Industry Dynamics and Productivity II

Paul T. Scott
Toulouse School of Economics
ptscott@gmail.com

Empirical IO
Spring 2014

"The Dynamics of Productivity in the Telecommunications Equipment Industry"
Olley and Pakes (1996)

Overview

- Analyzes effects of deregulation in telecommunications equipment industry.
- Deregulation increases productivity, primarily through reallocation toward more productive establishments.
- Estimation approach deals with simultaneity and selection issues.

Background I

- AT&T had a monopoly on telecommunications services in the US throughout most of the 20th century (note: a telecommunications network is a classic example of a natural monopoly).
- Before the regulatory change, AT&T required that equipment attached to their network must be supplied by the AT&T, and virtually all of their equipment was supplied by their subsidiary, Western Electric. Thus, they leveraged their network monopoly to a monopoly on phones.

Background II

- A change in technology opened up new markets for telecommunications equipment (e.g., fax machines)
- Meanwhile, the FCC (regulatory agency) decided to begin allowing the connection of privately-provided devices to AT&T's network.
- A surge of entry into telecommunications equipment manufacturing followed in the late 1960's and 1970's.

TABLE I

CHARACTERISTICS OF THE DATA

Year	Plants	Firms	Shipments (billions 1982 $)	Employment
1963	133	104	5.865	136899
1967	164	131	8.179	162402
1972	302	240	11.173	192248
1977	405	333	13.468	192259
1982	473	375	20.319	222058
1987	584	481	22.413	184178

Background III

- AT&T continued purchasing primarily from Western Electric into the 1980's (although consumers were free to purchase devices from other companies).
- The divestiture (breakup) of AT&T created seven regional Bell companies that were no longer tied to Western Electric, and they were prohibited from manufacturing their own equipment.
- The divestiture was implemented in January 1984. Western Electric's share dropped dramatically.

TABLE II

BELL COMPANY EQUIPMENT PROCUREMENT
(PERCENT PURCHASED FROM WESTERN ELECTRIC)

1982	1983	1984	1985	1986[E]
92.0	80.0	71.8	64.2	57.6

[E] Estimated for 1986.
Source: NTIA (1988, p. 336, and discussion pp. 335–337).

Entry

TABLE III

ENTRANTS ACTIVE IN 1987

	Number	Share of Number Active in 1987 (%)	Share of 1987 Shipments (%)	Share of 1987 Employment (%)
Plants: New since 1972	463	79.0	32.8	36.0
Firms: New since 1972	419	87.0	30.0	41.4
Plants: New since 1982	306	52.0	12.0	13.5
Firms: New since 1982	299	60.1	19.4	27.5

Exit

TABLE IV

INCUMBENTS EXITING BY 1987

	Number	Share of Number Active in Base Year (%)	Share of Shipments in Base Year (%)	Share of Employment in Base Year (%)
Plants active in 1972 but not in 1987	181	60.0	40.2	39.0
Firms active in 1972 but not in 1987	169	70.0	13.8	12.1
Plants active in 1982 but not in 1987	195	41.2	26.0	24.1
Firms active in 1982 but not in 1987	184	49.1	17.3	16.1

The model

- Incumbent firms (i) make three decisions:
  - Whether to exit or continue. If they exit, they receive a fixed scrap value Ψ and never return.
  - If they stay, they choose labor l_{it},
  - and investment i_{it}.

- Capital accumulation:

$$k_{t+1} = (1 - \delta) k_t + i_t$$

- Another state variable is age: $a_{t+1} = a_t + 1$

Production

- They assume the following Cobb-Douglas production function:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it}$$

  where y_{it} is output, k_{it} is capital, l_{it} is labor, ω_{it} is a persistant component of productivity, and ε_{it} is a transient shock to productivity.

- Productivity evolves according to a Markov process: $F(\cdot|\omega)$.
- η is either measurement error, or there is no information about it when labor decisions are made.

Equilibrium behavior

- They assume the existence of a Markov perfect equilibrium. Market structure and prices are state variables in the MPE, but they are common across firms, so they can be absorbed into time subscripts for the value function:

$$V_t(\omega_t, a_t, k_t) = \max\Big\{\Psi, \sup_{i_t \geq 0} \pi_t(\omega_t, a_t, k_t) - c(i_t) \\ +\beta E\left[V_{t+1}(\omega_{t+1}, a_{t+1}, k_{t+1}) \,|\, J_t\right]$$

  where J_t represents the information set at time t.

- Equilibrium strategies can be decribed by functions $\underline{\omega}_t(a_t, k_t)$ and $i_t(\omega_t, a_t, k_t)$.
  - A firm will continue if and only if $\omega \geq \underline{\omega}_t(a_t, k_t)$.
  - Continuing firms invest $i_t = i_t(\omega_t, a_t, k_t)$

Thinking about bias

- How does the simultaneity of the input decision bias the labor coefficient?
- How does selection due to exit bias the capital coefficient estimate?

Productivity inversion

- In a technical paper, Pakes (1994) shows that the investment rule $i_t(\omega_t, a_t, k_t)$ is monotonically increasing in ω_t, provided $i_t > 0$.
- Given monotonicity, optimal investment can be inverted for productivity:

$$\omega_{it} = h_t(i_{it}, a_{it}, k_{it}).$$

- We're going to talk more about the $i_t > 0$ requirement with Levinsohn and Petrin (2003).

First stage model

- Substituting in the inversion function,

$$y_{it} = \beta_l l_{it} + \phi_t \left(i_{it}, a_{it}, k_{it}\right) + \eta_{it}$$

where

$$\phi_t \left(i_{it}, a_{it}, k_{it}\right) = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + h_t \left(i_{it}, a_{it}, k_{it}\right)$$

- We can estimate this equation using a semiparametric regression. This may identify β_l, but not the other coefficients.
- With Ackerberg, Caves, and Frazer (2006), we will think more carefully about what's identifying β_l, but don't worry about it for now.

Selection

- Let $P_t = Pr\left(\chi_{t+1} = 1 | \underline{\omega}_{t+1}\left(k_{t+1}, a_{t+1}\right), J_t\right)$ be the propensity score
- As long as the conditional density of ω_{t+1} has full support, this can be inverted to express $\underline{\omega}_{t+1} = f\left(P_t, \omega_t\right)$

The second equation

- Write the expectation of $y_{t+1} - \beta_l l_{t+1}$ conditional on survival:

$$E\left[y_{t+1} - \beta_l l_{t+1} | a_{t+1}, k_{t+1}, \chi_{t+1} = 1\right]$$

$$= \beta_a a_{t+1} + \beta_k k_{t+1} + g\left(\underline{\omega}_{t+1}, \omega_t\right)$$

where $g\left(\underline{\omega}_{t+1}, \omega_t\right) = E\left[\omega_{t+1} | \omega_t, \chi_{t+1} = 1\right]$

- Using the inversion of the selection probability, we can write

$$g\left(\underline{\omega}_{t+1}, \omega_t\right) = g\left(f\left(P_t, \omega_t\right), \omega_t\right)$$

which can be written more simply as $g\left(P_t, \omega_t\right)$.

Final step

- Conditional on values of (β_a, β_k), we can construct an estimate of $\omega_t = \phi_t - \beta_a a_t - \beta_k k_t$
- Finally, write

$$\begin{aligned} y_{t+1} - \beta_l l_{t+1} = \quad & \beta_a a_{t+1} + \beta_k k_{t+1} + g\left(P_t, \phi_t - \beta_a a_t - \beta_k k_t\right) \\ & + \xi_{t+1} + \eta_{t+1} \end{aligned}$$

where

$$\xi_{t+1} = \omega_{t+1} - E\left[\omega_{t+1} | \omega_t, \chi_{t+1} = 1\right].$$

Final step

$$y_{t+1} - \beta_l l_{t+1} = \beta_a a_{t+1} + \beta_k k_{t+1} + g\left(P_t, \phi_t - \beta_a a_t - \beta_k k_t\right) + \xi_{t+1} + \eta_{t+1}$$

- This is a nonlinear estimation equation and we can estimate it using

$$E\left(\xi_{t+1}\begin{pmatrix} k_{t+1} \\ a_{t+1} \end{pmatrix}\right) = 0$$

noting that we should not impose $E\left(\xi_{t+1} l_{t+1}\right) = 0$.

Estimation steps

1. First stage semi-parametric regression:

$$y_{it} = \beta_l l_{it} + \phi_t \left(i_{it}, a_{it}, k_{it} \right) + \eta_{it}$$

2. Estimate propensity scores: $P_t = Pr\left(\chi_{t+1} = 1 | \underline{\omega}_{t+1}\left(k_{t+1}, a_{t+1} \right), J_t \right)$

3. Estimate remaining parameters:

$$\begin{aligned} y_{t+1} - \beta_l l_{t+1} = \; & \beta_a a_{t+1} + \beta_k k_{t+1} + g\left(P_t, \phi_t - \beta_a a_t - \beta_k k_t \right) \\ & + \xi_{t+1} + \eta_{t+1} \end{aligned}$$

   using exclusion restrictions on innovation term ξ_{t+1}.

Olley and Pakes (1996)

TABLE VI

ALTERNATIVE ESTIMATES OF PRODUCTION FUNCTION PARAMETERS[a]
(STANDARD ERRORS IN PARENTHESES)

Sample:	Balanced Panel		Full Sample[c, d]						
								Nonparametric F_ω	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation Procedure	Total	Within	Total	Within	OLS	Only P	Only h	Series	Kernel
Labor	.851	.728	.693	.629	.628			.608	
	(.039)	(.049)	(.019)	(.026)	(.020)			(.027)	
Capital	.173	.067	.304	.150	.219	.355	.339	.342	.355
	(.034)	(.049)	(.018)	(.026)	(.018)	(.02)	(.03)	(.035)	(.058)
Age	.002	−.006	−.0046	−.008	−.001	−.003	.000	−.001	.010
	(.003)	(.016)	(.0026)	(.017)	(.002)	(.002)	(.004)	(.004)	(.013)
Time	.024	.042	.016	.026	.012	.034	.011	.044	.020
	(.006)	(.017)	(.004)	(.017)	(.004)	(.005)	(.01)	(.019)	(.046)
Investment	—	—	—	—	.13	—	—	—	—
					(.01)				
Other Variables	—	—	—	—	—	Powers of P	Powers of h	Full Polynomial in P and h	Kernel in P and h
# Obs.[b]	896	896	2592	2592	2592	1758	1758	1758	1758

Productivity decomposition

- Estimate of productivity:

$$p_{it} = \exp\left(y_{it} - b_l l_{it} - b_k k_{it} - b_a a_{it}\right)$$

  where b's represent coefficient estimates.

- Aggregate productivity: $p_t = \sum_{i=1}^{N_t} s_{it} p_{it}$.
- Can be decomposed as follows:

$$\begin{aligned} p_t &= \sum_{i=1}^{N_t} \left(\bar{s}_t + \Delta s_{it}\right)\left(\bar{p}_t + \Delta p_{it}\right) \\ &= N_t \bar{s}_t \bar{p}_t + \sum_{i=1}^{N_t} \Delta s_{it} \Delta p_{it} \\ &= \bar{p}_t + \sum_{i=1}^{N_t} \Delta s_{it} \Delta p_{it} \end{aligned}$$

  where $\bar{p}_t$ are unweighted mean productivity and shares in the cross-section.
- Thus, aggregate productivity decomposes into an unweighted mean and a covariance term.

TABLE XI

DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_i \Delta s_{it} \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

[a] See text for details.

Quick Review of Dynamic Panel Data Estimation
Arellano and Bond (1991), Blundell and Bond (1998, 2000)

DP Setup

- Production function with fixed effects:

$$y_{it} = \beta x_t + \alpha_i + \omega_{it} + \eta_{it}$$

  where $x_t = (l_t, k_t)$, ω_{it} is the productivity term, and η_{it} is measurement error.
- Let $\psi_{it} = \alpha_i + \omega_{it} + \eta_{it}$. Then, with parametric assumptions about how ω_{it} evolves, we can estimate β using dynamic panel data methods.

DP Example

For example, assume $\omega_{it} = \rho\omega_{i,t-1} + \xi_{i,t}$.

- We can recover estimates of ψ_{it} as a function of a particular β: $\hat{\psi}_{it}(\beta) = y_{it} - \beta x_t$.
- Conditional on β, it is easy to estimate $\hat{\rho}(\beta)$ using $\hat{\psi}_{it}(\beta)$.
- We can then compute $\hat{\xi}_{it}(\beta) = \omega_{it} - \hat{\rho}(\beta)\,\omega_{i,t-1}$.
- Then, we can use moments to estimate β. If inputs are quasi-fixed, we could use

$$E\left(\hat{\xi}_{it}(\beta)\begin{pmatrix} l_t \\ k_t \end{pmatrix}\right) = 0.$$

- If labor is variable, perhaps

$$E\left(\hat{\xi}_{it}(\beta)\begin{pmatrix} l_{t-1} \\ k_t \end{pmatrix}\right) = 0.$$

Brief comparison

Advantages of DP methods:

- Can handle fixed effects together with an evolving productivity term.
- Does not rely on invertibility of input demand functions.

Disadvantages of DP methods:

- Standard dynamic panel data estimators don't deal with selection bias created by exit. Using the unbalanced panel deals reduces the magnitude of the selection problem, but we need an explicit treatment of it as in OP to eliminate it.
- DP methods are more restrictive in terms of process governing evolution of ω. (See Ackerberg, Caves, and Frazer)

"Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants"
Nina Pavcnik (2002)

Overview

- First application of OP, and early paper in what is now a massive structural literature on trade liberalization and productivity.
- Investigates effects of "massive trade liberalization" in Chile from late 70's to early 80's.
- The Pinochet regime was tumultuous, and there was a large recession in 82-83, so a simple before/after comparison wouldn't be plausible.
- Combines structural estimation with diff-in-diffs identification strategy
  - before vs. after trade liberalization
  - sectors affected by trade liberalization vs. non-traded goods industries

Findings

- Consistent with OP, selection and simultaneity bias substantially bias estimates of the coefficients of the production function
- Substantial within-plant productivity improvements
- There was massive exit during the period of liberalization, and exiting plants tended to be less productive

TABLE 1

Plants active in 1979 *but not in* 1986

Trade orientation	Share of plants	Share of labour	Share of capital	Share of investment	Share of value added	Share of output
Exiting plants of a given trade orientation as a share of all plants active in 1979						
All trade orientations	0·352	0·252	0·078	0·135	0·155	0·156
Export-oriented	0·045	0·049	0·009	0·039	0·023	0·023
Import-competing	0·141	0·108	0·029	0·047	0·068	0·065
Nontraded	0·165	0·095	0·040	0·049	0·064	0·067
Exiting plants of a given trade orientation as a share of all exiting plants						
Export-oriented	0·129	0·194	0·117	0·289	0·149	0·148
Import-competing	0·401	0·429	0·369	0·350	0·436	0·419
Nontraded	0·470	0·377	0·513	0·361	0·415	0·432
Exiting plants of a given trade orientation as a share of all plants active in 1979 in the corresponding trade sector						
Export-oriented	0·416	0·298	0·030	0·172	0·121	0·128
Import-competing	0·383	0·263	0·093	0·149	0·183	0·211
Nontraded	0·316	0·224	0·104	0·107	0·147	0·132

Note: This figure also includes plants that exited after the end of 1979, but before the end of 1980 and were excluded in the estimation because of missing capital variable.

Some details

- Methodologically almost identical to Olley and Pakes.
- One difference: while OP use value added as output, Pavcnik uses sales and includes materials on the right-hand side:

$$y_{it} = \beta_0 + \beta x_{it} + \beta_k k_{it} + e_{it}$$

where x includes unskilled labor, skilled labor, and material inputs.

- In the first-stage regression, she estimates β, i.e., the coefficients on the labor and materials variables.
- Zero investment is a significant phenomenon in the data, and she finds it doesn't matter whether she drops observations with $i_{it} = 0$ or if she ignores the monotonicity issue and includes them.

More details

- Sales deflated using price indices for four-digit industry codes. Note that this leaves A LOT of room for price heterogeneity. Things that are four-digit industries:
  - Manufacture of malt liquors and malt
  - Manufacture of consumer electronics
  - Manufacture of motor vehicles
- When estimating relationship between trade and productivity, she controls for heterogeneous prices/markups using plant-specific fixed effects.
- Estimates model separately for each 2- or 3-digit industry.

TABLE 2

Estimates of production functions

		Balanced panel				Full sample					
		OLS		Fixed effects		OLS		Fixed effects		Series	
		(1)		(2)		(3)		(4)		(5)	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Food processing	Unskilled labour	0·152	0·007	0·185	0·012	0·178	0·006	0·210	0·010	0·153	0·007
	Skilled labour	0·127	0·006	0·027	0·008	0·131	0·006	0·029	0·007	0·098	0·009
	Materials	0·790	0·004	0·668	0·008	0·763	0·004	0·646	0·007	0·735	0·008
	Capital	0·046	0·003	0·011	0·007	0·052	0·003	0·014	0·006	0·079	0·034
	N	6432				8464				7085	
Textiles	Unskilled labour	0·187	0·011	0·240	0·017	0·229	0·009	0·245	0·015	0·215	0·012
	Skilled labour	0·184	0·010	0·088	0·014	0·183	0·009	0·088	0·012	0·177	0·011
	Materials	0·667	0·007	0·564	0·011	0·638	0·006	0·558	0·009	0·637	0·097
	Capital	0·056	0·005	0·015	0·012	0·059	0·004	0·019	0·011	0·052	0·034
	N	3689				5191				4265	
Wood	Unskilled labour	0·233	0·016	0·268	0·026	0·247	0·013	0·273	0·022	0·195	0·015
	Skilled labour	0·121	0·015	0·040	0·021	0·146	0·012	0·047	0·018	0·130	0·014
	Materials	0·685	0·010	0·522	0·014	0·689	0·008	0·554	0·011	0·679	0·010
	Capital	0·055	0·007	0·023	0·018	0·050	0·006	−0·002	0·016	0·101	0·051
	N	1649				2705				2154	
Paper	Unskilled labour	0·218	0·024	0·258	0·033	0·246	0·021	0·262	0·029	0·193	0·024
	Skilled labour	0·190	0·018	0·022	0·027	0·180	0·016	0·050	0·023	0·203	0·018
	Materials	0·624	0·013	0·515	0·025	0·597	0·011	0·514	0·021	0·601	0·014
	Capital	0·074	0·010	0·031	0·025	0·085	0·009	0·031	0·023	0·068	0·018
	N	1039				1398				1145	

Diff-in-diffs

- After estimating productivity, she estimates the following regression:

  $$pr_{it} = \alpha_0 + \alpha 1(Time)_{it} + \alpha_2(Trade)_{it} + \alpha_3(Trade * Time)_{it} + \alpha_4 Z_{it} + \nu_{it}$$

- Idea is that year dummies capture omitted macroeconomic variables. We're hoping that different sectors don't have heterogeneous responses to macroeconomic shocks.

TABLE 4

Estimates of equation 12

	(1) Coef.	(1) S.E.	(2) Coef.	(2) S.E.	(3) Coef.	(3) S.E.	(4) Coef.	(4) S.E.	(5) Coef.	(5) S.E.	(6) Coef.	(6) S.E.
Export-oriented	0·106	0·030**	0·106	0·030**	0·112	0·031**	0·098	0·048**	0·095	0·048**	0·100	0·046**
Import-competing	0·105	0·021**	0·105	0·021**	0·103	0·021**	−0·024	0·040	−0·025	0·040	−0·007	0·039
ex_80	−0·054	0·025**	−0·053	0·025**	−0·055	0·025**	−0·071	0·026**	−0·068	0·026**	−0·071	0·026**
ex_81	−0·099	0·028**	−0·097	0·028**	−0·100	0·028**	−0·117	0·027**	−0·110	0·027**	−0·119	0·027**
ex_82	0·005	0·032	0·007	0·032	0·003	0·032	−0·054	0·028*	−0·042	0·028	−0·055	0·028*
ex_83	0·021	0·032	0·023	0·032	0·021	0·032	−0·036	0·029	−0·025	0·030	−0·038	0·029
ex_84	0·050	0·031	0·051	0·031	0·050	0·031	0·007	0·028	0·017	0·028	0·007	0·028
ex_85	0·030	0·030	0·032	0·031	0·028	0·030	−0·001	0·029	0·013	0·030	−0·003	0·029
ex_86					0·043	0·036					−0·008	0·034
im_80	0·011	0·014	0·011	0·014	0·010	0·014	0·013	0·014	0·013	0·014	0·013	0·014
im_81	0·047	0·015**	0·047	0·015**	0·046	0·015**	0·044	0·014**	0·044	0·014**	0·044	0·014**
im_82	0·033	0·016**	0·034	0·017**	0·030	0·016*	0·024	0·015*	0·024	0·015*	0·025	0·015*
im_83	0·042	0·017**	0·043	0·017**	0·043	0·017**	0·040	0·015**	0·041	0·015**	0·042	0·015**
im_84	0·062	0·017**	0·062	0·017**	0·063	0·017**	0·059	0·015**	0·059	0·015**	0·061	0·015**
im_85	0·103	0·017**	0·104	0·017**	0·104	0·017**	0·101	0·015**	0·102	0·016**	0·101	0·015**
im_86					0·071	0·019**					0·073	0·017**
Exit indicator	−0·081	0·011**	−0·076	0·014**			−0·019	0·010**	−0·010	0·013		
Exit_export indicator			−0·021	0·036					−0·069	0·035*		
Exit_import indicator			−0·007	0·023					−0·005	0·021		
Industry indicators	yes		yes		yes		yes		yes		yes	
Plant indicators	no		no		no		yes		yes		yes	
Year indicators	yes		yes		yes		yes		yes		yes	
R^2 (adjusted)	0·057		0·058		0·062		0·498		0·498		0·488	
N	22983		22983		25491		22983		22983		25491	

Note: ** and * indicate significance at a 5% and 10% level, respectively. Standard errors are corrected for heteroscedasticity. Standard errors in columns 1–3 are also adjusted for repeated observations on the same plant. Columns 1, 2, 4, and 5 do not include observations in 1986 because one cannot define exit for the last year of a panel.

TABLE 8

Relationship between productivity and tariffs, real exchange rate, and import competition

	(1)	(2)	(3)	(4)
Real exchange rate		0·0005**		0·0005**
		(0·0001)		(0·0001)
Tariff	−0·2790**	−0·2377**		−0·2376**
	(0·0280)	(0·0286)		(0·0285)
Imports/output			0·0023**	0·0023**
			(0·0006)	(0·0006)
Plant indicators	yes	yes	yes	yes
R^2 (adjusted)	0·48	0·48	0·48	0·48

Note: ** and * indicate significance at a 5% and 10% level, respectively. Standard errors are corrected for heteroskedasticity. All regressions also include a time trend. N is 25,491.

"Estimating Production Functions Using Inputs to Control for Unobservables"
Levinsohn and Petrin (2003)

Main idea

- Same general framework as Olley and Pakes (1996)
- Main idea: rather than use investment to control for unobserved productivity, use materials inputs.
- Two proposed benefits:
  - Investment proxy isn't valid for plants with zero investment. Zero materials inputs typically an issue in the data.
  - Investments may be "lumpy" and not respond to some productivity shocks.

Downsides of investment

- We need to drop observations with zero investment, which can lead to a substantial efficiency loss. Zero investments happen at a non-trivial rate in annual production data.
- Firms might face non-convex capital adjustment costs leading to flat regions in the $i(\omega)$ function even at positive levels of investment.
- What if investment actually happens with only partial information about productivity and then labor is set once the productivity realization is fully observed?

OP equations

- Production function:

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \omega_t + \eta_t.$$

- First stage regression:

$$y_t = \beta_l l_t + \phi_t \left(i_t, k_t\right) + \eta_t$$

with $\phi_t \left(i_t, k_t\right) = \beta_0 + \beta_k k_t + \omega_t \left(i_t, k_t\right)$.

- Final regression:

$$y_t^* = y_t - \beta_l l_t = \beta_0 + \beta_k k_t + E\left[\omega_t | \omega_{t-1}\right] + \eta_t^*$$

where $\eta_t^* = \eta_t + \left(\omega_t - E\left(\omega_t | \omega_{t-1}\right)\right)$.

LP equations

- Production function:

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_m m_t + \omega_t + \eta_t$$

- First stage regression:

$$y_t = \beta_l l_t + \phi_t \left(m_t, k_t \right) + \eta_t$$

with $\phi_t \left(m_t, k_t \right) = \beta_0 + \beta_k k_t + \beta_m m_t + \omega_t \left(m_t, k_t \right)$.

- Final regression:

$$y_t^* = y_t - \beta_l l_t = \beta_0 + \beta_k k_t + E \left[\omega_t | \omega_{t-1} \right] + \eta_t^*$$

where $\eta_t^* = \eta_t + \left(\omega_t - E \left(\omega_t | \omega_{t-1} \right) \right)$.

Invertability

- Just as OP require $i_t(\omega_t, k_t)$ be an invertible function of productivity, LP require that input use $m_t(\omega_t, k_t)$ is an invertible function of productivity.
- LP's monotonicity result relies on easily checked properties of the production function, and some may find this more appealing than a result which relies on a Markov perfect equilibrium.
- Unobserved input price variation may be a problem for the LP invertability condition (but of course it could be for OP, too).

Checking invertability

- LP claim that

$$\text{sign}\left(\frac{\partial m}{\partial \omega}\right) = \text{sign}\left(f_{ml} f_{l\omega} - f_{ll} f_{m\omega}\right).$$

- To see this, apply the Implicit function theorem to the FOC's to get

$$\begin{pmatrix} \frac{\partial m}{\partial \omega} \\ \frac{\partial l}{\partial \omega} \end{pmatrix} = -\begin{pmatrix} f_{mm} & f_{ml} \\ f_{lm} & f_{ll} \end{pmatrix}^{-1} \begin{pmatrix} f_{m\omega} \\ f_{l\omega} \end{pmatrix}.$$

- Inverting and solving,

$$\Rightarrow \frac{\partial m}{\partial \omega} = \frac{f_{ml} f_{l\omega} - f_{ll} f_{m\omega}}{\begin{vmatrix} f_{mm} & f_{ml} \\ f_{lm} & f_{ll} \end{vmatrix}}.$$

- By the second-order condition for profit maximization, $\begin{vmatrix} f_{mm} & f_{ml} \\ f_{lm} & f_{ll} \end{vmatrix}$ must be negative semidefinite. This means it has exactly two negative eigenvalues, which means its determinant is positive. Therefore, the numerator controls the sign.

Zero inputs

TABLE 2

Per cent of non-zero observations

Industry (ISIC)	Investment	Fuels	Materials	Electricity
Food products (311)	42·7	78·0	99·8	88·3
Metals (381)	44·8	63·1	99·9	96·5
Textiles (321)	41·2	51·2	99·9	97·0
Wood products (331)	35·9	59·3	99·7	93·8

Note: in OP's industry, it was only 8% zeros.

Differences from OP

- LP use a slightly different first stage:
  - First, they estimate $E\left(z_t|k_t\right)$ for $z_t = y_t, l_t^u, l_t^s, e_t, f_t$
  - They then use no-intercept OLS to estimate:

$$\begin{aligned} y_t - E\left(y_t|k_t, m_t\right) = \quad & \beta_s\left(l_t^s - E\left(l_t^s|k_t, m_t\right)\right) \\ & +\beta_s\left(l_t^u - E\left(l_t^u|k_t, m_t\right)\right) \\ & +\beta_e\left(e_t - E\left(e_t|k_t, m_t\right)\right) \\ & +\beta_f\left(f_t - E\left(f_t|k_t, m_t\right)\right) + \eta_t \end{aligned}$$

- Second stage is similar, but they have to estimate two coefficients (β_m, β_k), so they need two moments:

$$E\left(\xi_t\begin{pmatrix} k_t \\ m_{t-1} \end{pmatrix}\right) = 0$$

TABLE 6

Comparisons across estimators P-value for H_0: $\beta_1 = \beta_2$

	Industry (ISIC code)			
Comparison	311	381	321	331
Levinsohn–Petrin vs.				
OLS	<0·01	0·20	0·58	0·21
Fixed effects	<0·01	<0·01	<0·01	<0·01
Instrumental variables	<0·01	0·22	0·09	<0·01
Olley–Pakes	<0·01	0·54	0·20	0·89
Levinsohn–Petrin (i > 0 only)	<0·01	0·02	0·27	0·93
Olley–Pakes vs.				
OLS	<0·01	0·04	0·19	0·46
Fixed effects	<0·01	<0·01	<0·01	<0·01
Instrumental variables	<0·01	<0·01	<0·01	<0·01
Levinsohn–Petrin (i > 0 only)	0·56	0·47	0·85	0·55
Fixed effects vs.				
OLS	<0·01	<0·01	<0·01	<0·01
Instrumental variables	<0·01	<0·01	<0·01	<0·01
No. obs.	6115	1394	1129	1032

Note: The cells in the table contain the *P*-value for a standard Wald test for "no differences between the (vector of) parameter estimates for estimators 1 and 2". <0·01 indicates a *P*-value that is less than 0·01.

"Reallocation, Firm Turnover, and Efficiency:
Selection on Productivity or Profitability"
Foster, Haltiwanger, and Syverson (2008)

Overview

- They look at some rare industries where quantity data is available, allowing them to separate physical and revenue productivity
- Findings:
  - Physical productivity is inversely correlated with price
  - Young producers charge lower prices than incumbents, meaning the literature understates entrants' productivity advantages

Measurement

- Productivity is measured as follows:

$$tfp_{it} = y_{it} - \alpha_l l_{it} - \alpha_k k_{it} - \alpha_m m_{it} - \alpha_e e_{it}$$

- Coefficients (α) are just taken from input shares by industry.
- Different measures use different output measures y:
  - TFPQ uses physical output
  - TFP uses deflated sales (using standard industry-level deflators from NBER)
  - TFPR are sales deflated by mean prices observed in their data

Industries

TABLE 2—ESTIMATING PRICE ELASTICITIES BY PRODUCT

	IV estimation			OLS estimation	
Product	Price coefficient (α_1)		Income coefficient (α_2)	Price coefficient (α_1)	Income coefficient (α_2)
Boxes	−3.02		−0.03	−2.19	−0.03
	(0.17)	[0.61]	(0.02)	(0.12)	(0.02)
Bread	−3.09		0.12	−0.89	0.07
	(0.42)	[0.33]	(0.05)	(0.15)	(0.04)
Carbon black	−0.52		−0.21	−0.57	−0.21
	(0.38)	[0.50]	(0.11)	(0.21)	(0.11)
Coffee	−3.63		0.22	−1.03	0.20
	(0.98)	[0.41]	(0.14)	(0.32)	(0.13)
Concrete	−5.93		0.13	−0.83	0.15
	(0.36)	[0.10]	(0.01)	(0.09)	(0.01)
Hardwood flooring	−1.67		−0.20	−0.87	−0.24
	(0.48)	[0.61]	(0.18)	(0.47)	(0.18)
Gasoline	−1.42		0.23	−0.16	0.23
	(2.72)	[0.20]	(0.07)	(0.80)	(0.07)
Block ice	−2.05		0.00	−0.63	0.16
	(0.46)	[0.32]	(0.11)	(0.20)	(0.07)
Processed ice	−1.48		0.18	−0.70	0.16
	(0.27)	[0.37]	(0.03)	(0.13)	(0.03)
Plywood	−1.21		−0.23	−1.19	−0.23
	(0.14)	[0.89]	(0.10)	(0.13)	(0.10)
Sugar	−2.52		0.76	−1.04	0.72
	(1.01)	[0.15]	(0.13)	(0.55)	(0.12)

Correlations

TABLE 1—SUMMARY STATISTICS FOR OUTPUT, PRICE, AND PRODUCTIVITY MEASURES

Correlations								
Variables	Trad'l. output	Revenue output	Physical output	Price	Trad'l. TFP	Revenue TFP	Physical TFP	Capital
Traditional output	1.00							
Revenue output	0.99	1.00						
Physical output	0.98	0.99	1.00					
Price	−0.03	−0.03	−0.19	1.00				
Traditional TFP	0.19	0.18	0.15	0.13	1.00			
Revenue TFP	0.17	0.21	0.18	0.16	0.86	1.00		
Physical TFP	0.17	0.20	0.28	−0.54	0.64	0.75	1.00	
Capital	0.86	0.85	0.84	−0.04	0.00	−0.00	0.03	1.00
Standard deviations								
	1.03	1.03	1.05	0.18	0.21	0.22	0.26	1.14

Notes: This table shows correlations and standard deviations for plant-level variables for our pooled sample of 17,669 plant-year observations. We remove product-year fixed effects from each variable before computing the statistics. All variables are in logs. See the text for definitions of the variables.

Persistence

TABLE 3—PERSISTENCE OF PRODUCTIVITY, PRICES AND DEMAND SHOCKS

	Five-year horizon		Implied one-year persistence rates	
Dependent variable	Unweighted regression	Weighted regression	Unweighted regression	Weighted regression
Traditional TFP	0.249	0.316	0.757	0.794
	(0.017)	(0.042)		
Revenue TFP	0.277	0.316	0.774	0.794
	(0.021)	(0.042)		
Physical TFP	0.312	0.358	0.792	0.814
	(0.019)	(0.049)		
Price	0.365	0.384	0.817	0.826
	(0.025)	(0.066)		
Demand shock	0.619	0.843	0.909	0.966
	(0.013)	(0.021)		

Entry and exit

TABLE 4—EVOLUTION OF REVENUE PRODUCTIVITY, PHYSICAL PRODUCTIVITY, PRICES AND DEMAND SHOCKS

	Unweighted regression		Weighted regression	
Variable	Exit dummy	Entry dummy	Exit dummy	Entry dummy
Traditional TFP	−0.0209	0.0014	−0.0164	−0.0032
	(0.0042)	(0.0040)	(0.0126)	(0.0188)
Revenue TFP	−0.0218	0.0110	−0.0197	−0.0005
	(0.0044)	(0.0042)	(0.0135)	(0.0183)
Physical TFP	−0.0186	0.0125	−0.0142	0.0383
	(0.0050)	(0.0047)	(0.0144)	(0.0177)
Price	−0.0033	−0.0015	−0.0055	−0.0388
	(0.0031)	(0.0028)	(0.0080)	(0.0141)
Demand shock	−0.3586	−0.3976	−0.5903	−0.2188
	(0.0228)	(0.0224)	(0.0968)	(0.1278)

TABLE 5—EVOLUTION OF PRODUCTIVITY, PRICE AND DEMAND WITH AGE EFFECTS

Variable	Plant age dummies: Exit	Entry	Young	Medium
	Unweighted regressions			
Traditional TFP	−0.0211	0.0044	0.0074	0.0061
	(0.0042)	(0.0044)	(0.0048)	(0.0048)
Revenue TFP	−0.0220	0.0133	0.0075	0.0028
	(0.0044)	(0.0047)	(0.0051)	(0.0053)
Physical TFP	−0.0186	0.0128	0.0046	−0.0039
	(0.0050)	(0.0053)	(0.0058)	(0.0062)
Price	−0.0034	0.0005	0.0029	0.0067
	(0.0031)	(0.0034)	(0.0038)	(0.0042)
Demand shock	−0.3466	−0.5557	−0.3985	−0.3183
	(0.0227)	(0.0264)	(0.0263)	(0.0267)
	Weighted regressions			
Traditional TFP	−0.0156	−0.0068	−0.0156	−0.0234
	(0.0127)	(0.0203)	(0.0171)	(0.0132)
Revenue TFP	−0.0191	−0.0038	−0.0180	−0.0165
	(0.0136)	(0.0200)	(0.0198)	(0.0131)
Physical TFP	−0.0142	0.0383	0.0056	−0.0050
	(0.0144)	(0.0186)	(0.0142)	(0.0135)
Price	−0.0049	−0.0421	−0.0236	−0.0114
	(0.0079)	(0.0147)	(0.0114)	(0.0096)
Demand shock	−0.5790	−0.2785	−0.3133	−0.3164
	(0.0972)	(0.1459)	(0.1695)	(0.1197)

TABLE 6—SELECTION ON PRODUCTIVITY OR PROFITABILITY?

Specification:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Traditional TFP	−0.073						
	(0.015)						
Revenue TFP		−0.063					
		(0.014)					
Physical TFP			−0.040			−0.062	−0.034
			(0.012)			(0.014)	(0.012)
Prices				−0.021		−0.069	
				(0.018)		(0.021)	
Demand shock					−0.047		−0.047
					(0.003)		(0.003)
Controlling for plant capital stock							
Traditional TFP	−0.069						
	(0.015)						
Revenue TFP		−0.061					
		(0.013)					
Physical TFP			−0.035			−0.059	−0.034
			(0.012)			(0.014)	(0.012)
Prices				−0.030		−0.076	
				(0.018)		(0.021)	
Demand shock					−0.030		−0.029
					(0.004)		(0.004)
Capital stock	−0.046	−0.046	−0.046	−0.046	−0.023	−0.046	−0.023
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)

"Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity"
Jan De Loecker (2011)

Overview

- Studies effects of trade liberalization on Belgian textiles producers
- Develops strategy to disentangle price and productivity effects
- We see only 2% productivity gains rather than 8% after separating out price effects.

Disappearing Quotas

TABLE I

NUMBER OF QUOTAS AND AVERAGE QUOTA LEVELS (IN MILLIONS)

	Number of Quota Protections	kg		No. of Pieces	
		No. of Quotas	Level	No. of Quotas	Level
1994	1,046	466	3.10	580	8.58
1995	936	452	3.74	484	9.50
1996	824	411	3.70	413	7.95
1997	857	413	3.73	444	9.28
1998	636	329	4.21	307	9.01
1999	642	338	4.25	304	10.53
2000	636	333	4.60	303	9.77
2001	574	298	5.41	276	11.06
2002	486	259	5.33	227	12.37
Change	−54%	−44%	72%	−60%	44%

- Meanwhile, Belgian textile prices declined by 15%

Model

- Cobb-Douglas production function:

$$Q_{it} = L_{it}^{\alpha_l} M_{it}^{\alpha_m} K_{it}^{\alpha_k} \exp\left(\omega_{it} + u_{it}\right)$$

- As usual, Q_{it} is not observed, but sales R_{it} is.
- Assumed demand system:

$$Q_{it} = Q_{st} \left(\frac{P_{it}}{P_{st}}\right)^{\eta_s} \exp\left(\xi_t\right)$$

where Q_{st} is a sectoral aggregate demand shifter

Model

- Demand is CES with monopolistic competition for each sector with markup $\left(\frac{\eta_s}{\eta_s+1}\right)$. Revenue is $R_{it} = Q_{it}P_{it}$, and at the optimal price,

$$R_{it} = Q_{it}^{(\eta_s+1)/\eta_s} Q_{st}^{-1/\eta_s} P_{st} \left(\exp\left(\xi_{it}\right)^{-1/\eta_s}\right).$$

- Expanded revenue equation (in logs):

$$\tilde{r}_{it} = \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_s Q_{st} + \omega_{it}^* + \xi_{it}^* + u_{it}$$

- Estimating equation:

$$\tilde{r}_{it} = \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_s Q_{st} + \delta D + \tau q r_i t + \omega_{it}^* + u_{it}$$

  where D is a vector of demand-shifting dummy variables and $qr_{it} \in [0, 1]$ is a measure of exposure to quota protection.

- See paper for treatment of multi-product firms

Quotas and inversion

- $\omega_{it} = g_t\left(\omega_{i,t-1}, qr_{i,t-1}\right) + \nu_{it}$
- Inversion:
$$\omega_{it} = h_t\left(k_{it}, m_{it}, qr_{it}, q_{st}, D\right)$$
- Checking the monotonicity condition for a static input (as LP) is straightforward, but verifying the monotonicity of investment (OP) is harder.
- Estimation based on exclusion restrictions on innovation in productivity (what was ξ in previous papers but ν here:
$$E\left\{\nu_{i,t+1}\left(\beta_m, \beta_k, \beta_s, \tau, \delta\right)\begin{pmatrix} m_{it} \\ k_{i,t+1} \\ q_{st} \\ qr_{i,t+1} \\ D \end{pmatrix}\right\} = 0$$

Separation

- This framework allows for separate effects of quotas on productivity through g and through demand throuh τ
- Identifying assumption: protection can only affect productivity with a lag (note $g_t\left(\omega_{i,t-1}, qr_{i,t-1}\right)$, while current quota protection can impact prices through residual demand.

Results

TABLE VIII

IMPACT OF TRADE LIBERALIZATION ON PRODUCTIVITY[a]

Approach	Description	Estimate	Support
I	OLS levels	−0.161* (0.021)	n.a.
II.1	Standard proxy-levels	−0.153* (0.021)	n.a.
II.2	Standard proxy-LD	−0.135* (0.030)	n.a.
III	Adjusted proxy	−0.086 (0.006)	[−0 129 −0 047]
IV	Corrected	−0.021 sd: 0.067	[−0 27 0 100]
V	Corrected LD	−0.046** (0.027)	n.a.

[a] I report standard errors in parentheses for the regressions, while I report the standard deviation (sd) of the estimated nonparametric productivity effect in my empirical model (given by $g(\cdot)$). * and ** denote significant at 5 or lower and 10 percent, respectively. LD refers to a 3 year differencing of a two-stage approach where Approach II.2 relies on standard productivity measures, as opposed to Approach V, which relies on my corrected estimates of productivity.