$$\begin{aligned} M_{it} &= \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it}) \\ &= \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, h(\omega_{it-1}, \mathbf{X}_{it-1}) + \zeta_{it}) \\ &= \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \varphi(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \zeta_{it}), \end{aligned} \tag{3.8}$$

is a function of the following observables $(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$. Comparing the latter set of variables with those that enter the right-hand side of (3.7) and self-instrument, it is evident that M_{it} has an extra source of relevant exogenous variation that is excluded from the estimating equation contained in $\mathbf{S}_{it}$, which can therefore be used as a valid instrument for M_{it}. Using the evolution process for $\mathbf{S}_{it}$, I have that

$$\begin{aligned} \mathbf{S}_{it} &= \mathbf{S}_{it-1} + \Delta\mathbf{S}_{it-1} \\ &= \mathbf{S}_{it-1} + \mathbb{S}(K_{it-1}, L_{it-1}, X_{it-1}, \mathbf{S}_{it-1}, \omega_{it-1}, \xi_{it-1}) \\ &= \mathbf{S}_{it-1} + \mathbb{S}(K_{it-1}, L_{it-1}, X_{it-1}, \mathbf{S}_{it-1}, \mathbb{M}^{-1}(K_{it-1}, L_{it-1}, X_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), \xi_{it-1}), \end{aligned} \tag{3.9}$$

from where it can be easily seen that, conditional on the self-instrumenting variables $(K_{it}, L_{it}, \varrho_{it}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), K_{it-1},\ L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$ in model (3.7), $\mathbf{S}_{it}$ provides relevant independent (excluded) exogenous variation for M_{it} originating from the exogenous, albeit unobservable, ξ_{it-1}. Therefore, I can use $\mathbf{S}_{it}$ as an excluded valid instrument for M_{it} which ultimately enables me to identify (up to a constant) the production function $f_{s,t}(\cdot)$ in (3.7) on the basis of the following moment conditions:

$$\mathbb{E}\left[\zeta^*_{it} + \eta_{it} \left|\, \begin{matrix} \varrho_{it}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), \\ K_{it}, L_{it}, \mathbf{S}_{it}, L_{it-1}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}, \end{matrix} \Xi_{it} = 1 \right.\right] = 0. \tag{3.10}$$

While, owing to their predeterminedness, neither M_{it-1} nor the second- and higher-order lags can also be added to the conditioning set because these additional instruments are irrelevant for predicting M_{it}, conditional on the already included instrument set. As can be seen from (3.8), conditional on $(K_{it}, L_{it}, \mathbf{S}_{it},\ \varrho_{it}, K_{it-1},\ L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$, the endogenous M_{it} has only one source of "free" variation left which is the endogeneity-inducing unobservable ζ_{it}.[5]

[5] Another critique of proxy-based estimators (Ackerberg, Caves & Frazer, 2015) concerns *separable* identifiability of additive components in the model after conditioning the estimating equation on the instrument set. In this paper's particular case, such a concern would be applicable to two additive unknown functions $f_{s,t}(\cdot)$ and $\psi(\cdot)$ in (3.7). Such a problem may arise in the wake of perfect functional dependence between *endogenous* freely varying inputs

In order to identify persistent productivity of fully specialized firms, I go back to (3.2). Firm productivity can be identified (up to a constant) by subtracting the already identified $f_{s,t}(\cdot)$ from both sides of equation:

$$y^*_{it} \equiv y_{it} - \widehat{f}_{s,t}(K_{it}, L_{it}, M_{it}) = \nu(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it}) + \eta_{it}, \tag{3.11}$$

where $\nu(\cdot) \equiv \mathbb{M}^{-1}(\cdot)$, on the basis of the following moment condition:[6]

$$\mathbb{E}[\eta_{it}|\ K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it}] = 0. \tag{3.12}$$

Thus, the estimation methodology is to be implemented for non-exporters and complete exporters separately in three steps: *(i)* estimate export status propensity in (3.5); *(ii)* estimate production functions in (3.7) via method of moments based on (3.10); *(iii)* estimate firm productivity from (3.11) via least squares. If desired, all stages can be cast as a sequential multi-equation semi-nonparametric GMM estimator (with the instrument sets varying across equations) along the lines of Newey (1984) and Wooldridge (2009). Not only would this improve efficiency but it also permits the derivation of the covariance matrix that accounts for the use of generated covariates in the second and third stages. The resulting estimator is consistent and asymptotically normal under relatively mild regularity conditions (Chen, Hahn, Liao & Ridder, 2015).

3.2 Partial exporters

To identify partial exporters' latent productivity, one needs to recover unobserved input allocations ρ_{it} within their two-product technology, which is a mix of the already identified $f_{\mathrm{d},t}(\cdot)$ and $f_{\mathrm{e},t}(\cdot)$. For this, I adapt De Loecker et al.'s (2016) methodology designed for multi-product firms.

I first identify transitory productivity shocks $\eta_{\mathrm{d},it}$ and $\eta_{\mathrm{e},it}$ from the following *two* "reduced-form" intended-market-specific production functions (in logs) for each partial exporter:

$$y_{s,it} = f_{s,t}(K_{s,it}, L_{s,it}, M_{s,it}) + \omega_{it} + \eta_{s,it} \equiv g_{s,t}(K_{it}, L_{it}, M_{it}) + \omega_{it} + \eta_{s,it} \qquad \forall\ s \in \{\mathrm{d}, \mathrm{e}\}, \tag{3.13}$$

where the reduced-form function $g_{s,t}(\cdot)$ differs from its structural counterpart $f_{s,t}(\cdot)$ in that it is modeled as a function of observable *total* inputs. I estimate (3.13) by applying the three-step methodology outlined in Section 3.1 *twice* using different output $y_{s,it}$ on the left-hand side. See Appendix B for details. The procedure identifies transitory shocks via $\widehat{\eta}_{\mathrm{s},it} = y_{s,it} - \widehat{g}_{s,t}(K_{it}, L_{it}, M_{it}) - \widehat{\nu}_s(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it})\ \forall\ s \in \{\mathrm{d}, \mathrm{e}\}$. The reduced-form estimation however does *not* identify firm-

inside $f_{s,t}(\cdot)$ and the arguments of the proxy function $\psi(\cdot)$. In my model, I am able to separably identify these two functions because, conditional on the instrument set $(K_{it}, L_{it}, \mathbf{S}_{it},\ \varrho_{it}, K_{it-1},\ L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$ in (3.7), M_{it} still has a source, albeit unobservable, of *independent* variation, namely ζ_{it} as can be seen in (3.8). Thus, my model is immune to the perfect functional dependence problem.

[6] In practice, $\mathbf{X}_{it}$ is to be omitted from (3.11) and the moment condition because it contains no variation within a group of non-exporters or complete exporters since the firms' qualification for the group is based on $\mathbf{X}_{it}$ being the same across units.

specific persistent productivity because its proxy $\widehat{\nu}_s(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it})$, now s-indexed, produces different estimates of ω_{it} in the third step for $s \in \{\mathrm{d}, \mathrm{e}\}$. To obtain estimates of ω_{it}, equally applicable to the production of both $Y_{\mathrm{d},it}$ and $Y_{\mathrm{e},it}$, I numerically solve the following nonlinear system of two "structural" production functions for *each* partial-exporter firm with respect to (ω_{it}, ρ_{it}):

$$y_{\mathrm{d},it} - \widehat{f}_{\mathrm{d},t}(\rho_{it}K_{it}, \rho_{it}L_{it}, \rho_{it}M_{it}) - \omega_{it} - \widehat{\eta}_{\mathrm{d},it} = 0 \tag{3.14a}$$

$$y_{\mathrm{e},it} - \widehat{f}_{\mathrm{e},t}((1-\rho_{it})K_{it}, (1-\rho_{it})L_{it}, (1-\rho_{it})M_{it}) - \omega_{it} - \widehat{\eta}_{\mathrm{e},it} = 0, \tag{3.14b}$$

where $\widehat{f}_{\mathrm{d},t}(\cdot)$ and $\widehat{f}_{\mathrm{e},t}(\cdot)$ are the estimates of production technologies obtained using fully specialized firms in Section 3.1 to be evaluated using partial exporters' inputs; $\widehat{\eta}_{\mathrm{d},it}$ and $\widehat{\eta}_{\mathrm{e},it}$ are the transitory shocks estimated using reduced-form equations in (3.13).

4 Data

The data come from the Chinese Industrial Enterprises Database survey conducted by China's National Bureau of Statistics (NBS). This database covers all government-owned firms and all non-government-owned firms with sales above 5 million yuan (about 0.6 million in U.S. dollar). The firms in the database account for more than 90% of the gross industrial output value of the entire country. The covered industries include mining, manufacturing and public utilities. In this paper, I focus on one of China's most export-oriented manufacturing sectors, the "toy, stationery, sports and cultural goods" industry (CSIC 24), as classified by the two-digit China Standard Industrial Classification (CSIC) system for manufacturing industries. It includes manufacturing of toys, stationery goods (pens, ink, teaching tools and other related products), sports goods (balls, fitness equipment, protective tools and other related products), music instruments (Chinese and Western instruments, electronic music instruments and parts) and the game equipment including that for outdoor and indoor facilities. The sample period runs from 1999 to 2006 with the information for 1998 being used up to instrument for variables in the year 1999. The working sample includes 10,397 observations of which about 25%, 36% and 40% are non-, complete and partial exporters.

For the firm's total inputs: physical capital (K) is defined as the net fixed assets deflated by the price index of investment in fixed assets; labor (L) is measured as the total wage bill plus benefits deflated by the GDP deflator; and materials (M) are the total intermediate inputs, including raw materials and other production-related inputs, deflated by the purchasing price index for industrial inputs. The differentiated for-export (Y_{e}) and for-domestic-sale (Y_{d}) outputs are defined as the gross sales in respective markets deflated by the producer price index. The price indices are obtained from NBS and the World Bank. All variables are measured in thousands of real RMB. Following Guariglia, Liu & Song (2011) and Malikov et al. (2020), I drop observations in the top and the bottom 1st percentiles of the production function variables to rule out outliers, including the observations with missing or negative values as well as those few for which export intensity exceeds one.

Table 1. **Data Summary Statistics**

Variable	5th Perc.	Median	95th Perc.	Mean	SD
Total Output (Y)	54.12	190.52	1,134.03	339.17	403.44
Total Capital (K)	2.93	28.92	265.73	65.74	103.13
Total Labor (L)	3.89	18.69	119.38	34.27	45.27
Total Materials (M)	35.80	134.75	837.78	243.58	296.24
Export Intensity (x)	0.000	0.872	1.000	0.616	0.431
Government Equity Share (G)	0.000	0.000	0.000	0.011	0.074
Foreign Equity Share (D)	0.000	0.000	0.591	0.072	0.201
Subsidy Recipient (B)				0.150	
Foreign-Invested				0.162	
Government-Invested				0.030	
Exporter (Partial or Complete)				0.753	
Partial Exporter				0.395	
Complete Exporter				0.358	

NOTES: Y, K, L and M are measured in thousands of real RMB. x, D and G are unit-free proportions. "Exporter", "Partial Exporter", "Complete Exporter", "Government-Invested", "Foreign-Invested" and "Subsidy Recipient" are binary indicators.

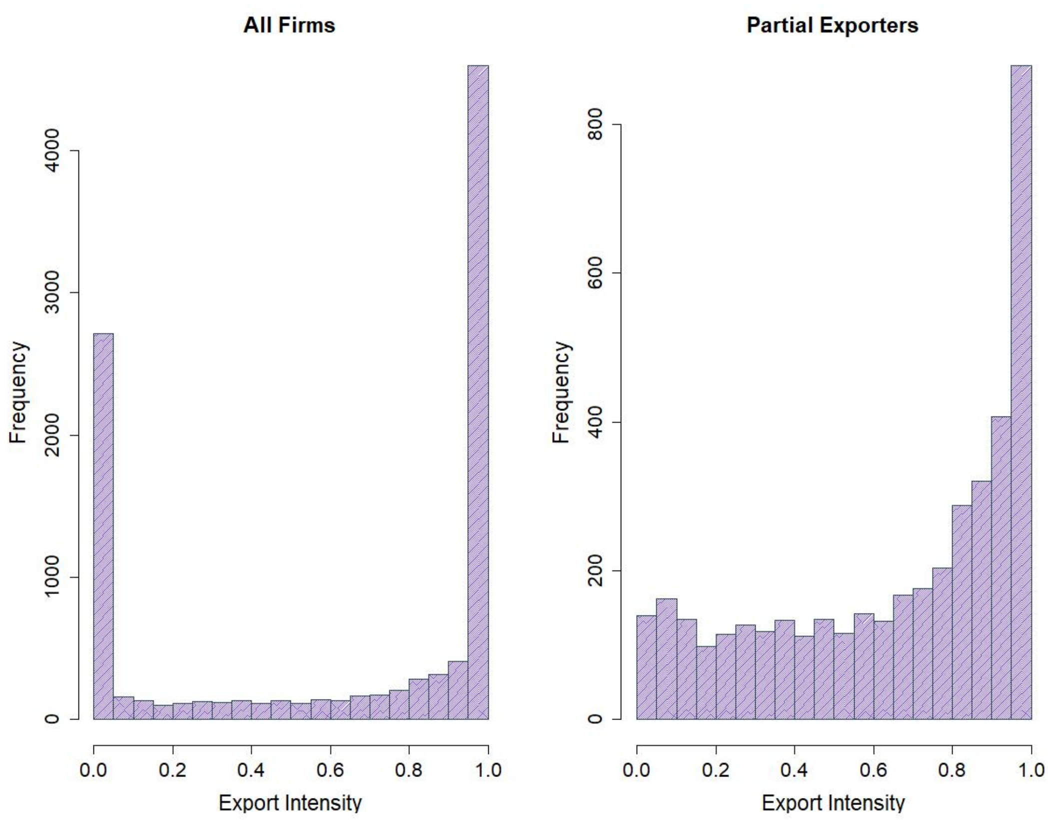

Figure 1. **Distribution of Export Intensity**

The export intensity (x) is defined as the ratio of the firm's total export value to its total output value. The variable is bounded and lies between zero and one by construction. Figure 1 plots two histograms of export intensity: *(i)* for the entire sample, which is expectedly bimodal indicating that the industry is dominated by fully specialized firms, and *(ii)* for partial exporters only. Analogously, the government and foreign equity shares (G and D, respectively) are defined as the ratios of the firm's government- and foreign-owned equity to its total equity. I also differentiate between subsidized and non-subsidized firms via a binary indicator (B) defined to take a unit value if the firm is reported to have received government subsidies and zero otherwise. Table 1 reports the data summary statistics.

5 Empirical Results

I estimate the model for each of the three types of firms—non-exporters, partial and complete exporters—following the multi-step estimation methodology outlined in Section 3. The time-variability of production functions is modeled by adding the time trend as an additional exogenous argument to function $F_{s,t}(\cdot)$ $\forall$ $s \in \{\mathrm{d}, \mathrm{e}\}$. To allow for structural changes in the market conditions ensuing from China's WTO accession, I also include a WTO dummy (equals one from 2002 onward and zero otherwise) as an additional exogenous state variable in $\mathbf{S}$ during the estimation. All unspecified functions (in second stages of the estimation procedure and beyond) are approximated using second-degree log-polynomial sieves. For the first-stage export status selection equations, I use the recently developed semiparametric distribution-free binary-choice sieve nonlinear least-squares estimator by Khan (2013).[7] In what follows, I study the firm productivity and its interplay with the firm's export behavior, which is this paper's primary focus.

I begin by examining the entire distributions of firm productivity across non-, partial and complete exporters. Figure 2 features a box-plot of empirical distributions of the estimated (log) persistent productivity ω_{it} for each of the three categories of firms. Not only do these estimates account for the export dependence of productivity (via learning by exporting) but they also do it while taking care of its endogeneity attributable to firms' selection of the degree of their export orientation (complete or partial). Comparing distributions of partial and complete exporters' productivity estimates with that of non-exporters' productivity as opposed to merely focusing on marginal moments, I am able to take a more holistic look at productivity differences between exporting and non-exporting firms. The results suggest that exporter firms (both partial and complete) are starkly more productive than non-exporters who sell their outputs exclusively in the domestic market. While this finding is largely consistent with the assumed conceptual framework whereby more productive firms self-select into exporting activities, it is important to recognize that these results are *not* a "by-construction" artifact of the model as discussed in Section 2.

[7] I use a probit criterion function with the coefficient in front of the WTO dummy normalized to unity inside the single-index mean function. The heteroskedastic standard deviation function is approximated using second-degree log-polynomial sieves.

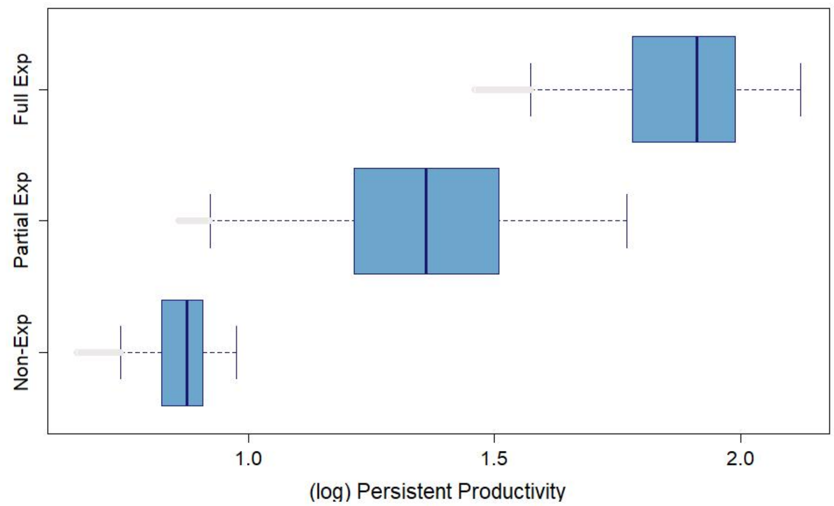

Figure 2. **Distribution of (log) Persistent Productivity**

I formally test for the existence of a systematic exporter productivity differential in the industry by testing for the (first-order) stochastic dominance of exporters' productivity over non-exporters'. I use Linton, Maasoumi & Whang's (2005) generalization of the Kolmogorov-Smirnov test which permits testing dominance over multiple variables (in my case, more than one type of exporters) and allows these variables to be estimated latent quantities as opposed to observables from the data and to also share dependence (in my case, the dependence due to endogenous self-selection). Specifically, let $G_{\mathrm{d}}(\omega)$, $G_{\mathrm{de}}(\omega)$ and $G_{\mathrm{e}}(\omega)$ respectively represent the cumulative distribution functions of (log) productivity $\omega \in \Omega$ for non-, partial and complete exporters with the subscripts representing markets they cater to. The null hypothesis that non-exporters' productivity is stochastically dominated by that of exporters (regardless whether the export orientation of the latter firms is complete or partial) is then given by

$$\mathrm{H}_0: \min_{\mathrm{s}=\{\mathrm{e},\mathrm{de}\}} \sup_{\omega\in\Omega} [G_{\mathrm{s}}(\omega) - G_{\mathrm{d}}(\omega)] \leq 0 \quad \text{vs.} \quad \mathrm{H}_1: \min_{\mathrm{s}=\{\mathrm{e},\mathrm{de}\}} \sup_{\omega\in\Omega} [G_{\mathrm{s}}(\omega) - G_{\mathrm{d}}(\omega)] > 0$$

with the corresponding test statistic being defined as

$$\mathcal{D} = \min_{\mathrm{s}=\{\mathrm{e},\mathrm{de}\}} \max_{1\leq j\leq (n_{\mathrm{s}}+n_{\mathrm{d}})} \sqrt{\frac{n_{\mathrm{s}} n_{\mathrm{d}}}{n_{\mathrm{s}}+n_{\mathrm{d}}}} \left[G_{\mathrm{s},n_{\mathrm{s}}}(\widehat{\omega}_j) - G_{\mathrm{d},n_{\mathrm{d}}}(\widehat{\omega}_j)\right],$$

where $G_{\mathrm{s},n_{\mathrm{s}}}$ is the empirical distribution function of the estimated (log) productivity $\widehat{\omega}$ for the sth category of firms of the sample size n_{s}, and $s \in \{\mathrm{d}, \mathrm{de}, \mathrm{e}\}$ designates non-, partial and complete exporters, respectively. I use the sub-sampling procedure suggested by Linton et al. (2005) to perform the test.[8] The corresponding p-value is virtually 1. Hence, consistent with the visual

[8]I use $r_N = 199$ equidistant sub-sample sizes $B_n = \{b_1, \ldots, b_{r_N}\}$, where $b_1 = [\log\log N]$ and $b_{r_N} = [N/\log\log N]$, each of which produces a p-value. Reported are the means of these r_N p-values.

Table 2. Summary of Estimated Models

Model	*Domestic/Export Production Technology*	*Export-Dependent Productivity?*	*Export-Dependent Material Demands?*	*Endog. Export Market Entry*
This Paper	Heterogeneous	Yes: $\omega_{it} = \mathbb{E}[\omega_{it}\vert\, \omega_{it-1}, \mathbf{X}_{it-1}] + \zeta_{it}$	Yes: $M_{it} = \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$	Accounted
Alt (a)	Common	Yes: $\omega_{it} = \mathbb{E}[\omega_{it}\vert\, \omega_{it-1}, \mathbf{X}_{it-1}] + \zeta_{it}$	Yes: $M_{it} = \mathbb{M}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, \omega_{it})$	Ignored
Alt (b)	Common	No: $\omega_{it} = \mathbb{E}[\omega_{it}\vert\, \omega_{it-1}] + \zeta_{it}$	No: $M_{it} = \mathbb{M}(K_{it}, L_{it}, \mathbf{S}_{it}, \omega_{it})$	Ignored

The conceptual assumptions underlying alternative model (a) are based on De Loecker (2013). However, this model can also be rationalized based on Van Biesebroeck's (2005) framework in which productivity evolution is assumed to be exogenous: $\omega_{it} = \mathbb{E}[\omega_{it}|\, \omega_{it-1}] + \zeta_{it}$. Both frameworks are however observationally equivalent for the estimation purposes and hence produce identical results.

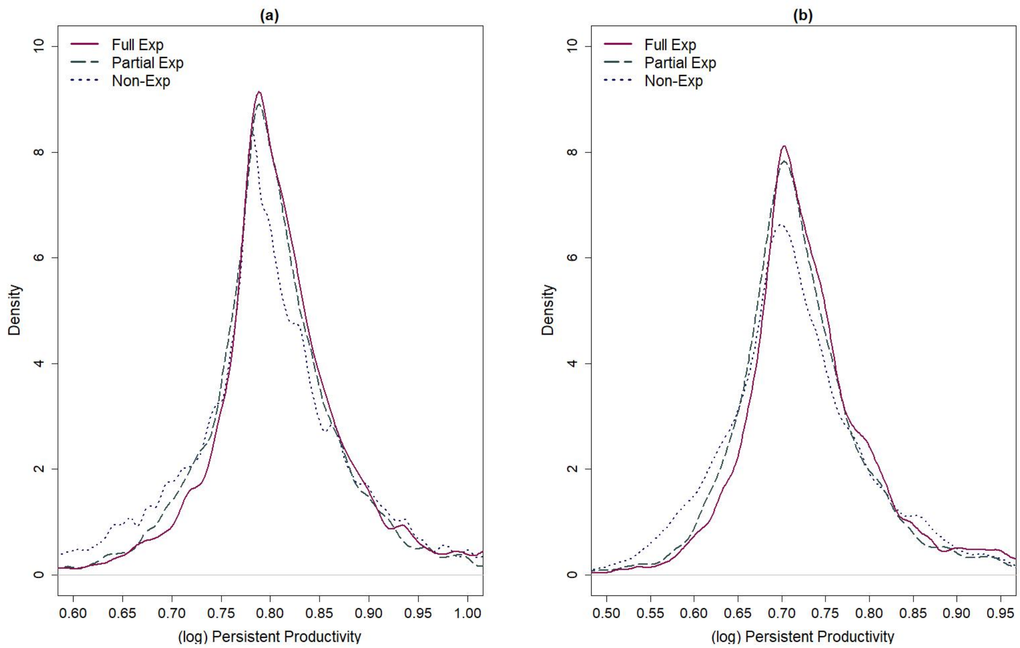

Figure 3. **Distribution of (log) Persistent Productivity using Alternative Models**

evidence from Figure 2, the data suggest a statistically significant (unconditional) productivity premium enjoyed by the exporting firms along the entire distribution.

Importantly, the conclusion would have been completely different had I not explicitly accommodated the technological heterogeneity across exporters and non-exporters subject to their endogenous self-selection during the estimation of latent firm productivity. Namely, I also estimate firm productivity using two alternative proxy-based production function estimators popularly used in the empirical literature on the export–productivity nexus: model (a) builds on Van Biesebroeck (2005) and De Loecker (2013) whose estimators, while somewhat different in the ways they incorporate information about firms' export behavior, are however observationally equivalent for the estimation purposes; model (b) is a conventional estimator à la Olley & Pakes (1996) and Levinsohn & Petrin (2003) that ignores exports information completely. Table 2 summarizes key features of these models as well as their differences from mine. Both these models assume a common production function for outputs sold domestically and abroad. Figure 3 plots kernel densities of the cross-tabulated (by

Table 3. **Productivity Ratios**

Exporter Type	75/25	90/10	95/5
Complete Exporters	1.265	1.610	1.706
Partial Exporters	1.388	1.946	2.494
Non-Exporters	1.104	1.160	1.387

Reported (by column from left ro right) are the ratios of the 75th to 25th, 90th to 10th and 95th to 5th percentiles of the empirical productivity distribution.

the firm's export status) estimates of ω_{it} obtained from these models. Evidently, due to the inability of these alternative models to accommodate potentially heterogeneous production technologies and/or endogenous self-selection into exporting activities, the estimated productivity distributions hardly differ across exporters and non-exporters and in fact, using formal tests, one misleadingly cannot reject their equality. The latter speaks in support of a richer modeling framework capable of allowing for endogenously determined cross- and within-firm technological heterogeneity based on the firms' export behavior like the one proposed in this paper. I now resume my analysis of the export–productivity relation based on the results from my model.

The differences in productivity estimates across the export status categories uncovered in Figure 2 go beyond "locational" differences. Evidently, the partial exporters are also more heterogeneous (in terms of productivity levels) than the non- or complete exporters are. Table 3 reports the within-group productivity ratios of persistent firm productivity $\exp(\omega_{it})$, which provide easy-to-interpret measures of the dispersion of corresponding productivity distributions. For instance, I find that, while a complete exporter at the 90th percentile is just 61% more productive than a complete exporter at the bottom 10th percentile, the corresponding productivity gap for partial exporters is one and a half times that at 95%. Thus, partial exporters exhibit a substantially larger degree of variation in productivity levels, which may be attributed to the fact that this category includes firms from a long continuum of the export intensity and hence a bigger potential for heterogeneity whereas both non- and complete exporters represent relatively more homogeneous groups operating in one market only.

To supplement the above analysis of the exporter productivity premium at the distribution level, Table 4 also reports the estimates of productivity differentials across exporters and domestically-oriented non-exporter firms at selected moments of the said (log) productivity distribution. Consistent with one's expectations, I find that the exporting firms in the toy industry command both statistically and economically significant productivity premium over non-exporters. Furthermore, the results point to a pronounced distributional heterogeneity in this premium whereby relatively more productive firms exhibit a premium of larger magnitudes. The latter is especially the case among partial exporters whose productivity premium estimates range from 28% at the bottom 10th quantile to 72% at the top 10th quantile. For the complete exporters, the size of premium is considerably less heterogeneous and, on average, they are heftily estimated to be twice as productivity as

Table 4. **(Unconditional) Exporter Productivity Premia, %**

Exporter Type	Mean	10th Q.	25th Q.	Median	75th Q.	90th Q.
All Exporters	73.64	40.63	51.99	73.26	99.39	105.93
	(0.40)	(0.80)	(0.70)	(0.74)	(0.43)	(0.36)
Complete Exporters Only	100.88	86.99	94.81	103.48	108.44	111.56
	(0.41)	(0.96)	(0.68)	(0.44)	(0.21)	(0.40)
Partial Exporters Only	48.98	28.36	38.29	48.52	61.29	71.88
	(0.50)	(1.03)	(0.76)	(0.50)	(0.51)	(0.67)

Estimates are relative to non-exporters. Bootstrap standard errors in parentheses.

non-exporters; the corresponding mean estimate for partial exporters is a more modest 49%. Such differences between complete and partial exporters are a novel insight which, to my knowledge, has not been previously documented in the literature. More generally, putting the distinction between partial and complete exporters aside, the estimated median exporter productivity premium is 73%.

Until now, the presented measurements of the exporter productivity premium in China's toy industry have relied on the comparison of *unconditional* moments of firm productivity. To allow for a more meaningful analysis of exporter productivity differentials, it is imperative to also account for cross-firm differences in characteristics other than the exporter status alone. The latter would enable me to study the exporter productivity premium *conditional* on relevant firm characteristics. Here, for simplicity, I focus on the binary non-exporter vs. exporter (be it partial or complete) taxonomy.

I first examine the exporter productivity differential conditional on the firm's ownership type. Specifically, I estimate the following quantile regression:

$$\begin{aligned}\mathbb{Q}_\tau\left[\omega_{it}|\cdot\right] = {} & \alpha_{\tau,0} + \alpha_{\tau,1}\mathrm{EXP}_{it} + \alpha_{\tau,2}\mathrm{GOVT}_{it} + \alpha_{\tau,3}\mathrm{FOREIGN}_{it} + \alpha_{\tau,4}\mathrm{SUBSIDY}_{it} + \\ & \alpha_{\tau,12}\mathrm{EXP}_{it}\times\mathrm{GOVT}_{it} + \alpha_{\tau,13}\mathrm{EXP}_{it}\times\mathrm{FOREIGN}_{it} + \alpha_{\tau,14}\mathrm{EXP}_{it}\times\mathrm{SUBSIDY}_{it} + \\ & \alpha_{\tau,5}\log(Y_{it}) + \alpha_{\tau,6}\log(Y_{it})^2 + \lambda_t,\end{aligned}$$

where $\mathrm{EXP}_{it} = \mathbb{1}\{x_{it} > 0\}$, $\mathrm{GOVT}_{it} = \mathbb{1}\{G_{it} > 0\}$, $\mathrm{FOREIGN}_{it} = \mathbb{1}\{D_{it} > 0\}$ and $\mathrm{SUBSIDY}_{it} = B_{it}$ are indicator variables for the "exporter", "government-invested", "foreign-invested" and "subsidy recipient" types of firms, respectively. An unsubsidized private domestic firm is the reference group. In addition to time fixed effects λ_t, I also include the log (and its square) of total output Y_{it} to control for the firm size. I opt to estimate a conditional quantile regression, as opposed to a traditional mean regression, in order to minimize distortionary impacts of outliers to which the quantile regression is fairly robust as well as to investigate the potential distributional heterogeneity in export productivity premia. The regressions are estimated for $\tau = \{0.25, 0.50, 0.75\}$ corresponding to the first, second (median) and third conditional quartiles of the (log) productivity distribution across firms. Based on the results from these regressions, I compute the (conditional) exporter productivity premia for each firm type as defined by the tabulation of the indicator variables.

Table 5. **Exporter Productivity Premia by Ownership, %**

Productivity Quant.	*Private Domestic* Non-Sub	Sub	*Government-Inv.* Non-Sub	Sub	*Foreign-Inv.* Non-Sub	Sub
25th Q.	51.82	57.36	–15.54	*4.68*	61.68	62.93
Median	70.58	68.31	*3.64*	*–6.06*	73.87	58.61
75th Q.	91.23	86.66	*7.48*	*3.42*	92.86	65.75

Reported are the conditional estimates of the exporter productivity premium at the 25th, 50th and 75th productivity quantiles, tabulated by three equity type categories — private domestic, government-invested and foreign-invested firms, which, in turn, are sub-tabulated by the subsidy indicator. Estimates are for all exporters (complete and partial) and are relative to the corresponding category of non-exporters. All estimates are statistically significant at the 5% level except those in italic.

Table 5 reports these premium estimates. Perhaps rather expectedly, I find that the government-invested exporters generally are not more productive than their non-exporting counterparts. In fact, among the least productive firms at the bottom of the distribution, the government-backed exporters may even exhibit negative productivity premium. Having said that, the empirical evidence does suggest significant productivity gains among privately owned domestic and foreign-invested exporters. Among the privately owned exporting firms, those that receive government subsidies and those that do not, largely, exhibit the same productivity premium. Interestingly however, the evidence suggests that the participation of foreign capital in the firm's equity provides a much bigger boost to the exporter premium if firms are unsubsidized by the domestic government, especially in the upper half of the productivity distribution. For instance, the median premium estimate for the unsubsidized foreign-invested exporters is 74% which drops notably to 59% if the firm is a subsidy recipient. The corresponding estimates for private domestic exporters are 71% and 68%, respectively.

I also examine if the (conditional) exporter productivity premium varies with the firm's degree of capital and export intensity. I do so by estimating the following two quantile models:

$$\begin{aligned}\mathbb{Q}_{\tau}\left[\omega_{it}|\cdot\right] = \beta_{\tau,0} + \beta_{\tau,1}\mathrm{EXP}_{it} + \beta_{\tau,2}\mathrm{GOVT}_{it} + \beta_{\tau,3}\mathrm{FOREIGN}_{it} + \beta_{\tau,4}\mathrm{SUBSIDY}_{it} + \\ \sum_{q}\left[\beta_{\tau,5q}\mathrm{D}^{K/L}_{q,it} + \beta_{\tau,6q}\mathrm{EXP}_{it}\times\mathrm{D}^{K/L}_{q,it}\right] + \beta_{\tau,7}\log(Y_{it}) + \beta_{\tau,8}\log(Y_{it})^2 + \lambda_t\end{aligned}$$

for the capital intensity (K/L), and

$$\begin{aligned}\mathbb{Q}_{\tau}\left[\omega_{it})|\cdot\right] = \gamma_{\tau,0} + \gamma_{\tau,1}\mathrm{GOVT}_{it} + \gamma_{\tau,2}\mathrm{FOREIGN}_{it} + \gamma_{\tau,3}\mathrm{SUBSIDY}_{it} + \\ \sum_{q}\gamma_{\tau,4q}\mathrm{D}^{x}_{q,it} + \gamma_{\tau,5}\mathrm{FULLEXP}_{it} + \gamma_{\tau,6}\log(Y_{it}) + \gamma_{\tau,7}\log(Y_{it})^2 + \lambda_t\end{aligned}$$

for the export intensity (x). In the above, $\mathrm{D}^{K/L}_{q,it} \equiv \mathbb{1}\{\mathbb{Q}_{0.2(q-1)}[(K/L)] < (K/L)_{it} \leq \mathbb{Q}_{0.2q}[(K/L)]\}\ \forall$ $q = 2,\ldots,5$ is the indicator variable that takes a unit value if the firm's capital intensity falls between the $(q-1)$th and qth quintiles of the empirical distribution of the capital-to-labor ratio;

Table 6. **Exporter Productivity Premia by Capital Intensity, %**

	Quintiles of Capital Intensity				
Productivity Quant.	1st	2nd	3rd	4th	5th
25th Q.	65.17	59.44	56.11	48.74	42.52
Median	87.08	77.01	68.60	57.57	50.09
75th Q.	97.27	94.66	92.54	86.03	78.88

Reported are the conditional estimates of the exporter productivity premium at the 25th, 50th and 75th productivity quantiles, tabulated by quintiles of K/L. Estimates are for all exporters (complete and partial) and are relative to the corresponding category of non-exporters. All estimates are statistically significant at the 5% level.

Table 7. **Exporter Productivity Premia by Export Intensity, %**

	Quintiles of Partial Export Intensity					
Productivity Quant.	1st	2nd	3rd	4th	5th	Complete Exporters
25th Q.	40.61	45.43	44.14	47.38	46.34	104.90
Median	40.62	43.84	45.83	46.49	46.52	98.24
75th Q.	46.83	4.782	50.90	49.60	51.93	98.53

Reported are the conditional estimates of the exporter productivity premium at the 25th, 50th and 75th productivity quantiles, tabulated by quintiles of $x|\ 0 < x < 1$ as well as the complete exporter ($x = 1$) dummy. All estimates are statistically significant at the 5% level except those in italic.

$\mathrm{D}^x_{q,it} \equiv \mathbb{1}\{\mathbb{Q}_{0.2(q-1)}[x|\ 0 < x < 1] < x_{it} \leq \mathbb{Q}_{0.2q}[x|\ 0 < x < 1]\}\ \forall\ q = 2, \dots, 5$ is the indicator variable that takes a unit value if the partial exporter firm's export intensity x falls between the $(q-1)$th and qth quintiles of the empirical distribution of the partial export intensity; and $\mathrm{FULLEXP}_{it} \equiv \mathbb{1}\{x = 1\}$ is the complete exporter dummy. In both instances, I convert the continuous (K/L) and x variables into a discrete dummy-variable representation to allow for a higher degree of nonlinearity in the relationship between the firm's productivity and its capital/export intensity.

Tables 6 and 7 report conditional exporter productivity premium estimates by the capital and export intensity quintiles measured for different quartiles of productivity distribution. The results suggest that the premium generally decreases with the firm's capital-to-labor ratio implying that the firms exporting more capital-intensive products command a smaller productivity premium over their non-exporter counterparts. This seems to suggest that, in a capital-intensive industry such as the toy manufacturing, the potential for productivity gains from exporting, to a great extent, originates from improvements in labor and its quality. This is plausible, arguably, because the new capital-centered technologies from abroad are more easily accessible for non-exporters due to their uniformity and superior codifiability than the labor-facilitated productivity improvements that are mainly based on informal learning by exporting. Thus, the more capital-intensive the production becomes, the less room there is for the productivity differential between exporters and non-exporters. In the case of export intensity, I find that the productivity premium tends to rise with the degree of firm's export orientation. The premium is larger among partial exporters at the top quintiles of the export intensity distribution, but the complete exporters command the largest

Table 8. LBE Productivity Effects of a Percent Change in Export Intensity, %

	Point Estimate	*95% Conf.* Lower	*Bounds* Upper
10th Q.	*–0.028*	–0.102	0.176
25th Q.	*0.032*	–0.033	0.344
Median	0.276	0.038	0.434
75th Q.	0.708	0.463	1.447
90th Q.	0.931	0.139	0.970
Mean	0.369	0.138	0.343

Reported are the moments of empirical distribution of $\partial\omega_{it}/\partial x_{it-1}$ point estimates along with 95% percentile bootstrap confidence intervals. All point estimates are statistically significant at the 5% level except those in italic.

productivity differential over non-exporters.

Lastly, I consider the export-dependent evolution of the firm's persistent productivity (in logs). In line with (2.2), I estimate the following nonparametric regression:[9]

$$\omega_{it} = h(\omega_{it-1}, x_{it-1}) + \zeta_{it},$$

where $h(\cdot)$ is an unspecified conditional mean function. The estimation of this controlled Markovian process for ω_{it} presents a natural way to measure learning-by-exporting effects on the evolution of firm productivity over time via $\partial\mathbb{E}[\omega_{it}|\cdot]/\partial x_{it-1} = \partial h(\omega_{it-1}, x_{it-1})/\partial x_{it-1}$. The (firm-specific) gradient estimates along with their percentile bootstrap confidence intervals are summarized in Table 8. Figure 4 plots their empirical distribution. The results yield a positively skewed distribution of the estimated learning-by-exporting effects. While some firms exhibit insignificant near-zero effects, for the majority of firms the learning-by-exporting effect is both statistically and economically significant. The average effect size is estimated at statistically significant 0.37% generally providing the evidence in favor of productivity-enhancing effects of the export engagement. Having said that, the fact that the learning-by-exporting estimates are not uniformly (statistically) positive indirectly suggests that the large productivity premium enjoyed by exporters in China's toy industry may be attributed, to a greater extent, to their *a priori* higher productivity that enticed their self-selection into exporting as opposed to the *ex post* export-driven productivity improvements.

[9] I use local-linear kernel fitting with least-squares cross-validated bandwidths. Given the bimodal distribution of export intensity, to avoid over-smoothing in dense ranges of the support of the data while under-smoothing in sparse ranges, I employ an adaptive k-nearest-neighbor bandwidth capable of adapting to the local distribution of the data.

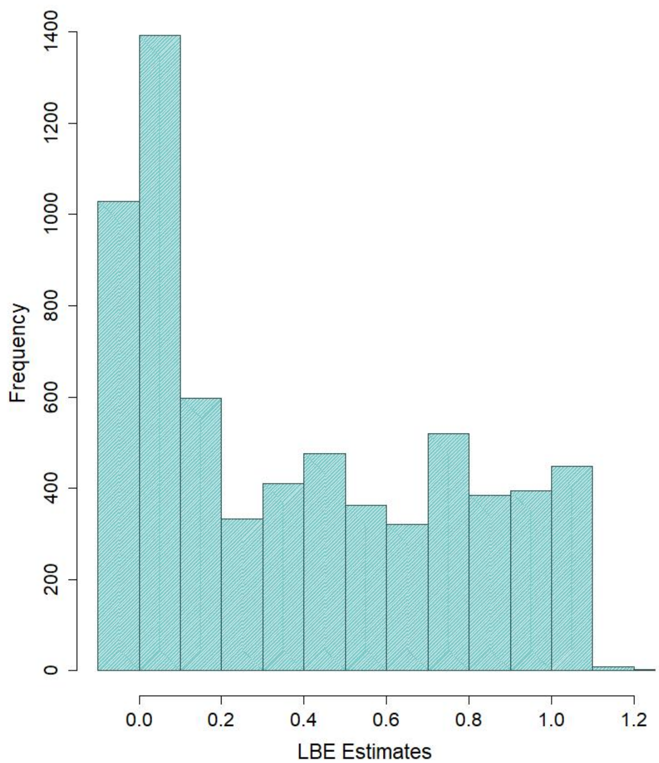

Figure 4. **Distribution of LBE Estimates**

6 Conclusion

Exporters are generally more productive than non-exporters, and uncovering the extent of these productivity differentials as well as causes thereof has long been one of the core topics in the literature. This paper focuses on the issues pertaining to robust empirical modeling of the intricate endogenous relationship between the firm's export behavior and its productivity. I provide a new synthesized methodology for the structural control-function-based identification of unobserved firm productivity in the presence of *(i)* endogenous self-selection into export markets, *(ii)* dependence of productivity evolution on the firm's past export experiences and *(iii)* potential heterogeneity in the production of for-export and for-domestic-sale outputs. The model offers a useful addition to the toolkit of practitioners particularly interested in studying the nexus between firm productivity and export behavior.

I showcase the proposed methodology by applying it to the firm-level data on China's toy industry, one of the country's most export-oriented manufacturing sectors, during the 1999–2006 period. I find overwhelming evidence in support of statistically and economically significant exporter productivity premium both at fixed marginal moments as well as across the entire distribution of firms. However, when employing less robust conventional estimators popularly used in the literature, I find that the estimated productivity hardly differs across exporters and non-exporters and one misleadingly cannot reject their equality. The latter exemplifies the benefits of a richer modeling framework capable of allowing for endogenously determined cross- and within-firm heterogeneity dependent on the firms' export behavior.

Appendix

A Derivation of (3.4)–(3.5)

The expectation of (3.3) conditional on the information available to a fully specialized firm i at time t and its decision to continue its current export behavior is given by

$$\begin{aligned}
\mathbb{E}[y_{it}|\,\mathcal{I}_{it-1},\Xi_{it}=1] &= \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\,\mathcal{I}_{it-1}] + \varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) + \mathbb{E}[\zeta_{it}|\,\mathcal{I}_{it-1},\Xi_{it}=1] \\
&= \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\,\mathcal{I}_{it-1}] + \varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) + \mathbb{E}[\zeta_{it}|\,\Xi_{it}=1] \\
&= \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\,\mathcal{I}_{it-1}] + \varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) \\
&\quad + \begin{cases} \mathbb{E}[\zeta_{it}|\,\zeta_{it}\leq \underline{\omega}_{it}-\varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})] & \text{for non-exporter} \\ \mathbb{E}[\zeta_{it}|\,\zeta_{it}\geq \overline{\omega}_{it}-\varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})] & \text{for complete exporter} \end{cases} \\
&\equiv \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\,\mathcal{I}_{it-1}] + \varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) \\
&\quad + \begin{cases} \lambda(\underline{\omega}_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for non-exporter} \\ \lambda(\overline{\omega}_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for complete exporter} \end{cases} \\
&\equiv \mathbb{E}[f_{s,t}(K_{it},L_{it},M_{it})|\,\mathcal{I}_{it-1}] + \begin{cases} \psi(\underline{\omega}_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for non-exporter} \\ \psi(\overline{\omega}_{it},K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for complete exporter} \end{cases}
\end{aligned} \tag{A.1}$$

where I use the Markovian process for ω_{it} inside the expectation in the second equality, let the expectation term (a selection bias) be an unknown function $\lambda(\cdot)$ in the third equality and, since $\varphi(\cdot)$ and $\lambda(\cdot)$ are not separably identified, define their sum $\psi(\cdot)\equiv\varphi(\cdot)+\lambda(\cdot)$ in the last equality.

The unobservable productivity thresholds $\underline{\omega}_{it}$ and $\overline{\omega}_{it}$ appearing inside $\psi(\cdot)$ can be proxied by inverting the firm's export status self-selection probability (propensity score) given by

$$\begin{aligned}
\Pr[\Xi_{it}=1|\,\mathcal{I}_{it}] &= \begin{cases} \Pr[\omega_{it}\leq \underline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it})|\,\mathcal{I}_{it}] & \text{for non-exporter} \\ \Pr[\omega_{it}\geq \overline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it})|\,\mathcal{I}_{it}] & \text{for complete exporter} \end{cases} \\
&= \begin{cases} \Pr[\zeta_{it}\leq \underline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it})-\varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})|\,\mathcal{I}_{it}] & \text{for non-exporter} \\ \Pr[\zeta_{it}\geq \overline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it})-\varphi(K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1})|\,\mathcal{I}_{it}] & \text{for complete exporter} \end{cases} \\
&\equiv \begin{cases} \mathcal{P}(\underline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it}),K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for non-exporter} \\ \mathcal{P}(\overline{\omega}_{it}(K_{it},L_{it},\mathbf{X}_{it},\mathbf{S}_{it}),K_{it-1},L_{it-1},\mathbf{X}_{it-1},\mathbf{S}_{it-1},M_{it-1}) & \text{for complete exporter} \end{cases}
\end{aligned}$$

$$= \mathcal{P}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) \equiv \varrho_{it}, \tag{A.2}$$

where, in the fourth equality, I recognize that dynamic state variables $(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it})$ are implicit functions of the already included lags. For instance, from $K_{it} = (1-\delta)K_{it-1} + I_{it-1}$ and $I_{it-1} = \mathbb{I}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, \omega_{it-1}) = \mathbb{I}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, \mathbb{M}^{-1}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}))$, it is evident that $K_{it}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$. Similar logic applies to other state variables.

B Estimation of reduced-form production functions

Consider the identification of transitory productivity shocks $\eta_{\mathrm{d},it}$ and $\eta_{\mathrm{e},it}$ from the partial exporter's "reduced-form" intended-market-specific production functions (in logs):

$$y_{s,it} = f_{s,t}(K_{s,it}, L_{s,it}, M_{s,it}) + \omega_{it} + \eta_{s,it} \equiv g_{s,t}(K_{it}, L_{it}, M_{it}) + \omega_{it} + \eta_{s,it} \qquad \forall\ s \in \{\mathrm{d}, \mathrm{e}\}, \tag{3.13}$$

where the reduced-form function $g_{s,t}(\cdot)$ differs from its structural counterpart $f_{s,t}(\cdot)$ in that it is modeled as a function of observable *total* inputs. I estimate $\eta_{\mathrm{d},it}$ and $\eta_{\mathrm{e},it}$ via a three-step methodology applied *twice* for each partial exporter firm using different output $y_{s,it}$ on the left-hand side.

Specifically, logging both sides of (3.13) and using the Markovian nature of ω_{it} from (2.2) yields

$$y_{s,it} = g_{s,t}(K_{it}, L_{it}, M_{it}) + h_s\left(\omega_{it-1}, \mathbf{X}_{it-1}\right) + \zeta_{s,it} + \eta_{s,it}, \tag{B.1}$$

where I deliberately index the proxy function $h_s(\cdot)$ and the innovation in persistent productivity $\zeta_{s,it}$ with subscript s to emphasize that the reduced-form estimation procedure cannot identify product-invariant firm-specific ω_{it}. Like in Section 3.1, I proxy for unobservable productivity using total materials:

$$y_{s,it} = g_{s,t}(K_{it}, L_{it}, M_{it}) + \varphi_s\left(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}\right) + \zeta_{s,it} + \eta_{s,it}, \tag{B.2}$$

where $\varphi_s(\cdot) \equiv h_s\left(\mathbb{M}^{-1}(\cdot), \cdot\right)$.

The expectation of (B.2) conditional on $\mathcal{I}_{it-1}$ and the partial-exporter firm's decision to continue its current export behavior ($\Xi_{it} = 1$) is given by

$$\begin{aligned}
\mathbb{E}[y_{s,it}|\ \mathcal{I}_{it-1}, \Xi_{it} = 1] &= \mathbb{E}[g_{s,t}(K_{it}, L_{it}, M_{it})|\ \mathcal{I}_{it-1}] + \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \mathbb{E}[\zeta_{s,it}|\ \Xi_{it} = 1] \\
&= \mathbb{E}[g_{s,t}(K_{it}, L_{it}, M_{it})|\ \mathcal{I}_{it-1}] + \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \\
&\quad \mathbb{E}[\zeta_{s,it}|\ \underline{\omega}_{it} - \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})] < \zeta_{s,it} < \overline{\omega}_{it} - \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})] \\
&\equiv \mathbb{E}[g_{s,t}(K_{it}, L_{it}, M_{it})|\ \mathcal{I}_{it-1}] + \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \\
&\quad \lambda_s(\underline{\omega}_{it} - \overline{\omega}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})
\end{aligned}$$

$$\equiv \mathbb{E}[g_{s,t}(K_{it}, L_{it}, M_{it})|\ \mathcal{I}_{it-1}] + \psi_s(\underline{\omega}_{it} - \overline{\omega}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), \tag{B.3}$$

where I use the Markovian process for ω_{it} inside the expectation in the second equality, let the expectation term (a selection bias) be an unknown function $\lambda_s(\cdot)$ in the third equality and, since $\varphi_s(\cdot)$ and $\lambda_s(\cdot)$ are not separably identified, define their sum $\psi_s(\cdot) \equiv \varphi_s(\cdot)+\lambda_s(\cdot)$ in the last equality. The unobservable productivity thresholds $\underline{\omega}_{it}$ and $\overline{\omega}_{it}$ are not separably identified either thereby rendering $\lambda_s(\cdot)$ and, hence, $\psi_s(\cdot)$ a function of their difference.[10] The difference in thresholds can be proxied using the firm's export status self-selection propensity score given by:

$$\begin{aligned}
\Pr[\Xi_{it} = 1|\ \mathcal{I}_{it}] &= \Pr[\underline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}) < \omega_{it} < \overline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it})|\ \mathcal{I}_{it}] \\
&= \Pr[\underline{\omega}_{it} - \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})] < \zeta_{s,it} < \overline{\omega}_{it} - \varphi_s(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) \\
&\equiv \mathcal{P}(\underline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}) - \overline{\omega}_{it}(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}), K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) \\
&= \mathcal{P}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) \equiv \varrho_{it}.
\end{aligned} \tag{B.4}$$

Then, inverting for an unobserved scalar $(\underline{\omega}_{it}-\overline{\omega}_{it})$ via $\underline{\omega}_{it}-\overline{\omega}_{it} = \mathcal{P}^{-1}(\varrho_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1})$ from the third equality in (B.4), from (B.3) I get

$$\mathbb{E}[y_{s,it}|\ \mathcal{I}_{it-1}, \Xi_{it} = 1] = \mathbb{E}[g_{s,t}(K_{it}, L_{it}, M_{it})|\ \mathcal{I}_{it-1}] + \psi_s(\varrho_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), \tag{B.5}$$

with the corresponding selection-bias-corrected analogue of (B.2) given by

$$y_{s,it} = g_{s,t}(K_{it}, L_{it}, M_{it}) + \psi_s(\varrho_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}) + \zeta^*_{s,it} + \eta_{s,it}, \tag{B.6}$$

where the new innovation $\zeta^*_{s,it} \equiv \zeta_{s,it} - \lambda_s(\cdot)$ is such that $\mathbb{E}[\zeta^*_{s,it}|\ \mathcal{I}_{it-1}, \Xi_{it} = 1] = 0$ by construction. Thus, the reduced-form production function $g_{s,t}(\cdot)$ can then be identified (up to a constant) based on

$$\mathbb{E}\left[\zeta^*_{s,it} + \eta_{s,it} \left|\ \begin{matrix} \varrho_{it}(K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}), \\ K_{it}, L_{it}, \mathbf{S}_{it}, K_{it-1}, L_{it-1}, \mathbf{X}_{it-1}, \mathbf{S}_{it-1}, M_{it-1}, \end{matrix} \quad \Xi_{it} = 1 \right.\right] = 0, \tag{B.7}$$

where $\mathbf{S}_{it}$ instruments for M_{it}.

Then, analogously to (3.11), the transitory productivity shock $\eta_{s,it}$ is then identified (up to a constant) from

$$y^*_{s,it} \equiv y_{s,it} - \widehat{g}_{s,t}(K_{it}, L_{it}, M_{it}) = \nu_s(K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it}) + \eta_{s,it} \tag{B.8}$$

based on $\mathbb{E}[\eta_{s,it}|\ K_{it}, L_{it}, \mathbf{X}_{it}, \mathbf{S}_{it}, M_{it}] = 0$.

To recap, the estimation methodology is to be implemented for each partial exporter twice (using exported vs. domestically sold outputs) in three steps: *(i)* estimate export status propensity

[10] In fact, this happens to be a good news since invertability of propensity score function requires *scalar* unobservability condition.

in (B.4); *(ii)* estimate reduced-form production functions in (3.13) via method of moments based on (B.7); *(iii)* estimate transitory shocks from (B.8) via least squares.

References

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

Amiti, M. & Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, *97*, 1611–1638.

Aw, B.-Y., Chung, S., & Roberts, M. J. (2000). Productivity and turnover in the export market: Micro-level evidence from the Republic of Korea and Taiwan (China). *World Bank Economic Review*, *14*, 65–90.

Aw, B.-Y. & Hwang, A. R. (1995). Productivity and the export market: A firm-level analysis. *Journal of Development Economics*, *47*, 313–332.

Baldwin, J. R. & Gu, W. (2003). Export-market participation and productivity performance in Canadian manufacturing. *Canadian Journal of Economics*, *36*, 634–657.

Bernard, A. B. & Jensen, J. B. (1995). Exporters, jobs, and wages in U.S. manufacturing: 1976–1987. *Brookings Papers on Economic Activity: Microeconomics*, 67–119.

Chen, X., Hahn, J., Liao, Z., & Ridder, G. (2015). Asymptotic properties of nonparametric two-step sieve estimates. Working Paper, University of California Los Angeles.

Clerides, S. K., Lach, S., & Tybout, J. R. (1998). Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *Quarterly Journal of Economics*, *113*, 903–947.

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

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

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., & Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, *84*, 445–510.

Delgado, M. A., Farinas, J. C., & Ruano, S. (2002). Firm productivity and export markets: A non-parametric approach. *Journal of International Economics*, *57*, 397–422.

Doraszelski, U. & Jaumandreu, J. (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies*, *80*, 1338–1383.

Gandhi, A., Navarro, S., & Rivers, D. (2017). On the identification of gross output production functions. *Journal of Political Economy*. forthcoming.

Girma, S., Greenaway, D., & Kneller, R. (2004). Does exporting increase productivity? A microeconometric analysis of matched firms. *Review of International Economics*, *12*, 855–866.

Guariglia, A., Liu, X., & Song, L. (2011). Internal finance and growth: Microeconometric evidence on Chinese firms. *Journal of Development Economics*, *96*, 79–94.

HKTDC, H. K. T. D. C. (2019). China's toy market. Market Report, The HKTDC Research: China Trade.

Khan, S. (2013). Distribution free estimation of heteroskedastic binary response models using Probit/Logit criterion functions. *Journal of Econometrics*, *172*, 168–182.

Levinsohn, J. & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, *70*, 317–341.

Linton, O., Maasoumi, E., & Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *Review of Economic Studies*, *72*, 735–765.

Malikov, E., Zhao, S., & Kumbhakar, S. C. (2020). Estimation of firm-level productivity in the presence of exports: Evidence from China's manufacturing. *Journal of Applied Econometrics*. forthcoming.

Mayer, T., Melitz, M. J., & Ottaviano, G. I. P. (2014). Market size, competition, and the product mix of exporters. *American Economic Review*, *104*, 495–536.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*, 1695–1725.

Melitz, M. J. & Ottaviano, G. I. P. (2008). Market size, trade, and productivity. *Review of Economic Studies*, *75*, 295–316.

Newey, W. K. (1984). A method of moments interpretation of sequential estimators. *Economics Letters*, *14*(2), 201–206.

Olley, G. S. & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*, 1263–1297.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *Review of Economic Studies*, *69*, 245–276.

Van Biesebroeck, J. (2005). Exporting raises productivity in sub-Saharan African manufacturing firms. *Journal of International Economics*, *67*, 373–391.

Wagner, J. (2007). Exports and productivity. *World Economy*, *30*, 60–82.

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