Supplementary Web Appendix to

Estimation of Firm-Level Productivity in the Presence of Exports: Evidence from China's Manufacturing*

Emir Malikov[1] Shunan Zhao[2] Subal C. Kumbhakar[3]

[1]University of Nevada, Las Vegas
[2]Oakland University
[3]Binghamton University

January 8, 2020

Abstract

Motivated by the longstanding interest of economists in understanding the nexus between firm productivity and export behavior, this paper develops a novel structural framework for control-function-based nonparametric identification of the gross production function and latent firm productivity in the presence of endogenous export opportunities that is robust to recent unidentification critiques of proxy estimators. We provide a workable identification strategy, whereby the firm's degree of export orientation provides the needed (excluded) relevant independent exogenous variation in endogenous freely varying inputs, thus allowing us to identify the production function. We estimate our fully nonparametric IV model using the Landweber-Fridman regularization with the unknown functions approximated via artificial neural network sieves with a sigmoid activation function which are known for their superior performance relative to other popular sieve approximators, including the polynomial series favored in the literature. Using our methodology, we obtain robust productivity estimates for manufacturing firms from twenty eight industries in China during the 1999–2006 period to take a close look at China's exporter productivity puzzle, whereby exporters are found to exhibit lower productivity levels than non-exports.

Keywords: ANN sieves, control function, export, nonparametric, productivity, proxy, regularized estimation, TFP

JEL Classification: D24, F10, L10

* *Email*: emir.malikov@unlv.edu (Malikov), shunanzhao@oakland.edu (Zhao), kkar@binghamton.edu (Kumbhakar).

We would like to thank the co-editor, Edward Vytlacil, and three anonymous referees for many insightful comments and suggestions that greatly helped improve this paper. We are also thankful to Yiguo Sun, Nic Ziebarth, seminar participants at Auburn and Oakland as well as conference participants at the 2016 North American Productivity Workshop in Quebec City for helpful comments. Any remaining errors are our own.

Appendix

A Under-identification Issues

Gandhi et al. (2017) show that the traditional control-function-based production function estimators, which utilize freely varying inputs (such as materials M_{it}) to proxy for unobserved latent productivity, suffer from an inherent under-identification. As shown in Section 3, our estimator is immune to this problem owing to our ability to employ the firm's export intensity X_{it} as an additional source of *relevant independent* exogenous variation from outside the production function to help us identify the latter. In what follows, we show why our model would have been under-identified *if* we had adopted the standard conceptual framework by ignoring the firm's export decisions altogether. In that case, *(i)* the persistent productivity would evolve according to the exogenous Markov process with no learning by exporting effects whereby $\omega_{it} = \mathbb{E}[\omega_{it}|\ \omega_{it-1}] + \zeta_{it}$, and *(ii)* the conditional material demand function would be given by $M_{it} = \mathbb{M}_t(K_{it}, L_{it}, \omega_{it})$. The analogue of the production function in (3.3) is then:

$$y_{it} = f(K_{it}, L_{it}, M_{it}) + \varphi(K_{it-1}, L_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it}. \tag{A.1}$$

Next, note that the endogenous M_{it} entering the production function $f(\cdot)$:

$$M_{it} = \mathbb{M}_t(K_{it}, L_{it}, \omega_{it}) = \mathbb{M}_t(K_{it}, L_{it}, \varphi(K_{it-1}, L_{it-1}, M_{it-1}) + \zeta_{it}), \tag{A.2}$$

is a function of the following observables $(K_{it}, L_{it}, K_{it-1}, L_{it-1}, M_{it-1})$ and the unobservable ζ_{it}. Comparing these arguments of M_{it} with the variables entering $f(\cdot)$ directly as well as appearing inside the proxy for productivity $\varphi(\cdot)$, it is evident that the only extra source of variation for M_{it}, which has not already been included on the right-hand side of (A.1), is the unobservable ζ_{it}. In other words, conditional on the already included self-instrumenting variables $(K_{it}, L_{it}, K_{it-1}, L_{it-1}, M_{it-1})$, there is *no* other relevant exogenous variable that may be used to instrument for the endogenous M_{it} because, for it to be relevant in predicting M_{it}, it would have to correlate with ζ_{it}, which is the only source of "free" variation left in M_{it}. The correlation with ζ_{it} would however violate the exogeneity requirement thereby invaliding the instrument. Notably, this conundrum also applies to (weakly exogenous) excluded lags (of the second and higher order) of inputs which, following the bulk of the literature, one may be tempted to utilize in hopes of identifying the model. As evident from (A.2), any such additional (excluded) lag is *ir*relevant for predicting M_{it} once conditioned on the already included $(K_{it}, L_{it}, K_{it-1}, L_{it-1}, M_{it-1})$. Therefore, M_{it} still lacks an excluded relevant instrument from outside of the equation, and the production function $f(\cdot)$ in (A.1) is unidentified under the *standard* structural assumptions.

B Separable Identifiability Issues

Ackerberg et al.'s (2015) critique of control-function-based production function estimators effectively boils down to one's potential inability to *separably* identify the production and productivity proxy functions due to the presence of endogenous freely varying inputs inside the production function. This problem occurs primarily when *(i)* labor is assumed to be a freely varying input (like M_{it}) thereby becoming endogenous and *(ii)* one attempts to "control" for contemporaneous ω_{it} as opposed to lagged ω_{it-1}, as originally done by Olley & Pakes (1996) and Levinsohn & Petrin (2003). If we were to go this route, *(i)* the conditional material demand function would no longer be a function of L_{it}, i.e., $M_{it} = \mathbb{M}_t(K_{it}, X_{it}, \omega_{it})$, and *(ii)* unobserved ω_{it} would be proxied by the contemporaneous inverted demand for materials without making use of its Markovian nature. In this case, the analogue of the production function in (3.3) is given by

$$\begin{aligned} y_{it} &= f(K_{it}, L_{it}, M_{it}) + \omega_{it} + \eta_{it} \\ &= f(K_{it}, L_{it}, M_{it}) + \mathbb{M}_t^{-1}(K_{it}, X_{it}, M_{it}) + \eta_{it} \\ &= f(K_{it}, \mathbb{L}_t(K_{it}, X_{it}, \omega_{it}), M_{it}) + \mathbb{M}_t^{-1}(K_{it}, X_{it}, M_{it}) + \eta_{it} \\ &= f\left(K_{it}, \mathbb{L}_t\left(K_{it}, X_{it}, \mathbb{M}_t^{-1}(K_{it}, X_{it}, M_{it})\right), M_{it}\right) + \mathbb{M}_t^{-1}(K_{it}, X_{it}, M_{it}) + \eta_{it} \\ &= g_t(K_{it}, X_{it}, M_{it}) + \mathbb{M}_t^{-1}(K_{it}, X_{it}, M_{it}) + \eta_{it}, \end{aligned} \tag{B.1}$$

where we have made use of the firm's conditional demand for now freely varying labor $L_{it} = \mathbb{L}_t(K_{it}, X_{it}, \omega_{it})$ in the third line. From (B.1), it is obvious that the two unknown functions $g_t(\cdot)$ and $\mathbb{M}_t^{-1}(\cdot)$ are *not* separably identified since they depend on the same covariates.

However, this separable unidentifiability problem no longer applies if, instead of ω_{it}, one were to proxy for ω_{it-1} (like we do in our paper). Specifically, from

$$\begin{aligned} y_{it} &= f(K_{it}, L_{it}, M_{it}) + h_t[\omega_{it-1}, X_{it-1}] + \zeta_{it} + \eta_{it} \\ &= f(K_{it}, L_{it}, M_{it}) + \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it} \\ &= f(K_{it}, \mathbb{L}_t(K_{it}, X_{it}, \omega_{it}), \mathbb{M}_t(K_{it}, X_{it}, \omega_{it})) + \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it} \\ &= f\left(K_{it}, \mathbb{L}_t\left(K_{it}, X_{it}, \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it}\right), \mathbb{M}_t\left(K_{it}, X_{it}, \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it}\right)\right) + \\ &\quad \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it} \end{aligned} \tag{B.2}$$

it is evident that, conditional on exogenous covariates $(K_{it}, X_{it}, K_{it-1}, X_{it-1}, M_{it-1})$, the production function $f(\cdot)$ is separably identified from the proxy function $\varphi(\cdot)$ because both endogenous inputs L_{it} and M_{it} still vary independently from all arguments of $\varphi(\cdot)$ owing to the presence of ζ_{it}.

Based on the above results, it might appear that, in our model, we too could have allowed labor to be freely varying. Unfortunately, while our model would still have been immune to Ackerberg et al.'s (2015) critique in this case, we would however have faced the under-identification problem à la Gandhi et al. (2017) not only with respect to M_{it} but also L_{it}. Since we only have *one* valid excluded instrument (X_{it}), we generally would have been unable to identify our model in the presence of *two* endogenous freely varying inputs due to the failure to meet the order condition. We are therefore bound to the assumption of quasi-fixity of labor.

C Bootstrap Algorithm

We approximate sampling distributions of the estimators via wild residual block bootstrap that takes into account a panel structure of the data, with all the steps bootstrapped jointly owing to a sequential nature of our estimation procedure. To make matters more concrete, we use the following wild residual panel bootstrap algorithm:

1. Compute the three steps of our estimation procedure using the original data. Denote the obtained estimates as $\widehat{\psi}_{it} = \widehat{\psi}(K_{it}, L_{it}, M_{it}, K_{it-1}, L_{it-1}, X_{it-1}, M_{it-1})$ in the first step, $\widehat{f}_{it} = \widehat{f}(K_{it}, L_{it}, M_{it})$ and $\widehat{\varphi}_{it} = \widehat{\varphi}(K_{it-1}, L_{it-1}, X_{it-1}, M_{it-1})$ in the second step, and $\widehat{\omega}_{it} = \widehat{g}(K_{it}, L_{it}, X_{it}, M_{it})$ in the third step, for all $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Let the residuals from each of these three sequential steps be respectively denoted by $\{\widehat{u}_{it} = y_{it} - \widehat{\psi}_{it}\}$, $\{\widehat{e}_{it} = y_{it} - \widehat{f}_{it} - \widehat{\varphi}_{it} - \widehat{u}_{it}\}$ and $\{\widehat{\eta}_{it} = y_{it} - \widehat{f}_{it} - \widehat{\omega}_{it}\}$. Recenter the residuals.

2. Generate bootstrap weights ξ_i^b for all cross-sectional units $i = 1, \ldots, n$ from the Mammen (1993) two-point mass distribution:

$$\xi_i^b = \begin{cases} \frac{1+\sqrt{5}}{2} & \text{with prob. } \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \frac{1-\sqrt{5}}{2} & \text{with prob. } \frac{\sqrt{5}+1}{2\sqrt{5}}. \end{cases} \tag{C.1}$$

   Next, for each observation (i, t) with $i = 1, \ldots, n$ and $t = 1, \ldots, T$, generate a new bootstrap first-step disturbance: $u_{it}^b = \xi_i^b \widehat{u}_{it}$.

3. Generate a new bootstrap outcome variable. From the first step, we have $y_{it}^b = \widehat{\psi}_{it} + u_{it}^b$ for all $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Recompute the first-step estimator using $\{y_{it}^b\}$ in place of $\{y_{it}\}$ and denote the obtained estimates as $\{\widehat{\psi}_{it}^b\}$ with the corresponding residuals $\{\widehat{u}_{it}^b = y_{it} - \widehat{\psi}_{it}^b\}$. Use these residuals in the next step of the algorithm.

4. Recompute the second-step estimator using $\{y_{it}^b\}$ in place of $\{y_{it}\}$ and $\{\widehat{u}_{it}^b\}$ in place of $\{\widehat{u}_{it}\}$. Denote the obtained estimates as $\{\widehat{f}_{it}^b\}$ and $\{\widehat{\varphi}_{it}^b\}$. Use these estimates in the next step of the algorithm.

5. Recompute the third-step estimator using $\{y_{it}^b\}$ in place of $\{y_{it}\}$ and $\{\widehat{f}_{it}^b\}$ in place of $\{\widehat{f}_{it}\}$ and denote the obtained estimates as $\{\widehat{\omega}_{it}^b\}$.

6. Repeat steps 2 through 5 of the algorithm B times. Use the empirical distribution of B bootstrap estimates of f_{it} and ω_{it} as well as the functionals thereof to construct accelerated bias-corrected percentile confidence intervals as described in Section 4.1.

D Monte-Carlo Experiments

In this section, we describe Monte Carlo experiments that evaluate the performance of our proposed estimator and demonstrate its ability, owing to its use of the excluded exogenous variation in the firm's degree of export orientation, to successfully identify the production function. We then compare its performance to that of the traditional proxy estimator à la Levinsohn & Petrin (2003) based on the lagged instrumentation of endogenous materials. In addition, we also investigate the finite-sample performance of our proposed procedure for bootstrap inference.

Our data generating process (DGP) is essentially a fusion of those used by Grieco et al. (2016) and Gandhi et al. (2017). More specifically, we consider a balanced panel of $n = \{50, 100, 200, 400, 800\}$ firms operating during 10 time periods.[20] Each panel is simulated 1,000 times. To simplify matters, we dispense with labor and consider the production process with two inputs only: a quasi-fixed capital and freely varying materials. We also let that the true technology take a simple constant-input-elasticity Cobb-Douglas form:

$$Y_{it} = K_{it}^{\alpha_1} M_{it}^{\alpha_2} \exp\{\omega_{it} + \eta_{it}\}, \tag{D.1}$$

where we assume decreasing returns to scale and set $\alpha_1 = 0.25$ and $\alpha_2 = 0.65$.

The productivity components are generated as follows. We model the persistent productivity as a controlled AR(1) process:

$$\omega_{it} = \rho_0 + \rho_1 \omega_{it-1} + \rho_2 X_{it-1} + \zeta_{it}, \tag{D.2}$$

where we set $\rho_0 = 0.2$, $\rho_1 = \rho_2 = 0.8$, and $\zeta_{it} \sim$ i.i.d. $\mathbb{N}(0, \sigma_\zeta^2)$ with $\sigma_\zeta = 0.04$. The initial level of productivity ω_{i1} is drawn from $\mathbb{U}(1, 3)$ identically and independently distributed over i. The random transitory shocks η_{it} are drawn from $\eta_{it} \sim$ i.i.d. $\mathbb{N}(0, \sigma_\eta^2)$ with $\sigma_\eta = 0.07$.

The exogenous export-associated cost shifter $\mathcal{V}_{it}$ is assumed to exhibit first-order Markov persistence, i.e.,

$$\mathcal{V}_{it} = \varrho_0 + \varrho_1 \mathcal{V}_{it-1} + \nu_{it}, \tag{D.3}$$

where $\varrho_0 = 0.2$, $\varrho_1 = 0.5$, and $\nu_{it} \sim$ i.i.d. $\mathbb{N}(0, \sigma_\nu^2)$ with $\sigma_\nu = 0.5$. Note that the time-persistence in $\mathcal{V}_{it}$ is unimportant, and making it be a white noise would have sufficed.

We assume the following about evolution of the firm's state variables. Capital and the export intensity, both quasi-fixed, are set to respectively evolve according to

$$K_{it} = I_{it-1} + (1 - \delta_i) K_{it-1} \tag{D.4}$$

$$X_{it} = \mathcal{X}_{it-1} + X_{it-1}, \tag{D.5}$$

[20] We have also experimented with 5 and 50 time periods. The results are qualitatively unchanged.

with the corresponding decision rules taking the following forms:[21]

$$I_{it-1} = K_{it-1}^{\beta_1} X_{it-1}^{\beta_2} [\exp\{-\mathcal{V}_{it-1}\}]^{\beta_3} [\exp\{\omega_{it-1}\}]^{\beta_4} \tag{D.6}$$

$$\mathcal{X}_{it-1} = \Phi\left(\gamma_0 + \gamma_1 \log K_{it-1} + \gamma_2 X_{it-1} - \gamma_3 \mathcal{V}_{it-1} + \gamma_4 \omega_{it-1}\right) - X_{it-1}, \tag{D.7}$$

where $\beta_1 = 0.8$, $\beta_2 = \beta_3 = \beta_4 = 0.1$, $\gamma_0 = -1$, $\gamma_1 = 0.1$, $\gamma_2 = 0.5$, $\gamma_3 = \gamma_4 = 0.1$ and $\Phi(\cdot)$ is the standard normal cdf. The firm-specific depreciation rates $\delta_i \in \{0.05, 0.075, 0.10, 0.125, 0.15\}$ are distributed uniformly across i. The initial levels of capital K_{i1} and export intensity X_{i1} are respectively drawn from $\mathbb{U}(10, 200)$ and $\mathbb{U}(0, 1)$ identically and independently distributed over firms. A particular form of the export adjustment rule in (D.7) along with the chosen $\{\gamma_j; j = 0, \ldots, 4\}$ ensures that the firm's export intensity stays bounded within the unit interval during the considered time periods, given our data generating process. Also, note that (D.6)–(D.7) imply an intuitively negative effect of the export-associated cost shifter $\mathcal{V}_{it}$ on both the investment and export intensity.

The optimal materials series are generated solving the firm's static restricted profit maximization problem along the lines of (2.5) after having already generated the series of $(K_{it}, X_{it}, \omega_{it})$ for each firm and time period. The conditional demand for M_{it} is given by

$$\begin{aligned} M_{it} &= \arg\max_{\mathcal{M}_{it}} \left\{ \left[P_t^X X_{it} + P_t^D (1 - X_{it})\right] K_{it}^{\alpha_1} \mathcal{M}_{it}^{\alpha_2} \exp\{\omega_{it}\} \mathcal{E} - P_t^M \mathcal{M}_{it} \right\} \\ &= \left(\frac{1}{P_t^M} \left[P_t^X X_{it} + P_t^D (1 - X_{it})\right] \alpha_2 K_{it}^{\alpha_1} \exp\{\omega_{it}\} \mathcal{E} \right)^{1/(1-\alpha_2)}, \end{aligned} \tag{D.8}$$

where we normalize $P_t^M = \mathcal{E}\ \forall\ t$ and intentionally assume away any temporal variation in the output prices: $P_t^X = 2$ and $P_t^D = 1$ for all t. Our decision to impose zero time variation in prices is dictated by Gandhi et al.'s (2017) recent findings whereby, at least in theory, the time-series variation in prices may help identify the production function even in the absence of an excluded relevant exogenous variable (in our case, X_{it} subject to the exogenous variation from $\mathcal{V}_{it}$).[22] By suppressing any such variation in prices, we are therefore able to restrict the source of identification of the production technology exclusively to the new element in the model, namely X_{it}.

Estimator. In Step 1 of our estimation procedure, we use the log-polynomial series of degree 2 to estimate r, $\mathbb{T}$ and $\mathbb{T}^*$. We do not cross-validate to minimize computational burden of our simulations. For the same reason, in Step 2, we approximate $f(\cdot)$ and $\varphi(\cdot)$ also using second-degree log-polynomial sieves thereby ensuring the second-step estimator has a closed-form solution, which obviates the need to numerically solve the least-squares optimization problem (as in the case of nonlinear ANN sieves used in our empirical application). For each simulation repetition, we compute the median, the root mean squared error (RMSE) and the mean absolute deviation

[21] For the sake of tractability, we follow Grieco et al. (2016) in postulating such policy rules instead of numerically deriving them from *each* firm's Bellman equation.

[22] Although their simulations suggest that the identification strategy based on temporal price variability performs poorly in practice.

Table D.1. Second-Step Estimates of Input Elasticities

	IV in Step 1: X_{it}		IV in Step 1: M_{it-2}	
	Capital	*Material*	*Capital*	*Material*
$n = 50$				
Median	0.2944	0.6031	0.7525	0.1107
RMSE	0.2614	0.1901	0.5283	0.5412
MAD	0.2142	0.1576	0.5016	0.5376
$n = 100$				
Median	0.2673	0.6367	0.8141	0.0546
RMSE	0.1807	0.1261	0.5744	0.5962
MAD	0.1481	0.1061	0.5638	0.5948
$n = 200$				
Median	0.2545	0.6486	0.8463	0.0288
RMSE	0.1226	0.0848	0.6017	0.6210
MAD	0.1004	0.0713	0.5956	0.6206
$n = 400$				
Median	0.2463	0.6522	0.8610	0.0150
RMSE	0.0822	0.0587	0.6131	0.6348
MAD	0.0671	0.0484	0.6111	0.6347
$n = 800$				
Median	0.2473	0.6528	0.8693	0.0082
RMSE	0.0546	0.0390	0.6201	0.6417
MAD	0.0455	0.0326	0.6188	0.6415

Notes: The true values of input elasticities are 0.25 for capital and 0.65 for materials. $T = 10$ throughout.

(MAD) of observation-specific[23] input elasticities (across firm-years) and report the medians of these metrics over 1,000 simulations.

We first estimate the production function via our proposed estimator using X_{it} to instrument for the endogenous M_{it}. The two-input analogues of (3.3) and (3.7) are

$$y_{it} = f(K_{it}, M_{it}) + \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it},\ \mathbb{E}[\zeta_{it} + \eta_{it}|\ K_{it}, X_{it}, K_{it-1}, X_{it-1}, M_{it-1}] = 0,$$

with the corresponding second-step simulation results reported in the left panel of Table D.1. The results are very encouraging and show that our approach recovers the true parameters well, thereby lending strong support to the validity of our identification strategy. As expected of a consistent estimator, the estimation becomes more stable with both the RMSE and MAD declining as the sample size grows.

Expectedly, the proxy approach however fails to identify the input elasticities if, in line with the convention in the literature, we employ exogenous but irrelevant higher-order lags such as M_{it-2} to instrument for M_{it} instead of using X_{it}. The right panel of Table D.1 reports the results for this

[23]The estimates are observation-specific despite that our DGP assumes a constant-elasticity Cobb-Douglas technology because our methodology estimates the unknown production function nonparametrically thereby delivering firm-year-specific estimates of the gradients of $f(\cdot)$.

Table D.2. Second-Step Estimates of Input Elasticities
for the DGP with no Exports

	Capital	*Material*
$n = 50$		
Median	0.6524	0.0760
RMSE	0.4255	0.5759
MAD	0.4036	0.5718
$n = 100$		
Median	0.6781	0.0391
RMSE	0.4408	0.6099
MAD	0.4264	0.6092
$n = 200$		
Median	0.6987	0.0199
RMSE	0.4514	0.6295
MAD	0.4477	0.6292
$n = 400$		
Median	0.7055	0.0105
RMSE	0.4572	0.6392
MAD	0.4553	0.6392
$n = 800$		
Median	0.7099	0.0046
RMSE	0.4613	0.6452
MAD	0.4599	0.6452

Notes: The true values of input elasticities are 0.25 for capital and 0.65 for materials. Both the DGP and model feature no export-related aspects; M_{it} is instrumented via M_{it-2} in Step 1.

model, i.e.,

$$y_{it} = f(K_{it}, M_{it}) + \varphi(K_{it-1}, X_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it},\ \mathbb{E}[\zeta_{it} + \eta_{it}|\ K_{it}, K_{it-1}, X_{it-1}, M_{it-1}, M_{it-2}] = 0.$$

This model grossly overestimates the capital elasticity and produces near-zero estimates of the material elasticity, showing no improvement in the estimation as n grows.

Lastly, to establish baseline for the under-identification problem in the model of firm production under the standard structural assumptions, we also examine the consequences of completely ignoring the exporting aspect of the firm's production behavior on the estimation of production function as oftentimes done in the literature. That is, we estimate the model assuming the more standard conceptual framework which ignores the firm's export decisions (including the learning by exporting effects) altogether. Appendix A discusses such a model in detail. The corresponding data generating process features no export-related aspects: we drop X_{it} from all equations and set $\rho_2 = \beta_2 = \beta_3 = 0$ and $P_t^X = P_t^X = 1$. The estimating equation where, following the tradition, M_{it-2} is employed to

Table D.3. Coverage Probability and Width of the 95% Bootstrap Confidence Intervals for Average Input Elasticities

	Capital	*Material*
$n = 200$		
Coverage Prob.	0.8767	0.8567
Interval Width	0.2726	0.2307
$n = 400$		
Coverage Prob.	0.9103	0.8970
Interval Width	0.1888	0.1597
$n = 600$		
Coverage Prob.	0.9467	0.9233
Interval Width	0.1535	0.1283

Notes: $T = 10$ throughout.

instrument for M_{it} is given by

$$y_{it} = f(K_{it}, M_{it}) + \varphi(K_{it-1}, M_{it-1}) + \zeta_{it} + \eta_{it},\ \mathbb{E}[\zeta_{it} + \eta_{it}|\ K_{it}, K_{it-1}, M_{it-1}, M_{it-2}] = 0,$$

with the corresponding results presented in Table D.2. In line with Gandhi et al.'s (2017) Monte-Carlo evidence, this under-identified—due to the conditional irrelevance of M_{it-2}—proxy estimator exhibits systematic biases in the estimates of production function, with the input elasticity estimates being nowhere near the true values and no improvement therein as n increases.

Bootstrap Inference. Using the same DGP, we also investigate a finite-sample performance of the accelerated bias-corrected bootstrap procedure that we employ for our proposed estimator. This is of interest because of the complexity of our multi-step regularized estimation procedure, which makes establishing the validity of such a bootstrap nontrivial. Just like earlier in this section, here our focus is on unknown input elasticities, which are estimated in the second step of our estimation procedure. Hence, we bootstrap only the first two stages, skipping step 5 of the bootstrap algorithm described in Section 4.1 of the paper. The estimation details are unchanged and the same as those used to populate the left panel of Table D.1.

We consider the 95% confidence intervals (i.e., $a = 0.05$) for the average elasticities of capital and labor. We focus on the *average* elasticities because our methodology estimates the unknown production function nonparametrically thereby delivering observation-specific estimates of the gradients of $f(\cdot)$ despite that our DGP assumes a constant-elasticity Cobb-Douglas technology. We average these estimates across firm-years to obtain scalar measures of elasticities denoted by $\bar{f}_K$ and $\bar{f}_M$ that we then study against the true $\beta_K = 0.25$ and $\beta_M = 0.65$.

Of interest are both the size and power. Table D.3 reports coverage probabilities and the median interval widths for $n = \{200, 400, 600\}$ and fixed $T = 10$. To conserve computational time, the number of simulations is set at $S = 300$ with $B = 200$ bootstrap replications per each simulation. The coverage probability is estimated as the relative frequency (over S simulations)

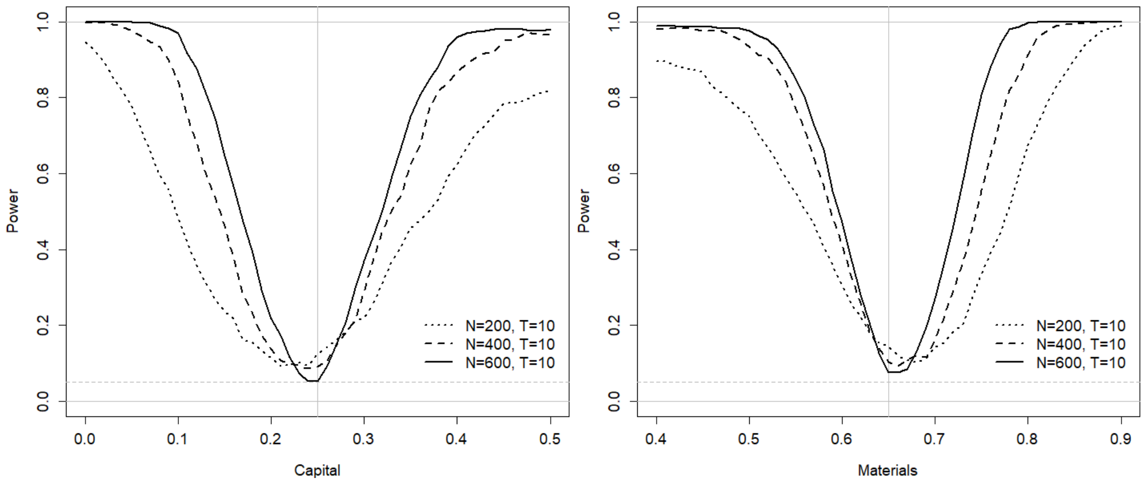

Figure D.1. Power of the Two-Sided Test for Average Input Elasticities
using the 95% Bootstrap Confidence Intervals

of the estimated 95% accelerated bias-corrected confidence interval containing the true elasticity parameter. The reported interval widths are the medians (across S simulations) of the distances between the upper and lower bounds of the estimated confidence intervals.

The results indicate that there may be size distortions for small n, which is quite common for nonparametric tests. However, with the sample size growing, the estimated coverage of our intervals approaches the correct coverage and, for n modestly large enough, the estimated rejection probabilities seem close to the nominal size. The employed bootstrap intervals exhibit good power that increases with the distance between the null and the true parameter value as can be seen in Figure D.1 which plots power curves for the estimated confidence intervals. Here, power is computed as the relative rejection frequency. Importantly, the power improves as the sample size increases which is anticipated of a consistent test. We see this both in the shrinking intervals widths (Table D.3) as well as in the increasing speed with which power curves attain unity (Figure D.1).

E Data

In this section, we provide additional details about construction of the data sample. The 28 two-digit China Standard Industrial Classification (CSIC) manufacturing industries that we include on our analysis are listed in Table E.1.

In 2002, the CSIC code system underwent slight changes at the two-digit level. The industry codes for the manufacturing of electrical machinery and equipment, manufacturing of communication equipment, computers and other electronic equipment, and manufacturing of measuring instruments and machinery for cultural activity and office work changed from 40, 41, 42 to 39, 40, 41, respectively. In addition, the original CSIC 43 industry ("other manufacturing") was split into two new, different industry groups. We match firm-year observations in the pre-2002 years with

Table E.1. Manufacturing Industries

CSIC	Description
13	Food Processing (from Agricultural Products)
14	Food
15	Beverages
16	Tobacco
17	Textile
18	Apparel and Other Fabric Products
19	Leather, Fur, Feather and Related Products
20	Timber
21	Furniture
22	Paper and Other Paper Products
23	Printing, Reproduction of Recording Media
24	(Office) Stationery, Educational and Sports Goods
25	Petroleum Refining, Coking and Processing of Nuclear Fuel
26	Raw Chemical Materials and Chemical Products
27	Medical and Pharmaceutical Products
28	Chemical Fiber
29	Rubber Products
30	Plastic Products
31	Nonmetal Mineral Products
32	Smelting and Pressing of Ferrous Metals
33	Smelting and Pressing of Nonferrous Metals
34	Metal Products
35	General Purpose Machinery
36	Special Purpose Machinery
37	Transport Equipment
39	Electric Machinery and Equipment
40	Communication Equipment, Computers and Other Electronics
41	Precision Machinery, Measuring Instruments and Meters

those classified in accordance with the new CSIC code (GB\T 4754 – 2002). Since the CSIC 42 and CSIC 43 industries did not exist in the old CSIC code (GB\T 4754 – 1994), we exclude them from our analysis.

Table E.2 reports basic summary statistics for our data, including the sample mean and standard deviation estimates for key variables tabulated by the CSIC industry groups. The table clearly points to significant heterogeneity across industries. For instance, the average firm size, as measured by the output value, ranges from 27.8 million RMB in the Printing industry to 97.4 million RMB in Petroleum Refining. Industries with the largest fraction of firms participating in exports include the manufacturing of Leather Products (64.6%), Apparel (66.1%) and Computers (53.3%). Interestingly, the heaviest presence of the state capital is seen in the production of tobacco goods (68.6%) with the second most state-invested industry being Printing (26.9%). Foreigners tend to invest more in the production of Computers, Office Goods and Apparel: more than 20% of firms in these industries reports having foreign capital in the equity.

Table E.2. Data Summary Statistics, 1999–2006

CSIC	Y	K	L	M	$X \mid X > 0$	Exporter	State	Foreign	Subsidy
13	55,683	11,726	1,581	38,895	0.601	0.216	0.195	0.098	0.128
	(91,049)	(24,692)	(2,870)	(64,771)	(0.412)				
14	48,688	15,812	2,458	33,089	0.548	0.248	0.171	0.143	0.127
	(85,595)	(32,403)	(4,330)	(58,834)	(0.439)				
15	59,485	27,064	2,799	38,696	0.420	0.135	0.250	0.104	0.118
	(99,480)	(48,397)	(4,812)	(66,165)	(0.434)				
16	121,697	64,351	7,597	74,720	0.179	0.080	0.686	0.018	0.152
	(158,099)	(67,079)	(7,413)	(102,147)	(0.278)				
17	49,571	15,143	2,944	35,107	0.646	0.410	0.069	0.098	0.139
	(76,274)	(31,366)	(4,534)	(55,229)	(0.379)				
18	36,638	6,334	3,455	25,158	0.859	0.661	0.035	0.209	0.122
	(59,243)	(13,612)	(4,478)	(41,807)	(0.299)				
19	48,011	7,035	3,751	33,401	0.856	0.646	0.024	0.190	0.118
	(76,481)	(14,259)	(5,557)	(54,737)	(0.316)				
20	34,127	10,448	1,728	23,464	0.719	0.293	0.082	0.096	0.155
	(56,708)	(28,551)	(2,762)	(40,317)	(0.375)				
21	41,420	9,511	2,851	28,548	0.787	0.463	0.048	0.157	0.113
	(65,747)	(21,140)	(4,726)	(45,953)	(0.367)				
22	39,756	13,388	1,984	27,864	0.482	0.132	0.083	0.065	0.121
	(63,213)	(30,754)	(3,209)	(44,484)	(0.407)				
23	27,795	13,307	2,151	18,268	0.498	0.107	0.269	0.060	0.110
	(53,211)	(25,723)	(3,697)	(35,124)	(0.417)				
24	39,341	7,955	3,726	27,340	0.852	0.749	0.034	0.208	0.149
	(61,771)	(16,777)	(5,423)	(43,882)	(0.304)				
25	97,379	28,138	2,724	68,461	0.305	0.059	0.121	0.060	0.108
	(145,996)	(53,831)	(4,681)	(106,756)	(0.375)				
26	52,981	15,410	2,293	36,941	0.415	0.225	0.134	0.088	0.163
	(87,251)	(35,219)	(4,027)	(62,066)	(0.359)				
27	58,718	22,182	3,351	37,543	0.395	0.222	0.207	0.105	0.181
	(87,346)	(38,701)	(5,346)	(58,173)	(0.365)				
28	77,280	32,057	2,687	56,694	0.444	0.197	0.086	0.105	0.176
	(113,607)	(62,971)	(4,322)	(83,797)	(0.378)				
29	44,127	12,376	2,946	30,326	0.601	0.341	0.081	0.116	0.129
	(78,259)	(28,431)	(4,616)	(55,164)	(0.409)				
30	38,863	11,468	2,154	27,304	0.669	0.330	0.051	0.125	0.126
	(63,048)	(25,916)	(3,413)	(45,200)	(0.402)				
31	36,878	15,987	2,389	24,809	0.553	0.159	0.140	0.058	0.160
	(57,787)	(33,493)	(3,591)	(39,350)	(0.420)				
32	83,951	16,410	2,478	59,758	0.403	0.097	0.083	0.040	0.111
	(130,908)	(33,855)	(4,249)	(95,700)	(0.365)				
33	83,154	17,405	2,362	59,761	0.389	0.178	0.088	0.062	0.171
	(126,831)	(41,760)	(4,221)	(93,077)	(0.368)				
34	40,393	9,029	2,279	28,335	0.689	0.331	0.061	0.106	0.126
	(71,140)	(20,513)	(3,650)	(51,348)	(0.386)				
35	39,959	10,252	2,618	27,354	0.485	0.268	0.108	0.094	0.144
	(69,351)	(23,179)	(4,227)	(48,675)	(0.397)				
36	43,249	11,787	3,033	29,317	0.357	0.239	0.167	0.099	0.148
	(74,656)	(24,839)	(5,080)	(52,616)	(0.362)				
37	54,702	15,739	3,479	38,035	0.457	0.222	0.155	0.101	0.160
	(93,397)	(32,781)	(5,546)	(66,975)	(0.397)				
39	59,328	11,622	3,087	41,718	0.621	0.326	0.079	0.121	0.166
	(99,034)	(24,713)	(4,929)	(70,752)	(0.410)				
40	76,803	20,297	5,249	53,144	0.678	0.533	0.104	0.247	0.178
	(121,183)	(40,980)	(7,087)	(88,139)	(0.394)				
41	45,410	11,724	3,885	30,396	0.654	0.444	0.136	0.188	0.194
	(79,157)	(25,411)	(5,739)	(55,665)	(0.412)				
All	48,418	13,357	2,756	33,531	0.646	0.314	0.110	0.112	0.144
	(82,518)	(29,920)	(4,487)	(58,731)	(0.403)				

Notes: Reported are the sample means with sample standard deviations in parentheses. Y – the gross output; K – capital stock; L – labor; M – materials; X – the export intensity. Y, K, L and M are measured in thousands of RMB. X is unit-free proportions. “Exporter”, “State”, “Foreign” and “Subsidy” are binary indicators equal to one if the firm is an exporter ($X_{it} > 0$), state-invested, foreign-invested and subsidized, respectively.

F Additional Results

F.1 Returns to Scale

We first take a look at the scale efficiency of manufacturing firms in China, where it is defined as the industry's attainment of the efficient firm scale size associated with (unitary) constant returns to scale. Table F.1 reports the average estimates of the returns to scale exhibited by firms across industries and years. We compute the firm-year-specific returns to scale as the sum of log-derivatives of the gross production function $f(\cdot)$ estimated in the second step: $\text{RTS}_{it} = \sum_{a\in\{k,l,m\}} \partial\widehat{f}(K_{it}, L_{it}, M_{it})/\partial a_{it}$, where $\widehat{f}(\cdot)$ is the estimated sieve approximation given in (4.13). The instance when the RTS estimate is significantly less than/equal to/greater than one corresponds to decreasing/constant/increasing returns to scale.

From Table F.1, we observe that all confidence intervals for the industry-level mean RTS estimates include unity suggesting that an *average* firm in all manufacturing industries operated at the constant returns to scale (at the 5% significance level) in 2000–2006. Thus, the manufacturing sector in China is generally scale efficient. There is also some evidence that, in most industries, the returns to scale have been increasing (albeit statically insignificantly) over the course of our sample period. We now proceed to the analysis of firm-level productivity, which constitutes the primary focus of our paper.

F.2 Industry-Level Productivity Growth

To explore potentially heterogeneous experiences of individual industry groups, we perform a productivity growth decomposition for each individual two-digit CSIC industry.

Table F.2 reports the results, and they generally bolster the narrative whereby China's manufacturing industries have experienced no tangible reallocation of factors from less to more productive firms during the 2000–2006 period. The only notable exceptions are the Apparel and Metals (both ferrous and nonferrous) industries which exhibited positive reallocation effects that contributed at least 3.5% points to the cumulative aggregate productivity growth in these industries over 6 years. The Nonferrous Metals industry (CSIC 33) is the leader in this respect having experienced a cumulative reallocation effect of 7.7% which amounted to slightly less than a fifth of the total productivity improvement observed in the industry. In contrast, the majority of industries (17 out of 28) experienced a worsening in allocation of productive factors with a cumulative *decline* in the sample covariance of firm-level output and productivity, which attests to the reallocation of output from *more to less* productive firms in these industries. For instance, in industries such as Printing, Timber, Special Purpose Machinery, Communication Equipment and Measuring Instruments, this reallocation effect shaved 8% points or more off the cumulative cumulative productivity growth. In case of the Computer industry (CSIC 40), the total contribution of negative reallocation effects is estimated at a non-negligible 11.4% which has brought the cumulative growth in the industry productivity from 14.5 down to 3.1% only.

Table F.1. Mean Returns to Scale Estimates across Industries, 2000–2006

CSIC	2000	2001	2002	2003	2004	2005	2006
13	0.885	0.887	0.889	0.890	0.889	0.888	0.888
	(0.438, 1.561)	(0.433, 1.545)	(0.434, 1.547)	(0.432, 1.532)	(0.430, 1.508)	(0.429, 1.532)	(0.427, 1.520)
14	0.893	0.894	0.895	0.896	0.896	0.897	0.897
	(0.668, 1.331)	(0.658, 1.337)	(0.637, 1.342)	(0.626, 1.346)	(0.611, 1.352)	(0.617, 1.358)	(0.610, 1.361)
15	0.881	0.881	0.884	0.885	0.891	0.890	0.891
	(0.522, 1.313)	(0.527, 1.306)	(0.529, 1.309)	(0.530, 1.312)	(0.534, 1.311)	(0.534, 1.320)	(0.535, 1.318)
16	0.900	0.906	0.912	0.911	0.906	0.901	0.899
	(0.748, 2.689)	(0.751, 2.725)	(0.758, 2.618)	(0.765, 2.722)	(0.753, 2.809)	(0.746, 2.812)	(0.747, 2.886)
17	0.891	0.892	0.894	0.894	0.896	0.894	0.895
	(0.282, 1.457)	(0.283, 1.457)	(0.285, 1.466)	(0.287, 1.469)	(0.270, 1.473)	(0.279, 1.497)	(0.281, 1.487)
18	0.906	0.906	0.906	0.905	0.903	0.901	0.902
	(0.669, 1.286)	(0.671, 1.283)	(0.674, 1.281)	(0.673, 1.281)	(0.673, 1.284)	(0.676, 1.270)	(0.671, 1.277)
19	0.883	0.886	0.886	0.889	0.889	0.889	0.889
	(0.678, 1.219)	(0.682, 1.218)	(0.683, 1.219)	(0.682, 1.217)	(0.680, 1.225)	(0.681, 1.228)	(0.683, 1.224)
20	0.881	0.884	0.889	0.892	0.893	0.895	0.896
	(0.385, 1.346)	(0.382, 1.337)	(0.392, 1.347)	(0.393, 1.339)	(0.394, 1.344)	(0.394, 1.358)	(0.399, 1.346)
21	0.902	0.903	0.904	0.904	0.904	0.905	0.907
	(0.634, 2.304)	(0.635, 2.322)	(0.637, 2.268)	(0.640, 2.272)	(0.640, 2.327)	(0.641, 2.397)	(0.651, 2.363)
22	0.912	0.913	0.914	0.914	0.914	0.914	0.914
	(0.770, 1.225)	(0.770, 1.229)	(0.771, 1.230)	(0.772, 1.232)	(0.767, 1.235)	(0.775, 1.232)	(0.774, 1.230)
23	0.909	0.910	0.911	0.911	0.913	0.911	0.912
	(0.852, 1.008)	(0.844, 1.013)	(0.840, 1.017)	(0.827, 1.028)	(0.824, 1.043)	(0.823, 1.037)	(0.821, 1.048)
24	0.899	0.901	0.903	0.905	0.906	0.905	0.906
	(0.719, 1.269)	(0.720, 1.271)	(0.733, 1.280)	(0.734, 1.282)	(0.734, 1.286)	(0.731, 1.280)	(0.734, 1.283)
25	0.896	0.900	0.899	0.902	0.902	0.903	0.904
	(0.268, 1.487)	(0.271, 1.487)	(0.272, 1.479)	(0.276, 1.464)	(0.281, 1.457)	(0.292, 1.459)	(0.289, 1.447)
26	0.893	0.894	0.897	0.899	0.901	0.899	0.901
	(0.409, 1.723)	(0.402, 1.730)	(0.405, 1.743)	(0.406, 1.740)	(0.406, 1.743)	(0.412, 1.738)	(0.414, 1.732)
27	0.879	0.881	0.883	0.883	0.884	0.884	0.884
	(0.689, 1.161)	(0.686, 1.168)	(0.688, 1.168)	(0.680, 1.177)	(0.669, 1.191)	(0.676, 1.185)	(0.676, 1.187)
28	0.918	0.917	0.915	0.920	0.920	0.919	0.920
	(0.681, 1.338)	(0.681, 1.336)	(0.681, 1.342)	(0.685, 1.356)	(0.685, 1.351)	(0.689, 1.352)	(0.687, 1.364)
29	0.905	0.904	0.907	0.909	0.909	0.909	0.910
	(0.639, 1.419)	(0.637, 1.425)	(0.639, 1.421)	(0.639, 1.423)	(0.638, 1.433)	(0.644, 1.432)	(0.642, 1.434)
30	0.894	0.894	0.894	0.896	0.898	0.894	0.896
	(0.367, 1.553)	(0.366, 1.550)	(0.366, 1.554)	(0.367, 1.555)	(0.368, 1.554)	(0.369, 1.562)	(0.371, 1.558)
31	0.901	0.902	0.903	0.903	0.904	0.904	0.905
	(0.460, 1.986)	(0.460, 1.971)	(0.453, 1.922)	(0.452, 1.913)	(0.446, 1.847)	(0.449, 1.828)	(0.448, 1.840)
32	0.902	0.904	0.905	0.907	0.909	0.907	0.908
	(0.390, 1.712)	(0.389, 1.666)	(0.410, 1.746)	(0.423, 1.720)	(0.426, 1.714)	(0.424, 1.749)	(0.425, 1.717)
33	0.903	0.905	0.906	0.906	0.908	0.904	0.907
	(0.378, 1.508)	(0.380, 1.505)	(0.371, 1.499)	(0.376, 1.505)	(0.373, 1.505)	(0.370, 1.512)	(0.375, 1.504)
34	0.901	0.903	0.905	0.906	0.907	0.906	0.908
	(0.680, 1.203)	(0.682, 1.204)	(0.683, 1.205)	(0.683, 1.204)	(0.682, 1.207)	(0.684, 1.208)	(0.683, 1.206)
35	0.902	0.905	0.908	0.910	0.912	0.908	0.910
	(0.101, 1.751)	(0.102, 1.739)	(0.088, 1.732)	(0.083, 1.720)	(0.055, 1.723)	(0.073, 1.757)	(0.055, 1.737)
36	0.899	0.900	0.903	0.906	0.906	0.906	0.907
	(0.572, 1.310)	(0.572, 1.309)	(0.574, 1.311)	(0.571, 1.312)	(0.566, 1.314)	(0.573, 1.321)	(0.573, 1.320)
37	0.906	0.907	0.908	0.909	0.909	0.908	0.908
	(0.608, 1.439)	(0.597, 1.443)	(0.601, 1.441)	(0.600, 1.444)	(0.589, 1.445)	(0.590, 1.448)	(0.589, 1.445)
39	0.904	0.905	0.906	0.908	0.910	0.908	0.909
	(0.755, 1.287)	(0.755, 1.288)	(0.758, 1.290)	(0.756, 1.294)	(0.757, 1.297)	(0.768, 1.291)	(0.759, 1.294)
40	0.884	0.887	0.890	0.891	0.892	0.889	0.891
	(0.401, 1.376)	(0.402, 1.379)	(0.415, 1.376)	(0.417, 1.371)	(0.421, 1.359)	(0.406, 1.404)	(0.417, 1.391)
41	0.901	0.901	0.904	0.905	0.904	0.904	0.904
	(0.599, 1.259)	(0.600, 1.261)	(0.600, 1.262)	(0.601, 1.259)	(0.592, 1.266)	(0.605, 1.265)	(0.601, 1.262)
All	0.898	0.899	0.901	0.902	0.903	0.901	0.902
	(0.843, 1.411)	(0.841, 1.423)	(0.845, 1.425)	(0.846, 1.432)	(0.848, 1.460)	(0.846, 1.509)	(0.844, 1.504)

Notes: Reported are the average nonparametric estimates of the returns to scale with the 95% accelerated bias-corrected percentile bootstrap confidence intervals in parentheses.

Table F.2. (Weighted) Aggregate Productivity Growth Decomposition across Industries

Year	Δp_t	$\Delta \overline{p}_t$	Δcov_t	Δp_t	$\Delta \overline{p}_t$	Δcov_t	Δp_t	$\Delta \overline{p}_t$	Δcov_t	Δp_t	$\Delta \overline{p}_t$	Δcov_t
		CSIC 13			**CSIC 14**			**CSIC 15**			**CSIC 16**	
2001	0.015	0.029	–0.013	–0.017	0.009	–0.027	0.026	0.013	0.013	–0.067	–0.007	–0.060
2002	0.026	0.039	–0.013	0.030	0.044	–0.014	0.003	0.033	–0.030	0.072	0.058	0.014
2003	0.066	0.069	–0.003	0.042	0.046	–0.004	0.058	0.051	0.007	0.035	–0.030	0.064
2004	0.085	0.065	0.020	0.041	0.058	–0.017	0.057	0.080	–0.023	0.090	0.093	–0.003
2005	0.043	0.069	–0.027	0.089	0.084	0.006	0.088	0.087	0.000	0.034	0.037	–0.003
2006	0.041	0.045	–0.004	0.060	0.059	0.000	0.060	0.058	0.002	0.059	0.051	0.007
Cum.	0.277	0.316	–0.040	0.245	0.300	–0.055	0.292	0.323	–0.030	0.223	0.203	0.020
		CSIC 17			**CSIC 18**			**CSIC 19**			**CSIC 20**	
2001	0.002	–0.009	0.011	–0.021	–0.030	0.010	–0.016	–0.007	–0.009	–0.006	0.003	–0.009
2002	0.004	0.005	–0.001	–0.030	–0.031	0.001	–0.041	–0.028	–0.013	–0.027	0.017	–0.045
2003	0.034	0.022	0.012	0.003	–0.001	0.004	0.027	0.000	0.027	0.009	0.031	–0.022
2004	0.036	0.041	–0.004	0.045	0.039	0.006	0.026	0.027	–0.001	0.054	0.045	0.009
2005	0.062	0.067	–0.006	0.058	0.054	0.004	0.041	0.041	0.000	0.044	0.070	–0.027
2006	0.044	0.043	0.001	0.029	0.019	0.009	0.038	0.030	0.008	0.051	0.053	–0.002
Cum.	0.181	0.169	0.013	0.085	0.050	0.035	0.074	0.063	0.011	0.124	0.220	–0.096
		CSIC 21			**CSIC 22**			**CSIC 23**			**CSIC 24**	
2001	0.007	0.009	–0.002	0.012	0.009	0.003	0.009	0.009	0.000	–0.035	–0.024	–0.011
2002	–0.011	–0.007	–0.004	0.006	0.019	–0.013	0.018	0.034	–0.016	–0.023	–0.019	–0.003
2003	–0.007	0.007	–0.014	0.048	0.036	0.011	0.018	0.044	–0.026	0.019	0.007	0.012
2004	0.071	0.049	0.022	0.061	0.056	0.005	0.039	0.062	–0.023	0.015	0.039	–0.023
2005	0.003	0.044	–0.040	0.069	0.057	0.013	0.066	0.086	–0.019	0.065	0.046	0.018
2006	0.027	0.033	–0.006	0.042	0.044	–0.002	0.036	0.047	–0.011	0.011	0.015	–0.004
Cum.	0.090	0.134	–0.044	0.237	0.220	0.017	0.185	0.280	–0.094	0.052	0.064	–0.011
		CSIC 25			**CSIC 26**			**CSIC 27**			**CSIC 28**	
2001	0.007	0.020	–0.012	0.020	0.021	–0.001	–0.005	0.020	–0.024	0.007	–0.019	0.026
2002	0.050	0.015	0.035	0.018	0.023	–0.005	0.000	0.003	–0.003	0.016	0.028	–0.012
2003	0.084	0.096	–0.013	0.064	0.047	0.016	0.021	0.037	–0.016	0.060	0.083	–0.022
2004	0.054	0.093	–0.038	0.077	0.069	0.008	0.080	0.061	0.018	0.054	0.058	–0.004
2005	0.053	0.038	0.015	0.071	0.077	–0.006	0.080	0.066	0.014	0.092	0.073	0.019
2006	0.051	0.060	–0.009	0.062	0.055	0.007	0.049	0.037	0.012	0.033	0.045	–0.011
Cum.	0.299	0.322	–0.023	0.311	0.291	0.020	0.224	0.224	0.001	0.262	0.268	–0.005
		CSIC 29			**CSIC 30**			**CSIC 31**			**CSIC 32**	
2001	–0.015	–0.026	0.011	0.001	0.000	0.000	0.008	0.003	0.005	0.005	0.003	0.002
2002	0.013	0.028	–0.015	–0.011	–0.011	0.000	0.022	0.026	–0.004	0.034	0.038	–0.004
2003	0.041	0.041	–0.001	0.005	0.032	–0.027	0.066	0.048	0.018	0.100	0.077	0.023
2004	0.043	0.032	0.012	0.040	0.054	–0.014	0.066	0.066	0.000	0.118	0.104	0.014
2005	0.043	0.064	–0.021	0.055	0.057	–0.003	0.069	0.076	–0.007	0.081	0.052	0.029
2006	0.076	0.065	0.011	0.052	0.037	0.015	0.069	0.059	0.010	0.036	0.054	–0.018
Cum.	0.202	0.205	–0.003	0.142	0.170	–0.028	0.300	0.277	0.023	0.374	0.328	0.046
		CSIC 33			**CSIC 34**			**CSIC 35**			**CSIC 36**	
2001	0.024	0.022	0.002	0.003	0.002	0.001	0.011	0.009	0.002	–0.004	0.010	–0.015
2002	0.027	0.019	0.008	–0.005	0.007	–0.012	0.023	0.023	–0.001	0.027	0.037	–0.010
2003	0.080	0.038	0.042	0.041	0.038	0.003	0.060	0.057	0.003	0.008	0.069	–0.061
2004	0.087	0.099	–0.012	0.051	0.051	0.000	0.061	0.074	–0.013	0.064	0.072	–0.008
2005	0.093	0.057	0.037	0.059	0.054	0.005	0.062	0.061	0.002	0.070	0.074	–0.003
2006	0.101	0.099	0.001	0.041	0.041	0.001	0.055	0.052	0.003	0.065	0.055	0.010
Cum.	0.411	0.334	0.077	0.190	0.193	–0.003	0.272	0.275	–0.003	0.230	0.316	–0.086
		CSIC 37			**CSIC 39**			**CSIC 40**			**CSIC 41**	
2001	0.025	0.020	0.005	0.005	0.003	0.002	–0.007	0.013	–0.019	–0.037	0.002	–0.039
2002	0.034	0.027	0.007	–0.002	0.004	–0.007	–0.060	–0.019	–0.041	–0.017	0.004	–0.021
2003	0.026	0.052	–0.027	0.045	0.044	0.001	0.009	0.038	–0.029	0.054	0.044	0.010
2004	0.048	0.065	–0.017	0.053	0.051	0.002	0.066	0.070	–0.003	0.047	0.079	–0.032
2005	0.053	0.064	–0.011	0.067	0.054	0.012	0.014	0.017	–0.003	0.044	0.065	–0.021
2006	0.054	0.043	0.010	0.063	0.049	0.014	0.008	0.026	–0.018	0.059	0.037	0.022
Cum.	0.240	0.272	–0.032	0.230	0.205	0.025	0.031	0.145	–0.114	0.151	0.231	–0.080

Notes: Reported are the annual growth rates of the (weighted) aggregate productivity p_t and of its two components: the (unweighted) average productivity $\overline{p}_t$ and the covariance between the firm-level output and productivity cov_t.

F.3 (Conditional) Exporter Productivity Differentials by Export Intensity

Table F.3. Median (Log) Exporter Productivity Premia by Export Intensity, 2000–2006

CSIC	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
13	–0.053	–0.051	–0.067	–0.075	–0.111	–0.132	–0.149	–0.163	–0.183	–0.120
14	–0.070	–0.089	–0.075	–0.082	–0.088	–0.089	–0.093	–0.135	–0.119	–0.129
15	*–0.029*	*0.003*	*–0.022*	*–0.034*	–0.081	–0.062	–0.082	–0.040	–0.048	–0.071
16	*–0.005*	*–0.006*	*–0.045*	*0.010*	*0.042*	*–0.067*	*–0.106*	*–0.095*	*–0.054*	*–0.195*
17	–0.046	–0.073	–0.090	–0.101	–0.118	–0.129	–0.133	–0.149	–0.152	–0.159
18	–0.067	–0.119	–0.148	–0.157	–0.167	–0.170	–0.178	–0.173	–0.175	–0.179
19	–0.034	–0.106	–0.133	–0.168	–0.183	–0.188	–0.197	–0.195	–0.205	–0.204
20	–0.046	–0.059	–0.085	–0.116	–0.128	–0.152	–0.145	–0.144	–0.148	–0.153
21	–0.099	–0.107	–0.129	–0.146	–0.149	–0.195	–0.201	–0.203	–0.229	–0.204
22	–0.034	*–0.020*	–0.045	–0.045	–0.076	–0.089	–0.098	–0.151	–0.170	–0.170
23	–0.064	–0.069	*–0.021*	–0.094	–0.073	–0.091	–0.137	–0.199	–0.247	–0.192
24	–0.075	–0.120	–0.140	–0.174	–0.175	–0.190	–0.197	–0.191	–0.191	–0.189
25	*–0.057*	–0.168	–0.165	–0.143	*–0.046*	*–0.047*	–0.122	–0.100	–0.317	–0.119
26	–0.080	–0.063	–0.081	–0.087	–0.090	–0.107	–0.123	–0.135	–0.156	–0.176
27	–0.071	–0.066	–0.056	–0.071	–0.108	–0.133	–0.115	–0.118	–0.115	–0.051
28	–0.058	*–0.021*	–0.078	–0.062	*–0.040*	–0.096	–0.055	–0.107	–0.092	–0.126
29	–0.084	–0.060	–0.103	–0.105	–0.148	–0.142	–0.152	–0.184	–0.198	–0.192
30	–0.088	–0.091	–0.094	–0.125	–0.145	–0.178	–0.202	–0.201	–0.210	–0.224
31	*–0.007*	–0.018	–0.038	–0.054	–0.064	–0.099	–0.115	–0.128	–0.166	–0.158
32	*–0.030*	–0.056	–0.077	–0.096	–0.113	–0.137	–0.155	–0.152	–0.170	–0.145
33	–0.120	–0.097	–0.133	–0.124	–0.145	–0.170	–0.177	–0.172	–0.226	–0.193
34	–0.086	–0.087	–0.116	–0.134	–0.151	–0.164	–0.175	–0.175	–0.189	–0.195
35	–0.087	–0.094	–0.098	–0.107	–0.128	–0.130	–0.141	–0.156	–0.170	–0.158
36	–0.099	–0.111	–0.105	–0.107	–0.086	–0.110	–0.123	–0.133	–0.152	–0.177
37	–0.081	–0.074	–0.057	–0.087	–0.089	–0.094	–0.110	–0.134	–0.158	–0.154
39	–0.087	–0.103	–0.107	–0.135	–0.154	–0.188	–0.194	–0.210	–0.201	–0.226
40	–0.069	–0.097	–0.131	–0.150	–0.176	–0.189	–0.221	–0.217	–0.236	–0.242
41	–0.091	–0.082	–0.106	–0.144	–0.127	–0.175	–0.179	–0.203	–0.224	–0.242
All	–0.071	–0.073	–0.096	–0.115	–0.138	–0.160	–0.177	–0.189	–0.203	–0.207

Notes: Reported are the conditional median estimates of the exporter productivity premium tabulated by deciles of the empirical distribution of non-zero export intensity. All point estimates, except those in *italic*, are statistically significant at the 5% level.