Supplemental Appendix for "Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence" (NOT FOR PUBLICATION)

Hiroyuki Kasahara[*]and Joel Rodrigue[†]

This version: September 3, 2007

1 Estimation Procedures: Selection and Adjustment Costs

1.1 The issue of selection in the LP approach

In this section we outline how we control for endogeneous selection while using LP intermediate proxy approach. The idea is essentially the same as the one used in Olley and Pakes (1996); namely, we first identify the state variables that are relevant for endogneous exiting decisions and approximate the survival probabilities using the polynomials in the observable variables. Then, we can control for the endogenous exiting decision by including the polinomials in the survival probabilities when the moment conditions are constructed.

First, the state variables that are relevant for the plant exit decision are the predetermined level of capital k_{it} and the past import decision $d_{i,t-1}$. The model in section 2 implies that a plant chooses to continue to produce if the current realization of productivity term ω_{it} is higher than the threshold value $\underline{\omega}_t(k_{it}, d_{i,t-1})$. One might think that the intermediate proxy approach is not applicable to control for the selection bias because we cannot "recover" ω_{it} from observables given that we do not observe the current period intermediates if the plant chooses to exit. Note, however, that ω_{it} follows the first order Markov process $\omega_{it} = \xi_t + \gamma d_{i,t-1} + \omega_{i,t-1} + u_{it}$ (equation (8) in the main text) and, thus, it is possible to approximate ω_{it} using the observable variables $(d_{i,t-1}, \omega_{i,t-1})$, where $\omega_{i,t-1}$, in turn, can be proxied by the past value of intermediates, the past capital, and the past import decision so that $\omega_{i,t-1} = \omega^*_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1})$ (equation (10) in the main text).[1]

[*]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]The linearity assumption can be relaxed. For instance, suppose that ω_{it} follows the stochastic process ω_{it} =

Specifically, the plant chooses to stay if $\omega_{it} \geq \underline{\omega}_t(k_{it}, d_{i,t-1})$, or using $\omega_{it} = \xi_t + \gamma d_{i,t-1} + \omega_{i,t-1} + u_{it}$ and $\omega_{i,t-1} = \omega^*_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1})$, the plant stays if

$$
\begin{aligned}
u_{it} &\geq \underline{\omega}_t(k_{it}, d_{i,t-1}) - \xi_t - \gamma d_{i,t-1} - \omega^*_{t-1}(x_{i,t-1}, k_{i,t-1}, d_{i,t-1}) \\
&\equiv \underline{u}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1}),
\end{aligned}
$$

where $\underline{u}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})$ is the threshold value of u_{it} that induces a plant to exit at t. Since a plant continues in operation if $u_{it} \geq \underline{u}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})$, the survival probabilities are given by

$$
\begin{aligned}
Pr\{\chi_{it} = 1 | \underline{u}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})\} &= 1 - F_u(\underline{u}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})) \\
&= \mathcal{P}_t(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1}) \\
&\equiv P_{it}, \qquad (1)
\end{aligned}
$$

where $F_u(\cdot)$ is the cumulative distribution of u_{it}. Equation (1) corresponds to equation (10) in Olley and Pakes (1996). We can approximately estimate (1) by probit using the polinomials in $(k_{it}, k_{i,t-1}, d_{i,t-1}, x_{i,t-1})$ as explanatory variables for the survival decision.

Once the survival probabilities are estimated in terms of the observables, the rest of the procedure for controlling for the selection bias is essentially the same as that of the OP approach. By inverting (1), 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

$$
E[\omega_{it} | \omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1] = \xi_t + \gamma d_{i,t-1} + \rho \omega_{i,t-1} + E[u_{it} | u_{it} \geq \underline{u}^*(P_{it})]. \qquad (2)
$$

Here, the term $E[u_{it} | u_{it} \geq \underline{u}^*(P_{it})]$ controls for the selection bias. For instance, if we know u_{it} is normally distributed, this term becomes the inverse Mill's ratio.

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 (1) which approximates the term $E[u_{it} | u_{it} \geq \underline{u}^*(P_{it})]$. Here, $\hat{\phi}_t(\cdot)$ 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}) while the survival probability is estimated 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. In estimating (2), we also allow for year-specific constant terms, ξ_t, to control for the year-specific aggregate productivity shocks.

Establishing the procedure to estimate $E[\omega_{it} | \omega_{i,t-1}, d_{i,t-1}, \chi_{it} = 1]$, we consistently estimate β_k, β_x,

---

$h_t(d_{i,t-1}, \omega_{i,t-1}, u_{it})$. Even in this case, we may apply the similar logic to control for the selection bias as long as $h_t(d_{i,t-1}, \omega_{i,t-1}, u_{it})$ is strictly increasing in u_{it}.

and β_d while controlling for the selection bias as discussed in section 4.1 of the main text.

1.2 Alternative Estimators: the Within-Groups and the GMM estimators

To address the simultaneity issue, we also consider the following two alternatives: the within-groups estimator and the system GMM estimator.

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 their slow adjustment over time. The within-estimator may lead to imprecise estimates especially for capital and imported intermediates. This issue becomes more pronounced when there is idiosyncratic measurement error in inputs; within-transformation lowers signal to noise ratio and magnifies the bias induced by measurement errors (cf., Griliches and Hausman, 1986).

In order to control for simultaneity in panel data, Blundell and Bond (1998, 2000) propose the system GMM estimator by extending the first-differenced GMM estimator (cf., Arellano and Bond, 1991). Consider the equation (7) in the main text with the following stochastic process of ω_{it}:

$$\omega_{it} = \xi_t + \gamma d_{i,t-1} + \rho\omega_{i,t-1} + \alpha_i + v_{it}, \tag{3}$$

where ξ_t is a year-specific effect, α_i is a plant-specific effect, v_{it} is an i.i.d. productivity shock. Using a dynamic common factor representation, equation (7) in the main text with (3) can be rewritten as:

$$\begin{aligned} y_{it} = {} & \beta_k k_{it} - \rho\beta_k k_{i,t-1} + \beta_s l^s_{it} - \rho\beta_s l^s_{i,t-1} + \beta_u l^u_{it} - \rho\beta_u l^u_{i,t-1} + \beta_e e_{it} - \rho\beta_e e_{i,t-1} \\ & + \beta_x x_{it} - \rho\beta_x x_{i,t-1} + \beta_d d_{it} + (\gamma - \rho\beta_d) d_{i,t-1} + \rho y_{i,t-1} + \xi_t + \alpha_i + \mu_{it} \end{aligned} \tag{4}$$

where $\mu_{it} = \eta_{it} - \rho\eta_{i,t-1} + v_{it}$.

Following Blundell and Bond (2000), we first estimate the unrestricted parameter vector of (4) 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 (cf., Chamberlain, 1982). The following moment conditions are used:

$$E[z_{i,t-s}\Delta\mu_{it}] = 0 \quad \text{for } s = 2, 3, \tag{5}$$

$$E[\Delta z_{i,t-s}(\alpha_i^* + \mu_{it})] = 0 \quad \text{for } s = 1, \tag{6}$$

where $z_{it} = (y_{it}, k_{it}, l^s_{it}, l^u_{it}, x_{it}, d_{it})$ and $\Delta z_{it} = z_{it} - z_{i,t-1}$. The first set of the moment conditions (5) comes from the first differenced equations with lagged levels of the variables as instruments. Blundell and Bond (1998) find that exploiting the additional moment conditions (6), 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 have found that even the system GMM estimator could lead to imprecise and possibly biased estimates due to weak instruments (e.g., Griliches and Mairesse, 1998; Mulkay, Hall, and Mairesse, 2000; Levinsohn and Petrin, 2003).

1.3 Adjustment Costs: Ackerberg, Caves and Fraser (2005)

Suppose that skilled labor, unskilled labor, and energy are subject to adjustment costs so that the past variables for skilled labor, unskilled labor, and energy are also this period's state variables. Denote $s_{it} = (l^s_{it}, l^u_{it}, e_{it})$. Then, the demand functions are written as: $l^s_{it} = l^{s*}_t(k_{it}, d_{it}, \omega_{it}, s_{i,t-1})$, $l^u_{it} = l^{u*}_t(k_{it}, d_{it}, \omega_{it}, s_{i,t-1})$, and $e_{it} = e^*_t(k_{it}, d_{it}, \omega_{it}, s_{i,t-1})$. Or we can write

$$s_{it} = s^*_t(k_{it}, d_{it}, \omega_{it}, s_{i,t-1}),$$

where $s^*_t(\cdot)$ is a vector-valued function. We add the subscript t since the demand functions depend on prices which are time-dependent.

We maintain the assumption that materials are not subject to adjustment costs. Since x_{it} is a freely variable input and, thus, the adjustment costs for skilled labor, unskilled labor, and energy affect the demand for materials only through their effects on the choices of s_{it}, we may consider the demand function for materials *conditioned on* $s_{it} = (l^s_{it}, l^u_{it}, e_{it})$:[2] $x_{it} = x^*_t(k_{it}, d_{it}, \omega_{it}, s_{it})$. In the Cobb-Douglas case, it is straightforward to verify that this demand function is strictly increasing in ω_{it} and we get the function $\omega^*_t(k_{it}, d_{it}, x_{it}, s_{it})$ which corresponds to the equation (10) in the main text.[3]

The rest of the estimation procedure is similar to the one discussed in the main text. In particular, we have an estimate of the residual for each candidate parameter vector as:

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

Based on the residual, we can construct the GMM estimator using the instrument $Z_{it} = (k_{it}, k_{i,t-1}, d_{i,t-1}, d_{i,t-2}, x_{i,t-1}, x_{i,t-2}, l^s_{i,t-1}, l^s_{i,t-2}, l^u_{i,t-1}, l^u_{i,t-2}, e_{i,t-1}, e_{i,t-2})$. That is, the parameters $\beta^* = (\beta^*_k, \beta^*_s, \beta^*_u, \beta^*_e, \beta^*_x, \beta^*_d)$

[2] Alternatively, as Ackerberg, Caves and Fraser (2005) suggest, we may also consider the demand function for materials as $x_{it} = x^*_t(k_{it}, d_{it}, \omega_{it}, s_{i,t-1})$ but, in this case, it is not easy to verify the strict monotonicity condition.

[3] For example, consider a simplified version of production function: $Y_{it} = e^{\omega_{it}} K^{\beta_k}_{it} L^{\beta_l}_{it} X^{\beta_x}_{it}$. In this case, $s_{it} = l_{it}$ and we omit d_{it}. Then, from the first order condition for X_{it}, we have $x(k_{it}, \omega_{it}, l_{it}) = const_t + (\beta_k/(\beta_x - 1))k_{it} + (\beta_l/(\beta_x - 1))l_{it} + \omega_{it}$, where "$const_t$" depends on prices and parameters. It's clear that this function is strictly increasing in ω_{it}. Inverting it with respect to ω_{it}, we have $\omega^*_t(k_{it}, x_{it}, l_{it}) = const_t - (\beta_k/(\beta_x - 1))k_{it} - (\beta_l/(\beta_x - 1))l_{it} + x_{it}$.

are estimated by minimizing the GMM criterion function $Q(\beta^*) = \sum_{h=1}^{12}[\sum_{i=1}^{N}\sum_{t=1981}^{T_i}(\nu_{it} \hat{+} \eta_{it})(\beta^*)Z_{it,h}]^2$, where T_i is the last year the i^{th} firm is observed. Note that, by using *past*, rather than current, labor and energy variables as instruments to identify β_s, β_u and β_e we allow for the possibility that a plant makes these decisions *after* observing this period's innovation in productivity.

2 Additional Estimation Results

2.1 Ackerberg, Caves, and Frazer (2005)

There are several reasons to suspect the coefficients estimated in the first stage of the OP/LP technique. One possibility is that the level of skilled labor cannot be altered without incurring extra adjustment costs; in this case, they are more properly defined as state variables in the firm's problem, rather than freely variable inputs. Columns (1) and (3) of Table 1 present the results from the OP/LP estimator where skilled labor is treated as a state variable for the Basic and Extended Sample, respectively. Because skilled labor is a state variable its coefficient is estimated in the second stage. The results show that the coefficients across all variables are reasonably similar to those found in the original experiment and again indicate the substantial static and dynamic effects from using imports.

Recently, Ackerberg, Caves and Frazer (2005) show that the OP/LP estimation technique may not properly identify the coefficients estimated in the first stage. Their critique argues that the demand functions for skilled labor, unskilled labor, and energy can be written as functions of the state variables $(k_{it}, d_{it}, \omega_{it})$ and, since we have $\omega_{it} = \omega_t^*(x_{it}, k_{it}, d_{it})$ in equation (10), they are fully written as functions of (x_{it}, k_{it}, d_{it}). Then, looking at the first stage regression (11), we notice that $l_{it}^s - E[l_{it}^s|x_{it}, k_{it}, d_{it}] = 0$ etc.. That is, there would be no variability left in the regressors to identify the first stage coefficients.

To deal with this identification issue, we estimate *all of the coefficients in the second stage* of the OP/LP technique, extending a method proposed by Ackerberg et al. (see also Bond and Söderbom, 2005). Our method assumes that skilled labor, unskilled labor, and energy are also subject to adjustment costs. Then, their demand functions depend on their past values, which in turn provide variations in their current values that are independent of $(k_{it}, d_{it}, \omega_{it})$ to identify their coefficients. In this case, β_s, β_u, and β_e are identified from the moment conditions $E[(\nu_{it} + \eta_{it})l_{it-1}^s] = 0$, $E[(\nu_{it} + \eta_{it})l_{it-1}^u] = 0$, and $E[(\nu_{it} + \eta_{it})e_{it-1}] = 0$, respectively, while $(\beta_x, \beta_k, \beta_d)$ are identified from the moment conditions $E[(\nu_{it} + \eta_{it})x_{it-1}] = 0$, $E[(\nu_{it} + \eta_{it})k_{it}] = 0$, and $E[(\nu_{it} + \eta_{it})d_{it-1}] = 0$. We also include six over-identifying conditions using the predetermined variables $(k_{i,t-1}, d_{i,t-2}, x_{i,t-2}, l_{i,t-2}^s, l_{i,t-2}^u, e_{i,t-2})$.

The results are reported in columns (2) and (4) of Table 1. In column (2) we see that the coefficient for unskilled labor is close to zero and the standard errors are considerably wider on skilled labor and energy. Similarly, in column (4) the coefficient on skilled labor is close to zero and the standard errors

Table 1: Additional Estimates of Production Function: Discrete Import Variable

The Data Set	Basic Sample		Extended Sample	
Estimators	OP/LP Proxy (1)	ACF (2)	OP/LP Proxy (3)	ACF (4)
Skilled labor	0.111	0.049	0.076	0.0001
	(0.032)	(0.181)	(0.010)	(0.084)
Unskilled labor	0.147	0.0001	0.138	0.107
	(0.008)	(0.017)	(0.009)	(0.021)
Energy	0.041	0.199	0.048	0.097
	(0.004)	(0.143)	(0.004)	(0.096)
Capital	0.060	0.036	0.075	0.058
	(0.009)	(0.012)	(0.010)	(0.013)
Materials	0.603	0.528	0.632	0.655
	(0.023)	(0.049)	(0.020)	(0.034)
Disc. Import (β_d)	0.260	0.228	0.178	0.162
	(0.037)	(0.047)	(0.035)	(0.031)
γ	0.035	0.051	0.024	0.028
	(0.009)	(0.007)	(0.008)	(0.007)
ρ	0.815	0.868	0.816	0.821
	(0.024)	(0.035)	(0.024)	(0.028)
Implied $\frac{\gamma}{1-\rho}$	0.190	0.388	0.130	0.157
P-value for over-identification test	0.744	0.719	0.975	1.000
No. of Obs.	33200		45518	

Notes: Standard errors are in parentheses. Columns (1)-(2) use the "Basic Sample" that excludes plants for which the initial capital stock are not reported. Columns (3)-(4) 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 OP/LP estimators in columns (1) and (3) specify the stochastic process of ω_t of the equation (8) in the main text. The OP/LP estimators in column (1) and (3) estimate the coefficient of skilled labor in the second stage by treating it as an additional state variable. The ACF estimator in column (2) and (4) treats skilled labor, unskilled labor, and energy as additional state variables and, hence, estimates all of these coefficients in the second stage.

are again wider on skilled labor, unskilled labor and energy. While this may be indicative of model misspecification, it could likely to point to a lack of good instruments for those variables. However, the size and significance of the coefficients measuring the static and dynamic effect from using imports does not change qualitatively across the estimation procedure. Both β_d and γ are positive and significant even when all coefficients are estimated in the second stage of the OP/LP procedure.

2.2 Industry-level Results

We also check how the results change across industries. We estimate the production functions based on the Basic Sample for two of the largest 3-digit level industries (ISIC codes): Food (311) and Metals (381). Table 2 presents the results from the OP/LP estimators, where ω_{it} processes are specified using either AR(1) or a third-order series approximation. 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 are reported in columns (1) and (3) and they indicate a large positive static effect on productivity (24.3-25.7 percent). When we specify ω_{it} processes by series in columns (2) and (4), the estimates are slightly lower but still large (19.1-22.7 percent) although they are marginally significant at a 10 percent level. The estimated values of γ for the discrete import variable are positive and significant

Table 2: 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.

for both industries, suggesting a positive dynamic effect of the usage of imported materials.

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 to 42.2 percent, but the estimates from the Metal industry are not as significant; for the Metal industry, the estimate from AR(1) specification in column (7) is barely significant at a 10 percent level while the estimate from series in column (8) is not significant even at a 10 percent level. Although the relatively large standard errors for the Metal industry might be due to its small sample size, the insignificance of the static productivity effect adds a 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 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.

Table 3: Panel OP/LP Estimates: Energy vs. Materials

The Data Set	Basic Sample							
Import Variable	Discrete				Continuous			
The Proxy	Materials		Energy		Materials		Energy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ω_{it} process	AR(1)	Series	AR(1)	Series	AR(1)	Series	AR(1)	Series
Skilled labor	0.137		0.137		0.138		0.138	
	(0.006)		(0.022)		(0.006)		(0.006)	
Unskilled labor	0.145		0.143		0.148		0.147	
	(0.008)		(0.024)		(0.008)		(0.007)	
Energy	0.043		0.057	0.061	0.044		0.098	0.084
	(0.005)		(0.016)	(0.022)	(0.005)		(0.024)	(0.022)
Capital	0.058	0.064	0.026	0.029	0.066	0.074	0.047	0.041
	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.014)	(0.011)
Materials	0.549	0.509	0.643		0.616	0.577	0.643	
	(0.025)	(0.029)	(0.101)		(0.021)	(0.027)	(0.006)	
Import	0.214	0.220	0.431	0.448	0.246	0.270	0.520	0.495
	(0.035)	(0.038)	(0.039)	(0.045)	(0.052)	(0.061)	(0.158)	(0.086)
θ	—	—	—	—	3.505	3.115	2.237	1.170
					(1.428)	(2.361)	(0.035)	(0.064)
γ	0.041	—	0.072	—	0.030	—	0.086	—
	(0.011)		(0.011)		(0.010)		(0.035)	
ρ	0.892	—	0.808	—	0.822	—	0.837	—
	(0.027)		(0.027)		(0.031)		(0.052)	
Implied $\frac{\gamma}{1-\rho}$	0.380	—	0.373	—	0.169	—	0.528	—
	(0.166)		(0.107)		(0.086)		(0.246)	
P-value for over-identification test	0.593	0.759	0.930	0.980	0.824	0.759	1.000	0.915
No. of Obs.	33200							

Notes: Standard errors are in parentheses. Columns (1)-(8) use the "Basic Sample" that excludes plants for which the initial capital stock are not reported. Columns (1), (2), (5) and (6) use the materials variable as the OP/LP proxy, while columns (3), (4), (5) and (6) use energy as a proxy. The OP/LP estimators in columns (2), (4), (6) and (8) specify the stochastic process of ω_t using the third order polynomials in (ω_{t-1}, d_{t-1}).

2.3 Energy As A Proxy

Table 3 presents the results from the OP/LP Proxy estimator using energy as a proxy (instead of materials) on the Basic Sample. Columns (1)-(4) use the discrete import variable, while columns (5)-(8) present the results for the continuous import variable. We have included the results for both energy and materials to ease comparison. Columns (1),(2),(5) and (6) use materials as the proxy for productivity in the OP/LP estimation, whereas columns (3),(4),(7) and (8) use energy as the proxy. Columns (1), (3), (5) and (7) present the results under the assumption that the plant-specific productivity shock, ω_{it}, follows an AR(1) process. Columns (2), (4), (6) and (8) use the OP/LP estimation technique without making any assumptions on the structure of ω_{it} and estimate ω_{it} as a third order polynomial in $(\omega_{i,t-1}, P_{it}, d_{i,t-1})$ where $d_{i,t-1}$ is the lagged decision to import.

The most important finding is the significance and large size of the current discrete import variable across proxies. Comparing columns (1) and (3), (2) and (4), (5) and (7), and, (6) and (8) it is clear that using the energy proxy increases the estimated impact of switching to imported materials. Similarly, a positive and significant γ coefficient across proxies strengthens the evidence of a dynamic import effect.

As a consequence of the strong dynamic effect, the long run effect, $\gamma/(1-\rho)$, is also substantially higher using the energy proxy compared to the materials proxy. None of the specifications are rejected

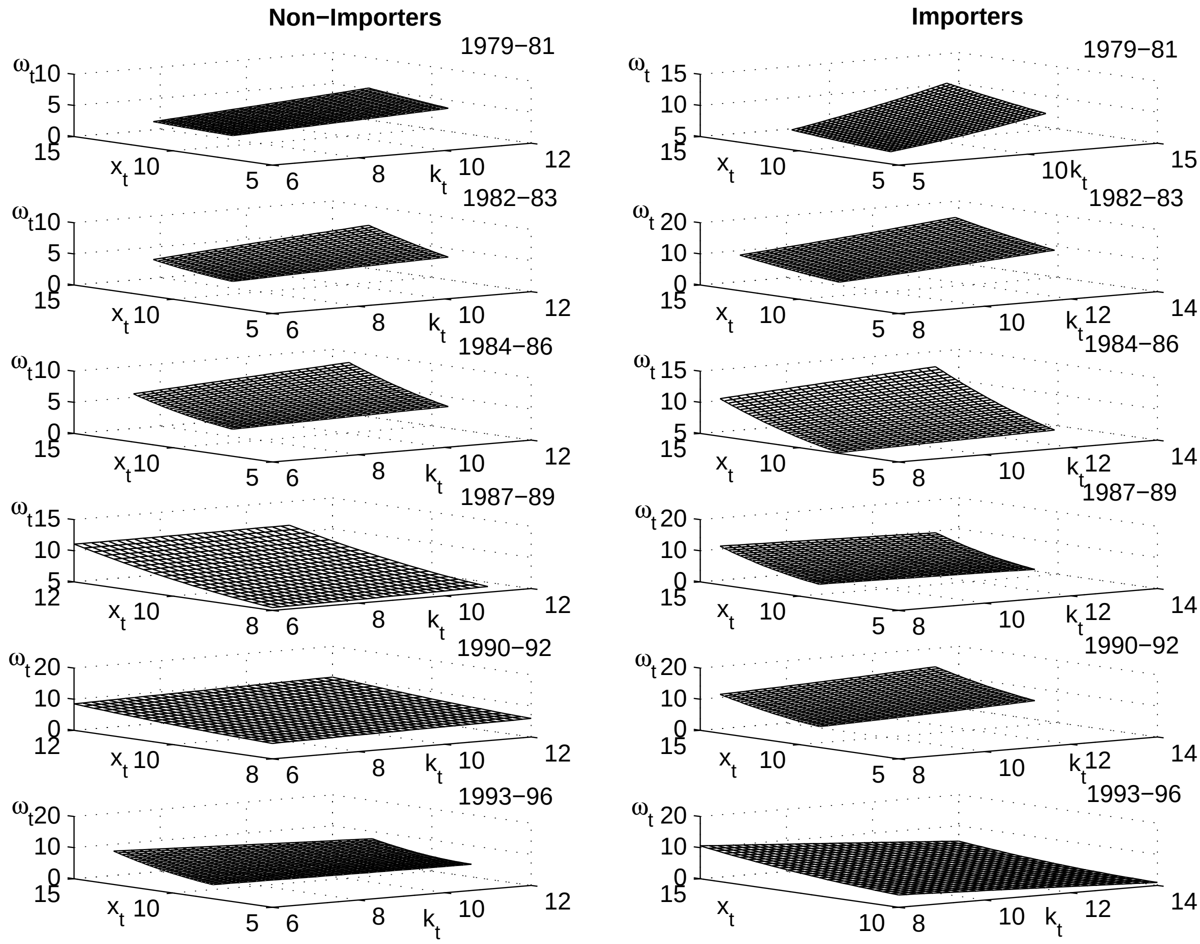

Figure 1: Monotonicity Condition

by the bootstrapped overidentification test.

2.4 Monotonicity Condition

In this section we graphically examine the monotonicity condition required for the OP/LP estimation procedure to be valid. The monotonicity condition is essentially the same as that in Levinsohn and Petrin (2002): conditional on capital and the decision to import intermediates, profit maximizing behavior must lead more productive firms to use more intermediate materials.

Figure 1 shows the relationship between productivity, capital and materials for each group of years. On the left hand side we show this relationship for non-importers and on the right hand side we show this relationship for importers. In all cases we see that the materials variable is increasing in both capital and productivity indicating that the monotonicity condition is satisfied.

Table 4: Descriptive Statistics in 1980 (Extended Sample)

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

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.

2.5 Other Omitted Results

In the main text, we omitted some results from the Extended Sample because these results are qualitatively very similar to those from the Basic Sample. In this section we report the omitted results.

Table 4 and 5, corresponding to Table 1 and 2 in the main text, report descriptive statistics for variables in the year of 1980 and transition rates across import status together with exit rates from the Extended Sample. The descriptive statistics as well as the transition probabilities across import status from the Extended Sample are similar to those from the Basic Sample.

Table 6 shows the frequency of export/import status change over the sample period of 1990-1996 among continuosly operating plants. More than 80 percent of plants did not change export/import status throughout the sample period.

For brevity, 90 % bootstrap confidence intervals are omitted in Figure 2 in the main text. As a example, Figure 1 plots the productivity dynamics before and after plants start importing together with 90 % bootstrap confidence interval for the 50 percentile importing plants with 22.2 percent import shares ($n_t = 0.251$). Note that the solid line in Figure 1 of this appendix corresponds to the dotted line in Figure 2 of the main paper.

Table 5: 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.

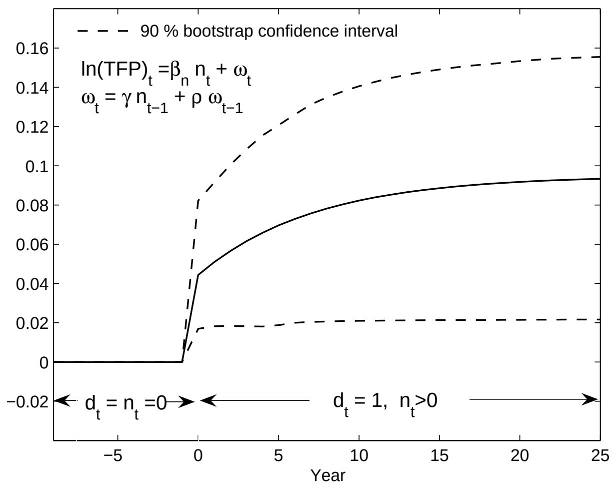

Figure 1: Productivity Dynamics for the 50 Percentile Importing Plants

Table 6: Export and Import Status Change

	Exports		Imports	
Status Changes	No.	%	No.	%
0	1837	0.835	1765	0.802
1	207	0.094	190	0.086
2	109	0.050	163	0.074
3	33	0.015	55	0.025
4+	15	0.007	28	0.013
No. of Plants	2,201			

Notes: Based on the sample of plants that continuously operated over the 1990-1996 period.

References

[1] Ackerberg, Daniel A., Benkard, C. Lanier, Berry, Steven, and Ariel Pakes (2006) "Econometric Tools for Analyzing Market Outcomes," mimeo.

[2] Ackerberg, Daniel A., Caves, Kevin, and Garth Frazer (2005) "Structural Identification of Production Functions," mimeo.

[3] 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.

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

[5] 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.

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

[7] Bond, S. and M. Söderbom M. (2005) "Adjustment Costs and the Identification of Cobb Douglas Production Functions," working paper.

[8] Griliches, 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.

[9] 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.

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

[11] 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.