8-20-2007

Exploiting Sequential Learning to Estimate Establishment-Level Productivity Dynamics and Decision Rules

David Greenstreet
Oxford University, dgreenst@andrew.cmu.edu

ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

EXPLOITING SEQUENTIAL LEARNING TO ESTIMATE ESTABLISHMENT-LEVEL PRODUCTIVITY DYNAMICS AND DECISION RULES

David Greenstreet

Number 345

August 2007

Manor Road Building, Oxford OX1 3UQ

Exploiting Sequential Learning to Estimate Establishment-Level Productivity Dynamics and Decision Rules

David Greenstreet [1,2]

August 20, 2007

[1]The author thanks Kai-Uwe Kühn, Francine Lafontaine, James Levinsohn, and Serena Ng for extensive advice and guidance reflected throughout this paper. Jagadeesh Sivadasan and Ivan Kandilov provided documentation and advice on using the Chilean establishment data. Daniel Ackerberg, Patrick Bajari, Susanto Basu, Steven Berry, Stephen Bond, Matias Busso, Steven Klepper, Aviv Nevo, John Van Reenen, Margaret Slade, Jan Sokolowsky, Gary Solon, and Kyoung-Soo Yoon also provided useful discussion. Earlier versions of this paper were presented at the 2005 International Industrial Organization Conference, the 2007 Annual Conference of the Royal Economic Society, and in seminars at the University of Michigan, UCLA, U.S. Federal Reserve Board of Governors, University of Oxford, Northwestern University, LSE Centre for Economic Performance, Carnegie Mellon University, and Warwick University. The author gratefully acknowledges financial support in the form of One-Term Dissertation Grant and Predissertation Fellowship from the Horace H. Rackham School of Graduate Studies, University of Michigan. Any remaining errors or omissions are the author's sole responsibility.

[2]Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ; david.greenstreet@economics.ox.ac.uk

Abstract

This paper develops a sequential learning estimator of production functions and productivity dynamics for unbalanced establishment panels. Extending an idea from the literature on dynamic industry models, establishments are uncertain about their own idiosyncratic productivities and update productivity beliefs using information revealed by their production experience. The estimator relies on the structure of this iterative learning process and thereby avoids placing any restriction on establishment strategic behavior. Consequently, the estimator is suitable for comparative studies of the behavioral sources of technological change across all types of industry. Estimation of productivity dynamics and of behavioral decision rules are separated into recursive stages. Using sequential learning estimates of productivity beliefs from the first stage, decision rules for exit, investment, and innovation effort can be estimated in a second stage. A test application with four Chilean industries confirms that the estimator produces plausible estimates with small standard errors. Decision rule estimates show that productivity beliefs affect investment and exit hazards in the expected direction.

JEL Classification: D24, L25, C23, D83, L60.
Keywords: microeconomic productivity dynamics, unbalanced panels, sequential learning, exit hazard, investment rates, Chilean manufacturing.

Establishment productivity levels and dynamics, their causes, and their effects on behavior are crucial in the analysis of many important economic phenomena. Microeconomic productivity growth, entry and selective exit, and innovative activities motivated by productivity improvement are the sources of aggregate technological change. Most explanations of within-industry heterogeneity, turnover, and evolution of industrial structure rely upon differences in capability and firm responses to those differences. How productive entrants are in comparison with incumbents and how establishments' productivity improves in response to experience, R&D, and other forms of innovation effort are important issues in studies of entrepreneurship and innovation. Many assessments of the benefits of policies in trade, development, and regulation use estimates of firm or establishment productivity with and without the policies.

Unfortunately, productivity[1] is not directly observable and must be estimated before any of these questions can be addressed empirically. From the very beginning, the literature on estimating production functions and productivity has recognized the econometric difficulties arising when firms' knowledge of their own productivities influences input choices and when only revenues and not physical outputs are observed (i. e. Marschak and Andrews [1944]). Since panel datasets have become readily available, selective exit has also been recognized as a potential source of bias. This paper develops and illustrates a new technique for addressing both the endogenous input and selective exit sources of bias.

The main contribution is an estimator of production functions and productivity dynamics based on sequential learning by firms. It begins with the observation that firms are uncertain about their own idiosyncratic (establishment-specific) productivity. Consequently, they base their decisions about entry/exit, investment, production inputs, and innovation effort on imprecise beliefs about what their actual productivity will be in the present and future periods. At the end of each period actual productivity can be computed from observed inputs and production. The difference between actual and expected productivity is a forecast error. Independence of the productivity forecast errors from predetermined input and other decisions provides orthogonality conditions that can be used in estimation. Olley and Pakes (1996) and Levinsohn and Petrin (2003) also use productivity forecast errors, but they infer an establishment's productivity expectations from investment or intermediate input decision rules, respectively. Here the approach is to model the learning process that produces those productivity beliefs. As an establishment observes its productivity forecast error at the end of each period, it uses the new information contained in that error to update its expected productivity in future periods. This is very close in spirit to Jovanovic's (1982) seminal dynamic industry model, which is driven by firm learning about an idiosyncratic cost factor. However, this is the first time a model of learning has been used in an empirical study of productivity.

Conditional on the values of production and productivity parameters, the sequential learn-

[1]Throughout this paper, productivity refers to the production function residual variously referred to as (firm or establishment) total factor productivity, multi-factor productivity, or (in some papers concerned with micro to macro aggregation) technology. As detailed below, this paper is focuses on the persistent, establishment-specific portion of productivity.

ing estimator reconstructs the iterative productivity learning process in order to recover the forecast errors used in estimation. When output and inputs are observed and the assumptions described below, particularly regarding information and decision timing, are satisfied, this model of learning about idiosyncratic productivity provides sufficient structure for identification of parameters.[2] No use is made of first order conditions, or of dynamic industry equilibria characterizations that limit the nature of potentially complex strategic interactions. Furthermore, the form of productivity dynamics can be specified in very flexible ways, including multi-dimension processes and effects of observable measures of innovation effort.

In contrast, many productivity estimators use characterizations of equilibrium firm behavior in order to create identifying structure and to solve selection and endogeneity (Arellano and Bond [1991], Arellano and Bover [1995],[3] Olley and Pakes [1996], and Levinsohn and Petrin [2003]). These characterizations are often introduced implicitly through first-order conditions.[4] Some other estimators dispose of selection and endogeneity issues by imposing a specific form on the dynamics of idiosyncratic productivity (e. g. fixed effects and the estimator proposed by Cornwell *et al* [1990]). Existing estimators that do not restrict firm strategic behavior or limit productivity dynamics are either infeasible in the microeconomic productivity applications considered here (e. g. traditional, strictly exogenous instruments with panel level variation are almost impossible to find) or suffer from selection and/or endogeneity biases (e. g. OLS, random effects, and Blundell and Bond's [1998, 2000] proposed extension of dynamic panel GMM estimators[5]).

With the sequential learning estimator, establishments' strategic decision rules on exit, investment, innovation effort and other behaviors can be estimated in a separate second stage. The sequential learning estimator not only avoids any restrictions on establishment behavior, it also produces estimates of each firm's belief about establishment productivity in addition to estimates of productivity itself. Thus it enables recursive estimation of production functions and productivity dynamics first, followed by empirical decision rules for any behavior, including sources of technological change. Previously these had to be addressed as parts of a single simultaneous dynamic system.[6]

[2] When revenues rather than outputs are observed, the sequential learning estimator still remains applicable for industries with uniform output prices. In other industries a revenue function interpretation can be applied to the estimates, although the information assumptions become less plausible. Simultaneous application of sequential learning and demand system estimation awaits further research.

[3] On their surface, dynamic panel GMM estimators as epitomized by Arellano and Bond and Arellano and Bover do not appear to rely on characterizations of firm behavior. Nevertheless, their constructed instruments — lagged levels of inputs instrumenting current input differences — only have identifying power when firm input choices exhibit adjustment lags or some other form of momentum.

[4] As in for example, cost function and factor demand techniques such as Slade (1989), Nadiri and Prucha (2001), and references therein; nonparametric extensions of the Solow residual reviewed in Hulten (2001); and macroeconomic estimates of cyclical productivity, scale economies, and utilization rates including Hall (1988), Basu (1996), Basu and Kimball (1997), and Basu and Fernald (2002).

[5] The instruments Blundell and Bond add to the dynamic panel estimator — lagged input differences as instruments for levels — are not uncorrelated with the fixed effects portion of the error term when endogeneity takes the complex strategic form with lags and leads described below.

[6] Such a simultaneous system approach has been adopted in a number of empirical implementations of dynamic industry models. In the past these have been calibrations of strongly simplified (but still complex)

The remainder of this introduction is a more detailed discussion of how the sequential learning estimator is implemented and resolves the issues of selection and endogeneity. Selection and endogeneity are standard econometric challenges when estimating production functions and productivity from unbalanced panels of establishments. Selection occurs because establishments are more likely to close when they believe their production capabilities are low. Thus, small to mid-size establishments that do stay in the panel will have greater idiosyncratic productivity on average than erstwhile establishments of the same size that exit the panel.

Endogeneity issues are especially challenging in the context of a dynamic strategic equilibrium among firms in an industry. The idiosyncratic component of productivity is part of the production function residual. At the same time, it is a state variable in the dynamic industry game and can therefore be related to capital, which is also a state variable, in very complex ways. Specifically, capital and any other quasi-fixed inputs may be correlated with idiosyncratic productivity not just contemporaneously, but at any lag and across establishments. Because investment is forward-looking, capital may also be related to the forecastable component of future productivity at any lead. These relationships will generally be nonlinear and may even be non-monotonic.[7] I refer to this particularly challenging form of endogeneity as strategic endogeneity because it emerges from strategic interactions among firms.

The sequential learning estimator addresses selection and strategic endogeneity by explicitly modelling establishments' productivity forecast errors. Collectively, the assumptions made about how establishments form productivity beliefs imply that the econometrician can reproduce those beliefs conditional on parameter values. There are six major assumptions. The first is an information assumption that firms update beliefs about their own productivity (as if) using the same input and production data available to the econometrician. The second is additive separability of the deterministic portion of the production function from the persistent portion of productivity and from transient noise. This implies that learning about an establishment's own productivity is passive.[8] The signal contained in a period's forecast error will be the same regardless of any other establishment's actions and for any actions, except entry and exit, taken by the establishment itself. The third is a timing assumption that all production inputs for a period are predetermined before any information about that period's forecast error is revealed. Fourth, all of the random components of the model are normally distributed. Normality will be used in formal justification of the estimator, but its practical significance is modest. In the absence of normality the Kalman updating formula used will still provide the best linear estimates of idiosyncratic productivity. Fifth,

versions (Hopenhayn and Rogerson [1993] and Jovanovic and MacDonald [1994a, 1994b]) or have faced immense computational burdens (Ericson and Pakes [1995], Gowrisankaran and Town [1997], Pakes and McGuire [2001], Doraszelski and Judd [2004], Pakes [2000] and references therein). The sequential learning estimator is actually complementary with the more recent literature on two-step semiparametric estimation of industry games (Berry and Pakes [2000], Pakes *et al* [2004], Bajari *et al* [2005], Ackerberg *et al* [2005]) because these rely upon observability of firm and industry states, including productivity or productivity beliefs.

[7] For examples of specific dynamic equilibria that show non-monotonicity between an index of competitive ability (e.g. productivity) and innovation effort see Budd *et al* (1993) and Ericson and Pakes (1995).

[8] This does not preclude active or strategic learning about other objects in the dynamic industry game.

in order to restrict the firms' learning problem to the value of their establishments' persistent idiosyncratic productivity, they are assumed to know the true values of all parameters. Knowledge of true parameters is a common assumption in dynamic industry models and in structural estimation of industry games. The sixth is a rational expectations assumption that entrants' initial productivity beliefs have the same distribution as the data generating process producing actual initial productivities for the entrants' cohort. Several of these can be relaxed.

The information and timing assumptions are strong, but have a testable implication that all input decisions (investment, employment, etc.) will be independent of all information about productivity revealed by data from after the decision is made. If not, the establishment is somehow acquiring and using this information sooner than the sequential learning estimator assumes.[9] Tests of this implication can be implemented by estimating an establishment decision rule as a function of estimated productivity belief and control variables. The difference between the best estimate of productivity using all available data and the estimated productivity belief should have no statistically significant effect when inserted into the estimated decision rule.[10] This provides a way to assess the validity of the assumptions above in any specific application.

The sequential learning estimator is implemented by simulating establishments' learning process. An establishment's input and output data enter beliefs[11] via realizations of the productivity forecast error. Specifically, at the start of each period an establishment has a mean belief about what its idiosyncratic productivity will be. The establishment knows its level of capital and selects variable input quantities. At this point the establishment can forecast its output. Note that this forecast is conditional on all the factor inputs, however chosen, as well as on any past information embodied in the belief. Once actual output is observed, the difference between predicted and realized output is the forecast error, which the establishment uses to update its belief about the persistent component of its productivity. Then this updated belief is used to forecast productivity in the next period. To begin simulation of the learning process, initial beliefs of entrants are parameterized by cohort. These parameterized initial beliefs are identified from cross-sectional variation using the rational expectations assumption.

This process of iteratively updated beliefs forms a Kalman filter, a widely studied type of stochastic process. In the sequential learning estimator the Kalman filter plays a dual role. First, it is a model of the establishments' learning and productivity beliefs. In addition

[9]If not correctly specified, innovation effort or embodied technology can also break this independence. In these cases the direction of causality is reversed — behavior to productivity.

[10]The Kalman filter provides a method, known as smoothing, for computing the best estimate of productivity using all available data. This will be described in the section on estimation.

[11]With the normality assumption (see footnote **??**), means and variances are sufficient to characterize an establishment's uncertain belief about its productivity. Furthermore, from the establishment's perspective belief variances are independent of the data because they are determined by parameters assumed known to the establishment (again, see footnote **??**). Thus only belief means are computed directly from the data. Variances enter indirectly through the belief updating process.

to the important example of Jovanovic (1982), the Kalman filter has been used in several other fields as a model of agents' learning. Empirically, several recent papers on signaling versus human capital in wage determination use a filter to represent firms' learning about employee ability (Farber and Gibbons [1996], Altonji and Pierret [2001], Lange [2003]), but these papers do not actually compute the filter.[12] Second, the sequential learning estimator uses the Kalman filter to implement its econometrics. Kalman filters have often been used to estimate latent variables, including at least one application (Slade [1989]) to aggregate industry productivity. However, in these applications it is only the econometrician, and not any modelled agent, who is learning about the latent variables.

The application of the Kalman filter in this paper has four unusual features. First, the econometrician and the modelled establishments are both learning about unobserved true productivity from the same data processed through the same filter. However, the econometrician has an additional layer of uncertainty because model parameters, which are known to the establishments, must be estimated. Second, separate filters for each establishment in the panel are pooled to estimate common parameters. Third, the distribution initializing an entrant's filter has a structural interpretation as the distribution of productivities in the entrant's cohort. This is identified from cross-sectional, within-cohort variation and need not be related to an unconditional expectation of the stochastic process governing incumbents' productivity dynamics. Finally, attributing the filter forecast errors to establishments' productivity expectations provides a means of resolving selection and strategic endogeneity.

In the estimation algorithm, forecast errors from each establishment's filter are joint functions of the model parameters and its history of inputs and outputs through the current period. The identifying condition is that these forecast errors are information shocks that must be independent of all previous establishment decisions, including (lack of) exit and production inputs from the same period. Orthogonality of forecast errors to inputs and exit is the basis of the sequential learning estimator's solution to strategic endogeneity and selection. Assuming normality, this can be expressed as a likelihood in terms of successive forecast errors and their variances. Since the Kalman filter repeatedly conditions on past information, as embodied in productivity beliefs, estimation is implemented by inserting these likelihoods into a prediction error decomposition maximum likelihood algorithm.

This approach to estimation gives the sequential learning estimator several major advantages. First, equilibrium of the dynamic industry game does not have to be characterized or restricted in any way. No use is made of equilibrium solutions, establishment Bellman values, best-response policy functions, or establishment first-order conditions of any kind. In fact, the estimator is robust to multiple equilibria, collusion, and out-of-equilibrium behavior because the forecast errors are independent from inputs no matter how they are chosen. Second, productivity dynamics can be specified in a wide variety of ways, including, if desired, the effects of variables measuring innovation efforts. Third, the cohort structure preserves

[12] Sargent's work on relaxing the assumption of rational expectations in estimation and simulation of macroeconomic equilibria (e. g. Sargent [1993], Sargent *et al* [2004]) is another notable example of an empirical model with learning agents.

more inter-establishment variation than alternative solutions to endogeneity do. Griliches and Mairesse (1995) conclude that loss of valid variability is the main reason existing production function estimation methods often produce unsatisfactory results. Finally, this method produces predictions of both establishment beliefs and persistent components of productivity, which can be used in modelling of establishment strategic behavior or in decompositions of aggregate industry productivity changes.

These advantages make the sequential learning estimator particularly well-suited to be the first part of a multi-stage approach to empirical modelling of co-evolving industry structure and technology. A second stage can model a firm's chosen actions as functions of productivity beliefs, the distribution of establishment productivities, industry characteristics, and information about exogenous economic conditions (e. g. demand). In a general dynamic industry game framework, actions include variable inputs, investment, entry/exit, and purposeful innovation and imitation (R&D, patents, entrepreneurship, business process redesign, etc.). Second-stage estimates may be exploratory or impose/test restrictions implied by specific equilibrium solutions or characterizations. In particular, comparative studies of strategic behavior become feasible across industries with varying structures, or when conditions vary over time within a given industry.

Application of the sequential learning estimator to panel data for four Chilean manufacturing sectors — food products, apparel, paper products, and fabricated metals — confirms that the estimator produces plausible parameter estimates with standard errors that compare favorably to other estimators. The results section will also show that the estimated productivity beliefs are indeed useful in exploratory second-stage estimation of decision rules for entry/exit, investment, and labor inputs.[13] The validity tests for the major assumptions are passed in three out of the four industries. Finally, a condition closely related to Olley and Pakes' (1996) investment invertibility condition is frequently violated in all four industries when sequential learning is used to estimate productivity beliefs. This accentuates the usefulness of a productivity estimator that does not rely on characterization of equilibrium behavior.

The next section presents the model, focusing in particular on the Kalman filter representation of sequential learning by firms. Section 2 discusses implementation of the estimator. The Chilean panel data is introduced in Section 3. Results in Section 4 include production function and productivity dynamics estimates from the sequential learning estimator as well as exploratory decision rule estimates for exit, investment, and labor. The conclusion summarizes major results and discusses advantages of the sequential learning estimator.

[13]Unfortunately, the Chilean data set used as an example lacks information on innovation effort.

1 The Sequential Learning Model

1.1 Production Function and Productivity Dynamics

Let i be an index of establishments and t indicate year. Establishments that began production in the same year belong to the same cohort, c, labelled by their first year of production. Let $t = 1$ be the first year in the panel data set, and let $t = T$ be the last period in the data. Define $T_i = \min[\text{last year of establishment } i\text{'s operation}, T]$ as the last year of observed data for establishment i. Abusing notation, partition the set of establishments $i \in \{1, 2, \ldots, I\}$ into subsets of cohorts starting at or after $c = 1$, i. e. $I_1, I_2, \ldots, I_T$, plus incumbent establishments already operating before the first period of data, $I_0 \equiv \{i|c < 1\}$.

Each active establishment will produce output, q_{it}, in year t using variable inputs, $\boldsymbol{\ell}_{it}$, such as labor, and quasi-fixed inputs, k_{it}, such as capital. Investment, i_{it}, will change capital in future periods and may incur adjustment costs. The input vector, $\mathbf{x}_{it}$, contains both $\boldsymbol{\ell}_{it}$ and k_{it}. In general terms, specify the production function as

$$q_{it} = F(\mathbf{x}_{it}; \Gamma_t^F) + \mathbf{z}'\boldsymbol{\mu}_{it} + \xi_{it} \qquad\qquad \xi_{it} \sim \text{NIID}(0, \sigma^2) \tag{1}$$

$F(.)$ is the deterministic portion of the production function and Γ_t^F is a parameter vector that may vary over time to reflect technological change. The total error, $\epsilon_{it} \equiv \mathbf{z}'\boldsymbol{\mu}_{it} + \xi_{it}$, includes a persistent productivity component, $\mathbf{z}'\boldsymbol{\mu}_{it}$, and a normal iid noise term, ξ_{it}. $\boldsymbol{\mu}_{it}$ may be a vector if idiosyncratic productivity has several components following separate dynamics. $\mathbf{z}$ is an $m \times 1$ aggregation vector and will generally be a column of 1's. Due to the dynamics of the $\boldsymbol{\mu}_{it}$ as described below, the total error, ϵ_{it}, will not be iid. In Kalman filter terminology this is known as the measurement equation.

The $\boldsymbol{\mu}_{it}$'s are unobserved stochastic state variables. Each period the establishment chooses a vector of innovation efforts, $\mathbf{e}_{it}$, such as R&D, to influence the future values of $\boldsymbol{\mu}_i$. $\boldsymbol{\mu}_{it}$ evolves according to the equation

$$\begin{aligned}
\boldsymbol{\mu}_{it} &= g(\boldsymbol{\mu}_{it-1}, \mathbf{e}_{it-1}, i_{it-1}, \mathbf{x}_{it-1}, \eta_{it}) \\
&\equiv R\boldsymbol{\mu}_{it-1} + G(\mathbf{e}_{it-1}, i_{it-1}, \mathbf{x}_{it-1}; \Gamma_t^G) + S\boldsymbol{\eta}_{it} \qquad \text{for } t > c \\
\boldsymbol{\eta}_{it} &\sim \text{NIID}(0, Q)
\end{aligned} \tag{2}$$

The stochastic productivity transition function $g(.)$ has three components. R is an $m \times m$ transmission matrix that captures the period-to-period persistence of $\boldsymbol{\mu}$. The $m \times 1$ vector function $G(.)$ describes how non-stochastic variables, including innovation effort, investment, and production inputs, affect next period's productivity.[14] $\boldsymbol{\eta}_{it}$ is an $n \times 1$ vector of shocks to persistent productivity, and the $m \times n$ matrix S permits the shocks on elements of $\boldsymbol{\mu}_{it}$ to be correlated. The covariance matrix, Q $(n \times n)$, also permits correlated structural shocks.

[14] Including investment and production inputs as arguments of the productivity transition equation permits a wider variety of productivity dynamics, such as capital-embodied technology or learning-by-doing.

Together, Equations 1 and 2 constitute the state space form (SSF) representation of the model.

Part of the power of an SSF representation is that alternative versions of the transition equation can accommodate a wide variety of productivity dynamics without modifying the econometric strategy. Furthermore, when variables such as R&D or production experience are included in the transition equation the productivity consequences of innovation effort can be incorporated directly into the specification. Appendix A provides a variety of examples.

Initial values of $\boldsymbol{\mu}_{ic}$ are assumed to come from a data generating process with distribution

$$\boldsymbol{\mu}_{ic} \sim \mathrm{N}(\mathbf{w}_c + G_c(\mathbf{e}_{ic-1}, \mathbf{s}^e_{ic-1}; \Gamma^c), W_c) \tag{3}$$

With $\mathbf{w}_c$ $(m \times 1)$ representing a cohort-specific mean of the entrants' initial productivity and W_c $(m \times m)$ representing the variance of initial productivities. $G_c(.)$ captures any effects on entrants' expected initial productivity of pre-entry efforts, $\mathbf{e}_{ic-1}$, such as preproduction training, or of observable characteristics, $\mathbf{s}^e_{ic-1}$, such as corporate ties or prior experience in related industries. The random components ξ_{it}, $\boldsymbol{\eta}_{it}$, and $\boldsymbol{\mu}_{ic} - G_c(.)$ are not only iid, they are also assumed to be independent from each other.

1.2 Timing of the Dynamic Industry Game

Each establishment is participating in a dynamic industry game. Firms' decisions about entry/exit, variable inputs, investment, and innovation efforts to improve productivity or product quality are motivated by expectations of future gain. These expectations, in turn, depend on the firm's perception of its current capabilities (including capital and belief about own productivity), the industry's current state (including distributions of capital and productivity), competitors' efforts, and the industry's external environment (market, policy, and technological). Dynamic industry equilibrium emerges as an internally consistent balance among co-evolution of industry structure and productivity; firms' incentive to enter, exit, invest, and innovate; and the actions they select. There is an extensive theoretical literature on dynamic industry games.[15]

Figure 1 describes the timing of events in a generic industry game among firms with heterogeneous capabilities. Three points about this game will be relevant for development of the sequential learning estimator. The most important of these is the timing of firm decisions and information revelation described below. Second, the nature of dynamic industry

[15] Jovanovic (1982) and Ericson and Pakes (1995) are especially notable. Also see Spence (1981), Lippman and Rumelt (1982), Jovanovic and Rob (1987), Klepper and Graddy (1990), Lambson (1991, 1992), Beggs and Klemperer (1992), Hopenhayn (1992), Hopenhayn and Rogerson (1993), Jovanovic and MacDonald (1994a, 1994b), Pakes and McGuire (1994), Klepper (1996), Sutton (1998), Petrakis and Roy (1999), and Asplund and Nocke (2002). Some papers on more specialized topics including dynamic duopolies (Budd *et al* [1993] and Cabral and Riordan [1994]), repeated patent races (Reinganum [1985]), and innovation in endogenous growth models (Grossman and Helpman [1991], Aghion and Howitt [1998], and Aghion *et al* [2001]) can be interpreted as special cases.

equilibrium provides guidance about which variables should enter firms' decision functions for entry, exit, investment, innovation, and variable inputs. Third, because the sequential learning estimator does not rely on any description of strategic behavior the dynamic industry equilibrium does not need to be solved or characterized. This means that details of the industry game can remain unspecified and may be almost arbitrarily complex, facilitating comparisons across industries with widely varying characteristics.

Let $\mathbf{s}_{it} \in \mathcal{S}$ be a vector state variable indexing establishment i's competitive capabilities at period t. The elements of $\mathbf{s}_{it}$ include capital stock, k_{it}, and persistent idiosyncratic productivity, $\boldsymbol{\mu}_{it}$. Any other persistent establishment characteristic affecting profitability, such as product attributes, inventories, or owner's characteristics, are also included. Variation in $\mathbf{s}_{it}$ is the source of all heterogeneity. Summarize the state of the entire industry with a measure ν_t of all establishment capabilities over $\mathcal{S}$, the set of all possible capability levels. The productivity component of $\mathbf{s}_{it}$ is not directly observable. Therefore, beliefs about the values of $\mathbf{s}_{it}$ and ν_t will replace the true values as the state variables on which strategic decisions will be based.[16] Unlike in many dynamic industry models, $\mathbf{s}_{it}$ and specifically productivity are of direct interest, rather than simply a means for explaining heterogeneity in establishment behavior and characteristics.

The time-line in Figure 1 illustrates the assumed sequence of events during a single period. First, outputs, q_{it} and gross period payoffs, $\pi_{it} = \pi(\boldsymbol{\ell}_{it}, \boldsymbol{\ell}_{-it}, \mathbf{s}_{it}, \nu_t, \mathbf{n}_t, \xi_{it}, \xi_{.t})$, are realized. These payoffs are a function of actions with immediate consequences, $\boldsymbol{\ell}_{it}$ and $\boldsymbol{\ell}_{-it}$, taken by the establishment and its competitors, respectively. Such actions will include the level of variable inputs, such as labor and materials, but may also include price-setting, advertising, any effort-free variations in product mix, and so forth. Payoffs also depend on the (true) state of the establishment and the industry. The vector $\mathbf{n}_t$ indicates relevant aspects of the industry's environment, such as demand, input supplies, and government regulations. Finally, ξ_{it} is the establishment's own independent random shock from Equation 1 and $\xi_{.t}$ is a vector of its competitors' shocks.

Returning to the time-line, once payoffs are realized establishments update beliefs about the components of $\mathbf{s}_{it}$ and ν_t, such as productivity, that are not directly observable. The model of how establishments update beliefs about the productivity component, $\boldsymbol{\mu}_{it}$, of their own state, $\mathbf{s}_{it}$, is at the core of the sequential learning estimator. The next sub-section will provide details. Immediately after productivity beliefs are updated any news in the information set, Ω_t, describing exogenous conditions is revealed. Beliefs about the environment, $\mathbf{n}$, in future periods are revised using Ω_t. Ω_t may or may not be larger than $\mathbf{n}_t$.

Next, establishments make simultaneous strategic decisions. The actions selected include the $\boldsymbol{\ell}_{it+1}$'s that enter directly into next period's payoff function, but they also include decisions that affect future payoffs indirectly by changing the establishment and industry states. These

[16] Beliefs about ν_t need not be common in order to estimate productivity with the sequential learning estimator or to fit empirical decision rules in a second stage. However, a common belief is the simplest informational assumption permitting a policy function interpretation of estimated decision rules.

include entry, exit ($X_{it} = 1$), investment in quasi-fixed capital, innovation effort, and any other costly activity motivated by its likely effect on future profitability. Effort will have a favorable impact on evolution of the establishment's productivity via the productivity transition function, $g(.)$, defined in Equation 2.

Completing the time-line, at the end of period t transition of establishments' states, $\mathbf{s}_i$, industry state, ν, and industry environment $\mathbf{n}$ to their values in $t+1$ occurs. The transition of establishments' states is defined by $g(.)$ and the standard law-of-motion for capital. Aggregating establishment transitions and adjusting for entry and exit produces the change in industry state. Industry environment transitions are exogenous.

1.3 Establishment Learning with a Kalman Filter

The establishments will form beliefs about their ability through a learning process consisting of iterative Bayes updating from production experience, beginning with a prior initial belief held before production begins. Assume:

Knowledge of True Parameters: All establishments know the true values of the parameters — Γ^F_t for all t, Γ^G, $\mathbf{z}$, R, S, σ^2, and Q.[17]

Passive Learning: The signal about own productivity received from a period's production experience will be the same regardless of any other establishment's actions and for any actions, except entry and exit, by the establishment itself. This is already implicit in the additively separable way productivity enters the production function in Equation 1.

Information: All establishments update their productivity beliefs (as if) their information is the same as that revealed in the data reported through t. This implies that the econometrician can reproduce those beliefs conditional on parameter values.

Rational Expectations: All entrants in a cohort have a common prior for their first period productivity consistent with the data generating process for that cohort's $\boldsymbol{\mu}_{ic}$'s. Implicitly this extends knowledge of true parameters to Γ^c, $\mathbf{w}_c$, and W_c.[18]

Two other assumptions have already been introduced:

Decision Timing: Establishments' decisions on exit, investment, and innovation effort between $t-1$ and t, and on all inputs during t, are taken before any information from period t is revealed. In econometric terminology, these decisions are predetermined.

[17]This does not include the $\boldsymbol{\mu}_{it}$'s, which are not parameters.
[18]This assumption is common in dynamic industry models.

Normality: Noise, productivity shocks, and the data generating process for entrants' initial productivities are each distributed iid normal.

With normal errors and the linear form assumed in Equations 1 and 2, each establishment's iterative productivity learning process is a univariate Kalman filter.[19] The decision timing and normality assumptions together with linearity are sufficient to endow the Kalman filter with some very useful properties. In particular, Kalman productivity predictions will be conditionally Gaussian and the minimum mean square estimator (MMSE) of $\boldsymbol{\mu}_{it}$. If the normality assumption is dropped Kalman filter predictions are no longer MMSE, but they are still the best (conditional) linear predictions of productivity. This is not quite enough to formally justify the sequential learning estimator. There may be a better nonlinear estimator of productivity, and decision rules for current period inputs are also nonlinear functions of the information set. In principle these could be correlated, but there is no obvious reason that their nonlinearities should be related. Therefore, in practice the consequences of non-normality will usually be small.[20]

The remaining four assumptions — knowledge of true parameters, passive learning, information, and rational expectations — allow the econometrician to reconstruct establishment productivity beliefs conditional on parameter values. This reconstruction of establishment beliefs is what gives the Kalman filter its dual nature in the sequential learning estimator. It is both a model of establishments' learning process and a means to compute their beliefs as a function of the parameters to be estimated.

Knowledge of true parameters ensures that establishment learning about productivity is not confounded by parameter uncertainty. Parameter uncertainty would make the learning problem nonlinear and introduce issues about what additional information the establishment uses to infer parameter values. Passive learning separates productivity learning from all strategic considerations. However, this is stronger than what is actually required. The key is that the information set an establishment uses to form productivity beliefs is well-defined. Alternative assumptions about the information set could have been used.[21] The information assumption is crucial. It states that the information set each establishment uses in forming productivity beliefs is observable. Rational expectations provides a means to identify entrants' initial productivity beliefs by connecting those beliefs to estimable features of the productivity data generating process in Equation 3. This is the point at which having a panel of establishments is extremely useful.

[19] Harvey [1989], especially Chapter 3, provides a detailed introduction to Kalman filters.

[20] When non-normality is a serious concern a GMM estimator could replace the MLE estimator developed here. GMM would use orthogonality conditions between the productivity forecast error and predetermined establishment decisions. Alternatively, the quasi-maximum likelihood estimator would remain equivalent to weighted nonlinear least squares. Both versions of the sequential learning estimator embed the Kalman filter representation of establishment learning.

[21] For example, correlated productivity shocks with a specified degree of observability of competitors' experience, or active learning because productivity is not additively separable in the production function

To apply the Kalman filter to an establishment's productivity learning problem, define $E_t[\boldsymbol{\mu}_{it}] \equiv \mathbf{u}_{it}$ and $E_{t-1}[\boldsymbol{\mu}_{it}] \equiv \mathbf{u}_{it|t-1}$ where the subscript on the expectation operator indicates conditioning on information available up to and including t or $t-1$, respectively. Also, let P_{ct} and $P_{ct|t-1}$ denote corresponding variances. All establishments in the same cohort will have identical belief variances so the cohort index c replaces the i index. Refer to the data with $y_{it} \equiv \{q_{it}, \mathbf{x}_{it}, \mathbf{e}_{it-1}, i_{it-1}\}$ and $\tilde{y}_{it} \equiv \{\mathbf{x}_{it}, \mathbf{e}_{it-1}, i_{it-1}\}$ for establishment i in period t. Define $Y_{it} \equiv \{y_{ic}, y_{ic+1}, \ldots, y_{it}\}$, summarizing i's data for all periods through t, and $Y_t \equiv \{Y_{1t}, Y_{2t}, \ldots, Y_{It}\}$, summarizing all data through t. Finally let $X_{it-1} = 0$ indicate the decision not to exit before t.

Because the entire Kalman filter is conditional on the establishment's initial beliefs, identification of $\mathbf{w}_c$, Γ^c, and W_c is required. The rational expectations assumption provides the necessary link between initial beliefs and observed data.

$$\mathbf{u}_{ic|c-1} = \mathbf{w}_c + G_c(\mathbf{e}_{ic-1}, \mathbf{s}^e_{ic-1}; \Gamma^c) \qquad\qquad P_{cc|c-1} = W_c \tag{4}$$

For establishments in I_0 learning, as well as production, began before the observed data. Thus for a cohort starting with a common initial belief at $c < 1$, each establishment will have its own updated belief about its idiosyncratic ability, $\mathbf{u}_{i0}$, derived from experience before period 1. This problem is avoided here by dropping these old establishments from the estimation. This is legitimate because sequential learning estimates are not biased by selection based on any predetermined or exogenous characteristic. A subsequent paper will deal with the issues raised by including pre-existing establishments in the estimation procedure.[22]

Turning to the standard Kalman filter equations, each iteration from $t-1$ to t, separately for each establishment i, begins with prediction equations

$$\mathbf{u}_{it|t-1} = R\mathbf{u}_{it-1} + G(\mathbf{e}_{it-1}, i_{it-1}, \mathbf{x}_{it-1}; \Gamma^G_t) \tag{5}$$

With variance

$$\mathrm{var}(\boldsymbol{\mu}_{it} - \mathbf{u}_{it|t-1}) \equiv P_{ct|t-1} = RP_{ct-1}R' + SQS' \qquad\qquad \text{for } i \in I_c \tag{6}$$

Compare this to Equation 2. These are the best linear predictions and, with normal disturbances, also the MMSE's. Therefore, conditional on the available information unobserved productivity is distributed

$$(\boldsymbol{\mu}_{it} | Y_{it-1}, X_{it-1} = 0, \tilde{y}_{it}) \sim N(\mathbf{u}_{it|t-1}, P_{ct|t-1}) \tag{7}$$

Unlike the common data generating process in Equation 3 shared by all members of a cohort, this distribution is unique to each establishment because learning based on random realizations of the data has begun.

[22] The alternative approach is to treat the u_{i0}'s and P_{c0}'s as estimable parameters. One serious drawback to parameterizing beliefs of old establishments at $t = 1$ is that additional parameters must be estimated for each establishment added to the sample.

Once production experience in period t is observed, the establishment can compute its prediction error

$$\begin{aligned} v_{it} &\equiv q_{it} - \mathrm{E}_{t-1}[q_{it}] \\ &= [F(\mathbf{x}_{it}; \Gamma_t^F) + \mathbf{z}'\boldsymbol{\mu}_{it} + \xi_{it})] - [F(\mathbf{x}_{it}; \Gamma_t^F) + \mathbf{z}'\mathbf{u}_{it|t-1}] \\ &= (\mathbf{z}'\boldsymbol{\mu}_{it} - \mathbf{z}'\mathbf{u}_{it|t-1}) + \xi_{it} \end{aligned} \tag{8}$$

Viewed as a random variable, v_{it} has a variance

$$f_{ct} = \mathbf{z}'P_{ct|t-1}\mathbf{z} + \sigma^2 = \mathbf{z}'RP_{ct-1}R'\mathbf{z} + \mathbf{z}'SQS'\mathbf{z} + \sigma^2 \tag{9}$$

Given the normality assumptions, $v_{it} \sim \mathrm{N}(0, f_{ct})$. This result can be re-expressed in terms of q_{it}.

$$(q_{it}|Y_{it-1}, X_{it-1} = 0, \tilde{y}_{it}) \sim N(F(\mathbf{x}_{it}; \Gamma_t^F) + \mathbf{z}'\mathbf{u}_{it|t-1}, f_{ct}) \tag{10}$$

Finally, the prediction error is used to update beliefs about the current value of the productivity state variable

$$\mathbf{u}_{it} = \mathbf{u}_{it|t-1} + (P_{ct|t-1}\mathbf{z}/f_{ct})v_{it} \tag{11}$$

With new variance

$$P_{ct} = P_{ct|t-1} - P_{ct|t-1}\mathbf{z}\mathbf{z}'P_{ct|t-1}/f_{ct} \tag{12}$$

Even though true productivity, $\boldsymbol{\mu}_{it}$, is unobserved Equation 11 is a linear projection from the forecast error, v_{it}, onto the productivity. f_{ct} is the forecast error variance and $P_{ct|t-1}\mathbf{z}$ equals the covariance between the error and productivity.[23] Thus $\mathbf{u}_{it}$ is a new predicted value for $\boldsymbol{\mu}_{it}$ computed in the same fashion as a linear regression predicted value. Equation 11 can also be interpreted as apportioning the prediction error to each component of $\mathbf{u}_{it|t-1}$ and the noise term ξ_{it} in proportion to their variances. To see this, compare the $P_{ct|t-1}\mathbf{z}$ factor in the second term to the formula for f_{ct} in Equation 9.[24]

Notice that the variances $P_{ct|t-1}$, f_{ct}, and P_{ct} are unconditional in the sense that they do not depend on the data realizations. This justifies the lack of an i index. However, the entire Kalman filter remains conditional on the parameter values, including σ^2, Q, and W_c, that are to be estimated as discussed in the next section.

The key to estimation will be that capital, k_{it}, other inputs such as labor, $\boldsymbol{\ell}_{it}$, investment, i_{it}, innovation effort, $\mathbf{e}_{it}$, and exit decisions, X_{it-1}, are all uncorrelated with the forecast error, v_{it}. ξ_{it} is independent white noise, so this conclusion depends on absence of correlation between the choice variables and $(\mathbf{z}'\boldsymbol{\mu}_{it} - \mathbf{z}'\mathbf{u}_{it|t-1})$. From the Kalman filter $(\mathbf{z}'\boldsymbol{\mu}_{it} - \mathbf{z}'\mathbf{u}_{it|t-1})$ and $\mathbf{u}_{it|t-1}$ are uncorrelated (Lipster and Shiryaev [p. 74, 2001]). Likewise, any function

[23] Using Equation 8 and independence of v_{it} from $\mathbf{u}_{it|t-1}$ and ξ.

[24] And recall that $\mathbf{z}$ is just an aggregation vector.

of the information, Y_{it-1}, used in computing $\mathbf{u}_{it|t-1}$ is uncorrelated with $(\mathbf{z}'\boldsymbol{\mu}_{it} - \mathbf{z}'\mathbf{u}_{it|t-1})$, because $\mathbf{u}_{it|t-1}$ is already the MMSE of $\boldsymbol{\mu}_{it}$ conditional on Y_{it-1}. Whatever form choice variable decision rules take, they can only be functions of the available information, which consists of Y_{t-1} and Ω_{t-1}. Therefore they are uncorrelated with v_{it}.[25] No matter how inputs are chosen and exit decisions made, v_{it} will be distributed $\mathrm{N}(0, f_{ct})$.

The MMSE conclusion is a standard result for a Kalman filter applied to Gaussian SSFs. The SSF model here is only conditionally Gaussian because previous decisions, which are functions of previous beliefs and therefore stochastic, are part of the information set used to form the beliefs.[26] This is a remanifestation of the endogeneity and selection issues in slightly modified guise. However, the MMSE result for Kalman filters remains valid for such conditionally Gaussian processes (Harvey [1989, pp156-160]; Lipster and Shiryaev [2001, Chapter 13]). This is the econometric foundation of the sequential learning solution to strategic endogeneity and selection.

The assumption that establishments learn from the same production experience that the econometrician can observe is crucial to this line of reasoning. The expression E_{t-1} is shorthand for conditioning on a given information set. If the econometrician cannot perform the same conditioning that the establishment uses to learn, then the econometrician's prediction error $v^e_{it} = (\mathbf{z}'\boldsymbol{\mu}_{it} - \mathbf{z}'\mathbf{u}^e_{it|t-1}) + \xi_{it}$ may well not be independent of the establishment's prediction $\mathbf{u}_{it|t-1}$. Consequently, v^e_{it} need not be independent of establishment behavior in period t based on $\mathbf{u}_{it|t-1}$.

The top two panels of Figure 2 illustrate the productivity learning process using actual estimates for an establishment in the Chilean data used later in the paper. In the top panel the black line tracks the total production residual, $\mathbf{z}'\boldsymbol{\mu}_{it} + \xi_{it}$. The gray line tracks predicted productivity, $\mathbf{z}'\mathbf{u}_{it|t-1}$. In the first year, there is a large difference because predicted productivity starts at $\mathbf{z}'\mathbf{w}_c$ (which equals zero in this case because of identification restrictions). After that, predicted productivity tracks the total residual fairly well. The total residual moves around because of the productivity shocks and noise.

The middle panel focuses on updating of the productivity belief using the forecast error. The forecast error is the black line and equals the difference between total residual and predicted productivity from the first panel. The gray line shows the change in belief, $\mathbf{z}'\mathbf{u}_{it} - \mathbf{z}'\mathbf{u}_{it|t-1}$, due to updating. The change in belief is always in the same direction as the error. The absolute magnitude of the change in belief is always smaller than the error because some of the error each period is attributed to noise. Changes in belief are more responsive to the error in early years because the belief variance, $P_{ct|t-1}$, is larger then.

[25] Conditional on parameter values known to the establishment, the model implies that v_{it} is independent of Ω_{t-1} and Y_{jt-1} for $j \neq i$.

[26] For the same reason the Kalman filter is only conditionally linear when applied to the model here.

1.4 An Illustration with AR(1) Plus Establishment Effects

For simplicity, let the production function take a Cobb-Douglas form linear in logs.

$$q_{it} = a_t + \mathbf{x}'_{it}\boldsymbol{\beta} + \mu^1_{it} + \mu^2_i + \xi_{it} \qquad \xi_{it} \sim \mathrm{NIID}(0, \sigma^2) \tag{13}$$

The vector of inputs, $\mathbf{x}_{it}$, will contain logs of capital, labor, energy, and materials. Although a_t is an index of productivity, it is not the average productivity because the $\boldsymbol{\mu}_{it}$'s will not generally have zero mean. Instead, changes in a_t represent the portion of productivity growth that is common, both permanent and transitory.

Persistent idiosyncratic productivity, $\boldsymbol{\mu}_{it}$, is assumed to have two components. The first is a first order auto-regressive term and the second is an unchanging establishment effect.[27] Therefore, $R = \left(\begin{smallmatrix} R^{11} & 0 \\ 0 & 1 \end{smallmatrix}\right)$, η_{it} and Q are scalars, and $S' = (1, 0)$. Assuming the $G(.)$ functions are null, the resulting transition equation can be written in scalar form as

$$\mu^1_{it} = R^{11}\mu^1_{it-1} + \eta_{it} \tag{14a}$$

$$\mu^2_{it} = \mu^2_{it-1} \equiv \mu^2_i \tag{14b}$$

For entrants' data generating process, let $\mathbf{w}'_c = (w^1_c, w^2_c)$ and $W_c = \left(\begin{smallmatrix} W^{11}_c & 0 \\ 0 & W^{22}_c \end{smallmatrix}\right)$.

Appendix B provides scalar versions of the Kalman filter equations (5 – 12) that result from this illustrative specification.

Three identifying restrictions will be required when this example is estimated. These apply to mean entrant productivity for the first and last cohorts. The first restriction, $w^2_1 = 0$, is the type of normalization always required when establishment effects are combined with constants or a complete set of dummies. The second, $w^1_1 = 0$, arises from the block triangular cohort structure of the panel. This can be interpreted as filling in for the absence of incumbent establishments with a mean productivity to define w^1_1 as a deviation from. The third, $w^1_T = 0$, arises because no dynamics are observed in the last cohort to separate the components of $w^1_T + w^2_T$.

2 Estimation

2.1 Prediction Error Decomposition

Label the set of parameters $\Gamma \equiv \{\ \{(\Gamma^F_t, \Gamma^G_t)\forall t \in [1, T]\},\ \Gamma^c,\ R,\ S,\ \sigma^2,\ Q,\ \{(\mathbf{w}_t, W_t)\forall t \in [1, T]\}\}$. The elements of the parameter set beginning with R are sometimes referred to

[27]These have often been referred to as fixed effects. The focus here on establishment beliefs and forecast errors rather than structural shocks somewhat obscures the distinction between fixed and random effects. However, the spirit of this specification is more consistent with random effects and correlated regressors. Compare Hsaio's (1986) discussion of Mundlak (1978).

as hyperparameters because they describe the stochastic process. Conditional on the initial belief and the sequential learning process up to and including time $t-1$, the exact conditional log-likelihood for establishment i's output in period t is

$$\log[\Pr(q_{it}|Y_{it-1}, \tilde{y}_{it}; \Gamma)] = -(1/2)\log 2\pi - (1/2)\log f_{ct} - (1/2)(v_{it}^2/f_{ct}) \tag{15}$$

The likelihood of the establishment's entire data process Y_{iTi} can be written as

$$\begin{aligned} L(Y_{iT_i}; \Gamma, \Gamma^B) &= \Pr(y_{ic_i}) \prod_{t=c_i+1}^{T_i} \Pr(y_{it}|Y_{it-1}) \\ &= \left(\prod_{t=c_i+1}^{T_i} \Pr(X_{it-1}=0|Y_{it-1}, X_{it-2}=0)\right) \times \left(\Pr(\tilde{y}_{ic_i}) \prod_{t=c_i+1}^{T_i} \Pr(\tilde{y}_{it}|Y_{it-1}, X_{it-1}=0)\right) \\ &\quad \times \left(\Pr(q_{ic_i}|\tilde{y}_{ic_i}) \prod_{t=c_i+1}^{T_i} \Pr(q_{it}|Y_{it-1}, X_{it-1}=0, \tilde{y}_{it})\right) \end{aligned} \tag{16}$$

Where $c_i \equiv c|i \in I_c$ and Γ^B contains parameters governing input, investment, innovation effort, and exit decisions.

Equation 15 can be used to fill in the elements in the last part of this likelihood. A model of establishment decision rules for input, investment, innovation effort, and exit would be required to operationalize the remainder of the likelihood. However, the first two terms provide no information about Γ and reintroduce all the complications of characterizing the dynamic strategic equilibrium. So it is desirable to estimate Γ from only the third term. Cox (1975) introduced terms of this form and named them partial likelihoods[28]

$$PL(Y_{iT_i}; \Gamma) \equiv \left(\Pr(q_{ic_i}|\tilde{y}_{ic_i}) \prod_{t=c_i+1}^{T_i} \Pr(q_{it}|Y_{it-1}, X_{it-1}=0, \tilde{y}_{it})\right) \tag{17}$$

Furthermore, this partial likelihood is "dynamically complete" because each element is constructed from the density of q_{it} conditional on all previous realizations of q_{is} and $\tilde{y}_{is}$ $(s < t)$ (Wooldridge [2002, p.408]). Under standard regularity conditions, estimators derived from maximizing dynamically complete partial likelihoods have the same asymptotic properties as they would if the partial likelihood was a standard likelihood (Cox [1975], Wooldridge [2002]).

In the Kalman filter literature, which generally abstracts from regressors such as $\tilde{y}$, the (partial) likelihood constructed by substituting the conditional densities of the prediction errors from Equation 15 into 17 is known as a "prediction error decomposition" (Harvey [1989, p.125]). This pools the T_i conditional log-likelihoods for establishment i. Therefore

$$\log PL(\Gamma : Y_{iT_i}) = -\frac{1}{2}\sum_{t=c_i}^{T_i}\left[\log 2\pi + \log f_{ct} + \frac{v_{it}^2}{f_{ct}}\right] \tag{18}$$

[28]Notice that even though the individual elements of $PL(Y_{iT_i}; \Gamma)$ are conditional likelihoods, $PL(Y_{iT_i}; \Gamma)$ is not a conditional likelihood because the $\tilde{y}_{it}$ are functions of previous realizations of q, which affect the establishment's belief about its productivity.

Because parameters in the likelihoods are common across establishments in the panel data set, the log-likelihood for the entire data set can now be written as

$$\log PL(\Gamma : Y_t) = -\frac{1}{2} \sum_{i=1}^{I} \sum_{t=c_i}^{T_i} \left[\log 2\pi + \log f_{ct} + \frac{v_{it}^2}{f_{ct}} \right] \tag{19}$$

Where $Y_T \equiv \{Y_{1T_1}, Y_{2T_2}, \ldots, Y_{IT_I}\}$ refers to the entire panel data set. The v_{it}'s and f_{ct}'s are computed as functions of the parameter values from the Kalman filter, Equations 5 – 12, with initial beliefs, Equation 4. By pooling cross-sectionally, maximization of Equation 19 allows identification of entrants initial beliefs when these are associated with the data generating process that produces the actual $\boldsymbol{\mu}_{ic}$'s. This is unusual for Kalman filter processes, which are generally estimated in a strictly time series context. In that case the usual options are to estimate conditional on a maintained prior, assume a concentrated prior (which may in some cases be estimated), or assume a diffuse prior (de Jong [1988], Harvey [1989, pp. 121-122, 137-140]). The sequential learning estimator does not require an unconditional distribution from a stationary stochastic process to initialize the Kalman filter.

2.2 Analytic Gradient and Hessian

Exact analytic gradients of the log-likelihood function can be computed for Kalman filters (Harvey [1989, pp 140-143]). Letting γ_j label a parameter from the parameter set Γ, the formula is

$$\frac{\partial \log L(\Gamma : Y_T)}{\partial \gamma_j} = -\frac{1}{2} \sum_{i=1}^{I} \sum_{t=c_i}^{T_i} \left\{ \left[\left(\frac{\partial f_{ct}}{\partial \gamma_j} \right) \frac{1}{f_{ct}} \right] \left[1 - \frac{v_{it}^2}{f_{ct}} \right] + 2 \left(\frac{\partial v_{it}}{\partial \gamma_j} \right) \frac{v_{it}}{f_{ct}} \right\} \tag{20}$$

The partial derivatives $\partial f_{ct}/\partial \gamma_j$ and $\partial v_{it}/\partial \gamma_j$ are computed in the same iterative sequence as the Kalman filter itself by differentiating Equations 4–12.

There is also an analytic formula for the empirical information matrix that takes advantage of the Kalman filter's iterative conditioning. Terms multiplied by second (and cross-) derivatives of the prediction error and prediction error variance with respect to the parameters are zero asymptotically.

$$IM_{jk}^e = \sum_{i=1}^{I} \sum_{t=c_i}^{T_i} \left\{ \left(\frac{\partial f_{ct}}{\partial \gamma_j} \right) \left(\frac{\partial f_{ct}}{\partial \gamma_k} \right) \frac{1}{2f_{ct}^2} + \left(\frac{\partial v_{it}}{\partial \gamma_j} \right) \left(\frac{\partial v_{it}}{\partial \gamma_k} \right) \frac{1}{f_{ct}} \right\} \tag{21}$$

Where IM^e divided by the number of observations is the empirical realization of the information matrix with the asymptotic restrictions imposed.

The formula for IM^e is used in two ways. As the best empirical approximation of the information matrix, its inverse is used as the estimator of the asymptotic covariance matrix for the maximum likelihood parameter estimates, $V(\Gamma)$. As a computationally fast approximation to the exact (negative) Hessian it will also be used in place of the exact Hessian in the nonlinear search algorithm maximizing the log-likelihood.

2.3 Algorithm and Implementation

Any standard nonlinear optimization algorithm may be used to maximize the likelihood in Equation 19. A subroutine to compute the likelihood conditional on parameter values passed from the optimizer must be programmed. To do this, the subroutine must compute separate Kalman filters for each establishment based on the parameter values and the establishment's data. The process is much more efficient if the subroutine also computes analytic gradients and Hessians from Equations 20 and 21. These also require computing the derivative of each component of each establishment's filter with respect to each parameter. Smoothing requires one more pass of the subroutine, supplemented with the smoothing equations 22 and 23, after the parameters have been estimated. Appendix C describes the programming of this procedure in more detail.

Maximizing the likelihood from the AR(1) plus establishment effect specification in Appendix B poses two practical challenges. First, there is a plateau of undesirable local maxima when the AR coefficient $\widehat{R^{11}} \approx 1$. In this case for each cohort, c, mean initial values of the productivity components, w_c^1 and w_c^2, are not separately identified. They both contribute to the initial value of a single random walk process. Consequently, the search algorithm will be free to pick $\widehat{w_c^1}$'s with no relation to the true w_c^1. Once the search algorithm is at arbitrary $\widehat{w_c^1}$'s changing $\widehat{R^{11}}$ is apt to decrease likelihood at a first step, trapping the search algorithm at the local maximum. Fortunately, there are some diagnostics available to help avoid erroneous $\widehat{R^{11}} \approx 1$ estimates. The first is to estimate an AR(1) with noise specification without the establishment effect. Since this is a nested specification, if its likelihood is greater than the full model at $\widehat{R^{11}} \approx 1$ then that cannot be the global maximum. Furthermore, estimates from the nested AR(1) model can then be used as initial values in the full model in order to avoid returning to the local maxima plateau. In Monte Carlo, initializing the search at the true simulated parameters will work in a similar fashion.

Second, while maximizing Equation 19 guarantees that all forecast error variances, f_{ct}, are positive, the estimated variance parameters, $\widehat{\sigma^2}$, $\hat{Q}$, and $\widehat{W_c}$ for all c, could be negative unless actively constrained. The f_{ct}'s are functions of all of these parameters, as well as R, S, and $\mathbf{z}$.[29] Furthermore, the variance parameters play a second role in determining the allocation of the forecast error to noise and the components of $\mathbf{u}_{it}$ in forming updated beliefs. (Compare Equations 6, 9, and 11.) Thus, in finite samples or misspecified models the likelihood may increase if the updating allocation weight were decreased, or even negative, on noise (negative $\widehat{\sigma^2}$), on u^1 or u^2 in early periods (negative $\widehat{W_c^{11}}$ or $\widehat{W_c^{22}}$, respectively), or on u^2 in later periods (negative $\hat{Q}$). Finally, the usual downward bias of maximum likelihood variance estimators plays a role, especially for the $\widehat{W_c}$'s, which are identified from variation in only one cohort.[30]

For the variance of entrants' initial beliefs, the second problem is almost always solved by

[29] Recall that the $\widehat{W_c}$ are not set equal to the unconditional variance of the stochastic process described by Q, σ^2, R, and S.

[30] And primarily from early years in that cohort if the eigenvalues of R are less than one.

pooling the cohorts, i. e. imposing $W_c = W\ \forall c$. This increases the effective sample size. Some guides to likelihood maximization (e. g. Gould *et al* [2003, p.56]) recommend directly estimating a transformed version (e. g. square root or log) of variance parameters that can range freely over the real line and transform back, via the delta method, to a non-negative variance estimate. This solution is largely illusory. If the likelihood is decreasing in the variance parameter (generically γ^v) in the neighborhood of zero, the optimizer will pick a value for the transformed parameter at (or approaching) a point that will transform back to a zero γ^v. At such a point, any of the standard transformations introduce a singularity[31] that will produce an infinite variance and standard error. Using the delta method post-estimation cannot undo this. With true $\gamma^v > 0$ this will not be an issue asymptotically. In finite samples γ^v may be estimated at a binding constraint $\widehat{\gamma^v} = 0$. In this case the gradient component $\partial \log(L)/\partial\gamma^v \neq 0$ and regardless of any transformation the standard proofs of asymptotic normality and distributions of significance tests (in Cramer [1986] for example) do not hold. The constrained estimate $\widehat{\gamma^v} = 0$ may be reported without standard errors and the remaining parameter estimates interpreted as conditional on the constraint.[32] This occurs occasionally in Monte Carlo simulations and once in the Chilean results described below.

2.4 Decision Rules and Some Validity Tests

Using estimated beliefs computed from the sequential learning process, it is possible to examine whether and how establishment productivity beliefs influence dynamic strategic behaviors such as exit or investment. This amounts to second-stage estimation of behavior decision rules $X_{it-1} = X(\mathbf{s}_{it-1}, \nu_{t-1}, \Omega_{t-1}, \epsilon^X_{it-1})$ and $i_{it-1} = i(\mathbf{s}_{it-1}, \nu_{t-1}, \Omega_{t-1}, \epsilon^i_{it-1})$ implied by optimization of the establishments' continuation values as in Figure 1.[33] Stage game behaviors, such as choice of variable inputs $\boldsymbol{\ell}_{it} = \boldsymbol{\ell}(\mathbf{s}_{it-1}, \nu_{t-1}, \Omega_{t-1}, \epsilon^k_{it-1})$ can also be examined. The error terms, ϵ^X_{it-1}, ϵ^X_{it-1}, and ϵ^k_{it-1}, can be attributed to additional private information about the economic environment, random decision-making errors, or different beliefs about the dynamic equilibrium. With common knowledge of a dynamic equilibrium and common beliefs over industry state, these decision rules can be interpreted as best response policy functions. Theory says little about the specific form of these relationships other than that they may be nonlinear or even nonmonotonic in the case of investment or innovation effort.

The results section will describe empirical estimates of decision rule linear approximations for exit hazard, investment rate (investment over current capital), and employment rate (la-

[31] For example, the square root transformation requires multiplying the variance parameter's gradient term by $2\sqrt{\gamma^v}$ and its second derivative in the diagonal of the Hessian by $4\gamma^v$. Also notice that the transformed Hessian element goes to zero an order of magnitude faster than the transformed gradient when approaching the relevant point in parameter space.

[32] See Gill and King (2004) for some alternative Bayesian solutions to this issue.

[33] Recall that $\mathbf{s}_{it}$ is vector of establishment i's state variables including capital, k_{it}, and, in this case, productivity belief $\mathbf{u}_{it}$. ν_t is the industry state summarizing $\mathbf{s}_{jt}$ for all active establishments and Ω_t is the environment information set.

bor over capital). An establishment's own capital, productivity belief (estimated in the first stage), and other state variables, such as age, can be used directly as regressors. However, industry state and the environment information set must be summarized somehow. Pursuing this is an important avenue of research, but in this paper there would not be enough variation to produce meaningful results even if measures of industry state and economic environment were included. Since these are aggregate measures there would only be one set of values for each year in the data. In order to account for variation in industry state and economic environment year dummy variables will be included. Variation in the coefficients of these dummies will be an indicator that variation in industry structure and/or economic environment is affecting the establishment behavior represented in the decision rule.

Failure to find a significant relationship between behavior and estimated beliefs would not automatically invalidate the proposed sequential learning estimator. The difficulties could arise from small samples, incorrect functional forms, incorrect summary indices of industry state or environment, or from the added complication that the estimated beliefs are generated variables. However, repeated failure to find any relationship would raise questions about the motivations for the proposed estimator. If establishment beliefs really have little or no influence on behavior then there is little potential for selection or strategic endogeneity biases in the first place. If no empirical strategic relationships can be established, robustness with respect to industry structure has little value. Finally, the possibility that sequential learning is an inadequate model of belief formation would arise.

As mentioned in the introduction, the information and decision timing assumptions taken together imply that any decision by an establishment should be independent of productivity information revealed in the data after the moment the decision is made. If this is not true, either the establishment is using sources of information other than the reported production data to learn its productivity more quickly, or it is modifying its decisions after some of its production experience for the year has been revealed. In either case, the production inputs are no longer predetermined and justification of the sequential learning estimator would break down. Therefore, it is important to construct validity tests using this joint implication.

The Kalman filter provides a way to compute the best estimate of an establishment's productivity using data from its entire history through T_i, including years after the decision is made. These are known as smoothed estimates and will be labelled $\mathbf{u}_{it|T_i}$. Smoothed estimates contain more information about productivity than the forward-looking predicted productivities, $\mathbf{u}_{it|t-1}$, computed from only the establishment's previous history. The difference $\mathbf{u}_{it|T_i} - \mathbf{u}_{it|t-1}$ summarizes information about current productivity revealed for the first time by production experience in the current and future periods. Smoothing algorithms are a standard augmentation of Kalman filters. When the underlying stochastic process is Gaussian the resulting estimates are MMSE for the information in all the years used.

The algorithm to smooth for the entire period from c_i to T_i is known as fixed-interval smoothing. After the Kalman algorithm (Equations 5–12) has been run through T_i for each establishment, the smoothed estimates, $\mathbf{u}_{it|T_i}$, are calculated with a backwards recursion from

T_i to c_i (Harvey, 1989, pp.154-155). Starting with $\mathbf{u}_{iT_i|T_i} = \mathbf{u}_{iT_i}$ and $P_{iT_i|T_i} = P_{cT_i}$ for $i \in I_c$

$$\mathbf{u}_{it|T_i} = \mathbf{u}_{it} + P_{ct}R'(P_{ct+1|t})^{-1}(\mathbf{u}_{it+1|T_i} - R\mathbf{u}_{it} - G(\mathbf{e}_{it}, i_{it}, \mathbf{x}_{it}; \Gamma_t^G; \Gamma^G)) \quad (22)$$

$$P_{it|T_i} = P_{ct} + P_{ct}R'(P_{ct+1|t})^{-1}(P_{it+1|T_i} - P_{ct+1|t})(P'_{st+1|t})^{-1}RP'_{st} \quad (23)$$

Notice that unlike other variances, $P_{it|T_i}$ has an i subscript. Although it is not conditional on the data, it is conditional on time of exit. This makes sense — fewer years of data to form the estimate $\mathbf{u}_{it|T_i}$ implies a larger MSE.

Appendix B has scalar versions of Equations 22 and 23 for the AR(1) plus establishment effect empirical example used later in the paper.

The third panel of Figure 2 compares updated productivity beliefs using production information through the end of the period with the smoothed estimate using the complete data history. The dark updated belief and gray smoothed estimate lines match at the last year. During the rest of the interval covered by the graph, updated beliefs and smoothed estimates follow similar courses. But the smoothed estimates are just that — smoother — because they place equal weight on information from previous and future periods.

The validity tests can now be expressed in terms of a null hypothesis that $\mathbf{u}_{it|T_i} - \mathbf{u}_{it|t-1}$ will have a zero coefficient if inserted into any of the empirical decision rules described above.[34] When this null is not rejected in any of the modified decision rules the sequential learning estimator passes the validity test for that data set.

3 Chilean Manufacturing Panel Data

Many papers on the Chilean economy in general, and Chilean productivity trends in particular, have been published in recent years. There are two primary reasons for this. First, beginning in the 1970s Chile has experienced extensive structural, trade, and macroeconomic reforms.[35] Second, Chile's national statistical institute, *Instituto Nacional de Estadisticas* (INE), has been conducting a high-quality annual manufacturing census, *Encuesta Nacional Industrial Anual* (ENIA), of all establishments with at least 10 employees since 1979.

This Chilean manufacturing census data was compiled into a panel and documented for English-language users as part of a World Bank multi-country study of productivity, development, and trade (see especially the volume edited by Roberts and Tybout [1996] and documentation in Liu [1991]). Since that time, many economists have used this panel and

[34] In practice exit cannot be used because exit happens at the beginning of $T_i + 1$ and $\mathbf{u}_{iT_i+1|T_i} = R\mathbf{u}_{iT_i} + G(\mathbf{e}_{it}, i_{it}, \mathbf{x}_{it}; \Gamma^G) = R\mathbf{u}_{iT_i|T_i} + G(\mathbf{e}_{it}, i_{it}, \mathbf{x}_{it}; \Gamma_t^G)$

[35] See discussions in Tybout *et al* (1991); Liu (1993); Tybout (1996); Levinsohn (1999); Pavcnik (2002); Bergoeing *et al* (2002); Hsieh and Parker (2002); Bergoeing, Hernando, and Repetto (2003); and Kandilov (2005).

its updates in their studies.[36] The version of the Chilean panel used here covers the years 1979 through 1996. The census questionnaire includes employment by type; labor compensation; consumption of electricity and fuels; costs of raw materials, intermediate inputs, and purchased services; value of sales, intra-corporate shipments, and other sources of income; inventories; capital investment and depreciation charges by type; and capital stocks for some years. The only notable gap is the absence of any measure of innovation effort, such as R&D expenditure. Appendix D has detailed documentation of sources and preparation of the analysis variables.

Establishment entry and exit are deduced from the panel structure. Here the sequential learning estimator is applied to subsets of the panel comprising cohorts with complete histories. That is, cohorts starting in 1980 or later. This avoids the difficulties created by having to estimate initial beliefs of incumbents in the first year of data. The 10 employee threshold raises a concern that entry and exit will sometimes be inferred erroneously.[37] False entry is a particular problem for the sequential learning algorithm because the establishment's learning process does not actually start in the year the model assumes. Therefore, selecting the most useful sectors for demonstration involves a trade-off between full-history cohort sample sizes on the one hand, and the risk of false entry on the other. An observable indicator of the risk of false entry is the frequency of apparent entrant employment near the 10 employee threshold.

Three large sectors — food products, apparel, and fabricated metals products — are selected because of their size and diversity. However, establishments with initial employment less than 15 are excluded to reduce the risk of false entry.[38] In the fourth industry, pulp and paper products, small entrants with employment between 10 and 15 are not excluded. Paper products has comparatively few small entrants. Among industries with a 25th percentile of entrant employment of at least 20, paper has far more full-history cohort observations (373 before scrubbing for reporting gaps and missing data) than any other. Fabricated metals will be featured in the discussion of results below.

Table 1 provides establishment and observation counts for each of the four industries. The top panel describes all observations in the ENIA source data and the bottom panel lists the analysis data of full-history cohorts. The analysis data also excludes observations with missing data, history gaps, and other assorted data problems. On all four dimensions (establishments or observations by source or analysis data) the sample size is largest for food products, followed by fabricated metals, apparel, and paper products. For food, apparel, and fabricated metals there are 17 usable cohorts. The final column of the second panel displays the average number of entrants in a cohort of the analysis data. Cohort size is important because it provides the variation for identification of initial productivity beliefs. The paper industry is restricted to 13 cohorts because the 1981 and 1982 cohorts are empty, the 1980

[36]See the introduction of Appendix D for a list of some of the papers and a genealogy of the version of the panel used here.

[37]Appendix D discusses this problem at greater length.

[38]This is legitimate with the sequential learning estimator, since initial employment is predetermined from the perspective of any period that the establishment would have been in the data.

cohort is short-lived, and the 1985 cohort is a short-lived singleton. Paper also has a much smaller average cohort size.

The right-hand portion of the first panel in Table 1 provides information on establishment turnover. Averaged over 17 years, annual entry rates (5.4% to 7.2%) and exit rates (5.0% to 8.5%) are moderate. Other than paper's low exit rate, industry rankings are the same by entry and exit rates. Table 2 provides year by year detail on turnover in fabricated metal products. The left side covers all observations in the source data and the right side includes analysis data only. Starting from the total number of establishments in the previous year, subtracting establishments exiting and switching sectors gives the number of observed incumbents. Adding entrants and inward sector switches[39] gives the total observed establishments in the current year. Both exit and entry vary considerably by year. In the analysis data, the number of exits tends to rise because the number of establishments observed (therefore at risk) increases over time.

Table 3 provides some detail on the cohorts in fabricated metals. The number of plants, establishments that survive through 1996, and number of observations are listed for the entire data set and the analysis data. The "0" cohort is an aggregate of all establishments present in 1979, whose true cohort cannot be identified. Four of the cohorts (1981, 1982, 1985, and 1990) are quite small in the analysis data. Estimates of the distribution of initial productivities in these cohorts especially benefit from pooling of the initial variance, $W_c = W \ \forall c$.

The analysis uses six variables: gross output, labor, capital services, energy inputs, material inputs, and investment. Table 4A contains descriptive statistics for all of these variables for fabricated metals. The variables are reported in levels and natural logarithms for both the analysis data set and for entrants only. All variables except labor are measured in 1985 Chilean pesos. Labor is a compensation-weighted aggregate of white- and blue-collar employees measured in blue-collar person-year equivalents. Both employment and capital services are converted to annual equivalents by adjusting for days of operation. The capital services variable is computed from three types of capital stock (buildings, machinery, and vehicles). These capital stocks are perpetual inventories constructed from annual investment and initial capital inferred from first-year depreciation (see Appendix D for details). Each type of capital stock is multiplied by a type-specific implicit rental rate. Summing the resulting implicit capital rents along with actual capital rents paid gives the capital services aggregate. Energy includes purchased electricity and a variety of fuels. Materials include both raw materials and intermediate inputs.

Gross output is a more natural concept than value added for production function and productivity studies at the establishment and industry levels. Demand and strategic interactions are in terms of output and output prices. Therefore, energy and materials efficiency can be just as important as labor or capital efficiency. Using output also avoids various measurement problems specific to value added (see discussion in Appendix D).[40]

[39]Inward sector switches and establishments returning from a data history gap are always excluded from the analysis data because their complete learning process cannot be modelled.

[40]The output variable used here is deflated nominal value of production, including changes in inventory and

Returning to Table 4A, notice that entrants are smaller in every dimension (both means and medians). Also the entrants' variances are less than for all observations in the analysis data set. This is a common finding and is a prediction of some dynamic industry model specifications.[41] The positive skew of the variables measured in levels is also typical. There are many small to medium-sized establishments and just a handful of large establishments. This also explains means substantially greater than medians. However, after taking logarithms the analysis variables are distributed almost symmetrically. Table 4B provides a more concise summary of the analysis variables for food products, apparel, and paper products. Establishments in the paper industry are substantially larger than establishments in the other three industries. Apparel and fabricated metals establishments are the smallest.

Table 5 has aggregate revenue and cost shares of expenditures on capital services, labor, energy, materials, and miscellaneous services for each industry. Paper is by far the most capital intensive and apparel is the most labor intensive. Energy is more important in food and paper than in apparel and fabricated metals. Raw materials and intermediate inputs account for over half of expenditures in all four industries. The total expenditure shares of revenue should be viewed with some caution. Close examination of the data reveals a tendency to under-report expenditures on energy, miscellaneous services, and rental payments.[42] The intra-firm shipments component of materials is subject to the usual valuation issues. Finally the capital services aggregate is constructed with a cost of capital based on commercial bank lending rates averaged over the period.[43] This may well overstate cost of capital in establishments with access to international capital markets or internally generated investment funds.[44] On balance, it is plausible that total expenses are overstated in the paper industry, which has the largest, most capital-intensive establishments, and under-stated in the other three industries.

Table 6 has simple correlations among the analysis variables in logarithms. The correlations are rather low for this type of data in fabricated metals, quite high in paper, and intermediate in food and apparel. In all four industries the highest correlation is between output and materials. The correlations among inputs imply enough variation in input ratios so that separate identification of their output elasticities should not be a problem.

intracorporate shipments, rather than an index of physical production. Thus this variable actually includes some residual cross-sectional variation in price, raising some complex issues regarding quality adjustment, specification of demand systems, and revenue function interpretations of the empirical results below. These will not be pursued in this paper.

[41]But see Pakes and Ericson (1998) for a counter-example.

[42]In particular, these items will sometimes be reported as zero (i. e. potentially missing) for observations that are preceded and followed by non-zero entries and with no other indication of large discontinuities.

[43]The imputed rental rate on capital stocks has three components: the opportunity cost of capital, depreciation, and expected capital gains. See Appendix D for details.

[44]Nevertheless, it is important to use the same implicit rental rates for all establishments. See discussion in Appendix D.

4 Results

The results in this section are primarily intended as a proof of concept. The following subsections show that the sequential learning estimator works, both as an estimator of production functions and productivity dynamics and as a first stage in multistage empirical modelling of industry strategic dynamics with endogenous productivity. The subsections are organized by object of estimation. Production function parameters and productivity dynamics are estimated with the sequential learning estimator itself in the first two subsections. The third subsection discusses second stage estimation of decision rules for establishment behaviors such as exit, investment, and employment. Estimates of productivity beliefs, generated by the sequential learning estimator in the first stage, are used as regressors. The third subsection also presents the validity tests on the information and decision timing assumptions. These tests use modifications of the decision rule specifications. The fourth subsection shows that an investment behavior characterization related to the invertibility condition used in Olley and Pakes' (1996) estimator is frequently incompatible with productivity belief estimates from the sequential learning estimator.

Specific results support several themes that will span these subsections. Evidence of the quality and reliability of sequential learning estimates takes four general forms. First, estimates of input elasticities, returns to scale, and productivity dynamics are plausible but differ in nontrivial ways from results obtained using alternative estimators. Second, the standard errors of production function parameter estimates are smaller than for previous estimators that address endogeneity. Third, in second-stage decision rule estimates, productivity beliefs almost always take the expected sign and are often statistically significant. Finally, there is indirect evidence that establishments are in fact uncertain about their own productivities and do update their beliefs using production experience.

Three additional themes are noteworthy. First, estimated productivity and decisions rules confirm that selection and strategic endogeneity do occur in these four industries. Second, there is evidence that certain alternative estimators would have been misspecified in these industries. Most importantly, there are a variety of preliminary results that contribute to the agenda of studying productivity dynamics and behavioral sources of technological change. These include evidence of persistent productivity shocks, estimates of the impact of productivity beliefs on strategic behavior, and preliminary indications that industry structure and/or economic environment have estimable effects on establishment behaviors that change aggregate industry productivity.

4.1 Production Function Estimates

Sequential learning estimates of Cobb-Douglas production function parameters and productivity dynamics parameters for fabricated metals are displayed in Table 7. The dependent variable is the natural logarithm of output. The table lists four nested specifications of the dynamics of establishments' idiosyncratic productivity. The left column, labelled DP for double process, is the primary specification with an AR(1) process, establishment effects, and noise. The second column, labelled SP for single process, keeps the AR(1) and noise specifications but drops the establishment effects. RW indicates a random walk process and can be considered a special case of SP in which the idiosyncratic productivity shocks are permanent. In other words, the AR(1) coefficient is constrained $R11 = 1$. Imposing the additional constraint that idiosyncratic productivity shocks are zero, $Q = 0$, makes the single process an establishment effect (EE) drawn at entry.

Parameter estimates are divided into blocks. The first block has production function elasticities. The second block has returns to scale, which is an ancillary parameter computed as the sum of elasticities, and the AR(1) coefficient. The variance parameters describe the variability of entrant's initial idiosyncratic productivities ($W11$ for the AR(1) component and $W22$ for the establishment effect component), the size of establishment-level productivity shocks (Q), and the noise (sigma2). The remainder of the first page of the table has coefficients on year dummies, which can be interpreted as the common component of productivity. The second page has estimates of mean initial idiosyncratic productivity ($w1$ for the AR(1) component and $w2$ for the establishment effect component) by cohort. By the rational expectations assumption these can also be interpreted as the entrants' expected idiosyncratic productivity in their first year. Asymptotic standard errors are in parentheses. Various significance tests are discussed below. Table 8 displays results of the double process specification for the other three industries using the same format.

Focus first on the production function elasticity estimates. Materials has the largest elasticity in every estimate reported on Tables 7 and 8. These are greater than 0.5 in every case except apparel, where the estimated materials elasticity is nearly that large. At the other end, energy elasticities are always the smallest. Estimated labor elasticities are always greater than capital elasticities, even in the paper industry where capital has a larger expenditure share than labor.[45]

These elasticity results are in rough conformity with the revenue and cost shares reported in Table 5 above. If the establishments in an industry are static cost minimizers with no quasi-fixed inputs and no strategic dynamics, then the ratio of an elasticity over the returns to scale parameter would equal that input's cost share. This equality is statistically rejected at 5% except for materials and energy in paper and capital in fabricated metals and food products. Nevertheless, the estimated ratio of elasticity over returns to scale is within 10% of the cost share in 23 of 28 cases.[46]

Next consider returns to scale. Point estimates indicate slightly decreasing returns in fabricated metals (all specifications) and food products. Apparel and paper show returns to

[45] Low capital elasticity estimates are a common result in microeconomic production function studies (Griliches and Mairesse [1995]). In preliminary fixed effects estimates for the paper industry, the capital services aggregate used here did produce greater capital elasticity estimates than capital stock aggregates.

[46] Exceptions are labor in the RW specification for fabricated metals, labor and materials in apparel, and capital and labor in paper.

scale that are just barely increasing. Table 9 displays the results of Wald tests on the null hypothesis of constant returns to scale (i. e. equal to 1) for all four industries and all four specifications. For food, apparel, and paper the specification does not matter — in food products constant returns is rejected at high levels of significance, but in apparel and paper products it is accepted. In fabricated metals the result is less clear-cut, but under the most general, double process, specification constant returns is rejected at the 5% level.

If the establishments in an industry are static profit maximizers with no quasi-fixed inputs and no strategic dynamics, returns to scale will equal the mark-up (price over marginal cost) times the revenue share of total input expenditures as in Table 5. Assuming static profit maximization and taking the results in Tables 5, 7, and 8 at face value, the implied mark-ups are 28.4% for fabricated metals, 13.3% for food products, 18.4% for apparel, and 1.4% for paper products. These are likely to be modestly understated in paper and overstated for the other industries.[47] Nevertheless, the relative magnitudes are plausible. Paper products are a major Chilean export and face highly elastic international demand. Fabricated metals products are highly varied and thus almost certainly face more strongly differentiated demand than the other industries. Notice that increasing returns, low mark-ups, and high demand elasticities in the paper industry are consistent with the large size of its establishments.

Accurate estimates of returns to scale can be extremely important in analyses of the productivity implications of policies regarding competition, trade, and so forth. In a static snapshot of establishments' contribution to industry or macroeconomic productivity, misestimated returns to scale will be compensated by shifts in attributed establishment productivities.[48] Overestimated productivity of large establishments would accompany underestimated returns to scale. But policy analysis involves counterfactuals. Are large establishments productive (or not) because they are large, or large because they are productive? When a group of small establishments appear to be unproductive, would they remain so if encouraged to grow? If a policy change appears to improve productivity, is it because large incumbents, which erroneously appear to be more productive, are favored? Or because establishment-level productivity actually improves in response to the new policy?

One of the reasons many null hypotheses regarding production function parameters are rejected is the precision of the estimates. Asymptotic standard errors of the elasticity estimates are 2.0% or less in fabricated metals and food, and only slightly larger for apparel and paper, which have smaller sample sizes. Comparison with the alternative fabricated metals production function estimates listed in Table 10 is instructive. The first panel lists results from some popular, but inconsistent, estimators. The second panel has results from a selection of estimators that are consistent in T or N. The sequential learning standard errors are quite similar to standard errors of the random effects estimator. In contrast, standard errors from the various consistent estimators are larger, sometimes considerably. Also notice that compared to the sequential learning estimate of returns to scale, the OLS estimate is

[47] See the discussion of data quality with respect to Table 5.

[48] See Basu and Fernald [2002] for a thorough exposition of this type of aggregate productivity decomposition.

8% larger, most of the consistent estimates are considerably smaller, but the random effects (and Levinsohn Petrin energy proxy) estimate is within 2%.

Similarity between the sequential learning and random effects estimates is not coincidental. The consistent estimators solve the problem of endogeneity by removing variation, especially between establishments, that could be correlated with the input variables. For example, the fixed effects estimator is also known as the within estimator because identification comes entirely from time series variation within establishments. As Griliches and Mairesse (1995) argue, by removing so much of the variation reflecting production relations, these estimators become more vulnerable to other types of undesirable variation such as measurement error. This usually results in implausibly low estimates of capital elasticity and returns to scale. In contrast, the random effects estimator retains both within- and between-establishment variation at the price of susceptibility to endogeneity bias. The sequential learning estimator's cohort structure allows it to use a large portion of the between-establishment variation. Among entrants in a cohort all between-establishment variation is part of the forecast error. Since establishments learn about their idiosyncratic productivities over time, some portion of between-establishment, within-cohort variation will also appear in forecast errors of subsequent periods. However, unlike random effects, sequential learning explicitly controls the forecastable portion of productivity variation that can cause endogeneity bias.

4.2 Productivity Dynamics

This paper's opening paragraph posed questions about the sources of technological change and incentives for establishment behaviors that create productivity growth. This subsection will use the sequential learning estimator results in Tables 7 and 8 to examine the dynamics of idiosyncratic and common productivity. The following subsections will estimate decision rules for behaviors that affect aggregate industry productivity (e. g. exit and investment).[49]

First consider the form of idiosyncratic productivity's dynamics. Since the four alternative specifications of productivity dynamics (DP, SP, RW, and EE) are nested, they can be compared using likelihood ratio tests with the more restrictive specification as the null hypothesis. Table 11 shows results for all four industries.[50] The establishment effect only (EE) specification is always strongly rejected. Incumbent establishments in these four industries definitely experience some kind of dynamics in their idiosyncratic productivities. The random walk (RW) specification is also rejected, usually strongly as well. In one instance, versus DP in apparel, the significance level is only 9.4%. This is unimportant since RW is strongly rejected in a test against the more restrictive SP alternative. Comparison between

[49] If data on innovation effort had been available, the estimates in this subsection could have included the effects of innovation effort on idiosyncratic productivity without any methodological changes. Likewise, the following subsections could have estimated establishment decision rules for the level of innovation effort.

[50] In the paper industry tests involving the double process specification are not applicable because the productivity shock variance parameter estimate ($\hat{Q}$) is at a binding non-negativity constraint with non-zero gradient, invalidating the test statistic.

the double process (DP) with establishment effects and the single process (SP) without establishment effects depends upon the industry. In fabricated metals SP is rejected against the DP alternative, but in food products and apparel SP is accepted. In food products and apparel there may be no permanent establishment effects. Finally, the distinction among the four specifications makes very little difference to the production function parameters, as can be seen be comparing columns in the first block of Table 7.

Returning to Tables 7 and 8, estimates of the productivity shock variance, $\hat{Q}$, and the AR(1) coefficient, $\widehat{R11}$, for fabricated metals, food, and apparel also confirm that productivity dynamics include shocks that are persistent, but not permanent. The constrained estimate of $\hat{Q} = 0$ in the DP specification of the paper industry is not an indication of the absence of productivity shocks, since the EE specification was rejected in contrast to SP or RW. Instead it is symptomatic of difficulty separating the establishment effect component from the AR(1) component, most probably because AR(1) is not the ideal specification of productivity dynamics in the paper industry.

A number of comparisons can be made among variance components in an industry. In fabricated metals temporary differences account for less variability in entrants' persistent productivity than permanent differences do ($\widehat{W11} = 0.035$ versus $\widehat{W22} = 0.049$). Similarly, the persistent component of productivity shocks is somewhat less than half of the total shock ($\hat{Q} = 0.015$ versus sigma2 $= 0.022$). Yearly shocks (Q+sigma2) have one third the variability that occurs in entrants' productivity ($W11 + W22$+sigma2). Similar comparisons can be made for the other industries using Table 8. For example, in food products temporary differences account for more variability than permanent differences in entrants' persistent productivity. In apparel transitory noise is smaller than persistent productivity shocks.

Comparisons of productivity dynamics parameters may also be made between industries. Productivity shocks are more persistent in fabricated metals and food products (AR(1) coefficients of 0.726 and 0.737, respectively) than in apparel and paper products (0.504 and 0.556). On the other hand, apparel has the largest persistent productivity shocks. Entrants' persistent productivity is more variable ($W11 + W22$) in food products and apparel than in fabricated metals and paper.

Taken together, these results constitute strong evidence for persistent establishment productivity shocks. Furthermore, these shocks to idiosyncratic productivity are not permanent. These have several important implications. First, fixed and random effects estimators would be misspecified in these industries. Second, since they are unpredictable, the presence of persistent productivity shocks and noise supports the contention that establishments have uncertainty about their productivities. Third, evidence of these shocks contributes to the project of characterizing establishment-level productivity dynamics. Fourth, there are differences among industries in the size of productivity shocks, their persistence, and the form of establishment productivity dynamics. Explaining the causes of these differences is an important topic for future exploration. One candidate hypothesis is that various sources of idiosyncratic productivity — management, human capital, labor relations, equipment quality — exhibit different degrees of variability and persistence and are more or less important in different industries. For example, absence of permanent productivity advantages in the food and apparel industries industries may due to easy imitation and less importance of human capital. Finally, random initial productivity and subsequent idiosyncratic shocks are the sources of heterogeneity in establishment productivity. This is a fundamental cause of other types of heterogeneity (e. g. size or profitability) and therefore of industry structure.

Regarding the last point, previous research has shown that productivity varies by as much as a factor of 2 or 3 across establishments within various industries.[51] The productivity distributions estimated here exhibit a slightly smaller, but similar range. Table 12 reports the distribution of idiosyncratic persistent establishment productivities (the μ's) using the estimates from all available data (i. e. smoothed in Kalman terminology). Because common productivity trends are captured separately in year dummies, these distributions represent cross-sectional variation pooled over years in the panel. Recall the productivity estimates are additive in logarithms. So a 0.01 difference represents approximately a one percent change in productivity. The non-zero means reflect selection effects due to less productive establishments' tendency to exit. The table presents two measures of productivity variability, standard deviations and differences between tenth and ninetieth percentiles. Expressed as a ratio, establishments at the ninetieth productivity percentile are roughly twice as productive as establishments at the tenth percentile. Comparing industries, the larger mean in apparel suggests more rigorous selection. This is consistent with its higher average exit rate reported in Table 1A. Nevertheless, productivity variability is quite similar across the four industries.

Remaining coefficients from the sequential learning estimates provide information on common and cohort-level productivity trends. Beginning with the year dummies, one might ask whether they provide any evidence of common productivity growth.[52] Table 13 shows the results of Wald tests on three common-productivity null hypotheses for each industry. The first null hypothesis, labelled H0a, is that all year dummies equal each other. This is strongly rejected in all four industries, indicating that there is some year-to-year variation in common productivity. The next two hypotheses assess whether there is a trend in the year dummies that could be interpreted as productivity growth. H0b compares the last year dummy (1996) to the first (1980, except for paper which is 1983). Equality is only rejected for apparel. Returning to Table 8, this indicates a reduction of common productivity in apparel and no significant change in other industries. Coefficients on first-year dummy variables may be confounded by the identifying restrictions on the first cohort's initial productivity. Therefore, H0c uses fifth year common productivity as the basis of comparison to the final year. With this comparison food and fabricated metals show modest, but not statistically significant, productivity increases. In apparel the productivity reduction is still present, but no longer significant. Apparel's decreasing common productivity is an anomaly. Changes in labor quality do not appear to be the explanation, because growth in real compensation per employee over the period is almost exactly the same as for total manufacturing. Another

[51] Bartlesman and Doms [2000] review the findings of previous establishment panel productivity studies.

[52] Recall that the year dummies capture the elements of productivity change common to all establishments in an industry. Also note that output and input data were deflated before estimation. Therefore, the coefficient estimates do not conflate average price-level changes with productivity.

candidate explanation is mismeasurement of the sector's output or input deflators.

The last two panels of Tables 7 and 8 report components of entrants' mean idiosyncratic productivity by cohort. Recall that $w1$ describes the component of productivity subject to AR(1) decay and $w2$ describes unchanging establishment effects.[53] For the first cohort $w1$ and $w2$ are set to zero as identifying restrictions. In the DP specification w1 is also set to zero for the last cohort. Table 14 reports two groups of hypothesis tests regarding these coefficients. H0a through H0c are comparisons of cohorts' initial productivities and the last two tests concern learning from experience. H0a tests whether mean initial idiosyncratic productivities are equal across all cohorts. This is strongly rejected in the food, apparel, and paper industries but not even weakly in fabricated metals. H0b tests for a significant difference between the second and final cohorts. The only significant result indicates better productivity in the last cohort in the apparel industry. H0c, comparing the fifth cohort with the last, produces quite different results. The test is highly significant (last cohort more productive) in paper, marginally significant (last cohort less productive) in food products, and now insignificant in apparel as well as fabricated metals. In light of the coefficient estimates in Tables 7 and 8, the correct conclusion for all industries appears to be no distinct trend in entrants' initial productivities.

Previous studies (see Bartlesman and Doms [2000]) have found evidence that entrants have systematically lower productivities than incumbents. In the four Chilean industries studied here, this is due to the effect of selection on incumbents and not to learning from experience in entrants' early years. The hypotheses in the bottom half of Table 14 test whether the decaying component ($w1$) of entrants' initial productivities tends to be negative. H0d tests if all all estimated w1 coefficients could equal zero. H0e is a similar test beginning with the fifth cohort. These hypotheses are only rejected in the paper industry. Examination of tests on individual coefficients shows that the significance is coming almost entirely from a single year (1988).

4.3 Decision Rules and Validity Tests

Decision rules can be estimated and validity tests performed with predicted productivity beliefs and smoothed productivity estimates derived from the sequential learning estimates reported in the previous two subsections. The types of behavior modelled here are exit (as a hazard rate), investment rates (in proportion to the current capital stock), and employment rates (in proportion to the current capital services aggregate). The estimated decision rules are exploratory in nature and specified in linear form. The objective is to establish whether these behaviors are related to estimated productivity beliefs, not to determine the correct functional form. Under appropriate conditions, primarily existence of a stationary industry equilibrium generating the observed establishment behavior, these decision rules can be interpreted as linear approximations of dynamic strategic policy functions.

[53]In the EE specification on Table 7, parameter restrictions convert w1 into an establishment effect.

Table 15A provides the basic decision rule estimates for establishments in the fabricated metals industry.[54] The variable *pred* contains productivity predictions at the beginning of each period (i. e. $\mathbf{z}'\mathbf{u}_{it|t-1}$) from each establishment's learning filter. Each 0.01 change in *pred* represents approximately a one percent change in productivity. $pred(t+1) - pred(t)$ measures the change in productivity belief going into the next period compared to the belief at the start of the current period. These changes in belief are due to productivity dynamics and information about idiosyncratic productivity revealed by production experience in the current period. Age measured in years is inferred from the panel. The capital regressor is the same capital services aggregate used as a production input above. Here it is expressed in millions of 1985 pesos, which is roughly equal to $7,820 in U.S. dollars at 2004 prices (see the introduction of Appendix D).

Exit Hazards

After excluding sector-switching and data gaps, the average exit rate for the fabricated metals sample is 5.8%, the same as for all establishments in the industry (compare Tables 1A and 2). The exit decision is estimated as an exponential hazard rate equation with a linear index function. Data from the last period of production is the only way to observe characteristics of exiting establishments. Therefore, exit is treated as a forward-looking decision made using all information available at the end of the current (observable) period. If the last observable period is t, the last mean productivity belief before exit is $pred(t+1)$. This is decomposed into $pred(t+1) - pred(t)$ and $pred(t)$ in order to assess the relative importance of productivity belief shocks versus levels in exit decisions.

For interpretation, the reported exit coefficients are transformed into relative hazard rates and affect the exit hazard multiplicatively.[55] A regressor with no effect on exit would have a relative hazard coefficient of 1 (rather than 0). For example, an increase in $pred(t)$ reduces the exit hazard because its relative hazard rate coefficient is less than 1. A one unit increase in $pred(t)$ (that is, an increase in productivity belief by a factor of $e \approx 2.718$) decreases the exit hazard by 82.8% (1-0.172) on average. A one percent increase of the productivity belief in $pred(t)$ reduces the exit hazard by 1.7% and a one percent increase represented in $pred(t+1) - pred(t)$ reduces the hazard by 2.0%. These are both statistically significant as verified by the tests on H0a and H0b reported in the bottom panel of Table 15A. In fabricated metals, the null that the coefficients on $pred(t)$ and $pred(t+1) - pred(t)$ are equal (H0c) cannot be rejected. So it appears that it is a low level of productivity belief, rather

[54] The standard errors in Tables 15A, 15B, and 16 are not corrected for the generated regressors $pred(t)$ and $smooth(t)$. Nevertheless, the hypothesis tests on coefficients of expressions involving $pred(t)$ and $smooth(t)$ are the correct size. Likewise the asymptotic standard errors do not require correction when the null hypotheses $\beta_{pred(t)} = 0$, $\beta_{pred(t+1)-pred(t)} = 0$, and $\beta_{smooth(t+1)-pred(t+1)} = 0$ hold (Newey and McFadden [Theorem 6.2, 1994]). Furthermore, because $pred(t)$, $pred(t+1) - pred(t)$, and $smooth(t+1) - pred(t+1)$ are orthogonal, a hypothesis test on the coefficient of any one is the correct size even if the true value of the other coefficients is not 0 (Newey and McFadden [p. 2181, 1994])

[55] Standard errors are estimated for the linear indicator function and transformed into the relative hazard metric using the delta method.

than adverse shocks in the last year per se, that induces exit.

Continuing with the exit decision rule, even controlling for productivity beliefs and size as measured by capital, age also reduces exit hazard. Each year an establishment ages reduces the hazard by 10.8%, which is a statistically significant result. This is a typical finding among panels of young establishments. One interesting aspect is that the result persists after conditioning on learned productivity beliefs. This suggests there is an additional mechanism at work beyond the response to learned belief offered by Jovanovic (1982) as an explanation for higher exit rates of young establishments.

The effect of capital on exit (a 0.3% hazard reduction for a million peso increase in capital service) appears small and is not statistically significant. However, measured in these units there is considerable variability in capital (see Table 4A). A one standard deviation change in capital (34.9 million 1985 pesos) would reduce the exit hazard by 9.6%.

In all of the estimated decision rules, coefficients on time dummies have substantive significance as representations of the variation in industry state and economic environment. For exit hazard, 1980, 1984, and 1986 relative hazard coefficients are exactly zero because no establishments exited the sample between these and the subsequent years (see Table 2).[56] The null hypothesis of no year-to-year variation in hazard (H0d) is rejected at a very high level of significance.

Table 15B summarizes results for the same three decision rules — exit hazard, investment rate, and employment rate — for the food, apparel, and paper industries. The exit hazard results for these three industries are generally similar to the fabricated metals results. Increases in productivity belief reduce exit hazards, although significantly only in apparel and for $pred(t+1) - pred(t)$ in food. Age significantly reduces the exit hazard in food, but is insignificant in paper and (insignificantly) increases the exit hazard in apparel. Temporal variation in exit hazard is strongly significant in all three industries, as was the case for fabricated metals. Finally, the results on capital are stronger in these industries. More capital decreases the exit hazard, but here the result is significant for paper and at the 10% level for food. Paper is much more capital intensive than the other three industries so it is not surprising that capital has a clearer effect on survival in this industry than the others.

These exit hazard results are important for three reasons. First, they confirm that estimated productivity beliefs from the sequential learning estimator can indeed be related to establishment behavior in second stage estimates. Furthermore, the coefficients on productivity belief all have the expected sign and are often statistically significant. Second, both productivity and capital reduce exit hazard, confirming the presence selection that will bias results from estimators that do not take this into account. Third, the finding of significant variation among the annual coefficients is consistent with industry state and economic environment affecting exit behavior. This is especially noteworthy because it indicates that future work

[56]No estimate is reported for 1996 because there is no way of inferring whether any establishments operating that year exited before 1997.

to include explicit measures of industry structure and economic environment in estimated decision rules has a chance of success.

Investment Rates

Results for the investment rate decision rule are in Table 15A's second column. Investment is also forward-looking, since it changes capital inputs in future periods. Consequently, $pred(t)$ and $pred(t+1) - pred(t)$ are again used as regressors along with age and year dummies. Investment is divided by current capital stock to produce investment rates. The investment rate is highly skewed with a median of 5.7% and a mean of 21.3%, and also highly variable with a standard deviation of 56.6%.

As expected, the relation between productivity belief and investment rate is positive. For $pred(t)$ it is also statistically significant. The coefficients on $pred(t)$ and $pred(t+1)-pred(t)$ are elasticities between productivity and capital stock. A one percent increase in the productivity belief at the beginning of the current period increases investment before the next period by an amount sufficient to increase capital by 0.21%. This suggests that capital stock has not fully adjusted to the current productivity belief. Although the difference is not significant (see H0c), the smaller coefficient on the change in productivity belief is also consistent with partial capital adjustment.[57] On average, a one percent change in productivity belief results in a 0.15% increase in next year's capital stock. Together, these results confirm the presence of strategic endogeneity between capital inputs and establishment productivity.

Age has a substantial and significant negative effect on investment rate. On average each year of age reduces investment by 1.4% of capital stock (i. e. 25% of mean investment rate). This is conditional on the level and current change in productivity belief, again indicating that establishment age affects strategic behavior through some means independent of learning about productivity. The large magnitude of age's affect on investment rate is partly a reflection of the youth of establishments in the analysis sample.

Unlike exit, investment rate behavior does not vary significantly across years in the sample (H0d). This may well be due to large variation in investment rate rather than absence of an underlying relationship with industry structure or economic environment. The estimated coefficients vary widely, but the extreme values occur in the third and fourth years when sample sizes are small.

For the other three industries in Table 15B, results on the investment rate decision rules are not as clear-cut as they were for fabricated metals. None of the productivity belief coefficients are significant. However, five of the six coefficients take the expected positive

[57]Evidence of partial capital adjustment suggests that the type of dynamic panel instruments proposed by Arellano and Bond (1991) and Arellano and Bover (1995) — lagged levels instrumenting input differences — would have some identifying power. However, inspection of Table 10 shows that this estimator still produces an implausibly low estimate of returns to scale.

sign. As before, age reduces investment rate, but none of the coefficients is significant for these three industries. Neither are differences among coefficients on year dummies significant in any of these industries.

Employment Rates

The third decision rule in Table 15A is for employment rate. This is a labor-capital ratio measured as blue-collar equivalent person years per million 1985 pesos worth of capital services.[58] These are the same labor and capital variables used in estimation of the production function above. The mean employment rate is 19.2, which works out to one blue-collar job-year per 52,100 1985 pesos. In 2004 U.S. dollars this is roughly one employee per \$407 in capital services.

There is no strong *a priori* basis to expect any particular sign for the coefficients in the employment rate decision rule. Productivity has been specified as Hicks neutral. Therefore if all inputs were freely adjustable and establishments were static cost-minimizers there would be no relation between the productivity and the labor-capital ratio in employment rate. With capital quasi-fixed and not fully adjusted to productivity beliefs, there may be a non-zero relationship between productivity and employment rate. With constant elasticity residual demand and increasing short-run marginal cost this relationship will be positive — increases in production outweighing substitution of productivity for labor.

Employment rate affects production inputs in the current period. Therefore, assumptions of the sequential learning model imply that $pred(t+1)$ should not be included as a regressor. For fabricated metals, the sign on $pred(t)$ is negative but not significant. Evaluated at the mean, the implied elasticity of employment rate with respect to productivity is -0.36. Age also has a negative but insignificant effect. The employment rate decision rule has statistically significant variation from year to year (H0d).[59] Since employment is a more variable input than investment, this significant temporal variation is more likely a response to demand conditions than to strategic considerations. Employment and other variable inputs can respond to contemporaneous information about the economic environment without violating any assumption of the sequential learning estimator.

The other three estimated employment rate decision rules present a mixed picture (Table 15B). The coefficient on productivity belief is positive, but not significant in the food and apparel industries. In the paper industry the effect of productivity belief is negative and just significant at the 5% threshold. Unlike the result in fabricated metals, year-to-year variation is not significant in these three industries.

[58] See Appendix D for detailed discussion of construction of the labor and capital services variables.

[59] This result should be considered tentative. The extreme coefficient values in 1983 and 1984 are attributable to extreme employment rate outliers in three observations from two establishments.

Validity Tests

The investment rate and employment rate decision rules are augmented to perform validity tests on the information and decision timing assumptions, as described in the estimation section. Recall that smoothed productivity estimates are the Kalman filter estimates using data from an establishment's entire history. Because idiosyncratic productivity is persistent, future observations are just as informative as past observations when inferring current productivity. The information and decision timing assumptions imply that establishment decisions should not be affected by this additional information revealed by subsequent production experience. In other words, decision rules for current inputs, such as employment, should be independent of $smooth(t) - pred(t)$ and forward-looking decision rules, such as for investment rate, should be independent of $smooth(t+1) - pred(t+1)$.[60]

Table 16 presents validity test results for all four industries. The equations are specified in exactly the same manner as the decision rule equations reported in Tables 15A and 15B except that $smooth(t+1) - pred(t+1)$ or $smooth(t) - pred(t)$ are included. The validity test is passed if the null hypothesis that the coefficient on these additional variables equals 0 is not rejected. The sequential learning validity test is passed in seven out of eight cases, including all of the forward-looking investment rate decision rules. The one instance of a significant coefficient on $smooth(t) - pred(t)$ is in the food industry employment rate decision rule. Notice that coefficient estimates for the other variables are nearly unchanged from Tables 15A and 15B in all eight decision rules. This is indirect confirmation that $smooth(t) - pred(t)$ is orthogonal to $pred(t)$ (as well as age), as it should be if the sequential learning model has been correctly estimated.

In three out of four industries the data do not reject validity of the sequential learning assumptions in tests using investment and employment behavior. The validity test based on employment rate can be quite stringent since it is easier for establishments to adjust employment in response to production experience during the period. Therefore, it is encouraging that the sequential learning estimator passed this test in three out of the four industries. But it is also reassuring that the validity test has enough power to occasionally reject appropriateness of the estimator for specific industries.

The main use here of the employment rate decision rule is as a potentially more stringent validity test on the sequential learning estimator's assumptions.

4.4 An Indirect Test of Olley and Pakes' Invertibility Condition

Olley and Pakes' (1996) semiparametric production function estimator relies upon strict monotonicity of investment with respect to productivity belief, conditional on age, current

[60] Exit hazard does not provide a meaningful validity test because in the year before exit $smooth(t+1) - pred(t+1) = 0$ by construction of $smooth(t+1)$.

capital, and positive investment. They invert this relation to nonparametrically infer productivity beliefs from investment. Pakes (1994) develops sufficient conditions for this strict monotonicity condition to hold. He assumes productivity, μ_{it}, evolves according to an exogenous Markov process and gross current payoffs can be expressed as a reduced form of the firm's current capital and productivity, $\pi_{it} = \pi(\mu_{it}, k_{it})$.[61] Then the firm's dynamic problem can be summarized in a Bellman value function

$$V(\mu_{it}, k_{it}) = \max \left\{ V_x, \max_{i_{it}} \left\{ \pi(\mu_{it}, k_{it}) - c(i_{it}, k_{it}) + \phi \mathrm{E}_t[V(g(\mu_{it}), k_{it}(1-\delta) + i_{it})] \right\} \right\} \quad (24)$$

Where V_x is an exit scrap value, ϕ is the discount rate, δ is rate of depreciation, i_{it} is investment, $c(i_{it}, k_{it})$ is the capital adjustment cost, $g(\mu_{it})$ is the stochastic Markov productivity transition function, and the expectation is over next period's realization of μ_{it+1} conditional on μ_{it}. Then supermodularity of $\pi(\mu_{it}, k_{it})$, convexity of capital adjustment cost, supermodularity of the negative of capital adjustment cost, and stochastic dominance and regularity conditions on $g(\mu_{it})$ are sufficient for investment to be strictly increasing in current productivity belief among firms with positive investment (Pakes [1994]).

Since sequential learning estimates of productivity beliefs are based on a different set of assumptions, the compatibility of the Olley-Pakes invertibility condition with these estimates can be assessed. In principle, the invertibility condition and the sequential learning assumptions could both hold. In that case apparent violations of invertibility should be entirely due to random errors in the productivity belief estimates.[62] If the frequency of apparent violations of the invertibility condition are more frequent than can be accounted for by random estimation error, this would be evidence that the invertibility condition, the sequential learning assumptions, or both are not applicable in that industry. So this amounts to a joint test of the simultaneous correctness of both models.

Because the invertibility is conditional on current capital and the chances of observing two establishments with exactly equal capital are exceedingly small, a direct test is impractical. However, Pakes' sufficient conditions also imply a more readily observed capital dominance persistence result.[63]

$$u_{it} \geq u_{jt} \text{ and } k_{it} \geq k_{jt} \Rightarrow k_{it+1} \geq k_{jt+1} \quad (25)$$

Table 17 shows the results when this condition is checked using estimated productivity beliefs from the sequential learning estimator. As an example consider fabricated metals in the last

[61] Absence of strategic interactions and full response of variable inputs to the current state permit this simplification.

[62] Or mismeasurement of capital. But that would cause problems in both the sequential learning and Olley-Pakes estimators.

[63] Let $V^T(\mu_{it}, k_{it})$ denote the value function for a finite T period problem. Pakes shows that $V^T(.)$ is supermodular. It follows from the stochastic dominance of $g(\mu_{it})$ that $\mathrm{E}[V^{T-1}|\mu_{it}]$ is also supermodular in μ_{it} and k_{it+1}. Redefine the choice variable as $k_{it+1} \equiv i_{it} + (1-\delta)k_{it}$. Since k_{it+1} is monotonic in i_{it}, negative capital adjustment cost is also supermodular in k_{it+1} and k_{it}. Therefore, the action-specific value function $V^T(\mu_{it}, k_{it}, k_{it+1})$ is supermodular in k_{it+1} and $s_{it} \equiv (\mu_{it}, k_{it})$. Thus $\mu'_{it} \geq \mu''_{it}$ and $k'_{it} \geq k''_{it}$ implies $k'_{it+1} \geq k''_{it+1}$. Taking limits as T goes to infinity completes the argument.

column. Pooling years, there are 82,294 ordered establishment pairs or 41,147 distinct pairs if order is ignored. The same pairs usually re-appear in several years and each occurrence is counted because each is an opportunity for dominance. There were 22,984 instances of dominance as defined in the first part of Equation 25. Of these, comparison in the subsequent year could not be made in 2,495 cases because at least one establishment in the pair switched sectors or had a gap in reported data. This leaves 20,489 cases to evaluate the implication in Equation 25. The implication was violated 2,647 times because the dominated establishment had more capital in the following year — a "capital reversal." The implication was violated an additional 819 times when the dominating establishment exited (i. e. capital=0) but the dominated establishment continued in operation — an "exit reversal." Combined, the reversals account for 16.9% of instances of dominance that can be checked. Olley and Pakes (1996) also condition on age in their nonparametric productivity inversion. The bottom panel of Table 17 limits the analysis to pairs of establishments coming from the same age cohort. This substantially reduces the number of observed instances of dominance, but has only a small effect on the percentage of reversals.

Scanning across the four industries, the percentage of reversals ranges from 9.4% in the paper industry when conditioning on age to 26.3% in apparel when considering all establishment pairs. The lower rate of reversals in paper is not surprising since that industry is both more capital intensive and has lower average mark-ups. The first increases the importance of capital decisions while the second is generally an indicator of weaker product differentiation. There are a number of reasons that the implication in Equation 25 may be violated so often. Capital adjustment costs may be non-convex. Productivity transitions may not exhibit first-order stochastic dominance. This may especially be true when embodied technology implies that the act of investment itself changes productivity. Establishments may not be dynamically optimizing. Firms may limit capacity as a collusive device. An important candidate explanation is that establishments do not face uniform output or input prices. This could be due to product differentiation and/or from pooling geographically distinct markets.

It should be re-emphasized that this is an indirect test. First, although similar in spirit, the condition in Equation 25 is not the invertibility condition itself. Rather, it can be derived from a set of assumptions that are also sufficient for the invertibility condition. Second, these dominance reversal results depend on productivity belief estimates produced by the sequential learning estimator. Nevertheless, Table 17 is a strong reminder that there are industries, perhaps many, that do not conform to the various characterizations of establishment behavior used by many estimators. This confirms the importance of having a productivity estimator that does not rely on behavioral restrictions.

5 Conclusion

The sequential learning estimator developed in this paper has several useful properties. It replaces explicit or implicit assumptions about firm behavior and industry equilibrium with a model of productivity beliefs formed through repeated learning from production experience. Second, the form of productivity dynamics can be specified with a great deal of flexibility. In particular, the estimator can accommodate effects of endogenous innovation effort on productivity. Third, by simulating establishments' productivity beliefs the sequential learning estimator computes forecast errors that are independent of all predetermined input choices. Consequently, it resolves the econometric issues of selection and strategic endogeneity that arise because establishment productivity and capital are both state variables in the dynamic industry game. Because of these features, the sequential learning estimator allows productivity dynamics to be studied in non-competitive industries where assumptions underlying alternative estimators do not hold.

The sequential leaning estimator also offers advantages as a first stage in empirical modelling of dynamic industry equilibria, especially when productivity and innovation effort are important contributors to the evolution of industry structure. Establishment behavior decision rules can be estimated in a second stage after sequential learning generates productivity dynamics and productivity belief estimates in a first stage. Previously, establishment behavior had to be estimated simultaneously with productivity. In principle, after introducing some structure into the dynamic industry game one could take a third step and estimate value functions and properties of the equilibrium. In this way the sequential learning estimator is likely to prove complementary to the emerging literature on structural estimation of industry games.

Another benefit of the sequential learning estimator is that it produces small standard errors on production function parameter estimates. This is a consequence of using cohorts to model entrants' initial productivity beliefs. Some between-establishment within-cohort variation is preserved in the forecast errors, especially in a cohort's early years. This produces standard errors closer to the random effects estimator than to previous estimators that address endogeneity. Preserving more inter-establishment variation also appears to push estimated returns to scale closer to constant returns.

To produce all of these advantages the sequential learning estimator makes several assumptions about the process establishments use to form beliefs about their own productivities. Establishment knowledge of true parameter values, entrants' rational expectations about their initial productivities, and passive learning about own productivity are each common aspects of the literature on dynamic industry models. Normality of the entrant productivity data generating process, productivity shocks, and noise is a strong assumption. But the practical consequences of its violation will usually be small, since Kalman updating still provides the best conditional linear estimate of productivity.

The two remaining assumptions are that establishments learn from the same information contained in the econometrician's data and that all input decisions are predetermined at the time the forecast error is revealed. These are strong and crucial to justification of the sequential learning estimator. Fortunately, they have a testable joint implication. Establishment decision rules should be independent of information about productivity revealed after the decisions are presumed to be taken. This additional information is summarized in the difference between estimated productivity belief and the Kalman filter's smoothed productivity estimate using an establishment's entire history in the panel. In three out of four industries this joint implication was not rejected in estimated decision rules for investment rate and employment rate. In those industries the sequential learning assumptions pass this validity test. The one instance (employment rates in the food industry) where the joint implication was rejected demonstrates that the sequential learning estimator will not be valid in every instance. However, it also demonstrates that these validity tests have power.

As a proof of concept, production functions, productivity dynamics, and decision rules were estimated for four industries using a panel of Chilean manufacturing establishments. First and foremost, this illustrates the feasibility of using the sequential learning estimator to estimate productivity and then decision rules in separate stages.

Several additional results support the conclusion that sequential learning produces good production function estimates. Estimated returns to scale are within 4% of constant in each industry. Factor elasticity shares are usually close to cost shares. The ranking of imputed mark-ups among the industries is plausible. Industry average establishment sizes and capital intensities are consistent with estimated returns to scale, capital elasticity, and imputed mark-up, especially in the paper industry.

There is indirect evidence that establishments are uncertain about their own productivity and learn from production experience as assumed. For entrants average establishment size and variability in size are less than for incumbents, which have had an opportunity to learn about their productivities. Estimates of productivity dynamics strongly support the presence of persistent productivity shocks, which contribute to uncertainty. Most importantly, estimated productivity beliefs are indeed related to establishment behavior. In estimated decision rules they almost always have the anticipated sign and are statistically significant in three out four exit hazard equations and in one investment rate equation.

Furthermore, empirical results indicate that assumptions made by alternative estimators are not satisfied for the four industries examined. In each industry a capital dominance persistence property derived from the same sufficient conditions used to justify Olley and Pakes' (1996) investment invertibility is frequently incompatible with estimated productivity beliefs. The restrictive productivity dynamics of fixed and random effects estimators are strongly rejected.

Future applications of the sequential learning estimator have the potential to address many questions about microeconomic productivity dynamics and its behavioral sources. The results here provide preliminary evidence that variations in industry state and economic environment do affect establishment behavior. These unmeasured variations are captured by annual dummy variables, which differ significantly in several of the estimated decision rules, especially for exit hazard. Identifying which changes in industry characteristics and economic condition affect establishment behaviors, explaining how these changes modify establishments' incentives for such behaviors, and deriving consequences for the course of industry-level productivity and technology are important directions for additional research. The sequential learning estimator will be a useful tool to explore these questions in a wide variety of industries by allowing estimation of productivity separately from description of establishment behavior and dynamic industry equilibrium.

References

[1] Ackerberg, Daniel A., Lanier Benkard, Steven Berry, and Ariel Pakes. 2005. "Econometric Tools for Analyzing Market Outcomes." Prepared for J. Heckman, ed., *Handbook of Econometrics*.

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

[3] Aghion, Phillippe, Christopher Harris, Peter Howitt, and John Vickers. 2001. "Competition, Imitation and Growth with Step-by-Step Innovation." *Review of Economic Studies*, 68:467-492.

[4] Aghion, Philippe, Chritopher Harris, and John Vickers. 1997. "Competition and Growth with Step-by-Step Innovation: An Example." *European Economic Review*, 41:771-782.

[5] Aghion, Philippe and Peter Howitt. 1998. *Endogenous Growth Theory*. MIT Press.

[6] Alvarez, Roberto and Ricardo A. López. 2004. "Exporting and Performance: Evidence from Chilean Plants." Manuscript.

[7] Altonji, Joseph G. and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics*, 116(1):313-350.

[8] Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2):277-297.

[9] Arellano, Manuel and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Component Models." *Journal of Econometrics*, 68:29-51.

[10] Asplund, Marcus and Volker Nocke. 2002. "Imperfect Competition, Market Size, and Firm Turnover." WP.

[11] Baltagi, Badi H. and Ping X. Wu. 1999. "Unequally Spaced Panel Data Regressions with AR(1) Disturbances." Disturbances." *Econometric Theory* 15:814-823.

[12] Baraji, Patrick, Lanier Benkard, and Jonathan Levin. 2005. "Estimating Dynamic Models of Imperfect Competition." Competition." Manuscript.

[13] Bartelsman, Eric J. and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." Microdata." *Journal of Economic Literature*, 38(3):569-594.

[14] Basu, Susanto. 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" *Quarterly Journal of Economics*, 111(3):719-751.

[15] Basu, Sussanto and John G. Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, 105(2):249-283.

[16] Basu, Susanto and John G. Fernald. 2002. "Aggregate Productivity and Aggregate Technology." *European Economic Review*, 46:963-991.

[17] Basu, Susanto and Miles S. Kimball. 1997. "Cyclical Productivity with Unobserved Input Variation." National Bureau of Economic Research. Working Paper 5915.

[18] Beggs, Alan and Paul Klemperer. 1992. "Multi-Period Competition with Switching Costs." *Econometrica*, 60(3):651-666.

[19] Bergoeing, Raphael, Patrick J. Kehoe, Timothy J. Kehoe, and Raimundo Soto. 2002. "A Decade Lost and Found: Mexico and Chile in the 1980s." *Review of Economic Dynamics*, 5(1):166-205.

[20] Bergoeing, Raphael, Andrés Hernando, and Andrea Repetto. 2003. "Idiosyncratic Productivity Shocks and Plant-Level Heterogeneity." Manuscript.

[21] Berry, Steven and Ariel Pakes. 2000. "Estimation from the Optimality Conditions for Dynamic Controls." Manuscript.

[22] Besecker, Ronald T. n.d. "Can Monetary Correction be Explained, and Even Better Yet, Managed?" AmCham Journal. Available at The Chilean American Chamber of Commerce website www.amchamchile.cl.

[23] Blundell, Richard and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics*, 87:115-143.

[24] Blundell, Richard and Stephen Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions." *Econometric Reviews*, 19(3):321-340.

[25] Bruno, Michael. 1978. "Duality, Intermediate Inputs and Value-Added." in Melvyn Fuss and Daniel McFadden, eds. *Production Economics: A Dual Approach to Theory and Applications, Volume 2: Applications of the Theory of Production.* North Holland, pp. 3-16.

[26] Budd, Christopher, Christopher Harris, and John Vickers. 1993. "A Model of the Evolution of Duopoly: Does Does the Asymmetry between Firm Tend to Increase or Decrease?" *Review of Economic Studies*, 60(3):543-573.

[27] Canadian Institute of Chartered Accountants, Colegio de Contadores de Chile A.G., Instituto Mexicano do Contadores Públicos A.G., Financial Accounting Standards Board of the United States. 1998. "Significant Differences in GAAP in Canada, Chile, Mexico, and the United States: An Analysis of Accounting Pronouncements as of June 1998."

[28] Cabral, Luis M. B. and Michael H. Riordan. 1994. "The Learning Curve, Market Dominance, and Predatory Pricing." *Econometrica*, 62(5):1115-1140.

[29] Cornwell, Christopher, Peter Schmidt, and Robin C. Sickles. 1990. "Production Frontiers with Cross-Sectional Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics*, 46:185-200.

[30] Cox, D. R. 1975. "Partial Likelihood." *Biometrika*, 62(2):269-276.

[31] Cramer, J. S. 1986. *Econometric Applications of Maximum Likelihood Methods*. Cambridge University Press.

[32] de Jong, Piet. 1988. "The Likelihood for a State Space Model." *Biometrika*, 75(1):165-169.

[33] Doraszelski, Ulrich and Kenneth J. Judd. 2004. "Avoiding the Curse of Dimensionality in Dynamic Stochastic Games." Manuscript.

[34] Ericson, Richard and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62(1):53-82.

[35] Fajnzylber, Pablo, William F. Maloney, and Eduardo Riberio. 2001. "Firm Entry and Exit, Labor Demand, and Trade Reform: Evidence form Chile and Columbia." Manuscript.

[36] Farber, Henry S. and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics*, 111(4):1007-1047.

[37] Gill, Jeff and Gary King. 2004. "What to Do When Your Hessian Is Not Invertible: Alternatives to Model Respecification in Nonlinear Estimation." *Sociological Methods & Research*, 32(4):1-34.

[38] Gould, William, Jeffrey Pitblado, and William Sribney. 2003. *Maximum Likelihood Estimation with Stata*. 2nd ed. Stata Press.

[39] Gowrisankaran, Gautam and Robert J. Town. 1997. "Dynamic Equilibrium in the Hospital Industry," *Journal of Economics and Management Strategy*, 6(1):45-74.

[40] Griliches, Zvi and Dale W. Jorgenson. 1966. "Sources of Measured Productivity Change: Capital Input." *American Economic Review*, 56(1/2):50-61.

[41] Griliches, Zvi and Jacques Mairesse. 1995. "Production Functions: The Search for Identification." National Bureau of Economic Research. Working Paper 5067.

[42] Grossman, Gene M. and Elhanen Helpman. 1991. "Quality Ladders in the Theory of Growth." *Review of Economic Studies*, 58(1):43-61.

[43] Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy*, 96(5):921-947.

[44] Harvey, Andrew C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

[45] Hopenhayn, Hugo. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5):1127-1150.

[46] Hopenhayn, Hugo and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy*, 101(5):915-938.

[47] Hsiao, Cheng. 1986. *Analysis of Panel Data*. Econometric Society Monographs No. 11. Cambridge University Press.

[48] Hsieh, Chang-Tai and Jonathan A. Parker. 2002. "Taxes and Growth in a Financially Underdeveloped Country: Evidence from the Chilean Investment Boom." Manuscript.

[49] Hulten, Charles R. 1978. "Growth Accounting with Intermediate Inputs." *Review of Economic Studies*, 45(3):511-518.

[50] Hulten, Charles R. 2001. "Total Factor Productivity: A Short Biography." pp. 1-47 in Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, eds. *New Developments in Productivity Analysis*. National Bureau of Economic Research, Studies in Income and Wealth, Vol. 63. University of Chicago Press.

[51] Jorgenson, Dale W. and Zvi Griliches. 1967. "The Explanation of Productivity Change." *Review of Economic Studies*, 34(3):249-283.

[52] Jorgenson, Dale W., Kevin J. Stiroh, Robert J. Gordon, and Daniel E. Sichel. 2000. "Raising the Speed Limit: U.S. Economic Growth in the Information Age." *Brookings Papers on Economic Activity*, 2000(1):125-235.

[53] Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica*, 50(3):649-670.

[54] Jovanovic, Boyan and Glenn M. MacDonald. 1994a. "Competitive Diffusion." *Journal of Political Economy*, 102(1):24-52

[55] Jovanovic, Boyan and Glenn M. MacDonald. 1994b. "The Life Cycle of a Competitive Industry." *Journal of Political Economy*, 102(2):322-347.

[56] Jovanovic, Boyan and Rafael Rob. 1987. "Demand-Driven Innovation and Spatial Competition Over Time." *Review of Economic Studies*, 54(1):63-72.

[57] Kandilov, Ivan. 2005. "Do Exporters Pay More? Plant-level Evidence from an Export Refund Policy in Chile." Manuscript.

[58] Klemperer, Paul. 1995. "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade." *Review of Economic Studies*, 62(4):515-539.

[59] Klepper, Steven. 1996. "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *American Economic Review*, 86(3):562-583.

[60] Klepper, Steven and Elizabeth Graddy. 1990. "The Evolution of New Industries and the Determinants of Market Structure." *Rand Journal of Economics*, 21(1):27-44.

[61] Lambson, Val Eugene. 1991. "Industry Evolution with Sunk Costs and Uncertain Market Conditions." *International Journal of Industrial Organization*, 9(2):171-196.

[62] Lambson, Val Eugene. 1992. "Competitive Profits in the Long Run." *Review of Economic Studies*, 59(1):125-142.

[63] Lambson, Val Eugene and Farrell E. Jensen. 1995. "Sunk Costs and the Variability of Firm Value Over Time." *Review of Economics and Statistics*, 77(3):535-544.

[64] Lambson, Val Eugene and Farrell E. Jensen. 1998. "Sund Costs and Firm Value Variability: Theory and Evidence." *American Economic Review*, 88(1):307-313.

[65] Lange, Fabian. 2003. "The Returns to Schooling and Ability During the Early Career: Evidence on Job Market Signaling, Employer Learning and Post-School Investments." Manuscript.

[66] Levinsohn, James. 1999. "Employment Responses to International Liberalization in Chile." *Jounral of International Economics*, 47:321-344.

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

[68] Lippman, S.A. and R. P. Rumelt. 1982. "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency Under Competition." *Bell Journal of Economics*, 13(2):418-438.

[69] Lipster, Robert S. and Albert N. Shiryaev. 2001. *Statistics of Random Processes. Vol II. Applications*. 2nd ed. Trans. A. B. Aries. Applications of Mathematics Vol. 6. Springer-Verlag.

[70] Liu, Lili. 1991. *Entry-Exit and Productivity Change: An Empirical Analysis of Efficiency Frontiers*. Dissertation, Department of Economics, University of Michigan.

[71] Liu, Lili. 1993. "Entry-Exit, Learning, and Productivity Change." *Journal of Development Economics*, 42:217-242.

[72] Marschak, Jacob and William H. Andrews, Jr. 1944. "Random Simultaneous Equations and the Theory of Production." *Econometrica*, 12(3/4):143-205.

[73] Muendler, Marc-Andreas. 2004. "Estimating Production Functions when Productivity Change is Endogenous." Department of Economics, University of California, San Diego, Working Paper 2004-05.

[74] Mundlak, Y. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica*, 46(1):69-85.

[75] Nadiri, M. Ishaq and Ingmar R. Prucha. 2001. "Dynamic Factor Demand Models and Productivity Analysis." pp.103-171 in Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, eds. *New Developments in Productivity Analysis*. National Bureau of Economic Research, Studies in Income and Wealth, Vol. 63. University of Chicago Press.

[76] Newey, Whitney K. and Caniel McFadden. 1994. “Large Sample Estimation and Hypothesis Testing.” pp. 2111-2245 in R. F. Engle and D. L. McFadden, eds. *Handbook of Econometrics, Volume IV*, Elsevier Science.

[77] Olley, G. Steven and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6):1263-1297.

[78] Pakes, Ariel. 1994. “Dynamic Structural Models, Problems and Prospects: Mixed Continuous Discrete Controls and Market Interactions.” pp. 171-259 in Christopher A. Sims, ed. *Advances in Econometrics: Sixth World Congress*, Volume II, Econometric Society Monographs No. 24. Cambridge University Press.

[79] Pakes, Ariel. 2000. “A Framework for Applied Dynamic Analysis in I.O.” National Bureau of Economic Research, WP8024.

[80] Pakes, Ariel and Richard Ericson. 1998. “Empirical Implications of Alternative Models of Firm Dynamics.” *Journal of Economic Theory*, 79:1-45.

[81] Pakes, Ariel and Paul McGuire. 1994. “Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model.” *Rand Journal of Economics*, 25(4):555-589.

[82] Pakes, Ariel and Paul McGuire. 2001. “Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the ’Curse’ of Dimensionality.” *Econometrica*, 59(5):1261-1281.

[83] Pakes, Ariel, Michael Ostrovsky, and Steven Berry. 2004. “Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples).” Manuscript.

[84] Pavcnik, Nina. 2002. “Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants.” *Review of Economic Studies*, 69:245-276.

[85] Petrakis, Emmanuel and Santanu Roy. 1999. “Cost-Reducing Investment, Competition, and Industry Dynamics.” *International Economic Review*, 40(2):381-401.

[86] Petrin, Amil and James Levinsohn. 2004. “Measuring and Mismeasuring Industry Productivity Growth Using Plant-Level Data.” Manuscript.

[87] Petrin, Amil and Jagadeesh Sivadasan. 2003. “Impact of Increases in Job Security on Employment and Output — Evidence from Chile.” Manuscript.

[88] Reinganum, Jennifer F. 1985. “Innovation and Industry Evolution.” *Quarterly Journal of Economics*, 100(1):81-99.

[89] Roberts, Mark J. and James R. Tybout, eds. 1996. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Oxford University Press for the World Bank.

[90] Sargent, Thomas J. 1993. *Bounded Rationality in Macroeconomics*. Clarendon Press.

[91] Sargent, Thomas J., Noah Williams, and Tao Zha. 2004. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." National Bureau of Economic Research. Working Paper 10764.

[92] Slade, Margaret E. 1989. "Modeling Stochasitc and Cyclical Components of Technical Change: An Application of the Kalman Filter." *Journal of Econometrics*, 41:363-383.

[93] Spence, Michael A. 1981. "The Learning Curve and Competition." *Bell Journal of Economics*, 12(1):49-70.

[94] Sutton, John. 1998. *Technology and Market Structure: Theory and History*. MIT Press.

[95] Tybout, James R. 1996. "Chile, 1979-86: Trade Liberalization and Its Aftermath." in Mark J. Roberts and James R. Tybout, eds. *Industrial Evolution in Developing Countries*. Oxford University Press for the World Bank. Pp. 200-226.

[96] Tybout, James R. 1991. "The Effects of Trade Reforms on Scale and Technical Efficiency: New Evidence from Chile." *Journal of International Economics*, 31:231-250.

[97] Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Appendices

The following appendices are available online at:

`http://www.econ.ox.ac.uk/members/david.greenstreet/`

Or by request to: david.greenstreet@economics.ox.ac.uk

A Illustrative Alternate Specifications

B Equations for AR(1) Plus Establishment Effect with Noise

C Estimation Algorithm

D Chilean Data Sources and Preparation

Table 1A
Establishments in Source Data (ENIA 1979-1996)

		Observations		Entrants		Exits			
Industry (ISIC)	Establishments	Total	Per Year	Total	Per Year	Total	Per Year	Entry Rate	Exit Rate
Food (311)	2,860	24,495	1,360.8	1,323	73.5	1,412	78.4	5.4%	5.8%
Apparel (322)	864	5,886	327.0	422	23.4	503	27.9	7.2%	8.5%
Paper (341)	145	1,194	66.3	75	4.2	60	3.3	6.3%	5.0%
Fabricated Metal (381)	921	7,105	394.7	462	25.7	414	23.0	6.5%	5.8%

Table 1B
Establishments in Analysis Data

		Observations			Average
Industry (ISIC)	Establishments	Total	Per Year	Cohorts	Cohort Size
Food (311)	831	3,740	220.0	1980-1996	48.9
Apparel (322)	297	1,141	67.1	1980-1996	17.5
Paper (341)	68	297	21.2	1983, 1984, 1986-1996	4.9
Fabricated Metal (381)	325	1,157	68.1	1980-1996	19.1

Table 2

Establishment Counts by Year and Status

(Fabricated Metal Products ISIC 381)

	Total Industry						Analysis Data				
Year	Exit	Outward Sector Switch or Data Gap	Incumbent	Entrant	Inward Sector Switch or Data Gap	Total	Exit	Outward Sector Switch or Data Gap	Incumbent	Entrant	Total
1979						459					
1980	42	58	359	16	72	447				7	7
1981	36	42	369	10	34	413	0	1	6	2	8
1982	48	39	326	7	32	365	1	1	6	2	8
1983	50	37	278	9	35	322	3	0	5	5	10
1984	24	18	280	48	30	358	3	0	7	30	37
1985	6	18	334	7	10	351	0	1	36	2	38
1986	23	9	319	18	10	347	4	0	34	13	47
1987	14	46	287	41	28	356	0	6	41	30	71
1988	15	34	307	12	29	348	4	9	58	12	70
1989	15	54	279	21	60	360	5	13	52	18	70
1990	16	24	320	4	27	351	4	6	60	3	63
1991	12	16	323	33	18	374	2	5	56	22	78
1992	15	19	340	29	36	405	5	5	68	18	86
1993	18	41	346	44	30	420	6	8	72	35	107
1994	20	21	379	36	29	444	5	5	97	23	120
1995	25	38	381	51	38	470	9	10	101	42	143
1996	35	2	433	76	6	515	8	2	133	61	194

Table 3
Cohorts
(Fabricated Metal Products ISIC 381)

	Total Industry			Analysis Data		
Cohort	Plants	1996 Survivors	Observations	Plants	1996 Survivors	Observations
0	459	204	4,966			
1980	16	3	106	7	2	47
1981	10	3	54	2	0	6
1982	7	3	68	2	0	3
1983	9	4	89	5	2	43
1984	48	24	434	30	12	232
1985	7	4	56	2	1	20
1986	18	9	133	13	3	69
1987	41	22	297	30	9	152
1988	12	4	53	12	2	33
1989	21	14	134	18	7	85
1990	4	2	21	3	0	10
1991	33	26	175	22	11	88
1992	29	22	118	18	10	63
1993	44	29	146	35	23	113
1994	36	21	83	23	14	53
1995	51	45	96	42	37	79
1996	76	76	76	61	61	61

Table 4A
Descriptive Statistics
(Fabricated Metal Products ISIC 381)

		Mean	Median	Variance	Skewness
Output	All	232.9	99.2	2.20E+05	6.53
	Entrant	155.1	75.5	1.07E+05	8.57
Log Output	All	11.6	11.5	1.18	0.53
	Entrant	11.3	11.2	0.95	0.67
Labor	All	55.8	34.3	4.37E+03	4.39
	Entrant	43.8	29.0	2.40E+03	4.40
Log Labor	All	3.7	3.5	3.67	0.77
	Entrant	3.5	3.4	0.49	0.81
Capital Service	All	15.3	5.0	1.22E+03	5.92
	Entrant	11.5	4.3	6.56E+02	6.56
Log Capital Service	All	8.5	8.5	2.14	0.04
	Entrant	8.3	8.4	2.11	-0.20
Energy	All	4.0	1.6	7.57E+01	8.19
	Entrant	3.0	1.1	7.34E+01	9.81
Log Energy	All	7.4	7.4	1.90	-0.09
	Entrant	7.0	7.0	1.83	-0.03
Materials	All	113.7	42.1	7.98E+04	8.14
	Entrant	72.4	31.3	2.59E+04	7.14
Log Materials	All	10.8	10.6	1.41	0.51
	Entrant	10.4	10.4	1.20	0.54
Investment	All	0.1	0.0	2.91E-01	9.01
	Entrant	0.1	0.0	2.88E-01	4.04
Log Investment	All	3.6	3.7	3.26	-0.11
	Entrant	3.6	3.7	2.88	-0.27

Notes: Analysis data (full cohort histories) only.
Labor in blue-collar equivalent person years.
All other levels in millions 1985 Chilean pesos.
All other logarithms of thousands of 1985 Chilean pesos.

Table 4B
Descriptive Statistics

		Mean	Median	Variance
Food:	Output	397.6	84.4	6.59E+05
	Labor	94.7	45.3	2.06E+04
	Capital Service	33.5	4.2	8.40E+03
	Energy	19.2	3.1	2.48E+03
	Materials	234.0	49.5	2.22E+05
	Investment	0.3	0.0	2.47E+00
Apparel:	Output	137.3	52.1	6.88E+04
	Labor	58.1	30.4	6.91E+03
	Capital Service	8.9	2.8	3.38E+02
	Energy	1.6	0.5	1.62E+01
	Materials	72.7	28.5	1.85E+04
	Investment	0.0	0.0	5.62E-02
Paper:	Output	1,918.0	177.6	1.23E+07
	Labor	185.0	52.4	7.34E+04
	Capital Service	555.7	20.9	5.25E+06
	Energy	108.9	4.7	8.52E+04
	Materials	1,020.1	91.0	3.40E+06
	Investment	5.7	0.0	2.92E+03

Notes: Analysis data (full cohort histories) only.
Labor in blue-collar equivalent person years.
All other levels in millions 1985 Chilean pesos.
All other logarithms of thousands of 1985 Chilean pesos.

Table 5
Aggregate Revenue and Cost Shares

	Food	Apparel	Paper	Fabricated Metals
Revenue Shares:				
Capital	8.4%	6.5%	29.0%	6.6%
Labor	8.7%	15.9%	6.9%	12.8%
Energy	4.8%	1.2%	5.7%	1.7%
Materials	58.8%	53.0%	53.2%	48.8%
Misc. Services	5.2%	9.0%	5.5%	5.0%
Total	86.0%	85.5%	100.2%	74.9%
Imputed Profit Share	14.0%	14.5%	-0.2%	25.1%
Cost Shares:				
Capital	9.8%	7.6%	28.9%	8.8%
Labor	10.1%	18.6%	6.8%	17.1%
Energy	5.6%	1.4%	5.7%	2.3%
Materials	68.4%	61.9%	53.1%	65.2%
Misc. Services	6.1%	10.5%	5.5%	6.6%

Note: Based on analysis data.

Table 6
Simple Correlations Among Analysis Variables
(In Natural Logaritms)

Food

	Output	Capital	Labor	Energy	Materials
Output	1.00				
Capital	0.80	1.00			
Labor	0.77	0.71	1.00		
Energy	0.79	0.74	0.68	1.00	
Materials	0.95	0.73	0.71	0.72	1.00

Apparel

	Output	Capital	Labor	Energy	Materials
Output	1.00				
Capital	0.74	1.00			
Labor	0.83	0.74	1.00		
Energy	0.69	0.66	0.71	1.00	
Materials	0.91	0.64	0.70	0.57	1.00

Paper

	Output	Capital	Labor	Energy	Materials
Output	1.00				
Capital	0.88	1.00			
Labor	0.92	0.83	1.00		
Energy	0.84	0.84	0.83	1.00	
Materials	0.97	0.85	0.87	0.80	1.00

Fabricated Metal

	Output	Capital	Labor	Energy	Materials
Output	1.00				
Capital	0.72	1.00			
Labor	0.76	0.59	1.00		
Energy	0.68	0.58	0.63	1.00	
Materials	0.92	0.60	0.65	0.58	1.00

Table 7
Sequential Learning Estimator
(Fabricated Metal Products ISIC 381)

	DP	SP	RW	EE
Production Function:				
Capital	0.100	0.103	0.097	0.084
	(0.011)	(0.011)	(0.011)	(0.010)
Labor	0.256	0.265	0.262	0.271
	(0.020)	(0.020)	(0.020)	(0.020)
Energy	0.049	0.049	0.050	0.056
	(0.010)	(0.010)	(0.010)	(0.010)
Materials	0.557	0.558	0.555	0.570
	(0.014)	(0.013)	(0.014)	(0.014)
Returns to Scale	0.962	0.975	0.964	0.980
	(0.018)	(0.018)	(0.018)	(0.017)
R11 (AR(1) coefficient)	0.726	0.914	1.000	1.000
	(0.120)	(0.029) .	.	
Variances:				
W11 (entrants' AR(1) component)	0.035	0.079	0.072	0.067
	(0.012)	(0.008)	(0.007)	(0.007)
W22 (unchanging establishment effect)	0.049			
	(0.012)			
Q (persistent AR(1) shock)	0.015	0.011	0.007	0.000
	(0.005)	(0.003)	(0.001)	
sigma2 (noise)	0.022	0.027	0.030	0.040
	(0.004)	(0.003)	(0.002)	(0.002)
Year Dummies:				
yr80	3.392	3.327	3.417	3.287
	(0.179)	(0.179)	(0.179)	(0.177)
yr81	3.472	3.419	3.517	3.378
	(0.184)	(0.183)	(0.187)	(0.181)
yr82	3.354	3.319	3.422	3.306
	(0.183)	(0.181)	(0.187)	(0.179)
yr83	3.504	3.249	3.322	3.179
	(0.194)	(0.183)	(0.193)	(0.177)
yr84	3.379	3.154	3.187	2.992
	(0.197)	(0.186)	(0.200)	(0.174)
yr85	3.333	3.191	3.223	3.023
	(0.193)	(0.180)	(0.201)	(0.175)
yr86	3.351	3.282	3.311	3.112
	(0.195)	(0.175)	(0.203)	(0.176)
yr87	3.304	3.268	3.287	3.084
	(0.195)	(0.170)	(0.204)	(0.176)
yr88	3.427	3.426	3.444	3.244
	(0.198)	(0.167)	(0.205)	(0.177)
yr89	3.320	3.350	3.372	3.170
	(0.200)	(0.165)	(0.207)	(0.178)
yr90	3.356	3.399	3.426	3.225
	(0.200)	(0.161)	(0.208)	(0.179)
yr91	3.323	3.369	3.399	3.199
	(0.201)	(0.158)	(0.208)	(0.179)
yr92	3.380	3.421	3.452	3.250
	(0.202)	(0.156)	(0.209)	(0.179)
yr93	3.499	3.542	3.574	3.372
	(0.204)	(0.154)	(0.210)	(0.179)
yr94	3.464	3.504	3.536	3.337
	(0.204)	(0.152)	(0.210)	(0.179)
yr95	3.491	3.544	3.579	3.379
	(0.205)	(0.150)	(0.211)	(0.180)
yr96	3.501	3.583	3.618	3.419
	(0.206)	(0.148)	(0.211)	(0.180)

Initial Productivity continued…

Notes:
Asymptotic standard errors in parentheses. See text for explanation.
Variance estimates transformed from square roots with delta method.
DP: Double process - AR(1) and establishment effect.
SP: Single process - AR(1).
RW: Random walk.
EE: Establishment effect.

Table 7 continued
Sequential Learning Estimator
(Fabricated Metal Products ISIC 381)

	DP	SP	RW	EE
Initial Productivity: w1				
(AR(1) component)				
cohort81	0.507	-0.234	-0.249	-0.237
	(0.457)	(0.254)	(0.247)	(0.238)
cohort82	2.927	-0.134	-0.165	-0.147
	(1.603)	(0.257)	(0.255)	(0.251)
cohort83	-0.306	0.124	0.159	0.203
	(0.225)	(0.192)	(0.194)	(0.171)
cohort84	-0.358	-0.011	0.049	0.127
	(0.171)	(0.145)	(0.158)	(0.130)
cohort85	-0.371	-0.022	0.069	0.172
	(0.258)	(0.250)	(0.254)	(0.224)
cohort86	-0.064	0.017	0.071	0.130
	(0.155)	(0.144)	(0.172)	(0.144)
cohort87	-0.192	-0.079	-0.002	0.084
	(0.136)	(0.122)	(0.161)	(0.131)
cohort88	-0.045	-0.034	0.042	0.102
	(0.191)	(0.136)	(0.176)	(0.149)
cohort89	0.055	-0.090	-0.024	0.040
	(0.119)	(0.119)	(0.168)	(0.138)
cohort90	0.425	-0.078	-0.041	0.026
	(0.333)	(0.199)	(0.230)	(0.204)
cohort91	0.038	-0.003	0.059	0.129
	(0.114)	(0.105)	(0.166)	(0.135)
cohort92	-0.079	-0.096	-0.030	0.049
	(0.137)	(0.106)	(0.170)	(0.138)
cohort93	-0.061	-0.083	-0.017	0.068
	(0.117)	(0.089)	(0.163)	(0.130)
cohort94	-0.424	0.008	0.073	0.163
	(0.240)	(0.094)	(0.168)	(0.136)
cohort95	-0.260	0.022	0.081	0.161
	(0.221)	(0.079)	(0.163)	(0.130)
cohort96		-0.028	0.028	0.107
		(0.072)	(0.162)	(0.129)
Initial Productivity: w2				
(establishment effect component)				
cohort81	-0.673			
	(0.435)			
cohort82	-2.913			
	(1.577)			
cohort83	0.207			
	(0.209)			
cohort84	0.180			
	(0.171)			
cohort85	0.281			
	(0.260)			
cohort86	0.088			
	(0.189)			
cohort87	0.133			
	(0.177)			
cohort88	0.080			
	(0.226)			
cohort89	-0.027			
	(0.182)			
cohort90	-0.324			
	(0.349)			
cohort91	0.075			
	(0.183)			
cohort92	0.090			
	(0.197)			
cohort93	0.086			
	(0.187)			
cohort94	0.496			
	(0.282)			
cohort95	0.387			
	(0.267)			
cohort96	0.123			
	(0.164)			

Table 8
Sequential Learning Estimator
(Double Process - AR(1) plus Establishment Effect)

	Food	Apparel	Paper
Production Function:			
Capital	0.104	0.105	0.094
	(0.007)	(0.014)	(0.022)
Labor	0.161	0.366	0.315
	(0.011)	(0.024)	(0.042)
Energy	0.090	0.069	0.030
	(0.006)	(0.012)	(0.017)
Materials	0.620	0.473	0.577
	(0.007)	(0.013)	(0.026)
Returns to Scale	0.975	1.013	1.016
	(0.009)	(0.019)	(0.027)
R11 (AR(1) coefficient)	0.737	0.504	0.556
	(0.062)	(0.143)	(0.132)
Variances:			
W11 (entrants' AR(1) component)	0.068	0.045	0.029
	(0.010)	(0.013)	(0.018)
W22 (unchanging establishment effect)	0.034	0.056	0.045
	(0.009)	(0.009)	(0.014)
Q (persistent AR(1) shock)	0.022	0.037	0
	(0.004)	(0.012)	
sigma2 (noise)	0.038	0.017	0.034
	(0.003)	(0.010)	(0.003)
Year Dummies:			
yr80	2.452	4.077	
	(0.095)	(0.193)	
yr81	2.392	3.824	
	(0.097)	(0.217)	
yr82	2.334	3.390	
	(0.092)	(0.205)	
yr83	2.344	2.827	3.032
	(0.090)	(0.212)	(0.318)
yr84	2.331	2.879	3.022
	(0.088)	(0.226)	(0.298)
yr85	2.327	2.924	2.891
	(0.087)	(0.241)	(0.278)
yr86	2.369	2.988	3.156
	(0.087)	(0.254)	(0.272)
yr87	2.351	3.006	2.889
	(0.088)	(0.262)	(0.277)
yr88	2.354	3.021	2.898
	(0.088)	(0.265)	(0.276)
yr89	2.344	3.026	2.945
	(0.090)	(0.268)	(0.280)
yr90	2.353	3.057	2.991
	(0.090)	(0.270)	(0.280)
yr91	2.396	3.082	3.089
	(0.092)	(0.272)	(0.281)
yr92	2.440	3.161	3.108
	(0.093)	(0.273)	(0.282)
yr93	2.452	3.067	3.072
	(0.094)	(0.273)	(0.282)
yr94	2.484	2.965	2.959
	(0.094)	(0.274)	(0.281)
yr95	2.447	2.969	2.949
	(0.095)	(0.275)	(0.281)
yr96	2.418	2.778	2.849
	(0.096)	(0.275)	(0.280)

Initial Productivity continued…

Notes: Asymptotic standard errors in parentheses. See text for explanation.
Variance estimates transformed from square roots with delta method.
Paper industry estimated Q at binding constraint. Remaining estimates are conditional.

Table 8 continued
Sequential Learning Estimator
(Double Process - AR(1) plus Establishment Effect)

	Food	Apparel	Paper
Initial Productivity: w1			
(AR(1) component)			
cohort81	0.124	-0.354	
	(0.141)	(0.317)	
cohort82	0.067	-0.441	
	(0.136)	(0.287)	
cohort83	0.146	0.061	
	(0.110)	(0.205)	
cohort84	0.130	0.185	0.140
	(0.092)	(0.142)	(0.221)
cohort85	0.104	-0.406	
	(0.100)	(0.432)	
cohort86	-0.013	0.011	-0.245
	(0.089)	(0.117)	(0.161)
cohort87	-0.056	-0.062	0.116
	(0.071)	(0.087)	(0.127)
cohort88	-0.005	0.010	-0.893
	(0.077)	(0.130)	(0.238)
cohort89	0.208	-0.156	-0.239
	(0.097)	(0.115)	(0.133)
cohort90	0.061	-0.107	0.532
	(0.088)	(0.104)	(0.435)
cohort91	-0.136	-0.022	-0.230
	(0.092)	(0.079)	(0.147)
cohort92	0.128	-0.128	-0.165
	(0.086)	(0.097)	(0.213)
cohort93	-0.228	-0.076	0.121
	(0.123)	(0.167)	(0.198)
cohort94	-0.157	0.055	-0.507
	(0.159)	(0.132)	(0.236)
cohort95	0.217	0.153	0.243
	(0.187)	(0.172)	(0.369)
Initial Productivity: w2			
(establishment effect component)			
cohort81	0.054	0.076	
	(0.105)	(0.293)	
cohort82	0.120	0.642	
	(0.109)	(0.298)	
cohort83	0.067	0.622	
	(0.095)	(0.263)	
cohort84	0.051	0.485	0.183
	(0.088)	(0.256)	(0.206)
cohort85	0.104	0.884	
	(0.099)	(0.487)	
cohort86	0.069	0.558	0.036
	(0.094)	(0.269)	(0.214)
cohort87	0.078	0.455	0.143
	(0.085)	(0.261)	(0.201)
cohort88	0.068	0.524	0.225
	(0.090)	(0.279)	(0.276)
cohort89	0.116	0.687	0.441
	(0.107)	(0.276)	(0.204)
cohort90	0.081	0.433	-0.412
	(0.101)	(0.272)	(0.411)
cohort91	0.145	0.629	0.322
	(0.103)	(0.264)	(0.214)
cohort92	0.121	0.536	0.446
	(0.105)	(0.273)	(0.252)
cohort93	0.309	0.596	0.257
	(0.124)	(0.301)	(0.245)
cohort94	0.193	0.466	0.520
	(0.155)	(0.282)	(0.267)
cohort95	-0.013	0.633	-0.019
	(0.191)	(0.298)	(0.373)
cohort96	0.052	0.589	0.115
	(0.082)	(0.262)	(0.218)

Table 9
Wald Tests of Constant Returns to Scale
(Sequential Learning Estimator)

		Food	Apparel	Paper	Fabricated Metals
DP	chi2	7.81	0.46	0.35	4.43
	df	1	1	1	1
	prob	0.005	0.497	0.552	0.035
SP	chi2	6.37	0.62	0.73	1.97
	df	1	1	1	1
	prob	0.012	0.430	0.393	0.160
RW	chi2	18.14	0.54	0.07	3.96
	df	1	1	1	1
	prob	0.000	0.461	0.796	0.047
EE	chi2	6.92	1.26	0.02	1.32
	df	1	1	1	1
	prob	0.009	0.262	0.880	0.251

Notes: DP: Double process - AR(1) and establishment effect.
SP: Single process - AR(1).
RW: Random walk.
EE: Establishment effect.

Table 10
Comparison Estimators
(Fabricated Metal Products ISIC 381)

Inconsistent Estimators:

	OLS	Random Effects	Panel AR(1) Random Effects
Capital	0.107	0.084	0.098
	(0.009)	(0.010)	(0.010)
Labor	0.261	0.269	0.263
	(0.017)	(0.020)	(0.020)
Energy	0.067	0.056	0.055
	(0.009)	(0.010)	(0.010)
Materials	0.608	0.571	0.570
	(0.011)	(0.014)	(0.013)
Returns to Scale	1.043	0.980	0.987
	(0.012)	(0.017)	(0.016)
AR(1)			0.246

Estimators Consistent in T or N:

	Fixed Effects	Panel AR(1) Fixed Effects	Between Differences	Arellano Bond	Levinsohn Petrin Energy Proxy	Levinsohn Petrin Materials Proxy
Capital	0.056	0.048	0.110	0.052	0.082	0.071
	(0.012)	(0.017)	(0.031)	(0.019)	(0.044)	(0.048)
Labor	0.253	0.171	0.299	0.146	0.251	0.240
	(0.024)	(0.029)	(0.046)	(0.033)	(0.029)	(0.025)
Energy	0.039	0.041	-0.045	0.028	0.029	0.078
	(0.012)	(0.013)	(0.024)	(0.015)	(0.029)	(0.017)
Materials	0.517	0.492	0.509	0.497	0.588	0.485
	(0.018)	(0.022)	(0.033)	(0.024)	(0.016)	(0.138)
Returns to Scale	0.865	0.752	0.873	0.723	0.950	0.874
	(0.025)	(0.032)	(0.052)	(0.037)	(0.073)	(0.153)
AR(1)		0.246				

Notes: All regressions run on full history cohort analysis data n=1,157.
All regressions include year dummies.
Standard errors in parentheses.
Between differences is standard between estimator run on first differences as recommended by Griliches and Mairesse (1995).
Panel AR(1) estimator is from Baltagi and Wu (1999).

Table 11
LR Tests of Nested Specifications
(Sequential Learning Estimator)

		Food	Apparel	Paper	Fabricated Metals
SP vs DP	chi2	12.86	14.08	NA	27.92
	df	16	16		16
	prob	0.683	0.593		0.032
RW vs DP	chi2	105.93	25.02	NA	39.46
	df	17	17		17
	prob	0.000	0.094		0.002
EE vs DP	chi2	214.75	65.04	NA	76.99
	df	18	18		18
	prob	0.000	0.000		0.000
RW vs SP	chi2	93.06	10.93	8.09	11.54
	df	1	1	1	1
	prob	0.000	0.001	0.004	0.001
EE vs SP	chi2	201.89	50.96	13.18	49.07
	df	2	2	2	2
	prob	0.000	0.000	0.001	0.000
EE vsRW	chi2	108.83	40.03	5.09	37.53
	df	1	1	1	1
	prob	0.000	0.000	0.024	0.000

Notes: DP: Double process - AR(1) and establishment effect.
SP: Single process - AR(1).
RW: Random walk.
EE: Establishment effect.
Tests are not valid for paper industry DP specification due to binding parameter constraint.

Table 12
Distributions of Persistent Idiosyncratic Establishment Productivities
(Smoothed Estimates)

		Food	Apparel	Paper	Fabricated Metals
Mean		0.114	0.529	0.207	0.074
Standard Deviation		0.271	0.317	0.264	0.260
Percentiles	10th	-0.173	0.132	-0.116	-0.242
	90th	0.451	0.908	0.504	0.413
	Difference	0.624	0.776	0.620	0.655
	As Ratio	1.87	2.17	1.86	1.93

Table 13

Common Technological Progress

(Sequential Learning Estimator)

		Food	Apparel	Paper	Fabricated Metals
H0a	chi2	34.46	125.24	40.50	45.61
	df	16	16	13	16
	prob	0.005	0.000	0.000	0.000
H0b	chi2	0.14	24.77	0.75	0.45
	df	1	1	1	1
	prob	0.704	0.000	0.386	0.501
H0c	chi2	2.29	2.10	0.38	1.78
	df	1	1	1	1
	prob	0.130	0.148	0.538	0.182

Notes: Double process specification - AR(1) and establishment effect.
Wald tests.
H0a: B_yr80=B_yr81=…=B_yr96, except paper - B_yr83=…=B_yr96.
H0b: B_yr80=B_yr96, except paper - B_yr83=B_yr96.
H0c: B_yr85=B_yr96, except paper - B_yr88=B_yr96.

Table 14

Initial Productivities:

(Cohort Comparisons and Learning from Experience)

		Food	Apparel	Paper	Fabricated Metals
Cohort Comparisons:					
H0a	chi2	42.80	34.57	28.90	7.28
	df	15	15	11	15
	prob	0.000	0.003	0.002	0.949
H0b	chi2	0.82	6.37	0.66	1.03
	df	1	1	1	1
	prob	0.366	0.012	0.415	0.310
H0c	chi2	3.07	0.24	11.43	0.66
	df	1	1	1	1
	prob	0.080	0.626	0.001	0.416
Learning from Experience:					
H0d	chi2	17.25	12.93	26.68	11.61
	df	15	15	11	15
	prob	0.304	0.608	0.005	0.709
H0e	chi2	15.08	5.94	21.00	7.11
	df	11	11	8	11
	prob	0.179	0.878	0.007	0.790

Notes: Double process specification - AR(1) and establishment effect.
Wald tests.
H0a: (B_[w1]cohort81+B_[w2]cohort81)=(B_[w1]cohort82+B_[w2]cohort82)=…
=(B_[w1]cohort96+B_[w2]cohort96), paper excludes 1980-1982 and 1985.
H0b: (B_[w1]cohort81+B_[w2]cohort81)=(B_[w1]cohort96+B_[w2]cohort96),
paper (B_[w1]cohort84+B_[w2]cohort84)=(B_[w1]cohort96+B_[w2]cohort96)
H0c: (B_[w1]cohort85+B_[w2]cohort85)=(B_[w1]cohort96+B_[w2]cohort96),
paper (B_[w1]cohort88+B_[w2]cohort88)=(B_[w1]cohort96+B_[w2]cohort96)
H0d: B_[w1]cohort81=0, B_[w1]cohort82=0, … B_[w1]cohort95=0.
paper excludes 1981, 1982 and 1985.
H0e: B_[w1]cohort85=0, B_[w1]cohort86=0, … B_[w1]cohort95=0.
paper excludes 1985-1987.

Table 15A
Decision Rules: Linear Approximation
(Fabricated Metal Products ISIC 381)

	Exit Hazard	Investment Rate	Employment Rate
pred(t)	0.172	0.205	-7.034
	(0.129)	(0.092)	(12.064)
pred(t+1)-pred(t)	0.134	0.153	
	(0.099)	(0.110)	
age	0.892	-0.014	-0.323
	(0.049)	(0.005)	(0.712)
capital	0.997		
	(0.005)		
Year Dummies			
yr80	0.000	0.234	12.902
	(0.000)	(0.213)	(27.871)
yr81	0.110	0.273	8.827
	(0.106)	(0.199)	(26.123)
yr82	0.192	-0.023	6.023
	(0.162)	(0.201)	(26.117)
yr83	0.217	0.622	157.998
	(0.127)	(0.179)	(23.405)
yr84	0.000	0.301	83.684
	(0.000)	(0.094)	(12.303)
yr85	0.131	0.203	35.552
	(0.061)	(0.093)	(12.117)
yr86	0.000	0.238	24.045
	(0.000)	(0.084)	(10.914)
yr87	0.077	0.394	23.538
	(0.039)	(0.068)	(8.932)
yr88	0.112	0.402	20.515
	(0.051)	(0.070)	(9.105)
yr89	0.092	0.210	17.267
	(0.048)	(0.071)	(9.187)
yr90	0.057	0.234	19.237
	(0.043)	(0.075)	(9.847)
yr91	0.110	0.179	18.150
	(0.051)	(0.068)	(8.921)
yr92	0.121	0.254	17.774
	(0.050)	(0.066)	(8.602)
yr93	0.075	0.262	16.362
	(0.031)	(0.059)	(7.747)
yr94	0.136	0.233	15.824
	(0.056)	(0.058)	(7.470)
yr95	0.110	0.235	13.390
	(0.430)	(0.054)	(7.001)
yr96		0.196	14.626
		(0.048)	(6.170)
Prob Values:			
H0a	0.019	0.026	0.560
H0b	0.007	0.166	
H0c	0.813	0.702	
H0d	0.000	0.231	0.000

Notes:

Exit hazard coefficients are exponentiated hazard ratios.
Investment rate is linear regression on investment/capital stock.
Employment rate is linear regression on employment/capital service - blue collar-equivalent person-years per million 1985 pesos.
Pred is establishment's prediction of idiosyncratic productivity.
Capital is service aggregate including rental payments in millions 1985 pesos.
Standard errors in parentheses. Not corrected for generated regressors - see text.
H0a: B_pred(t)=0 for investment rate and employment rate. B_pred(t)=1 for exit hazard.
H0b: B_{pred(t+1)-pred(t)]=0 for investment rate and employment rate. B_[pred(t+1)-pred(t)]=1 for exit hazard.
H0c: B_pred(t)=B_[pred(t+1)-pred(t)].
H0d: B_yr80=B_yr81=…B_yr96.

Table 15B
Decision Rules: Linear Approximation
(Key Results for Other Industries)

	Exit Hazard	Investment Rate	Employment Rate
Food:			
Coefficients:			
pred(t)	0.643	0.064	50.207
	(0.226)	(0.126)	(42.717)
pred(t+1)-pred(t)	0.278	-0.189	
	(0.145)	(0.169)	
age	0.936	-0.010	-4.660
	(0.021)	(0.008)	(2.679)
capital	0.998		
	(0.001)		
Prob Values:			
H0a	0.209	0.613	0.240
H0b	0.014	0.264	
H0c	0.129	0.185	
H0d	0.000	0.248	0.406
Apparel:			
Coefficients:			
pred(t)	0.342	0.394	10.625
	(0.160)	(1.800)	(37.629)
pred(t+1)-pred(t)	0.297	0.628	
	(0.140)	(2.293)	
age	1.011	-0.033	-3.785
	(0.040)	(0.147)	(3.124)
capital	0.989		
	(0.007)		
Prob Values:			
H0a	0.021	0.827	0.778
H0b	0.010	0.784	
H0c	0.823	0.931	
H0d	0.000	0.174	0.990
Paper:			
Coefficients:			
pred(t)	0.342	0.151	-5.020
	(0.641)	(0.124)	(2.549)
pred(t+1)-pred(t)	0.844	0.181	
	(2.060)	(0.236)	
age	0.981	-0.013	0.028
	(0.096)	(0.009)	(0.207)
capital	0.991		
	(0.004)		
Prob Values:			
H0a	0.567	0.223	0.050
H0b	0.945	0.443	
H0c	0.740	0.906	
H0d	0.000	0.293	0.745

Notes:
Exit hazard coefficients are exponentiated hazard ratios.
Investment rate is linear regression on investment/capital stock.
Employment rate is linear regression on employment/capital service
- blue collar-equivalent person-years per million 1985 pesos.
Pred is establishment's prediction of idiosyncratic productivity.
Capital is service aggregate including rental payments in millions 1985 pesos.
Standard errors in parentheses. Not corrected for generated regressors - see text.
Models also include time dummies.
H0a: B_pred(t)=0 for investment rate and employment rate. B_pred(t)=1 for exit hazard.
H0b: B_[pred(t+1)-pred(t)]=0 for investment rate and employment rate. B_[pred(t+1)-pred(t)]=1 for exit hazard.
H0c: B_pred(t)=B_[pred(t+1)-pred(t)].
H0d: B_yr80=B_yr81=…B_yr96.

Table 16
Decision Rule Validity Tests
(of Sequential Learning)

	Food	Apparel	Paper	Fabricated Metals
Investment Rate:				
Coefficients:				
smooth(t+1)-pred(t+1)	-0.218	0.992	-0.307	0.088
	(0.187)	(2.246)	(0.380)	(0.153)
pred(t)	0.063	0.390	0.147	0.207
	(0.126)	(1.801)	(0.124)	(0.092)
pred(t+1)-pred(t)	-0.188	0.634	0.160	0.153
	(0.169)	(2.294)	(0.237)	(0.110)
age	-0.010	-0.033	-0.013	-0.014
	(0.008)	(0.147)	(0.009)	(0.005)
Prob Values:				
H0	0.244	0.659	0.420	0.566
Employment Rate:				
Coefficients:				
smooth(t)-pred(t)	179.719	50.166	-2.673	9.807
	(45.119)	(34.530)	(4.137)	(12.231)
pred(t)	50.686	10.339	-5.065	-6.893
	(42.632)	(37.611)	(2.552)	(12.068)
age	-4.648	-3.783	0.028	-0.326
	(2.674)	(3.12)	(0.21)	(0.71)
Prob Values:				
H0	0.000	0.147	0.519	0.423

Notes:

Investment rate is linear regression on investment/capital stock.
Employment rate is linear regression on employment/capital service - blue collar-equivalent person-years per million 1985 pesos.
Pred is establishment's prediction of idiosyncratic productivity.
Smooth is best idiosyncratic productivity estimate using complete establishment history.
Standard errors in parentheses. Not corrected for generated regressors - see text.
Models also include time dummies.
Sequential learning information and decision timing assumptions imply H0.
H0: B_(smooth(t)-pred(t))=0.

Table 17
Frequency of Capital Rank Reversals

		Food	Apparel	Paper	Fabricated Metals
All Plants					
Frequency of:	Establishment Pairs	866,848	82,668	6,092	82,924
	Dominance:				
	Total	248,871	20,501	1,804	22,984
	With Data in t+1	238,134	19,027	1,558	20,489
	Capital Reversals	30,014	3,794	166	2,647
	Exit Reversals	15,251	1,216	42	819
Percentage Reversals:	Capital Only	12.6%	19.9%	10.7%	12.9%
	Capital and Exit	19.0%	26.3%	13.4%	16.9%
Conditional on Age					
Frequency of:	Establishment Pairs	108,130	15,560	966	15,616
	Dominance:				
	Total	21,718	2,977	191	2,737
	With Data in t+1	20,624	2,763	159	2,426
	Capital Reversals	2,592	513	13	293
	Exit Reversals	1,195	152	2	102
Percentage Reversals:	Capital Only	12.6%	18.6%	8.2%	12.1%
	Capital and Exit	18.4%	24.1%	9.4%	16.3%

Notes:

Dominance: Capital and productivity belief greater for one establishment than the other.
Capital Reversal: Capital in the subsequent period is greater for the dominated establishment.
Exit Reversal: The dominated establishment survives in the next period but the dominating establishment exits.
With Data in t+1 excludes establishments departing the panel in the next period for sector switches or data problems.
Capital based on capital services aggregate.

Table D1
Inferred versus Reported Capital for Entrants

		Type of Fixed Asset		
		Building	Machinery, etc	Vehicles
	Mean Reported	148,182	402,467	23,101
	Correlation	0.60	0.84	0.09
Regression:	Constant	11,002	103,606	20,914
		(24,861)	(46,842)	(4,623)
	Coefficient	1.310	0.929	0.131
		(0.038)	(0.013)	(0.031)

Notes: In thousands of nominal Chilean pesos.
Based on entrants in all manufacturing sectors 1992-1996 n=2118.
See Appendix D text for computation of inferred capital.

Figure 1
Time-Line for Dynamic Industry Game During Period t

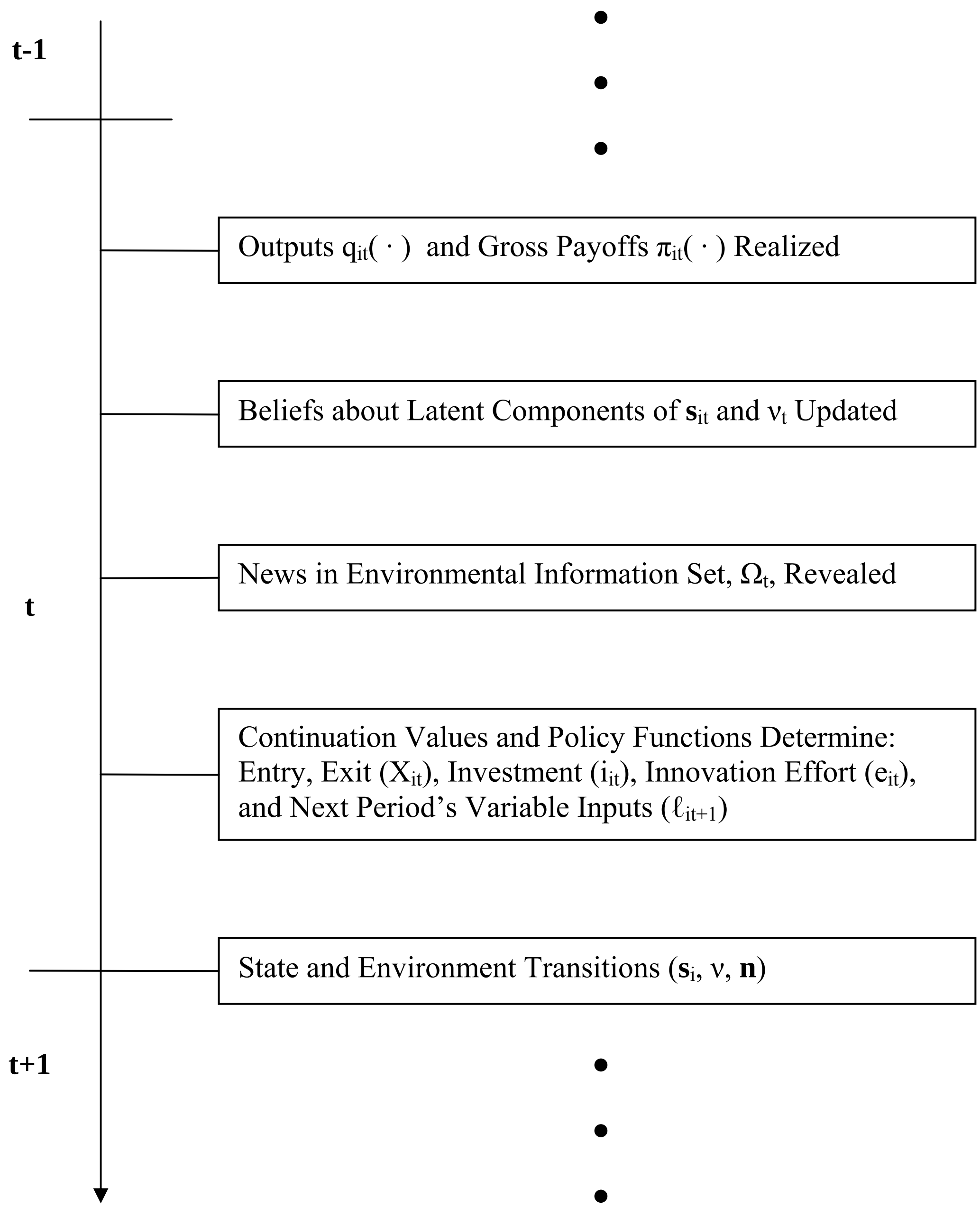
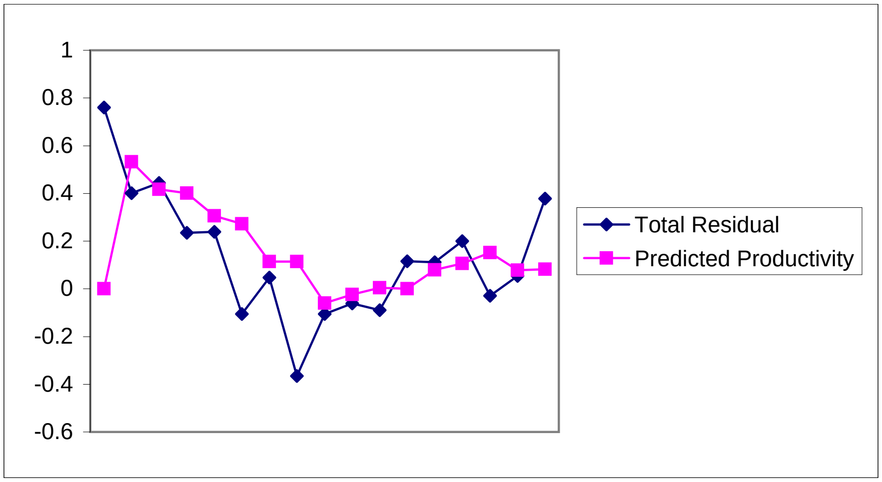
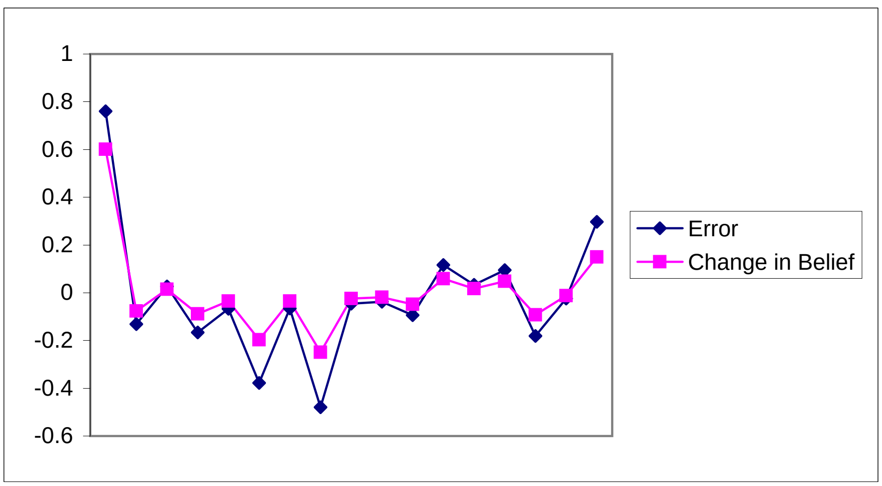


Figure 2
An Illustration of Productivity Belief Dynamics





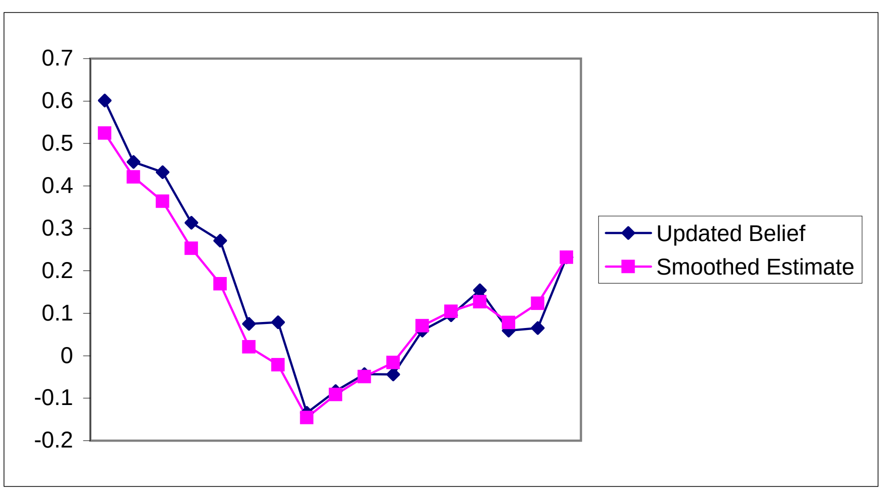


Figure C-1
Sequential Learning Algorithm