

Business Cycles and the Failure of Marginal Firms

Michael Horvath*

Department of Economics, Stanford University

June 1999

Abstract

This paper considers the effects of plant entry, learning, and exit on such aggregate measures as productivity, output, and job creation/destruction. The working hypothesis is that of Schumpeter: a portion of productivity and output fluctuations associated with business cycles is due to the ongoing birth and death of plants and the consequent reallocation of resources between plants of differing abilities this engenders. The economy is modeled as a continuum of plants with different levels of productivity. The composition of plants in the economy evolves through plant learning. Through experience in producing some plants deduce that they are unproductive and exit the industry while other plants are reaffirmed as productive plants, and remain in the industry. The resolution of uncertainty regarding plant productivity endogenously generates persistent aggregate productivity fluctuations resembling observations from U.S. data in response to completely transitory shocks.

JEL Classification: E1, E32, L16.

Keywords: Aggregate Fluctuations, Productivity, Entry, Exit.

*Address: Department of Economics, Stanford University, Stanford, CA 94305, (650) 723-4116, mhorvath@stanford.edu. An earlier version of this paper was circulated under the title *Business Cycles Without Shocks: A Model of Firm Selection*. Financial support from the National Science Foundation (SBR-9507978) is gratefully acknowledged. This paper was written, in part, while I was enjoying the hospitality of the University of Wisconsin-Madison, department of economics. Rishi Goyal provided very helpful research assistance. I have benefited from discussions about this topic with Susanto Basu, Bob Hall, Pierre-Olivier Gourinchas, Chad Jones, Jonathan Parker, Randy Verbrugge, Michael Woodford and seminar participants at the Federal Reserve Bank of Minneapolis, Stanford University, UBC, UCLA, UVA, UW-Madison, VPI & SU and the NBER Impulse and Propagation Group. With respect to these, the usual disclaimer applies yet again.

1 Introduction

Research efforts aimed at explaining large macroeconomic fluctuations have focused for several decades on exogenous “shocks” as driving forces for the upturns and downturns evident in the data. Technology shocks or total factor productivity residuals (Solow residuals) have driven fluctuations in a host of dynamic macroeconomic models since the pioneering work of Brock and Mirman (1972) and Kydland and Prescott (1982). The lack of universal acceptance of the Real Business Cycle (RBC) framework for explaining macroeconomic fluctuations can be traced, in part, to the extreme difficulty experienced in identifying plausible candidates for sources of technology shocks (Cochrane [1994]). Without an explanation of the driving forces behind productivity variation, models that generate business cycles from Solow-residual-type shocks seem to avoid the very important question: What causes business cycles?

This has led some authors to investigate the reverse causality: output fluctuations give rise to productivity fluctuations.¹ While these models provide plausible mechanisms to link output movements with productivity movements, they do not explain the persistence in either output or productivity, *per se*. As with the RBC model, an exogenous driving process in these models which is itself inherently persistent causes persistent movements in output which, in turn, give rise to persistent movements in productivity. A third strand of the literature on productivity investigates possible mismeasurement of productivity due to cyclical capacity utilization.² While correcting for these measurement errors typically reduces the correlation of productivity with output it does not appreciably reduce persistence in the corrected productivity series. Finally, the recent literature on creative destruction³ ignores productivity fluctuations altogether and focuses solely on the timing of job creation and destruction in response to deterministic persistent fluctuations in demand.

This paper provides an explanation for *persistent* productivity variation in a dynamic stochastic general equilibrium (DSGE) model. Cyclical fluctuations in productivity and output are viewed as the outcome of the process of creative destruction envisioned by Schumpeter (1942). Productivity variation over the business cycle is explained as the aggregate manifestation of this process rather than appealing to an exogenous source of productivity variation to explain cyclical fluctuations as in the RBC literature. In this sense, persistence in productivity movements is an economic feature of the model rather than an assumed attribute of exogenous disturbances that are fed into the model.

The economy is modeled as a continuum of plants⁴ with different levels of productivity. There are two classes or types of plants, differing by their average level of productivity. Plant

¹See Basu and Fernald (1997, 1998) and Basu and Kimball (1994).

²See Burnside, Eichenbaum, and Rebelo (1995) and Basu (1996).

³See Caballero and Hammour (1994, 1996).

⁴The distinction between firm and plant is important when looking at data. In this theory there is no difference between firm, plant, or establishment, hence the term *plant* will be used when discussing the basic unit of production in the economy.

managers are uncertain about their plants' types. Through experience in producing a homogeneous good, plant managers update beliefs over their plants' types; some plants deduce that they are bad or unproductive and exit the industry while other plants are reaffirmed as good, or productive plants, and remain in the economy. Aggregate productivity is a simple function of the distribution of plants producing in the economy at any given point in time. In this view, aggregate productivity variation is a composition effect; productivity variation is explained by the endogenous variation in the distribution of plant types. This feature of the model is consistent with recent plant-level empirical studies on the reallocation of resources between plants of different productivity levels and the effect this has on industry productivity.⁵

In the model, productivity increases with the natural selection of the more productive plants as incumbents. Productivity declines with net exit by plants since inputs must then be reallocated to the remaining plants, each exhibiting decreasing returns to scale or limited *span of control* (Lucas [1978]). Persistence of aggregate productivity comes through the learning process at the plant level. Because it takes time for low productivity plants to learn that they are inefficient and exit, periods of low productivity caused by a relative abundance of low productivity plants can persist. Because high productivity plants tend to learn that they are efficient and hence survive longer than low productivity plants, periods of high productivity caused by a relative abundance of high productivity plants can persist. Therefore, learning by plants constitutes an endogenous propagation mechanism capable of generating persistent fluctuations in aggregate productivity and output, even in response to *i.i.d.* shocks to the distribution of plants.

There are three stark implications of the model that are not found in standard representative agent models of business cycles. First, productivity variation at the aggregate level appears to be exogenous to other macro variables in the model. Only an econometrician with knowledge of the evolution of the distribution of plants would be able to detect the endogenous nature of productivity fluctuations. This feature would explain why econometric efforts such as Cochrane (1994) to uncover the origins of cyclical variation from observations of aggregate time series are largely unsuccessful. Second, the model implies that *any* exogenous shock to the distribution of plants has a persistent effect on productivity. Uncorrelated, serially independent, random disturbances from a multitude of sources that result in plant entry or exit will engender persistent aggregate productivity variations. This feature implies that the identification of shocks that give rise to productivity variation is perhaps of secondary importance. The shocks themselves are not the source of persistence, and hence not the explanation for persistent business cycle fluctuations. The third implication of the model is an extension of the second; exogenous demand shocks that affect plant entry or exit will affect aggregate productivity. Several commonly used econometric

⁵See Foster, Haltiwanger, and Krizan (1998), Aw, Chen, and Roberts (1997), Olley and Pakes (1996), Bahk and Gort (1993), Gort, Bahk, and Wall (1993), Baily, Hulten and Campbell (1992), and Bartlesman and Dhrymes (1993). This literature is discussed briefly in section 5.2 of this paper.

techniques are dependent on the separation of demand shocks from productivity movements. For example, Hall (1988, 1990) and Basu and Fernald (1997) estimate returns to scale in production using demand shift instruments which they assume are independent of productivity. Galí (1996) and Kiley (1997) decompose macroeconomic fluctuations into a temporary and permanent component following the Blanchard and Quah (1989) methodology and then use the separation of demand shocks from productivity shocks to associate demand shocks with the temporary component and productivity shocks with the permanent component. If demand shocks cause fluctuations in productivity, returns to scale estimates using demand shift instruments will be biased upward and the temporary component of Blanchard-Quah decompositions may contain fluctuations in productivity, thus overstating the importance of demand shocks.

The model in this paper and that of Shleifer (1986) are similar in that both share the Schumpeterian view that the process of creative destruction is itself a major source of economic fluctuations. The distinctive feature of the model in this paper is that the process of learning at the plant level endows the model with an endogenous propagation mechanism capable of generating very high levels of persistence in aggregate productivity and output, even in response to *i.i.d.* aggregate shocks being fed into the model. Several other papers use entry and exit of production units to *magnify* the effects of persistent exogenous shocks.⁶ Caballero and Hammour (1994, 1996) employ the machinery of creative destruction in conjunction with persistent exogenous shocks to demand to explain manufacturing census data on job creation and destruction. Gilchrist and Williams (1998) explore the effect of low short-run substitutability between capital and labor in a vintage capital model and are able to generate increased persistence of aggregate endogenous variables in response to persistent shocks to technology or factor prices. Campbell (1998) also studies the effect of entry and exit of plants in a vintage capital model where all technological progress is embodied in new plants. The distinguishing feature of this paper is that the exogenous driving process need not be persistent in order to generate persistent responses of endogenous aggregate variables.

Two recent papers have employed learning to generate persistence in a macro model. Pries (1998) explains persistence in unemployment rates with a model in which a worker and a firm learn about the worker's productivity over time. The main distinction in this paper is its focus on explaining aggregate productivity fluctuations and providing an economic explanation for why these fluctuations are persistent enough to generate business cycles. Cooper and Johri (1999) explore the propagation of shocks via learning-by-doing in an otherwise standard neoclassical one-sector model. Their emphasis is on explaining persistence in economic output rather than productivity and therefore, they must assume serial correlation in the exogenous technology shock process in order to match data on total factor productivity. However, the models of this paper, Pries (1998), and Cooper and Johri (1999) all suggest that learning over time by agents should more generally

⁶See, for example, Campbell (1998), Portier (1995), Galí (1995), and Cooper and Haltiwanger (1993).

be understood as an inherent source of persistence in economic models. There are two channels for this. First, learning takes time and therefore responses to shocks become drawn out. Second, learning changes the state of an economic system by changing a stock (*e.g.* of good plants, of good job matches, of know-how) so that the economy experiencing the same shock in two different periods will respond differently in each.⁷

This paper also ties into a growing literature on the relevance of heterogeneity in explaining aggregate phenomenon. Equilibria in models with heterogeneous agents are often complicated to determine because agents must keep track of the distribution over types in order to form expectations over future equilibrium prices. The typical finding in several previous papers, including Caballero, Engel, and Haltiwanger (1997), Krusell and Smith (1998), Thomas (1998), and Veracierto (1998), is that micro-level heterogeneity has a small impact on aggregate dynamics. For example, Veracierto (1998) considers a model where individual plants must undertake irreversible investment projects that give rise to micro-level (S, s) policy rules. Though this modeling feature generates very different plant-level responses to aggregate shocks, the aggregate economy behaves similarly to a standard DSGE economy.⁸ In this model it is also true that the plant-level productivity shocks have no effect on aggregate productivity, given the existing distribution of good and bad plants. However, since plants learn from these shocks over time and make exit decisions based on their evolving beliefs, aggregate productivity is not constant over time. In the basic model of this paper plant-level learning is the only source of dynamics for the model's aggregates and therefore, heterogeneity is important for the model's aggregate behavior.

The paper proceeds as follows. Section 2 presents empirical evidence from the U.S. economy on productivity variation and discusses the various ways in which this evidence has been incorporated into macro models of business cycles. Section 3 describes the model without capital, placing emphasis on describing the roles played by plant heterogeneity and learning. Section 4 presents selected simulation results from the basic model without capital highlighting the endogenous persistence generated by the learning mechanism in response to *i.i.d.* shocks. Section 5.1 adds capital to the model and provides an extended set of simulation results. Further empirical evidence in support of the model is also provided using the Census Bureau's County Business Patterns dataset and data in the Longitudinal Research Database. Section 6 provides three extensions of the basic model. Section 7 concludes.

⁷See also Jovanovic and Nyarko (1995) for more applications of learning models.

⁸Thomas (1998) considers a similar model with plant-level lumpy investment decisions and finds similar results. Krusell and Smith (1998) find that heterogeneity in wealth among households has little effect on the aggregate behavior of a DSGE model. Caballero, Engel, and Haltiwanger (1997) document the relevance of idiosyncratic, plant-level shocks in explaining plant-level employment dynamics and the near irrelevance of these micro shocks in explaining aggregate employment dynamics. Also see Abadir and Talmain (1998) on the role heterogeneity plays in increasing persistence in a time-series setting.

2 Productivity Data and Business Cycle Models

Figure 1 plots quarterly data (1955:3-1984:1) on total factor productivity (TFP) and private sector, non-farm output (GDP) detrended by the HP-filter.⁹ Vertical lines denote NBER recession dates. The plots suggest two broad features of TFP movements. First, TFP and GDP move together. The correlation between the series is .93 with a standard deviation of .10. Recessions are periods of below trend productivity and output while booms are periods of above trend productivity and output. Second, TFP is highly persistent. The first autocorrelation of TFP is .74 with a standard deviation of .10. Movements of TFP above trend tend to be followed by further movements of TFP above trend.¹⁰

While empirical facts such as these are widely known, the ways in which they are incorporated into existing macroeconomic models of business cycles vary considerably. Broadly, the literature on business cycles and productivity variation can be divided into two categories: models in which exogenous productivity fluctuations gives rise to output fluctuations and models in which output fluctuations from some other exogenous source such as demand shocks give rise to productivity fluctuations. Dividing existing models of business cycles into these two broad camps helps simplify the discussion, although there exist examples of models that resist easy classification along these lines.¹¹

The standard real business cycle (RBC) model¹² is a member of the first category. The story is a familiar one. Assuming constant returns to scale in the production of a single homogenous good, the economy can be described by a representative plant choosing variable inputs capital, k , and labor, n , from competitive factor markets with rental rates r and wages w so as to maximize profits subject to exogenous Hicks neutral (multiplicative) stochastic productivity variation. Letting $F(n_t, k_t)$ denote the production function and A_t denote the level of total factor productivity in period t , the representative plant's problem can be written as

$$\max_{n_t, k_t} A_t F(n_t, k_t) - r_t k_t - w_t n_t \quad (1)$$

⁹Data sources are as follows: output is real non-farm private sector output from Citibase; TFP is calculated data on capital inputs, labor inputs, and the factor payments of these as a share of output according to the method of Solow (1957).

¹⁰The correlation of TFP with output is not sensitive to the use of HP-filtering; the correlation between TFP growth and output growth is .95. TFP growth is not persistent, however; the first autocorrelation of TFP growth is .10 with a standard deviation of .10.

¹¹For example, some business cycle models use more than one type of shock (*e.g.* a productivity shock and a demand side shock as in Christiano and Eichenbaum (1992)). Sunspot models made popular by the work of Farmer and Guo (1994) and Benhabib and Farmer (1997) are also difficult to classify along these lines though they typically do not focus on explaining productivity variation. Finally, models of cyclical capacity utilization help explain the procyclicality of productivity by appealing to mismeasurement in the calculation of the TFP series presented in figure 1.

¹²While there are many variants of RBC models, the prototypical example of this class is most often associated with King, Plosser, and Rebelo (1987).

Table 1: Output Persistence and Productivity Persistence in the KPR Model

ρ_A	ρ_{yA}	σ_y	ρ_y
.96	.98	2.58	.93
.75	.98	3.64	.71
.5	.99	3.77	.42
.25	1.00	3.89	.15
0	1.00	4.01	-.11

Note: Table reports simulated moments from the KPR (1987) model. Parameters except for ρ_A and σ^2 , the variance of the productivity shock, are set at the published values. The first row reports results for the published value of $\rho_A = .96$. Subsequent rows lower the value of ρ_A and set σ^2 so as to keep the population variance of productivity equal to the published value. The columns ρ_{yA} , σ_y and ρ_y report the correlation of output and productivity, the standard deviation of output (in percent) and the first autocorrelation of output from the model. Source: author's calculations

Given empirical evidence such as figure 1, the data generating process for A is assumed to be an autoregression, frequently of order one, with a root near or at unity. In place of Solow (1957) residual accounting, the RBC literature favors estimating the autocorrelation coefficient and variance of A with Generalized Method of Moments by imposing an AR(1) structure to detrended data on output and factor inputs. Typical estimates from detrended quarterly data on aggregate GDP, aggregate capital stocks, and aggregate labor hours return an autoregressive root ρ_A around 0.96 with a shock standard error around 0.008.¹³

The unifying feature of these real business cycle models is that output fluctuations, as well as fluctuations in labor hours, consumption, and net investment are consistent with maximizing behaviour on the part of consumers and plants in the face of this exogenous productivity variation. With even modest levels of intertemporal substitution in the specification of preferences of consumers, the optimal response of the economy following a positive movement in productivity is to increase the level of the capital stock. The reason for this is that productivity movements are assumed to be so persistent. If productivity is $x\%$ above its mean in its current period then it is expected to be $\rho_A x\%$ above its mean in the next period, $\rho_A^2 x\%$ above its mean in the following period, and so on. With ρ_A near one this amounts to a substantial incentive to expand capacity today by increasing the capital stock (and forgoing consumption today) to reap the benefits of higher expected productivity in the future. However, this story, and the business cycles that the model is capable of generating, are *critically* dependent on a highly persistent process for aggregate TFP. Consumers would not be willing to forgo consumption in the current period if higher productivity today did not signal higher expected productivity in the future. Consequently, the model-generated time series look much like the exogenous process of productivity being fed in to drive the model. Table 1 provides simulated moments on the standard deviation and first autocorrelation of output

¹³A number of authors find similar numbers. See, for example, Hansen (1985) and King, Plosser and Rebelo (1987).

generated from the King, Plosser and Rebelo (1987) model under alternative values for ρ_A .¹⁴ As the assumed persistence of the productivity process declines (holding its variance constant), output becomes more volatile and even more correlated with productivity movements but the persistence of aggregate output declines. At $\rho_A = 0$, *i.i.d.* productivity movements generate *i.i.d.* movements in output. The model provides no persistence beyond the persistence that is contributed by the exogenous process for aggregate productivity. Without an explanation of the driving forces behind productivity variation, the RBC model seems mute on the very question it tries to answer: What causes business cycles?¹⁵

The sometimes angry commentary¹⁶ regarding the RBC paradigm stems from its conclusion that policy has no role to play in attenuating the cycle. Policy measures aimed at boosting output during recessions are not welfare improving under the RBC paradigm since output is *optimally* low, given the low state of aggregate productivity. Since the RBC theory depends so critically on persistent exogenous productivity variation, these authors argue that, without candidate causes for Solow-type productivity shocks, the real business cycle model is a non-starter for truly *explaining* business cycles. Exhaustive attempts by Cochrane (1994) at finding candidates for productivity shocks have turned up little convincing evidence. Furthermore, the RBC model can not explain why productivity should move with output in response to demand-side shocks that are unrelated to changes in the technology feasibility frontier at either the aggregate or sectoral level.

Basu (1996), Basu and Fernald (1997, 1998) and Basu and Kimball (1994) investigate the reverse causality: output fluctuations give rise to productivity fluctuations. These authors posit models that generate observed patterns in measured aggregate productivity from other exogenous shocks to output such as demand or sectoral shocks. The mechanism employed in these works is one of resource reallocation over downturns in the business cycle from industries with high degrees of returns to scale to industries with low degrees of returns to scale. Measured aggregate productivity declines as a result, though the aggregate technology feasibility frontier is unchanged. While these models provide plausible mechanisms to link output movements with productivity movements, they do not explain the persistence in either output or productivity, *per se*. As with the RBC model, an exogenous driving process in these models which is itself inherently persistent causes persistent movements in output which, in turn, give rise to persistent movements in productivity. However, the approach of these authors is insightful and related to the approach taken here in that it views aggregate productivity variation as the outcome, in part, of compositional changes in production and thus partly endogenous to the economic environment.

¹⁴Model simulations were carried out under the parameterizations reported in the authors' published version for all parameters except ρ_A and σ^2 . See notes for table 1.

¹⁵King, Plosser, and Rebelo (1987) and Cogley and Nason (1995) level a slightly different criticism on the RBC model. They point to the fact that the RBC model is unable to: 1) generate output *growth rates* that have a positive first-order autocorrelation in response to productivity shocks and 2) generate output and hours impulse responses following demand shocks that are hump shaped.

¹⁶See, for example, Summers (1986) and Mankiw (1989)

3 The Basic Model without Capital

This section presents a DSGE model of aggregate productivity variation. The model in this section abstracts from decisions over other, storable, inputs such as physical capital for the sake of simplicity, so as to highlight the role played by the evolution of the distribution of plant types in generating the model's dynamics. Section 5.1 considers an extension of the model with capital in the production function. The source of productivity variation is variation in the distribution of plants in the economy caused by plant entry, learning, and exit. In the basic model the centralized and decentralized solutions are equivalent and all resources are allocated efficiently. Furthermore, the aggregate representation of this economy is observationally equivalent to the standard RBC model. Therefore, this model can be viewed as a justification for specifying exogenous persistent productivity variation in a business cycle model that is focusing on other features of macroeconomic fluctuations in response to movements in aggregate productivity.¹⁷

The model is described in three parts: the behavior of the consumer/shareholder, the behavior of the individual productive unit or plant, and finally, the evolutionary path of the distribution of plants and the behavior of the aggregate economy. The equivalence between the heterogeneous plant model and a standard one-sector model is shown in theorem 1. A closed-form representation for aggregate productivity is derived which highlights the factors that affect persistence in the model. This section ends with the definition of a competitive equilibrium and steady-state for the model economy. Conditions for the existence and uniqueness of these are established in theorems 2 and 3.

3.1 Consumption and labor supply

Time is discrete. The economy is populated by a continuum (with mass one) of identical, infinitely lived individuals. Each period, the representative individual derives utility from consuming a single good in the amount C and from leisure, L . The individual's time endowment is fixed at T , and all time is spent either in leisure or working N hours, hence $L = T - N$. The lifetime utility of the representative individual is given by

$$\sum_{t=0}^{\infty} \beta^t U(C_t, T - N_t) \quad , \quad 0 < \beta < 1 \quad . \quad (2)$$

The period utility function, $U(.,.)$ satisfies the usual properties: increasing, quasi-concave, smooth as needed.

¹⁷An extension of the basic model in section 6.2 permits inefficient allocation of resources in a decentralized equilibrium. Inefficiencies arise from informational constraints: households can not observe plant-level productivity data prior to deciding labor supply but must instead base their labor supply decisions on *beliefs* about plant productivity which may be overly optimistic or pessimistic compared with the actual distribution of plant types. The resulting equilibrium will entail allocations that are different from the unconstrained case studied in sections 3 and 5.1.

Each period, the representative individual receives labor income, $w_t N_t$, and the profits from all the plants in the economy, Π_t . Therefore, (2) is maximized subject to the series of budget constraints

$$C_t \leq w_t N_t + \Pi_t \quad , \quad t = 0, 1, 2, \dots \quad . \quad (3)$$

Since the individual has no means of accumulating wealth, and given the assumptions on U , maximization of (2) subject to (3) is accomplished by maximizing utility each period through the choice of labor supply. Since the individual takes wages and corporate profits as given, this results in the familiar intratemporal optimality condition, written without time subscripts,

$$-\frac{U_2(wN + \Pi, T - N)}{U_1(wN + \Pi, T - N)} = w \quad , \quad (4)$$

where U_1 and U_2 denotes the derivative of U with respect to its first and second arguments, respectively. Equation 4 can be solved implicitly for aggregate labor supply $N^s(w)$.

3.2 Production

Production of the single good occurs in units called plants. A plant consists of a production technology taking labor, n , as its only variable input, operating in competitive labor and final goods markets. Output of a single plant in a period is given by

$$y_t = a_t n_t^\alpha \quad (5)$$

where $0 < \alpha < 1$ and a_t is an idiosyncratic productivity shock. The parameter α controls the *span of control* of a single plant, familiar from the work of Lucas (1978). The assumption of decreasing returns to scale in variable factors is based on the existence of unmodeled fixed factors such as managerial ability or land that give rise to capacity constraints.¹⁸ Since plants exhibit decreasing returns to scale, the social optimum will involve production by all active plant in the economy rather than production by only the most efficient plant, as would be the case if $\alpha = 1$. The output price is normalized to unity in each period.

Realizations of the idiosyncratic a_t take values in the set $\mathcal{A} \equiv \{A_1, A_2, \dots, A_J\}$ with $A_1 < A_2 < \dots < A_J$. It is assumed that plant productivity is independent across plants though it may be serially correlated across time at a given plant. More specifically, plant productivities are independent first-order Markov processes with transition probabilities specified below.

¹⁸Another formulation of decreasing returns to scale that is used frequently in the literature on reallocations (see Caballero and Hammour (1994, 1996)) is an explicit capacity constraint assumption that a plant can hire only one worker or $y_t = a_t$. Results in this paper would hold with this formulation as well although the aggregate representation of the model would not match that of the standard neoclassical growth model.

Each period, after realizing productivity level a , plants maximize profits within the period, taking wages paid to labor, w_t , as given as in

$$\max_{n_t} a n_t^\alpha - w_t n_t \quad (6)$$

Therefore, a plant with productivity a in the period will demand $n^d(a; w_t)$ labor units given by

$$n^d(a; w_t) = w_t^{\frac{1}{\alpha-1}} [\alpha a]^{\frac{1}{1-\alpha}}. \quad (7)$$

Profits accruing to a plant experiencing productivity a are always positive and amount to

$$\Pi(a; w_t) = \frac{1-\alpha}{\alpha} w_t^{\frac{\alpha}{\alpha-1}} [a]^{\frac{1}{1-\alpha}} \quad (8)$$

3.3 The evolution of plants' beliefs and aggregation

3.3.1 Plant types and learning

There are two types of plants in the economy, indexed by i : high average productivity or *good* plants ($i = 2$) and low average productivity or *bad* plants ($i = 1$). Let $\pi_i(a, a')$ denote the probability with which a plant of type i receives productivity $a_{t+1} = a'$ conditional on $a_t = a$.¹⁹ Good plants have a greater conditional probability of receiving higher values of a than bad plants. Formally, the conditional distribution of productivity in good plants stochastically dominates that of bad plants or $\sum_{j=1}^k \pi_1(a, A_j) \geq \sum_{j=1}^k \pi_2(a, A_j)$ for $k = 1, \dots, J$ and $a \in \mathcal{A}$.

As in Jovanovic (1982), plants enter the economy uncertain about their type and learn through producing which type they are likely to be.²⁰ Here, this is implemented as Bayesian learning.²¹ A plant starts a period with a prior probability q that it is good and an experience a in the previous period. A plant's values for $z \equiv (q, a)$ will completely characterize its state at the beginning of the period when it decides whether to remain in operation or to exit.²² If it decides to operate, it costlessly draws a productivity level a' from the distribution associated with its unknown type, demands labor according to (7), and updates its prior according to the Bayesian posterior

¹⁹The transition probabilities satisfy $\sum_{j=1}^J \pi_i(a, A_j) = 1$ and parameterize the degree of persistence of plant-level productivity. When $\pi_i(a, A_j) = \pi_i(A_j) \forall j = 1, \dots, J$ a plant's productivity is not dependent on its past productivity.

²⁰More precisely, plant managers are uncertain about their plants' types and learn through producing about this characteristic. However, the personification of plants simplifies the description of the model.

²¹Plants in this model passively learn about their abilities simply by operating, in contrast to the plants in Ericson and Pakes (1995, 1998) which actively learn about their environment. Other models of passive learning include Hopenhayn (1992a, 1992b), Atkeson and Kehoe (1993), and Jovanovic (1982). Also see Bahk and Gort (1993).

²²For completeness, $z \in \mathcal{Z} \equiv (0, 1) \times \mathcal{A}$.

function

$$b(z, a') = \frac{q\pi_2(a, a')}{q\pi_2(a, a') + (1 - q)\pi_1(a, a')} \quad . \quad (9)$$

Therefore, its state at the end of the current period and at the beginning of the next is $z' \equiv (b(z, a'), a')$.

Let $\lambda_t^1(z)$ denote the measure of low productivity plants at state z and let $\lambda_t^2(z)$ denote the measure of high productivity plants at state z *operating* during period t for $z \in \mathcal{Z}$. Furthermore, let $L_t(a) \equiv \int_0^1 \lambda_t^1(q, a) dq$ and $H_t(a) \equiv \int_0^1 \lambda_t^2(q, a) dq$ denote the mass of low productivity and high productivity plants, respectively, operating in period t with prior realization of productivity a . It is important to stress that the mass $L_t(a) + H_t(a)$ for $a \in \mathcal{A}$ constitutes plants that chose to operate rather than exit in period t .

Two comments are appropriate here. First, the measures λ_t^1 and λ_t^2 and the quantities $L_t(a)$ and $H_t(a)$ are not static measures, but evolve from period to period, as described below. In fact, given the absence of storable factors or wealth in the basic model of this section, the evolution of these measures provides the only source of quantity dynamics in the model. Second, since plants choose labor demand *after* observing their draw of productivity in a period, beliefs are not material for the behavior of aggregate quantities. Labor is allocated efficiently, given the realization of productivities across plants. Beliefs matter only when describing the exit decision of plants. In an alternate formulation of the consumer's problem where labor supply is based on plants reported expectations of their productivity, beliefs affect aggregate quantities and may lead to a misallocation of resources (labor). This alternate formulation will be considered in section 6.2.

3.3.2 Aggregation

Since there exists a continuum of plants, there is no aggregate productivity uncertainty, even though productivity is stochastic at the plant level. Illuminating the deterministic nature of aggregate productivity requires aggregating from plant-level decision rules to aggregate quantities.

Labor demand in period t by type i plants at z is given by

$$n^{d(i)}(z; w_t) = \lambda_t^i(z) \alpha^{\frac{1}{1-\alpha}} w_t^{\frac{1}{\alpha-1}} \tilde{A}_i(a) \quad , \quad (10)$$

where $\tilde{A}_i(a) \equiv \int_{\mathcal{A}} \pi_i(a, a') (a')^{\frac{1}{1-\alpha}} da'$. Aggregate labor demand is then

$$N^d(w_t) = \int_{\mathcal{Z}} \left[n^{d(1)}(z; w_t) + n^{d(2)}(z; w_t) \right] dz = \left[\frac{\alpha A_t}{w_t} \right]^{\frac{1}{1-\alpha}} \quad , \quad (11)$$

where A_t denotes aggregate productivity and is given by

$$A_t = \left[\int_{\mathcal{A}} (L_t(a)\tilde{A}_1(a) + H_t(a)\tilde{A}_2(a))da \right]^{1-\alpha} \quad (12)$$

Analagous aggregation results in aggregate profits of

$$\Pi(w_t) = (1 - \alpha)A_t \left[N^d(w_t) \right]^\alpha, \quad (13)$$

and aggregate output of

$$Y(w_t) = A_t \left[N^d(w_t) \right]^\alpha. \quad (14)$$

Goods market clearing requires that

$$C_t = w_t N_t + \Pi_t = Y_t \quad (15)$$

Finally, the labor market clearing quantity, N_t , can be determined by solving (11) for w and equating this to (4).

Viewing (14) as the aggregate production function for the economy it is clear that aggregate productivity at date t does not depend on the entire distribution of plants over beliefs, but is simply a function of the mass of high and low productivity plants with prior realizations a operating in period t given by (12). This result permits the following theorem.

Theorem 1 *The heterogeneous plant economy described by the masses $L_t(a)$ and $H_t(a)$ and their underlying distributions $\lambda_t^1(z)$ and $\lambda_t^2(z)$, with plants individually solving (6), yields aggregate allocations of labor and output equivalent to those of the following social planner's problem for an economy comprised of a representative plant and consumer*

$$\begin{aligned} (\text{SP1}) \quad & \max_{C_t, N_t} U(C_t, T - N_t) \\ \text{s.t.} \quad & C_t \leq Y_t \\ & Y_t = A_t N_t^\alpha \\ & A_t = \left[\int_{\mathcal{A}} (L_t(a)\tilde{A}_1(a) + H_t(a)\tilde{A}_2(a))da \right]^{1-\alpha}. \end{aligned}$$

Proof: see Appendix A

Theorem 1 establishes the relationship between the decentralized economy and a centralized social planner economy where the social planner knows the mass of high and low productivity plants operating in the economy. Theorem 1 is useful because it establishes a dichotomy between the solution for aggregate equilibrium quantities on the one hand and plant behavior and learning which affects the distribution of plants in the economy on the other. Given A_t , aggregate allocations are trivial to compute from (SP1). This permits greater attention to be focused on the evolution

of the distributions of plant types in the economy which determine the movements in A_t .

3.3.3 Entry and exit

Since time is discrete, something must be said about the sequence of events within a period, specifically entry, exit, and production. Obviously, there is little empirical evidence to suggest an ordering of these events. In the real economy entry, exit and production are continuous processes. Therefore, to proceed it is assumed that entry and exit decisions are made at the beginning of the period, prior to the realization of productivity, and that production and the updating of plant priors occurs at the end of the period. From (12) it is clear that exit by either low or high productivity plants decreases aggregate productivity. However, exit by high productivity plants decreases aggregate productivity by more than an equal mass of exit by low productivity plants since $\tilde{A}_2(a) \geq \tilde{A}_1(a) \forall a \in \mathcal{A}$. Finally, it is assumed that all plants that exist at the beginning of a period face an *exogenous* probability of failure prior to producing of δ with $0 \leq \delta < 1$.

In the model, the exit decision by a plant involves weighing the value of remaining in production during the current period, and deciding optimally whether to exit in the subsequent period, against scrapping the plant (exiting) and receive a value ψ^s .²³ Formally, the exit choice can be represented as the solution to the following recursive problem for a plant with state z taking aggregate wage w_t as given

$$V(z; w_t) = \max \left[\psi^s, \tilde{V}(z; w_t) \right] \quad (16)$$

where

$$\tilde{V}(z; w_t) = \Pi^e(z; w_t) + \beta(1 - \delta) \int_{\mathcal{A}} \pi(z, a') V(z'; w_{t+1}^e) da' \quad (17)$$

In (16), $\Pi^e(z; w) \equiv \int_{\mathcal{A}} \pi(z, a') \Pi(a'; w) da'$ is the expected profit from operating the plant during the current period and $\pi(z, a') \equiv (q\pi_2(a, a') + (1 - q)\pi_1(a, a'))$ is the perceived probability of receiving productivity a' by a plant with state z in the current period which would give rise to profits $\Pi(a'; w)$. This would cause updating of the plant's state to $z' = (b(z, a'), a')$ in the subsequent period at which time the wage is expected to be w_{t+1}^e .

A *marginal* plant is one for whom $\tilde{V}(z; w_t) = \psi^s$. In general, there is a measure of plants at z^* defined by $\tilde{V}(z^*; w_t) = \psi^s$. Since $\Pi^e(z; w)$ and $b(z, a')$ are monotonically increasing in q for fixed (a, a') , $V(z; w)$ will be monotonically increasing in q ensuring that a unique $q^*(a) \in [0, 1)$ exists for each $a \in \mathcal{A}$.²⁴ Therefore, $q^*(a)$ serves as a convenient means of summarizing the exit/no-exit

²³The models of plant exit of Hopenhayn (1992a,b) use a similar structure to the one presented here.

²⁴An interior solution ($q^*(a) \in (0, 1)$) is not guaranteed. $q^*(a) = 0$ implies plants with any prior and previous productivity a are viable. $q^*(a) = 1$ implies all plants with previous productivity a exit. The latter can be ruled out.

decisions of all plants that exist at the begining of a period that previously experienced productivity a . All plants with z such that $q < q^*(a)$ will chose, by (16), to exit. All plants with z satisfying $q \geq q^*(a)$ will chose to remain in production for at least one more period.²⁵ Note that since the cutoff priors influence the mass of plants that operate during the period, aggregate productivity and the aggregate wage will depend on them. The solution for $q^*(a)$ must take this into account though the individual plant takes w_t as given.

Entry of new plants occurs in period t with mass $M_t \geq 0$. Entering cohorts are comprised of a fraction h high productivity plants and a fraction $1 - h$ low productivity plants. Entrants must pay a cost $\psi^b > \psi^s$ to acquire a plant. It is assumed, in line with the evidence presented in Bartlesman and Dhrymes (1993) and Bailey, Hulten, and Campbell (1992) that new entrants have priors which are uniformly distributed on $\mathcal{Z}$.²⁶ A potential entrant will enter if $\int_{\mathcal{Z}} V(z; w_t) dz \geq \psi^b$ implying that total entry mass M_t satisfies the zero expected profit condition

$$\int_{\mathcal{Z}} V(z; w_t) dz = \psi^b \quad (18)$$

The distance $\psi^b - \psi^s$ represents the degree to which resources inside a plant are specific to that plant. When $\psi^b - \psi^s$ is close to zero, plant specificity is low. An entrant is able to recoup a significant portion of the cost of establishing the plant upon exit because the value of the plant is not specific to the current management. When $\psi^b - \psi^s$ is high, plant specificity is high and little of the cost of establishing the plant is recouped upon exit. The equilibrium of the model will be characterized by higher gross entry as specificity declines and lower gross entry as specificity rises.

3.3.4 The evolution of the distributions λ^1 and λ^2

In order to complete the description of the model, all that remains is to describe how the distributions of plant types, λ_t^1 and λ_t^2 evolve from period to period. Plants must keep track of the movements in these distributions in order to forecast future prices when making exit decisions according to (16) and when making entry decisions according to (18). It will be convenient to refer to the measure of type $i = 1, 2$ plants at z at the *begining* of a period when entry and exit decisions are made by $\rho_t^i(z)$. Recall, plants that operated in period $t - 1$ with productivity a updated their priors based on (9). Then ρ_t^i is the distribution that results from the updating by plants described by λ_{t-1}^i .

See appendix A.

²⁵It is possible to implement other exit rules, such as rules that account for the uncertainty inherent in the continuation value in (16), and have all the subsequent results go through. Jovanovic (1982) considers such alternate exit rules.

²⁶This evidence is discussed further in section 5.2. All results go through with more burdensome notation if initial priors are distributed according to some other, known, distribution.

Another useful concept for the ensuing discussion is the inverse-Bayesian function

$$g(z, a_{-1}) = \frac{q\pi_1(a_{-1}, a)}{q\pi_1(a_{-1}, a) + (1-q)\pi_2(a_{-1}, a)} \quad . \quad (19)$$

The interpretation of (19) is that a plant which experienced productivity a_{-1} in period $t-1$ and holds the prior q at the beginning of period t must have held prior $g(z, a_{-1})$ while operating during period $t-1$. In other words, (19) is the inverse of (9).

With (19) it is possible to write the evolution process from λ_{t-1}^i to ρ_t^i as

$$\rho_t^i(z) = \int_{\mathcal{A}} \pi_i(a_{-1}, a) \frac{\pi_i(a_{-1})}{\pi_i(a)} \lambda_{t-1}^i(g(z, a_{-1})) da_{-1} \quad i = 1, 2 \quad (20)$$

The integrand in (20) “counts” the type i plants that arrived at state z in period t by way of prior $g(z, a_{-1})$ and experience a_{-1} in period $t-1$. $\pi_i(a)$ denotes the unconditional probability of drawing productivity level a for a type i plant.

From $\rho_t^1(z)$ and $\rho_t^2(z)$ arriving at $\lambda_t^1(z)$ and $\lambda_t^2(z)$, the measures of plants operating during period t , requires removing plants that exit and adding plants that enter. Exogenous exit reduces $\rho_t^i(z)$ everywhere proportionally by $(1-\delta)$. Entry with total mass hM_t is spread uniformly over $(0, 1) \times \mathcal{A}$ in ρ_t^2 and entry with total mass $(1-h)M_t$ is spread uniformly over $(0, 1) \times \mathcal{A}$ in λ_t^1 . Simultaneously with entry, endogenous exit involves finding a $q_t^*(a)$ satisfying $\tilde{V}((q_t^*(a), a), w_t) = \psi_t^s$ and removing all plants in both distributions with priors $q < q_t^*(a)$. Summarizing this transformation from $\rho_t^i(z)$ to $\lambda_t^i(z)$:

$$\begin{aligned} \lambda_t^1(z) &= (1-\delta)\rho_t^1(z) + (1-h)M_t & q \in [q_t^*(a), 1) \\ \lambda_t^2(z) &= (1-\delta)\rho_t^2(z) + hM_t & q \in [q_t^*(a), 1) \\ \lambda_t^i(z) &= 0 & q \in (0, q_t^*(a)) \quad i = 1, 2. \end{aligned} \quad (21)$$

In general it is difficult to characterize the evolution of the entire distributions $\lambda^1(q)$ and $\lambda^2(q)$ described by (20) and (21). However, from (12) it should be clear that of primary interest in the basic model is the evolution of the masses of good and bad plants, H_t and L_t . Integrating (21) over $\mathcal{Z}$ results in the following simple vector autoregressive representation for L_t and H_t

$$\begin{bmatrix} L_t \\ H_t \end{bmatrix} = (1-\delta) \begin{bmatrix} L_{t-1} \\ H_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^L \\ \epsilon_t^H \end{bmatrix} \quad (22)$$

where ϵ_t^L denotes the net mass of transitioning low productivity plants and ϵ_t^H denotes the net mass of transitioning high productivity plants. These quantities are functions of the distributions ρ_t^1 and

ρ_t^2 and the mass of entrants M_t :

$$\begin{aligned}\epsilon_t^L &\equiv (1-h)M_t \sum_{j=1}^J (1-q_t^*(A_j)) - \sum_{j=1}^J \int_0^{q_t^*(A_j)} \rho_t^1(z) dq \\ \epsilon_t^H &\equiv hM_t \sum_{j=1}^J (1-q_t^*(A_j)) - \sum_{j=1}^J \int_0^{q_t^*(A_j)} \rho_t^2(z) dq\end{aligned}\tag{23}$$

The first terms represent the endogenous entry of plants of each type, uniformly distributed on $\mathcal{Z}$ that choose not to exit in their period of entry. The second terms represent the mass of plants of each type endogenously choosing to exit. Letting $N_t^f = (L_t \ H_t)'$, $\epsilon_t = (\epsilon_t^L \ \epsilon_t^H)'$, (22) can be written in matrix form as $N_t^f = (1-\delta)N_{t-1}^f + \epsilon_t$.

One can interpret ϵ_t as summarizing the extent of reallocation activity in the economy. Recall that aggregate productivity is a simple function of L_t and H_t . By (22), these are autoregressive processes with reallocation “shocks” ϵ_t^L and ϵ_t^H . Therefore, aggregate productivity is directly influenced by reallocation, by the birth and death of plants. However, the shocks ϵ_t are endogenous, being governed by the mechanics of (20) and (21). Since aggregate productivity influences aggregate variables like the equilibrium wage, it will affect the exit decision of plants via (16). So it is also true that aggregate productivity affects reallocation.

3.4 Persistent productivity variation

The exposition up to this point has allowed for *exogenous* persistent productivity variation at the plant level by specifying a Markov transition matrix with typical element $\pi_i(a, a')$. For the remainder of the paper, it is assumed that plant productivity draws are *i.i.d.* to make the point as clearly as possible that persistence at the aggregate level is an *endogenous* feature of the model stemming from plant learning rather than from the assumption of an exogenous persistent driving process at the micro level. This additional assumption imposes the restriction $\pi_i(a, A_j) = \pi_i(A_j) \equiv \pi_{ij}$ for all $a \in \mathcal{A}$. Consequently, a sufficient statistic for denoting the state of a plant is the plant’s prior q . Furthermore, the dependence of the distributions λ_t^i and ρ_t^i , and the cutoff priors q_t^* on $a \in \mathcal{A}$ is dropped. In this case (12) simplifies to

$$A_t = \left[\tilde{A}_2 H_t + \tilde{A}_1 L_t \right]^{1-\alpha}, \tag{24}$$

where $\tilde{A}_i \equiv \sum_{j=1}^J \pi_{ij}(A_j)^{\frac{1}{1-\alpha}}$, $L_t \equiv \int_0^1 \lambda_t^1(q) dq$, and $H_t \equiv \int_0^1 \lambda_t^2(q) dq$. Bayes’ rule simplifies to $b_j(q) = \frac{q \pi_{2j}}{q \pi_{2j} + (1-q) \pi_{1j}}$ for a plant experiencing productivity A_j holding prior q . Finally, the expressions for

ϵ_t^L and ϵ_t^H simplify to

$$\begin{aligned}\epsilon_t^L &\equiv (1-h)(1-q_t^*)M_t - \int_0^{q_t^*} \rho_t^1(q) dq \\ \epsilon_t^H &\equiv h(1-q_t^*)M_t - \int_0^{q_t^*} \rho_t^2(q) dq\end{aligned}\tag{25}$$

Equations (24) and (25) illuminate the key parameters responsible for determining the time series properties of aggregate productivity in the model in response to “shocks” ϵ_t^L and ϵ_t^H . To a first-order approximation²⁷, the log deviation from mean of A_t is a linear function of the the log deviation from mean of N_t^f

$$\hat{A}_t = (1-\alpha) \left[\tilde{A}_2 \hat{H}_t + \tilde{A}_1 \hat{L}_t \right] = (1-\alpha) \tilde{a}' \hat{N}_t^f\tag{26}$$

where $\hat{X}$ denotes the deviation of the variable X from its mean and $\tilde{a} = (\tilde{A}_1 \ \tilde{A}_2)'$. Suppose, for the time being, that $\hat{\epsilon}_t$ were *i.i.d.*. Then, since $\hat{N}_t^f = (1-\delta)\hat{N}_{t-1}^f + b\hat{\epsilon}_t$, where b is a diagonal matrix with $(\frac{1}{L_{ss}} \ \frac{1}{H_{ss}})'$ on its diagonal, as δ approaches zero, aggregate productivity approaches a unit root process and persistence increases. The intuition for this result is simple. Lowering δ makes deviations of N^f from its steady-state level²⁸ persist longer since plants are less likely to leave the economy exogenously. If, for example, a temporary shock results in a higher percentage of low productivity plants compared with the steady-state level, then a low δ implies that the distribution of plants returns to steady-state through learning or endogenous exit, rather than through exogenous death of the low productivity plants. The speed of learning is controlled by the values of π_{ij} which enter through ϵ_t .

In general, $\hat{\epsilon}_t$ may not be *i.i.d.* and therefore, the covariance properties of $\hat{\epsilon}_t$ will also influence the covariance properties of $\hat{A}_t$. Recursive substitution using (20)-(21) and (25) described in appendix C achieves the following representation for ϵ_t

$$\epsilon_t = (1-q_t^*)M_t v_h - \sum_{k=1}^{\infty} (1-\delta)^k \theta_t^k v_h M_{t-k} \quad .\tag{27}$$

The θ_t^k in (27) are diagonal matrices. The diagonals of these matrices contain the probability of exiting in period t of type 1 and type 2 plants of age k , conditional on not having exited at ages $k-1, k-2, \dots, 1$, or the age- k hazard rate for the two types of plants in period t . The representation (27) implies that ϵ_t follows an infinite order process that depends on variables determined in the model's equilibrium, namely $\{q_{t-k}^*, M_{t-k}\}_{k=0}^{\infty}$.

Taking the model as a true representation of the actual economy leads to three stark implications. First, the VAR representation (22) reveals the origins of productivity variation in the

²⁷This assumes the mean value for $A = 1$, an arbitrary normalization of the economy.

²⁸The definition of steady-state and proof of its existence and uniqueness will be dealt with shortly.

economy: shocks ϵ_t . To the observer of aggregate variables the shocks ϵ_t appear exogenous. In other words, an econometrician with knowledge of (20) and (21) capable of tracking the evolution of the distribution of plant types would realize the endogenous nature of ϵ_t^L and ϵ_t^H . However, the observer of only aggregate variables cannot distinguish the effect on aggregate productivity of ϵ_t^L and ϵ_t^H from that of exogenous shocks from some unnamed source. If this were a correct representation of aggregate productivity, then it would explain why econometric efforts such as Cochrane (1994) to uncover the origins of cyclical variation from observation of aggregate time series are largely unsuccessful.

Second, the model implies that *any* exogenous shock to the economy that triggers a mass of exit of plants will affect aggregate productivity through (22). Uncorrelated, serially independent, random disturbances from a multitude of sources that result in plant exit will get turned into persistent aggregate productivity variations via (22). Therefore, the identification of shocks that give rise to plant exit is perhaps of secondary importance to the study of productivity. The shocks themselves are not the source of *persistent* productivity variation. Rather, it is the dynamics of plant learning and the consequent evolution of the distribution of plant types that is chiefly responsible for persistent fluctuations in productivity.

Third, a further implication of this work is that if demand shocks affect plant entry and exit then they will affect aggregate productivity. This implies that the assumption of zero correlation between demand shifters and productivity employed by Hall (1988, 1990) and Basu and Fernald (1997) to estimate returns to scale at the industry and aggregate level is invalid. The likely direction of the bias is positive; an exogenous increase in demand is likely to result in net entry of plants implying a positive correlation between the demand shock and productivity. Using demand shift instruments that are positively correlated with productivity movements in the data would result in an overestimate of markups and/or returns to scale. The correlation between demand shocks and productivity movements also affects the interpretation of results from Blanchard and Quah (1989) decompositions. Galí (1996) and Kiley (1997) decompose macroeconomic fluctuations into a temporary and permanent component following the Blanchard and Quah methodology and then use the separation of demand shocks from productivity shocks to associate demand shocks with the temporary component and productivity shocks with the permanent component. If demand shocks cause fluctuations in productivity the temporary component of Blanchard-Quah decompositions may contain fluctuations in productivity, thus overstating the importance of demand shocks.

3.5 Competitive equilibrium and steady state

The next task at hand is defining a competitive equilibrium, given a sequence of values $\{\psi_t^s, \psi_t^b\}_{t=0}^\infty$ and a steady-state equilibrium for this economy, when $\psi_t^s = \psi^s$ and $\psi_t^b = \psi^b$ for all $t = 0, 1, \dots$. Theorem 2 establishes the existence and uniqueness of the competitive equilibrium for a given sequence $\{\psi_t^s, \psi_t^b\}_{t=0}^\infty$. Theorem 3 states that the model has either no steady-state or a unique

steady-state equilibrium.

Definition: A **competitive equilibrium** is comprised of sequences of measures $\{\lambda_t(q)\}_{t=0}^\infty$, sequences of prices $\{w_t, \psi_t^s, \psi_t^b\}_{t=0}^\infty$, cutoff priors $\{q_t^*\}_{t=0}^\infty$, allocation rules $\{n_t(q, w_t)\}_{t=0}^\infty$, and aggregate quantities $\{N_t, \Pi_t, Y_t, C_t, L_t, H_t\}_{t=0}^\infty$ such that:

1. given prices and aggregate sequences, plants maximize expected value (16) by exiting if their priors satisfy $q_t < q_t^*$ and using $n_t(w_t)$ if $q_t \geq q_t^*$
2. given prices and aggregate profits, consumers maximize utility (2) by choosing N_t^s, C_t subject to (3) and $N_t = T - L_t$;
3. prices are such that aggregate sequences and allocation rules satisfy labor market clearing and goods market clearing in every period;
4. the measures $\{\lambda_t^1(q), \lambda_t^2(q)\}_{t=0}^\infty$ satisfy (20), (21).

Theorem 2 *Given sequences $\{\psi_t^s, \psi_t^b\}_{t=0}^\infty$ and initial distributions $\lambda_0^1(q)$ and $\lambda_0^2(q)$, a competitive equilibrium exists and is unique.*

Proof: see Appendix A

The definition of competitive equilibria permits ψ_t^s and/or ψ_t^b to vary exogenously according to some known process.²⁹ If these do not vary but are fixed then the economy can be described by its *steady state equilibrium*, defined below.

Definition: A **steady state equilibrium** is a competitive equilibrium under $\psi_t^s = \psi^s$, and $\psi_t^b = \psi^b \forall t$. which must satisfy³⁰

$$\begin{aligned}
1) \quad & \epsilon_{ss} = \delta N_{ss}^f \\
2) \quad & \begin{pmatrix} \rho_{ss}^1(q) \\ \rho_{ss}^2(q) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^J \pi_{1j} \lambda_{ss}^1(g_j(q)) \\ \sum_{j=1}^J \pi_{2j} \lambda_{ss}^2(g_j(q)) \end{pmatrix} \\
3) \quad & \begin{pmatrix} \lambda_{ss}^1(q) \\ \lambda_{ss}^2(q) \end{pmatrix} = \begin{pmatrix} (1-\delta)\rho_{ss}^1(q) + (1-h)(1-q_{ss}^*)M_{ss} \\ (1-\delta)\rho_{ss}^2(q) + h(1-q_{ss}^*)M_{ss} \end{pmatrix} \quad q \in [q_{ss}^*, 1) \\
4) \quad & \tilde{V}(q_{ss}^*; w_{ss}) = \psi^s \text{ and } \int_0^1 V(q; w_{ss}) dq = \psi^b \quad .
\end{aligned}$$

Condition 1) simply ensures that the number of low and high productivity plants is constant in steady-state. Conditions 2) and 3) are necessary for the distributions of low and high productivity plants to be fixed in steady-state. Finally, condition 4) states that q_{ss}^* is the steady-state cutoff prior and that the zero-expected-profit condition is met for potential entrants.

²⁹Other types of shocks could be allowed as well so long as the definition of competitive equilibrium was changed accordingly. For example, shocks to the probabilities π_{ij} could be included in the model.

³⁰Vector integration denotes element-by-element integration.

3.5.1 Analysis of steady-state equilibria

Since the model is non-standard it is interesting to explore issues of existence and uniqueness of steady-states. The issues involved here do not concern the aggregate economy. Existence of a steady-state equilibrium in N and w for a given value of A_{ss} is guaranteed by Theorem 1 and standard results on existence of equilibrium in the one-sector neoclassical growth model. Rather, the relevant question is whether steady-state distributions exist which satisfy conditions 1) through 3) above for a q_{ss}^* satisfying condition 4) at the steady-state wage. It turns out that the existence of a steady-state is not guaranteed. This is shown in Theorem 3.

Theorem 3 *Let $\theta^k(q)$ denote the age- k hazard rate for a plant with prior q in steady-state and let $\Theta(q) \equiv \sum_{k=0}^{\infty} (1 - \delta)^k \theta^k(q)$ be the cumulative hazard rate for all age plants with prior q . Then, a steady-state equilibrium does not exist if*

$$\frac{1 - q_{ss}^*}{1 - \delta} < \Theta_{ii}(q_{ss}^*), \quad i = 1, 2 \quad .$$

A steady state equilibrium exists and is unique if

$$\frac{1 - q_{ss}^*}{1 - \delta} \geq \Theta_{ii}(q_{ss}^*), \quad i = 1, 2 \quad .$$

Proof: *see Appendix A*

The diagonal matrix $\theta^k(q_{ss}^*)$ in theorem 3 represents the age- k hazard rate in a steady-state for plants on the exit margin. Therefore, the steady-state mass of endogenous exiting plants is given by $\int_0^{q_{ss}^*} \rho_{ss}(q) dq = \Theta(q_{ss}^*) v_h M_{ss}$. This mass must be non-negative. Intuitively, when the parameter ψ^s is sufficiently close to ψ^b for a given rate of learning determined by the probabilities π_{ij} , the steady-state q_{ss}^* is so high that low productivity plants exit the economy faster than they are replaced, or $N_{1,ss}^f < 0$. Therefore, no steady-state equilibrium is possible in this case. Under such a parameterization the economy would continually fluctuate between a near zero mass of low productivity plants and a higher mass of low productivity plants, even with fixed ψ^s and ψ^b . Though it would be interesting to pursue this point further, the solution methodology described in the next section depends on the existence of a fixed steady-state. Therefore, it would not be possible to use the simulation techniques described below under such a parameterization.

4 Model Simulations

Two types of simulation results are reported in this section. In the first, impulse responses from a one-time shock to the measure of specificity $\Delta\psi \equiv (\psi^b - \psi^s)$ are presented for two different parameterizations of the learning process, fast and slow. Fast learning occurs when signals are

informative about plant type, when $|\pi_{1j} - \pi_{2j}|$ is large for many j . Slow learning occurs when signals are relatively uninformative, when π_{1j} is close to π_{2j} for many j . The second type of results report sample moments from time series generated by the model assuming *i.i.d.* shocks to $\Delta\psi$.³¹ Before turning to the simulation results, the solution methodology and parameterization of the model are described.

4.1 Solution Methodology

The dynamic path of the economy is calculated as follows. Given aggregate productivity determined by the exit and entry decisions of plants, aggregate quantities are determined by the necessary first order conditions to **(SP1)** in theorem 1. The determination of the equilibrium cut-off prior q_t^* in each period is carried out using (17). This requires knowledge of the expected continuation value $V_{t+1}^e(q) \equiv \sum_j \pi_j(q) V_j(b_j(q); w_{t+1}^e)$ for a plant with prior q in the current period who decides to remain in production. This continuation value is calculated by assuming that the economy will be at its steady-state point τ periods hence.³² The expected continuation value of a plant with prior q is known at steady-state. Each period of a simulation, the sequences of policies $\{q_{t+j}^*\}_{j=1}^\tau$ and value functions $\{V_{t+j}^e(q)\}_{j=1}^\tau$ are determined using standard policy and value function iteration.³³ The time to reach steady-state, τ is set so that simulation results are not different if it took $\tau + k$, $k = 1, 2, \dots$ periods to reach steady-state. For the simulation results presented below a value of $\tau = 20$ is used.

4.2 Parameters

The economy is parameterized with three possible levels of productivity ($J = 3$). The values π_{ij} under slow and fast learning are chosen to make the steady-state plant survival probabilities of the model match empirically observed survival probabilities reported in Dunne, Roberts and Samuelson (1989), for $\delta = .01$.³⁴ A comparison of the steady-state survival probabilities from the model and the empirical values from U.S. data are given in table 2. The values for A_j under slow and fast learning are chosen so that $\sum_j \pi_{2j} A_j / \sum_j \pi_{1j} A_j = 1.15$ implying that high productivity plants are

³¹In implementing these shocks it is assumed that ψ^b is held fixed and ψ^s varies to yield the desired movements in their difference.

³²Veracierto (1998) uses a similar method to calculate equilibrium outcomes.

³³The policies and value functions for τ periods are iterated over until convergence is achieved.

³⁴Higher values of δ reduce persistence in the model. Lower values of δ raise persistence in the model.

Table 2: Empirical and Parameterized Survival Probabilities

	S_5	S_{10}	S_{15}
U.S. Manufacturing Plants (1966-77)	.62	.47	.37
Model: Slow Learning	.65	.51	.42
Model: Fast Learning	.55	.42	.29

Note: Table reports empirical estimates and model analogs of S_j , the average plant survival probability to year j . Values for the U.S. manufacturing sector are calculated from results presented in Dunne, Roberts and Samuelson (1989), table VI by taking an unweighted average across the reported results for single and multi-plant plants and size classes. Values for the Model rows are calculated as $S_j = \prod_{i=1}^j (1 - \theta^i)$ where θ^i is the year i hazard probability at steady-state across all plants.

15 % more productive than low productivity plants, on average.³⁵

<i>Slow Learning</i>			<i>Fast Learning</i>		
$\pi_{11} = .4$	$\pi_{12} = .4$	$\pi_{13} = .2$	$\pi_{11} = .5$	$\pi_{12} = .3$	$\pi_{13} = .2$
$\pi_{21} = .25$	$\pi_{22} = .35$	$\pi_{23} = .4$	$\pi_{21} = .2$	$\pi_{22} = .3$	$\pi_{23} = .5$
$A_1 = .75$	$A_2 = 1$	$A_3 = 1.5$	$A_1 = .75$	$A_2 = 1$	$A_3 = 1.25$

Other parameter values used are $\delta = .01$, $h = .50$, $\beta = .993$, $\alpha = .85$. Under slow learning $\psi^s = 5.5$ while under fast learning $\psi^s = 16$ with ψ^b set, in each case, to normalize $A_{ss} = 1$. The choice of β conforms to standard practice in the RBC literature for associating the period length with a quarter year. The choice of α is guided by the average ratio of before tax corporate profits to GDP in U.S. data over the post-war period. The model indicates that this ratio should equal $(1 - \alpha)$.

The utility function functional form is set to $U(C, T - N) = \frac{C^{1-\sigma}(T-N)^{\gamma(1-\sigma)}}{1-\sigma}$, with $\sigma = .70$ and γ set so that $\frac{N}{T-N} = 1/2$ ($\gamma = 1.70$). In the simple model without capital, this utility function yields a vertical labor supply schedule. Consequently, output and wages move one-for-one with productivity.

Figure 2 plots impulse responses under the fast learning parameterization following a one-time negative shock to $\Delta\psi$ in period 2. Panels *a* and *b* show the behavior of the combined distribution $\lambda(q)$ in periods 1, 2, and 10, and the percentage deviation in productivity $\hat{A}$ from steady state following a 1% shock to $\Delta\psi$. Panels *c* and *d* do the same following a 5% shock to $\Delta\psi$. First, note that the percentage movement in A is small under either panel *b* or *d*. A 1% shock to $(\psi^b - \psi^s)$ results in less than a .2% increase in A while a 5% shock increases A by .4% on impact. The productivity response is positive since the degree of specificity is lowered by the shock resulting in above steady-state levels of plant entry. Second, note that the distribution $\lambda(q)$ returns quickly

³⁵Increasing the average productivity differences between good and bad plants increases the magnitude of the responses to shocks but has no effect on their persistence.

Table 3: Select Simulated Moments: Model without Capital

U.S. Data		Model Simulation: $\sigma_{\Delta\psi} = .01$		Model Simulation: $\sigma_{\Delta\psi} = .05$	
		Fast	Slow	Fast	Slow
<i>a. Log-Level Data</i>					
σ_a	—	0.14	0.29	1.26	0.29
σ_ϵ	—	0.24	0.15	1.75	0.15
ρ_a	—	0.37	0.65	0.72	0.85
ρ_ϵ	—	-0.28	-0.06	-0.11	-0.12
$\text{corr}(a, \epsilon)$		0.53	0.42	0.32	0.15
<i>b. HP-Filtered Data</i>					
σ_a	0.97	0.13	0.25	1.03	0.99
σ_ϵ	—	0.24	0.15	1.74	0.15
ρ_a	0.74	0.29	0.55	0.59	0.75
ρ_ϵ	—	-0.29	-0.08	-0.13	-0.15
$\text{corr}(a, \epsilon)$	—	0.57	0.48	0.41	0.49
<i>c. First Differenced Data</i>					
σ_a	0.82	0.15	0.24	0.94	0.96
σ_ϵ	—	0.39	1.74	1.74	1.74
ρ_a	0.10	-0.27	-0.08	-0.07	-0.13
ρ_ϵ	—	-0.57	-0.14	-0.14	-0.48
$\text{corr}(a, \epsilon)$	.41	0.78	0.89	0.98	0.96

Note: Table reports selected moments from U.S. data and simulations from the basic model without capital under fast and slow learning as well as two different size shock standard deviations. Correlations between total factor productivity and reallocation shocks, $\epsilon = (1 - h)\epsilon^L + h\epsilon^H$ (the weighted sum of reallocation shocks in the simulations) are denoted by $\text{corr}(a, \epsilon)$ and standard deviations and autocorrelations of total factor productivity are denoted σ_a and ρ_a , respectively. U.S. statistics are calculated from aggregate quarterly U.S. NIPA Data (1955:4-1984:1). Results are reported for data in log-levels, HP filtering, and first differences. Since U.S. data contain trends and are not stationary in levels, statistics on the level of these variables are not reported. See appendix B for further details on the calculation of productivity. A simulation consists of equilibrium responses of the model to *i.i.d.* Normal shocks to $(\psi^b - \psi_t^f)$ with standard deviation $\sigma_{\Delta\psi}$ for 140 periods.

back to its pre-shock shape. Consequently, productivity returns to steady-state in the period after the 1% shock and three periods after the 5% shock. Figure 3 plots results under the slow learning parameterization. Compared with figure 2, this figure shows that slow learning by plants draws out the responses of productivity following the shocks and increases the magnitude of the response on impulse. The 1% shock increases aggregate productivity by .4% while the 5% shock increases aggregate productivity by 3% on impact. In panel *d* of figure 3, productivity has returned only half-way to steady state eight periods after the shock.

Table 3 presents select moments from U.S. data and from model simulations. Since the basic model does not have capital or fluctuations in hours worked, the model is silent on fluctuations in investment or hours and productivity and output levels differ only by a multiplicative constant. Therefore, table 3 concentrates on the model's ability to generate persistent fluctuations

in productivity from *i.i.d.* shocks to $\Delta\psi$. Panel *a* reports statistics based on raw log-level data. Panels *b* and *c* report statistics for HP-filtered and first-differenced data.

Looking first at section *a.* of the table, it is evident that the model is capable of generating considerable persistence in productivity and output from purely *i.i.d.* shocks to $\Delta\psi$. Under $\sigma_{\Delta\psi} = .05$ and the slow learning parameterization, the first autocorrelation of productivity and output is .85. Persistence is increasing in the standard deviation of the shock and decreasing in the speed with which plants learn about their productivities. The autocorrelation of ϵ is slightly negative but close to zero.

Section *b.* of the table reports statistics for HP-filtered real and simulated data. The shocks to $\Delta\psi$ need to have a standard deviation in the vicinity of 5% in order to match the volatility of total factor productivity in the data. However, the next section will show that, as expected, the volatility in the model with capital and elastic labor supply is higher. The correlation between productivity and net reallocation is strongly positive.

Section *c.* of the table reports statistics based on log first-differenced data. Again, the basic model matches the near-zero first autocorrelation for aggregate productivity fairly well (.10 in the data versus $-.13$ to $-.07$ in the model and exceeds the standard deviation of aggregate productivity slightly (.82 in the data versus .94 to .96 in the model).

5 An Extended Set of Simulation Results and Further Empirical Evidence

5.1 The Model with Capital

In order to provide a fuller set of simulated moments by which to judge the model, physical capital is added as follows. The exposition will only consider *i.i.d.* plant-level shocks, hence a plant's q is a sufficient statistic for describing its state and exit decision.

The production function of an individual plant with productivity a is now a function of both labor input and capital input, $a[F(n, k)]^\alpha$.³⁶ The function F is assumed to exhibit constant returns to scale and unit elasticity of substitution between capital and labor. Because $0 < \alpha < 1$, plant production still exhibits decreasing returns to scale. Given their idiosyncratic realizations of a , plants choose labor and capital demands taking wages w and rental rates on capital r as given as in

$$\max_{n, k} a[F(n, k)]^\alpha - wn - rk \quad . \quad (28)$$

The resulting factor demands $n(a_j; w, r)$ and $k(a_j; w, r)$ and plant profits $\Pi(a_j; w, r)$ can be aggre-

³⁶Time subscripts are perfunctory here.

gated analagous to (10)-(11) resulting in aggregate labor demand $N^d(w, r)$ and aggregate capital demand $K^d(w, r)$ of

$$N^d(w, r) = (\alpha\gamma)^{\frac{1}{1-\alpha}} w^{\frac{1}{\alpha-1}} F\left(1, \frac{1-\gamma}{\gamma} \frac{w}{r}\right)^{\frac{\alpha}{1-\alpha}} [\tilde{A}_1 L + \tilde{A}_2 H] \quad (29)$$

$$K^d(w, r) = \frac{1-\alpha}{\gamma} \frac{w}{r} N^d(w, r) \quad (30)$$

where γ is the parameter controlling labor's share in the production function. Aggregate output $Y(w, r)$, and aggregate profits $\Pi(w, r)$ are given by

$$Y(w, r) = [\tilde{A}_1 L + \tilde{A}_2 H]^{1-\alpha} \left[F(N^d(w, r), K^d(w, r)) \right]^\alpha \quad (31)$$

$$\Pi(w, r) = (1 - \alpha)Y(w, r) \quad . \quad (32)$$

The plant's exit decision (16) is unchanged except for the change in expected profits to $\Pi^e(q; w, r) \equiv \sum_{j=1}^J \pi_j(q) \Pi(A_j; w, r)$. The cut-off prior q^* and the distribution of plant types is now endogenously determined with the economy-wide wage and capital rental rate. The learning mechanism in section 3.3.4 is unaffected by inclusion of capital.

The representative consumer's problem is augmented with decisions over investment and the budget constraint now includes returns on renting capital to plants. Since investment decisions involve intertemporal considerations, maximization of life-time utility is no longer accomplished by period-by-period maximization of $U(C, T - N)$. The problem can be written in the familiar form

$$\begin{aligned} & \max_{\{C_t, K_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} U(C_t, T - N_t) \\ \text{s.t. } & C_t + K_{t+1} - (1 - \mu)K_t = w_t N_t + r_t K_t + \Pi_t \\ & K_0 \text{ given} \end{aligned} \quad (33)$$

The first-order conditions for (33) are equivalent to those of the standard neoclassical growth model with competitive output and factor markets. Furthermore, since $w_t N_t + r_t K_t + \Pi_t = Y_t$, (33) is also equivalent to the social planner's problem with aggregate productivity determined according to $[\tilde{A}_1 L + \tilde{A}_2 H]^{1-\alpha}$

The end result of adding capital to the model is that it now resembles the neoclassical RBC model³⁷, as exemplified by King, Plosser and Rebelo (KPR) (1987). A natural question is whether the models exhibit similar dynamics in the vicinity of steady-state. It is possible to write the log

³⁷This model now resembles the neoclassical growth model in deviation from steady-state growth path, to be exact.

linearization around steady state of both this model and the RBC model as

$$\Phi_{s0}\hat{s}_t + \Phi_{a0}\hat{A}_t = \Phi_{s1}\hat{s}_{t+1} + \Phi_{a1}\hat{A}_{t+1} \quad , \quad (34)$$

where $\hat{s}_t \equiv (\hat{c}_t \ \hat{k}_t)'$, $\hat{x}$ denotes the deviation of x from its steady-state value, and Φ_{s0} and Φ_{s1} are (2×2) matrices and Φ_{a0} and Φ_{a1} are (2×1) vectors with elements that are functions of the parameters of the model.

The difference between this and the RBC model is that, instead of assuming an exogenous process for aggregate total factor productivity, aggregate productivity here is made endogenous to the entry and exit decisions of plants. In the context of (34), in the neoclassical RBC model $\hat{A}_{t+1} = \rho_A \hat{A}_t + \xi_t$ where ξ_t is assumed to be *i.i.d.*, mean zero, with variance σ_ξ^2 . Consequently, the RBC model can be written as a VAR in the state vector $\hat{x}_t \equiv (\hat{s}_t' \ \hat{A}_t)'$ with the single *i.i.d.* shock ξ_t .

$$\begin{pmatrix} \Phi_{s1} & \Phi_{a1} \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \hat{s}_{t+1} \\ \hat{A}_{t+1} \end{bmatrix} = \begin{pmatrix} \Phi_{s0} & \Phi_{a0} \\ 0 & \rho_A \end{pmatrix} \begin{bmatrix} \hat{s}_t \\ \hat{A}_t \end{bmatrix} + \begin{bmatrix} 0 \\ \xi_{t+1} \end{bmatrix} \quad (35)$$

The long-run dynamics of (35) are partly controlled by the eigenvalues of $\Phi_{s1}^{-1}\Phi_{s0}$. One additional eigenvalue at ρ_A is added to the system in the RBC case. Under the assumptions made on preferences and technology, the steady-state of the RBC model exhibits unique saddle path stability with a single eigenvalue greater than one in absolute value and the other two less than one in absolute value. In response to deviations from steady-state due to shocks ξ_t there exists a unique path for the vector $\hat{x}_t$ back to steady state.

Turning now to the model of this paper, aggregate productivity is a function of the mass of good and bad plants in the economy. Written in %-deviation from steady-state this becomes

$$\hat{A}_{t+1} = a_1' \begin{pmatrix} \hat{L}_{t+1} \\ \hat{H}_{t+1} \end{pmatrix} = (1 - \delta)a_1' \begin{pmatrix} \hat{L}_t \\ \hat{H}_t \end{pmatrix} + a_2' \begin{pmatrix} \hat{\epsilon}_{t+1}^L \\ \hat{\epsilon}_{t+1}^H \end{pmatrix} \quad , \quad (36)$$

where

$$a_1 \equiv (1 - \alpha) \begin{pmatrix} \bar{A}_1 L_{ss} \\ \bar{A}_2 H_{ss} \end{pmatrix} \quad a_2 \equiv (1 - \alpha) \begin{pmatrix} \bar{A}_1 \\ \bar{A}_2 \end{pmatrix} \quad .$$

Consequently, (34) becomes

$$\begin{bmatrix} \Phi_{s1} & \Phi_{a1}a_1' \\ 0 & I \end{bmatrix} \begin{pmatrix} \hat{s}_{t+1}' \\ \hat{L}_{t+1} \\ \hat{H}_{t+1} \end{pmatrix} = \begin{bmatrix} \Phi_{s0} & \Phi_{a0}a_1' \\ 0 & (1 - \delta)I \end{bmatrix} \begin{pmatrix} \hat{s}_t' \\ \hat{L}_t \\ \hat{H}_t \end{pmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{L_{ss}} & 0 \\ 0 & 0 & \frac{1}{H_{ss}} \end{bmatrix} \begin{pmatrix} 0 \\ \hat{\epsilon}_{t+1}^L \\ \hat{\epsilon}_{t+1}^H \end{pmatrix} \quad (37)$$

Evaluating the long-run dynamic behavior of (37) is more difficult than in the RBC case since

Table 4: Simulated Moments: Model with Capital

<i>a. Standard Deviations-HP Filtered Data</i>					
	U.S. Data	Model Simulation: $\sigma_{\Delta\psi} = .01$		Model Simulation: $\sigma_{\Delta\psi} = .05$	
		Fast	Slow	Fast	Slow
σ_y	2.25	0.19	1.30	1.49	2.67
σ_a/σ_y	0.43	0.56	0.48	0.36	0.30
σ_c/σ_y	0.39	0.27	0.54	0.55	0.47
σ_i/σ_y	1.95	1.35	2.13	3.49	3.95
σ_n/σ_y	0.84	0.84	0.91	1.04	1.18
σ_w/σ_y	0.43	0.17	0.85	0.78	0.73
σ_ϵ	—	0.22	0.07	0.27	0.07

<i>b. Key Correlations-HP Filtered Data</i>					
	U.S. Data	Model Simulation: $\sigma_{\Delta\psi} = .01$		Model Simulation: $\sigma_{\Delta\psi} = .05$	
		Fast	Slow	Fast	Slow
$\text{corr}(y,a)$	0.93	1.00	0.54	0.62	0.52
$\text{corr}(y,c)$	0.76	0.96	0.79	0.90	0.60
$\text{corr}(y,i)$	0.81	1.00	0.99	0.98	0.98
$\text{corr}(y,n)$	0.88	1.00	0.98	0.98	0.98
$\text{corr}(y,w)$	0.32	0.95	0.93	0.90	0.87
ρ_a	0.74	0.39	0.47	0.57	0.69
ρ_ϵ	—	-0.43	-0.38	-0.19	-0.18
$\text{corr}(a,\epsilon)$	.41	0.65	0.63	0.55	0.56

Note: Table reports selected moments from U.S. data and simulations from the basic model with capital under fast and slow learning as well as two different size shock standard deviations. Panel *a* reports standard deviations and panel *b* reports correlations based on HP-filtered data. σ_x denotes the standard deviation of variable x , $\text{corr}(y,x)$ denotes the contemporaneous correlation of output with variable x , and ρ_x denotes the first-order autocorrelation of variable x . U.S. statistics are calculated from aggregate quarterly U.S. NIPA Data as reported in Horvath (1999). See appendix B for further details on the calculation of productivity. The calculation of the correlation between productivity and ϵ in U.S. data is described under table 5 below. A simulation consists of equilibrium responses of the model to *i.i.d.* Normal shocks to $\Delta\psi$ with standard deviation $\sigma_{\Delta\psi}$ for 140 periods.

$\hat{\epsilon}_t$ is not exogenous to the vector $\hat{x}_t^m \equiv (\hat{s}_t' \hat{L}_t \hat{H}_t)'$. More precisely, (37) is not an equivalent representation to (35) since ϵ_t is dependent on the learning process described above and is jointly determined with $\hat{x}_t^m$ in equilibrium. However, it is instructive to consider the case of $\hat{\epsilon}_t$ *i.i.d* with $E(\hat{x}_t^m \hat{\epsilon}_t) = 0$. In this case, the long-run behavior of (37) is determined by the eigenvalues of $\Phi_{s1}^{-1} \Phi_{s0}$, as in the RBC model and by the eigenvalue $(1 - \delta)$. Thus, in this special case, the model of this paper will exhibit saddle path stability under parameterizations that yield stability in the RBC model. Equivalently, any deviations from the standard dynamics in this model stem from the dynamics embodied in ϵ_t from learning at the plant level.

Table 4 presents a broader set of moments from U.S. data and simulated data from the model with capital. The format of the table resembles closely that of table 3. Once again simulations

of the model are driven by shocks to the degree of plant specificity, $\Delta\psi$ that are *i.i.d.* over time. Panel *a* reports the standard deviation of output and the standard deviations of aggregate total factor productivity, consumption, investment, hours, wages, and reallocation shocks relative to the standard deviation of output. Panel *b* reports correlations of variables with output as well as the first-order autocorrelation statistics for aggregate productivity and reallocation shocks and the contemporary correlation between productivity and reallocations shocks.

Considering first the volatility measures, it appears that the model with capital is capable of generating sufficient volatility of aggregates in response to 5% shocks to the gap ($\psi^b - \psi_t^s$). As found in table 3 for the model without capital, volatility is higher in the slow learning parameterization than in the fast learning parameterization. With a shock standard deviation of 5%, output volatility is 1.49% in the fast learning case and 2.67% in the slow learning case compared with 2.25% for quarterly U.S. data. The relative volatilities of other variables conform to the U.S. experience as well. Aggregate productivity in the simulations has a volatility that is one-half to one-third that of output, depending on the shock size, which is in the vicinity of the figure for the U.S. post-war experience, 0.43. Consumption volatility is about half that of output while investment volatility is 1.5 to 3 times that of output, depending on shock size. Interestingly, hours are about as volatile as output while wages are smoother than output. This result is difficult to replicate in a standard RBC setting such as King, Plosser, and Rebelo (1987). Typically, the RBC type models generate labor hours that are too smooth relative to output and wages that are too volatile relative to output compared with the U.S. data.

Turning to the correlation statistics, it is evident that the model with capital replicates the U.S. correlations fairly well. All macro variables comove with output as in the data although wages are too correlated with output. Aggregate productivity has a first-order autocorrelation of between .39 and .69 depending on the speed of learning parameterization and the shock size. This matches the empirically observed productivity autocorrelation of .74 quite well. The reallocation shocks ϵ are slightly negatively autocorrelated due to the fact that larger than average entering cohorts tend to give rise to larger than average exiting cohorts in subsequent periods through exogenous death. Finally, the correlation between reallocation shocks ϵ and aggregate productivity is above .50 for all parameterizations. Net entry of plants increases productivity in the model. The empirical measure from the U.S. economy for this correlation is .41 and is described next.

5.2 Back to the Data

Testing the model empirically is a difficult task. It requires panel data with a long time-series dimension on productivity and factor payments at the individual plant level. In principle, such data exists. For example, the Longitudinal Research Database (LRD) for Manufacturing plants described in Davis, Haltiwanger and Schuh (1996) and used by Foster, Haltiwanger, and Krizen

Table 5: Correlation between U.S. TFP, Establishment Count, and Employment: 1962-1989

	TFP(-2)	TFP(-1)	TFP(0)	TFP(+1)	TFP(+2)
HP-Filter					
Establishments	-.37	.00	.41	.46	.19
Employment	.15	.42	.55	.26	-.19
First-Difference					
Establishments	-.42	-.06	.55	.49	-.27
Employment	.05	.23	.60	.21	-.48

Note: Table reports correlations between total factor productivity (TFP) and establishment and employment counts from the County Business Patterns dataset. The top panel reports correlations from HP-filtered data while the bottom panel reports correlations from data in log-first differences. Numbers in column TFP(j) denotes the correlation with TFP j periods from the variable in each row.

(1998), and Bartlesman and Dhrymes (1993) could be used for this purpose.³⁸ However, an extended empirical test would constitute a separate work beyond the scope and space limitations of this paper. Nevertheless, some rudimentary empirical tests can be performed with publicly available data. This section explores two pieces of empirical evidence that bear directly on the model's implications. While the empirical facts below are surely consistent with other theories of output and productivity variation, they also support the description of the economy presented in this model.

5.2.1 Empirical evidence on entry, exit, and productivity

The model implies that productivity should be positively correlated with the number of plants operating in the economy. Figure 4 plots annual data on TFP, employment and the number of establishments operating during the period 1962-1996. TFP is calculated as the cumulated percentage changes in weighted average sectoral productivity, as described in appendix B. The number of employed workers and the active establishment count are taken from the County Business Patterns (CBP) dataset collected by the Census Bureau.³⁹ An advantage of using this dataset over the Longitudinal Research Database (LRD) from the Census of Manufacturing is that it covers all sectors of the economy. Panel (a) of the figure reports the level of TFP, and employee and establishment counts in natural logarithms while panel (b) reports data detrended with the HP-filter. The figure also denotes NBER recession dates with vertical lines.

The correlation between TFP and the number of establishments is obvious. From panel (b), TFP and the establishment count start falling one year prior to the NBER-determined onset of recession and rise during and after the recession intervals. On the other hand, employment appears

³⁸Conditional on being a Census Fellow, of course.

³⁹An establishment is counted as active in a year if it paid employees during any part of the year. This feature of the data makes the count of establishments with 1-4 employees highly suspect. Consequently, the count of establishments with 1-4 employees is not included in figure 4. See <http://www.census.gov/epcd/cbp/view/genexpl.html>

to lag TFP and establishment count. Table 5 provides more systematic evidence on the timing of these movements, reporting the correlations of these series at different lags. The establishment count is negatively correlated with lagged TFP (based on either HP-filtered data or first differencing) but is positively correlated with contemporaneous TFP as well as with TFP movements up to two years ahead. In fact, the peak correlation of .46 occurs between establishment count movements and one-year-ahead movements in TFP. Employment is only strongly correlated with contemporaneous TFP but is more correlated with lagged TFP than with TFP leads.

From (26) the model implies that productivity should decline in periods of mass exit and should rise in periods of mass entry. In other words, besides being correlated with the number of plants operating in the economy, with the net flows shown in figure 4, productivity movements should be specifically related to the gross flows of establishments. Figure 5 plots annual data on the percentage change in exit (panel *a*) and entry (panel *b*) of numbers of *firms*⁴⁰ over the sample period 1948-83 along with the annual percentage change in aggregate productivity. In order to improve resolution, the productivity changes have been scaled up by a factor of 10. Vertical lines indicate beginning and ending months of recessions.⁴¹

Consistent with the model's predictions, the correlation between the exit series and aggregate productivity changes is $-.43$ while the correlation between the entry series and aggregate productivity changes is $.35$. Furthermore, the figure shows that plant entry and exit *lead* the cycle: exits rise sharply and entry drops sharply before recessions start and both series reverse direction before recessions end. Such evidence indicates that the conventional wisdom that plants fail during recessions and enter during booms may have the timing and causality reversed. In the literature on creative destruction exemplified Caballero and Hammour (1994, 1996) plants choose to enter and exit in conjunction with cyclical demand variation. These models would not be able to replicate the finding that entry and exit lead the cycle.

5.2.2 Empirical evidence on job creation/destruction

The model suggests that younger plants are less informed about their operating characteristics than older plants. Consequently, a bulk of the reallocation of labor should be associated the exit of younger plants and this should coincide with periods of lower aggregate productivity. Similarly, the model suggests that the creation of jobs through entry by new plants is also a major source of labor reallocation and should be associated with productivity increases.

Davis, Haltiwanger and Schuh (1996) provide summary data at the quarterly frequency on the rate of job creation and destruction⁴² for plants of different age groups. Figures 6 and 7 present

⁴⁰Data are taken from Dun and Bradstreet as reported in *U.S. Statistical Abstracts* for various years. Unfortunately, this data is unavailable at the plant level.

⁴¹While the time series are annual series, the recession dates are based on beginning and ending months established by the NBER.

⁴²These rates are calculated as the number of jobs created or destroyed divided by a measure of total employment

seasonally adjusted job creation and destruction rates from 1972:2-1988:4 for three groupings of plants: young (generally 1 year old or less), middle aged (generally 2-10 years of age) and old (generally more than 10 years old). Figure 6 presents these rates for all plants while figure 7 focuses only on creation by entering plants and destruction by exiting plants. In both sets of figures, NBER recession dates are marked with vertical lines.⁴³ From figure 6 the relative magnitudes of gross job flow rates is apparent. Job creation is greater in young plants while the job destruction rate is greater in middle aged plants. Old plants have relatively low rates of both job creation and destruction.⁴⁴ This suggests that with age of the production unit comes less need for reallocation of resources. Conversely, much of the reallocation activity during a recession is concentrated in production units that have a relatively short operating history.

From figure 7 it is clear that young entering plants account for nearly half of the job creation by young plants. It is supportive of the model in this paper that job creation by young entering plants seems to slow well before the NBER start dates and rises before NBER end dates. As with the evidence in figure 5, figure 7 suggest that a decline in the pace of entry of new plants is a leading indicator for cyclical fluctuations. A similar pattern *vis à vis* the NBER dates is not observed among middle aged or old transitioning plants. In these plants job creation slows mildly after the NBER start date and rises after the NBER end date.

5.3 Micro Evidence on Productivity Variation

Relatively little research has been performed to date to asses the implications of plant-level behavioral patterns on aggregate productivity. Foster, Haltiwanger, and Krizen (1998), Olley and Pakes (1996), Bartlesman and Dhrymes (1993) and Bailey, Hulten, and Campbell (1992) present micro-empirical evidence that productivity variation and reallocations (both extensive and intensive) are linked. These studies consider TFP movements in specific two- or four-digit SIC sectors. Their findings uniformly indicate that a significant portion of sectoral productivity change is due to the reallocation of resources (*e.g.* capital and labor) between entering and exiting plants. This finding suggests that plants within a sector may not be experiencing changes in their productivity (at least over intervals of several decades) but that “good” plants gain market share while “bad” plants shrink or exit their industry. A portion of variation in industry productivity is due, therefore, to compositional shifts in the types of plants producing in the industry.

Another finding in their research is that entry occurs rather uniformly at all levels of productivity within an industry but the probability of moving up or down in a productivity ranking of all plants is higher for the youngest plants. This again suggests the existence of a higher degree

at manufacturing plants. The data is available by anonymous FTP from <ftp://129.2.28.100>.

⁴³These series are not matched with data on productivity variation because of the short sample period and the lack of reliable quarterly data on manufacturing productivity.

⁴⁴See also Dunne, Roberts, and Samuelson (1989a).

of uncertainty surrounding the productivity of entering plants which is resolved through the first few years of experience in the industry.⁴⁵

The behavior of gross flows or reallocations, both within existing production units and by entering and exiting production units, has been the focus of intense research over the past decade, not only because it is an interesting phenomenon in itself, but because it is a necessary component for understanding such aggregate phenomena as labor fluctuations.⁴⁶ Dunne, Roberts, and Samuelson (1989a) present empirical evidence on plant failure rates. Their most relevant finding is that failure rates for manufacturing plants are declining in age, holding size constant, and declining in size, holding age constant. This evidence corroborates the learning models in this paper and in Jovanovic (1982) in which productive units are initially uncertain about some feature of their operating environment (*e.g.* productivity, marginal cost, *etc.*) and learn through experience about these features. In these models, learning induces exit of unfit productive units over time, generating exit rates that decline with a cohort's age as only fit productive units remain in operation.

Furthermore, Dunne, Roberts, and Samuelson (1988) find that plant and industry specific uncertainty dominates aggregate sources of uncertainty in influencing plant level dynamics. This is evidenced by substantial and persistent differences in entry and exit rates across industries and a high correlation between entry and exit rates within industries. Their findings also reveal that the long-run impact of entry or exit on market structure and performance depends on the characteristics of the entering and exiting plants. Consider, for example, their finding that small plants have significantly higher failure rates, even when controlling for age. This suggests the size distribution of plants within an industry will affect aggregate measures such as industry productivity.

6 Extensions

Two extensions of the model will be considered in this section. First, growth will be incorporated into the model in two separate ways. Under either or both of these modifications the simple relationship between aggregate productivity and the distributions of low and high productivity plants in the economy remains unchanged. In the second modification to the model, the consumer's problem is respecified so that decisions over labor supply must be made based on beliefs about plant productivity. Through income effects, this results in lower labor supply when beliefs are overly optimistic and higher labor supply when beliefs are overly pessimistic. A social planner with

⁴⁵Bartlesman and Dhrymes also find that larger plants have a lower probability of moving up or down in the productivity ranking, which is another manifestation of the resolution of uncertainty if entrants tend to be smaller plants.

⁴⁶See Dunne, Roberts, and Samuelson (1989b) and Davis, Haltiwanger, and Schuh (1996), Davis and Haltiwanger (1990, 1992) and Caballero, Engel, and Haltiwanger (1994). More recently, Cooper, Haltiwanger, and Power (1994) and Ramey and Shapiro (1998a,b,c) have begun to investigate reallocations involving durable capital goods.

knowledge about plant productivity would chose factor allocations that would be welfare improving relative to the competitive equilibrium.

6.1 Incorporating Growth

The model⁴⁷ can be modified to incorporate the observed growth in the level of aggregate productivity seen in figures 1 and 4. This can be done in two ways. Growth can enter the model through a deterministic trend in M_t , the mass of new entrants in period t in the no-growth economy. In this case, measured aggregate productivity grows since adding more plants helps overcome the decreasing returns to scale assumption on each individual plant. However, plant level productivity does not grow over time. Alternately, growth can take the form of traditional technological improvements in productivity at the plant level. Both versions will be considered in turn as will their implications for explaining empirical observations on productivity growth.

6.1.1 Growth in M_t

In the absence of shocks, the mass of new entrants each period, M_t is now assumed to follow the deterministic path

$$M_t = (1 + \gamma_M)M_{t-1}, \quad \gamma_M > 0 \quad . \quad (38)$$

This path is achieved through deterministic reductions in the degree of specificity $\Delta\psi$ at the plant level. As a consequence, the mass of low and high productivity plants in the economy will exhibit exponential growth. In order to study the model economy, it necessary to normalize the measures $\lambda_t^i(q)$, $\rho_t^i(q)$ for $i = 1, 2$ by M_t . With this normalization, (22) becomes

$$\begin{pmatrix} \tilde{L}_t \\ \tilde{H}_t \end{pmatrix} = (1 - \delta)(1 + \gamma_M)^{-1} \begin{pmatrix} \tilde{L}_{t-1} \\ \tilde{H}_{t-1} \end{pmatrix} + \begin{pmatrix} \tilde{\epsilon}_t^L \\ \tilde{\epsilon}_t^H \end{pmatrix} \quad (39)$$

where $\tilde{x}_t \equiv \frac{x_t}{M_t}$. Aggregate productivity can be written as

$$A_t = M_t^{1-\alpha} \left[\tilde{A}_1 \tilde{L}_t + \tilde{A}_2 \tilde{H}_t \right]. \quad (40)$$

implying that, in the absence of shocks, aggregate productivity exhibits deterministic growth at the rate $(1 + \gamma_M)^{1-\alpha}$. The aggregate variables Y_t , w_t , Π_t , and C_t will all grow at the rate $(1 + \gamma_M)^{1-\alpha}$ as well, implying that growth is balanced.⁴⁸

⁴⁷Growth can be incorporated into the model with or without capital in an equivalent fashion.

⁴⁸The model abstracts from deterministic population growth. Growth in the measure of households/workers would modify the above growth accounting. Output growth would be the sum of the growth in productivity and the growth in the labor force.

While measured aggregate productivity is growing, it is still the case that average plant-level productivity is $\tilde{A}_1$ for low productivity plants and $\tilde{A}_2$ for high productivity plants. The source of aggregate growth is in overcoming the decreasing returns to scale at the plant level by the addition of more plants in the economy each period.

6.1.2 Growth in $\tilde{A}_i$

Technological advancement that leads to growth can be incorporated into the model by the assumption

$$A_{j,t} = (1 + \gamma_A)A_{j,t-1}, \quad \gamma_A > 0 \quad . \quad (41)$$

In this case, productivity at all plants improves by $(1 + \gamma_A)$ each period. This will have no effect on the distributions of plants in the economy. Measured aggregate productivity is still

$$A_t = \left[\tilde{A}_{1,t}L_t + \tilde{A}_{2,t}H_t \right] \quad , \quad (42)$$

where now $\tilde{A}_{i,t} = \sum_{j=1}^J \pi_{ij} A_{j,t}$. Using (41) it is clear that $A_t/A_{t-1} = (1 + \gamma_A)$. Again, all aggregate variables will also grow at the rate $(1 + \gamma_A)$.

6.1.3 Decomposing measured productivity growth

Both versions of growth can be incorporated into the model simultaneously. The resulting growth rate of measured aggregate productivity is given by

$$\ln \frac{A_t}{A_{t-1}} \equiv \gamma_{AP} = (1 - \alpha)\gamma_M + \gamma_A \quad . \quad (43)$$

Equation (43) decomposes growth in measured aggregate productivity (γ_{AP}) into a portion due to overcoming decreasing returns to scale $((1 - \alpha)\gamma_M)$ and a portion due to technological advancement (γ_A). Using data from figure 4, the mean annual growth rate of measured aggregate productivity is 0.51% over the sample 1948-89. The mean annual growth rate of the number of plants⁴⁹ in the economy is 2.48%. The value of α can be inferred from the average ratio of corporate profits to aggregate output. The model suggests that $\frac{\Pi_t}{Y_t} = (1 - \alpha)$. The average ratio of corporate profits to aggregate output⁵⁰ over the sample 1948-89 is 0.133 implying a value of $\alpha = 0.867$.⁵¹ Therefore,

⁴⁹Based County Business Pattern data reported in figure 4.

⁵⁰The series used here are as follows: corporate profits before taxes and adjustments (Citibase PBT) and GDP less government, housing and farm output (Citibase GDPQ).

⁵¹The value of α could also be estimated from (43) by an uncentered regression of the annual percentage change in measured aggregate productivity onto the percentage change in the number of plants. The residual from this

the model suggest that nearly two-thirds of aggregate productivity growth over the sample period has come simply from the increase in the number of plants in the economy.

6.2 Beliefs and factor misallocation

In the model of section 3, labor is always allocated efficiently given the distribution of plants existing in the economy. Theorem 1 showed that allocations in the decentralized economy are equivalent to the allocations made by a social planner with knowledge of the number of low and high productivity plants. Furthermore, the beliefs held by plants about their type only affect the exit decision of marginal plants and do not affect aggregate variables. In this section⁵² it will be shown that misallocation of labor results when households can not gather information from plants in a timely fashion. The heuristic story of informational constraints being modeled is this: consumers receive quarterly reports from plants describing past performance and expectations of future profitability. In other words, consumers receive an updated assesment of expected productivity in period t , $A_t(q) \equiv q_t \tilde{A}_2 + (1 - q_t) \tilde{A}_1$, from each plant at the end of period $t - 1$. However, plants are unable to communicate the value of A_j they receive to the consumers within period t . In this case decisions over labor supply, which by (4) require the consumer to know or form expectations over corporate profits, are based on the distribution of beliefs held by plants, as is the equilibrium wage. When beliefs communicated to households from plants are overly optimistic, households undersupply labor relative to the social planner optimum because of expected income effects. The resulting equilibrium wage is above what would obtain in the economy without informational constraints causing more plants to exit and a level of production lower than is socially optimal. By (26), aggregate productivity falls. A similar but opposite picture obtains when beliefs communicated to households from plants are overly pessimistic.

The modifications to the model of section 3 described above enter solely in the problem of the representative consumer. Now the consumer must solve

$$\max_{C_t, N_t} U(C_t, T - N_t) \quad \text{s.t.} \quad C_t \leq w_t N_t + \Pi_t^e(w_t; \lambda_t), \quad 0 \leq N_t \leq T \quad . \quad (44)$$

regression is interpreted as the annual percentage change in aggregate technology and the coefficient being estimated is $(1 - \alpha)$. Doing this, the OLS estimate is $\alpha = .953$ with a standard error of 0.098. A two-sided test of the null hypothesis $\alpha = .867$ has a p-value of .35 and thus can not be rejected. Regression by OLS may result in a biased estimate of $(1 - \alpha)$ since, according to the model, growth in the number of plants should be negatively correlated with growth in technology. When technology improves, the aggregate wage rises and net entry of plants falls. Given this, the bias will result in an underestimate of $(1 - \alpha)$ or an overestimate of α . Therefore, OLS can be considered a conservative estimation approach when attempting to determine the share of measured productivity growth accounted for by the growth in the number of plants.

⁵²In the interest of conserving space, the presentation of this modification is made in the context of the model without capital. However, the results would extend to the model with capital as well.

where

$$\begin{aligned}\Pi^e(w_t; \lambda_t) &= (1 - \alpha) \alpha^{\frac{\alpha}{1-\alpha}} w_t^{\frac{\alpha}{\alpha-1}} A_{\Pi,t} \\ A_{\Pi,t} &\equiv \int A(q)^{\frac{1}{1-\alpha}} \lambda_t(q) dq \quad .\end{aligned}\tag{45}$$

In (45), $\lambda_t(q) = \lambda_t^1(q) + \lambda_t^2(q)$. Aggregate labor demand is as before, given by (11). Labor supply is determined by (4) with $\Pi^e(w_t; \lambda_t)$ replacing $\Pi(w_t)$. Labor market clearing determines equilibrium aggregate labor input $N(w_t; \lambda_t)$, and the equilibrium wage $w_t(\lambda_t)$. Aggregate output as a function of the equilibrium wage is still given by (14). Therefore, measured aggregate productivity is still $A_t = [\tilde{A}_1 L_t + \tilde{A}_2 H_t]$. Aggregate profits (observed rather than expected) as a function of the equilibrium wage are still given by (13). Thus the only difference between the economy with informational constraints and the unconstrained economy in section 3 is that the equilibrium wage depends on the distribution of beliefs λ_t held by operating plants in period t .

As in theorem 1, the social planner's problem involves solving

$$\max_{C_t, N_t} U(C_t, T - N_t) \quad \text{s.t.} \quad C_t \leq A_t N_t^\alpha, \quad 0 \leq N_t \leq T \quad .\tag{46}$$

Because beliefs do not enter (46), the solution to the social planner's problem does not depend on $A_{\Pi,t}$. Consequently, Pareto efficient equilibrium allocations will be the same as in the model of section 3.⁵³

Comparing (45) with (13) it is clear that

$$r_{\Pi,t} \equiv \frac{\Pi^e(w_t; \lambda_t)}{\Pi(w_t; \lambda_t)} = A_{\Pi,t} A_t^{\frac{1}{\alpha-1}} \quad .\tag{47}$$

When λ_t is shifted to the right with more plants holding high beliefs, the economy is characterized by excess optimism relative to the actual number of low and high productivity plants L_t and H_t . In this case $r_{\Pi,t} > 1$. Since U is assumed to be quasi-concave, the excess expected profits translates into a lower labor supply at every wage offer. If the labor demand curve is unaffected, equilibrium in this case is characterized by a lower labor input and a higher equilibrium wage. When more mass in λ_t is concentrated closer to the exit margin, the economy is characterized by excess pessimism relative to the actual numbers of L_t and H_t that determine measured productivity A_t and $r_{\Pi,t} < 1$. Consequently, the income effect works in the opposite direction shifting the labor supply curve out to the right. With a stable labor demand curve, equilibrium is characterized by higher labor input

⁵³Of course, this compares the outcome from a social planner with knowledge of L_t and H_t when setting N_t to a decentralized economy where consumers are uncertain about L_t and H_t when determining labor supply. If the social planner had the same information that the representative consumer had (Π_t^e) when setting N_t , the social planner allocations would match those from the decentralized solution. However, the former is the correct comparison to make since, by Theorem 1 it is also the comparison between the decentralized informational constraints economy and the unconstrained decentralized economy.

and lower equilibrium wage.

The story is not so simple, however. The description of equilibrium above ignored the effect of equilibrium wages on the exit decisions of marginal plants. Recall from (16) that when wages are higher, q^* tends to move to the right inducing more plants to exit. Therefore, periods of excess optimism (higher wages) are also periods in which a larger mass of plants exit. This will shift the labor demand curve to the left. The result is an equilibrium with even lower labor input but higher wages compared to the description given above. Empirical observations on wages and hours worked suggest that the former move much less over the business cycle than the latter. The standard RBC model parameterized to micro observations on labor supply elasticities imply the opposite: volatile wages and smooth labor hours. Consequently, rectifying the RBC paradigm with this stylized fact has been the focus of many RBC-type models including Hansen (1985), Rogerson (1988), and Boldrin and Horvath (1995). It appears that the model with informational constraints may provide an explanation for smooth wages and volatile hours distinct from explanations that focus on preferences of consumers or labor market mechanisms. Furthermore, the dynamic implications of the model with informational constraints are potentially very interesting since it seems intuitively possible to generate endogenous fluctuations.

7 Conclusion

This paper has presented a candidate source for aggregate productivity fluctuations and an economic explanation for why these fluctuations are persistent. Aggregate productivity is shown to depend on the composition of plants in the economy. The essence of the model is the process of creative destruction envisioned by Schumpeter (1942). In this view, entry and exit of plants are the “shocks” to aggregate productivity since they alter the distribution of plant types. Mass plant failure causes aggregate productivity to fall as resources are reallocated across remaining plants. Mass entry causes aggregate productivity to rise as more productive units are brought on line. Persistence of aggregate productivity comes through the learning process at the plant level. Because it takes time for low productivity plants to learn that they are inefficient and exit, periods of low productivity caused by a relative abundance of low productivity plants can persist. Because high productivity plants tend to learn that they are efficient and hence survive longer than low productivity plants, periods of high productivity caused by a relative abundance of high productivity plants can persist. Simulations of the model are encouraging in its ability to generate time series for productivity and other macro variables that are consistent with empirical observations. In particular, the model responses of aggregate productivity are highly persistent even in response to *i.i.d.* shocks. The simple relationship between aggregate productivity and the turnover of plants suggested by the model is supported by empirical time series on plant entry and exit and aggregate TFP.

References

- [1] Abadir, Karim and Gabriel Talmain. 1998. "Aggregation and Persistence in a Macromodel," manuscript, University of York.
- [2] Atkeson, Andrew and Patrick J. Kehoe. 1993. "Industry Evolution and Transition: The Role of Information Capital" manuscript, University of Chicago and University of Pennsylvania.
- [3] Aw, Bee Yan, Chen Xiaomin, Roberts, Mark J. 1997. "Firm-level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing," NBER Working Paper: 6235.
- [4] Bahk, Byong and Michael Gort. 1993. "Decomposing Learning by Doing in New Plants," *Journal of Political Economy*, **101**, 561-83.
- [5] Bailey, Martin, Charles Hulten and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity: Microeconomics*, 187-267.
- [6] Bartlesman, Eric J. and Phoebe J. Dhrymes. 1993. "Productivity Dynamics: US Manufacturing Plants, 1972-1986" manuscript, Federal Reserve Board and Columbia University.
- [7] Basu, Susanto. 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization," *Quarterly Journal of Economics*, **111**, 719-752.
- [8] Basu, Susanto and John G. Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy* **105**, 249-83.
- [9] Basu, Susanto and John G. Fernald. 1998. "Aggregate Productivity and Aggregate Technology," manuscript, University of Michigan.
- [10] Basu, Susanto, and Miles S. Kimball. 1994. "Cyclical Productivity with Unobserved Input Variation," Hoover Institution, manuscript.
- [11] Benhabib, Jess and Roger E.A. Farmer. 1997. "Indeterminacy and Sunspots in Macroeconomics," manuscript prepared for the *Handbook of Macroeconomics*.
- [12] Blanchard, Olivier Jean and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review* **79**, 655-73.
- [13] Boldrin, Michele and Michael Horvath. 1995. "Business Cycles and Labor Contracts," *Journal of Political Economy* **103**, 982-1018.
- [14] Brock, William and Leonard Mirman. 1972. "Optimal Economic Growth and Uncertainty: the Discounted Case," *Journal of Economic Theory* **4**, 479-513.
- [15] Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 1995. "Capital Utilization and Returns to Scale," in Bernanke, B. and Rotemberg, J. (eds.), *NBER Macroeconomics Annual 1995*, Cambridge: MIT Press, 67-110.
- [16] Caballero, Ricardo, Eduardo Engel, and John Haltiwanger. 1997. "Aggregate Employment Dynamics: Building from Microeconomic Evidence," *American Economic Review* **87**, 115-137.
- [17] Caballero, Ricardo and Mohamad Hammour. 1994. "The Cleansing Effect of Recessions," *American Economic Review* **84**, 1350-1368.
- [18] Caballero, Ricardo and Mohamad Hammour. 1996. "On the Timing and Efficiency of Creative Destruction," *Quarterly Journal of Economics* **111**, 805-854.
- [19] Campbell, Jeffrey R. 1998. "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics* **1**, 371-408.
- [20] Christiano, Lawrence J. and Martin Eichenbaum. 1992. "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, **82**, 430-450.
- [21] Cochrane, John. 1994. "Shocks" *Carnegie-Rochester Conference Series on Public Policy* **41**, 295-364.
- [22] Cogley, Timothy and James M. Nason. 1995. "Output Dynamics in Real-Business Cycle Models," *American Economic Review*, **85**, 492-511.

- [23] Cooper, Russell and John Haltiwanger. 1993. "The Aggregate Implications of Machine Replacement: Theory and Evidence," *American Economic Review* **83** 360-82.
- [24] Cooper, Russell and John Haltiwanger and Laura Power. 1994. "Machine Replacement and the Business Cycle: Lumps and Bumps," manuscript.
- [25] Cooper, Russell and Alok Johri. 1999. "Learning By Doing and Aggregate Fluctuations," NBER Working Paper 6898.
- [26] Davis, Steven J. and John Haltiwanger. 1990. "Gross Job Creation and Destruction: Microeconomic Evidence and Macroeconomic Implications," *NBER Macroeconomics Annual*, 123-168.
- [27] Davis, Steven J. and John Haltiwanger. 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, **107**, 819-863.
- [28] Davis, Steven J., John Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*, Cambridge, Mass: MIT Press.
- [29] Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1989a. "The Growth and Failure of U.S. Manufacturing Plants," *Quarterly Journal of Economics* **104**: 671-698. November.
- [30] Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1989b. *Journal of Labor Economics*, **7**, 48-71.
- [31] Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1988. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, **19**, 495-515.
- [32] Ericson, Richard and Ariel Pakes. 1987. "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory* **79**, 1-45.
- [33] Ericson, Richard and Ariel Pakes. 1988. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies* **62**, 53-82.
- [34] Farmer, Roger E.A. and Jang-Ting Guo. 1994. "Real Business Cycles and the Animal Spirits Hypothesis," *Journal of Economic Theory* **63**, 42-72.
- [35] Foster Lucia, John Haltiwanger, and C.J. Krizen. 1998. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," NBER Working Paper: 6803.
- [36] Galí, J. 1995. "Non-Walrasian Employment Fluctuations," NBER Working Paper: 5337.
- [37] Galí, J. 1996. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" NBER Working Paper: 5721.
- [38] Gilchrist, Simon and John C. Williams. 1998. "Putty-Clay and Investment: A Business Cycle Analysis," Board of Governors of the Federal Reserve FEDS Working Paper 1998-30.
- [39] Gort, Michael, Byong Bahk, and Richard Wall. 1993. "Decomposing Technical Change," *Southern Economic Journal*, **60**, 220-34.
- [40] Hall, Robert E. 1988. "The Relation Between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, **96**, 921-47,
- [41] Hall, Robert E. 1990. "Invariance Properties of Solow's Productivity Residual," In *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday*, Peter Diamond, ed. Cambridge, Mass.:MIT Press.
- [42] Hansen, Gary. 1985. "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics* **16**, 309-328.
- [43] Hopenhayn, Hugo A. 1992a. "Exit, Selection and the Value of Firms," *Journal of Economic Dynamics and Control*, **16**, 621-653.
- [44] Hopenhayn, Hugo A. 1992b. "Entry, Exit and Firm Dynamics in Long Run Equilibrium," *Econometrica*, **60**, 1127-1150.
- [45] Horvath, Michael T.K. 1999. "Sectoral Shocks and Aggregate Fluctuations," *forthcoming, Journal of Monetary Economics*.
- [46] Jorgenson, Dale W., Frank Gollop, and Barbara Fraumeni. 1987. *Productivity and U.S. Economic Growth*. Cambridge: Harvard University Press. 1987.

- [47] Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry" *Econometrica* **50**: 649-670.
- [48] Jovanovic, Boyan and Yaw Nyarko. 1995. "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," *Brookings Papers on Economic Activity*: 247-299
- [49] Kiley, Michael T. 1997. "Labor Productivity in U.S. Manufacturing: Does Sectoral Comovement Reflect Technology Shocks?"
- [50] King, Robert G., Charles I. Plosser, and Sergio Rebelo. 1987. "Production, Growth and Business Cycles: I. The Basic Neoclassical Model" *Journal of Monetary Economics* **21**.
- [51] Krusell, Per and Anthony A. Smith. 1998 "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy* **106**, 867-896.
- [52] Kydland, Finn, and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations" *Econometrica* **50**, 1345-1370.
- [53] Lucas, Robert E. 1978. "On the Size Distribution of Business Firms" *The Bell Journal of Economics* **9**:508-523.
- [54] Mankiw, N. Gregory. 1989. "Real Business Cycles: A New Keynesian Perspective," *Journal of Economic Perspectives*, **3**, 79-90.
- [55] Olley, Steven G. and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, **64**, 1263-97.
- [56] Portier, Franck. 1995. "Business Formation and Cyclical Markups in the French Business Cycle," *Annales d'Economie et de Statistique*, Jan-June 1995, 411-440.
- [57] Pries, Michael J. 1998 "Persistence of Employment Fluctuations: A Model of Recurring Job Loss," manuscript, Stanford University.
- [58] Ramey, Valerie and Matthew D. Shapiro. 1998a. "Costly Capital Reallocation and the Effects of Government Spending," *Carnegie-Rochester Conference Series on Public Policy*, forthcoming 1998.
- [59] Ramey, Valerie and Matthew D. Shapiro. 1998b. "Displaced Capital," manuscript.
- [60] Ramey, Valerie and Matthew D. Shapiro. 1998c. "Capital Churning," manuscript.
- [61] Rogerson, Richard. 1988. "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics* **21**, 3-16.
- [62] Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper, 1942.
- [63] Shleifer, Andrei. 1986. "Implementation Cycles," *Journal of Political Economy* **94**, 1163-90.
- [64] Solow, Robert M. 1957. "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics* **39**: 312-320.
- [65] Summers, Lawrence H. 1986. "Some Skeptical Observations on Real Business Cycle Theory," Federal Reserve Bank of Minneapolis *Quarterly Review*, 10, 23-27.
- [66] Thomas, Julia. 1998. "Lumpy Investment, Partial Adjustment and the Business Cycle: A Reconciliation," manuscript, Carnegie Mellon University.
- [67] Veracierto, Marcelo. 1997. "Plant Level Irreversible Investment and Equilibrium Business Cycles," Federal Reserve Bank of Minneapolis Discussion Paper 115.

A Proofs

A.1 Proof Theorem 1

The first-order conditions for a solution to **SP1** match those of the aggregated decentralized heterogeneous plant economy. Specifically, given A_t , equilibrium in **SP1** is achieved with $C_t = Y_t$, $w_t = \alpha A_t N_t^{\alpha-1}$ and N_t equates (11) solved for w_t to (4), just as in the heterogeneous plant economy. Therefore, aggregate allocations in the two economies must be identical. ■

A.2 Proof Theorem 2

Standard proofs for the existence and uniqueness of a centralized equilibrium for the aggregate economy (**SP1**) and extensions to the competitive one-sector economy can be invoked to prove existence and uniqueness of aggregate quantities N_t, Π_t, Y_t, C_t , given aggregate productivity A_t . Furthermore, given distributions of plants $\lambda_{t-1}^i(q)$ $i = 1, 2$ in the heterogeneous plant economy, aggregate productivity in period t is uniquely determined by q_t^* and M_t . Therefore, all that is required to prove existence and uniqueness of a competitive equilibrium in the heterogeneous plant economy is to show that given ψ_t^s, ψ_t^b , and distributions of plants $\lambda_{t-1}^i(q)$ $i = 1, 2$, there exists a unique $q^* \in [0, 1]$ satisfying $\tilde{V}(q^*; w_t) = \psi_t^s$ and a unique $M \in R_+$ such that $\int_0^1 V(q; w_t) dq = \psi_t^b$. Since $\beta(1 - \delta) < 1$, standard dynamic programming results ensure that $\tilde{V}(q; w_t)$ is continuous, bounded and monotonically increasing in q so long as $\Pi^e(q; w_t)$ is continuous, bounded and monotonically increasing in q and $b_j(q)$ is continuous and increasing in q for all j . $\Pi^e(q; w_t)$ is continuous, linearly increasing in q , and bounded for $w_t < \infty$. Therefore, there can be only one q^* satisfying $\tilde{V}(q^*; w_t) = \psi_t^s$. Since q^* must stay on the open unit interval, the only problem with existence of a unique $q^* \in [0, 1]$ occurs when $\tilde{V}(0; w_t) > \psi_t^s$ or $\tilde{V}(1; w_t) < \psi_t^s$. In the first case, plants with any prior $q \in (0, 1)$ are viable while in the second no plants are viable. The second case can be ruled out directly since in the limit, as all plants exit the economy, $A_t \rightarrow 0$. Hence, $w_t \rightarrow 0$, $\Pi^e(q; w_t) \rightarrow \infty$ and $\tilde{V}(q; w_t) \rightarrow \infty$. The first case (all plants viable) is permissible and handled by simply letting $q^* = 0$ whenever $\tilde{V}(0; w_t) > \psi_t^s$. Given $q^*(w_t)$ satisfying $\tilde{V}(q; w_t) = \psi_t^s$ (ignoring the corner solution at $q^* = 0$ for the remainder of the proof), all that remains is to show that there is a unique value M_t such that $\int_0^1 V(q; w_t) dq = \psi_t^b$ and that w_t is consistent with labor market clearing at the implied value for A_t under q_t^* and M_t . Let $\tilde{N}_t^f = N_t^f - (1 - q^*)v_h M_t$. Then $A_t = (\tilde{a}\tilde{N}_t^f + (1 - q_t^*)\tilde{a}v_h M_t)^{1-\alpha}$. Given q_t^* , $\tilde{N}_t^f$ is uniquely determined from $\tilde{N}_t^f = (1 - \delta)N_{t-1}^f - (1 - \delta)\int_0^{q_t^*} \rho_t(q) dq$ and does not depend on M_t . Therefore, for a given q_t^* , A_t is strictly increasing in M_t . Given q_t^* , the zero-expected-profit entry condition can be written

$$0 = \psi_t^b - \int_0^1 V(q; w_t) dq = q_t^* \psi_t^s + w_t^{\frac{\alpha}{\alpha-1}} \int_{q_t^*}^1 \Pi_x(q) dq + \beta(1 - \delta) \int_{q_t^*}^1 E(V_{t+1}(q)) dq \quad (48)$$

where $\Pi_x(q) \equiv (1 - \alpha)\alpha^{\frac{\alpha}{1-\alpha}}(q\tilde{A}_2 + (1 - q)\tilde{A}_1)$ and $E(V_{t+1}(q))$ denotes the expected value in period $t + 1$ of a plant with prior q in period t . From theorem 1, it must be the case that $w_t = \alpha A_t N_t^{\alpha-1}$. Substituting this expression in for w_t in (48) yields the following equation in M_t .

$$0 = \psi_t^b - q_t^* \psi_t^s + \alpha^{\frac{\alpha}{\alpha-1}} N_t^{\alpha} \Pi_x(q_t^*) \left[\tilde{a}\tilde{N}_t^f + (1 - q_t^*)\tilde{a}v_h M_t \right]^{-\alpha} + \beta(1 - \delta) \int_{q_t^*}^1 E(V_{t+1}(q)) dq \quad (49)$$

where $\Pi_x(q_t^*) = \int_{q_t^*}^1 \Pi_x(q) dq = (1 - \alpha)\alpha^{\frac{\alpha}{1-\alpha}} \{ \tilde{A}_1(1 - q_t^*) + (\tilde{A}_2 - \tilde{A}_1)\frac{1}{2}(1 - (q_t^*)^2) \}$. The right hand side of (49) is monotonically decreasing in M_t so long as labor supply is not so wage elastic that equilibrium wages fall

with an increase in A_t . Therefore, given q_t^* , a unique M_t exists satisfying (49), subject to the boundary condition $M_t \geq 0$. M_t will be bounded so long as $\psi_t^b > \psi_t^s$. To see this note that as $M_t \rightarrow \infty$, $A_t \rightarrow \infty$ and $w_t \rightarrow \infty$. Consequently, $q_t^* \rightarrow 1$ and $E(V_{t+1}(q)) \rightarrow 0$. Since labor supply is bounded, substituting this into (49) yields $\psi_t^b - \psi_t^s$. If this is positive, $M_t = \infty$ can not be the solution to (49). ■

A.3 Proof Theorem 3

A preliminary step used in proving theorem 3 is given in Lemma 1.

Lemma 1. $\Theta(q)$ and $\theta^k(q)$ defined in theorem 3 are continuous, strictly concave on $q \in (0, 1)$ and satisfy $\Theta(0) = \Theta(1) = 0$, $\theta^k(0) = \theta^k(1) = 0$.

Proof Lemma 1. Let $q_{j_1 j_2 \dots j_k}^m \equiv \max[b_{j_1}(q), b_{j_2 j_1}(q), \dots, b_{j_k j_{k-1} \dots j_1}(q)]$. By Baye's Rule, $q_{j_1 j_2 \dots j_k}^m = \frac{q P_2}{q P_2 + (1-q) P_1}$ for some $0 < P_2 < 1$ and $0 < P_1 < 1$. Furthermore, $q_{j_1 j_2 \dots j_k}^m \geq q$ if and only if $P_2 \geq P_1$, with equality at $P_2 = P_1$. In the k -fold sum of $\theta^k(q)$, any $q_{j_1 j_2 \dots j_k}^m$ that exceed q ($P_2 > P_1$) contribute 0 to the value of $\theta^k(q)$. Therefore, for any non-zero values of $(q - q_{j_1 j_2 \dots j_k}^m)$ in the k -fold sum of $\theta^k(q)$, $P_2 < P_1$. For any P_2 and P_1 satisfying $P_2 < P_1$, $q_{j_1 j_2 \dots j_k}^m$ is continuous with second derivative that is strictly negative for $q \in (0, 1)$ implying that $(q - q_{j_1 j_2 \dots j_k}^m)$ is continuous and strictly concave in q for $q \in (0, 1)$. Since $\theta^k(q)$ is the (weighted) sum of continuous, strictly concave functions on $q \in (0, 1)$, $\theta^k(q)$ is continuous and strictly concave on $q \in (0, 1)$. Since $\Theta(q)$ is also the (weighted) sum of continuous, strictly concave functions on $q \in (0, 1)$, it is also continuous and strictly concave on $q \in (0, 1)$ and satisfies $\Theta(0) = \Theta(1) = 0$. ■

Proof Theorem 3. Using the definitions of $\Theta(q_{ss}^*)$ and $\theta^k(q_{ss}^*)$ from theorem 3, it is possible to show that ϵ_{ss} satisfies

$$\epsilon_{ss} \equiv \{1 - q_{ss}^* - (1 - \delta)\Theta(q_{ss}^*)\}v_h M_{ss} \quad (50)$$

By condition 1) in the definition of steady-state equilibrium, $\epsilon_{ss} = \delta N_{ss}$ implying that

$$N_{ss}^f = \frac{1}{\delta} \{1 - q_{ss}^* - (1 - \delta)\Theta(q_{ss}^*)\}v_h M_{ss} \quad (51)$$

If $\frac{1-q_{ss}^*}{1-\delta} \leq \Theta_{ii}(q_{ss}^*)$ then $N_{i,ss}^f$ is negative or infinite. Hence a steady-state equilibrium does not exist under such a parameterization of the model. If $\frac{1-q_{ss}^*}{1-\delta} > \Theta_{ii}(q_{ss}^*)$ then $N_{i,ss}^f$ is strictly positive for $i = 1, 2$. By Lemma 1, $\Theta_{ii}(q)$ crosses the line $\frac{1-q}{1-\delta}$ at most once on $q \in [0, 1]$. Therefore, all that remains is to solve for q_{ss}^* and M_{ss} . Since A_{ss} can, without loss of generality, be normalized to unity by appropriate choice of parameters, it is possible to compute w_{ss} from the labor market clearing condition without first calculating N_{ss}^f . By definition, $\tilde{V}(q; w_{ss})$ also does not depend on N_{ss}^f so long as w_{ss} is determined by the normalization $A_{ss} = 1$. Therefore, by theorem 2, a unique q_{ss}^* exists satisfying $\tilde{V}(q_{ss}^*, w_{ss}) = \psi^s$. And M_{ss} can be determined by $M_{ss} = \frac{1}{\text{frac} \{1 - q_{ss}^* - (1 - \delta)\Theta(q_{ss}^*)\}v_h \bar{a}}$. Furthermore, ψ^b is the free parameter that normalizes $A_{ss} = 1$. Consequently, the expression for ψ^b that achieves this normalization is

$$\psi^b = q_{ss}^* \psi^s + \alpha^{\frac{\alpha}{\alpha-1}} N_{ss}^\alpha \Pi_x(q_{ss}^*) + \beta(1 - \delta) \int_{q_{ss}^*}^1 E(V_{ss}(q)) dq \quad (52)$$

In (52) the expression $E(V_{ss}(q))$ denotes the one-period hence expected value of a plant with prior q in steady state and is calculated as $\sum_{j=1}^J \pi_j(q) V(b_j(q); w_{ss})$. ■

B Data and Methodology for TFP Data

Figures 1 and 4 are based on annual data on sectoral output and factor inputs collected and analyzed by Jorgenson *et al.* (1987) (the **KLEM** dataset). The basic accounting methodology used here is that of Solow (1957) as updated by Hall (1988, 1990) and Basu and Fernald (1998). Under an assumption of constant returns to scale⁵⁴, TFP change in a sector is constructed as the difference between a Divisia index of output change, dy_t and a Divisia index of (weighted⁵⁵) factor inputs change dx_t . Data on sectoral sales and prices is used to construct dy_t while data on sectoral uses of capital, labor, energy and materials and their respective prices is used to construct dx_t .⁵⁶ A measure of aggregate productivity variation can be constructed from sectoral productivity variation as the weighted sum of sectoral TFP change with weights constructed from the output share (of total output) represented by each sector.

C Derivation of (27)

First, with the assumption of state independence at the plant level implies that the Bayesian posterior function (9) becomes

$$b_j(q) = \frac{q\pi_{2j}}{q\pi_{2j} + (1-q)\pi_{1j}} \quad (53)$$

and equation (20) becomes

$$\rho_t^i(q) = \sum_{j=1}^J \pi_{ij} \lambda_{t-1}^i(g_j(q)) \quad i = 1, 2 \quad . \quad (54)$$

From (20)-(21) and (25), and letting $\tilde{q}_t^{m_1} = \min_j [q_t^*, b_j(q_{t-1}^*)]$ it is possible to write ϵ_t as

$$\epsilon_t = (1 - q_t^*)M_t v_h - (1 - \delta)^2 \sum_{j=1}^J \pi_j^d \int_{\tilde{q}_t^{m_1}}^{q_t^*} \rho_{t-1}(g_j(q)) dq - (1 - \delta)\theta_t^1 v_h M_{t-1} \quad , \quad (55)$$

where π_j^d is a diagonal matrix with $(\pi_{1j} \ \pi_{2j})'$ on its diagonal, $v_h \equiv ((1 - h) \ h)'$, and

$$\theta_t^1 = \sum_{j=1}^J \pi_j^d \max[0, (q_t^* - b_j(q_{t-1}^*))] \quad (56)$$

In (55), integration of the vector $\rho_{t-1}(q) \equiv (\rho_{t-1}^1(q) \ \rho_{t-1}^2(q))'$ denotes element-by-element integration. The first term in (55) captures the mass of new entrants in period t , a fraction h of which are high productivity and a fraction $1 - h$ of which are low productivity. The second term in (55) corresponds to the mass of exit in period t comprised of plants that operated during any period $t - k$, $k = 2, 3, \dots$ and period $t - 1$ but whose priors and experiences led them to decide to exit in period t . The third term in (55) corresponds to exit in period t of plants that entered during period $t - 1$ but whose priors and experiences led them to decide to exit in period t . The term θ_t^1 represents the period t probability of exiting one period after entering, or the age-1 hazard rate in period t .

⁵⁴Results are similar allowing for non-constant returns to scale.

⁵⁵The analysis here uses revenue based shares for weighting each input. Cost based shares yield virtually identical results.

⁵⁶See Basu and Fernald (1998) for a full development of this methodology.

Continuing the recursive substitutions in (55) for $\rho_{t-1}(q)$ using (21), achieves the stated representation for ϵ_t

$$\epsilon_t = (1 - q_t^*)M_t v_h - \sum_{k=1}^{\infty} (1 - \delta)^k \theta_t^k v_h M_{t-k} \quad . \quad (57)$$

where

$$\begin{aligned} \theta_t^k \equiv & \sum_{j_1=1}^J \sum_{j_2=1}^J \dots \sum_{j_k=1}^J \pi_{j_1}^d \pi_{j_2}^d \dots \pi_{j_k}^d \\ & \times \max[0, \{q_t^* - \max[b_{j_1}(q_{t-1}^*), b_{j_2 j_1}(q_{t-2}^*), \dots, b_{j_k j_{k-1} \dots j_1}(q_{t-k}^*)]\}] \end{aligned} \quad (58)$$

In (58), $b_{j_k j_{k-1} \dots j_1}(q_{t-k}^*)$ represents the period t prior held by a plant whose period $t - k$ prior was q_{t-k}^* and which experienced productivities $\{A_{j_k}, A_{j_{k-1}}, \dots, A_{j_1}\}$ in periods $t - k, t - k + 1, \dots, t - 1$, respectively. Successive applications of Baye's Rule achieves

$$\begin{aligned} b_{j_k j_{k-1} \dots j_1}(q_{t-k}^*) & \equiv b_{j_1}(b_{j_2}(\dots(b_{j_{k-1}}(b_{j_k}(q_{t-k}^*)))) \\ & = \frac{q_{t-k}^* \pi_{2j_k} \pi_{2j_{k-1}} \dots \pi_{2j_2} \pi_{2j_1}}{q_{t-k}^* \pi_{2j_k} \pi_{2j_{k-1}} \dots \pi_{2j_2} \pi_{2j_1} + (1 - q_{t-k}^*) \pi_{1j_k} \pi_{1j_{k-1}} \dots \pi_{1j_2} \pi_{1j_1}} \end{aligned} \quad (59)$$

Fig. 1: Aggregate TFP and GDP: 1955:3–1984:1 (HP-Filter)
NBER Recessions Marked

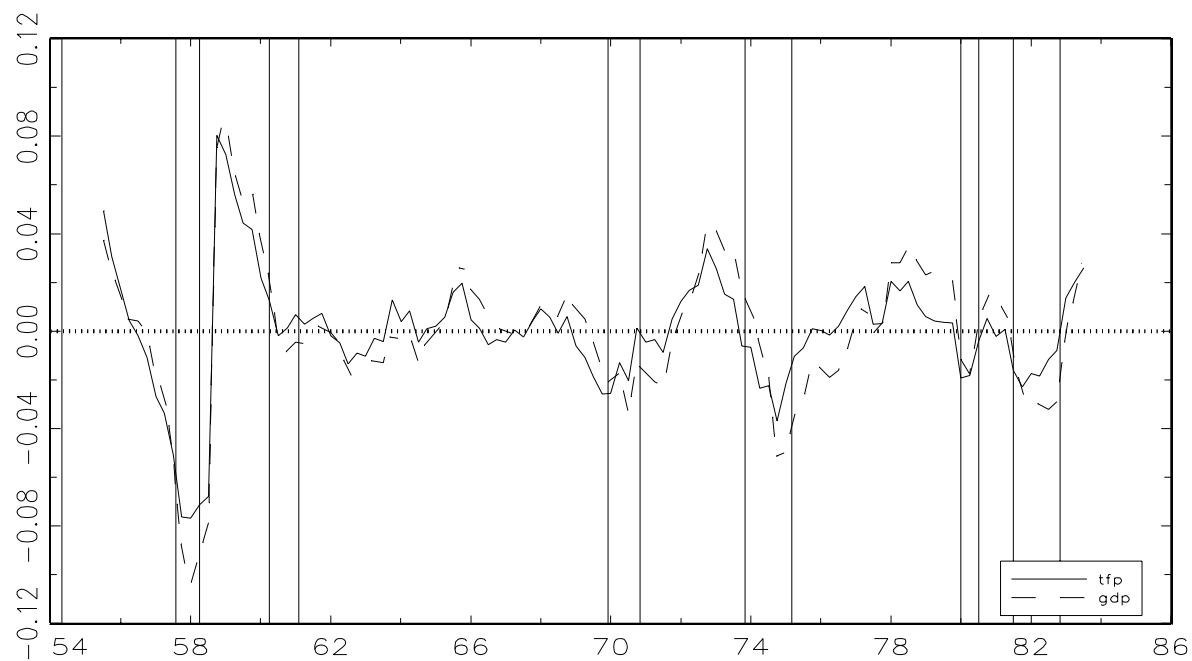
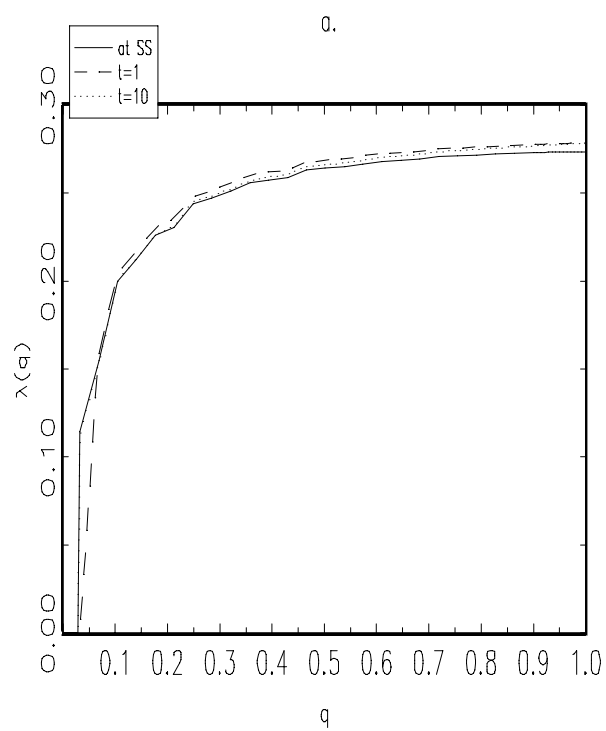
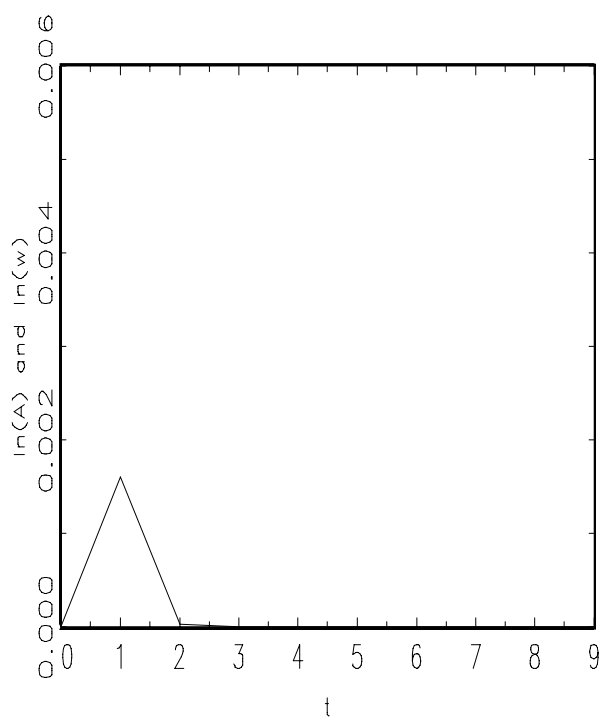


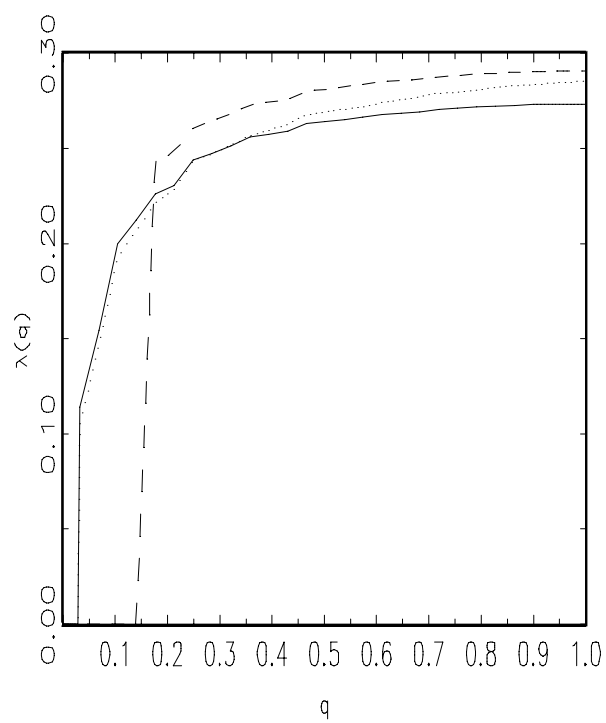
Fig 2a. Impulse Responses from Transitory Shock to ψ^S
Fast Learning, $\% \Delta \psi^S = .01$



Fast Learning, $\% \Delta \psi^S = .01$
b.



Fast Learning, $\% \Delta \psi^S = .05$
c.



Fast Learning, $\% \Delta \psi^S = .05$
d.

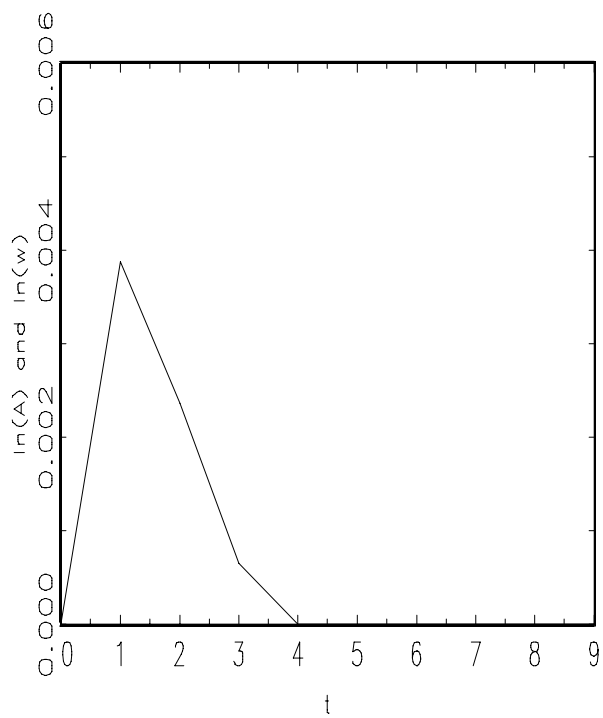
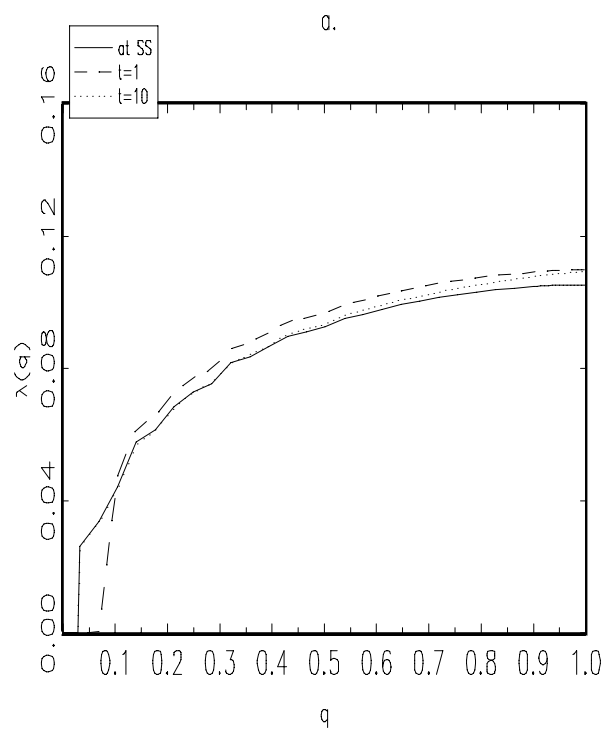
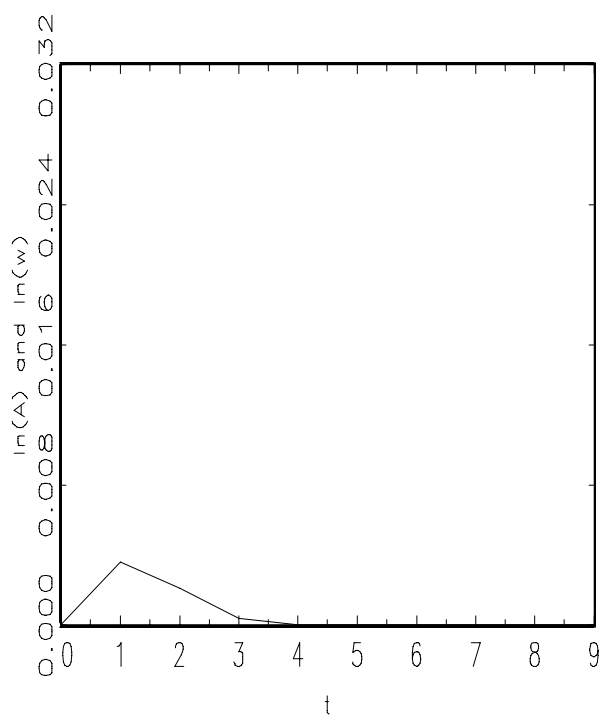


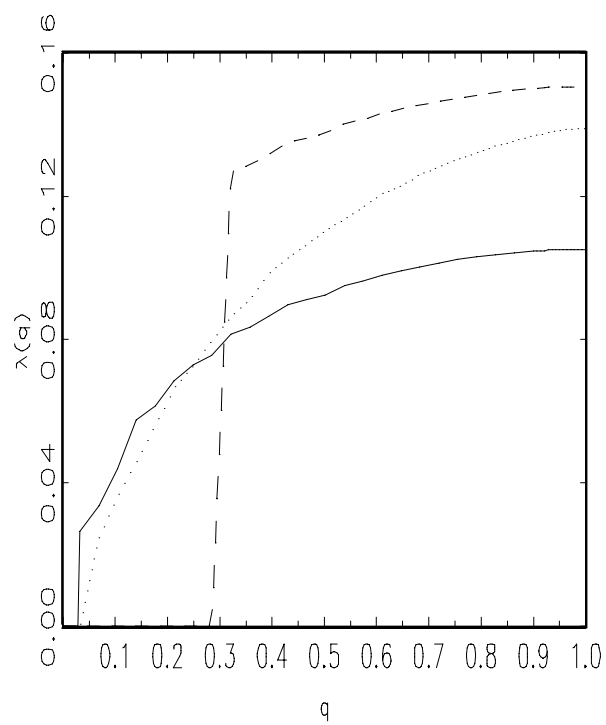
Fig 3a. Impulse Responses from Transitory Shock to ψ^s
Slow Learning, $\% \Delta \psi^s = .01$



Slow Learning, $\% \Delta \psi^s = .01$
b.



Slow Learning, $\% \Delta \psi^s = .05$
c.



Slow Learning, $\% \Delta \psi^s = .05$
d.

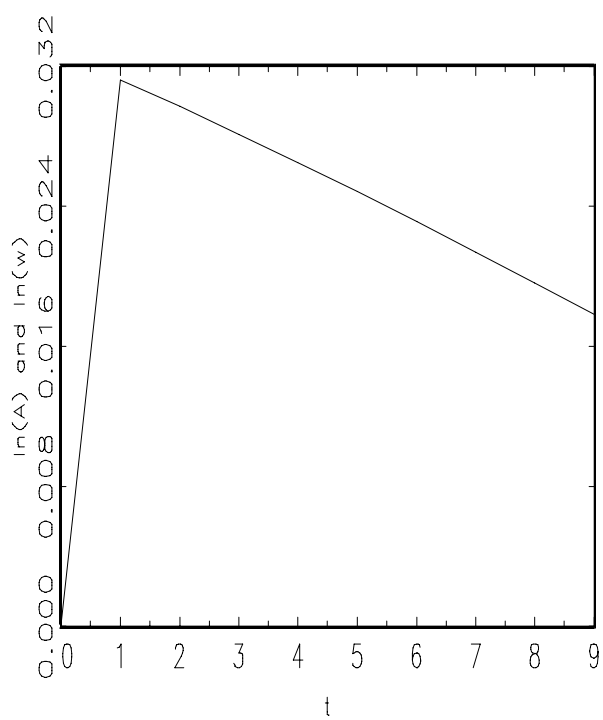


Fig. 4a: Total Employment, Establishments, and TFP
CBP (log) 1962–1996

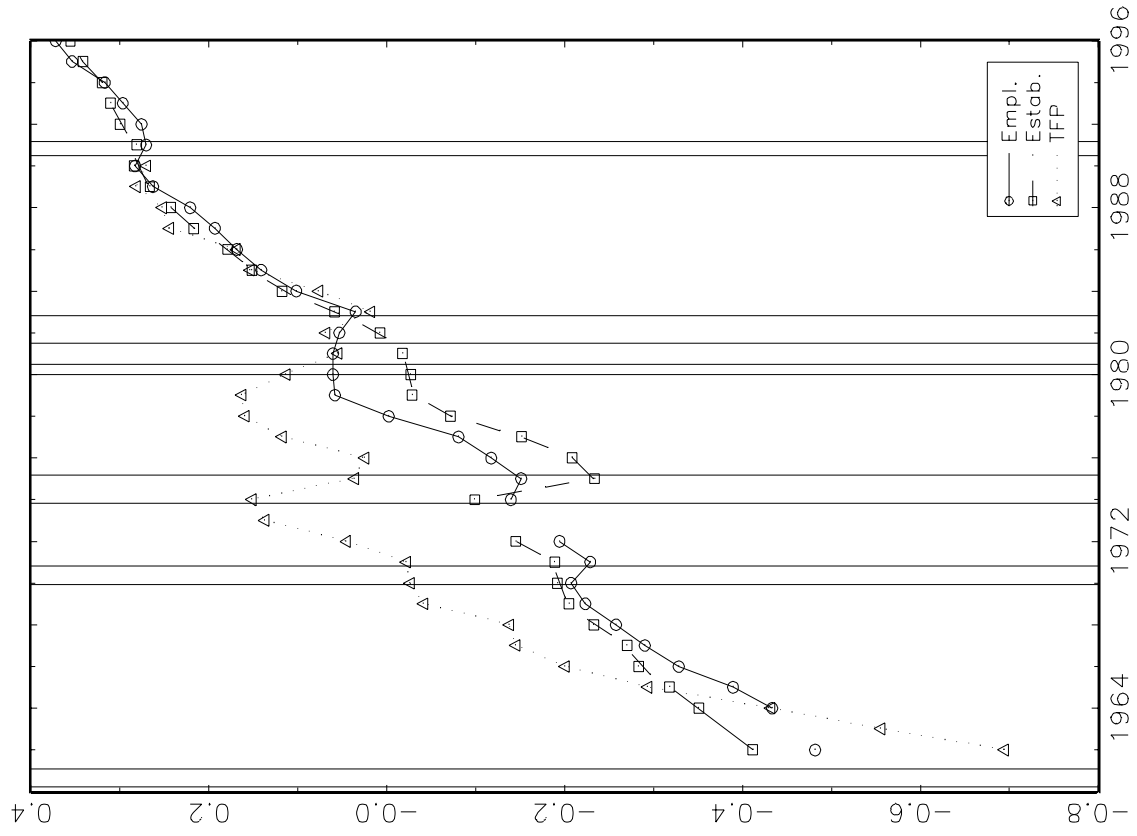


Fig 4b: Total Employment, Establishments, and TFP
CBP (HP) 1962–1996

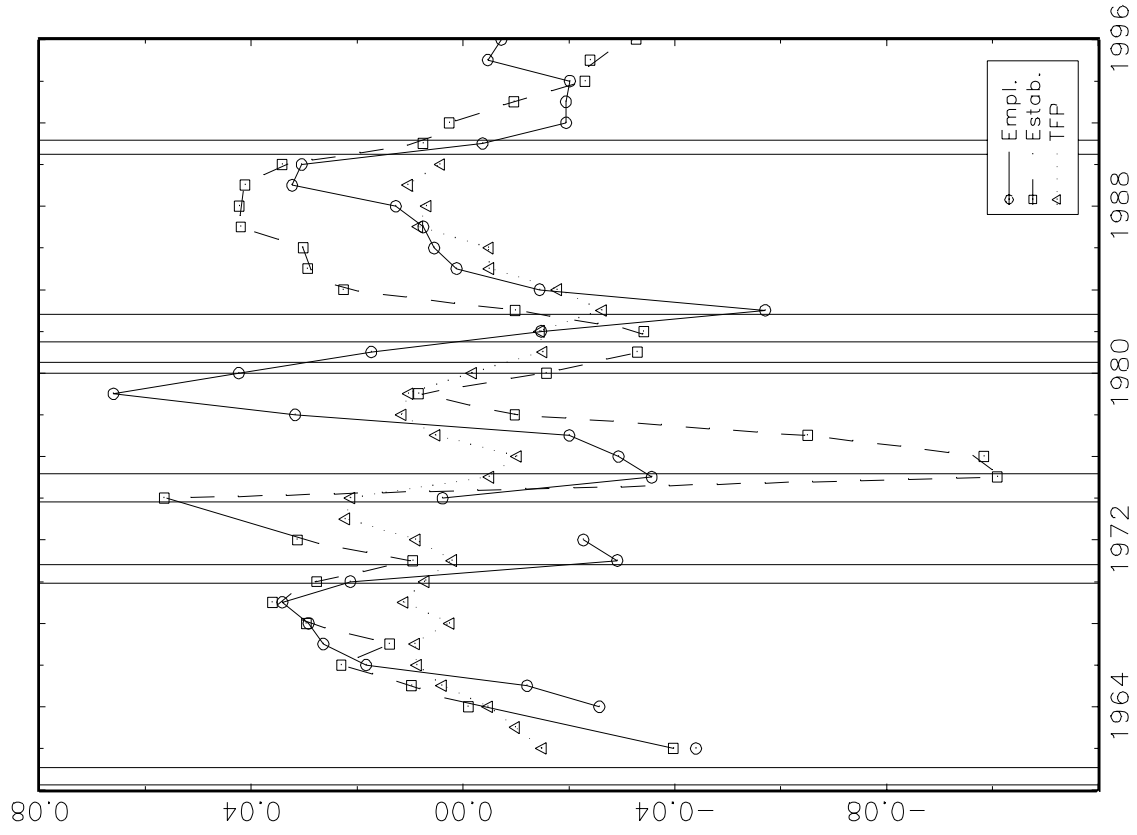
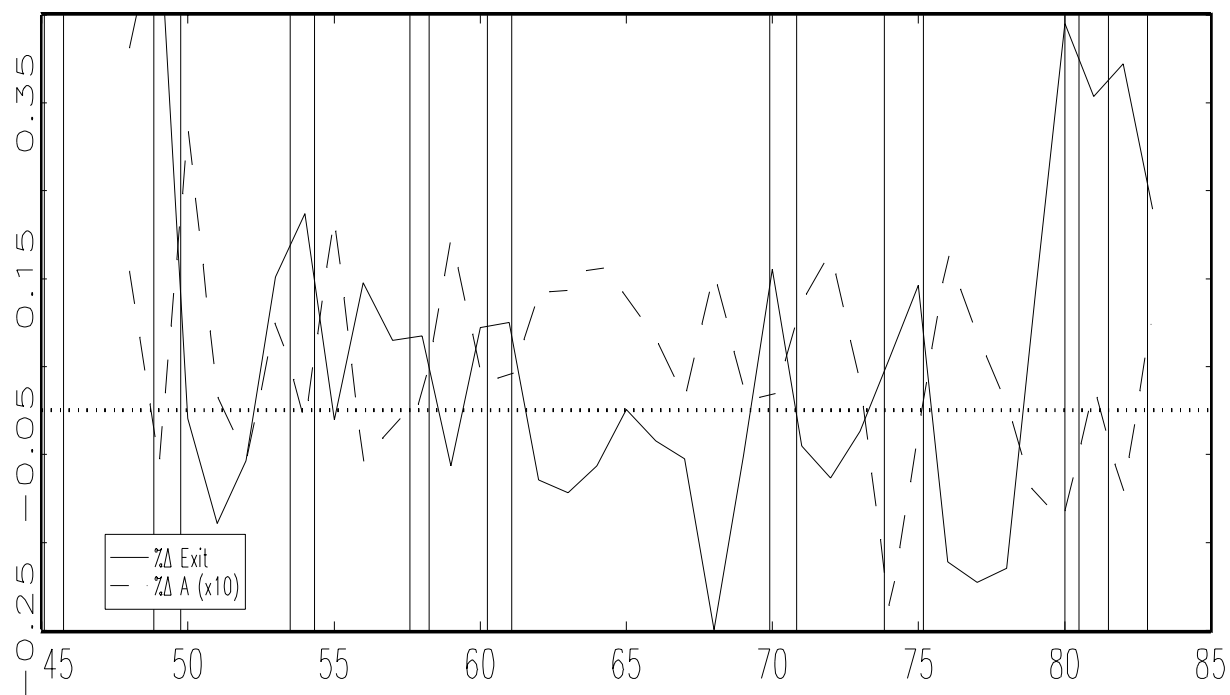


Fig 5a. % Δ Exits and Aggregate Solow Residual: D&B 1948–83



% Δ Entry and Aggregate Solow Residual: D&B 1948–83

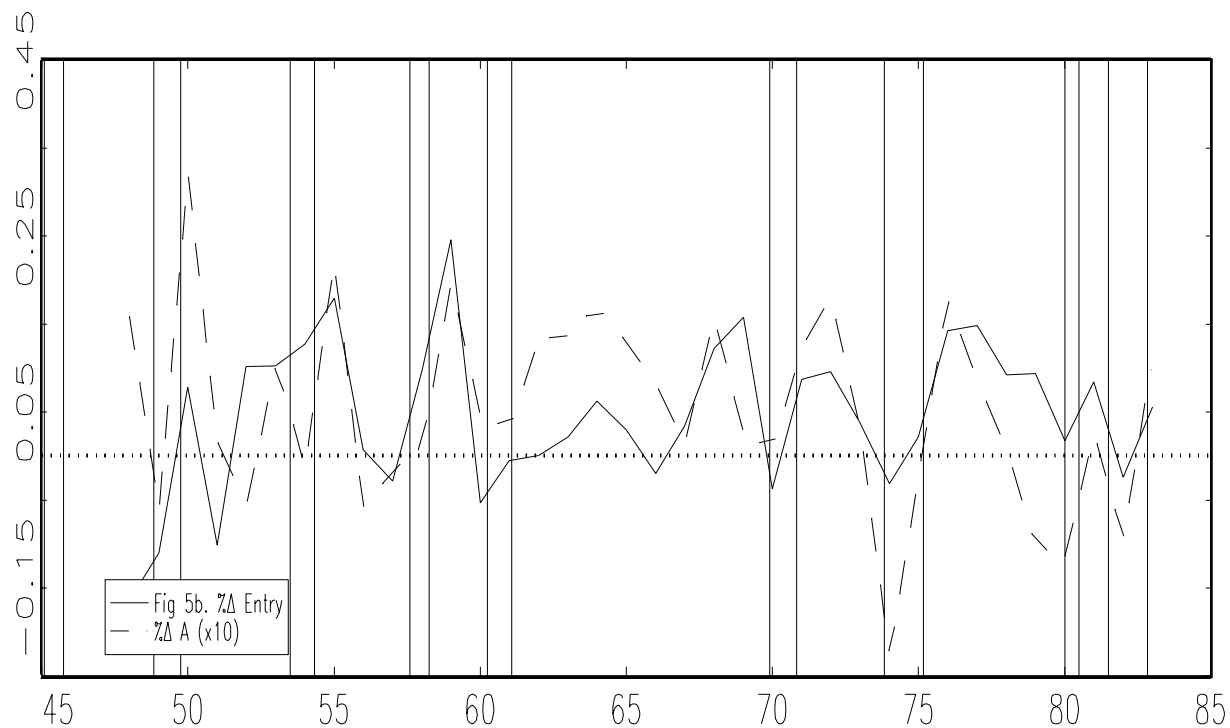
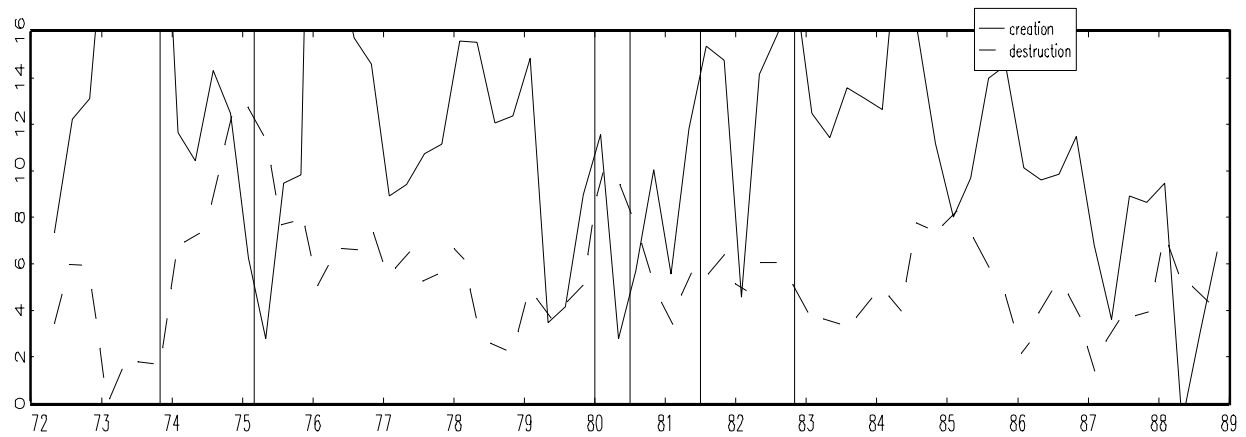
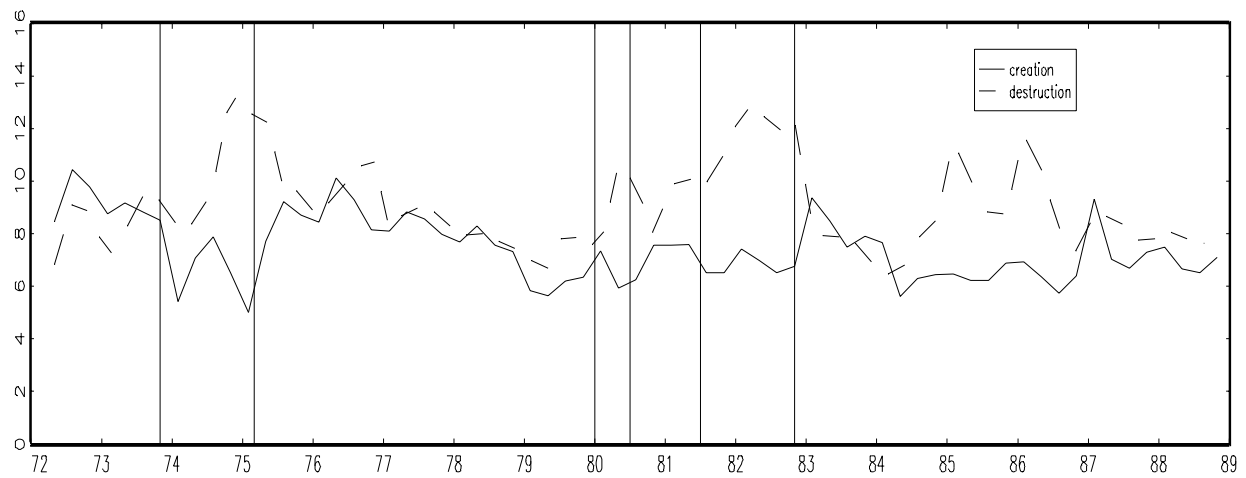


Fig 6: Job Creation/Destruction Rates: U.S. Manuf. LRD 1972–1988

a. Young Plants (all)



b. Middle Aged Plants (all)



c. Old Plants (all)

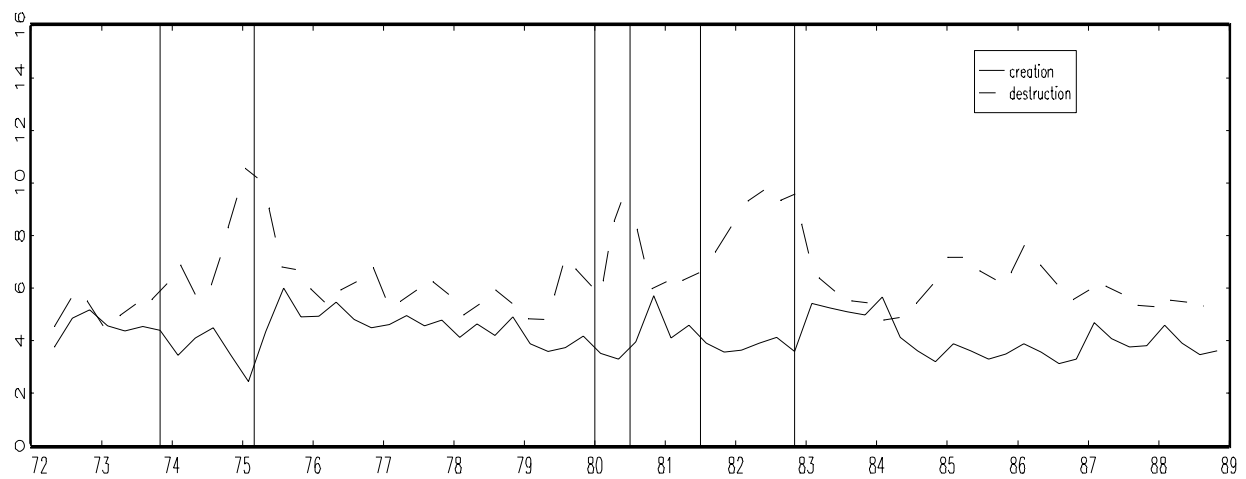
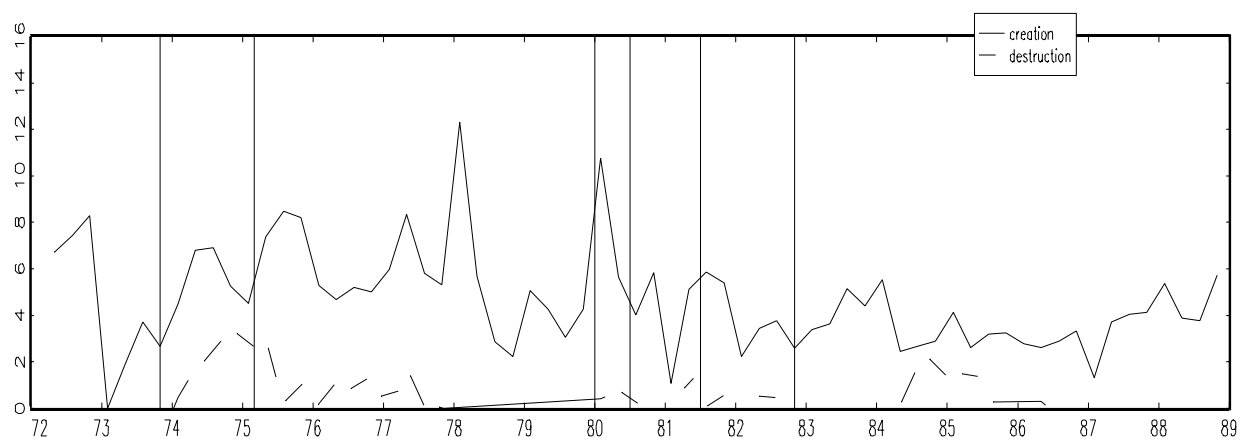
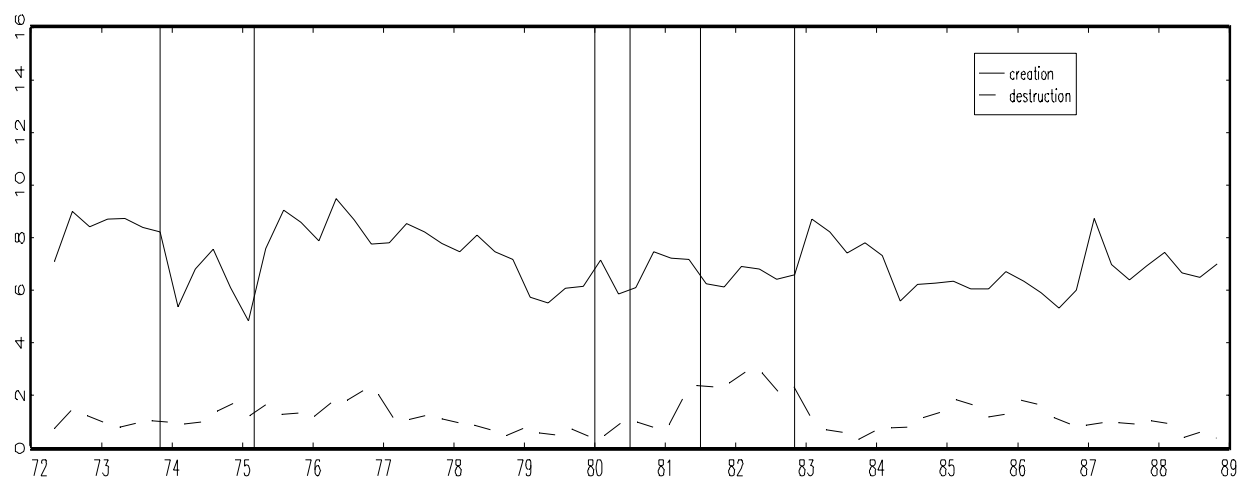


Fig 7: Job Creation/Destruction Rates: U.S. Manuf. LRD 1972–1988

a. Young Plants (transitioning)



b. Middle Aged Plants (transitioning)



c. Old Plants (transitioning)

