Preliminary draft
Please do not cite

Experience, Expectations, and Export Dynamics

Mita Das
Indian Statistical Institute

Mark Roberts
Pennsylvania State University and the NBER

James Tybout
Georgetown University

December 4, 1998

Abstract

The responses of industrial exports to regime changes are notoriously hard to predict. In this paper we investigate whether the predictability problem traces partly to micro phenomena that are undetectable with aggregated data. We begin by developing a dynamic model of exporting behavior that allows for uncertainty, heterogenous firms, and one-shot entry costs for firms breaking into foreign markets. Then we fit the model to plant-level panel data on Colombian chemical producers. Finally, using the results, we simulate aggregate export trajectories under alternative exchange rate regimes to quantify the effects of entry costs, regime credibility, and heterogeneity. Each proves to matter, but a substantial portion of the randomness in aggregate exports remains unexplained.

This paper was partially funded by the World Bank research project "Micro Foundations of Successful Export Promotion," RPO 679-20. The views expressed herein are our own and do not necessarily reflect those of the World Bank. Without implicating them, we would like to thank Behzad Diba, Martin Evans, Andrew Levin, John Rust, and Ariel Pakes for helpful discussions.

I. Overview

In developing countries, industrial exports are famously unpredictable. Seemingly similar stimuli have given rise to very different export responses in different countries and time periods, leaving analysts to wonder whether the next reform package will generate a surge or a trickle. When foreign exchange reserves are dwindling or domestic demand is slack, the stakes are substantial.

It is not hard to identify micro explanations for the predictability problem. First, a strong export response often means convincing non-exporters to initiate foreign sales. But to break into foreign markets, firms must establish marketing channels, learn bureaucratic procedures, and develop new packaging or product varieties.[1] Exchange rate movements that are viewed as temporary may not induce firms to bear these start-up costs, while credible regime shifts may trigger strong responses. Second, even within narrowly defined industries, firms are quite heterogeneous in terms of their production costs and their product characteristics. For some, small perturbations to the return from exporting may entice them into foreign markets, while for others, dramatic changes in the incentive structure may be needed. The more heterogeneity, the less likely it is that an export boom can be triggered by moving the exchange rate past some critical threshold. Finally, history matters. If a core of exporting firms has already been established, it is less important to induce non-exporters to retool for foreign markets, and export responses may largely reflect volume adjustment among these incumbents.

[1] Start-up costs are the focus of the analytical literature on export hysteresis (Baldwin and Krugman, 1989; Dixit, 1989; Krugman, 1989). For evidence on the role of new exporters see Roberts and Tybout (1996).

Unfortunately, we have little sense for which of these factors is quantitatively important, or what to look for when assessing potential export responsiveness in a particular context. Early attempts to model export supplies were based on macro data, so they did not quantify the effects of sunk costs, expectations, or firm heterogeneity in shaping export responsiveness.[2] Several more recent studies have used micro panel data to test the hypothesis that sunk costs matter, and have concluded that they clearly do (Roberts and Tybout, 1997a; Sullivan, 1996; Bernard and Jensen, 1996). But these second generation studies have been based on reduced form relationships between current exporting status, exporting history, and exogenous shocks. Thus they have not recovered the deep parameters that link behavior to stochastic processes, and they have not addressed the question of what would happen if regime changes were to stabilize the exchange rate, or change its mean realization. Nor have they shed much light on the role of firm-level distributions of marginal costs or foreign demand conditions. The purpose of the present paper is to estimate a structural model that quantifies the role of each of these factors using firm-level panel data.

II. An Empirical Model of Exporting Decisions with Sunk Costs

Previous micro studies of export dynamics have avoided structural models because they are computationally complex. Each period, firms decide whether to export on the basis of observable information about uncertain future market conditions, as well as entry costs, exit costs, and their current exporting status. Hence, for each possible decision, the evaluation of

[2] Roberts and Tybout (1997a) briefly review studies that use macro data to assess the effects of start-up costs on exports.

future export profit trajectories involves multi-dimensional integration over realizations on all the state variables.[3] If there are more than several state variables, the problem quickly becomes computationally intractable.

Given these difficulties, we approach structural estimation by, first, keeping the number of state variables in the model small, and second, using techniques recently developed by Rust (1997) that afford new flexibility in estimation. Our key assumptions are the following:

- *The foreign and domestic market for each firm's product are monopolistically competitive.* This eliminates strategic competition, but it ensures that each firm faces a downward-sloping marginal revenue function in each market. It seems like a very reasonable assumption for most manufactured products in a semi-industrialized country.

- *Producers are heterogeneous in terms of their marginal production costs and the foreign demand schedules they face for their products*. Heterogeneity is certainly present in the data, and it is a potentially important explanation for unpredictable export supply responses.

- *Future realizations on the exchange rates, marginal costs, and foreign demand shifters are unknown, but each evolves according to a known Markov process.* Uncertainty matters in sunk-cost models (e.g., Dixit, 1989), so it is critical to allow for it.

- *Plants are risk-neutral and maximize the expected discounted sum of real profits.* This standard assumption eliminates the need to deal with parameters that measure risk aversion.

- *Marginal costs do not respond to output shocks.* This assumption implies that shocks that shift the domestic demand schedule do not affect the optimal level of exports, so it allows us to focus on the export market only. The assumption appears to be reasonable for the industry, country, and time period we will study, since some excess capacity was present.[4]

[3] Because the decision to export is a discrete choice, optimal behavior cannot be characterized using Euler equations. Hence integration cannot be avoided.

[4] Estimates of average variable cost functions revealed little dependence on within-plant temporal output fluctuations.

1.2 The static optimization problem

Let us now specify the plant-level demand and cost functions, and derive the associated operating profit function.[5] First, conditioned on the real peso/dollar exchange rate (e_t) and a firm-specific stochastic intercept (m_{it}^f), foreign demand is isoelastic, so in logarithms we may write:

$$q_{it}^f = -\eta^f p_{it}^f + \psi_1 e_t + m_{it}^f, \tag{1}$$

where q_{it}^f is the log of the i^{th} firm's exports in year t, $\eta^f > 1$ is the elasticity of foreign demand, p_{it}^f is the log of the firm's real price of exports in domestic currency, and the random variable m_{it}^f captures factors that shift the demand curve like foreign income fluctuations and changes in the prices of competing products.[6] Analogously, we write the home (h) market demand schedule as:

$$q_{it}^h = -\eta^h p_{it}^h + m_{it}^h. \tag{2}$$

Next, let the logarithm of marginal costs be specified as:

$$c_{it} = a_{it} + \psi_2 e_t, \tag{3}$$

[5] The static optimization problem developed here is based on Roberts, Sullivan and Tybout (1995) and Clerides, Lach and Tybout (1998).

[6] This demand equation can be derived from the Dixit/Stiglitz (1977) utility function for differentiated products under the assumption that each firm is atomistic in the foreign market. The coefficient on the exchange rate (ψ_1) is distinct from the own elasticity of demand (η^f) because exchange rate movements change the foreign currency price of *all* Colombian exports, while changes in p only change the price of the individual firm's product.

where a_{it} is a random variable that evolves over time with shocks to real factor prices and technology. The exchange rate appears in our marginal cost expression because it affects imported input prices and real wages.

For any firm that has decided to export, equating marginal revenue to marginal cost and solving for $\boldsymbol{q}_{it}^{f}$ yields the optimal volume of foreign sales (in logarithms):

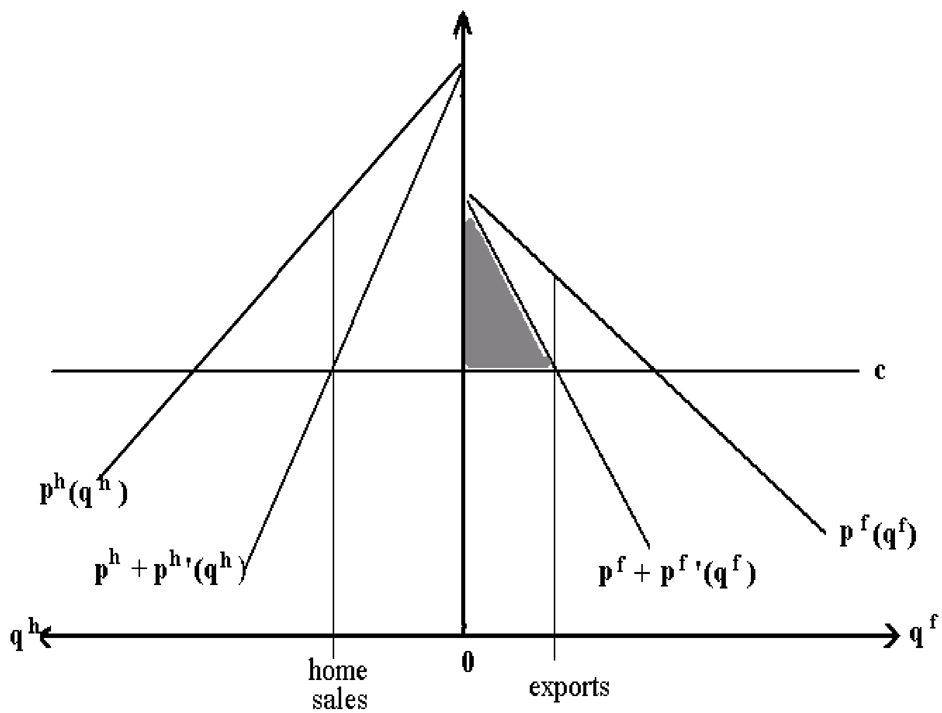
$$\boldsymbol{q}_{it}^{f*} = \eta^{f}\boldsymbol{log}\left[1 - \frac{1}{\eta^{f}}\right] + \boldsymbol{m}_{it}^{f} - \eta^{f}\boldsymbol{a}_{it} + \left(\psi_1 - \eta^{f}\psi_2\right)\boldsymbol{e}_t. \tag{4}$$

Accordingly, using (1), (3) and (4), the period t operating profits $\pi(m_{it}^{f},\ a_{it},\ e_t)$ from exporting may be written as:

$$\pi(\boldsymbol{m}_{it}^{f},\ \boldsymbol{a}_{it},\ \boldsymbol{e}_t) = \left(\frac{1}{\eta^{f}-1}\right)\exp\left[\eta^{f}\log\left(1-\frac{1}{\eta^{f}}\right) + \boldsymbol{m}_{it}^{f} + (1-\eta^{f})\boldsymbol{a}_{it} + \left(\psi_1 + (1-\eta^{f})\psi_2\right)\boldsymbol{e}_t\right] . \tag{5}$$

These profits are depicted graphically as the shaded area in figure 1 below, where $P^{h}+P^{h}{'}(Q^{h})$ and $P^{f}+P^{f}{'}(Q^{f})$ are marginal revenue in the home and foreign market, respectively. The foreign demand schedule approaches the vertical axis below the domestic demand schedule because transport costs and foreign tariffs reduce revenue per unit sale in foreign markets. (That is, $\psi_1\boldsymbol{e}_t + \boldsymbol{m}_{it}^{f}$ is typically less than $\boldsymbol{m}_{it}^{h}$.)

Figure 1: Gross Operating Profits from Exports



2.2 Dynamic optimization

If there were no sunk costs associated with becoming an exporter, firms would sell abroad whenever their marginal costs were low enough, or foreign demand shocks were high enough to generate operating profits in excess of the fixed costs, Γ_F , associated with maintaining a presence in foreign markets: $\pi(\boldsymbol{m}_{it}^{f}, \boldsymbol{a}_{it}, \boldsymbol{e}_t) > \Gamma_F$. (Fixed costs include dealing with customs, minimum freight and insurance charges, and monitoring foreign standards.) But those firms that are not already exporting must pay an additional start-up cost, Γ_S , to establish distribution channels, learn bureaucratic procedures, and adapt their products and packaging for foreign markets. Defining the binary variable y_t to take a value of one during periods when the firm exports and zero otherwise, and assuming that sunk costs are borne in the first year of exporting, net current profits from exporting, $u(\cdot)$, may be written as:

$$\boldsymbol{u}(\cdot) = \begin{cases} \pi^f(\boldsymbol{m}_{it}^f, \boldsymbol{a}_{it}, \boldsymbol{e}_t) - \Gamma_F + \epsilon_{1it} & \textbf{\textit{if }} y_{it} = 1 \textbf{\textit{ and }} y_{it-1} = 1 \\ \pi^f(\boldsymbol{m}_{it}^f, \boldsymbol{a}_{it}, \boldsymbol{e}_t) - \Gamma_F - \Gamma_S + \epsilon_{2it} & \textbf{\textit{if }} y_{it} = 1 \textbf{\textit{ and }} y_{it-1} = 0 \\ 0 & \textbf{\textit{if }} y_{it} = 0 \end{cases} \tag{6}$$

Here we have added the noise components ε_{1it} and ε_{2it} to allow for unobserved transitory shocks that induce managers to deviate from expected profit-maximizing behavior. These variables also pick up transitory variation in the fixed and sunk costs of exporting.

One might object that (6) oversimplifies behavior because it implies that firms completely lose their investment in start-up costs if they are absent from the export market for a single year. However, earlier studies suggest that these investments depreciate very quickly, and that firms which most recently exported two years ago must pay nearly as much to re-enter foreign markets as firms that never exported (Roberts and Tybout, 1997; Bernard and Jensen, 1996). In light of these findings, and given that more general representations make structural estimation intractable, we consider (6) to be a reasonable abstraction.

To characterize the decision to export, let us collapse our notation by writing the exogenous random state variables as the vector $\mathrm{x}_{it} = (m_{it}^f,\ a_{it},\ e_t)$ and the complete parameter vector of interest as $\theta = (\Lambda, \Gamma_F, \Gamma_S, \sigma_{\varepsilon 1}, \sigma_{\varepsilon 2})$, where $\sigma_{\varepsilon 1}^2 = var(\varepsilon_{1it})$, $\sigma_{\varepsilon 2}^2 = var(\varepsilon_{2it})$, and Λ is the vector of parameters that govern the evolution of $\pi(\boldsymbol{x}_{it})$. Then, suppressing *i* subscripts, at time *t* the decision problem of a plant is to choose the sequence of decision rules $\boldsymbol{Y} = \{\boldsymbol{y}_\tau = \boldsymbol{f}_\tau(\boldsymbol{x}_\tau, \boldsymbol{y}_{\tau-1}, \epsilon_\tau, \theta)\}_{\tau=t}^{\infty}$ to maximize the expected discounted sum of net profits:

$$\max_{\boldsymbol{Y}} \quad \boldsymbol{E}_t\{\sum_{\tau=t}^{\infty} \delta^\tau \boldsymbol{u}(\boldsymbol{x}_\tau, \boldsymbol{y}_{\tau-1}, \boldsymbol{y}_\tau, \epsilon_\tau, \theta)\} \tag{7}$$

where E_t is the expectation operator conditioned on information available at time *t*, and δ is the discount rate, $0 < \delta < 1$.

This expression is the value function $\boldsymbol{V}(\boldsymbol{x}_t,\ \boldsymbol{y}_{t-1},\ \epsilon_t,\ \theta)$, which is the unique solution to the Bellman equation:

$$\boldsymbol{V}(\boldsymbol{x}_t,\ \boldsymbol{y}_{t-1},\ \epsilon_t,\ \theta) = \max_{\boldsymbol{y}_t \in \{0,1\}} [\boldsymbol{u}(\boldsymbol{x}_t, \boldsymbol{y}_{t-1}, \boldsymbol{y}_t, \epsilon_t, \theta) + \delta \boldsymbol{EV}(\boldsymbol{x}_t, \boldsymbol{y}_t, \epsilon_t, \theta)], \tag{8}$$

where

$$\boldsymbol{EV}(\boldsymbol{x}_t, \boldsymbol{y}_t, \epsilon_t, \theta) = \int_{\boldsymbol{x}_{t+1}} \int_{\epsilon_{t+1}} \boldsymbol{V}(\boldsymbol{x}_{t+1}, \boldsymbol{y}_t, \epsilon_{t+1}, \theta) \boldsymbol{dF}(\boldsymbol{x}_{t+1}, \epsilon_{t+1} | \boldsymbol{x}_t, \epsilon_t). \tag{9}$$

Under quite weak regularity conditions (see Bhattacharya and Majumdar (1989)), there exists a unique $V(\cdot)$ that solves (8). Moreover the optimal policy is stationary, i.e., $Y = \{f, f, f, ..\}$ and

$$\boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{y}_{t-1}, \epsilon_t, \theta) = \operatorname*{argmax}_{\boldsymbol{y}_t \in \{0,1\}} \{\boldsymbol{u}(\boldsymbol{x}_t, \boldsymbol{y}_{t-1}, \boldsymbol{y}_t, \epsilon_t, \theta) + \delta \boldsymbol{EV}(\boldsymbol{x}_t, \boldsymbol{y}_t, \epsilon_t, \theta)\}. \tag{10}$$

At any point in time, the i^{th} firm's behavior is determined by its own vector of realizations on the state variables $(\boldsymbol{x}_{it}, \boldsymbol{y}_{it-1}, \epsilon_{it})$. Using this information, one can calculate *total* export volumes by determining which firms are exporting (from equation 10) and how much each exporter supplies (from equation 4). One can also simulate counterfactual export trajectories by changing the characteristics of the exogenous $\boldsymbol{x}_{it}$ processes or by changing threshold costs. However, each of these exercises requires knowledge of the parameter vector θ. We now turn our attention to estimating these unknowns.

III. Econometric Analysis

Because exporting decisions are discrete, there are no closed-form solutions or first-order conditions for (10). Hence, to estimate the parameters that characterize dynamic behavior $(\Gamma_F, \Gamma_S, \sigma_{\varepsilon 1}, \sigma_{\varepsilon 2})$, one must solve the dynamic programming model numerically for each candidate vector and choose the one that yields the best match between predicted and observed y_t trajectories according to some metric. In our case we maximize the value of a simulated likelihood function.

Our estimation proceeds in two stages. First, before solving the dynamic programming model, we estimate the parameters Λ that govern the evolution of $\pi(\boldsymbol{x}_{it})$. Then, using the results we search over Γ_S, Γ_F, $\sigma_{\varepsilon 1}$ and $\sigma_{\varepsilon 2}$ values in the dynamic optimization problem. This approach limits the dimensionality of the maximization problem at the stage which is computationally intensive.[7]

To pursue this two-stage strategy, we assume that innovations in the state variables $\boldsymbol{x}_{it}$ are

[7]For details on the computational burden of such models, see Rust (1995).

independent of the optimization errors $\varepsilon_{it} = (\varepsilon_{1it}, \varepsilon_{2it})$. This assumption, which seems reasonable, allows us to estimate Λ while ignoring ε_{it}. Further, if we add the assumption that ε_{it} is serially uncorrelated, independence considerably simplifies computations in the second stage by making $EV(x_t, y_t, \varepsilon_t, \theta) = EV(x_t, y_t, \theta)$.[8]

3.1 First Stage Estimation

To estimate Λ, we first must be specific about the process that generates x_{it}. Let us begin with the simple assumption that each element of x_{it} follows a trend stationary AR1 process:

$$\boldsymbol{a}_{it} = \lambda_0 + \lambda_1 \boldsymbol{a}_{it-1} + \lambda_2 \boldsymbol{t} + \nu_{it} \tag{11a}$$

$$\boldsymbol{m}^{f}_{it} = \rho_0 + \rho_1 \boldsymbol{m}^{f}_{it-1} + \rho_2 \boldsymbol{t} + \omega_{it} \tag{11b}$$

$$\boldsymbol{e}_t = \gamma_0 + \gamma_1 \boldsymbol{e}_{t-1} + \gamma_2 \boldsymbol{t} + \xi_t \tag{11c}$$

where all three disturbances are white noise and orthogonal to one another.

If all three series were observable for each plant it would be a simple matter to test these assumptions, estimate the parameters of the three processes, and establish whether each variable is covariance stationary. Unfortunately, although the real exchange rate is readily observed, the other variables can only be crudely approximated with the available data. We do not observe plant-specific foreign demand shifters and marginal costs, only nominal export revenues and a nominal measure total variable costs. Good plant-specific price deflators are unavailable, so the best we can do is construct real export revenue measures and price-cost mark-up measures. Hence, we

---

[8] See Rust (1988) for the proof.

must recast the model in terms of these observables and extract the processes described by (11a) and (11b) indirectly.

To begin, note that our assumption of a constant, common demand elasticity for each firm implies that export profits are proportional to export revenues, and the factor of proportionality, $(\eta^f)^{-1}$, bears a simple relationship to total variable cost over total revenue: $(\eta^f)^{-1} = 1 - \frac{\boldsymbol{TVC}}{\boldsymbol{TR}}$. Hence, assuming that true variable costs are measured up to a zero-mean error by labor and intermediate input costs, we can easily estimate $(\eta^f)^{-1}$ as an average of the right-hand-side of this expression across observations.

After converting export revenues to export profits, we use equation (5) and the relationships $\boldsymbol{m}_{it}^f = (1-\boldsymbol{L}\rho_1)^{-1}(\rho_0 + \rho_2\boldsymbol{t} + \omega_{it})$ and $\boldsymbol{a}_{it} = (1-\boldsymbol{L}\lambda_1)^{-1}(\lambda_0 + \lambda_2\boldsymbol{t} + \nu_{it})$, to write the log of export profits as an ARMA(2,1) process, conditioned on $\boldsymbol{e}_{it}[\psi_1 - (1-\eta^f\psi_2)](1-\boldsymbol{L}\lambda_1)(1-\boldsymbol{L}\gamma_1)$ and a trend term:

$$
\begin{aligned}
\ln(\pi_{it}) = \beta_0 + \beta_1 \ln(\pi_{it-1}) + \beta_2 \ln(\pi_{it-2}) \\
+ \beta_3[e_t - \beta_1 e_{t-1} - \beta_2 e_{t-2}] + \beta_4 t + \xi_{it} + \mu\xi_{it-1}
\end{aligned}
\tag{12}
$$

Here $\beta_0 = [(\eta^f-1)\ln(\eta^f-1) - \eta^f\ln(\eta^f)](1-\rho_1)(1-\lambda_1) + \lambda_0(1-\eta^f)(1-\rho_1) + \rho_0(1-\lambda_1) + \rho_2\lambda_1 + \rho_1(1-\eta^f)\lambda_2$, $\beta_1 = \lambda_1 + \gamma_1$, $\beta_2 = -\lambda_1\cdot\gamma_1$, $\beta_3 = \psi_1 + (1-\eta^f)\psi_2$, $\beta_4 = \rho_2\cdot(1-\lambda_1) + (1-\eta^f)\cdot\lambda_2\cdot(1-\rho_1)$, and the MA(1) disturbance term has properties defined by the variances of ω_{it} and ν_{it}. Note that ρ_0 and λ_0 are not identified by the parameters of equation (12), so the same ARMA(2,1) process can be expressed by setting ρ_0 and λ_0 to zero and treating β_0 as an unconstrained parameter. Further, since β_4 absorbs the net effect of all trends on export profits, one can re-cast profits as a function of a deterministic trend and the de-trended vector:

$\tilde{x}_{it} = [(1-\eta^f)(a_{it} - \lambda_2 t),\ (m_{it}^f - \rho_2 t),\ (e_t - \gamma_2 t)] = [\tilde{x}_{it}^1, \tilde{x}_{it}^2, \tilde{x}_t^3]$. We will do so to facilitate stage 2 calculations.

There are several ways to estimate equation (12). One is by using the generalized method of moments (GMM) estimator described by Arellano and Bond (1991). This approach allows us to deal with correlation between lagged export profits and ξ_{it-1}, and it provides a convenient way to test our assumptions about the error structure and stationarity. However, non-linear parameter constraints cannot easily be handled with this estimator, so once we are satisfied that the presumed error structure is reasonable, we shall re-estimate (12) using maximum likelihood to impose the relevant constraints. The ML estimator is described in appendix 1.

Estimation of equation (12) is complicated by the fact that we only observe export revenues for exporters, and firms self-select into the export market partly on the basis of innovations in the state variables. To control for the truncation bias that this self-selection creates, we include a Mills ratio for purposes of estimation. The Mills ratio is based on a simple reduced form Probit equation that explains the probability of firms exporting as a function of strictly exogenous firm characteristics: location, business type, and initial capital stock.

Finally, our characterization of firm behavior implies that each manager knows the set of current realizations on $\tilde{x}_{it}$. But we do not observe $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ ourselves. Accordingly, we use our MLE estimates to impute Kalman-smoothed values of $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$, period by period, for each exporting plant.(Appendix 1 provides details.) Then we impute values of $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ for firms that are *not* currently exporting by regressing $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ on observable plant characteristics in our sub-sample of exporters, and we use the resulting parameters to impute $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ values for non-exporting firms. The set of observable plant characteristics includes lagged values of domestic sales, an average variable cost proxy and real capital stocks, as well as location dummies, and our Mills ratio, *M*. The logic of this approach is that non-exporters anticipate foreign demand, should they enter foreign markets, by looking at the demand that similar firms have encountered upon exporting. Clearly the imputed values of $(\tilde{x}_{it}^{1}, \tilde{x}_{it}^{2})$ will be noisy measures of the values that firms perceive, and ε_{2it} in equation (6) will include the effects of the discrepancy. We therefore expect that $\mathrm{var}(\varepsilon_{2it}) > \mathrm{var}(\varepsilon_{1it})$.

3.2 Second Stage Estimation

The remaining parameters, $(\Gamma_F,\ \Gamma_S,\ \sigma_{\varepsilon 1},\ \sigma_{\varepsilon 2})$, are estimated by maximizing the sample likelihood function:

$$\boldsymbol{L} = \prod_{i=1}^{N} \prod_{t=1}^{T} \boldsymbol{P}(\boldsymbol{y}_{it}=0|\boldsymbol{y}_{it-1}=0)^{(1-\boldsymbol{d}_{it})(1-\boldsymbol{d}_{it-1})} \boldsymbol{P}(\boldsymbol{y}_{it}=0|\boldsymbol{y}_{it-1}=1)^{(1-\boldsymbol{d}_{it})(\boldsymbol{d}_{it-1})} \tag{13}$$

$$\cdot \boldsymbol{P}(\boldsymbol{y}_{it}=1\,|\,\boldsymbol{y}_{it-1}=1)^{\boldsymbol{d}_{it}\boldsymbol{d}_{it-1}}\ \boldsymbol{P}(\boldsymbol{y}_{it}=1\,|\,\boldsymbol{y}_{it-1}=0)^{\boldsymbol{d}_{it}(1-\boldsymbol{d}_{it-1})}\,.$$

where d_{it} is a dummy that takes a value of 1 if the i^{th} firm exports in year t, and a value of zero if it does not. All probabilities are conditioned on the vector of state variables observed in period *t-1*.

To calculate the probabilities in (13), we must find the decision rule (10) for each firm at each point in time. In turn, this means we must specify the distribution of the unobserved vector ε_t, and use our information on transition probabilities for the vector of state variables from stage 1. Suppressing plant (*i*) subscripts, let V_{0t}, V_{10t} and V_{11t} respectively be the value from not exporting in period *t*, from beginning to export in period *t* after not exporting in period *t*-1, and from continuing to export in period t after exporting in period t - 1, each exclusive of transitory noise (ε_{it}):

$$\begin{aligned} \boldsymbol{V}_{11\boldsymbol{t}} &= \pi(\boldsymbol{x}_t) - \Gamma_{\boldsymbol{F}} + \delta \boldsymbol{EV}(\boldsymbol{x}_t, 1, \theta) \\ \boldsymbol{V}_{10\boldsymbol{t}} &= \pi(\boldsymbol{x}_t) - \Gamma_{\boldsymbol{F}} - \Gamma_{\boldsymbol{S}} + \delta \boldsymbol{EV}(\boldsymbol{x}_t, 1, \theta) \\ \boldsymbol{V}_{0\boldsymbol{t}} &= \delta \boldsymbol{EV}(\boldsymbol{x}_t, 0, \theta). \end{aligned} \tag{14}$$

Then the probabilities of observing the different exporting states are:

$$\begin{aligned} \boldsymbol{P}[\boldsymbol{y}_t{=}1|\boldsymbol{y}_{t-1}{=}0] &= \boldsymbol{P}[\boldsymbol{V}_{10}+\epsilon_2 > \boldsymbol{V}_0] = \Phi\left(\frac{\boldsymbol{V}_{10}-\boldsymbol{V}_0}{\sigma_{\epsilon 2}}\right) \\ \boldsymbol{P}[\boldsymbol{y}_t{=}1|\boldsymbol{y}_{t-1}{=}1] &= \boldsymbol{P}[\boldsymbol{V}_{11}+\epsilon_1 > \boldsymbol{V}_0] = \Phi\left(\frac{\boldsymbol{V}_{11}-\boldsymbol{V}_0}{\sigma_{\epsilon 1}}\right) \\ \boldsymbol{P}[\boldsymbol{y}_t{=}0|\boldsymbol{y}_{t-1}{=}0] &= \boldsymbol{P}[\boldsymbol{V}_{10}+\epsilon_2 > \boldsymbol{V}_0] = 1 - \Phi\left(\frac{\boldsymbol{V}_{10}-\boldsymbol{V}_0}{\sigma_{\epsilon 2}}\right) \\ \boldsymbol{P}[\boldsymbol{y}_t{=}0|\boldsymbol{y}_{t-1}{=}1] &= \boldsymbol{P}[\boldsymbol{V}_{11}+\epsilon_1 > \boldsymbol{V}_0] = 1 - \Phi\left(\frac{\boldsymbol{V}_{11}-\boldsymbol{V}_0}{\sigma_{\epsilon 1}}\right) \end{aligned} \tag{15}$$

Clearly, once V_{0t}, V_{10t} and V_{11t} are calculated up to the unknown parameters, maximum likelihood estimation becomes straightforward. However, these expressions are difficult to calculate because they involve the expected value of the period $t+1$ value function conditioned on period t information. We begin by writing this expectation as:

$$\boldsymbol{EV}(\boldsymbol{x}_t, \boldsymbol{y}_t, \theta) = \boldsymbol{y}_t \boldsymbol{E}_t \max(\boldsymbol{V}_{11t+1}+\epsilon_{1t+1}, \boldsymbol{V}_{0t+1}) + (1-\boldsymbol{y}_t)\boldsymbol{E}_t \max(\boldsymbol{V}_{10t+1}+\epsilon_{2t+1}, \boldsymbol{V}_{0t+1}) \tag{16}$$

where:

$$\boldsymbol{E_t}\max(\boldsymbol{V}_{11t+1}+\epsilon_{1t+1},\boldsymbol{V}_{0t+1}) = \boldsymbol{P}[\epsilon_{1t+1}>\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{11t+1}]\cdot[\boldsymbol{V}_{11t+1} + \boldsymbol{E}(\epsilon_{1t+1}|\epsilon_{1t+1}>\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{11t+1})]$$
$$+ \boldsymbol{P}[\epsilon_{1t+1}<\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{11t+1}]\cdot\boldsymbol{V}_{0t+1}$$

$$\boldsymbol{E_t}\max(\boldsymbol{V}_{10t+1}+\epsilon_{2t+1},\boldsymbol{V}_{0t+1}) = \boldsymbol{P}[\epsilon_{2t+1}>\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{10t+1}]\cdot[\boldsymbol{V}_{10t+1} + \boldsymbol{E}(\epsilon_{2t+1}|\epsilon_{2t+1}>\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{10t+1})]$$
$$+ \boldsymbol{P}[\epsilon_{2t+1}<\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{10t+1}]\cdot\boldsymbol{V}_{0t+1}$$

Then, under the assumption that (ε_{1t}, ε_{2t}) are jointly normally distributed, the conditional expectations above can be expressed as Mills ratios, and the probabilities (conditioned on x_{t+1}) can be obtained from the standard normal distribution function, $\Phi(\cdot)$:

(17*a*) $$\boldsymbol{E_t}\max(\boldsymbol{V}_{11t+1}+\epsilon_{1t+1},\boldsymbol{V}_{0t+1}) =$$
$$\int_{\boldsymbol{x}_{t+1}}[\Phi\left(\frac{\boldsymbol{V}_{11t+1}-\boldsymbol{V}_{0t+1}}{\sigma_{\epsilon 1}}\right)\cdot\left[\boldsymbol{V}_{11t+1}+\sigma_{\epsilon 1}\phi\left(\frac{\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{11t+1}}{\sigma_{\epsilon 1}}\right)\Big/\Phi\left(\frac{\boldsymbol{V}_{11t+1}-\boldsymbol{V}_{0t+1}}{\sigma_{\epsilon 1}}\right)\right]$$
$$+ \Phi\left(\frac{\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{11t+1}}{\sigma_{\epsilon 1}}\right)\cdot\boldsymbol{V}_{0t+1}]\,\boldsymbol{dF}(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$$

(17*b*) $$\boldsymbol{E_t}\max(\boldsymbol{V}_{10t+1}+\epsilon_{2t+1},\boldsymbol{V}_{0t+1}) =$$
$$\int_{\boldsymbol{x}_{t+1}}[\ \Phi\left(\frac{\boldsymbol{V}_{10t+1}-\boldsymbol{V}_{0t+1}}{\sigma_{\epsilon 1}}\right)\cdot\left[\boldsymbol{V}_{10t+1}+\sigma_{\epsilon 2}\phi\left(\frac{\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{10t+1}}{\sigma_{\epsilon 2}}\right)\Big/\Phi\left(\frac{\boldsymbol{V}_{10t+1}-\boldsymbol{V}_{0t+1}}{\sigma_{\epsilon 2}}\right)\right]$$
$$+ \Phi\left(\frac{\boldsymbol{V}_{0t+1}-\boldsymbol{V}_{10t+1}}{\sigma_{\epsilon 2}}\right)\cdot\boldsymbol{V}_{0t+1}]\,\boldsymbol{dF}(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$$

These expressions take care of integration over ε_{t+1}, but they still require multi-dimensional integration over x_{t+1}, and by equation (14), they still involve the unknown $EV(x_{t+1}, y_{t+1}, \theta)$. To deal with the latter problem we use the backward induction algorithm described by Rust (1995). This amounts to assuming that firms have a finite planning horizon of *H* years, so that *T* = *t* + *H* is the (distant) terminal period.[9] Then, in the terminal year there are no future periods to consider so $EV(x_T, y_T, \theta)$ is set to zero in (14), and each firm's exporting decision maximizes current payoffs, $u(x_T, \varepsilon_T, y_T, y_{T-1}|\theta)$. Accordingly, for period *T-1*, the expected value function is simply:

$$EV_{T-1}(x_{T-1}, y_{T-1}, \theta) = E_{T-1}\Big[\max_{y_T} u(x_T, \varepsilon_T, y_T, y_{T-1}|\theta)\Big]$$

where expectations are taken over both ε_T and x_T conditioned on x_{T-1}, as in (17). Once the expected value function for period *T-1* has been calculated for each realization on x_{T-1}, the expected value function for period *T-2* can be calculated for each realization on x_{T-2}:

$$EV_{T-2}(x_{T-2}, y_{T-2}, \theta) = E_{T-2} \max_{y_{T-1}}\Big[u(x_{T-1}, \varepsilon_{T-1}, y_{T-1}, y_{-2}|\theta) + \delta EV_{T-1}(x_{T-1}, y_{T-1}, \theta)\Big]$$

Clearly this calculation can be repeated, backing up one year at a time, until one reaches period *t*.

---

[9]So long as *H* is large and the discount rate is positive the resulting solution approximates the solution to the infinite horizon problem.

This generates the values needed for $\boldsymbol{V}_{10\boldsymbol{t}}$, $\boldsymbol{V}_{11\boldsymbol{t}}$, and $\boldsymbol{V}_{0\boldsymbol{t}}$ which in turn enter the likelihood function.

The final issue is how to take expected values over x_{t+1} realizations, given x_t. One approach is to discretize the vector x_t (Rust, 1987, Das, 1992). Then, since this vector is a stationary first-order process, transition probabilities from any realization in period *t* to any other in *t+1* can be summarized with a transition matrix, ***M***. For example, the probability of the k^{th} realization in period *T*, given that the j^{th} realization is observed in period *T-1*, is in row *k*, column *j* of ***M***.

The problem with this approach is that it involves a large number of calculations. With 3 state variables and, say, *r* different values per state variable, the dimension of ***M*** is 3*r* by 3*r*. With a reasonably long planning horizon (*T*), approximation errors compound, and it becomes necessary to use very large *r* values. This is the "curse of dimensionality" that made a supercomputer necessary for estimation of Das's (1992) model, which involved 5 state variables.

Recently Rust (1997) has developed an alternative approach that substantially reduces the computational burden. His technique has two key features. First, expected values of the value function are calculated using Monte Carlo integration. Second, the draws for the Monte Carlo calculations are done using a multi-grid algorithm. This yields a random Bellman operator.

We take advantage of Rust's (1996) insights in the present study. First we assume that the vector x_t is multivariate normal with all x_t realizations lying in the observed region of support. This ensures boundedness of x_t. Truncating the normal imposes very weak restrictions because almost the entire area under the normal curve lies in this range. Second, we make N uniform draws from within the bounds of x_t. In our case this implies N values of every V_{10}, V_{11}, and V_0. The multigrid approach means choosing small values of N (coarser grids) for periods in the more distant future, and larger values of N (finer grids) for periods close to the present. Finally, the transition probabilities in the discrete analog to $\boldsymbol{dF}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ are normalized to ensure that the random Bellman operator is a contraction mapping. In particular, for each $\boldsymbol{x}_{t-1}$, the imputed probabilities of moving to the various possible $\boldsymbol{x}_t$ values are scaled to sum to unity.

IV. Findings: The Colombian Chemicals Industry

4.1 Descriptive overview of the data

Although our framework should describe any industry in which exporting is potentially profitable for some firms, it will be easiest to identify parameters in those industries which have many exporters, and which exhibit substantial variation in the set of exporters over time. For these reasons, we choose to estimate the model using data on the Colombian chemicals industry for the period 1982 through 1991, which is summarized in table 1 below.

Note that the Colombian peso depreciated substantially in real terms during the sample period, and that chemical exports simultaneously grew. The expansion was partly due to an increase in the number of exporters, and partly due to increases in the magnitude of foreign sales at the typical exporting plant.[10] Colombian chemicals plants produced 34.92 (units?) worth of exports in 1991, of which 29.94 (units) came from plants that were exporting in 1984. So entry by new exporters contributed 5.02 out of the 27.02 expansion. Also, of the 62 plants that existed during the entire sample period, 18 exported in all ten years, 26 never exported, and 18 switched exporting status at least once. So, although there were a number of switches, the data exhibit

[10]The number of Chemical plants remains fixed at 62 during the sample because we have excluded producers who enter or exit to simplify the econometrics, so there is some potential for selectivity bias.

substantial persistence. This could be due to serial correlation in the plant-specific state variables, m_{it}^{f} and a_{it}, or it could be due to sunk entry costs, Γ_s, or some combination of both. Our estimates will shed light on the relative importance of these different forces.

Table 1: Colombian Chemical Producers, Exporters versus Non-exporters

Year	***Total Value of Exports***	***Number of Exporters***	***Number of Non-exporters***	***Number of Entrants***	***Number of Quitters***	***Real Exchange Rate***
1982	6.17	25	37	1	0	79.5
1983	8.49	27	35	3	1	80.5
1984	7.90	28	34	1	0	89.8
1985	11.78	24	38	2	6	102.2
1986	14.10	24	38	1	1	113.6
1987	15.40	23	39	1	2	113.7
1988	21.93	25	37	3	1	112.3
1989	20.63	27	35	2	0	115.2
1990	27.10	28	34	1	0	127.2
1991	34.92	28	34	0	0	121.1
Average	16.842	25.9	36.1	1.5	1.1	105.51

4.2 First Stage Estimation: Evolution of the State Variables

Our first task is to characterize the process generating export profits using the techniques described in section 3. Estimates of equation (12) obtained with the generalized method of moments are reported in the first column of table 2 below. Recall that time dummies control for the exchange rate and trend effects, which are common to all plants, and we include a Mills ratio to control for selectivity bias since export profits can only be observed for firms that are currently exporting.

Note first that our results are quite consistent with the stationary ARMA(2,1) process we posited. Using Arellano and Bond's (1991) tests for serial correlation in the residuals, the null of no first-order correlation is rejected with a p-value of .034, but the null of second order correlation is easily accepted (p-value .907). Further, adding a third lag on log profits to the equation does not significantly improve the fit.[11] Finally, note that the coefficients on lagged profits imply trend-stationarity, with roots of 0.941 and -0.751. An augmented Dickey-Fuller test confirms that we can reject the null hypothesis of a unit root with a p-value of .0001.[12] Of course, we cannot tell from our estimates which root corresponds to the cost process, $(1-\eta^f)\boldsymbol{a}_{it}$, and which to the demand-shifter process , $\boldsymbol{m}_{it}^f$.

Since the GMM results suggest that our assumptions about the process generating export profits are reasonable, we proceed to jointly estimate all the parameters of this process using maximum likelihood. The results are reported in the last column of table 2. The roots we obtain are quite similar to those obtained with GMM—in fact we easily accept the null that the MLE estimates are the same as those reported in the GMM column. But the MLE estimates have the advantage of imposing all relevant parameter constraints, and yielding values for $(\sigma_{\tilde{\nu}}^2, \sigma_{\omega}^2)$, the net trend effect in profits, $\tilde{\beta}_4$, and the elasticity of profits with respect to the exchange rate, β_3. (These latter two parameters are identified with only 10 years of data, but fortunately they turn out sensibly.) Finally, note that the Mills ratio has the expected sign although it is not highly

---

[11] The coefficient on the third lag was 0.032 with a standard error of 0.036. These results are not reported in Table 2.

[12] The asymptotic properties for this test are obtained by letting the number of plants go to infinity, so the standard errors from our GMM estimator are valid even under the null of a unit root, and no special critical values are needed. (See, for example, Breitung, 1995.)

significant. This implies that realization on the two latent variables are more favorable than average for firms that are currently exporting.

Using the MLE results, it is possible to calculate Kalman-smoothed realizations on the latent processes for all exporters, as discussed in section 3. Then expected values of these latent processes for non-exporters are constructed by regressing $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ on observable plant characteristics. This approach yields an excellent fit for the latent state variable with strong persistence (r^2 =.91), and a reasonable fit for the other variable (r^2 = .38).

Table 2: Stage 1 Parameter Estimates

(Asymtotic standard errors in parentheses)

	GMM [a]	MLE
β_0 (*intercept*)	1.196 (0.410)	3.121 (5.35)
β_1 (*log profits, t-1*)	0.195 (0.126)	0.410 (imputed)
β_2 (*log profits, t-2*)	0.701 (0.112)	0.520 (imputed)
β_3 (*log exchange rate*)	– [b]	1.428 (1.113)
β_4 (*trend*)	– [b]	0.071 (0.066)
β_5 (*Mills ratio*)	0.081 (0.150)	1.356 (0.732)
ρ_1	0.941 (imputed)	0.954 (0.294)
λ_1	-0.745 (imputed)	-0.544 (0.415)
$\text{var}(\tilde{\omega})$	--	0.708 (0.194)
$\text{var}(\nu)$	--	0.280 (0.371)
$\text{var}(\xi_1 + \mu\xi_2)$	0.596	--
log-likelihood	--	-243.73
1st order serial correlation test (*z*)	2.116	--
2nd order serial correlation test (*z*)	0.117	--

[a]The instrument set includes lagged values of export profits (two or more periods back), beginning of period capital stocks, location and 4-digit industry dummies, and time dummies.

[b] Year dummies included instead of these variables. (Note that the intercept is affected.) The $\chi^2(8)$ statistic for the null hypothesis that the year dummies are jointly insignificant is 32.28, which has a p-value of 0.00.

4.3 Second Stage Estimates: Sunk Costs and Fixed Costs

Using the parameters in table 2, we can use the techniques described in section 3.2 to calculate the value function (8) for each plant at each point in time, up to the parameter vector $(\Gamma_F,\ \Gamma_S,\ \sigma_{\varepsilon 1},\ \sigma_{\varepsilon 2})$. Hence, given any value for this vector, we can characterize the probability of the observed series of export market participation decisions for each producer using (16) and (17). Finally, searching over alternative values of the unknown parameter vector, we can find the set that maximizes the likelihood function (15). Because the likelihood function is not globally concave in the discount rate, δ, we obtain estimates at each δ value on the [0,1) interval, incrementing by .05. The results of this exercise are reported in table 3 below.

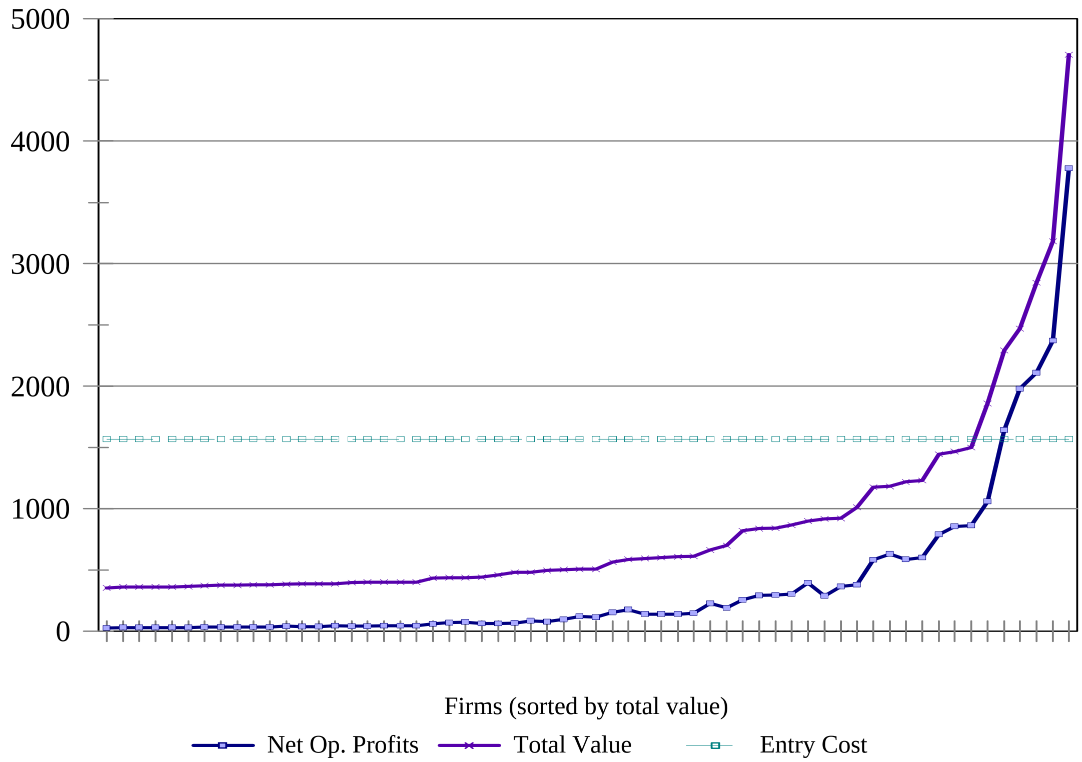
Note that the standard errors have not been corrected for the fact that the expected value function is itself stochastic, both because Rust's (1997) techniques have been used and because the parameters estimated in stage 1 are treated as given in stage 2. Experimentation with repeated grid draws revealed that when the grid includes at least 1,000 evaluation points the estimates of $(\Gamma_F,\ \Gamma_S,\ \sigma_{\varepsilon 1},\ \sigma_{\varepsilon 2})$ are fairly stable, but the estimates are sensitive to the way in which latent state variables are imputed for non-exporters contemplating entry.

Table 3: Sunk and Fixed Cost Estimates

Parameter	***Estimate and t-ratio***
Γ_s	1569 (2.12)
Γ_F	-26 (0.72)
$\sigma_{\epsilon 1}$	655 (2.02)
$\sigma_{\epsilon 2}$	521 (1.80)
log-likelihood	-146.3

Most fundamentally, the results indicate that even after controlling for serial correlation in the state variables, substantial persistence remains due to sunk start-up costs. Hence, as Roberts and Tybout (1997) find in their reduced-form estimates, the Dixit/Baldwin/Krugman hysteresis framework appears to have empirical relevance. Fixed are negative, but very small and insignificant. This could be a consequence of our simple functional forms, but it might also reflect export subsidy programs that were in place during the sample period. Finally, the variances in the optimization errors are substantial, but they are not estimated with much accuracy.

Figure 2:
Profits, Value and Entry Costs



The option value of avoiding entry costs next period is determined by expected operating profits from exporting and the entry costs themselves. (It can be no greater than the latter, of course.) This value, *EV*, may be inferred for each firm in figure 2 as the difference between the total value of being an exporter and current operating profits. The horizontal line represents sunk entry costs, Γ_s so non-exporters for whom the sum of operating profits and the option value is below this line do best to stay out. Incumbents, for their part, should stay in whenever the total value from doing so is positive.

Figure 3:
Export Probabilities and History

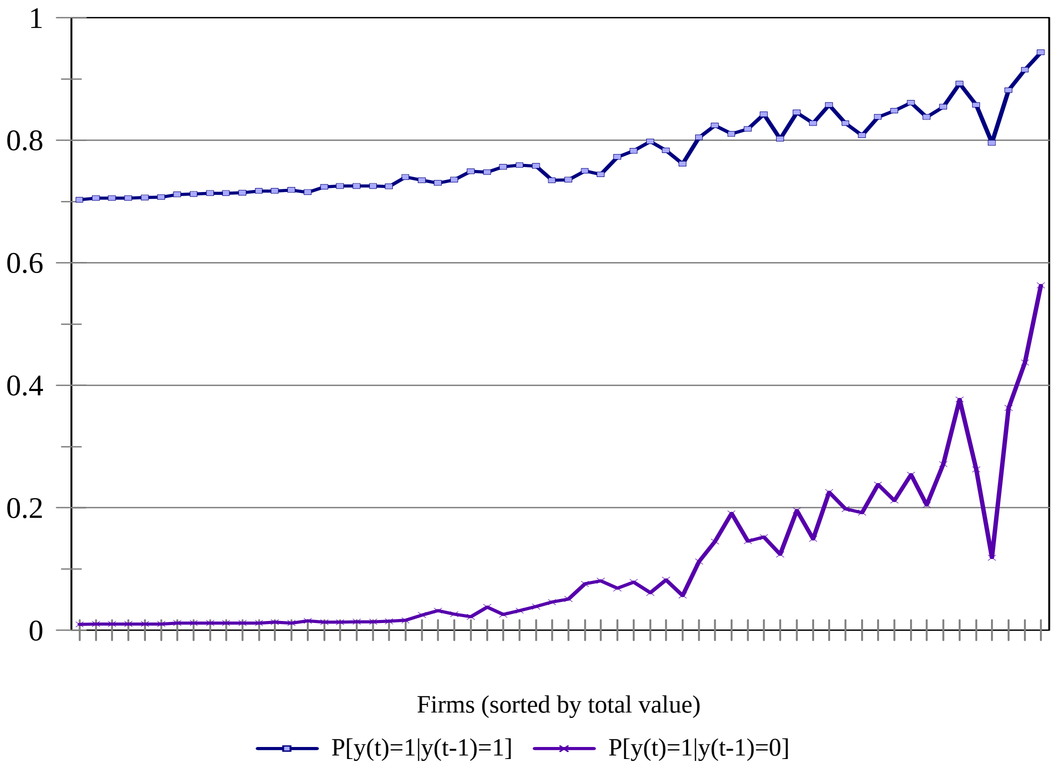


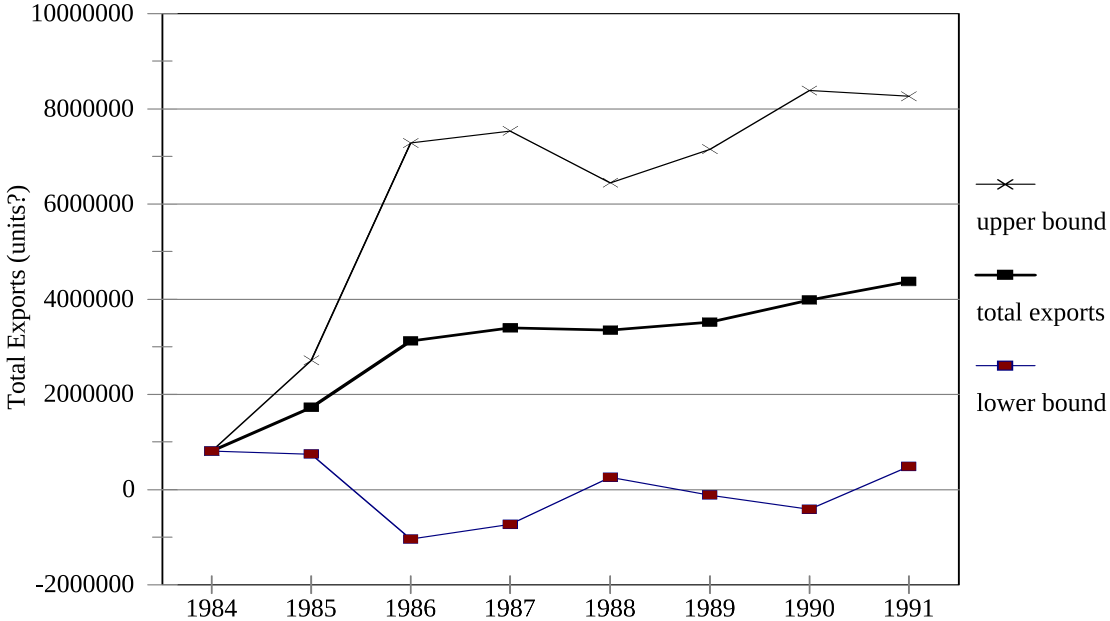
Figure 3 shows the effect of incumbency on the probability of continuing to export. As Roberts and Tybout (1997) found, those firms that have already paid the sunk start-up costs of breaking in are up to 70 percent more likely to continue exporting the next period than otherwise comparable firms that have not. Interestingly, if there were no errors in our profit functions, our estimates would imply that all firms in the sample would be exporters if they were already exporting in the previous period. This is because when $\varepsilon_{it} = 0$, net operating profits from exporting are positive for everyone. On the other hand, most firms anticipate a total value from exporting that is less than the entry cost (figure 2). These firms do best to stay out of the export market if they haven't already entered. In sum, our estimates imply a substantial hysteresis band and a very large role for history, as shown by figure 3. However, firms with small expected profits would be small-scale exporters if they were to enter, so one cannot conclude that a wide hysteresis band implies large *volume* effects of sunk costs.

3.4 Model Fit

To check the model's fit, we simulate optimal behavior using our parameter estimates, assuming that firms realize the observed values of the state variables in the base year (1984), but thereafter these variables evolved according to simulated realizations, based on the processes implied by the estimated Λ vector. Simulated realizations on x_{it} should not be expected to match the observed realizations, of course, but with repeated simulations they should at least bracket the range of plausible outcomes implied by our model, and this range should not be inconsistent with the particular outcomes observed in our data.

The results of this exercise are reported in Figures 3-5. These figures represent averages over 30 different sets of simulations. In each simulation we have randomly drawn x_{it} trajectories for each firm, based on the AR(1) processes estimated for these state variables. In figures 4 and 5, the upper and lower bounds are standard deviations in predicted totals across the 30 simulations. The simulations begin from observed values of the state variables in 1984 and use randomly drawn innovations in these variables for the remaining 7 years.

Figure 4:
Average Total Predicted Exports

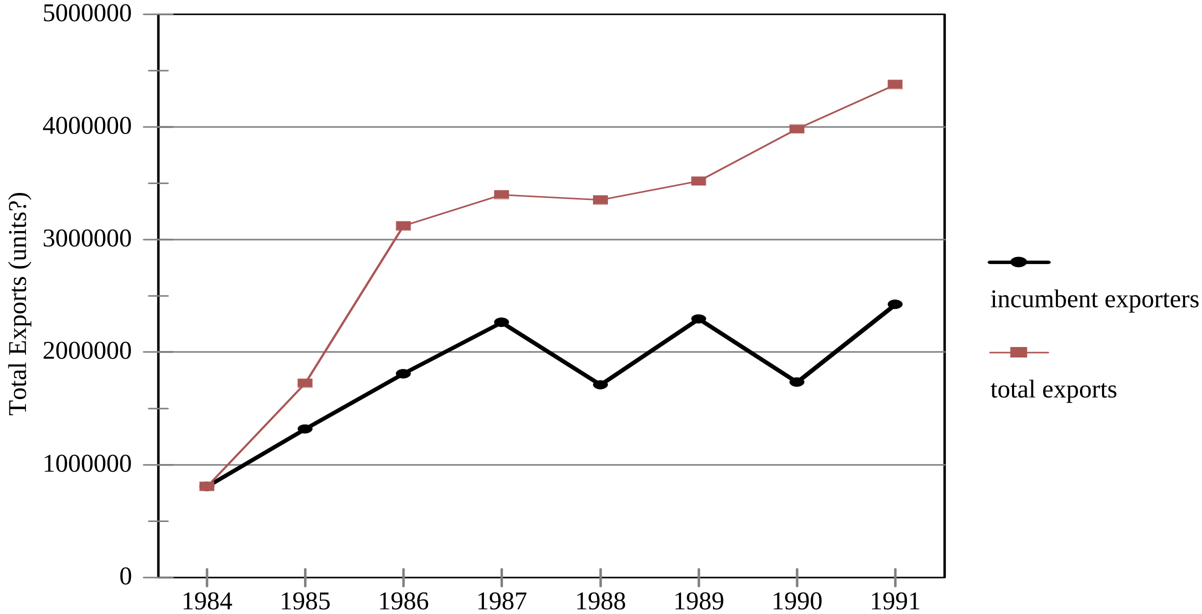


Note first that after one year, there is substantial variation in the predicted outcomes from simulation to simulation. This reflect the importance of random shocks in the exchange rate, the profit function, and the firm-specific foreign demand shifters. The unfortunate implication is that even when working with micro data sets like ours, a large fraction of the variation in exports is unpredictable. Notice, however, that *on average* the upward drift in total predicted export volumes matches the observed data rather well.

Consistent with the actual data, the amount of export growth coming from new exporters is substantial (figure 5). In fact, the role of new exporters is greater in our simulations than that which was actually observed. On average, total simulated exports grow from 7.9 to 43.7 (units?), and 19.4 (units?) of this expansion comes from firms that were not exporting in 1984. In the sample data, total exports grew from 7.9 million in 1984 to 34.9 in 1989, and of which came from firms not exporting in 1984.

Figure 5: Predicted Exports

Total versus from Incumbent Exporters



There is one sense in which our simulated trajectories fail to match the actual data. While the actual data show a dip and then a recovery in the number of exporting plants, our simulations imply a continued drop in the number of firms (Figure 6). This traces to overestimation of exit rates after 1985 (Figure 7). Table 1 shows that the number of plants exiting averaged 1.1 over the sample period, while our simulations predict exit rates in excess of 2. (Simulated entry rates, on the other hand, match the actual data quite well.) The excessive exit does not translate into under-prediction of gross export volumes because those firms that drop out are relatively small scale.

Figure 6:
Average Predicted Number of Exporters

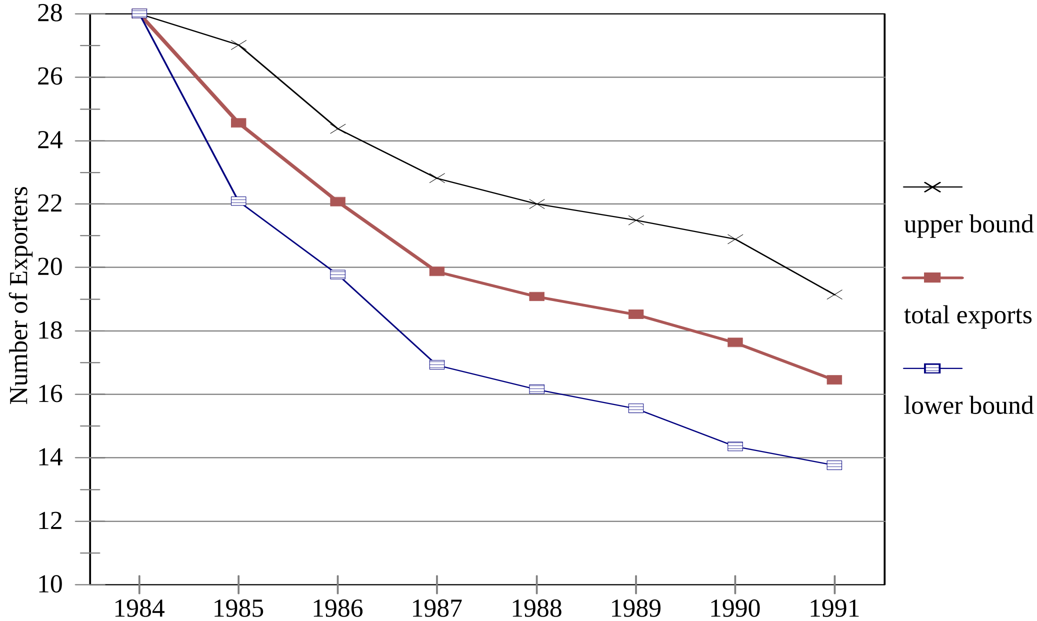
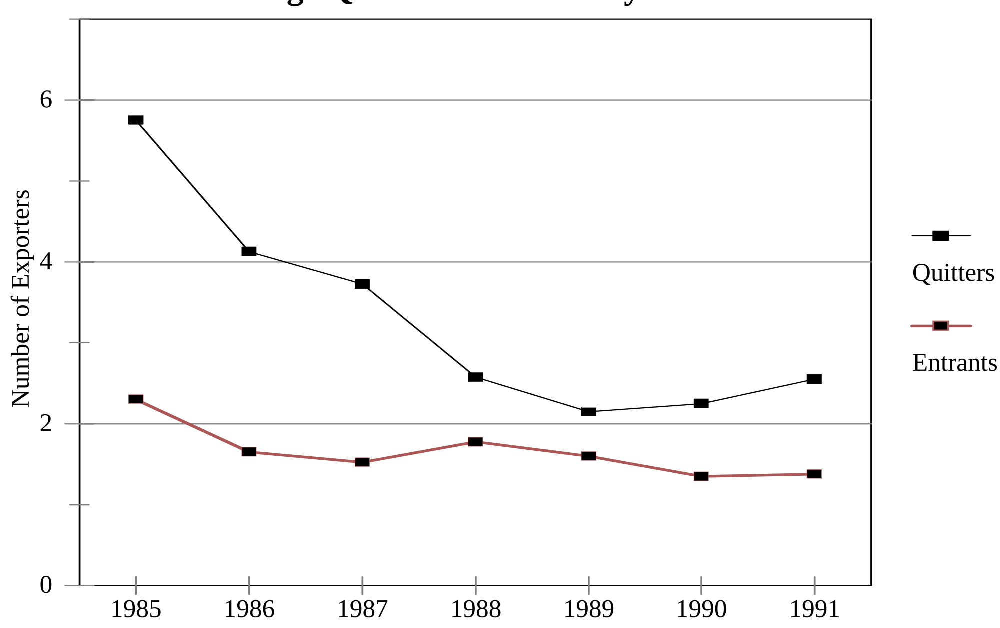


Figure 7:
Average Quit Rates and Entry Rates



4.5 Counterfactual Simulations (<u>this section under revision</u>)

The fundamental advantage of structural estimation over the reduced-form approach of Roberts and Tybout (1997) is that it allows us to perform counterfactual exercises. Whereas Roberts and Tybout's estimates are based on the presumption that all stochastic processes are stable, the present framework allows us to address the question of how different exchange rate *regimes*—that is different stochastic processes for *e*— might affect the responsiveness of aggregate export flows to given exchange rate movements. It also allows to determine how much persistence in exports is due to sunk costs, and to properly quantify the irreversibility effects that have been stressed in theoretical work Finally, it allows us to examine the effects of firm heterogeneity on aggregate responsiveness.

We will organize our analysis around three counterfactual exercises:.

- credible versus non-credible change in the exchange rate regime
- reductions in sunk costs
- reduction in firm heterogeneity

(Results forthcoming)

References

Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58, pp. 277-297.

Baldwin, Richard E. and Paul Krugman. 1989. "Persistent Trade Effects of Large Exchange Rate Changes." *Quarterly Journal of Economics* 104(4), 635-54.

Bhattacharya, R.N. and M. Majumdar. 1989. "Controlled Semi-Markov Models -- The Discounted Case," *Journal of Statistical Planning and Inference*, 21, 365-381.

Chow, Gregory. 1983. *Econometrics*. New York: McGraw-Hill.

Clerides, S., S. Lach, and J. Tybout.1998. "Is Learning by Exporting Important? Micro Dynamic Evidence from Colombia, Mexico and Morocco," *Quarterly Journal of Economics,* August, pp. 903-947.

Das, S. 1992. "A Micro-Econometric Model of Capital Utilization and Retirement: The Case of the U.S. Cement Industry," *Review of Economic Studies*, 59, 277-297.

Dixit, Avinash K. 1989. "Entry and Exit Decisions Under Uncertainty." *Journal of Political Economy* 97(3), 620-38.

Krugman, Paul R. 1989. *Exchange Rate Instability*. Cambridge, MA: The MIT Press.

Roberts, Mark J. and James R. Tybout. 1997a. "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Cost," *American Economic Review,* September, pp. 545-564.

Roberts, Mark and James Tybout. 1997b. *What Makes Exports Boom?* The World Bank. Directions in Development Monograph.

Roberts, Mark, Theresa.A. Sullivan and James.R. Tybout. 1995. "Micro Foundations of Export Booms," Pennsylvania State University, Department of Economics Working Paper.

Rust, John. 1997. "Using Randomization to Break the Curse of Dimensionality," *Econometrica* 65(3), May, pp. 487-516.

Rust, John. 1988. "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM Journal of Control and Optimization*, 26(5), 1006-1023.

Rust, John. 1995. "Numerical Dynamic Programming in Economics," in *Handbook of Computational Economics,* edited by H. Amman, D. Kendrik and J. Rust.

Appendix 1: Imputing $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$ realizations from profit trajectories

To impute values of $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$, firm by firm, we first re-state the profit function in terms of these de-trended variables, a trend term, and the exchange rate. Multiplying both sides of (12) by $(1-L\rho_1)^{-1}(1-L\lambda_1)^{-1}$ we obtain:

$$\ln(\pi_{it}) = \tilde{\beta}_0 + \beta_3 e_t + \tilde{\beta}_4 t + (\tilde{x}_{it}^1 + \tilde{x}_{it}^2), \qquad (12')$$

where

$$\begin{bmatrix} \tilde{x}_{it}^1 \\ \tilde{x}_{it}^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \rho_1 \end{bmatrix} \begin{bmatrix} \tilde{x}_{it-1}^1 \\ \tilde{x}_{it-1}^2 \end{bmatrix} + \begin{bmatrix} \tilde{v}_{it} \\ \omega_{it} \end{bmatrix} = M\tilde{x}_{it-1} + \upsilon_{it} \text{ , and}$$

$$\text{var}\begin{bmatrix} \tilde{v}_{it} \\ \omega_{it} \end{bmatrix} = \begin{bmatrix} (1-\eta^f)^2 \sigma_v^2 & 0 \\ 0 & \sigma_\omega^2 \end{bmatrix} = \begin{bmatrix} \sigma_{\tilde{v}}^2 & 0 \\ 0 & \sigma_\omega^2 \end{bmatrix}$$

Then, presuming that each firm begins with a random draw from the steady state distribution of $(\tilde{x}_{it}^1, \tilde{x}_{it}^2)$, $E[\tilde{x}_{it}^1 + \tilde{x}_{it}^2] = 0$ and $\text{cov}(\tilde{x}_{it}^1 + \tilde{x}_{it}^2,\ \tilde{x}_{it-k}^1 + \tilde{x}_{it-k}^2) = \ell' M^k \Gamma_0 \ell$, where

$$\ell = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \Gamma_0 = \begin{bmatrix} \dfrac{\sigma_{\tilde{v}}^2}{1-\lambda_1^2} & 0 \\ 0 & \dfrac{\sigma_\omega^2}{1-\rho_1^2} \end{bmatrix}$$

(e..g., Chow, 1983). Using these relationships, equation (12') can be estimated with maximum likelihood. This yields estimates of (λ_1, ρ_1), $(\sigma_{\tilde{v}}^2, \sigma_\omega^2)$, β_3 and β_4. As with our GMM estimates, we include a Mills ratio in this equation to control for selection effects.

Finally, using these estimates we impute Kalman-smoothed values of $\tilde{x}_{it} = (\tilde{x}_{it}^1, \tilde{x}_{it}^2)$, period by period, for each exporting plant (e.g., Chow, 1983). By repeated substitution,

$$\begin{aligned}\tilde{x}_{it} &= M\tilde{x}_{it-1} + \upsilon_{it} = M^2\tilde{x}_{it-2} + \upsilon_{it} + M\upsilon_{it-1} \\ &= M^{t-j}\tilde{x}_{ij} + M^{t-j-1}\upsilon_{ij+1} + \ldots + M\upsilon_{it-1} + \upsilon_{it}, \qquad 1 \le j < t.\end{aligned}$$

Similarly,

$$\begin{aligned}\tilde{x}_{it+s} &= Mx_{it+s-1} + \upsilon_{it+s} = M^2\tilde{x}_{it+s-2} + M\upsilon_{it+s-1} + \upsilon_{it+s} \\ &= M^s\tilde{x}_{it} + M^{s-1}\upsilon_{it+1} + \ldots + M\upsilon_{it+s-1} + \upsilon_{it+s}, \qquad s > 0.\end{aligned}$$

So the year t realization on $\tilde{x}_{it}$ satisfies the following system of t+s equations:

$$\begin{bmatrix} \ln(\pi^*_{i1}) \\ \ln(\pi^*_{i2}) \\ \cdot \\ \cdot \\ \ln(\pi^*_{it-1}) \\ \ln(\pi^*_{it}) \\ \ln(\pi^*_{it+1}) \\ \cdot \\ \cdot \\ \ln(\pi^*_{it+s-1}) \\ \ln(\pi^*_{it+s}) \end{bmatrix} = \begin{bmatrix} \ell M^{-t+1} \\ \ell M^{-t+2} \\ \cdot \\ \cdot \\ \ell M^{-1} \\ \ell \\ \ell M^{1} \\ \cdot \\ \cdot \\ \ell M^{s-1} \\ \ell M^{s} \end{bmatrix} \cdot \tilde{x}_{it} + \begin{bmatrix} -\ell M^{-1} & -\ell M^{-2} & \cdot & \cdot & -\ell M^{-t+1} & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & -\ell M^{-1} & & & -\ell M^{-t+2} & \cdot & & & & & \cdot \\ 0 & \cdot & \cdot & & \cdot & \cdot & & & \cdot & & \cdot \\ \cdot & & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & -\ell M^{-1} & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & \cdot & \ell & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & \cdot & \ell M & \ell & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & & & & & \ell & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & \ell M^{s-1} & \ell M^{s-2} & \cdot & \cdot & \ell M & \ell \end{bmatrix} \cdot \begin{bmatrix} \upsilon_{i2} \\ \upsilon_{i3} \\ \cdot \\ \cdot \\ \upsilon_{it-1} \\ \upsilon_{it} \\ \upsilon_{it+1} \\ \cdot \\ \cdot \\ \upsilon_{it+s-1} \\ \upsilon_{it+s} \end{bmatrix}$$

where $\ln(\pi^*_{it}) = \ln(\pi_{it}) - (\tilde{\beta}_0 + \beta_3 e_t + \tilde{\beta}_4 t)$. An Aitken estimator based on this relation, applied firm by firm and period by period, generates efficient estimates of $\tilde{x}_{it}$.

Finally, to impute values of $(\tilde{x}^1_{it}, \tilde{x}^2_{it})$ for firms that are *not* currently exporting, we regress $(\tilde{x}^1_{it}, \tilde{x}^2_{it})$ on observable plant characteristics in our sub-sample exporters, then we use the resulting parameters to impute $(\tilde{x}^1_{it}, \tilde{x}^2_{it})$ values for non-exporting firms. The set of observable plant characteristics includes lagged values of domestic sales, an average variable cost proxy and real capital stocks, as well as location dummies, and our Mills ratio.