2005

2005-07 Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence

Hiroyuki Kasahara

Joel Rodrigue

Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence

by

Hiroyuki Kasahara and Joel Rodrigue

Working Paper # 2005-7 **September 2005**

RBC Financial Group
Economic Policy Research Institute
EPRI Working Paper Series

Department of Economics
Department of Political Science
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence*

Hiroyuki Kasahara† and Joel Rodrigue‡

This version: August 30, 2005

Abstract

This paper examines whether importing intermediate goods improves plant performance. While addressing the issue of simultaneity of a productivity shock and the decision to import intermediates, we estimate the impact of the use of foreign intermediates on plants' productivity using plant-level Chilean manufacturing panel data. We found that switching from being a non-importer to being an importer of foreign intermediates can improve productivity by 2.3 to 22.0 percent. We also investigate the plant dynamic decisions to import, invest, and exit. The results show that having imported last year substantially increases the probability of importing this year, providing the evidence for sunk start-up costs of importing. We also found that importers accumulate more capital and are less likely to exit than non-importers, which indicates that importing intermediates may play an important role in reallocating resources across heterogenous plants.

KEYWORDS: productivity, imported intermediates, plant-level

JEL: F10, D21, D24

*The authors would like to thank Johannes Van Biesebroeck, Lee Branstetter, David Hummels, seminar participants at the Bank of Canada, CEA Annual Meeting, EIIT Conference, North America Econometric Society Meeting, and the University of Western Ontario for helpful comments.

†Department of Economics, the University of Western Ontario. E-mail: hkasahar@uwo.ca.

‡Department of Economics, Queen's University. E-mail: rodrigue@qed.econ.queensu.ca.

1 Introduction

International trade is one of the primary avenues for the diffusion and adoption of new technologies worldwide. This is particularly true and important for developing nations where it is believed that importing new technologies is a significant source of productivity and economic growth. Through adoption and imitation of imported technologies, countries can take advantage of research and development (R&D) abroad to improve the efficiency of domestic production.

The previous empirical work using aggregate cross-country data show that importing intermediate goods that embody R&D from an industrial country can significantly boost a country's productivity (c.f., Coe and Helpman, 1995; Coe, Helpman and Hoffmaister, 1997). Countries that are more open to trade benefit more from foreign R&D because they are better able to access improvements in technology by importing intermediate goods.[1] Aggregated data, however, do not capture heterogeneity across different plants in the economy. As empirically shown by Baily, Hulten, and David (1992), to understand changes in aggregate productivity levels it is vital to examine plant-level changes. Furthermore, recent developments in trade theory suggest that understanding the plant-level response to trade policy is a crucial factor in understanding its impact on aggregate productivity (e.g., Melitz, 2003; Bernard, Eaton, Jensen and Kortum, 2003). Yet, few empirical studies have been done on the impact of importing intermediate goods on productivity at plant level.[2]

The goal of this paper is to test whether the use of foreign intermediate goods increases plant productivity, using a detailed unbalanced panel data set on Chilean manufacturing plants from 1979-1996. The data set captures heterogeneity in terms of import status across plants and across time: some plants import most of the intermediate goods, some change their import status, others do not import at all. While importers are larger and more productive than non-importers in the data, the direction of causality between importing foreign intermediates and plant's performance is not clear. Does the use of foreign intermediate goods per se increase productivity? Or, do inherently high productivity plants tend to use foreign intermediate goods? To answer these questions, we estimate both static and dynamic effect of the use of imported intermediates

[1] Keller (2001) provides the industry-level empirical evidence for the role of R&D spillovers through imports.

[2] Exceptions are Muendler (2004) and Van Biesebroeck (2003). On the other hand, there is a growing literature on the impact of exporting on firm performance (c.f., Clerides, Lach, and Tybout, 1998; Bernard and Jensen, 1999).

on plant's productivity while addressing the important econometric issues of simultaneity and endogenous selection using the estimator developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003) (OP/LP Proxy estimator, hereafter). The results from the Within-Group estimator and the System GMM estimator (c.f., Blundell and Bond, 1998) are also provided.

The results from our analysis indicate substantial positive impact from the use of imported intermediates on plant productivity. The Within-Groups estimates show that the use of imported intermediates leads to an immediate increase in productivity by 2.3 percent. The point estimates from the System GMM and the OP/LP Proxy estimators suggest an even larger impact of the use of imported intermediates on productivity: 18.0-22.0 percent. In addition, we found some evidence for a positive dynamic effect of the use of imported intermediates (i.e., "learning by importing").

Our finding of a positive productivity effect from the use of foreign intermediates suggests that understanding the plant-level import decisions may be particularly important in understanding the impact of trade policy on aggregate productivity. Contrasting with a growing literature on plant-level export decisions (e.g., Roberts and Tybout, 1997; Bernard and Jensen, 1999), however, few empirical studies have examined the determinants of the plant-level import decisions and their implications. Is there a sunk start-up cost for importing intermediates? Does importing intermediates play any role in reallocating resources across heterogenous plants? To answer these questions, we further examine the plant dynamic decision to import and its dynamic interaction with investment and exit decisions. The results show that having imported last period substantially increases the probability of importing today, providing the evidence for sunk start-up cost for importing. We also found that importers tend to accumulate more capital and are less likely to exit than non-importers, which indicates that importing status may play an important role in resource reallocation across different plants.

The paper is organized as follows. The next section proceeds to describe the analytical framework used to study the relationship between productivity and imported intermediates. Section 3 outlines the empirical specification, while sections 4 and 5 explain the estimation procedure and data set, respectively. The sixth section presents the results. The seventh section empirically examines the plant decisions to import and exit and their implications. The last section concludes.

2 The Theoretical Framework

2.1 Production Function

For each period t, the i^{th} plant's production, Y_{it}, is given by:

$$Y_{it} = e^{\omega_{it}} K_{it}^{\beta_k} L_{it}^{s\beta_s} L_{it}^{u\beta_u} E_{it}^{\beta_e} \left[\int_0^{N(d_{it})} x(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\beta_x \theta}{\theta-1}}, \quad (1)$$

where ω_{it} represents a serially correlated productivity shock, K_{it} is capital input, L_{it}^s is skilled labor input, L_{it}^u is unskilled labor input, E_{it} is energy input, and intermediate materials are horizontally differentiated.[3] The elasticity of substitution between any two material inputs is given by $\theta > 1$. The variable $N(d_{it})$ denotes the range of intermediate inputs which are employed in the i^{th} plant; it is a function of a plant's discrete choice, denoted by d_{it}, to import from abroad or not:

$$N(d_{it}) = \begin{cases} N_{h,t}, & \text{for } d_{it} = 0 \\ N_{f,t}, & \text{for } d_{it} = 1 \end{cases}$$

where $N_{h,t}$ is the range of intermediate inputs produced in this country and $N_{f,t}$ is the range of intermediate inputs available in the world. There are a range of intermediate inputs that are not produced domestically in this country but are produced in foreign countries and thus available through imports. Therefore, $N_{f,t} > N_{h,t}$. The ratio

$$\frac{N_{f,t}}{N_{h,t}} \geq 1$$

represents the technological gap in ability to produce a variety of intermediate goods between the rest of the world and this country.

Consider the equilibrium in which all intermediate goods are symmetrically produced at level $\bar{x}$. Substituting $x(j) = \bar{x}$ into equation (1) leads to

$$Y_{it} = e^{\omega_{it}} N(d_{it})^{\frac{\beta_x}{\theta-1}} K_{it}^{\beta_k} L_{it}^{s\beta_s} L_{it}^{u\beta_u} E_{it}^{\beta_e} X_{it}^{\beta_x}, \quad (2)$$

where $X_{it} = N(d_{it})\bar{x}$.

[3] For most of our purposes, we do not need the assumption of the Dixit-Stiglitz specification of horizontally differentiated materials. What matters is the empirical specification in which the measured Total Factor Productivity depends on import decisions.

Total factor productivity (TFP) is defined as $A_{it} = \frac{Y_{it}}{K_{it}^{\beta_k} L_{it}^{s\beta_s} L_{it}^{u\beta_u} E_{it}^{\beta_e} X_{it}^{\beta_x}}$. Then, from equation (2),

$$\ln A(d_{it}, \omega) = \frac{\beta_x}{\theta - 1} \ln(N(d_{it})) + \omega_{it}. \tag{3}$$

This equation indicates that productivity is positively related to the range of employed intermediate inputs. In the view of equation (2), plants importing intermediate inputs from abroad employ a larger variety of intermediate inputs and hence exhibit higher productivity than those employing domestic intermediate inputs only; for example, had there been no difference in the value of ω across plants, then $\ln A(1, \omega) - \ln A(0, \omega) = \frac{\beta_x}{\theta - 1} \ln(N(1)/N(0)) > 0$.

2.2 Exit, Import, and Learning by Importing

The behavioral framework of Olley and Pakes (1996) is extended by incorporating import-decisions into their dynamic model. Consider a risk-neutral plant that maximizes the expected present value of the sum of net cash flows. At the beginning of every period, after observing the current productivity shock ω_t, the plant makes the following decisions. First, it makes a discrete decision to exit, χ_t, by comparing a sell-off value of Φ with its continuation value. If it continues in operation, it chooses the import status (d_t), and then variable factors (labour, materials, fuels) and investment level ($\imath_t$). The capital is accumulated as $K_{t+1} = (1-\delta)K_t + \imath_t$; it is assumed that this year's investment becomes productive the next year. Denote the logarithm of capital stock by k_t.

The past import status may have an impact on the evolution of productivity; importing materials may bring plants into close contact with foreign suppliers in developed countries, which may lead to the positive dynamic externalities, or "learning by importing". To examine the possibility of "learning by importing", we allow the distribution of ω_{t+1} conditional on information available at t to be dependent not only on the past productivity, ω_t, but also on the past import status, d_t.[4] Specifically, the distribution function of ω_{t+1} conditional on ω_t and d_t is given by $F(\cdot|\omega_t, d_t)$. In this context, we formalize the hypothesis of "learning by importing" by the condition that

$$E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1} = 1] > E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1} = 0]$$

[4] Ericson and Pakes (1995) consider the model in which the distribution of ω_{t+1} depends on the amount of *R&D* investment.

for every $\omega_{i,t-1} \in \Omega$. If this is the case, then the current import status d_t may have a dynamic effect on productivity in addition to a (potential) static effect that is implied in (3).

Consider a fixed cost for importing materials, which may depend not only on the current import choice but also on the past import status because of a sunk start-up cost of importing materials.[5] We denote the fixed import cost—which we may think as the sum of the per-period fixed cost and the sunk start-up cost—by $\Gamma(d_{t-1}, d_t)$. Since the profit and the value functions depend on the time specific factors, such as factor prices, we index the profit and the value functions by time. The Bellman equation for the plant can be written as

$$V_t(\omega_t, k_t, d_{t-1}) = \max \left\{ \varPhi_t, \max_{d_t, \imath_t} \{ \pi_t(\omega_t, k_t, d_t) - c(\imath_t, k_t) - \varGamma(d_{t-1}, d_t) + \beta E[V_{t+1}(\omega_{t+1}, k_{t+1}, d_t) | J_t] \} \right\},$$

where $\varPhi_t$ is the sell-off value of the plant, $\pi_t(\cdot)$ is the profit after maximizing out the variable factors, $c(\imath_t, k_t)$ is the cost of investment, $\Gamma(d_{t-1}, d_t)$ is the fixed cost of importing materials, and J_t represents information available at time t. The policy functions associated with the fixed point of the Bellman equation specify an exit rule, a discrete import decision rule, and an investment decision rule. In particular, when the profit function $\pi_t(\cdot)$ is strictly increasing in ω_t, the plant exit rule is characterized by the threshold value $\underline{\omega}_t(k_t, d_{t-1})$ as:

$$\chi_t = \begin{cases} 1, & \text{for } \omega_t \geq \underline{\omega}_t(k_t, d_{t-1}), \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

The discrete import decision rule and the investment demand equation are written, respectively, as:

$$d_t = d_t^*(\omega_t, k_t, d_{t-1}) \tag{5}$$

$$\imath_t = \imath_t^*(\omega_t, k_t, d_{t-1}) \tag{6}$$

Note that the decisions to exit, import, and invest crucially depend not only on the capital stock but also the past import status because the past import status is one of the state variables. Accordingly, we will modify the "standard" OP/LP estimation procedure by incorporating the past import status as an additional state variable.

[5] Section 7 provides empirical evidence for the presence of sunk start-up costs of importing.

3 Econometric Specification

The equation (2) suggests the following specification of the Cobb-Douglas production function augmented by the term representing the use of imported intermediates:[6]

$$y_{it} = \beta_k k_{it} + \beta_s l_{it}^s + \beta_u l_{it}^u + \beta_e e_{it} + \beta_x x_{it} + \beta_d d_{it} + \omega_{it} + \eta_{it}, \tag{7}$$

where $y_{it} = \ln Y_{it}$, $k_{it} = \ln K_{it}$, $l_{it}^s = \ln L_{it}^s$, $l_{it}^u = \ln L_{it}^u$, $e_{it} = \ln E_{it}$, and $x_{it} = \ln X_{it}$. A plant's discrete choice to import from abroad is denoted by d_{it}. ω_{it} is a serially correlated shock and η_{it} is an i.i.d. shock that is not known to plants at the time of input decisions.

We examine whether the use of imported intermediates leads to higher productivity by testing whether $\beta_d > 0$. A positive estimate of β_d provides plant-level evidence for R&D spillovers through trade in intermediate goods. It suggests that plants using the imported intermediates close the technological gap between the home country and the rest of the world, i.e., $\frac{\beta_x}{\theta-1}\ln(N(1)/N(0))$, and hence achieve a higher productivity relative to those only using domestic intermediates.

To examine the possibility of dynamic effect of import status on productivity through "learning by importing," we consider the following stochastic process of ω_{it}:

$$\omega_{it} = \xi_t + \gamma d_{i,t-1} + \rho\omega_{i,t-1} + u_{it}, \tag{8}$$

where ξ_t is a year-specific productivity shock, u_{it} is independent of $d_{i,t-1}$ and $\omega_{i,t-1}$ with the cumulative distribution $F_u(\cdot)$. A positive value of γ is evidence for "learning by importing" and the long-run effect of "learning by importing" is measured by $\frac{\gamma}{1-\rho}$. To examine the robustness, we also consider the flexible specification for the process of ω_{it} based on the polynomials in (ω_{t-1}, d_{t-1}).

The degree of the technological gap may differ across plants, for instance, if they produce different products. In fact, among plants that are using imported intermediates, we observe substantial differences in the ratios of imported intermediates to total intermediates. Assuming that all intermediate goods are symmetrically produced at level $\bar{x}$ and that plants only import a variety of intermediate goods that are not available in domestic market, we may use the ratio of

[6] We focus on the impact of imported intermediates on productivity while excluding the impact of exporting from the specification given that the typical finding from the literature is that, while good firms become exporters, becoming exporters do not necessarily improve firm performance (c.f., Clerides et al., 1998; Bernard and Jensen, 1999).

total intermediates to domestic intermediates as a measurement of the technological gap between world and domestic technology since

$$\frac{X_{it}}{X_{it}^h} = \frac{N(1)\bar{x}}{N(0)\bar{x}} = \frac{N(1)}{N(0)},$$

where X_{it} is total intermediates and X_{it}^h is domestic intermediates.

We examine whether a larger technological gap leads to higher productivity conditional on the use of imported intermediates by considering the following alternative specification:

$$y_{it} = \beta_k k_{it} + \beta_s l_{it}^s + \beta_u l_{it}^u + \beta_e e_{it} + \beta_x x_{it} + \beta_n n_{it} + \omega_{it} + \eta_{it}, \tag{9}$$

where

$$n_{it} = \ln \frac{X_{it}}{X_{it}^h}.$$

From the estimate of β_n and β_x, we may compute the elasticity of substitution across different varieties of intermediate goods, denoted by θ, using $\beta_n = \frac{\beta_x}{\theta - 1}$.

4 Estimation

4.1 The OP/LP Proxy Estimator

One of the main econometric issues in estimating the equations (7)-(9) is the simultaneity of a productivity shock ω_{it} and input decisions. For example, if inputs are chosen on the basis of the productivity shocks, a plant with a higher productivity shock may use more inputs; since the regressors are positively correlated with the error term, the coefficients estimated by ordinary least squares (OLS) may be upwardly biased, especially for variables that are more responsive to a contemporary productivity shock.

The selection due to endogenous exit decisions is another important issue. When a profit function is increasing in k_t and d_{t-1}, the threshold value of productivity that induces exiting, $\underline{\omega}_t(k_t, d_{t-1})$, is decreasing in k_t and d_{t-1}. Specifically, plants having larger capital stocks and importing materials expect larger future profits and hence stay in the market for lower realized values of ω_t; the OLS estimates may lead to biases in the coefficients of capital and imported materials [c.f., Olley and Pakes (1996)].

To deal with the issues of simultaneity and self-selection, we apply the framework developed by Olley and Pakes (1996) and extended by Levinsohn and Petrin (2003).[7] Our specification for production technology differs from those of Olley and Pakes (1996) and Levinsohn and Petrin (2003) in that we have an additional state variable of import status, d_{it}, and that the import status has a dynamic effect on productivity as specified by (8). Since these differences lead to subtle differences in estimating procedures, we discuss the details of our estimating procedure in the following.

Suppose that capital k_{it} and the import decision d_{it} are the state variables but l_{it}, x_{it}, and e_{it} are freely variable inputs. Then, the material's demand function is given as $x_{it} = x_t^*(\omega_{it}, k_{it}, d_{it})$, where the function $x_t^*(\cdot)$ is time-dependent, reflecting its dependence on time-specific common shocks in productivity and relative prices (e.g., changes in exchange rates and tariff rates). The material's demand function crucially depends on the import decision d_{it} for the following two reasons. First, if there is a sunk start-up cost of importing materials, then the current import choice d_{it} is not freely variable and hence it should be included in the material demand function.[8] Second, since plants using imported materials are likely to face a different material input market than those using only domestic materials, the material's demand function must be not only time-dependent but also import-status-dependent.

Assuming that $x_t^*(\cdot)$ is strictly increasing in ω_{it}, we can invert this function to obtain the productivity shock ω_{it} as a function of (x_{it}, k_{it}, d_{it}):[9]

$$\omega_{it} = \omega_t^*(x_{it}, k_{it}, d_{it}). \tag{10}$$

Replacing $\omega_t^*(x_{it}, k_{it}, d_{it})$ for ω_{it} in the equation (7) leads to a partial linear function:

$$y_{it} = \beta_s l_{it}^s + \beta_u l_{it}^u + \beta_e e_{it} + \phi_t(x_{it}, k_{it}, d_{it}) + \eta_{it}, \tag{11}$$

where

$$\phi_t(x_{it}, k_{it}, d_{it}) = \beta_k k_{it} + \beta_x x_{it} + \beta_d d_{it} + \omega_t^*(x_{it}, k_{it}, d_{it}).$$

[7] Levinsohn and Petrin's estimator is developed based on the investment proxy estimator of Olley and Pakes (1996). In the Chilean data, there are a substantial number of zero investment observations (perhaps due to the presence of fixed investment cost). For these observations, the investment proxy estimator of Olley and Pakes cannot be used because they do not satisfy the monotonicity condition (and thus the investment function is not invertible with respect to shocks). Given this feature of the Chilean data, we choose to use the Levinsohn and Petrin intermediate proxy estimator rather than the Olley and Pakes investment proxy estimator.

[8] Section 7 presents evidence for the presence of sunk start-up costs of importing materials.

[9] See Levinsohn and Petrin (2003) for the general conditions on the production technology under which a demand function for intermediate inputs is strictly increasing in productivity.

In the first stage, we obtain consistent estimates of β_s, β_u, and β_e. By subtracting the expectation of (11) conditional on (x_{it}, k_{it}, d_{it}) from (11), we obtain

$$y_{it} - E(y_{it}|x_{it}, k_{it}, d_{it}) = \beta_s(l^s_{it} - E(l^s_{it}|x_{it}, k_{it}, d_{it})) + \beta_u(l^u_{it} - E(l^u_{it}|x_{it}, k_{it}, d_{it})) + \beta_e(e_{it} - E(e_{it}|x_{it}, k_{it}, d_{it})) + \eta_{it}. \quad (12)$$

We first consistently estimate the conditional expectations, $E(y_{it}|x_{it}, k_{it}, d_{it})$, $E(l^s_{it}|x_{it}, k_{it}, d_{it})$, $E(l^u_{it}|x_{it}, k_{it}, d_{it})$, and $E(e_{it}|x_{it}, k_{it}, d_{it})$ by the OLS regressions of y_{it}, l^s_{it}, l^u_{it}, and e_{it}, respectively, on the power series of (x_{it}, k_{it}, d_{it}).[10] Using the estimates of the conditional expectations in place of the actual conditional expectations in (12), we estimate β_s, β_u, and β_e by OLS with no-intercept. Denote the estimates by $\hat{\beta}_s$, $\hat{\beta}_u$, and $\hat{\beta}_e$. Note that β_k, β_x, and β_d are not identified in the first stage.

In the second stage, we first obtain the estimate of $E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1]$ as follows (c.f., Olley and Pakes, 1996). From (4) and (8), define the threshold value of u_{it} that induces a plant to exit at t by $\underline{u}_{it} \equiv \underline{\omega}_t(k_{it}, d_{i,t-1}) - \gamma d_{i,t-1} - \rho\omega_{i,t-1}$. Since a plant continues in operation if $u_{it} \geq \underline{u}_{it}$, the survival probabilities are given by

$$Pr\{\chi_{it} = 1|\underline{u}_{it}, \omega_{i,t-1}, d_{i,t-1}\} = 1 - F_u(\underline{u}_{it}) \equiv P_{it}, \quad (13)$$

where $F_u(\cdot)$ is the cumulative distribution of u_{it}. By inverting (13), we may obtain $\underline{u}_{it}$ as a function of P_{it} and write this inverse function as $\underline{u}_{it} = \underline{u}^*(P_{it})$. Then, the conditional expectation of ω_{it} given $\omega_{i,t-1}$, $d_{i,t-1}$, and $\chi_{it} = 1$ can be expressed as

$$E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1] = \xi_t + \gamma d_{i,t-1} + \rho\omega_{i,t-1} + E[u_{it}|u_{it} \geq \underline{u}^*(P_{it})]. \quad (14)$$

We obtain the estimate of $E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1]$ by the pooled OLS regression of $(\omega_{it} \hat{+} \eta_{it})(\beta^*) \equiv y_{it} - \hat{\beta}_s l^s_{it} - \hat{\beta}_u l^u_{it} - \hat{\beta}_e e_{it} - \beta^*_k k_{it} - \beta^*_x x_{it} - \beta^*_d d_{it}$ on the past import status $d_{i,t-1}$, the estimate of the previous period's productivity shock $\hat{\omega}_{i,t-1}(\beta^*) \equiv \hat{\phi}_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1}) - \beta^*_k k_{i,t-1} - \beta^*_x x_{i,t-1} - \beta^*_d d_{i,t-1}$, and a third-order polynomial series of the survival probability (13) which approximates the term $E[u_{it}|u_{it} \geq \underline{u}^*(P_{it})]$.[11] In estimating (14), we also allow for year-specific constant terms,

[10] The results presented in this paper use a third order polynomial with a full set of interactions to approximate unknown functions. We have also tried a fourth order polynomial for some cases and found very similar estimates of the coefficients $(\beta_x, \beta_k, \beta_d)$.

[11] $\hat{\phi}_t(\cdot)$ is the estimate of $\phi_t(\cdot)$ obtained by the OLS regressions of $y_{it} - \hat{\beta}_s l^s_{it} - \hat{\beta}_u l^u_{it} - \hat{\beta}_e e_{it}$ on a third-order polynomial series of (x_{it}, k_{it}, d_{it}) as implied by equation (11). On the other hand, using (10), the survival

ξ_t, to control for the year-specific productivity shocks that are common across firms, such as the aggregate disembodied technical change and the shocks associated with business cycles.

Define the innovations in productivity conditional on last year's productivity, last year's import status, *and* survival:

$$\nu_{it} = \omega_{it} - E[\omega_{it}|\omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1],$$

where $\chi_{it} = 1$ if the i^{th} plant continues in operation at t and $\chi_{it} = 0$ if it exits. For each candidate parameter vector $\beta^* = (\beta^*_x, \beta^*_k, \beta^*_d)$, we may construct an estimate for the residual as:

$$(\nu_{it} \hat{+} \eta_{it})(\beta^*) = y_{it} - \hat{\beta}_s l^s_{it} - \hat{\beta}_u l^u_{it} - \hat{\beta}_e e_{it} - \beta^*_k k_{it} - \beta^*_x x_{it} - \beta^*_d d_{it} - \hat{E}[\omega_{it}|\omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1]. \quad (15)$$

To identify β_x, β_k, and β_d, we use the following three moment conditions:

$$E[(\nu_{it} + \eta_{it})x_{it-1}] = 0, \quad (16)$$

$$E[(\nu_{it} + \eta_{it})k_{it}] = 0, \quad (17)$$

$$E[(\nu_{it} + \eta_{it})d_{it-1}] = 0. \quad (18)$$

The identification of β_x comes from the moment condition (16) which is implied by the fact that the last period's material choice is uncorrelated with this period's innovation in productivity. The second moment condition (17) identifies β_k and comes from the fact that the past investment decisions are uncorrelated with this period's innovation in productivity.[12] The third moment condition (18) identifies β_d by using the fact that the past import decisions are uncorrelated with the innovation in productivity this period. Note that, by using the past, rather than the current, import variable as an instrument to identify β_d, we allow for the possibility that a plant makes an import decision this period *after* observing this period's innovation in productivity.

In addition to these three moment conditions, we also include six over-identifying conditions using the predetermined variables $(k_{i,t-1}, d_{i,t-2}, x_{i,t-2}, l^s_{i,t-1}, l^u_{i,t-1}, e_{i,t-1})$ as additional instruments. The parameters $\beta^* = (\beta_x, \beta_k, \beta_d)$ are estimated by minimizing the GMM criterion function

$$Q(\beta^*) = \sum_{h=1}^{9}[\sum_{i=1}^{N}\sum_{t=1981}^{T_i}(\nu_{it} \hat{+} \eta_{it})(\beta^*)Z_{it,h}]^2, \quad (19)$$

probability (13) is $P_{it} = 1 - F_u(\underline{\omega}_t(k_{it}, d_{i,t-1}) - \gamma d_{i,t-1} - \rho\omega^*_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1}))$ and hence is a nonlinear function of $(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})$. We estimate the survival probability by the probit with a third-order polynomial series in $(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})$ as regressors.

[12]The moment conditions (16)-(17) in this paper correspond to the moment conditions (13)-(14) in Levinsohn and Petrin (2003).

where $(\nu_{it} \hat{+} \eta_{it})(\beta^*)$ is given by (15), T_i is the last year the i^{th} firm is observed, $Z_{it,h}$ is the h^{th} element of the instrument vector $Z_{it} = (k_{it}, k_{i,t-1}, d_{i,t-1}, d_{i,t-2}, x_{i,t-1}, x_{i,t-2}, l^s_{i,t-1}, l^u_{i,t-1}, e_{i,t-1})$. The standard errors are obtained by the bootstrap.[13]

We test the hypothesis that a contemporary use of imported materials increases plant productivity by examining whether $\beta_d > 0$. We also test the hypothesis of "learning by importing" by examining whether $\gamma > 0$ or not.

4.2 Alternative Estimators

To address the simultaneity issue, we also consider the following two alternatives: the within-groups estimator and the system GMM estimator (c.f., Blundell and Bond, 1998).

The within-groups estimator only uses the within-plant variation so that it is robust against the simultaneity arising from the correlation between an unobserved plant-specific productivity shock and inputs. It is not robust, however, against the simultaneity due to the correlation between a transitory shock and inputs. Furthermore, the between-plant variation often plays an important role in identifying the parameters; this is especially true for coefficients of capital and imported intermediates where the within-plant variation is much less than the between-plant variation due to the slow adjustment of capital and the persistence in the import status over time. The within-estimator may lead to imprecise estimates especially for capital and imported intermediates.

To deal with the issue of simultaneity in panel data, Blundell and Bond (1998, 2000) propose the system GMM estimator by extending the first differenced GMM estimator (c.f., Arellano and Bond, 1991). Consider the equation (7) with the following additional structure on ω_{it}:

$$\begin{aligned} \omega_{it} &= \xi_t + \alpha_i + v_{it}, \\ v_{it} &= \gamma d_{i,t-1} + \rho v_{i,t-1} + \zeta_{it}, \end{aligned} \tag{20}$$

where ξ_t is a year-specific effect, α_i is a plant-specific effect, v_{it} is AR(1) productivity shock with $|\rho| < 1$, and ζ_{it} is MA(0). The hypothesis of "learning by importing" can be tested by

[13] Assuming that each set of plant-level observations $\mathbf{x_i} \equiv \{y_{it}, k_{it}, l^s_{it}, l^u_{it}, x_{it}, e_{it}, d_{it}\}_{t=1980}^{T_i}$ is independently and identically distributed across plants, we draw a bootstrap sample until the size of plant-time observations reaches $\sum_{i=1}^N T_i$ with replacement from the original sample $(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$. We generate 200 independent bootstrap samples and estimate the parameters for each sample using the recentered moment conditions (c.f., Horowitz, 2001). The bootstrap standard errors are then computed based on the B sets of parameter estimates.

examining whether $\gamma > 0$. Using a dynamic common factor representation, (7) with (20) can be rewritten as:

$$\begin{aligned} y_{it} &= \beta_k k_{it} - \rho\beta_k k_{i,t-1} + \beta_s l_{it}^s - \rho\beta_s l_{i,t-1}^s + \beta_u l_{it}^u - \rho\beta_u l_{i,t-1}^u + \beta_e e_{it} - \rho\beta_e e_{i,t-1} \\ &\quad + \beta_x x_{it} - \rho\beta_x x_{i,t-1} + \beta_d d_{it} + (\gamma - \rho\beta_d) d_{i,t-1} + \rho y_{i,t-1} + \xi_t^* + \alpha_i^* + \mu_{it} \end{aligned} \quad (21)$$

where $\xi_t^* = \xi_t - \rho\xi_{t-1}$, $\alpha_i^* = (1-\rho)\alpha_i$, and $\mu_{it} = \zeta_{it} + \eta_{it} - \rho\eta_{i,t-1}$.

Following Blundell and Bond (2000), we first estimate the unrestricted parameter vector of (21) by the one-step GMM and then obtain the restricted parameter vector $(\beta_k, \beta_s, \beta_u, \beta_e, \beta_m, \beta_d, \gamma, \rho)$ using minimum distance (c.f., Chamberlain, 1982). The following two set of moment conditions are used:

$$E[z_{i,t-s}\Delta\mu_{it}] = 0 \quad \text{for } s = 2, 3, \quad (22)$$

$$E[\Delta z_{i,t-s}(\alpha_i^* + \mu_{it})] = 0 \quad \text{for } s = 1, \quad (23)$$

where $z_{it} = (y_{it}, k_{it}, l_{it}^s, l_{it}^u, x_{it}, d_{it})$ and $\Delta z_{it} = z_{it} - z_{i,t-1}$. The first set of the moment conditions (22) comes from the first differenced equations (to eliminate the plant-specific effect) with lagged levels of the variables as instruments (c.f., Arellano and Bond, 1991). The first-differenced GMM estimator based only on these moment conditions (22) may have poor finite sample properties due to weak instruments. Blundell and Bond (1998) find that exploiting the additional moment conditions (23), based on the level equations with lagged differences of the variable as instruments, may lead to dramatic reductions in finite sample bias. Recently, however, some researchers found that even the system GMM estimator could lead to imprecise and possibly biased estimates due to weak instruments.[14]

5 Data

The data set is based on an annual census of Chilean manufacturing plants, which covers all plants with more than 10 workers, by Chile's Instituto Nacional de Estadistica (INE) for 1979-1996.[15] The data set includes gross revenue, the number of blue- and white-collar workers,

[14]For example, see Griliches and Mairesse (1998), Mulkay, Hall, and Mairesse (2000), and Levinsohn and Petrin (2003).

[15]The unit of observation in our empirical analysis is "plant" rather than "firm." This is due to limitations of our data set. Firm-level analysis might be particularly important to address the issue of "learning-by-importing"; the dynamic learning through importing might be more important at firm-level than at plant-level.

various types of investment, imported materials, total materials, electricity and fuels. Each variable is deflated by using the corresponding annual price deflator to real 1980 Chilean pesos.[16] We exclude plants for which any of the data for capital stocks, unskilled labor, skilled labor, energy, and domestic intermediates are either not available or reported as zero values. In particular, plants that do not report book values of their capital stocks in any year are excluded since constructing capital stocks for these plants is impossible.[17] After cleaning the data, the unbalanced panel data set contains 3598 plants. Since a substantial number of plants are dropped out of the sample due to the missing initial capital stock and this might lead to a sample selection problem, we also report the results under the extended sample of 4508 plants in which the capital stock in 1980, if it is missing, is imputed by a projected capital stock based on other reported plant observables. Hereafter, the sample that excludes the plants missing book values of their capital stocks is called the "Basic Sample" while the sample with imputed capital stocks is called the "Extended Sample." Since the main features of the descriptive statistics of these two samples are similar, we focus on the statistics of the Basic Sample in this section; the appendix provides the statistics of the Extended Sample.

Output is total revenue adjusted for inventory change. Real output variable (Y) is constructed by deflating nominal output using an industry output price deflator. Real domestic material (X^h) is constructed by subtracting the nominal value of imported materials from the total materials and then deflated using an industry price deflator.[18] Real imported materials is constructed by deflating the nominal imported materials by the import price index (in Chilean peso) reported in International Financial Statistics. The real value of total materials (X) is the sum of the real domestic materials and the real imported materials. The number of blue- and white-collar workers are used for skilled and unskilled labor input (L^s and L^u). The energy

[16]Previous empirical studies using (a subset of) this data set includes Lui (1993), Pavcnik (1999), Levinsohn and Petrin (2003).

[17]The book values of capital are only reported, if any, in 1980 and 1981. Some plants did not report the book values of capital in either 1980 or 1981. Since it is not possible to construct capital stock without these reports, the plants missing their book values of capital were excluded from the sample. Notably, plants enter into the market after 1982 are not included in the sample. We focus on the sample of plants that operated in both 1979 and 1980 so that we may use the variables that are two period lagged in our regression analysis.

[18]For both output price deflator and intermediate price deflator, we have used a 3-digit industry deflator for 1979-1986, which is contained in the original data set as described in Lui (1993), and a 2-digit industry deflator for 1987-1996 obtained from Yearbook of National Accounts by the Central Bank of Chile. As far as we know, the material price deflators at 3-digit levels are not available after 1987.

input (E) is the sum of the real purchased value of electricity and that of fuels. The value of imported materials is reported separately from the total value of materials.

The capital stock is constructed separately for buildings, machinery and equipment, and vehicles from the 1980 book value of capital (the 1981 book value if the 1980 book value is not available) using perpetual inventory method: $K_{t+1} = (1-\delta)K_t + \imath_t$.[19] The nominal net investment variable is constructed, separately for buildings, machinery and equipment, and vehicles, and then deflated using the construction deflator for buildings and the machinery deflator for machinery and equipment, and vehicles to obtain the real net investment ($\imath$).[20] Buildings are likely to be rented rather than owned by plants, since zero values are found frequently for buildings, especially for small plants. We add the capitalized rental value to current year capital stock.[21] The total capital stock (K) is the sum of the real capital stock for building, machinery and equipment, and vehicles, and the capitalized rental value. Note that the capital stock in year t does not include the investment in year t.

Table 1 reports descriptive statistics for variables in the year of 1980. A comparison between "Importing Plants," 'Non-Importing Plants," and "Switchers" in Table 1 reveals the substantial differences between the three types of plants. Importing plants are substantially larger and have higher labor productivity while 'Non-Importing Plants" are smaller and least productive among those three types of plants, although the direction of causality is not clear. On the other hand, as shown in the last two rows of Table 1, "Survivors" which do not exit before 1996 are larger, more productive, and tend to import more in 1980 than "Quitters" that exit within the sample period of 1980-1996.

Out of 3598 plants, 273 plants (7.6%) continuously import foreign intermediates throughout the sample period (i.e., "Importing Plants" in Table 1), while 2017 plants (56.1%) are "Non-

[19]Since the reported book values are evaluated at the end of year t, the book values of capital are deflated by the (geometric) average deflator of machinery and equipment for years t and t+1. Depreciation rates are set to 5 % for building, 10 % for machinery and equipment, and 20% for vehicles.

[20]The data contains information on five types of investments: (i) purchases of new capital, (ii) purchases of used capital, (iii) production of capital for own use, (iv) improvements in own capital by third parties, and (v) sales of capital. The net investment is the sum of (i)-(iv) minus (v).

[21]The data on rental rate is not available. To obtain a crude measure of rental rate, assuming the aggregate Cobb-Douglas production, we compute (rental rate)=(the share of capital)$\times$ GDP/(Capital Stock)-(depreciation rate)$\approx$ 0.15 on average for 1980-1996 using the data on Chilean GDP and Capital Stock constructed from the Chilean national accounting data with (the share of capital)=0.3 and (depreciation rate)=0.05. The capitalized rental value is computed as (rental value)/0.15.

Table 1: Descriptive Statistics in 1980

	Output	Capital	Labor	Energy	Interme-diates	Import Ratios	Output/Workers	No. of Plants
All Plants	95.58	45.20	54.45	3.35	49.36	0.08	1.16	3598
	(437.46)	(253.38)	(105.09)	(26.13)	(221.81)	(0.18)	(1.63)	—
Importing Plants	445.64	194.32	177.17	11.20	201.87	0.37	2.56	273
	(1021.01)	(431.44)	(256.26)	(34.72)	(407.35)	(0.25)	(3.73)	—
Non-Importing Plants	20.84	9.08	26.18	0.66	12.60	—	0.74	2017
	(38.29)	(51.95)	(29.51)	(5.52)	(25.28)	—	(0.72)	—
Switchers	137.77	69.78	72.44	5.86	74.23	0.13	1.50	1308
	(521.03)	(355.70)	(103.35)	(39.38)	(303.84)	(0.22)	(1.68)	—
Survivors	170.97	76.44	77.74	6.28	84.93	0.11	1.50	1348
	(79.34)	(376.54)	(139.50)	(40.20)	(338.10)	(0.21)	(1.96)	—
Quitters	50.41	26.48	40.50	1.60	28.05	0.05	0.95	2250
	(155.63)	(129.72)	(74.09)	(10.77)	(94.93)	(0.16)	(1.36)	—

Notes: Standard errors are in parentheses. The statistics are based on the "Basic Sample" that excludes plants for which the initial capital stock are not reported. "Importing Plants" are plants that continuously imported foreign intermediates in the sample. "Non-Importing Plants" are plants that never imported foreign intermediates in the sample. "Switchers" are plants that switched their import status in the sample. "Survivors" are plants that did not exit during the sample period (1980-1996) while "Quitters" exit during the sample period. "Output," "Capital," "Energy," and "Intermediates" are measured in millions of 1980 pesos. "Labor" is the number of workers. "Import Ratios" are the ratios of imported intermediate materials to total intermediate materials.

Table 2: Transition Probability of Import Status and Exit

Year t status	No Imports			Imports		
Year $t+1$ status	No Imports	Imports	Exit	No Imports	Imports	Exit
1981-1985 ave.	0.844	0.054	0.102	0.170	0.788	0.042
1986-1990 ave.	0.885	0.055	0.061	0.173	0.805	0.023
1991-1995 ave.	0.874	0.067	0.058	0.119	0.860	0.021
1981-1995 ave.	0.868	0.059	0.074	0.154	0.818	0.028

Notes: The statistics are based on the "Basic Sample" that excludes plants for which the initial capital stock are not reported.

Importing Plants" that never import intermediates from abroad. This suggests that plant import status is persistent over time. There are, nevertheless, 1308 plants (36.4%) that switch between importing and not importing over the period and, among them, 757 plants switch import status more than once. This within-plant variation of import status is, thus, an important source of identification of the import variable coefficient.

Table 2 presents transition rates across import status together with exit rates. The last row indicates the average transition rates for 1981-1995. The persistence in import status is also clear here; among the plants that did not import in year t, more than 85 percent of them did not import in year $t+1$, while, among the plants that did import in year t, about 82 percent of them did import in year $t+1$. Comparing plants across import status in year t, we notice that importers are more likely to survive than non-importers.

6 Results

Table 3 presents the results from various estimators using the discrete choice import variable; columns (1)-(5) report the results of the Basic Sample while columns (6)-(8) report those of the Extended Sample.[22] The most important finding is the significance and often large size of the current discrete import variable coefficient across different estimators. The OLS point estimate implies that a plant only using domestic intermediates can increase its productivity by 3.8 percent if it starts importing intermediates. The OLS estimate is, however, likely to be biased due to correlation between an unobserved plant productivity shock and inputs. The

[22]The results from the OLS and the Within estimators for the Extended Sample are similar to those for the Basic Sample reported in columns (1)-(2) of Table 3.

Table 3: Estimates of Production Function: Discrete Import Variable

The Data Set	Basic Sample					Extended Sample		
Estimators	OLS	Within	GMM	OP/LP Proxy		GMM	OP/LP Proxy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skilled labor	0.083	0.057	0.034	0.137		0.038	0.127	
	(0.004)	(0.004)	(0.031)	(0.006)		(0.027)	(0.008)	
Unskilled labor	0.167	0.173	0.251	0.145		0.243	0.142	
	(0.005)	(0.006)	(0.032)	(0.008)		(0.028)	(0.008)	
Energy	0.081	0.067	0.092	0.043		0.002	0.057	
	(0.003)	(0.003)	(0.025)	(0.005)		(0.002)	(0.006)	
Capital	0.079	0.049	0.108	0.058	0.064	0.139	0.065	0.076
	(0.003)	(0.003)	(0.019)	(0.009)	(0.010)	(0.016)	(0.011)	(0.012)
Materials	0.628	0.568	0.612	0.549	0.509	0.655	0.575	0.525
	(0.004)	(0.005)	(0.023)	(0.025)	(0.029)	(0.020)	(0.024)	(0.037)
Disc. Import	0.038	0.023	0.180	0.214	0.220	0.161	0.139	0.129
	(0.005)	(0.005)	(0.049)	(0.035)	(0.067)	(0.045)	(0.032)	(0.039)
γ	0.044	0.007	0.009	0.041	—	0.008	0.024	—
	(0.003)	(0.005)	(0.013)	(0.011)		(0.011)	(0.009)	
ρ	0.723	0.287	0.245	0.892	—	0.271	0.900	—
	(0.004)	(0.009)	(0.022)	(0.027)		(0.016)	(0.116)	
Implied $\frac{\gamma}{1-\rho}$	0.159	0.010	0.012	0.379	—	0.011	0.235	—
P-value for over-identification test				0.593	0.759		0.874	0.427
No. of Obs.	33200					45518		

Notes: Standard errors are in parentheses. Columns (1)-(5) use the "Basic Sample" that excludes plants for which the initial capital stock are not reported. Columns (6)-(8) use the "Extended Sample" in which a missing initial capital stock is imputed by a projected initial capital stock based on other reported plant observables. The System GMM estimator in columns (3) and (6) use a lag length of 2 and 3 for instruments in the first-differenced equations and a lag length of 1 in the level equations. The OP/LP estimators in columns (5) and (8) specify the stochastic process of ω_t using the third order polynomials in (ω_{t-1}, d_{t-1}) and (ω_{t-1}, n_{t-1}), respectively.

within-estimator is robust against the simultaneity between a permanent plant-specific shock and input decisions. Column (2) of Table 3 demonstrates that although estimate of β_d is smaller using the within-estimator relative to OLS, at 2.3 percent, it is still positive and significant.

While the within-estimator controls for correlation between inputs and a permanent shock, it does not address the simultaneity between inputs and the persistent shock that varies within-plant over time. To correct for such simultaneity in panel data, we further provide the results from two alternative estimators: the system GMM estimator and the OP/LP Proxy estimator. The system GMM estimates in columns (3) and (6) of Table 3 also indicate that imports have a strong, positive effect on plant productivity. The model finds 18.0 and 16.1 percent increases in productivity from a switch to imports for the Basic and the Extended Sample, respectively.

Finally, columns (4)-(5) and (7)-(8) of Table 3 provide the results of the OP/LP Proxy estimator which controls for both selection and correlation between inputs and an unobserved productivity shock by using intermediate inputs as proxies for unobserved productivity shocks.[23] The over-identification restrictions are not rejected for all four cases.[24] Column (4) reports the OP/LP estimates under the AR(1) specification for the ω_{it} process using the Basic Sample. It indicates a large productivity effect (21.4 percent) arising from the usage of imported intermediate goods. The estimate from the Extended Sample reported in column (7) is slightly smaller but still indicates a large productivity effect of 13.9 percent. To examine the robustness of the results with respect to the specification of the productivity process, we report the OP/LP estimates under the flexible specification for ω_t using a third-order polynomial in (ω_{t-1}, d_{t-1}) with selection. The results are reported in columns (5) and (8). Once again, it indicates a large productivity effect (22.0 and 12.9 percents for the Basic and the Extended Sample, respectively) from the usage of imported intermediates. Overall, the results suggest that the use of imported materials has a substantial positive static effect on productivity.

Another interesting finding is that, throughout all columns, the estimated values of γ are positive and often significant. This suggests a positive *dynamic* effect of the use of imported materials (i.e., "learning by importing") although the evidence is not as strong as the case for the static effect given the small and not significant estimates of γ for Within Groups estimator and GMM estimator as reported in column (2), (3), and (6). The "long-run" impacts of the use of imported materials on productivity, $\frac{\gamma}{1-\rho}$, implied by the OP/LP estimates in columns (4) and (7) are 37.9 and 23.5 percent, respectively.

Figure 1 shows what would happen to the total factor productivity, defined as $\beta_d d_{it} + \omega_{it}$, for a plant that is not importing at the steady state (i.e., $\omega_{i0} = 0$) before Year 0 and, for some exogenous reason, starts importing intermediates at Year 0.[25] The solid line indicates

[23]Since both the investment and the import policy functions (5)-(6) may differ across years, due to the macroeconomic cycles and the changes in trade policies, we allow for $\phi(\cdot)$ to differ across the following six periods: 1979-1981, 1982-1983, 1984-1986, 1987-1989, 1990-1992, and 1993-1996.

[24]In addition to the over-identification test, we conducted two other specification tests suggested by Levinsohn and Petrin (2003). First, to be consistent with the model, productivity shock should be monotonically increasing in the materials, holding state variables (i.e., capital and import status) constant. By plotting the productivity proxy ω_{it} as a function of capital and intermediate inputs separately for importers and non-importers, we found that this is indeed the case. Second, we use the energy variable in place of the materials as an input proxy and found that the estimated impact of imported materials on productivity is even larger, ranging from 43.1-44.8 percent, under the energy proxy.

[25]Idiosyncratic shocks, u_{it}, are set to zero for all periods. Or, alternatively, we may interpret the solid line as the path of "average" productivity among plants that start importing at Year 0 and keep importing after Year 0.

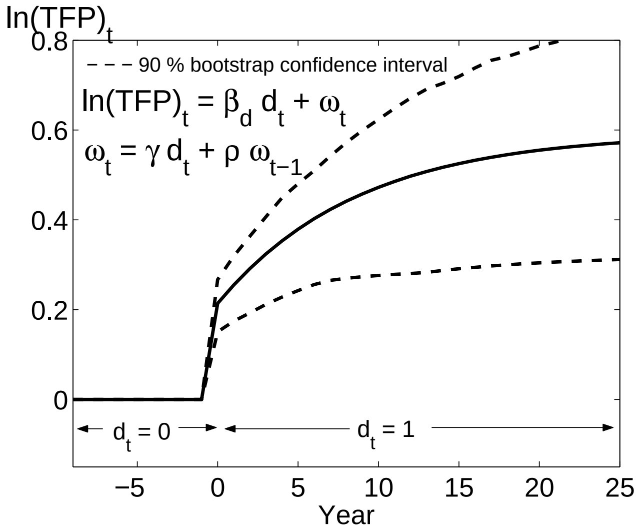

Figure 1: Productivity Dynamics and Import Status

the dynamic path of productivity implied by the point estimates reported in column (4) while the dashed lines represent a 90 percent bootstrap confidence interval.[26] At Year 0, a plant starts importing foreign intermediates, leading to an immediate increase in productivity by 21.4 percent (static effect). After Year 0, a plant gradually achieves additional 37.9 percent productivity increase (dynamic effect). Note, however, that a 90 percent bootstrap confidence interval suggests that there exits a substantial uncertainty regarding the precise magnitude of the dynamic effect of importing materials.

Table 4 presents estimates using the continuous measure of import usage, measured by the ratio of total intermediates to domestic intermediates. The results from the continuous import variable are largely similar to those from the discrete variable. Across various estimators, the coefficients for the continuous import variable are often significant and of large size across different estimators, indicating the importance of foreign intermediates in explaining productivity differences across plants and over time. The system GMM estimates in columns (3) and (6) imply that a 100 percent decrease in the share of domestic intermediates in total intermediates could increase productivity between 5.8 and 7.2 percent although the estimate from the Basic Sample in column (3) is not significant. The OP/LP estimates reported in columns (4)-(5) and

[26]To construct a 90 percent bootstrap confidence interval, we repeatedly compute the dynamic path of productivity under different bootstrap estimates for (β_d, γ, ρ) and take a 5^{th} and a 95^{th} percentile of $\beta_d d_{it} + \omega_{it}$ for each year.

Table 4: Estimates of Production Function: Continuous Import Variable

The Data Set	Basic Sample					Extended Sample		
Estimators	OLS	Within	GMM	OP/LP Proxy		GMM	OP/LP Proxy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skilled labor	0.103	0.057	0.056	0.138		0.046	0.128	
	(0.004)	(0.004)	(0.031)	(0.006)		(0.026)	(0.008)	
Unskilled labor	0.163	0.172	0.278	0.148		0.253	0.145	
	(0.005)	(0.006)	(0.033)	(0.008)		(0.028)	(0.009)	
Energy	0.081	0.067	0.074	0.044		0.002	0.056	
	(0.003)	(0.003)	(0.025)	(0.005)		(0.002)	(0.006)	
Capital	0.077	0.049	0.121	0.066	0.074	0.145	0.074	0.089
	(0.003)	(0.003)	(0.019)	(0.009)	(0.010)	(0.016)	(0.009)	(0.016)
Materials	0.636	0.569	0.603	0.616	0.577	0.653	0.608	0.548
	(0.004)	(0.005)	(0.024)	(0.021)	(0.027)	(0.020)	(0.023)	(0.026)
Cont. Import	0.052	0.037	0.058	0.246	0.270	0.072	0.177	0.182
	(0.006)	(0.007)	(0.042)	(0.052)	(0.061)	(0.032)	(0.043)	(0.062)
γ	0.036	0.007	0.004	0.030	—	0.001	0.026	—
	(0.004)	(0.005)	(0.016)	(0.010)		(0.014)	(0.008)	
ρ	0.723	0.287	0.241	0.822	—	0.271	0.871	—
	(0.004)	(0.009)	(0.023)	(0.031)		(0.016)	(0.027)	
Implied $\frac{\gamma}{1-\rho}$	0.130	0.010	0.005	0.169	—	0.001	0.199	—
Implied θ	13.23	16.38	11.40	3.51	3.12	10.07	4.43	4.01
P-value for over-identification test				0.824	0.759		0.834	0.995
No. of Obs.	33200					45518		

Notes: Standard errors are in parentheses. Columns (1)-(5) use the "Basic Sample" that excludes plants for which the initial capital stock are not reported. Columns (6)-(8) use the "Extended Sample" in which a missing initial capital stock is imputed by a projected initial capital stock based on other reported plant observables. The System GMM estimator in columns (3) and (6) use a lag length of 2 and 3 for instruments in the first-differenced equations and a lag length of 1 in the level equations. The OP/LP estimators in column (5) and (8) specify the stochastic process of ω_t using the third order polynomials in (ω_{t-1}, n_{t-1}).

(7)-(8) again support a substantial impact of an increase in the share of imported intermediates on productivity, finding that a 100 percent decrease in the share of domestic intermediates increases productivity by 17.7 to 27.0 percent.

The positive estimates of γ throughout all columns in Table 4 are suggestive of the positive dynamic effect of an increase in the usage of imported intermediates although, as in the case of the discrete import variable, the estimates from within estimator and GMM estimator are not significant. The OP/LP estimates in columns (4) and (7) indicate that the long-run dynamic effect captured by $\frac{\gamma}{1-\rho}$ are 16.9 and 19.9 percent, respectively. From the estimated coefficients of materials and continuous import variable, we can also compute an estimate of the elasticity of substitution as $\hat{\theta} = 1 + \frac{\hat{\beta}_x}{\hat{\beta}_n}$. Using the OP/LP estimates in Table 4, we obtain point estimates

Table 5: Estimates of Production Function for Food and Metal Industries

	Discrete Import Variable				Continuous Import Variable			
Industry	Food		Metals		Food		Metals	
ω_{it} process	AR(1)	Series	AR(1)	Series	AR(1)	Series	AR(1)	Series
Skilled labor	0.073		0.139		0.072		0.138	
	(0.009)		(0.017)		(0.008)		(0.017)	
Unskilled labor	0.088		0.199		0.091		0.203	
	(0.015)		(0.023)		(0.011)		(0.023)	
Energy	0.070		0.051		0.072		0.050	
	(0.010)		(0.011)		(0.007)		(0.010)	
Capital	0.051	0.050	0.074	0.075	0.051	0.050	0.099	0.092
	(0.009)	(0.009)	(0.032)	(0.033)	(0.009)	(0.010)	(0.036)	(0.036)
Materials	0.664	0.658	0.400	0.378	0.753	0.722	0.434	0.408
	(0.055)	(0.060)	(0.097)	(0.099)	(0.025)	(0.040)	(0.096)	(0.104)
Discrete Import	0.257	0.191	0.243	0.227				
	(0.103)	(0.111)	(0.110)	(0.124)				
Continuous Import					0.370	0.422	0.258	0.234
					(0.167)	(0.209)	(0.157)	(0.181)
γ	0.064	—	0.060	—	0.068	—	0.037	—
	(0.025)		(0.030)		(0.107)		(0.056)	
ρ	0.837	—	0.881	—	0.748	—	0.884	—
	(0.067)		(0.073)		(0.107)		(0.093)	
Implied $\frac{\gamma}{1-\rho}$	0.393	—	0.504	—	0.270	—	0.319	—
Implied θ					3.035	2.711	2.682	2.748
P-value for over-identification test	0.784	0.769	0.809	0.834	0.724	0.871	0.844	0.839
No. of Obs.	12273		3733		12273		3733	

Notes: Standard errors are in parentheses. The estimates are based on the "Basic Sample" that excludes plants for which the initial capital stock are not reported. The OP/LP estimators that specify ω_{it} processes by "Series" use the third order polynomials in (ω_{t-1}, d_{t-1}) for discrete import variable and (ω_{t-1}, n_{t-1}). for continuous import variable.

of θ of 3.12 to 4.43. These estimates are in line with those found by Feenstra, Markusen, and Zeile (1992).

The production functions might be different across industries. We estimate the production functions using the OP/LP estimators for two of the largest 3-digit level industries (ISIC codes): Food (311) and Metals (381). Table 5 presents the results based on the Basic Sample. There are significant differences in the coefficients on skilled/unskilled labor as well as capital across these two industries. Probably reflecting a difference in the sample sizes, the standard errors for Metal Industry are generally larger than those for Food Industry.

Using the discrete import variable, the estimated coefficients on the import variables under the AR(1) specification for ω_{it} process are significantly positive for both industries, indicating a large positive static effect on productivity (25.7 and 24.3 percents for Food and Metals, respectively). When we specify ω_{it} processes by "Series," the estimates are slightly lower but still large (19.1 and 22.7 percents for Food and Metals, respectively); however, they are marginally significant at a 10 percent level. The estimated values of γ for the discrete import variable are positive and significant for both industries, suggesting a positive dynamic effect of the usage of imported materials; the long-run effect of the usage of imported materials on productivity is estimated as 39.3 and 50.4 percents for Food and Metal industries, respectively.

As for the results from using the continuous import variable, all the estimated coefficients for the continuous import variable are positive and of large size, ranging from 23.4 percent to 42.2 percent, but the estimates from Metal industry are not so significant; for Metal industry, the estimate from AR(1) specification is barely significant at a 10 percent level while the estimate from "Series" is not significant even at a 10 percent. Although the relatively large standard errors for Metal industry might be due to its small sample size, the insignificance of the static productivity effect adds some caveat on the positive impact of changing the amount of imported materials on productivity.

The estimated values of γ for the continuous import variable are positive but not significant for both Food and Metal industries. Thus, the evidence for the positive dynamic effect of an increase in the usage of imported intermediates is, at best, weak. Compared to the result for the discrete import variables, the relative insignificance of the dynamic effect of the usage of imported materials in the regression using continuous import variables might indicate that, it is not the intensive margin of *how much* a plant imports but the extensive margin of *whether* a plant imports or not that determines the dynamic effect of importing materials. This could be the case, for instance, if importing intermediates from foreign countries per se—regardless of how much a plant imports—provides an opportunity to learn foreign technologies and thus leads to a positive dynamic effect on productivity.

7 The Decision to Import, Invest, and Exit

Our finding that the use of imported intermediates increases plant-level productivity suggests that understanding plant-level dynamic import decisions is particularly important to understand the impact of trade policy on aggregate productivity. This section empirically investigates the dynamic decisions to import intermediates, invest, and exit. In particular, we examine the issues of the impact of importing intermediates on the resource reallocation through investment and exits, and the existence of the sunk start-up cost of importing intermediates. All the reported estimates are based on the Basic Sample.[27]

Motivated by the decision rules (4)-(6) in the model, we consider the following reduced-form empirical specification for plants' decisions to invest, import and exit:

$$k_{i,t+1} - k_{it} = \alpha_t + \alpha_z Z_{it}^k + \alpha_1 d_{i,t-1} + \alpha_2 k_{it} + \alpha_3 \omega_{i,t-1} + \epsilon_{1,it}, \quad (24)$$

$$d_{it} = 1(\psi_t + \psi_z Z_{it}^d + \psi_1 d_{i,t-1} + \psi_2 k_{it} + \psi_3 \omega_{i,t-1} + \epsilon_{2,it} > 0) \quad (25)$$

$$\chi_{it} = 1(\vartheta_t + \vartheta_z Z_{it}^\chi + \vartheta_1 d_{i,t-1} + \vartheta_2 k_{it} + \vartheta_3 \omega_{i,t-1} + \epsilon_{3,it} > 0). \quad (26)$$

where α_t, ψ_t, and ϑ_t are year-specific constants; $(\epsilon_{1,it}, \epsilon_{2,it}, \epsilon_{3,it})'$ is jointly normally distributed with mean zero and variance-covariance matrix Σ. The logarithm of capital stock at the beginning of year t is denoted by k_{it} and thus $k_{i,t+1} - k_{it}$ in (24) represents the growth rate of capital constructed as investment minus depreciation of capital in year t. The equation (25) specifies a decision to import intermediates, where d_{it} is equal to one if a plant imports in year t and zero otherwise. The equation (26) is a selection equation, where the indicator χ_{it} is equal to one if a plant operates in year t and zero if a plant has exited. We also include the following additional explanatory variables: the average of the logarithm of plant capital stocks at 3-digit industry level, denoted by Z_{it}^k, in the capital accumulation equation (24), the last year's exit rate at 3-digit industry level, Z_{it}^χ, in the selection equation (26) and the last year's fraction of importing plants at 3-digit industry level, Z_{it}^d, in import equation (25). These industry-level explanatory variables are intended to capture the industry-specific shocks that are relevant to each of the three decisions.[28] By simultaneously estimating the exit decision with investment and import decisions, we deal with the important issue of endogenous sample selection.

The parameters in (24)-(26) are estimated by the Simulated Maximum Likelihood (SML) using the GHK recursive simulator which has been found to perform well in Monte Carlo studies for simulating multinomial probit probabilities [c.f., Hajivassiliou, MacFadden, and Ruud (1996);

[27] The results from the Extended Sample are similar to those from the Basic Sample and available upon request from authors.

[28] These additional industry-level explanatory variables provide exclusion restrictions and are especially important for the identification of selection equation (26) separately from import equation (25). Without these exclusion restrictions, its identification comes only from the non-linearity in probit specification since both import and selection equations would have the same set of explanatory variables [c.f., Chapter 17.4 of Wooldridge (2002)].

Geweke, Keane, and Runkle, (1997)].[29] If there are other important sources of unobserved heterogeneity that are not captured by the productivity ω_t (e.g., unobserved heterogeneity in the cost of importing), then the estimates may be biased especially for the coefficients for lagged endogenous variables (i.e., import status and capital stocks). To examine this issue, we also estimated the equations (24)-(26) using the random-effects specification, where $\epsilon_{n,it}$ is decomposed into idiosyncratic terms and plant-specific terms.[30] Note, however, that the presence of unobserved heterogeneity other than productivity ω implied by the random-effects specification violates the conditions for using the OP/LP estimator. We use 200 simulation draws to estimate the model without random effects and 800 simulation draws to estimate the model with random effects.[31]

Table 6 provides the estimates of equations (24)-(26). Standard errors are in parentheses and numbers in blankets represent changes in the probability evaluated at the means of the right-side variables. Panel A presents the estimates of three equations (24)-(26) without random effects and panel B presents the result with random effects. While the two results are qualitatively very similar, the estimates in panel B tend to be smaller than those in panel A.

Plant productivity is significantly and positively related to the capital accumulation, the import probability, and the survival probability. Notably, the inherently more productive plants tend to use imported intermediates; according to the estimate under random effects specification

[29]Let Λ be the unique lower triangular Cholesky decomposition: $\Sigma = \Lambda\Lambda'$. Then, $(\epsilon_{1,it}, \epsilon_{2,it}, \epsilon_{3,it})' \equiv (\lambda_{11}\eta_{1,it}, \lambda_{21}\eta_{1,it} + \lambda_{22}\eta_{2,it}, \lambda_{31}\eta_{1,it} + \lambda_{32}\eta_{2,it} + \lambda_{33}\eta_{3,it})'$, where $\lambda_{m,n}$ is the (m,n) element of Λ, and $\eta_{n,it}$ is independently distributed $N(0,1)$ for all n, i, t. For identification, we assume that $E[\epsilon_{n,it}] = 1$ for $n = 1, 2, 3$. We estimate the model by the Simulated Maximum Likelihood using the GHK simulator. The appendix provides the expression for the likelihood function.

[30]Plant-specific terms are denoted by $u_i = (u_{1i}, u_{2i}, u_{3i})$; using the Cholesky decomposition, $u_i = (\gamma_{11}\xi_{1i}, \gamma_{21}\xi_{1i} + \gamma_{22}\xi_{2i}, \gamma_{31}\xi_{1i} + \gamma_{32}\xi_{2i} + \gamma_{33}\xi_{3i})'$, where ξ_{ni} is independently distributed $N(0,1)$ for all n, i. For identification, since there is no within-plant variation in the exiting equation, we assume that $\gamma_{33} = 0$. The initial conditions problem is handled following the suggestion by Heckman (1981); specifically, we model the presample period's equations with the following exogenous variables: (i) the last year's plant productivity, (ii) the last year's fraction of importing plants at 3-digit industry level, and (iii) the last year's 3-digit industry average of the logarithm of capital stocks. This random effects specification is still restrictive in that it does not allow for persistent transitory shocks (e.g., AR(1)); but the results from the linear probability model reported below in Table 7 suggest that persistent transitory shocks may not be so important once we control for random effects.

[31]Note that the method of Simulated Maximum Likelihood leads to asymptotically biased estimates given a finite number of simulation draws. We examine the robustness of our results with respect to the number of simulation draws by estimating the model with different number of simulation draws. We found that 100 draws and 200 draws lead to very similar results for the model without random effects and that 400 draws and 800 draws lead to very similar results for the model with random effects. The results are available upon request from the authors.

Table 6: The Decision to Import, Invest and Exit

	(A) No Random Effects			(B) Random Effects		
	ΔCapital(t+1)	Import(t)	Stay/Exit(t)	ΔCapital(t+1)	Import(t)	Stay/Exit(t)
Capital(t)	-0.008	0.193	0.076	-0.010	0.302	0.094
	(0.001)	(0.007)	(0.007)	(0.001)	(0.010)	(0.010)
		[0.016]	[0.005]		[0.009]	[0.002]
Import(t-1)	0.017	2.030	0.300	0.014	1.347	0.397
	(0.005)	(0.023)	(0.036)	(0.004)	(0.032)	(0.063)
		[0.581]	[0.045]		[0.249]	[0.034]
Productivity(t-1)	0.090	0.436	0.455	0.086	0.581	0.978
	(0.004)	(0.031)	(0.044)	(0.004)	(0.034)	(0.081)
		[0.035]	[0.031]		[0.018]	[0.024]
Industry-Capital(t)	0.012			0.015		
	(0.002)			(0.002)		
Industry-Import(t-1)		1.538			2.761	
		(0.055)			(0.131)	
		[0.124]			[0.086]	
Industry-Exit(t-1)			-1.824			-2.563
			(0.377)			(0.501)
			[-0.124]			[-0.063]
No. of Observations		29740			29740	

Notes: Standard errors are in parentheses. Numbers in blankets represent changes in the probability evaluated at the means of the right-side variables. All equations include year dummies. "Productivity(t-1)" is computed as $\hat{\omega}_{i,t-1} \equiv \hat{\phi}_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1}) - \hat{\beta}_k k_{i,t-1} - \hat{\beta}_x x_{i,t-1} - \hat{\beta}_d d_{i,t-1}$. The data for $t = 1982, ..., 1996$ is used since an industry exit rate can be computed starting only from 1981. "Industry-Capital(t)" represents the average of the logarithm of plant capital stocks at 3-digit industry level in the beginning of year t. "Industry-Import(t-1)" is a fraction of importing plants at 3-digit industry level in year $t-1$. "Industry-Exit(t-1)" is the exit rate at 3-digit industry level in year $t-1$. The data set excludes plants without information on their initial capital stocks.

of panel B, a 100 percent increase in productivity would lead to a statistically significant 1.8 percent increase in the probability of importing intermediates. Together with the results of the previous section, this indicates that the direction of causality between productivity and import status goes both ways.

The coefficients of the import variable are significantly positive in the capital accumulation equation (24) as well as the survival equation (26). One interpretation of this result is that, through its effect on productivity, the use of imported intermediates increases the current and future profits as well as marginal products of capital, which in turn reduces the exiting probability and increases investment. The positive coefficients of the past import variable in the capital accumulation and the survival equations are indicative of resource reallocation across plants. Non-importing plants tend to accumulate less capital and are more likely to exit than importing plants, implying that resources tend to be reallocated from non-importing plants toward importing plants. Since importing plants are more productive—not only because more productive plants tend to import but also because importing intermediates per se increases plant productivity—resource reallocation from non-importing plants to importing plants may be a potentially important source of aggregate productivity gains.[32]

Another important finding in Table 6 is that the past use of imported intermediates has a positive and significant effect on the probability of importing even after controlling for plant-specific random effects. According to the estimate from no random effects specification of panel A, having imported last year increases the probability of importing this year by 58.1 percent. The estimated impact of imported intermediates declines to 24.9 percent after controlling for random effects but it is still large and significant.[33] This suggests that sunk start-up costs of importing intermediates play an important role in explaining the persistence of importing intermediates.

One major concern is, however, that its validity depends on the distributional assumptions we make; in particular, a poor fit of the plant-specific distribution may lead to an upward-biased

[32] This mechanism is similar to that of Melitz (2003) who emphasizes the role of endogenous export decisions for resource reallocation; the results in this paper suggests that importing intermediate may play a similar reallocative role as exporting.

[33] This is suggestive of other important sources of unobserved heterogeneity that are not captured by the productivity ω_t (e.g., unobserved heterogeneity in the cost of importing). Since the presence of unobserved heterogeneity other than productivity ω violates the conditions for using the OP/LP estimator, this provides an important caveat about the empirical results based on the OP/LP estimator in the previous section.

Table 7: The Decision to Import: Linear Probability Model

	(1) OLS	(2) Within	(3) GMM	(4) GMM-AR	(5) GMM	(6) GMM
Import(t-1)	0.6915	0.3924	0.2681	0.3383	0.2827	0.2829
	(0.0050)	(0.0082)	(0.0113)	(0.0356)	(0.0109)	(0.0108)
Capital(t)	0.0115	0.0079	0.0024	0.0085	-0.0157	-0.0199
	(0.0006)	(0.0007)	(0.0027)	(0.0028)	(0.0071)	(0.0064)
Productivity(t-1)	0.0098	0.0107	0.0073	0.0138	0.1215	0.1085
	(0.0017)	(0.0015)	(0.0041)	(0.0050)	(0.0324)	(0.0286)
$(\text{Capital(t)})^2$	—	—	—	—	0.0036	0.0039
					(0.0007)	(0.0006)
$(\text{Productivity(t-1)})^2$	—	—	—	—	-0.0087	-0.0076
					(0.0025)	(0.0022)
Wage(t-1)	—	—	—	—	—	0.0011
						(0.0033)
Nonproduction/total employment(t-1)	—	—	—	—	—	0.0142
						(0.0260)
AR(1) coefficient	—	—	—	0.0635	—	—
				(0.0479)		
No. of Observations	33200					

Notes: Dependent variable is discrete import variable. Standard errors are in parentheses. All equations include year dummies. The OLS estimator in column (1) includes 4-digit industry dummies. The System GMM estimator in columns (3) use a lag length of 2 and 3 for instruments in the first-differenced equations and a lag length of 1 in the level equations. The data set excludes plants without information on their initial capital stocks. The parameter ρ represents the coefficient of AR(1) error structure.

coefficient on the lagged import status in import equation because, in such a case, plant-specific effects would not be sufficiently controlled. To examine further the issue of sunk start-up costs of importing, we also estimate the import equation by using the linear probability model. While the linear probability model has disadvantage that the predicted probabilities may lie outside of 0 and 1, the linear probability model estimated by GMM is valid under much weaker distributional assumption on plant-specific effects than the probit model. We also examine the sensitivity of the estimated coefficient on the past import status with respect to inclusions of other potentially important plant-level variables.

Columns (1)-(4) in Table 7 present the estimates of the linear probability model of import decisions from the OLS, the Within-Group, the System GMM estimators, and the System GMM estimators with AR(1) error structure, respectively. Across all four estimators, the coefficient on the past import status is positive and significant. Specifically, the System GMM estimator in column (3), which deals with the issues of both the presence of plant-specific effect and the endogeneity of past import status, indicates that having imported last year increases the probability of importing this year by 26.8 percent, providing further evidence for the existence of sunk costs of importing. Allowing for AR(1) error shocks in column (4), we have a larger estimate of 33.8 percent while the estimate of AR(1) coefficient is not significantly different from zero. To further examine the robustness of the results with respect to misspecification, we estimated the linear probability model with the quadratic terms for capital and productivity in column (5) and with additional explanatory variables of past wage and the ratio of nonproduction employment to total employment in column (6); even with additional controls, the coefficients on the past import status in columns (5)-(6) are significant and of the similar magnitude to that of the system GMM estimator in column (3).

8 Conclusion

The results in this paper demonstrate significant plant-level evidence that imported intermediates improve a plant's productivity. We found that by switching from being a non-importer to an importer of foreign intermediates a plant can immediately improve productivity by 2.3 to 22.0 percent. We also found some evidence for a positive dynamic effect of the use of imported materials. Intermediate imports, therefore, allow plants to adopt technology from abroad and substantially benefit from foreign research and development. This result alone is important for both government policy and plant production strategy.

The paper also provides one of the first detailed empirical analyses in the literature on import decision and its interaction with investment and exiting decisions at plant-level. Estimating plant's dynamic decisions to import, invest, and exit, we found that inherently productive plants are more likely to import intermediates and, in turn, importing intermediates—probably through its positive productivity effect—increases plant's investment and decreases the probability of exiting. The result indicates that intermediate import may play an important role for aggregate productivity growth not only because it has a direct effect on plant productivity but also because it induces resource reallocation from low productivity plants to high productivity plants. We also found the evidence for a significant sunk start-up cost of importing intermediates.

The empirical findings of this paper shed new light on the issue of how trade policy affects aggregate productivity and suggest directions for future research. Recent developments in trade theory, focusing on heterogeneous plants, suggest that understanding the plant-level export response to trade policy is a crucial factor in understanding its impact on aggregate productivity (e.g., Melitz, 2003; Bernard, Eaton, Jensen and Kortum, 2003). Our results imply that understanding the plant-level *import* decisions may be particularly important in understanding the impact of trade policy on aggregate productivity and welfare. The exposure to trade may induce the more productive plants to start importing intermediates and subsequently lead to resource reallocation from productive importers to less productive non-importers.

References

[1] Arellano, M., and S.R. Bond (1991) "Some Test of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58: 277-297.

[2] Baily, Martin N., Hulten, Charles, and Campbell, David (1992) "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity: Microeconomics*, 67-119.

[3] Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003) "Plants and Productivity in International Trade," *American Economic Review*, 93(4): 1268-1290.

[4] Bernard, Andrew B. and J. Bradford Jensen (1999) "Exceptional Exporter Performance: Cause, Effect, and or Both?" *Journal of International Economics*, 47: 1-25.

[5] Bernard, Andrew B. and J. Bradford Jensen (2004) "Why Some Firms Export," *The Review of Economics and Statistics*, 86(2): 561-569.

[6] Blundell, R.W. and S.R. Bond (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 68: 29-51.

[7] Blundell, R.W. and S.R. Bond (2000) "GMM Estimation with Persistent Panel Data: an Application to Production Functions," *Econometric Reviews*, 19(3): 321-340.

[8] Chamberlain, G. (1982) "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, 18: 5-46.

[9] Clerides, Sofronis, Lach, Saul, and James R. Tybout (1998) "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *Quarterly Journal of Economics* 113: 903-947.

[10] Coe, David T. and Elhanan Helpman (1995) "International R&D Spillovers," *European Economic Review*, 39: 859-887.

[11] Coe, David T., Elhanan Helpman, and Alexander W. Hoffmaister (1997) "North-South R&D Spillovers," *Economic Journal*, 107: 134-149.

[12] Feenstra, Robert C. (1994) "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1): 157-177.

[13] Feenstra, Robert C., James R. Markusen, and William Zeile (1992) "Accounting for Growth with New Inputs: Theory and Evidence," *American Economic Review*, 82(2): 415-421.

[14] Fernandes, Ana M. (2003) "Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries," mimeo.

[15] Geweke, John F., Michael P. Keane, and David E. Runkle (1997) "Statistical Inference in the Multinomial Multiperiod Probit Model," *Journal of Econometrics*, 80: 125-165.

[16] Grilicches, Z. and J. Mairesse (1998) "Production Functions: The Search for Identification," in S.Stöm e., *Econometrics and Economic Theory in the 20th Century: The Ragnar Frish Centennial Symposium* Cambridge, England: Cambridge University Press, pp 169-203.

[17] Grossman, G.M. and E. Helpman (1991) *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press.

[18] Hajivassiliou, Vassilis, Daniel MacFadden, and Paul Ruud (1996) "Simulation of Multivariate Normal Rectangle Probabilities and their Derivatives Theoretical and Computational Results," *Journal of Econometrics*, 72: 85-134.

[19] Heckman, J.J. (1981) "The Incidental Parameters Problem and the Problem of Initial COnditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C.F. Manski and D. McFadden. Cambridge, MA: MIT Press, pp 179-195.

[20] Horowitz, Joel L. (2001) "The Bootstrap," in J.J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Vol. 5, Oxford: Elsevier Science, pp 3159-3228.

[21] Keller, Wolfgang. (2002) "Trade and the Transmission of Technology," *Journal of Economic Growth*, (1): 5-24.

[22] Lui, L. (1993) "Entry, Exit, and Learning in the Chilean Manufacturing Sector," *Journal of Development Economics*, 42: 217-242.

[23] Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70:317-341.

[24] Marschak, J. and Andrews, W.H. (1994) "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12(3,4): 143-205.

[25] Melitz, Mark J.(2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71: 1695-1725.

[26] Muendler, Marc-Andreas (2004) "Trade, Technology, and Productivity: A Study of Brazilian Manufacturers, 1986-1998," UCSD, mimeo.

[27] Mulkay, Benoit, Hall, Bronwyn H., and Jacques Mairesse (2000) "Firm Level Investment and R&D in France and the United States: A Comparison," National Bureau of Eocnomic Research, Working Paper No. 8038.

[28] Olley, S. and Pakes, A. (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 65(1): 292-332.

[29] Pavcnik, Nina (2002) "Trade Liberalization, Exit, and Productivity Improvements: Evidence From Chilean Plants," Review of Economic Studies, 69(1): 245-276.

[30] Roberts, Mark and Tybout, Robert (1997) "Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs" *American Economic Review*, 87: 545-564.

[31] Van Biesebroeck, Johannes (2003) "Revisiting some Productivity Debates," *NBER Working Paper* 10065.

[32] Wooldridge, Jeffery M. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

9 Appendix

In this appendix, we derive the likelihood function to estimate the model for plants' decisions to invest, import and exit, (24)-(26), which are rewritten as:

$$\begin{aligned} k_{i,t+1} - k_{it} &= X_{it}^k \alpha + \epsilon_{1,it}, \\ d_{it} &= 1(X_{it}^d \psi + \epsilon_{2,it} > 0), \\ \chi_{it} &= 1(X_{it}^\chi \vartheta + \epsilon_{3,it} > 0), \end{aligned}$$

where

$$\begin{aligned} X_{it}^k \alpha &= \alpha_t + \alpha_z Z_{it}^k + \alpha_1 d_{i,t-1} + \alpha_2 k_{it} + \alpha_3 \omega_{i,t-1} \\ X_{it}^d \psi &= \psi_t + \psi_z Z_{it}^d + \psi_1 d_{i,t-1} + \psi_2 k_{it} + \psi_3 \omega_{i,t-1} \\ X_{it}^\chi \vartheta &= \vartheta_t + \vartheta_z Z_{it}^\chi + \vartheta_1 d_{i,t-1} + \vartheta_2 k_{it} + \vartheta_3 \omega_{i,t-1}. \end{aligned}$$

See the main text for the variable definitions.

First, we derive the likelihood function for the model without random effects. We assume that $\epsilon_{it} = (\epsilon_{1,it}, \epsilon_{2,it}, \epsilon_{3,it})'$ is jointly normally distributed with mean zero and variance-covariance matrix Σ. Let Λ be the unique lower triangular Cholesky decomposition: $\Sigma = \Lambda\Lambda'$. Then, $\epsilon_{it} \equiv (\lambda_{11}\eta_{1,it}, \lambda_{21}\eta_{1,it} + \lambda_{22}\eta_{2,it}, \lambda_{31}\eta_{1,it} + \lambda_{32}\eta_{2,it} + \lambda_{33}\eta_{3,it})'$, where $\lambda_{m,n}$ is the (m,n) element of Λ, and $\eta_{n,it}$ is independently distributed $N(0,1)$ for all n, i, t. For identification, we assume that $E[\epsilon_{n,it}] = 1$ for $n = 1,2,3$. Denote the parameter vector to be estimated by β. Let

$$\hat{\eta}_{1,it}(\beta) = \frac{(k_{i,t+1} - k_{it}) - X_{it}^k \alpha}{\lambda_{11}}.$$

Then, the likelihood contribution from the observation (i,t) is given by

$$L_{it}(\beta) = \begin{cases} P(\chi_{it} = 0) & \text{for } \chi_{it} = 0, \text{ and} \\ (1/\lambda_{11})\phi(\hat{\eta}_{1,it}(\beta))P(d_{it}, \chi_{it} = 1 | \hat{\eta}_{1,it}(\beta)) & \text{for } \chi_{it} = 1 \end{cases}$$

where

$$P(\chi_{it} = 0) = 1 - \Phi(X_{it}^{\chi}\vartheta),$$

and

$$P(d_{it}, \chi_{it} = 1 | \hat{\eta}_{1,it}) = \tag{27}$$
$$\int\int 1(X_{it}^d \psi + \lambda_{21}\hat{\eta}_{1,it} + \lambda_{22}\eta_2' > 0) 1(X_{it}^{\chi}\vartheta + \lambda_{31}\hat{\eta}_{1,it} + \lambda_{32}\eta_2' + \lambda_{33}\eta_3' > 0)\phi(\eta_2')\phi(\eta_3')d\eta_2' d\eta_3'.$$

The parameter vector β is estimated by maximizing the log likelihood function

$$\max_{\beta} \sum_{i=1}^{N} \sum_{t=T_{i0}}^{T_{i1}} \ln L_{it}(\beta).$$

For the model with random effects, $\epsilon_{it} = (\epsilon_{1,it}, \epsilon_{2,it}, \epsilon_{3,it})'$ is decomposed into idiosyncratic terms and plant-specific terms. Plant-specific terms are denoted by $u_i = (u_{1i}, u_{2i}, u_{3i})'$ so that $\epsilon_{n,it} = u_{ni} + \eta_{n,it}$ for $n = 1,2,3$. Using the Cholesky decomposition, $u_i = (\gamma_{11}\xi_{1i}, \gamma_{21}\xi_{1i} + \gamma_{22}\xi_{2i}, \gamma_{31}\xi_{1i} + \gamma_{32}\xi_{2i} + \gamma_{33}\xi_{3i})'$, where ξ_{ni} is independently distributed $N(0,1)$ for all n, i. Since there is no within-plant variation in the exiting equation, for identification, we assume that $\gamma_{33} = 0$; this means that we can drop ξ_{3i} from our consideration. Denote $\xi_i = (\xi_{1i}, \xi_{2i})'$. For identification, we assume $E[\epsilon_{n,it}] = 1$ for $n = 1,2,3$.

Let

$$\hat{\eta}_{1,it}(\beta; \xi) = \frac{(k_{i,t+1} - k_{it}) - X_{it}^k \alpha - \gamma_{11}\xi_1}{\lambda_{11}}.$$

Then, the likelihood contribution from the observation (i,t), conditioned on the value of ξ, is given by

$$L_{it}(\beta;\xi) = \begin{cases} P(\chi_{it}=0|\xi) & \text{for } \chi_{it}=0, \text{ and} \\ (1/\lambda_{11})\phi(\hat{\eta}_{1,it}(\beta;\xi))P(d_{it},\chi_{it}=1|\hat{\eta}_{1,it}(\beta;\xi),\xi) & \text{for } \chi_{it}=1 \end{cases}$$

where

$$P(\chi_{it}=0|\xi) = 1-\Phi(X_{it}^{\chi}\vartheta+\gamma_{31}\xi_1+\gamma_{32}\xi_2),$$

and

$$P(d_{it},\chi_{it}=1|\hat{\eta}_{1,it},\xi) = \int\int 1(X_{it}^{d}\psi+\lambda_{21}\hat{\eta}_{1,it}+\lambda_{22}\eta_2'+\gamma_{21}\xi_1+\gamma_{22}\xi_2>0) \quad (28)$$
$$1(X_{it}^{\chi}\vartheta+\lambda_{31}\hat{\eta}_{1,it}+\lambda_{32}\eta_2'+\lambda_{33}\eta_3'+\gamma_{31}\xi_1+\gamma_{32}\xi_2>0)\phi(\eta_2')\phi(\eta_3')d\eta_2'd\eta_3'.$$

To deal with the initial conditions problem, we follow the suggestion by Heckman (1981); specifically, we model the initial year's equations for investment and discrete import decisions with the following variables:[34] (i) the last year's plant productivity, (ii) the last year's fraction of importing plants at 3-digit industry level, and (iii) the last year's 3-digit industry average of the logarithm of capital stocks. Denote the vector of those three variables by Z_{it}. The likelihood contribution from the initial year's observation for plant i, conditioned on a plant-specific effect ξ, is computed as:

$$L_{it}^{p}(\beta;\xi) = (1/\lambda_{11}^{p})\phi(\hat{\eta}_{1,it}^{p}(\beta;\xi))P^{p}(d_{it},\chi_{it}=1|\hat{\eta}_{1,it}^{p}(\beta;\xi),\xi)$$

where $\hat{\eta}_{1,it}^{p}(\beta;\xi) = \frac{(k_{i,t+1}-k_{it})-Z_{it}\alpha^{p}-\rho_1^{p}\gamma_{11}\xi_1}{\lambda_{11}^{p}}$ and

$$P^{p}(d_{it},\chi_{it}=1|\hat{\eta}_{1,it}^{p},\xi) = \int 1(Z_{it}\psi^{p}+\lambda_{21}^{p}\hat{\eta}_{1,it}^{p}+\lambda_{22}^{p}\eta_2'+\rho_2^{p}(\gamma_{21}\xi_1+\gamma_{22}\xi_2)>0)\phi(\eta_2')d\eta_2'.$$

Note that the parameter vector in the initial year's equations $(\alpha^{p},\psi^{p},\lambda_{11}^{p},\lambda_{21}^{p},\lambda_{22}^{p})$ is different from $(\alpha,\psi,\lambda_{11},\lambda_{21},\lambda_{22})$. We introduce the parameters ρ_1^{p} and ρ_2^{p} as suggested by Heckman (1981).

Then, the plant i's likelihood contribution is derived by integrating out the plant-specific effect as:

$$L_i(\beta) = \int\int L_{i,T_{i0}}^{p}(\beta;\xi')\prod_{t=T_{i0}+1}^{T_{i1}} L_{it}(\beta;\xi')\phi(\xi_1')\phi(\xi_2')d\xi_1'd\xi_2'. \quad (29)$$

[34] We do not observe the exit decision at the initial year.

The parameter vector β is estimated by maximizing the log likelihood function

$$\max_{\beta} \sum_{i=1}^{N} \ln L_i(\beta).$$

To evaluate the integrals of the right hand side of (28)-(29), we use the GHK simulator. Note that the method of Simulated Maximum Likelihood leads to asymptotically biased estimates given a finite number of simulation draws. Estimating the model with serially correlated errors accurately may require a large number of simulation draws as some Monte Carlo studies suggest [c.f., Geweke, Keane, and Runkle (1997) and Lee (1997)]. We examine the robustness of our results with respect to the number of simulation draws by estimating the model with different number of simulation draws. As shown in Table (8)-(9), we found that 100 draws and 200 draws lead to very similar results for the model without random effects and that 400 draws and 800 draws lead to very similar results for the model with random effects.

Table 8: Estimates of Import/Exit/Investment Decision (No Random Effects, Basic Sample)

Simulation Draws	R=20	R=100	R=200
Dependent Variable: $\Delta k(t+1)$			
d_{t-1}	0.0170	0.0170	0.0170
	(0.0048)	(0.0048)	(0.0048)
k_t	-0.0084	-0.0084	-0.0084
	(0.0013)	(0.0013)	(0.0013)
ω_{t-1}	0.0903	0.0903	0.0903
	(0.0041)	(0.0042)	(0.0042)
Industry-Capital(t)	0.0121	0.0121	0.0121
	(0.0017)	(0.0017)	(0.0017)
Dependent Variable: $d(t)$			
d_{t-1}	2.0995	2.0997	2.0997
	(0.0192)	(0.0221)	(0.0233)
	[0.5807]	[0.5807]	[0.5807]
k_t	0.1929	0.1930	0.1930
	(0.0058)	(0.0064)	(0.0067)
	[0.0156]	[0.0156]	[0.0156]
ω_{t-1}	0.4344	0.4364]	0.4361
	(0.0290)	(0.0300)	(0.0305)
	[0.0350]	[0.0352]	[0.0352]
Industry-Import(t-1)	1.5399	1.5382	1.5381
	(0.0548)	(0.0548)	(0.0548)
	[0.1242]	[0.1240]	[0.1240]
Dependent Variable: $\chi(t)$			
d_{t-1}	0.2999	0.3001	0.3001
	(0.0363)	(0.0364)	(0.0364)
	[0.0448]	[0.0448]	[0.0448]
k_t	0.0763	0.0763	0.0763
	(0.0065)	(0.0065)	(0.0065)
	[0.0052]	[0.0052]	[0.0052]
ω_{t-1}	0.4527	0.4547	0.4550
	(0.0441)	(0.0440)	(0.0440)
	[0.0306]	[0.0308]	[0.0308]
Industry-Exit(t-1)	-1.8355	-1.8247	-1.8235
	(0.3769)	(0.3768)	(0.3768)
	[-0.1242]	[-0.1236]	[-0.1236]
Variance-Covariance Matrix			
λ_{11}	0.3114	0.3114	0.3114
	(0.0005)	(0.0005)	(0.0005)
λ_{21}	0.0679	0.0679	0.0679
	(0.0122)	(0.0122)	(0.0122)
λ_{31}	-0.0140	-0.0141	-0.0141
	(0.1110)	(0.1110)	(0.1110)
λ_{32}	-0.0004	-0.0016	-0.0017
	(0.1932)	(0.3077)	(0.3471)
Log-Likelihood	-40509.1581	-40509.1457	-40509.1447
No. of Observations		29740	

Table 9: Estimates of Import/Exit/Investment Decision (No Random Effects, Extended Sample)

Simulation Draws	R=20	R=100	R=200
Dependent Variable: $\Delta k(t+1)$			
d_{t-1}	0.0126	0.0126	0.0126
	(0.0041)	(0.0041)	(0.0041)
k_t	-0.0060	-0.0060	-0.0060
	(0.0011)	(0.0012)	(0.0012)
ω_{t-1}	0.0773	0.0773	0.0773
	(0.0034)	(0.0034)	(0.0034)
Industry-Capital(t)	0.0203	0.0203	0.0203
	(0.0013)	(0.0013)	(0.0013)
Dependent Variable: $d(t)$			
d_{t-1}	2.1150	2.1150	2.1150
	(0.0182)	(0.0209)	(0.0217)
	[0.5743]	[0.5744]	[0.5744]
k_t	0.1959	0.1958	0.1958
	(0.0053)	(0.0056)	(0.0057)
	[0.0123]	[0.0123]	[0.0123]
ω_{t-1}	0.4262	0.4258	0.4258
	(0.0252)	(0.0269)	(0.0274)
	[0.0268]	[0.0268]	[0.0268]
Industry-Import(t-1)	1.4654	1.4647	1.4647
	(0.0518)	(0.0518)	(0.0518)
	[0.0920]	[0.0920]	[0.0920]
Dependent Variable: $\chi(t)$			
d_{t-1}	0.3234	0.3233	0.3233
	(0.0336)	(0.0337)	(0.0337)
	[0.0596]	[0.0596]	[0.0596]
k_t	0.0752	0.0752	0.0752
	(0.0063)	(0.0063)	(0.0063)
	[0.0079]	[0.0079]	[0.0079]
ω_{t-1}	0.3993	0.3998	0.3998
	(0.0379)	(0.0380)	(0.0380)
	[0.0418]	[0.0418]	[0.0418]
Industry-Exit(t-1)	-1.8691	-1.8605	-1.8606
	(0.3227)	(0.3227)	(0.3227)
	[-0.1956]	[-0.1947]	[-0.1947]
Variance-Covariance Matrix			
λ_{11}	0.2907	0.2907	0.2907
	(0.0005)	(0.0005)	(0.0005)
λ_{21}	0.0691	0.0691	0.0691
	(0.0111)	(0.0111)	(0.0111)
λ_{31}	-0.0225	-0.0225	-0.0225
	(0.0959)	(0.0959)	(0.0959)
λ_{32}	-0.0005	-0.0007	-0.0008
	(0.1623)	(0.2517)	(0.2769)
Log-Likelihood	-44904.2160	-44904.2148	-44904.2144
No. of Observations		29740	

Table 10: Estimates of Import/Exit/Investment Decision (Random Effects, Basic Sample)

	R=200	R=400	R=800
Dependent Variable: $\Delta k(t+1)$			
d_{t-1}	0.015	0.015	0.014
	(0.004)	(0.004)	(0.004)
k_t	-0.010	-0.010	-0.010
	(0.001)	(0.001)	(0.001)
ω_{t-1}	0.085	0.086	0.086
	(0.004)	(0.004)	(0.004)
Industry-Capital(t)	0.015	0.015	0.015
	(0.002)	(0.002)	(0.002)
Dependent Variable: $d(t)$			
d_{t-1}	1.357	1.347	1.347
	(0.033)	(0.032)	(0.032)
	[0.260]	[0.249]	[0.249]
k_t	0.288	0.302	0.302
	(0.010)	(0.010)	(0.010)
	[0.010]	[0.009]	[0.009]
ω_{t-1}	0.593	0.582	0.581
	(0.032)	(0.033)	(0.034)
	[0.020]	[0.018]	[0.018]
Industry-Import(t-1)	2.599	2.761	2.761
	(0.130)	(0.131)	(0.131)
	[0.088]	[0.086]	[0.086]
Dependent Variable: $\chi(t)$			
d_{t-1}	0.439	0.397	0.397
	(0.059)	(0.062)	(0.063)
	[0.045]	[0.034]	[0.034]
k_t	0.082	0.094	0.094
	(0.009)	(0.010)	(0.010)
	[0.003]	[0.002]	[0.002]
ω_{t-1}	0.724	0.978	0.978
	(0.075)	(0.080)	(0.081)
	[0.026]	[0.024]	[0.024]
Industry-Exit(t-1)	-1.980	-2.563	-2.563
	(0.464)	(0.498)	(0.501)
	[-0.072]	[-0.063]	[-0.063]
Variance-Covariance Matrix			
λ_{11}	0.305	0.305	0.305
	(0.001)	(0.001)	(0.001)
λ_{21}	0.067	0.063	0.063
	(0.014)	(0.015)	(0.015)
λ_{31}	-0.019	-0.023	-0.022
	(0.093)	(0.084)	(0.086)
λ_{32}	0.000	0.000	0.000
	(0.266)	(0.262)	(0.281)
γ_{11}	0.019	0.019	0.019
	(0.003)	(0.003)	(0.003)
γ_{21}	0.902	0.917	0.917
	(0.033)	(0.034)	(0.034)
γ_{22}	0.000	0.001	0.000
	(0.084)	(0.105)	(0.217)
γ_{31}	0.006	0.061	0.062
	(0.058)	(0.088)	(0.110)
γ_{32}	0.58	0.758	0.758
	(0.065)	(0.073)	(0.073)
Log-Likelihood	-42740.01	-42710.75	-42708.37
No. of Observations		29740	

Table 11: Descriptive Statistics in 1980 (Extended Sample)

	Output	Capital	Labor	Energy	Interme-diates	Import Ratios	Output/ Workers	No. of Plants
All	98.33	40.76	54.33	0.64	50.65	0.07	1.18	4502
Plants	(468.41)	(233.12)	(127.34)	(7.65)	(235.82)	(0.18)	(1.93)	
Importing	442.28	180.44	168.44	2.91	203.11	0.37	2.58	308
Plants	1003.26)	(430.15)	(231.34)	(14.46)	(404.55)	(0.26)	(3.69)	
Non-Importing	22.03	10.91	25.93	0.08	13.02	0.00	0.75	2626
Plants	(57.58)	(74.60)	(29.39)	(0.57)	(32.36)	(0.00)	(0.84)	
Switchers	158.55	63.32	79.48	1.13	83.72	0.13	1.62	1568
	(625.17)	(323.55)	(173.49)	(11.18)	(343.34)	(0.22)	(2.43)	
Survivors	201.20	77.06	84.34	1.46	99.60	0.11	1.65	1460
	(784.12)	(377.72)	(192.22)	(12.87)	(388.51)	(0.22)	(2.64)	
Quitters	48.95	23.34	39.93	0.25	27.16	0.05	0.95	3042
	(149.14)	(105.10)	(75.05)	(2.59)	(90.49)	(0.15)	(1.41)	

Notes: Standard errors are in parentheses. The statistics are based on the extended sample, where a missing initial capital stock is imputed by a projected initial capital stock based on other reported plant observables. "Importing Plants" are plants that continuously imported foreign intermediates in the sample. "Non-Importing Plants" are plants that never imported foreign intermediates in the sample. "Switchers" are plants that switched their import status in the sample. "Survivors" are plants that did not exit during the sample period (1980-1996) while "Quitters" exit during the sample period. "Output," "Capital," "Energy," and "Intermediates" are measured in millions of 1980 pesos. "Labor" is the number of workers. "Import Ratios" are the ratios of imported intermediate materials to total intermediate materials.

Table 12: Transition Probability of Import Status and Exit (Extended Sample)

Year t status	No Imports			Imports		
Year $t+1$ status	No Imports	Imports	Exit	No Imports	Imports	Exit
1981-1985 ave.	0.832	0.052	0.116	0.169	0.785	0.046
1986-1990 ave.	0.877	0.052	0.071	0.176	0.801	0.023
1991-1995 ave.	0.866	0.064	0.070	0.126	0.852	0.023
1981-1995 ave.	0.858	0.056	0.086	0.157	0.813	0.031

Notes: The statistics are based on the extended sample, where a missing initial capital stock is imputed by a projected initial capital stock based on other reported plant observables.

Table 13: The Decision to Import: Linear Probability Model (Extended Sample)

	(1) OLS	(2) Within	(3) GMM	(4) GMM-AR	(5) GMM	(6) GMM
Import(t-1)	0.7017	0.3993	0.2725	0.3568	0.2841	0.2846
	(0.0046)	(0.0077)	(0.0097)	(0.0296)	(0.0093)	(0.0092)
Capital(t)	0.0102	0.0076	0.0037	0.0061	-0.0197	-0.0243
	(0.0005)	(0.0006)	(0.0022)	(0.0024)	(0.0052)	(0.0045)
Productivity(t-1)	0.0104	0.0109	0.0071	0.0147	0.0962	0.0775
	(0.0016)	(0.0013)	(0.0038)	(0.0039)	(0.0239)	(0.0209)
$(\text{Capital(t)})^2$	—	—	—	—	0.0037	0.0041
					(0.0005)	(0.0004)
$(\text{Productivity(t-1)})^2$	—	—	—	—	-0.0070	-0.0057
					(0.0018)	(0.0016)
Wage(t-1)	—	—	—	—	—	-0.0042
						(0.0026)
Nonproduction/total employment(t-1)	—	—	—	—	—	0.0219
						(0.0207)
AR(1) coefficient	—	—	—	0.0523	—	—
				(0.0411)		
No. of Observations	45518					

Notes: The regressions are based on the extended sample, where a missing initial capital stock is imputed by a projected initial capital stock based on other reported plant observables. Standard errors are in parentheses. All equations include year dummies. The OLS estimator in column (1) includes 4-digit industry dummies. The System GMM estimator in columns (3) use a lag length of 2 and 3 for instruments in the first-differenced equations and a lag length of 1 in the level equations.