View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Digital Repository at the University of Maryland

ABSTRACT

Title of Dissertation / Thesis: RESOURCE REALLOCATION, PRODUCTIVITY DYNAMICS, AND BUSINESS CYCLES

Min Ouyang, Doctor of Philosophy, 2005

Dissertation / Thesis Directed By: Professor John Shea, Department of Economics
Professor John Rust, Department of Economics

This dissertation explores the interactions between resource reallocation, productivity dynamics and business cycles. A theory that combines two driving forces for resource reallocation, learning and creative destruction, is presented to reconcile several empirical findings of gross job flows. The theory suggests a scarring effect, in addition to the conventional Schumpeterian cleansing effect, of recessions on the allocation efficiency of resources. I argue that while recessions kill off some of the least productive businesses, they also impede the development of potentially good businesses -- the ones that might have proven to be efficient in the future are cleared out and lose the opportunity to realize their potential. Calibrations of the model using US manufacturing job flows suggest that the scarring effect is likely to dominate the cleansing effect and account for the observed pro-cyclical average productivity.

RESOURCE REALLOCATION, PRODUCTIVITY DYNAMICS, AND BUSINESS CYCLES

by

Min Ouyang

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park, in partial fulfillment of the requirements for the degree of Doctor of Philosophy
[2005]

Advisory Committee:

Professor John Shea, Chair
Professor John Rust, Co-chair
Professor John Haltiwanger
Professor John Horowitz
Professor Daniel Vincent

© Copyright by

[Min Ouyang]

[2005]

DEDICATION

To my family

Acknowledgments

I am grateful to John Shea, John Rust and John Haltiwanger, from whose instruction and encouragement I have gained immeasurably. I owe special thanks to John Shea for his particularly detailed comments that helped through this research, to John Rust for his C codes of multi-dimensional interpolation, and to John Haltiwanger for providing data on US manufacturing gross job flows. I am also indebted to the other members of my proposal and dissertation committees, Daniel Vincent, Michael Pries, and John Horowitz, for their advice and service. Comments from seminar participants at the department of Economics of University of Maryland and the Midwest Macroeconomics Meetings 2004 are greatly appreciated.

Contents

1 Resource Reallocation and Business Cycles **1**

2 Plant Life Cycle and Aggregate Employment Dynamics **8**

2.1 Introduction 8

2.2 Evidence: Gross Job flows and plant age 10

2.3 A Model of Learning and Creative Destruction 14

2.3.1 "All-Or-Nothing" Learning 16

2.3.2 Creative Destruction and Industry Equilibrium 19

2.4 Aggregate Employment Dynamics 23

2.4.1 Job Flows over the Plant Life Cycle 23

2.4.2 Decomposition of Job flows 25

2.4.3 The magnitude of job flows with no fluctuations 30

2.4.4 Cyclical job flows with industry fluctuations 33

2.5 Quantitative Implications 34

2.5.1 Calibration 34

2.5.2 Simulations of aggregate employment dynamics 36

2.6 Conclusion 45

3 The Scarring Effect of Recessions **46**

3.1 Introduction 46

3.2 A Renovating Industry with Learning 50

3.2.1 Firms 50

3.2.2 "All-Or-Nothing" Learning 52

3.2.3 The Recursive Competitive Equilibrium 56

3.3 Cleansing and Scarring 60
3.3.1 The Steady State 61
3.3.2 Comparative Statics: Cleansing and Scarring 63
3.3.3 Sensitivity Analysis 69
3.4 Quantitative Implications with Stochastic Demand Fluctuations . . . 73
3.4.1 Computational Strategy 74
3.4.2 Calibration 80
3.4.3 Response to a Negative Demand Shock and Simulations of U.S. Manufacturing Job Flows 84
3.4.4 Simulation of U.S. Manufacturing Job Flows 89
3.5 Sensitivity Analysis of the Dominance of Scarring over Cleansing . . . 90
3.6 Cost of Business Cycles with Heterogeneous Firms 93
3.7 Conclusion 97

4 Appendix **99**

1 Resource Reallocation and Business Cycles

Ever since the foundation of real business cycle theory in Kydland and Prescott (1982), the empirical regularities seen in productivity dynamics over business cycles have attracted a great amount of research attention. In recent years with longitudinal micro business data bases becoming more available, our understanding of aggregate productivity as well as its measurements have much improved.[1] We now know that the representative firm paradigm does not hold in the real world. As a matter of fact, economies across time and regions are characterized by a large and pervasive restructuring process due to entry, exit, expansion and contraction of businesses.[2] This gives the economy great flexibility and potentially allows economic resources to be used where they will be most productive. Businesses that use outdated technologies, or produce products flagging in popularity, experience employment decreases. And the displaced workers can then be re-employed by entrants or businesses that are expanding. Davis and Haltiwanger (1999) document that, in the U.S., roughly thirty percent of productivity growth over a ten-year horizon is accounted for by more productive entering businesses displacing less productive exiting ones.

A body of literature has arisen attempting to empirically synthesize the microeconomic and macroeconomic patterns of reallocation.[3] Much of them have centered on the creation and destruction of jobs, defined by Davis, Haltiwanger and Schuh (1996) (hereafter DHS) as gross job flows. A key stylized fact in this literature is that job reallocation exceeds that necessary to implement observed net job growth.

[1]The most heavily examined one is the Longitudinal Research Data (LRD) provided by U.S. Census of Bureau.

[2]Davis and Haltiwanger (1999) report that in most western economies roughly 1 in 10 jobs is created and 1 in 10 jobs is destroyed every year.

[3]Due to data limitations, most of the evidence comes from the manufacturing sectors.

This implies that jobs are continually being reallocated across businesses within the same industry. DHS document that this is true even when looking at very narrowly defined industries (four-digit) within specific geographic regions.[4] Hence, the large and pervasive job flows seem to reflect businesses' idiosyncratic characteristics and the resulting heterogeneity in their individual labor demand.

This dissertation is an attempt at providing a theoretical framework with heterogeneous businesses that relates resource reallocation to productivity dynamics over business cycles. I combine two driving forces for job flows – learning and creative destruction. There has been a long tradition in the profession of examining each force separately. The idea of creative destruction traces back to Schumpeter (1942), and has been formalized into a class of vintage models by Caballero and Hammour (1994 and 1996) and Campbell (1997).[5] Firm learning, originated by Jovanovic (1982), can be seen in Ericson and Pakes (1995) and more recently in Moscarini (2003) and Pries (2004).

Both theories on their own can match some of empirical evidence, but not all. The vintage models of creative destruction assume that new technology can only be adopted by constructing new businesses, so that technologically sophisticated businesses enter to displace older, out-moded ones. This is supported by the fact, as documented by DHS, that entry and exit of businesses account for a large fraction of job reallocation. However, while holding some appeal, this prediction runs counter to

[4] Davis and Haltiwanger (1999) document that, employment shifts among the approximately 450 four-digit industries in the U.S. manufacturing sector account for a mere 13% of excess job reallocation. Simultaneously cutting the U.S. manufacturing data by state and two-digit industry, region, size class, age class and ownership type, between-sector shifts account for only 39 percent of excess job reallocation. The same finding holds up in studies for other countries(e.g. Nocke 1994).

[5] Another important paper that formalizes Schumpeterian idea is Aghion and Hawitt (1992). They develop a theoretical model in which endogneous innovations drive creative destruction and growth.

the prevalent findings that failure rates decrease sharply with business age (Dunne, Roberts, and Samuelson 1989), and that productivity rises with business age (Aw, Chen and Roberts 1997, Jensen, McGuckin and Stiroh 2000). The learning models formalize the idea that businesses learn over time about initial conditions relevant to success and business survival. As learning diminishes with age, its contribution to job flows among businesses in the same birth cohort decreases. While providing an appealing interpretation of the strong and pervasive negative relationship between employer age and the magnitude of gross job flows, the learning models fail to explain the large gross job flows among mature businesses. Moreover, neither learning nor creative destruction alone can link business age with relative volatility of job destruction to creation, while these two have displayed a positive relationship in U.S. Manufacturing.[6]

In Chapter Two, I show that the empirical findings above can be potentially reconciled by a model that combines learning with creative destruction. I focus on two salient facts of gross job flows: the first is that young plants display greater turnover rates than old plants; the second is that, although job destruction is more volatile than job creation in general, this asymmetry is weaker for younger plants. I then present a framework where two forces interact together to drive micro-level job flows: creative destruction reallocates labor into technologically more advanced production units; while learning leads labor to production units with good businesses. With demand fluctuations, learning generates relative symmetric responses of creation and destruction, while the creative destruction force makes job destruction more responsive. Since old businesses are surer about their true idiosyncratic productivity,

[6] DHS document that in U.S. Manufacturing, creation and destruction are almost equally volatile for young businesses while old businesses features more volatile destruction.

the learning force weakens with age. Hence, Chapter Two interprets the observed cyclical pattern of job flows as the dominance of learning for young businesses and the dominance of creative destruction for old ones. I use the model to assess job-flow magnitude over a business's life cycle analytically and calibrate the model to match the data quantitatively. Calibration results show that my model does well in matching young businesses' higher job-flow magnitude as well as their relative symmetric volatility of job creation and destruction. However, it cannot fully account for the magnitude of job flows among mature businesses because of my assumption of a simplified all-or-nothing form of learning.

Chapter Three takes a further step to relate cyclical resource reallocation to cyclical productivity dynamics. It explores how recessions affect the allocative efficiency of resources and hence the average productivity. The conventional cleansing view argues that recessions promote more efficient resource allocation by driving out less productive units and freeing up resources for better uses. Using Chapter Two's framework of learning and creative destruction, I posit that recessions create an *additional* "scarring" effect by reducing the learning opportunities of "potentially good firms." I show that as a recession arrives and persists, the reduced profitability truncates the process of learning, limits the realization of truly good firms, and thus pulls down average labor productivity. Calibrating my model using data on job flows from U.S. manufacturing sector, I find that the scarring effect is likely to dominate the conventional cleansing effect, and can account for the observed procyclical average productivity.

To be consistent with the evidence provided by DHS, in Chapter Two I call the basic production unit underlying gross job flows "a plant", which refers to a physical location where production takes place. DHS argue that the plant represents the finest level of disaggregation available in the Longitudinal Research Data for calculating job creation and destruction statistics. To fit into the literature of industry organization, in Chapter Three a production unit is called "a firm". Chapter Three also features a simpler version of the model by assuming each firm employs only one worker, so that a job is created when a firm enters and a job is destroyed when a firm exits. Under this set-up, learning does not drive job creation but still does promote productivity growth. Since Chapter Three focuses on productivity dynamics, this simplification does not affect my main results.

Different theories of job flows have existed in the literature. One branch of theories emphasize the matching of employees and employers (see Mortensen and Pissarides 1994). In their analysis, job destruction is more volatile than job creation because, although job destruction takes place instantaneously, job creation can not due to the time-consuming matching process. Another branch focuses on nonconvex adjustment costs (see Caballero 1992, Campbell and Fisher 2000). In their environment, the cross-sectional distribution of production units, in terms of where they stand relative to their adjustment thresholds, may yield asymmetries in the cyclical dynamics of job creation and destruction. Foote (1997) gives another interesting story regarding the relative volatility of job creation and destruction. He connects (S, s) idiosyncratic productivity adjustments with trend employment growth and predicts that a growing industry features a more responsive creation margin while a declining industry a more responsive destruction margin. Nevertheless, none of these theories have incorporated the observed strong and persistent relationship between *business age* and job-flow patterns. My theory in Chapter Two builds on this relationship. Campbell and Fisher (2004) also links business age with job-flow volatility by modeling substitution between structured and unstructured jobs over a plant life cycle. While it is difficult to define structured and unstructured jobs empirically, their work does not feature the observed pro-cyclical entry rate and counter-cyclical exit rate. Chapter Two shows that these patterns are present in the cyclical response of my model.

Chapter Three posits that, with a scarring effect pulling down productivity by limiting businesses' learning opportunities, the observed intense reallocation during recessions may contribute to the procyclical behavior of productivity. Barlevy (2002) proposes a different story. He argues that during recessions, workers are more likely to stuck in mediocre matches with reduced worker flows, so that fewer high quality matches are created. Besides resource reallocation, the literature have provided other explanations for the cyclical behavior of productivity including cyclical technological shocks, increasing returns to scale, and factor utilization. Basu (1996) empirically investigate their merits using a panel on US manufacturing inputs and outputs from 1953 to 1984, and highlights the relative importance of factor utilization. But with industry-level data, he cannot assess the contribution of cross-business resource reallocation. Baily, Bartelsman, and Haltiwanger (2001) provide a more disaggregated empirical exploration. They find productivity to be more procyclical at the plant level than at the industry level, and posit that *short-run* reallocation yields a countercyclical contribution to productivity. However, as I elaborate in Chapter Three, the cleansing effect takes place immediately while the scarring effect takes place gradually. Can reallocation in the *longer run* yield a procyclical contribution to productivity with the stronger dominance of the scarring effect? This remains an interesting empirical question.

This dissertation tries to theoretically improve our understanding of the link between resource reallocation and productivity dynamics over business cycles. I hope it will be a base for future research that looks more intensively into this direction. There are many potential connections that have yet to be fully explored. For instance, Kydland and Prescott (1982) argue that a representative-agent real business cycle model with technological shocks can account for most of the observed aggregate fluctuations. However, later empirical work by Basu (1997) suggests that the technological residual interacts very little with output and input sequences once we control for increasing returns, cyclical utilization and resource reallocation. Can a heterogeneous-agent model with resource reallocation reconcile these papers by showing that the cyclical resource reallocation is a natural response of the economy to technological shocks? I believe there are many benefits to be gained from answering this question. The resulting findings will undoubtedly allow economists to learn more about the sources and consequences of business cycles.

2 Plant Life Cycle and Aggregate Employment Dynamics

2.1 Introduction

Research on aggregate employment dynamics has focused on two separate components: the number of jobs created at expanding and newly born plants (job creation) and the number of jobs lost at declining and dying plants (job destruction). A key stylized fact in this literature is that patterns of job creation and destruction differ significantly by plant age. In magnitude, job flow rates are larger for younger plants; In cyclical responses, job destruction varies more over time than job creation for old plants; but for young plants, their variations are much more symmetric.

This chapter proposes an explanation. I highlight the following relative advantages and disadvantages of young and old plants in market competition. Intuitively, old plants tend to be more productive since they have survived long; but they may be using out-dated technologies or producing products flagging in popularity. On the contrary, young plants, although lacking market experience, are more likely to be technologically updated. If these are true over the plant life cycle and plants' employment positively depends on their productivity, then there are multiple margins for a plant to create or destroy jobs as it ages. Cyclical aggregate employment dynamics involve the interactions of these difference driving forces.

I embody this intuition in an industry model whose employment dynamics are driven by two forces – learning and creative destruction. In my model, technology grows exogenously over time. Only entrants have access to the most advanced technology. Plants enter the market with the leading technology, but differ in idiosyncratic productivity. A plant's idiosyncratic productivity is not directly observable, but can be learned over time. A plant increases its employment (creates jobs) when it learns its true idiosyncratic productivity as high (a good plant); it exits (destroys jobs) when learning its true idiosyncratic productivity as low (a bad plant). Meanwhile, as newly born plants continually enter with more advanced technology, incumbents becomes more and more technologically outdated. They tend to destroy jobs and eventually leave the market at a certain age. This gives rise to a creative destruction process that allows technologically more advanced entering plants to replace outdated ones.

The resulting employment dynamics match the observed *magnitude* of job flows over the plant life cycle. Because learning diminishes with age, job creation and destruction decline with plant age; while large job flows still exist among mature plants with outdated plants being replaced due to creative destruction. The model also matches the observed *cyclical* pattern of job flows with plant age. The learning force generates relative symmetric responses to business cycles on the creation and destruction sides, while the creative destruction force makes job destruction more responsive. Since learning diminishes with plant age, the symmetric response of learning dominates for young plants and the asymmetric response of creative destruction dominates for old plants. Therefore, my model suggests that the variance ratio of job destruction over job creation increases with plant age, as shown in the data.

Other work has also studied the sources and macroeconomic implications of the relative variance of job creation and job destruction. Foote (1997) connects (S,s) idiosyncratic productivity adjustments with trend employment growth and predicts a tight relationship between trend growth and volatility of creation relative to destruction. In his analysis, a growing industry features a more responsive creation margin while a declining industry a more responsive destruction margin. Although his model succeeds in explaining the differences in relative gross-flow volatility across sectors, it cannot account for the high volatility of destruction within manufacturing or the high volatility of creation within service sector. This paper differs from Foote's work by emphasizing the within-sector differences in relative gross-flow volatility arising from plant-level heterogeneity. Campbell and Fisher (2004) link plant age with relative volatility of creation and destruction by modeling substitution between structured and unstructured jobs over a plant life cycle. While it is difficult to define structured and unstructured jobs empirically, their work does not feature the observed pro-cyclical entry rate and counter-cyclical exit rate. These patterns are present in the cyclical response of my model.

The remainder of this chapter is organized as follows. In the next section, I describe the differences in young and old plants' employment dynamics. In Section 3, I present my model, with which I analytically analyze the job flows over the plant life-cycle in Section 4. A calibrated version of the model is studied numerically in Section 5. I conclude in section 6.

2.2 Evidence: Gross Job flows and plant age

This section describes the observations of employment dynamics over the producer's life cycle that motivate my theory of learning and creative destruction. Two salient facts emerge from the analysis carried out by Davis and Haltiwanger (1999). The first is that young plants display greater turnover rates than old plants. The second is that, although job destruction is more volatile than job creation in general, this asymmetry is weaker for younger plants.

A. Means						
Plant type	$E(Cb)$	$E(Cc)$	$E(C)$	$E(Dd)$	$E(Dc)$	$E(D)$
all	0.42	4.77	5.20	0.64	4.89	5.53
young	1.52	6.00	7.52	1.24	5.33	6.56
old	0.12	4.42	4.54	0.47	4.77	5.24
B. Standard deviations						
Plant type	$\sigma(Cb)$	$\sigma(Cc)$	$\sigma(C)$	$\sigma(Dd)$	$\sigma(Dc)$	$\sigma(D)$
all	0.26	0.78	0.89	0.23	1.50	1.66
young	1.06	1.23	1.80	0.66	1.67	2.07
old	0.07	0.78	0.78	0.22	1.50	1.60

C. Variance ratio of job destruction to creation		
plant type	$\sigma(D)^2/\sigma(C)^2$	$\sigma(Dc)^2/\sigma(Cc)^2$
all	3.49	3.64
young	1.32	2.80
old	4.18	3.69

Table 1: Quarterly gross job flows from plant birth, plant death, and continuing operating plants in the US manufacturing sector: 1973 II to 1988 IV. Cb denotes job creation from plant birth, Dd job destruction from plant death, Cc and Dc job creation and destruction from continuing operating plants. C and D represent gross job creation and destruction. C=Cc+Cb, D=Dd+Dc. All numbers are in percentage points.

My data source is DHS's observations of job creation and destruction rates for the US manufacturing sector. For a given population of plants, the job creation rate in a period is defined as the total number of jobs added since the previous period at plants that increased employment, divided by the average of total employment in the current and previous periods. The job destruction rate is similarly defined in terms of employment losses at shrinking plants. The difference between job creation and destruction is the rate of job growth. As proposed by DHS, the sum of job creation and destruction rates is used as a measure of job reallocation across plants.

For my comparison of young and old plants' employment dynamics, I use DHS's quarterly job creation and destruction series for plants in three different age categories.

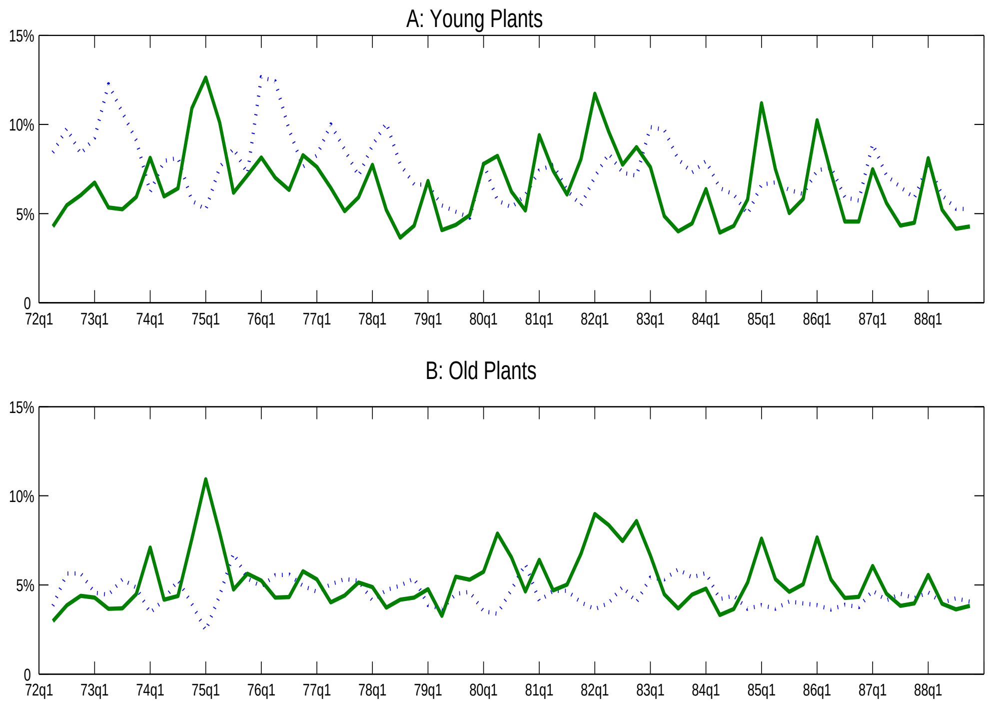


Figure 1: Job flows at young and old plants, 1972:2 – 1988:4. Dashed lines represent the job creation series; solid lines represent job destruction.

As recommended by DHS,[7] I aggregate the two categories that include the youngest plants and refer to this combination as "young". These plants are usually less than 10 years old and account for 22.5% of total employment on average over the sample period. The remaining are referred to as "old".[8]

Table 1A reports the sample means of the overall job creation (C), overall destruction (D), job creation from plant birth (Cb), job destruction from plant death (Dd), and job creation and destruction from continuing operating plants (Cc and

[7] See DHS, p.225.

[8] Because of the sample design, the threshold between young and old plants changes slightly over time. The minimum age of old plants is between 9 years and 13 years. See DHS, p. 225, for details.

Dc) for young and old plants separately, as well as for the US manufacturing sector as a whole.[9] Table 1B reports sample standard deviations. The sample covers the statistics from the second quarter of 1972 to the fourth quarter of 1988.

As shown in Table 1A and 1B, young plants' average job creation and destruction are both higher than those for old plants. So are the standard deviations. Table 1C reports the relative variability of job creation and destruction. For the US manufacturing sector as a whole, the variance ratio of job destruction to job creation equals 3.49, so job destruction fluctuates much more than job creation. The variance ratio is also considerably higher than one for old plants, 4.18. However, a more interesting finding, as noted by DHS, is that young plants' job creation and destruction rates have approximately equal variances.

Figure 1 reinforces the above impression of greater variability of job flows at young plants. It illustrates that job creation and destruction rates at young plants are visibly more volatile. Moreover, the time series variability of creation and destruction seems more *symmetric* for young plants than for old plants.

Because the observed frequency of plant exit declines with age and entering plants are young by definition, it is important to consider the possibility that young and old plants' different variances only reflect the concentration of entry and exit among young plants. If I exclude the contributions of plant birth and death to the job creation and destruction, the negative relationship between magnitude of job flows and plant age is still evident. As shown in Table 1A, the average job creation and destruction from

[9] Notice that average job creation from plant birth is positive even for old plants. It seems strange at the first sight, since by definition, an old plant cannot be newly born. DHS define plant birth and plant age differently. Plant age is caculated based on *the first time* a plant has positive employment, while plant birth is recorded as plants starting up. Most of the starting up plants age zero, but some old plants' employment can drop to zero temporarily and start to increase again. Also notice that the contribution of "old plant birth" to old plants' job creation is very small.

continuing young plants are higher than those of continuing old plants. So are the standard deviations shown in Table 1B. Table 1C suggests that, excluding plant birth and death, job destruction still varies more than job creation: the variance ratio of job destruction to creation is 2.80 for continuing young plants, and 3.69 for continuing old plants. Moreover, this asymmetry is weaker for continuing young plants, as it is for young plants' overall job flows that include the contributions of plant birth and death.

Table 1 and Figure 1 reveal a sharp relationship between plant age and job reallocation rates.[10] Davis and Haltiwanger (1999) report that this relationship exists in very narrowly defined (four-digit) manufacturing industries, even with detailed controls for size and other producer characteristics. This highlights the connection between a producer's life-cycle and its employment dynamics, which is modeled in the next section.

2.3 A Model of Learning and Creative Destruction

Consider an industry of plants that produce a single good for sale in a competitive product market. Plants use a single factor of production, labor, that they hire in a competitive labor market. I refer to an employee working one period as a job. Each plant can be thought of as an "institutional adoption of technology" with the following three characteristics:

1. vintage;

2. idiosyncratic productivity;

[10] Similar patterns have also been found in data on job flows in France, Canada, Norway, Netherlands, Germany and U.K. See Davis and Haltiwanger (1999).

3. a group of workers (jobs).

There is an exogenous technological progress $\{A_t\}_0^\infty$ with a positive growth rate γ so that:[11]

$$A_t = A_0 \cdot (1+\gamma)^t,$$

where A_0 is a constant. With a as plant age, apparently the vintage of a plant of age a in period t is,

$$A_{t-a} = A_0 \cdot (1+\gamma)^{t-a},$$

Assuming discrete time, each period a continuum of plants enter embodied with the latest technology. Incumbents do not have access to the latest technology.[12] so that young plants have technological advantages over old plants.

At the time of entry, a plant is endowed with idiosyncratic productivity θ, so that plants of the same vintage(age) cohort differ in idiosyncratic productivity. θ can represent the talent of the manager as in Lucas (1978), or alternatively, the location of the store, the organizational structure of the production process, or its fitness to the embodied technology.[13] The key assumption regarding θ is that its value, although fixed at the time of entry, is not directly observable.

[11] See Caballero and Hammour (1994).

[12] This serves as a short-coming of my theory, and of the whole branch of vintage models. Not allowing re-tooling, a business's technology level is fixed and even good plants exit eventually. But in the real world, some plants may stay long by keep updating their technology. DHS report that, in US manufacturing, a large fraction of labor is concentrated on a small number of *old* plants. Their finding is not present in my model, in which old plants tend to hire less labor due to out-moded technology. For a model that allows re-tooling, see Cooper, Haltiwanger, and Power (1999).

[13] Since a firm is identical to a job under this set-up, θ can also be interpreted as "match quality." See Pries (2004).

Production takes place through a group of workers. n_t represents the employment level of a plant in period t. The period-t output of this plant is given by

$$A_{t-a} \cdot X_t \cdot (n_t)^\alpha,$$

where α is between zero and one, and

$$X_t = \theta + \varepsilon_t.$$

The shock ε_t is an i.i.d. random draw from a fixed distribution that masks the influence of θ on output. I set the wage rate to 1 as a normalization, and let P_t denote the equilibrium output price in period t. Then the profit generated by a plant of age a and idiosyncratic productivity θ in period t equals $P_t \cdot A_{t-a} \cdot X_t \cdot (n_t)^\alpha - n_t$. Both output and profit are directly observable. Since the plant knows its vintage, it can infer the value of X_t. The plant uses its observations of X_t to learn about θ.

2.3.1 "All-Or-Nothing" Learning

Plants are price takers and profit maximizers. They attempt to resolve the uncertainty about θ to decide on an employment level and whether to continue or terminate production. The random component ε_t represents transitory factors that are independent of the idiosyncratic productivity θ. By assuming that ε_t has mean zero, I have $E_t(x_t) = E_t(\theta) + E_t(\varepsilon_t) = E_t(\theta)$.

Given knowledge of the distribution of ε_t, a sequence of observations of x_t allows the plant to learn about its θ. Although a continuum of potential values for θ is more realistic, for simplicity it is assumed here that there are only two values: θ_g for a good plant and θ_b for a bad plant. Furthermore, ε_t is assumed to be distributed uniformly on $[-\omega, \omega]$. Therefore, a good plant will have x_t each period as a random draw from a uniform distribution over $[\theta_g - \omega, \theta_g + \omega]$, while the x_t of a bad plant is drawn from an uniform distribution over $[\theta_b - \omega, \theta_b + w]$. Finally, θ_g, θ_b and ω satisfy $0 < \theta_b - \omega < \theta_g - \omega < \theta_b + \omega < \theta_g + \omega$.

Pries (2004) shows that the above assumptions give rise to an "all-or-nothing" learning process. With an observation of x_t within $(\theta_b + \omega, \theta_g + \omega]$, the plant learns with certainty that it is a good plant; conversely, an observation of x_t within $[\theta_b - \omega, \theta_g - \omega)$ indicates that it is a bad plant. However, an x_t within $[\theta_g - \omega, \theta_b + \omega]$ does not reveal anything, since the probabilities of falling in this range as a good plant and as a bad plant are the same (both equal to $\frac{2\omega + \theta_b - \theta_g}{2\omega}$).

This all-or-nothing learning simplifies my model considerably. Since it is θ^e instead of θ that affects plants' decisions, there are three types of plants corresponding to the three values of θ^e: plants with $\theta^e = \theta_g$, plants with $\theta^e = \theta_b$, and plants with $\theta^e = \theta_u$, the prior mean of θ. I define "unsure plants" as those with $\theta^e = \theta_u$. I further assume that the unconditional probability of $\theta = \theta_g$ is φ, and let $p \equiv \frac{\theta_g - \theta_b}{2\omega}$ denote the probability of the true idiosyncratic productivity being revealed every period. Hence a plant's life-cycle is incorporated into the model as follows. A flow of new plants enter the market as unsure; thereafter, every period they stay unsure with probability $1 - p$, learn they are good with probability $p \cdot \varphi$ and learn they are bad with probability $p \cdot (1 - \varphi)$. The evolution of θ^e from the time of entry is a Markov process with values

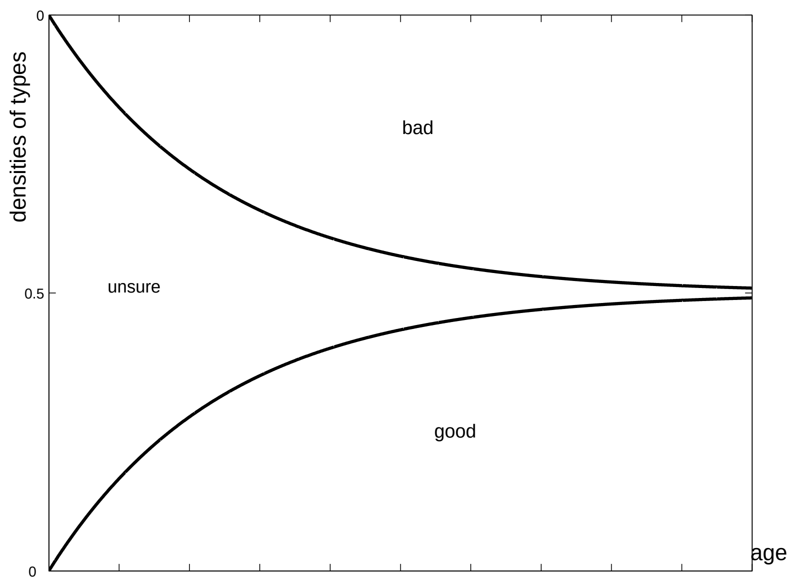


Figure 2: Dynamics of a Birth Cohort with Learning: the distance between the concave curve and the bottom axis measures the density of plants with $\theta^e = \theta_g$; the distance between the convex curve and the top axis measures the density of plants with $\theta^e = \theta_b$; and the distance between the two curves measures the density of unsure plants (plants with $\theta^e = \theta_u$).

$(\theta_g, \theta_u, \theta_b)$, an initial probability distribution $\begin{pmatrix} 0, & 1, & 0 \end{pmatrix}$,and a transition matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ p \cdot \varphi \quad , & 1-p \quad , & p \cdot (1-\varphi) \\ 0 & 0 & 1 \end{pmatrix}.$$

If plants were to live forever, eventually all uncertainty would be resolved because the market would provide enough information to reveal each plant's idiosyncratic productivity. The limiting probability distribution as a goes to ∞ is $\begin{pmatrix} \varphi, & 0, & (1-\varphi) \end{pmatrix}$.

Because there is a continuum of plants, it is assumed that the law of large numbers applies, so that both φ and p are not only the probabilities but also the fractions of unsure plants with $\theta = \theta_g$, and of plants who learn θ each period, respectively. Hence, *ignoring plant exit for now*, the densities of the three groups of plants in a cohort of age a as

$$\left(\varphi \cdot \left[1 - (1-p)^a\right], \quad (1-p)^a, \quad (1-\varphi) \cdot \left[1 - (1-p)^a\right] \right),$$

which implies an evolution of the idiosyncratic-productivity plant distribution within a birth cohort as shown in Figure 2, with the horizontal axis depicting the age of a cohort over *time*. The densities of plants that are certain about their idiosyncratic productivity, whether good or bad, grow as a cohort ages. Moreover, the two "learning curves" (depicting the evolution of densities of good plants and bad plants) are concave. This feature is defined as the decreasing property of marginal learning in Jovanovic (1982): the marginal learning effect decreases with plant age, which, in my model, is reflected by the fact that the marginal number of learners decreases with cohort age. The convenient feature of all-or-nothing learning is that, on the one hand, it implies that any single plant learns "suddenly", which allows us to easily keep track of the cross-section distribution of beliefs while, on the other hand, it still implies "gradual learning" at the cohort level.

2.3.2 Creative Destruction and Industry Equilibrium

I now turn to the supply and demand conditions in this model, and to the economics of creative destruction. I model a perfectly competitive industry in partial equilibrium. Plants of different vintages and beliefs may coexist. The mass of plants with age a and belief θ^e in period t is denoted $f_t(\theta^e, a)$.

The following sequence of events is assumed to occur within a period. First, entry and exit occur by observing the aggregate state and perfectly predicting the *current-period* price. Second, each surviving plant adjusts its employment and produces. Third, the aggregate price is realized. Fourth, plants observe revenue and update beliefs. Then, another period begins.

I assume costless employment adjustment each period so that a plant adjusts its employment to solve a *static* profit maximization problem. With θ^e as a plant's current belief of its idiosyncratic productivity and P_t as the equilibrium price, I denote firm's employment as $n_t(\theta^e, a)$. That is,

$$\begin{aligned} n_t(\theta^e, a) &= \arg\max_{n_t \geq 0} P_t \cdot A_{t-a} \cdot X_t \cdot (n_t)^{\alpha} - n_t \\ &= \left[\alpha \cdot P_t \cdot A_0 \cdot (1+\gamma)^{t-a} \cdot \theta^e\right]^{\frac{1}{1-\alpha}}. \end{aligned} \tag{2.1}$$

The corresponding expected value of the single-period profit maximized with respect to n_t is,

$$\pi_t(\theta^e, a; P, A) \equiv (\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}) \cdot \left[P_t \cdot A_0 \cdot (1+\gamma)^{t-a} \cdot \theta^e\right]^{\frac{1}{1-\alpha}}. \tag{2.2}$$

Let $W > 0$ be the expected present value of the plant's fixed factor (its "managerial ability" or "advantageous location") if employed in a different activity. The value of W is the same for all plants in the industry regardless of their vintages and idiosyncratic productivity. If a plant believes that the *expected* present discounted value of staying is less than W, it chooses to exit.

Thus, the exit decision of a plant is forward-looking: plants have to form expectations about both current and future profits. It is a dynamic problem with five state variables: 1) θ^e, the plant's belief about θ; 2) P, the expected price sequences under possible paths of demand realizations; 3) $A \equiv \{A_t\}_0^\infty$, the technology sequence; 4) time t, which determines where one is along the price sequence; 5) age a, which, combined with time t, gives the vintage A_{t-a}.[14] Let $V_t(\theta^e, a; P, A)$ be the value of staying in the market for t'th period for a plant with age a, when the plant's belief is θ^e, price sequence is P and technology sequence is A. Then a plant has V that satisfies:

$$V_t(\theta^e, a; P, A) = \pi(\theta^e, a; P, A) + \beta \cdot E_t\{\max[W, V_{t+1}(\theta^{e'}, a+1; P, A)]|\theta^e\}$$

I assume that parameters are such that $W > V_t(\theta_b, a; P, A)$ for any a, t: the present discounted value of life-time profit as a bad idiosyncratic productivity at any age is always lower than the outside option value. Therefore, bad plants always exit.

> Proposition 2.1: $V_t(\theta^e, a; P, A)$ is strictly decreasing in a, holding θ^e constant, and strictly increasing in θ^e, holding a constant; therefore, there is a cut-off age $\overline{a}_t(\theta^e; P, A)$ for each idiosyncratic productivity, such that firms of θ^e and age $a \geq \overline{a}_t(\theta^e; P, A)$ exit before production takes place in period t; furthermore, $\overline{a}_t(\theta_g; P, A) \geq \overline{a}_t(\theta_u; P, A)$.

See the appendix for proof. This follows from the fact that plants with smaller a and higher θ^e have higher expected value of staying. As V is strictly decreasing in a, plants with belief θ^e that are older than $\overline{a}_t(\theta^e; P, A)$ exit at the beginning of period t; as the expected value of staying is strictly increasing in θ^e, the exit age of good plants is older than that of unsure plants.

The industry also features continual entry. To fix the size of entry, I furthermore

[14] P is affected by demand parameter D, as well as the distribution of heterogeneous plants. See sub-section 3.2.3 for a more strict definition of the recursive competitive equilibrium.

assume that each entrant has to pay an entry cost c to enter the market:

$$c_t = c(f_t(\theta^e, 0)), \qquad c(\cdot) > 0, c'(\cdot) \geq 0.$$

I let the entry cost depend positively on the amount of entry to capture the idea that, for the industry as a whole, *fast* entry is costly and adjustment may not take place instantaneously. This can arise from a limited amount of land available to build production sites or an upward-sloping supply curve for the industry's specific capital. The free entry condition equates a plant's entry cost to its value of entry, and can be written as

$$V_t(\theta_u, 0; P, A) = c(f_t(\theta_u, 0; P, A)).$$

As more new plants enter, the entry cost is driven up until it reaches the value of entry. At this point, entry stops.

Let Q_t represent the period-t aggregate output level. An equilibrium in this industry is a path $\left\{P_t, Q_t, \{f_t(\theta^e, a)\}_{\theta^e=\theta_u \text{ or } \theta_g,\ a\geq 0}\right\}$, which satisfies the following conditions: 1) plants' entry, exit and employment decisions are optimal; 2) the evolution of $\{f_t(\theta^e, a)\}$ is generated by the appropriate summing-up of plants' entry, exit and learning: 3) P_t is such that

$$D_t = Q_t \cdot P_t, \forall t \tag{2.3}$$

, where D_t is an fully observable exogenous demand parameter that captures aggregate conditions. Industry cycles are driven by the fluctuations of D_t.

> Proposition 2.2: With time-invariant demand level $D_t = \overline{D}$, the value of $P_t \cdot A_t$ is also time-invariant; when demand fluctuates, $P_t \cdot A_t$ fluctuates with D_t.

Proposition 2.2 suggests that $P_t \cdot A_t$ moves with the value of D_t. Aggregate fluctuations affect individual plant decisions through the fluctuations of $P_t \cdot A_t$.

2.4 Aggregate Employment Dynamics

This section uses the model to assess the impact on industry-level employment dynamics of learning and creative destruction over the business cycle. Since industry-level employment dynamics are computed by aggregating the individual decisions of plants, I begin with the plant-level employment policy:

$$n_t(\theta^e, a) = [\alpha \cdot P_t \cdot A_0 \cdot (1+\gamma)^{t-a} \cdot \theta^e]^{\frac{1}{1-\alpha}} = [\alpha \cdot P_t \cdot \frac{P_t \cdot A_t}{(1+\gamma)^a} \cdot \theta^e]^{\frac{1}{1-\alpha}}$$

Because $n_t(\theta^e, a)$ depends positively on belief θ^e, a plant increases its employment (creates jobs) when it learns it is good, and exits (destroys jobs) once it learns it is bad. Hence, the evolution of θ^e captures the *learning effect*. With age a affecting n_t negatively, a plant tends to decrease its employment (destroys jobs) as it grows older (a increases), and eventually exits (destroys jobs). Therefore, $(1+\gamma)^a$ captures the *creative destruction effect*. Whether a plant creates or destroys jobs also depends further on $P_t \cdot A_t$, the product of equilibrium price and the industry-wide leading technology. I call the impact of $P_t \cdot A_t$ the *industry shock effect*. These three effects interact together to drive plant-level and thus aggregate employment dynamics.

2.4.1 Job Flows over the Plant Life Cycle

Aggregate dynamics in job creation and destruction reflect the number of plants choosing to adjust employment and the magnitude of their adjustment. In my model,

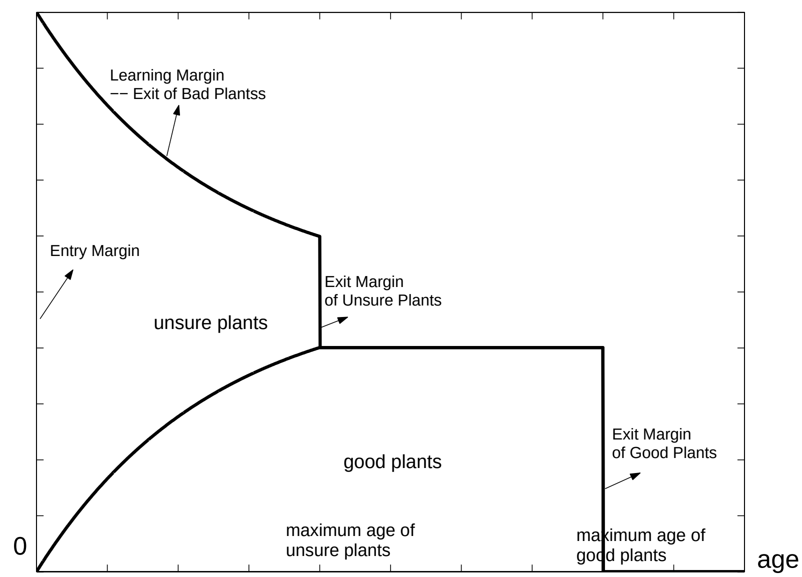


Figure 3: Dynamics of a Birth Cohort with both Learning and Creative Destruction. the distance between the lower curve (extended as the horizontal line) and the bottom axis measures the density of good firms; the distance between the two curves measures the density of unsure firms.

the response of plants varies systematically on both of these dimensions over the life-cycle. Proposition 2.1 suggests that, because of creative destruction, the evolution of the idiosyncratic-productivity distribution within a birth cohort shown in Figure 1 will be truncated by exit ages of unsure and good plants, as shown in Figure 3:

Figure 3 displays the life-cycle dynamics of a representative cohort with the horizontal axis depicting the cohort age across time. All plants enter as unsure. As the cohort ages and learns, bad plants are thrown out and good plants are realized. After a certain age, all unsure plants exit because their vintage is too old to survive with $\theta^e = \theta_u$. However, plants with $\theta^e = \theta_g$ stay. Subsequently, the cohort contains only good plants and its size remains constant because learning has stopped. Eventually, the vintage of the cohort will be too old even for good plants to survive.

Figure 3 also implies a job creation and destruction schedule over the plant life-cycle. First, because all newly born plants begin with zero employment, they begin their lives by job creation. As they age, the learning effect drives job creation among plants that discover they are good, and drives job destruction among plants that discover they are bad. Meanwhile, the creative destruction effect drives aging plants that do not learn to destroy jobs. Upon a certain age, all plants end their lives by job destruction.

As I have elaborated in Section 2.3, the concave learning curves suggest that the marginal number of learners decreases as a cohort ages. Hence, as plants grow older, *the learning effect weakens*. Fewer and fewer plants create or destroy jobs because of learning. Once all unsure plants have left, learning stops completely. On the contrary, *the creative destruction effect strengthens* with plant age. According to Proposition 2.1, older plants are more likely to exit (destroy jobs). At a certain age, all unsure plants destroy jobs by exit; as the remaining good plants grow older, eventually they destroy jobs too.

2.4.2 Decomposition of Job flows

To show the above intuition mathematically, I assume a variable x_t such that

$$x_t = \frac{P_t}{P_{t-1}}.$$

Let $C(a, x_t)$ denote the job creation rate of a cohort aged a in period t. I decompose $C(a, x_t)$ into the sum of job creation from entry, denoted $C^{entry}(a, x_t)$, job creation from learning, denoted $C^{learn}(a, x_t)$, and job creation from price increases, denoted

$C_t^{price}(a, x_t)$:

$$C(a, x_t) = C^{entry}(a, x_t) + C^{learn}(a, x_t) + C^{price}(a, x_t).$$

Apparently, both $C^{learn}(a, x_t)$ and $C^{price}(a, x_t)$ are zero for an entering cohort ($a = 0$), so that

$$C(0, x_t) = C^{entry}(0, x_t) = \frac{(\alpha P_t A_t \theta_u)^{\frac{1}{1-\alpha}}}{\frac{1}{2}\left[0 + (\alpha P_t A_t \theta_u)^{\frac{1}{1-\alpha}}\right]} = 2, \, \forall x_t.$$

For an incumbent cohort ($a > 0$), $C^{entry}(a, x_t) = 0$; its job creation from other components are

$$C^{learn}(a, x_t) = \frac{\left\{[x_t\theta_g]^{\frac{1}{1-\alpha}} - \theta_u^{\frac{1}{1-\alpha}}\right\} p\varphi f_{t-1}(\theta_u, a-1)}{\begin{array}{c}\frac{1}{2}x_t^{\frac{1}{1-\alpha}}\left[f_t(\theta_u, a)\theta_u^{\frac{1}{1-\alpha}} + f_t(\theta_g, a)\theta_g^{\frac{1}{1-\alpha}}\right] + \\ \frac{1}{2}\left[f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\right]\end{array}}, \quad (2.4)$$

and

$$C^{price}(a, x_t) = \frac{\left(\max\left\{0, x_t^{\frac{1}{1-\alpha}} - 1\right\}\right)\left[\begin{array}{c}(1-p)\, f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + \\ f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\end{array}\right]}{\begin{array}{c}\frac{1}{2}x_t^{\frac{1}{1-\alpha}}\left[f_t(\theta_u, a)\theta_u^{\frac{1}{1-\alpha}} + f_t(\theta_g, a)\theta_g^{\frac{1}{1-\alpha}}\right] + \\ \frac{1}{2}\left[f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\right]\end{array}}. \quad (2.5)$$

Here I have assumed that $x_t\theta_g > \theta_u$.[15] $C^{learn}(a, x_t)$ depends on the number

[15] This assumption, which remains valid in my calibration exercise in the next sub-section, has its support from U.S. manufacturing job-flow facts. As documented in DHS, quarterly job creation among continuing operating plants, modeled here as the sum of $POS_t^{learn}(a)$ and $POS_t^{price}(a)$, stayed strictly positive from 1972 to 1988. Notice that, if $x_t\theta_g \leq \theta_u$ instead so that $POS_t^{learn}(a)$

of plants who learn in period t they are good, represented by $p\varphi f_{t-1}(\theta_u, a-1)$, and how many jobs each of them creates, shown as $(x_t\theta_g)^{\frac{1}{1-\alpha}} - \theta_u^{\frac{1}{1-\alpha}}$. $C_t^{price}(a, x_t)$ captures possible job creation by plants not on the learning margin, including old good plants and unsure plants that have not learned, the number of which are denoted $f_{t-1}(\theta_g, a-1)$ and $(1-p)\, f_{t-1}(\theta_u, a-1)$. The term $C^{price}(a, x_t)$ is driven by industry shocks; if $x_t > 1$, the term will be zero.

Similarly, the job destruction rate for an incumbent cohort aged a ($a > 0$) in period t , denoted $D(a, x_t)$, can be decomposed as

$$D(a, x_t) = D^{learn}(a, x_t) + D^{exit}(a, x_t) + D^{price}(a, x_t),$$

where $D^{learn}(a, x_t)$ denotes job destruction from learning, $D^{exit}(a, x_t)$ denotes that from the exit of unsure and good plants, and $D^{price}(a, x_t)$ denotes that from decreases in price.

$$D^{learn}(a, x_t) = \frac{\theta_u^{\frac{1}{1-\alpha}} p(1-\varphi) f_{t-1}(\theta_u, a-1)}{\frac{1}{2} x_t^{\frac{1}{1-\alpha}} \left[f_t(\theta_u, a)\theta_u^{\frac{1}{1-\alpha}} + f_t(\theta_g, a)\theta_g^{\frac{1}{1-\alpha}} \right] + \frac{1}{2} \left[f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}} \right]} \tag{2.6}$$

drops to zero, $POS_t^{price}(a)$ would also be zero since $\theta_g > \theta_u$, together with $x_t\theta_g \leq \theta_u$, suggests $x_t \leq 1$. In that case, there would be no job creation from continuing operating plants and my model would not be able to match DHS's documented facts of gross job flows.

$$D^{price}(a,x_t) = \frac{\left(\max\left\{0, 1 - x_t^{\frac{1}{1-\alpha}}\right\}\right)\left[\begin{array}{c}(1-p)\, f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + \\ f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\end{array}\right]}{\begin{array}{c}\frac{1}{2}x_t^{\frac{1}{1-\alpha}}\left[f_t(\theta_u, a)\theta_u^{\frac{1}{1-\alpha}} + f_t(\theta_g, a)\theta_g^{\frac{1}{1-\alpha}}\right] + \\ \frac{1}{2}\left[f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\right]\end{array}}, \tag{2.7}$$

$D^{learn}(a,x_t)$ captures job destruction on the learning margin, with $p(1-\varphi)f_{t-1}(\theta_u, a-1)$ representing the number of plants who learn they are bad. It equals zero for cohorts without unsure plants in period $t-1$ ($f_{t-1}(\theta_u, a-1) = 0$) since learning among these plants has stopped. $D^{price}(a,x_t)$ represents possible job destruction by plants not on the learning margin. This job destruction is driven by industry shocks and occurs as long as $x_t < 1$. The plants affected include old good plants and unsure plants that have not learned.

The magnitude of $D^{exit}(a,x_t)$, job destruction from the exit of unsure or good plants, is more complicated due to shifts of the exit margins. Let $\overline{a_t}(\theta_u)$ represent the period-t exit age of unsure plants and $\overline{a_t}(\theta_g)$ that of good plants. When $\overline{a_t}(\theta_u) > \overline{a_{t-1}}(\theta_u)$, unsure exit margin extends to an older age and no unsure plants are exiting. Similarly, no good plants are exiting when $\overline{a_t}(\theta_g) > \overline{a_{t-1}}(\theta_g)$. If both margins extend to older ages, then no plants are exiting and $D^{exit}(a,x_t)$ must be zero for any cohorts.

On the contrary, with $\overline{a_t}(\theta_u) \leq \overline{a_{t-1}}(\theta_u)$, the unsure exit margin stays at the same age or shifts to a younger age, so that one or more cohorts of unsure plants are exiting. It can be shown that

$$\begin{aligned} D^{exit}(a,x_t) &= \frac{\theta_u^{\frac{1}{1-\alpha}} f_{t-1}(\theta_u, a-1)}{\frac{1}{2}x_t^{\frac{1}{1-\alpha}} f_t(\theta_g, a)\theta_g^{\frac{1}{1-\alpha}} + \frac{1}{2}\left[f_{t-1}(\theta_u, a-1)\theta_u^{\frac{1}{1-\alpha}} + f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}\right]} \\ \text{if } a &< \overline{a_{t-1}}(\theta_u) + 1 \text{ and } a \geq \overline{a_t}(\theta_u). \end{aligned}$$

Similarly, with $\overline{a_t}(\theta_g) \leq \overline{a_{t-1}}(\theta_g)$, the good exit margin stays at the same age or shifts to a younger age, so that one or more cohorts of good plants are exiting. For these cohorts, job destruction reaches its maximum value two, shown as follows:[16]

$$\begin{aligned} D^{exit}(a, x_t) &= \frac{\theta_g^{\frac{1}{1-\alpha}} f_{t-1}(\theta_g, a-1)}{\frac{1}{2} f_{t-1}(\theta_g, a-1)\theta_g^{\frac{1}{1-\alpha}}} = 2 \\ \text{if } a &< \overline{a_{t-1}}(\theta_g) + 1 \text{ and } a \geq \overline{a_t}(\theta_g). \end{aligned}$$

I conclude this sub-section by relating above decomposition of job flows to the underlying driving forces. I argued earlier that plant-level employment is affected by the learning effect, the creative destruction effect, and the industry shock effect. Apparently, C^{learn} and D^{learn} come from the learning effect. The creative destruction effect drives C^{entry} and D^{exit}: youngest plants enter, and the oldest plants exit. However, the industry shock effect and the creative destruction effect *together* drive C^{price} and D^{price}.

To see why, recall that the industry shock effect is defined as the impact of $P_t A_t$. With constant $P_t A_t$, there is no industry shock effect, but D^{price} would still be positive because

$$x_t = \frac{P_t}{P_{t-1}} = \frac{P_t A_t}{P_{t-1} A_{t-1}(1+\gamma)} = \frac{1}{(1+\gamma)} < 1, \text{ when } P_t A_t = P_{t-1} A_{t-1}.$$

This is due to the creative destruction effect: a plant becomes more technologically outdated by $\frac{1}{(1+\gamma)}$ as it ages for another period. To see how creative destruction affects C^{price}, let me assume that $P_t A_t$ increases so that the industry shock effect is present.

[16] Here I implicitly assume $\overline{a_{t-1}}(\theta_u) < \overline{a_t}(\theta_g) - 1$, so that any cohorts with exiting good plants included no unsure plants in period $t-1$. This will the case in my calibration exercises.

In that case, C^{price} may not be positive because x_t may not be greater than one: the impact of P_tA_t' s increase on x_t needs to overcome the impact of $\frac{1}{(1+\gamma)}$. Hence, the creative destruction effect affects C^{price} by dampening the industry shock effect.

2.4.3 The magnitude of job flows with no fluctuations

This sub-section establishes analytically a negative relationship between the average magnitude of job flows and plant age in a version of my model with no fluctuations. Variations in D_t serve as the source of economic fluctuations in my model. Proposition 2.2 establishes that time-invariant D_t implies a time-invariant P_tA_t. With time-invariant P_tA_t, the expected value of staying across a and θ^e would also be time-invariant, which implies time-invariant size of entry and time-invariant exit ages.

Moreover, with $P_tA_t = P_{t-1}A_{t-1}$, x_t equals a less-than-one value $\frac{1}{(1+\gamma)}$. Hence, $C^{price}(a, \frac{1}{(1+\gamma)})$ is zero so that job creation includes only C^{entry} and C^{learn}, while $D^{price}(a, \frac{1}{(1+\gamma)})$ stays positive. Let $C^{*}(a)$ denote the job creation rate of a cohort aged a with no fluctuations, $D^{*}(a)$ the job destruction rate, and $\overline{a_u}^{*}$ and $\overline{a_g}^{*}$ the exit ages of unsure and good plants. I have:

$$
\begin{aligned}
C^{*}(a) &= C^{entry}(a, \frac{1}{(1+\gamma)}) = 2, \text{ if } a = 0 \\
C^{*}(a) &= C^{learn}(a, \frac{1}{(1+\gamma)}) \\
&= \frac{\left\{\left[\frac{\theta_g}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}} - \theta_u^{\frac{1}{1-\alpha}}\right\} p\varphi}{\frac{1}{2}\left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}\left[(1-p)\theta_u^{\frac{1}{1-\alpha}} + \frac{\varphi(1-(1-p)^{a})}{(1-p)^{a-1}}\theta_g^{\frac{1}{1-\alpha}}\right] + \frac{1}{2}\left[\theta_u^{\frac{1}{1-\alpha}} + \frac{\varphi(1-(1-p)^{a-1})}{(1-p)^{a-1}}\theta_g^{\frac{1}{1-\alpha}}\right]}, \text{ if } 0 < a < \overline{a_u}^{*} \\
C^{*}(a) &= 0, \text{ otherwise}
\end{aligned}
$$

Proposition 2.3: without demand fluctuations, the job creation rate weakly decreases in cohort age.

Job creation strictly decreases in age for cohorts younger than $\overline{a_u}^*$, because $\frac{\varphi(1-(1-p)^a)}{(1-p)^{a-1}}$ increases in a. According to the all-or-nothing learning described in Sub-section 2.2, $\varphi(1-(1-p)^a)$ is the fraction of good plants in a cohort aged a, and $(1-p)^{a-1}$ the fraction of unsure plants. The ratio of good plants to unsure plants increases in a because of learning: for older cohorts, more plants have learned. Job creation drops to and stays at zero for the group of plants older than $\overline{a_u}^*$, since these plants have already learned that they are good.

I also have:

$$
\begin{aligned}
D^*(a) &= D^{learn}(a, \frac{1}{(1+\gamma)}) + D^{price}(a, \frac{1}{(1+\gamma)}) \\
&= \frac{\begin{array}{c} \theta_u^{\frac{1}{1-\alpha}} p(1-\varphi) + \\ \left(1 - \left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}\right) \left[\begin{array}{c} (1-p)\,\theta_u^{\frac{1}{1-\alpha}} + \\ \frac{\varphi(1-(1-p)^{a-1})}{(1-p)^{a-1}} \theta_g^{\frac{1}{1-\alpha}} \end{array}\right] \end{array}}{\begin{array}{c} \frac{1}{2}\left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}} \left[(1-p)\theta_u^{\frac{1}{1-\alpha}} + \frac{\varphi(1-(1-p)^a)}{(1-p)^{a-1}} \theta_g^{\frac{1}{1-\alpha}}\right] + \\ \frac{1}{2}\left[\theta_u^{\frac{1}{1-\alpha}} + \frac{\varphi(1-(1-p)^{a-1})}{(1-p)^{a-1}} \theta_g^{\frac{1}{1-\alpha}}\right] \end{array}}, \text{ if } 0 < a < \overline{a_u}^*
\end{aligned}
$$

$$D^*(a) = D^{price}(a, \frac{1}{(1+\gamma)}) + D^{exit}(a, \frac{1}{(1+\gamma)})$$

$$= \frac{\left(1 - \left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}\right)\left[\frac{\varphi(1-(1-p)^{a-1})}{(1-p)^{a-1}}\theta_g^{\frac{1}{1-\alpha}}\right] + \theta_u^{\frac{1}{1-\alpha}}}{\frac{1}{2}x_t^{\frac{1}{1-\alpha}}\left[\frac{\varphi(1-(1-p)^{a})}{(1-p)^{a-1}}\theta_g^{\frac{1}{1-\alpha}}\right] + \frac{1}{2}\left[\theta_u^{\frac{1}{1-\alpha}} + \frac{\varphi(1-(1-p)^{a-1})}{(1-p)^{a-1}}\theta_g^{\frac{1}{1-\alpha}}\right]}, \text{ if } a = \overline{a_u}^*$$

$$D^*(a) = D^{price}(a, \frac{1}{(1+\gamma)}) = \frac{1 - \left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}}{2\left(1 + \left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}\right)}, \text{ if } \overline{a_u}^* < a < \overline{a_g}^*$$

$$D^*(a) = D^{exit}(a, \frac{1}{(1+\gamma)}) = 2, \text{ if } a = \overline{a_g}^*$$

Proposition 2.4: For small enough γ, job destruction weakly decreases in cohort age for $a \neq \overline{a_u}^*$ and $a \neq \overline{a_g}^*$.

For cohorts younger than $a \leq \overline{a_u}^*$, $D^*(a)$ decreases in a because learning implies that the ratio of good to unsure plants $(\frac{\varphi(1-(1-p)^a)}{(1-p)^{a-1}})$ increases in a. For cohorts with $\overline{a_u}^* < a < \overline{a_g}^*$, although learning has stopped, plants gradually decrease employment (destroy jobs) due to technological obsolescence or creative destruction. Their job destruction rate is as shown in 2.7. Notice that for small enough γ, the value of $\frac{1-\left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}}{2\left(1+\left[\frac{1}{(1+\gamma)}\right]^{\frac{1}{1-\alpha}}\right)}$ is close to zero, so that $D^*(a)|_{a \leq \overline{a_u}^*} > D^*(a)|_{\overline{a_u}^* < a < \overline{a_g}^*}$.

Therefore, in the absence of industry shocks, the magnitude of job destruction declines with plant age, except for the cohorts with exiting unsure plants or good plants. These exceptions result from the simplified nature of all-or-nothing learning. With plants either learning nothing or learning everything suddenly, the number of beliefs is limited, so that upon a certain age, the force of creative destruction drives out the whole group of plants with a certain belief. However, strictly speaking, these exceptions do not contradict with the empirical evidence. Notice that here a is the number of quarters a plant has survived. But in Table 1A, the negative relationship between the magnitude and plant age is shown by *age categories.* Hence, my model can possibly produce patterns similar to these shown in Table 1A, even with these exceptions.

2.4.4 Cyclical job flows with industry fluctuations

I introduce industry fluctuations by allowing D_t to vary over time. As stated in Proposition 2.2, industry fluctuations affect plant-level employment decisions through the variations of P_tA_t. According to 2.4, 2.5, 2.6 and 2.7, variations of P_tA_t affect job creation and destruction through the variations of x_t. Now the question is: can my model generate variation in the cyclical responses of job creation and destruction across age categories similar to those shown in Table 1 and Figure 1?

According to 2.4 and 2.6, the value of C^{learn} and D^{learn} depend on the number of learners and how much each learner adjusts. Holding fixed the variations of individual adjustment caused by variations of x_t, less learners imply smaller employment variations at the cohort level. Since younger plants are more likely to be on the learning margin, their job flows from learning should vary more with cyclical shocks.

Since plant birth are concentrated among young plants, variations in entry add to the variations of young plants' job creation. As analyzed earlier, D^{exit}, job destruction from the exit of unsure and good plants, is affected by shifts of exit ages. When exit ages increase following a favorable shock, there is no exit of unsure or good plants for cohorts close to the previous exit margins. When exit ages decrease, several cohorts exit, giving rise to a jump in job destruction. Since exit ages apply only to older plants, job destruction may be more responsive to shocks for older plants.

Recall that, in the data, the variances of job flows decline in plant age; but job destruction is more responsive to shocks than job creation for older plants. The analysis of this sub-section suggests that my model has the potential to produce these patterns. The negative relationship between the magnitude of job flows and plant age can come from the cyclical job flows from learning (C^{learn} and D^{learn}). Since learning weakens with plant age, the cyclical job flows from exit may dominate older plants, implying a more responsive job destruction margin.

2.5 Quantitative Implications

This section numerically analyzes a stochastic version of my model in which the demand level follows a two-state Markov process with states $[D_h, D_l]$ and transition probability μ. My computational strategy follows Krusell and Smith (1998) by shrinking the state space into a limited set of variables and showing that these variables' laws of motion can approximate the equilibrium behavior of firms in the simulated time series.[17] After solving for the approximate value functions, I calibrate my model so that its equilibrium job flows mimic the observed patterns in U.S. manufacturing.

2.5.1 Calibration

The assigned parameter values are listed in Table 2. Some of the parameter values are pre-chosen. I allow a period to represent one quarter and let the quarterly discount rate $\beta = 0.99$. μ is set equal to 0.95, so that aggregate demand switches between D_h and D_l with a constant probability of 0.05 per quarter. This implies that a given

[17] See Section 3.4 for the details of applying Krusell and Smith approach with a one-worker-per-firm set-up. With multiple workers per plant in this chapter, the fit generally improves.

parameters (pre-chosen)	value
productivity of bad firms: θ_b	1
productivity of good firms: θ_g	3.5
quarterly technological pace: γ	0.0032
quarterly discount rate: β	0.99
entry cost function	$0.405 + 0.52 * f(0, \theta_u)$
persistence rate of demand: μ	0.95
high demand: D_h	10000
low demand: D_l	8500
parameters (calibrated)	**value**
prior probability of being a good firm: φ	0.05
quarterly pace of learning: p	0.12
Outside option value: W	5

Table 2: Parameterization of the Model

demand level will persist for five years on average, consistent with business cycle frequencies. The most significant parameters in this group are the relative productivity of good and bad plants. I follow the choices of Davis and Haltiwanger (1999), who assume a ratio of high-to-low productivity of 2.4 for total factor productivity and 3.5 for labor productivity based on the between-plant productivity differentials reported by Bartelsman and Doms (1997). Since labor is the only input in my model, I normalize productivity of bad plants as 1 and set productivity of good plants as 3.5. The quarterly pace of technological progress γ is set equal to 0.0032, based on the estimates of growth rates of TFP in US manufacturing sector by Basu, Fernald and Shapiro (2001). Caballero and Hammour (1994) assume a linear entry cost function $c_0 + c_1\ f(0, \theta_u)$ with $f(0, \theta_u)$ denoting the size of entry, which is also applied in my calibration exercises.

The values of p, the pace of learning, and φ, the probability of being a good firm, are chosen to match the mean job creation and destruction rates for young plants shown in Table 1A. I follow DHS in defining young plants as those younger than 40 quarters (10 years) and old plants as all others. According to 2.8 and 2.9, the mean job creation rate drops to zero for plants older than $\overline{a_u}^*$. But this is not the case for old plants in Table 1A. This suggests that $\overline{a_u}^* > 40$. Hence the value of φ and p are set so that $C^*(0 \leq a \leq 40) \approx 7.52\%$ and $D^*(0 \leq a \leq 40) \approx 6.56\%$. The implied calibrated values are as shown in Table 2.

With the value of all other parameters assigned as above, I adjust the value of W, the outside option value, to make sure the followings. First, the expected value of staying for bad plants is always lower than W, so that bad plants always exit. Second, the expected value of entry, or, in another word, the expected value of staying for age-zero unsure plants, is always positive, so that entry never stops. W is set equal to 5.

2.5.2 Simulations of aggregate employment dynamics

I simulate my calibrated model's response to random demand realizations of 1000 periods generated by the model's Markov chain. Figure 4 presents the time series of detrended price, entry size, and exit ages of good and unsure plants. Young and old plants' job flow series by learning are shown in Figure 5. Job flows by entry and by shifts of exit ages are in Figure 6. Figure 7 presents young and old plants' job flows by price variations. The related statistics are displayed in Table 3 and Table 4.

In Figure 4, all series move in the same direction: with higher demand, price increases, more plants enter, and both exit margins extend to older ages. In Figures 5, 6, and 7, job creation and destruction co-move *negatively*, which matches DHS's finding that job creation is pro-cyclical but job destruction counter-cyclical. In Figure 5, job destruction from learning is higher than job creation from learning. This comes from my calibrated value of a plant's probability of being good as 0.05. According to the all-or-nothing learning, a probability of being good of 0.05 implies that 95% of the learners exit by learning they are bad. It also suggests that, although learning *itself* affects job creation and destruction equally, destruction varies more with more plants affected. Notice that old plants' job destruction by unsure and good plants' exit is lower-bounded by zero in Figure 6. This matches my model's intuition: D^{exit} drops to zero when exit ages increase, and jumps up when exit ages decrease. Job flows by price variations are also lower-bounded by zero: D^{price} drops to zero when x_t is greater than one, and C^{price} drops to zero when x_t is less than one.

Table 3 displays statistics on gross job flows from my simulations together with those from the data. The table suggests that a reasonably calibrated version of my model reproduces, at least qualitatively, the observed differences in young and old plants' magnitude of job flows and the relative volatility in job creation and destruction. As shown in Table 3, the sample means of job creation and destruction decline in plant age; and the relative volatility in job creation and destruction is more symmetric for young plants. These relationships are still present when job flows from plant birth and death are excluded.

Table 4 takes one step further by reporting separately the decomposed sample statistics of job flows by plant birth, plant death, and continuing plants. Job creation by entry only contributes to young plants' job flows. Job creation from continuing plants (Cc) is decomposed into job creation by learning (C^{learn}) and that by price increases (C^{price}). Since my simulated exit ages never drop below 40 quarters, young plants' job destruction from plant death (Dd) comes only from learning (D^{learn}). Old plants' job destruction from plant death comes both from learning (D^{learn}) and from

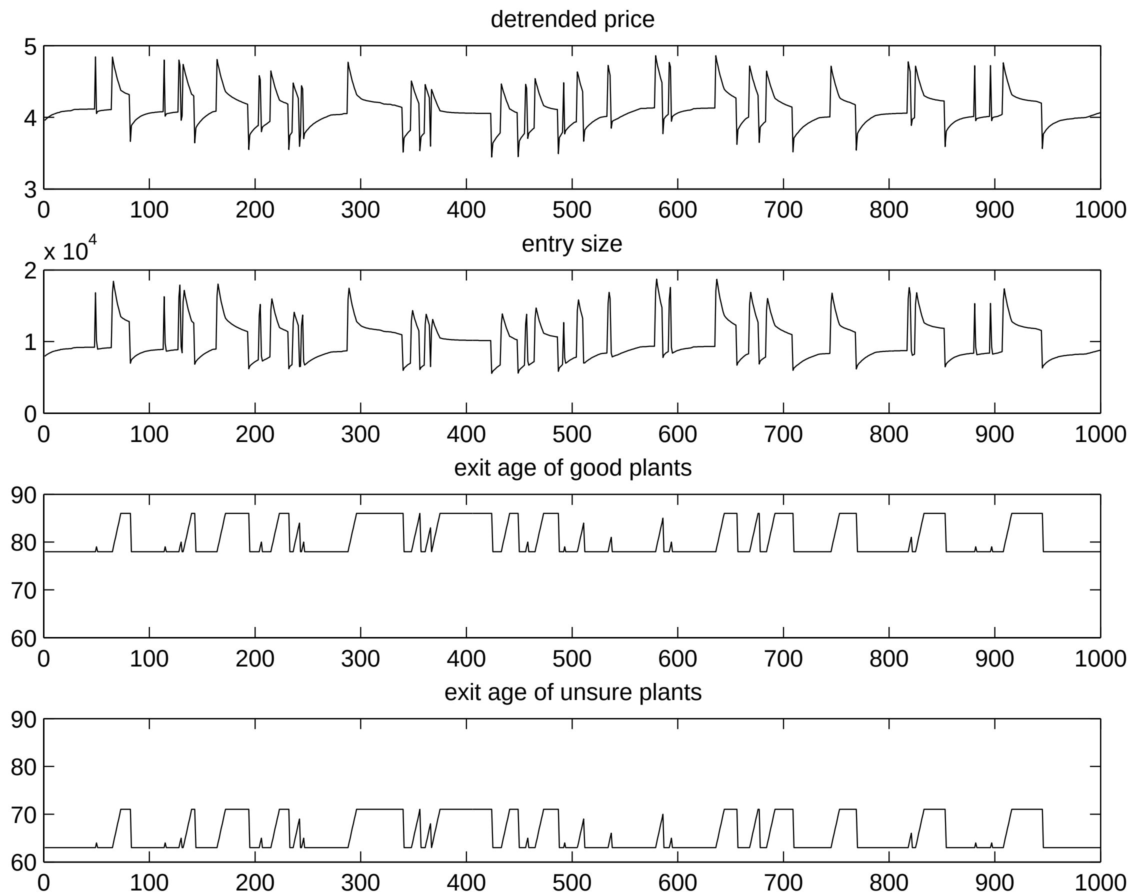


Figure 4: Simulated Time Series of Detrended Price, Entry Size, and Exit Ages Of Unsure and Good Plants. Detrended price = $P_t A_t \cdot (1+\gamma)$. Entry size is calculated as the amount of labor hired by entrants.

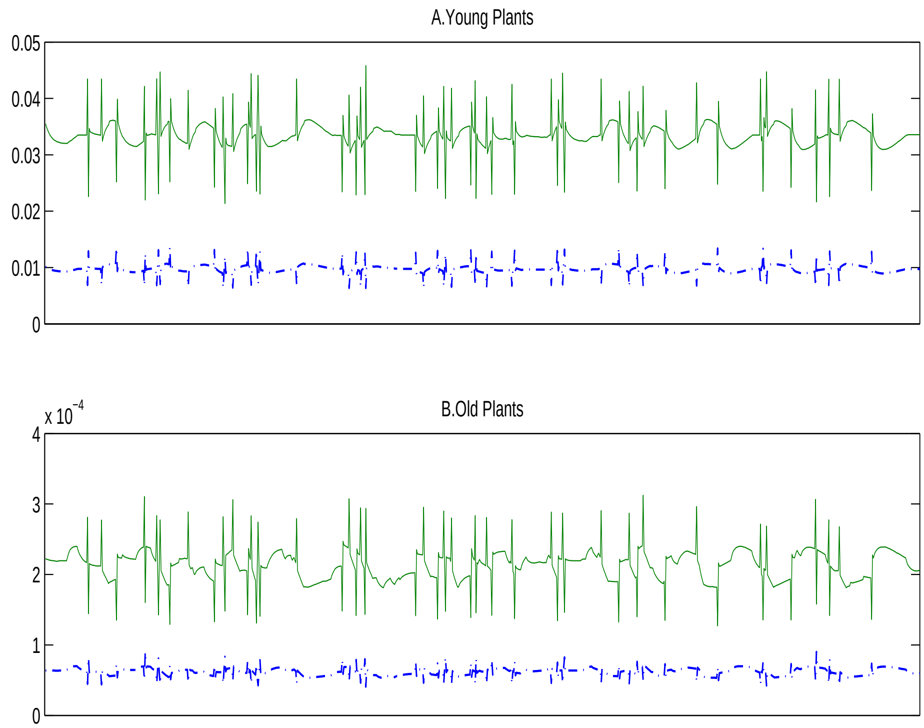


Figure 5: Job Flows By Learning. Dashed lines job creation. Solid lines job destruction.

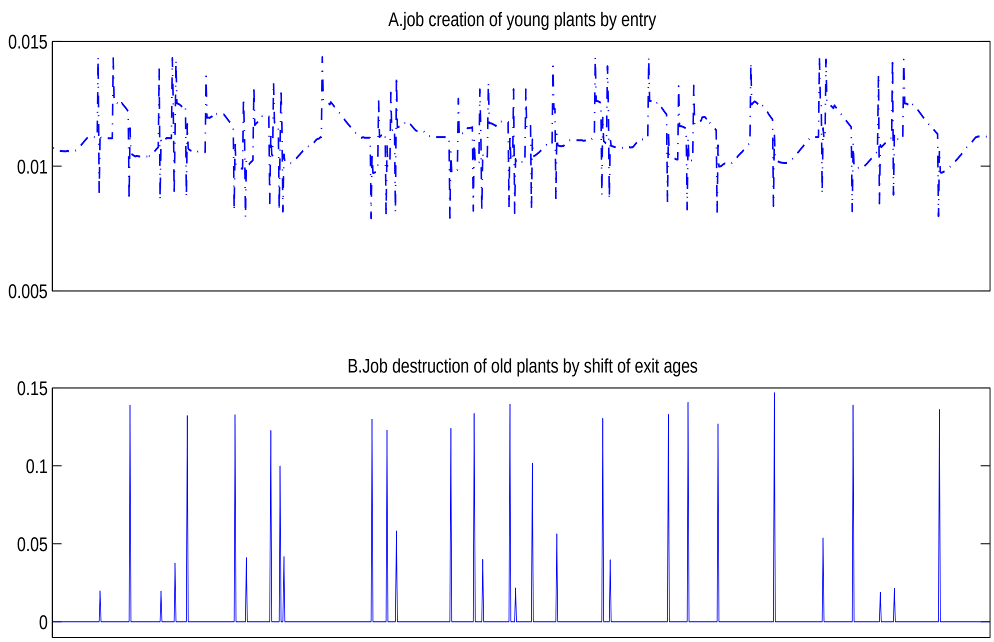


Figure 6: Young Plants' Job Creation by Entry and Old Plants' Job Destruction by Shift of Exit Ages.

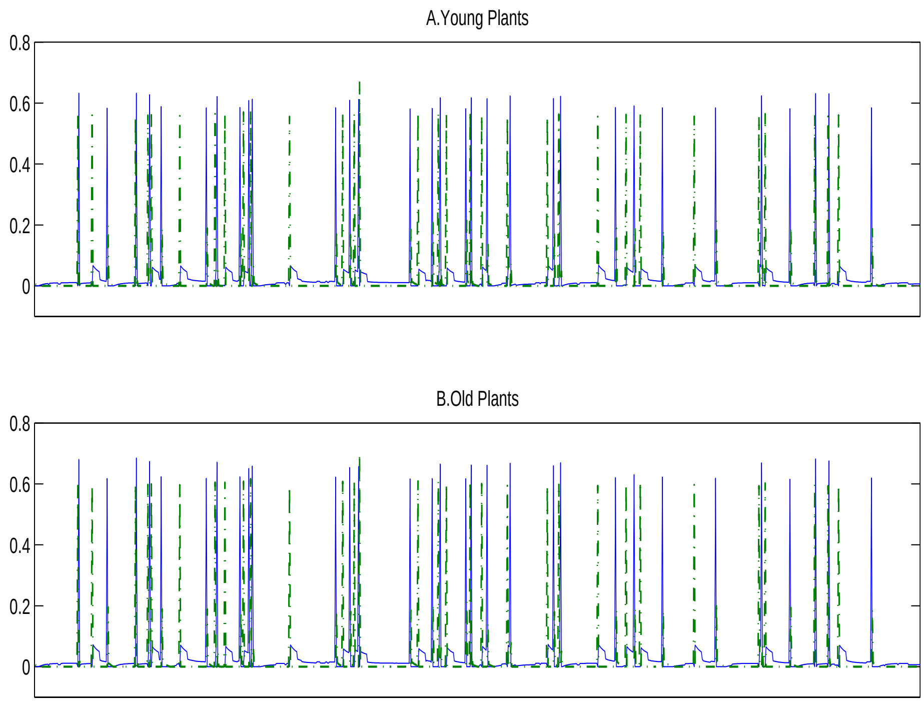


Figure 7: Job Flows by Price Variations. Dashed lines job creation. Solid lines job destruction.

A. Means						
Plant type	$E(Cb)$	$E(Cc)$	$E(C)$	$E(Dd)$	$E(Dc)$	$E(D)$
young (data)	1.52	6.00	7.52	1.24	5.33	6.56
old (data)	0.12	4.42	4.54	0.47	4.77	5.24
young (simu)	1.11	5.56	6.67	0.98	3.29	4.26
old (simu)	0	2.36	2.36	0.29	3.52	3.81
B. Standard deviations						
Plant type	$\boldsymbol{\sigma}(Cb)$	$\boldsymbol{\sigma}(Cc)$	$\sigma(C)$	$\sigma(Dd)$	$\sigma(Dc)$	$\sigma(D)$
young (data)	1.06	1.23	1.80	0.66	1.67	2.07
old (data)	0.07	0.78	0.78	0.22	1.50	1.60
young (simu)	0.10	10.24	10.29	0.10	10.37	10.44
old (simu)	0	10.71	10.71	2.18	11.09	11.54

C. Variance ratio of job destruction to creation		
Plant type	$\boldsymbol{\sigma}(\mathbf{D})^2/\boldsymbol{\sigma}(\mathbf{C})^2$	$\sigma(Dc)^2/\sigma(Cc)^2$
young (data)	1.32	2.80
old (data)	4.18	3.59
young (simu)	1.03	0.2658
old(simu)	1.16	0.2781

Table 3: Quarterly gross job flows in the calibrated model and in US manufacturing by plant age: 1973 II to 1988 IV. Notations are the same as Table 1. "data" indicates statistics from the data, "simu" statistics from my simulations. All numbers are in percentage points.

	Cb	Cc		Dd		Dc
Plant type	$E(C^{entry})$	$E(C^{learn})$	$E(C^{price})$	$E(D^{learn})$	$E(D^{exit})$	$E(D^{price})$
young	1.11	3.35	2.21	0.98	0	3.29
old	0	0.02	2.34	0.006	0.28	3.52
Plant type	$\boldsymbol{\sigma}(C^{entry})$	$\boldsymbol{\sigma}(C^{learn})$	$\boldsymbol{\sigma}(C^{price})$	$\boldsymbol{\sigma}(D^{learn})$	$\boldsymbol{\sigma}(D^{exit})$	$\boldsymbol{\sigma}(D^{price})$
young	0.10	0.28	10.06	0.10	0	10.37
old	0	0.002	10.71	0.0006	2.18	11.09
Plant type	$\sigma(D^{learn})^2/\sigma(C^{learn})^2$			$\sigma(D^{price})^2/\sigma(C^{price})^2$		
young	0.102			1.06		
old	0.092			1.17		

Table 4: Decomposed job flows in the calibrated model. C the job creation rate, D the job destruction rate. "entry" indicates job flows by entry; "learn" job flows by learning; "price" job flows by price variations; "exit" job flows by exit of unsure and good plants. All numbers are in percentage points.

unsure and good plants' exit (D^{exit}). Table 4 suggests that learning contributes to the negative relationship between the sample means of job flows and plant age in Table 3, since in Table 4, mean job flows by price variations and by unsure and good plants' exit *increase* in plant age. Hence, the steady-state predictions of Propositions 2.3 and 2.4 remain valid in the stochastic version of the model. Table 4 also shows that, although the relative volatility of job destruction to creation by learning *decrease* in plant age (a variance ratio of 0.102 for young, and 0.092 for old), old plants feature more variations in job destruction due to price variations (a variance ratio of 1.06 for young, and 1.17 for old) and exit of unsure and good plants.

In summary, Table 3 and Table 4 together confirm the conjectures posited in the previous section. Because learning weakens with plant age, the magnitude of job flows declines with plant age, and the cyclical patterns of job flows by exit and price variations dominate in old plants, which features a more responsive job destruction margin.

However, Table 3 also suggests that my calibrated model does not match the data well in two respects. First, it cannot fully account for the large magnitude of job flows in mature plants. In Table 3, the simulated sample means of job creation and destruction for old plants are well below those in the data. This is especially the case for job creation in old plants, with a sample mean of only 2.36% from the simulations compared to 5.56% in the data. Second, the simulated standard deviations of job creation and destruction are too high; moreover, although the standard deviation of job destruction *decreases* in plant age in the data, it *increases* in my simulations.

I conclude from Table 4 that two features of my calibrated model contribute to the above failings. The first is that learning in old plants seems too weak in my model.

This is shown in Table 4 as a very low sample average of old plants' job flow rates by learning (0.02% for job creation and 0.006% for job destruction). This feature not only drives down the magnitude of job flows in old plants in Table 3, but also causes the cyclical patterns of job flows by exit and price variations over-dominate in old plants (in Table 3, the standard deviation of job destruction *fails* to decline in plant age, although the standard deviation of job destruction by learning *does* decline in age). The second feature is that price variations generated by my model is too sharp, so that when plants not on the learning margin adjust their employment, they adjust too much. This can also be seen in Table 4: the standard deviations of old plants' C^{price} and D^{price} are all very high. This feature drives up the simulated standard deviations of job flow rates in Table 3.

Two of my simplifying assumptions may be responsible for these failures. First, the assumption that aggregate demand can take on only two values leads to sharp variations in price. Assuming that D_t follows a Markov chain with more states is likely to improve this feature. Second, the all-or-nothing learning process gives rise to an overly weak learning effect in old plants, since there is no learning for plants older than $\overline{a_t}(\theta_u)$. Making the learning process more complicated may improve this feature. For example, suppose that the random noise is normally distributed so that the signals received by good plants are normally distributed around θ_g and the signals received by bad plants are normally distributed around θ_b. In that case, a plant can never know *for sure* that it is good or bad; even the oldest continuing operating plants may adjust employment in the face of learning.

2.6 Conclusion

In this chapter, I present a framework where two forces interact together to drive micro-level job flows: creative destruction reallocates labor into technologically more advanced production units; while learning leads labor to good plants. Two salient stylized facts motivate my theory: The first is that young plants display greater turnover rates than old plants. The second is that, although job destruction is more volatile than job creation in general, this asymmetry weakens in younger plants.

The key of my explanation is that learning weakens with plant age. With this feature, my model generates the observed negative relationship between the mean magnitude of job flow rates and plant age. When demand fluctuates, the learning force generates relatively symmetric responses on the creation and destruction sides, while the creative destruction force makes job destruction much more responsive. Again, because learning weakens with age, the relatively symmetric response of learning dominates for young businesses and the asymmetric response of creative destruction dominates for old ones.

I use the model to assess job-flow magnitude over a plant's life cycle analytically and calibrate the model to match the data quantitatively. Calibration results show that my model does well in matching young businesses' higher job-flow magnitude as well as their relative symmetric volatility of job creation and destruction. However, it cannot fully account for the magnitude of job flows among mature businesses because of my assumption of a simplified all-or-nothing form of learning.

3 The Scarring Effect of Recessions

3.1 Introduction

How do recessions affect resource allocation? This question has long attracted the attention of economists. As far back as 1934, Schumpeter advanced the view of "cleansing": recessions are times when outdated or relatively unprofitable techniques and products are pruned out of the productive system. This view has been revived since the finding of Davis and Haltiwanger (1992) that job reallocation in the U.S. manufacturing sector is concentrated during recessions.[18] Attempting to explain these cyclical patterns, an assortment of theoretical work has arisen returning to the Schumpeterian cleansing view.[19] In their arguments, production units with different efficiency levels coexist due to certain reallocation frictions; when recessions drive down profitability, the least efficient units should cease to be viable and shut down,[20] which frees up resources for more productive uses. Therefore, setting aside the losses to particular businesses and individuals, reallocation during recessions leads to greater efficiency in resource allocation.[21]

Despite solid theoretical reasoning, the cleansing view deviates from empirical evidence in one important aspect — it implies countercyclical productivity, while average labor productivity is in fact procyclical. This was pointed out by Caballero

[18] Similar evidence has also been found in the manufacturing sectors of Canada, Denmark, Norway and Colombia. See Davis and Haltiwanger (1999).

[19] See Hall (1992, 2000), Mortensen and Pissarides (1994), Caballero and Hammour (1994, 1996), and Gomes, Greenwood, and Rebelo (2001).

[20] These models assume perfectly competitive markets so that, as price takers, less efficient units are also less profitable. However, with market power, a less efficient unit can be more profitable. See Foster, Haltiwanger and Syverson (2003).

[21] However, these papers not necessarily suggest that recessions lead to higher welfare. In other words, it is likely that higher *allocation efficiency* and lower *welfare efficiency* coexist during recessions.

and Hammour (1994), where who suggest that the cleansing effect may be dwarfed by other factors. Subsequent empirical work has challenged the cleansing view from the creation side. For example, Bowlus (1993) and Davis, Haltiwanger and Schuh (1996) find that jobs created during recessions tend to be short-lived, which inspired Barlevy (2002) to question whether recessions encourage the creation of the most efficient units. However, although job destruction has been documented to be more responsive to business cycles than job creation,[22] few have yet asked the question, "Are the production units cleared by recessions necessarily inefficient?" If not, then recessions might exacerbate the inefficiency of resource allocation instead of alleviating it as the conventional cleansing view suggests.[23]

In this chapter, I propose a "scarring effect" of recessions that plays against the conventional cleansing effect. I argue that while recessions drive out some of the least productive firms, they also kill off "potentially good firms"; firms that have the potential to be proven efficient in the future are forced to leave due to reduced profitability. The loss of potentially good firms leaves "scars" when a recession arrives, and the "scars" deepen as the recession persists. The presence of the scarring effect revises the conventional view of recessions as periods of solely healthy reallocation: the overall impact of recessions on allocative efficiency should depend on the relative magnitude of two competing effects — cleansing and scarring.

I offer my explanation by combining the vintage model of Caballero and Ham-

[22] Davis and Haltiwanger (1999) document that job destruction tends to be more volatile than job creation in manufacturing sectors. The variance of destruction divided by the variance of creation is 2.04 for the U.S., 1.49 for Canada, 1.0 for Denmark, 2.68 for the Netherlands, 1.69 for Germany, 0.68 in Colombia, and 18.19 for the U.K..

[23] Ramey and Watson (1997) and Caballero and Hammour (1999) argue that job destruction threshold in recessions can be socially inefficient. However, their cyclical implications on productivity are the same as in the models of the conventional cleansing effect: average job quality goes up during recessions.

mour (1994) with *learning* in the spirit of Jovanovic (1982). As in Caballero and Hammour (1994), exogenous technological progress introduces a force of creative destruction that drives in technologically sophisticated entrants to displace older, outmoded firms.[24] However, in my model, firms of the same vintage also differ in idiosyncratic productivity: some are good and others are bad. A firm's idiosyncratic productivity can represent the talent of the manager, or alternatively, the store location, the organizational structure of the production process, or its fitness to the embodied technology. More importantly, firms' idiosyncratic productivity are not observable ex ante, but can be learned through experience. As information arrives, firms choose to exit or stay, so that an additional learning force arises to keep good firms and select out bad firms. Variations in aggregate demand serve as the source of economic fluctuations. As a negative demand shock strikes and persists, the intensified creative destruction directs labor to younger, more productive vintage, causing a cleansing effect that raises average labor productivity; meanwhile, a truncated learning process shifts labor toward bad firms, creating a scarring effect that pulls down average labor productivity. The question then becomes, which effect dominates? In Section 4, I calibrate my model using data on U.S. manufacturing job flows and study its quantitative implications. My results suggest that the scarring effect dominates the cleansing effect in the U.S. manufacturing sector from 1972 to 1993, and can account for the observed procyclical average labor productivity.

My model stresses two frictions that stifle instantaneous labor reallocation. First, entry is costly, which allows different vintages to coexist. Second, learning takes time, so that good and bad firms both survive. Vintage and idiosyncratic productivity

[24] The phrase "creative destruction" comes from Schumpeter (1939). It refers to the birth and death of firms due to the introduction of new technology into the production process.

together can explain the observed heterogeneous firm-level productivity. The vintage component suggests that entering cohorts are more productive than incumbents.[25] The idiosyncratic productivity component implies that each vintage cohort is itself a heterogeneous group. Vintage and idiosyncratic productivity together also lead to the following productivity dynamics. Creative destruction perpetually drives in entrants with higher productivity. Learning selects out bad firms over time so that as a cohort ages, its average productivity rises but productivity dispersion declines. Data from the U.S. manufacturing sector provides large and pervasive empirical evidence to support these predictions.[26]

The existing empirical literature has advanced learning and creative destruction as powerful tools to understand the patterns of firm turnover and industrial dynamics.[27] The significance of their interaction has also been suggested. Davis and Haltiwanger (1999) note that "vintage effects may be obscured by selection effects; vintage and selection effects may also interact in important ways..." In my model, the interaction of these two forces generates the scarring effect of recessions.

The rest of the chapter is organized as follows. Section 2 lays out a model combining creative destruction with learning. The cleansing and scarring effects are mo-

[25] Although this is often true in the data, some authors such as Aw, Chen and Roberts (1997) find evidence that entrants are no more productive than incumbents. Foster, Haltiwanger and Syverson (2003) propose an explanation by separating two measures for plant-level productivity: a revenue-based measure and a quantity-based measure. They find that entrants are more productive than incumbents in terms of the quantity-based measure, but not in the revenue-based measure because entrants charge a lower price on average. Hence, more productive entrants can appear less profitable when prices are not observed.

[26] For evidence on the cross-cohort and within-cohort productivity distribution, see Baldwin (1995), Balk and Gort (1993), Foster, Haltiwanger and Syverson (2003). For evidence on cohort productivity dynamics, see Balk and Gort (1993) and Jensen, McGuckin and Stiroh (2000).

[27] See Hall (1987), Evans (1987), Montgomery and Wascher (1988), Dunne, Roberts and Samuelson (1989), Bresnahan and Raff (1991), Bahk and Gort (1993), Caves (1998), Davis and Haltiwanger (1999), and Jensen, McGuckin and Stiroh (2000).

tivated in Section 3 by comparative static exercises on the steady state equilibrium. Section 4 numerically solves the model with stochastic demand fluctuations and studies its quantitative implications for productivity using data on U.S. manufacturing job flows. I conclude in Section 5.

3.2 A Renovating Industry with Learning

This section describes a learning industry that experiences exogenous technological progress. New firms that capture the leading technology are continuously being created, and outdated firms are being destroyed. Firms enter with different idiosyncratic productivity. As time passes by, good firms survive and bad firms leave. Allocative inefficiency comes from costly entry and time-consuming learning.

3.2.1 Firms

I consider an industry where labor and capital combine in fixed proportions to produce a homogenous output. There is a continuum of firms, each hiring one worker, so that a job is created when a firm enters and a job is destroyed when a firm exits. Each firm is characterized by two components:

1. Vintage.

2. idiosyncratic productivity

There is an exogenous technological progress $\{A_t\}_0^\infty$ that grows at a constant rate $\gamma > 0$ so that

$$A_t = A_0 \cdot (1+\gamma)^t,$$

where A_0 is a constant. When a firm that enters the industry, it embodies the leading technology, which becomes its vintage and will affect its production afterward. I assume that, only entrants have access to the updated technology, incumbents cannot retool. Since technology grows exogenously, young firms are always technologically more advanced than old firms. With a as the firm age, the vintage of a firm of age a in period t is A_{t-a}. Apparently:

$$A_{t-a} = A_0 \cdot (1+\gamma)^{t-a}.$$

At the time of entry, a firm is endowed with idiosyncratic productivity θ. Hence, firms of the same vintage differ in idiosyncratic productivity. θ can represent the talent of the manager as in Lucas (1978), or alternatively, the location of the store, the organizational structure of the production process, or its fitness to the embodied technology.[28] The key assumption regarding θ is that its value, although fixed at the time of entry, is not directly observable. We can think of some real-world cases that reflect this assumption. For example, when a firm adopts new technology or introduces a new product, it needs to make many decisions, such as picking a manager to take charge of the production or choosing a location to sell the product. Although all firms try to make the best decisions possible, the outcome of their choices is uncertain and will be tested via market performance. Furthermore, their investments are irreversible; once a manager has signed the contract and a store is built, it becomes costly to make a new choice. Hence, the value of θ, as the consequence of a firm's random decisions, is unobservable and remains constant afterward.

[28] Since a firm is identical to a job under this set-up, θ can also be interpreted as "match quality." See Pries (2004).

A firm of age a and idiosyncratic productivity θ produces output in period t, according to

$$q_t(a, \theta) = A_{t-a} \cdot x_t = A_0 \cdot (1 + \gamma)^{t-a} \cdot x_t, \tag{3.1}$$

where

$$x_t = \theta + \varepsilon_t.$$

The shock ε_t is an i.i.d. random draw from a fixed distribution that masks the influence of θ on output. I set the operating cost of a firm (including wages) to 1 by normalization, and let P_t denote the output price in period t. Then the profit generated by a firm of age a and idiosyncratic productivity θ in period t is

$$\pi_t (a, \theta) = P_t \cdot A_0 \cdot (1 + \gamma)^{t-a} \cdot (\theta + \varepsilon_t) - 1. \tag{3.2}$$

Both $q_t(a, \theta)$ and $\pi_t (a, \theta)$ are directly observable. Since the firm knows its vintage, it can infer the value of x_t. The firm uses its observations of x_t to learn about θ.

3.2.2 "All-Or-Nothing" Learning

Firms are price takers and profit maximizers. They attempt to resolve the uncertainty about θ to decide whether to continue or terminate the production. The random component ε_t represents transitory factors that are independent of the idiosyncratic productivity θ. Assuming that ε_t has mean zero, we have

$$E_t(x_t) = E_t(\theta) + E_t(\varepsilon_t) = E_t(\theta).$$

Given knowledge of the distribution of ε_t, a sequence of observations of x_t allows the firm to learn about its θ. Although a continuum of potential values for θ is more realistic, for simplicity it is assumed here that there are only two values: θ_g for a good firm and θ_b for a bad firm. Furthermore, ε_t is assumed to be distributed uniformly on $[-\omega, \omega]$. Therefore, a good firm will have x_t each period as a random draw from a uniform distribution over $[\theta_g - \omega, \theta_g + \omega]$, while the x_t of a bad firm is drawn from an uniform distribution over $[\theta_b - \omega, \theta_b + w]$. Finally, θ_g, θ_b and ω satisfy $0 < \theta_b - \omega < \theta_g - \omega < \theta_b + \omega < \theta_g + \omega$.

Pries (2004) shows that the above assumptions give rise to an "all-or-nothing" learning process. With an observation of x_t within $(\theta_b + \omega, \theta_g + \omega]$, the firm learns with certainty that it is a good idiosyncratic productivity; conversely, an observation of x_t within $[\theta_b - \omega, \theta_g - \omega)$ indicates that it is a bad idiosyncratic productivity. However, an x_t within $[\theta_g - \omega, \theta_b + \omega]$ does not reveal anything, since the probabilities of falling in this range as a good firm and as a bad firm are the same (both equal to $\frac{2\omega + \theta_b - \theta_g}{2\omega}$).

This all-or-nothing learning simplifies my model considerably. I let θ^e represent the expected θ. Since it is θ^e instead of θ that affects firms' decisions, there are three idiosyncratic productivity of firms corresponding to the three values of θ^e: firms with $\theta^e = \theta_g$, firms with $\theta^e = \theta_b$, and firms with $\theta^e = \theta_u$, the prior mean of θ. I define "unsure firms" as those with $\theta^e = \theta_u$. I further assume that the unconditional probability of $\theta = \theta_g$ is φ, and let $p \equiv \frac{\theta_g - \theta_b}{2\omega}$ denote the probability of true idiosyncratic productivity being revealed every period. Firms enter the market as unsure; thereafter, every period they stay unsure with probability $1 - p$, learn they are good with probability $p \cdot \varphi$ and learn they are bad with probability $p \cdot (1 - \varphi)$. Thus, the evolution of θ^e from the time of entry is a Markov process with values $(\theta_g, \theta_u, \theta_b)$, an

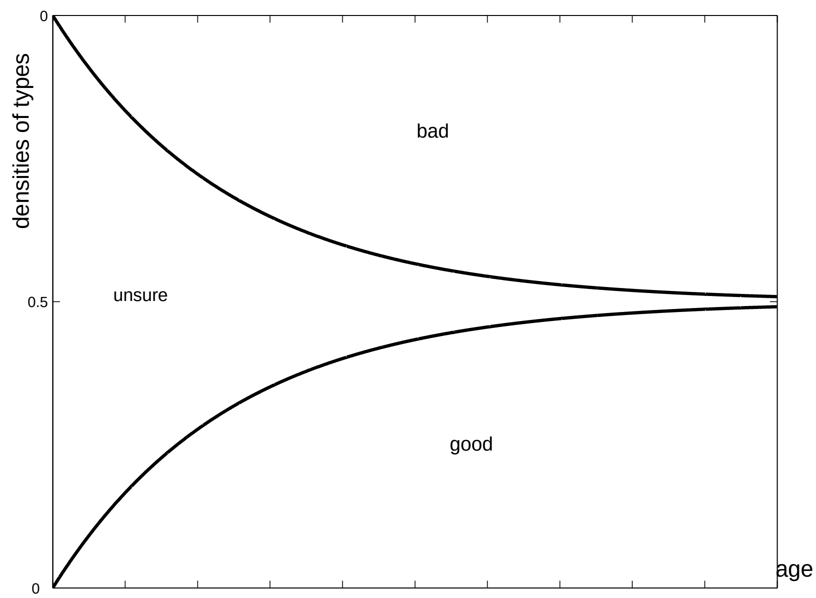


Figure 8: Dynamics of a Birth Cohort: the distance between the concave curve and the bottom axis measures the density of firms with $\theta^e = \theta_g$; the distance between the convex curve and the top axis measures the firms with $\theta^e = \theta_b$; the distance between the two curves measures the density of unsure firms (firms with $\theta^e = \theta_u$).

initial probability distribution:

$$\begin{pmatrix} 0, & 1, & 0 \end{pmatrix},$$

and a transition matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ p \cdot \varphi \ , & 1 - p \ , & p \cdot (1 - \varphi) \\ 0 & 0 & 1 \end{pmatrix}.$$

If firms were to live forever, eventually all uncertainty would be resolved because the market would provide enough information to reveal each firm's idiosyncratic productivity. The limiting probability distribution as a goes to ∞ is

$$\left(\varphi, \quad 0, \quad (1-\varphi) \right).$$

Because there is a continuum of firms, it is assumed that the law of large numbers applies, so that both φ and p are not only the probabilities but also the fractions of unsure firms with $\theta = \theta_g$, and of firms who learn θ each period, respectively. Hence, *ignoring firm exit for now*, I have the densities of three groups of firms in a cohort of age a as

$$\left(\varphi \cdot [1-(1-p)^a], \quad (1-p)^a, \quad (1-\varphi) \cdot [1-(1-p)^a] \right),$$

which implies an evolution of the idiosyncratic-productivity firm distribution within a birth cohort as shown in Figure 8, with the horizontal axis depicting the age of a cohort *across time*. The densities of firms that are certain about their idiosyncratic productivity, whether good or bad, grow as a cohort ages. Moreover, the two "learning curves" (depicting the evolution of densities of good firms and bad firms) are concave. This feature is defined as the decreasing property of marginal learning in Jovanovic (1982): the marginal learning effect decreases with firm age, which in my model is reflected by the fact that the marginal number of learners decreases with cohort age. The convenient feature of all-or-nothing learning is that, on the one hand, it implies that any single firm learns "suddenly", which allows us to easily keep track of the cross-section distribution of beliefs, while on the other hand, it still implies "gradual learning" at the cohort level.

However, there is more that Figure 8 can tell. If we let the horizontal axis depict the cross-sectional distribution of firm ages at any instant, then Figure 4 can be interpreted as the firm distribution across ages and idiosyncratic productivity of an industry that features constant entry but no exit. In this industry, cohorts continuously enter in the same size and experience the same dynamics afterward, so that at any one time, different life-stages of different birth cohorts overlap, giving rise to the distribution in Figure 4. Under this interpretation, Figure 4 indicates that at any instant older cohorts contain fewer unsure firms, because they have lived longer and learned more.

3.2.3 The Recursive Competitive Equilibrium

The following sequence of events is assumed to occur within a period. First, entry and exit occur after firms observe the aggregate state. Second, each surviving firm pays a fixed operating cost to produce. Third, the aggregate price is realized. Fourth, firms observe revenue and update beliefs. Then, another period begins.

With the above setup, this subsection considers a *recursive competitive equilibrium* definition which includes as a key component the law of motion of the aggregate state of the industry. The aggregate state is (F, D). F denotes the distribution (measure) of firms across vintages and idiosyncratic productivity. The part of F that measures the number of firms with belief θ^e and age a is denoted $f(\theta^e, a)$. D is an exogenous demand parameter; it captures aggregate conditions and is fully observable. The law of motion for D is exogenous, described by D's transition matrix. The law of motion for F is denoted H so that $F' = H(F, D)$. The sequence of events implies that H captures the influence of entry, exit and learning.

Three assumptions characterize the equilibrium: firm rationality, free entry and competitive pricing.

Firm Rationality: firms are assumed to have rational expectations; their decisions are forward-looking. In period t, a firm with age a and belief θ^e expects its profit in period $s \geq t$ to equal

$$A_{t-a} \cdot E(P_s|F_t, D_t) \cdot \theta^e - 1.$$

$E_t(P_s|F_t, D_t)$ implies that firms need to observe (F, D) to predict the sequence of prices from today onward. Therefore, the relevant state variables for a firm are its vintage, its belief about its true idiosyncratic productivity, and the aggregate state (F, D). I let $V(\theta^e, a; F, D)$ be the expected value, for a firm with belief θ^e and age a, of staying in operation for one more period and optimizing afterward, when the aggregate state is (F, D). Then V satisfies:

$$V(\theta^e, a; F, D) = E[\pi(\theta^e, a)|F, D] + \beta E[\max(0, V(\theta^{e\prime}, a+1; F', D'))|F, D] \quad (3.3)$$

subject to

$$F' = H(F, D)$$

and the exogenous laws of motion for D and θ^e (driven by all-or-nothing learning).

Since firms enter as unsure, firm rationality implies that entry occurs if and only if $V(\theta_u, 0; F, D) > 0$. Meanwhile, a firm with belief θ^e and age a exits if and only if $V(\theta^e, a; F, D) < 0$.

Free entry: new firms are free to enter at any instant, each bearing an entry cost c. The entry cost can be interpreted as the cost of establishing a particular location or the cost of finding a manager. Assuming $f(\theta_u, 0; F, D)$ represents the size of the entering cohort when the aggregate state is (F, D), and letting c represent the entry cost, I have

$$c = C\left(f\left(\theta_u, 0; F, D\right)\right), c > 0 \text{ and } C' \geq 0. \tag{3.4}$$

I let the entry cost depend positively on the entry size to capture the idea that, for the industry as a whole, *fast* entry is costly and adjustment may not take place instantaneously. This can arise from a limited amount of land available to build production sites or an upward-sloping supply curve for the industry's capital stock.[29] The free entry condition equates a firm's entry cost to its value of entry, and can be written as

$$V\left(\theta_u, 0; F, D\right) = C\left(f\left(\theta_u, 0; F, D\right)\right). \tag{3.5}$$

As more new firms enter, the entry cost is driven up until it reaches the value of entry. At this point, entry stops.

Competitive Pricing: the output price is competitive; the price level is given by

$$P\left(F, D\right) = \frac{D}{Q\left(F, D\right)} \tag{3.6}$$

Q represents aggregate output; it equals the the sum of production over heterogeneous firms. Given (3.1), the sequence of events implies that:[30]

$$Q\left(F, D\right) = Q\left(F'\right) = A \sum_a \sum_{\theta^e} (1+\gamma)^{-a} \cdot \theta^e \cdot f'\left(\theta^e, a\right), \tag{3.7}$$

where A represents the industry leading technology when the aggregate state is (F, D).

[29] See Subsection 3.3.1 for further discussion.

[30] Q is the sum of realized output rather than expected output, since the contribution to aggregate output by each firm depends on its true type θ rather than θ^e. However, with a continuum of firms, the law of large numbers implies that the random noises and the expectation errors cancel out in each cohort, so that the sum of realized output equals the sum of expected output.

$f'(\theta^e, a)$ measures the number of operating firms with θ^e and a *after entry and exit.* $f'(\theta^e, a)$ belongs to F', the updated firm distribution. Since $F' = H(F, D)$, Q is a function of (F, D).

(3.6) implies that high output drives down the price. (3.7) implies that Q depends not only on the number of firms in operation, but also on their distribution. More firms yield higher output and drive down the price; the more the distribution is skewed toward younger vintages and better idiosyncratic productivity, the higher the output and the lower the price.

With the above three conditions, I have the following:

> Definition: A *recursive competitive equilibrium* is a law of motion H, a value function V, and a pricing function P such that (i) V solves the firm's problem; (ii) P satisfies (3.6) and (3.7); and (iii) H is generated by the decision rules suggested by V and the appropriate summing-up of entry, exit and learning.

An additional assumption is made to simplify the model:

> Assumption: Given values for other parameters, the value of θ_b is so low that $V(\theta_b, a; F, D)$ is negative for any (F, D) and a.

This assumption implies that bad firms always exit, so that at any one time, there are only two idiosyncratic productivity of firms in operation – unsure and good.

The following proposition characterizes the value function V and the corresponding exit ages of heterogeneous firms.

> Proposition 3.1: $V(\theta^e, a; F, D)$ is strictly decreasing in a, holding θ^e constant, and strictly increasing in θ^e, holding a constant; therefore, there is a cut-

off age $\overline{a}(\theta^e; F, D)$ for each idiosyncratic productivity, such that firms of idiosyncratic productivity θ^e and age $a \geq \overline{a}(\theta^e; F, D)$ exit before production takes place; furthermore, $\overline{a}(\theta_g; F, D) \geq \overline{a}(\theta_u; F, D)$.

The proof for Proposition 3.1 presented in the appendix is not restricted to all-or-nothing learning. Hence, Proposition 3.1 holds for *any* learning process. It follows from the fact that firms with smaller a and higher θ^e have a higher expected value of staying. As V is strictly decreasing in a, firms with belief θ^e that are older than $\overline{a}(\theta^e; F, D)$ exit; as the expected value of staying is strictly increasing in θ^e, a good firm stays longer than an unsure firm.

3.3 Cleansing and Scarring

The firm distribution F enters the model as a state variable, which makes it difficult to characterize the dynamics generated by demand fluctuations. However, similar studies find that the effects of temporary changes in aggregate conditions are qualitatively similar to the effects of permanent changes.[31] Therefore, I begin in this section with comparative static exercises on the steady-state equilibrium. The comparative static exercises capture the essence of industry dynamics as well as how demand can affect the labor allocation, and thus provide a more rigorous intuition for the scarring and cleansing effects described in the introduction. In the next section, I will turn to a numerical analysis of the model's response to stochastic demand fluctuations and confirm that the results from the comparative static exercises carry over.

[31] See Mortensen and Pissarides (1994), Caballero and Hammour (1994 and 1996), and Barlevy (2003).

3.3.1 The Steady State

I define a steady state as a recursive competitive equilibrium with time-invariant aggregate states.[32] It satisfies two additional conditions, (i) D is and is perceived as time-invariant: $D' = D$. (ii) F is time-invariant: $F' = H(F, D)$. Since H is generated by entry, exit and learning, a steady state must feature time-invariant entry and exit for $F = H(F, D)$ to hold. Thus, a steady state equilibrium can be summarized by $\{f(0), \overline{a}_g, \overline{a}_u\}$, with $f(0)$ as the entry size, $\overline{a}_g$ as the maximum age for good firms, and $\overline{a}_u$ as the maximum age for unsure firms. The next proposition establishes the existence of a unique steady-state equilibrium. The proof is presented in the appendix.

Proposition 3.2: With D constant over time, there exists a unique time-invariant $\{f(0), \overline{a}_g, \overline{a}_u\}$ that satisfies the conditions of firm rationality, free entry and competitive pricing.

The steady-state labor distribution and job flows are illustrated in Figure 9. Like Figure 8, there are two ways to interpret Figure 9. First, it displays the steady-state life-cycle dynamics of a representative cohort with the horizontal axis depicting the cohort age *across time*. Firms enter in size $f(0)$ as unsure. As the cohort ages and learns, bad firms are thrown out so that the cohort size declines; good firms are realized, so that the density of good firms increases. After age $\overline{a_u}$, all unsure firms exit because their vintage is too old to survive with $\theta^e = \theta_u$. However, firms with $\theta^e = \theta_g$ stay. Afterwards, the cohort contains only good firms and the number of good firms remains constant because learning has stopped. Good firms live until $\overline{a_g}$.

[32] The term "steady state" follows Caballero and Hammour (1994). Despite its name, the steady-state price decreases while the steady-state average labor productivity increases over time due to technological progress.

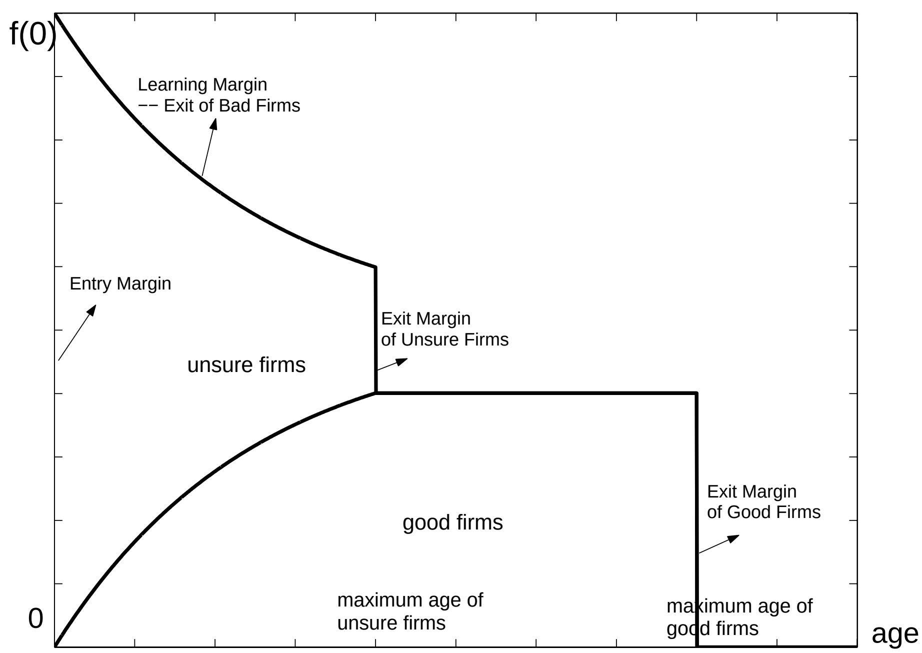


Figure 9: The Steady-state Labor Distribution and Job Flows: the distance between the lower curve (extended as the horizontal line) and the bottom axis measures the density of good firms; the distance between the two curves measures the density of unsure firms.

The vintage after $\overline{a_g}$ is too old even for good firms to survive.

Second, Figure 9 also displays the firm distribution across ages and idiosyncratic productivity at any one time, with the horizontal axis depicting the cohort age *cross section*. At the steady state, firms of different ages coexist. Since older cohorts have lived longer and learned more, their size is lower and their density of good firms is higher. Cohorts older than $\overline{a_u}$ are of the same size and contain only good firms. No cohort is older than $\overline{a_g}$.

Despite its time-invariant structure, the industry experiences continuous entry and exit. With entry, jobs are created; with exit, jobs are destroyed. From a pure accounting point of view, there are three margins for job flows: the entry margin, the exit margins of good firms and unsure firms, and the learning margin. Two forces – learning and creative destruction – interact together to drive job flows. At the entry margin, creative destruction drives in new vintages. At the exit margins, it drives out old vintages. At the learning margin, bad firms are selected out. Because of creative destruction, average labor productivity grows at the technological pace γ. Because of learning, the productivity distribution among older cohorts is more skewed toward good firms. For cohorts older than $\overline{a_u}$, labor is employed only at good firms.

3.3.2 Comparative Statics: Cleansing and Scarring

The previous subsection has shown that for a given demand level, there exists a steady-state equilibrium summarized by $\{f(0), \overline{a_g}, \overline{a_u}\}$. In this subsection, I establish that across steady states corresponding to different demand levels, the model delivers the conventional cleansing effect promoted in the previous literature, as well as an additional scarring effect. The two effects are formalized in Propositions 3.3 and 3.4.

Proposition 3.3: In a steady-state equilibrium, the exit age for firms with a given belief is weakly increasing in the demand level and the job destruction rate is weakly decreasing in the demand level.

A detailed proof is included in the appendix. To understand Proposition 3.3, compare two steady states with different demand levels, $D_h > D_l$. For any time t, (3.6) suggests that the steady state with D_l features either a lower price, or a lower output, or both. Now assume initially that the lower demand is fully reflected as a lower output and the prices of the two steady states are identical. Then firms' profitability in the two steady states would also be identical: $V_l(\theta^e, a) = V_h(\theta^e, a)$ for any θ^e and a. Free entry and the exit conditions suggest that identical value functions lead to identical entry size and exit ages, and thus an identical firm distribution. With firm-level output of a given age and idiosyncratic productivity independent of demand, identical cross-sectional distributions imply identical aggregate output, *which contradicts our assumption*. Therefore, we can conclude that the low-demand steady state must feature a lower price compared to the high-demand steady state, so that $V_l(\theta^e, a) < V_h(\theta^e, a)$ for any θ^e and a. Since $V(\theta^e, a)$ strictly decreases in a, the cut-off age that solves the $V(\theta^e, a) = 0$ must be lower for lower demand. Intuitively, lower demand tends to drive down the price so that some firms that are viable in a high-demand steady state are not viable when demand is low.

Moreover, the following equation is derived by combining the exit conditions for unsure and good firms:

$$\left(\frac{\theta_u}{\theta_g} + \frac{p\varphi\beta}{1+\gamma-\beta}\right)(1+\gamma)^{\bar{a}_g-\bar{a}_u} = 1 + \frac{p\varphi\beta}{1-\beta} - \frac{p\varphi\beta\gamma}{(1-\beta)(1+\gamma-\beta)}\beta^{\bar{a}_g-\bar{a}_u} \tag{3.8}$$

I prove in the appendix that (3.8) gives an unique solution for $\overline{a}_g - \overline{a}_u$ as long as $\theta_g > \theta_u$. Since D does not enter (3.8), $\overline{a}_g - \overline{a}_u$ is independent of demand: $\frac{d(\overline{a}_g - \overline{a}_u)}{dD} = 0$. (3.8) suggests that the demand level does not affect the gap between the exit ages of good and unsure firms.

The steady-state job destruction rate, denoted jd^{ss}, equals the following:[33]

$$jd^{ss} = \frac{1}{\overline{a_u} \cdot \varphi + [\frac{1-\varphi}{p} + (\overline{a_g} - \overline{a_u}) \cdot \varphi] \cdot [1 - (1-p)^{\overline{a_u}+1}]}. \tag{3.9}$$

Since $(\overline{a_g} - \overline{a_u})$ is independent of D, demand affects jd^{ss} only through its impact on $\overline{a_u}$: $\frac{d(jd^{ss})}{d(D)} = \frac{d(jd^{ss})}{d(\overline{a_u})} \cdot \frac{d(\overline{a_u})}{d(D)}$. I prove in the appendix that $\frac{d(jd^{ss})}{d(\overline{a_u})} \leq 0$, which, together with $\frac{d(\overline{a_u})}{d(D)} \geq 0$, implies $\frac{d(jd^{ss})}{d(D)} \leq 0$. Put intuitively, a high-demand steady state allows both unsure firms and good firms to live longer, so that fewer jobs are destroyed at the exit margins.

To summarize, Proposition 3.3 argues that the steady state with lower demand features younger exit ages and a higher job destruction rate. In other words, it suggests that more firms are cleared out in an environment that is more difficult for survival.

If the above story suggested by comparative statics carries over when D fluctuates stochastically over time, then my model delivers a conventional "cleansing" effect, in which average firm age falls during recessions so that recessions direct resources to

[33] According to Davis and Haltiwanger (1992), the job destruction rate at time t is defined as:

$$\frac{2 * \text{Jobs destroyed in period } t}{[(\text{number of jobs at the beginning of period } t) + (\text{number of jobs at the beginning of period } t+1)]}.$$

With constant total number of jobs, the steady-state job destruction rate equals the ratio of jobs destroyed at the learning and exit margins over the total number of jobs. The expression of jd^{ss} applies not only to a steady state, but also to any industry equilibrium that features time-invariant entry and exit. See Subsection 4.2 for further discussions on jd^{ss}.

younger, more productive vintages. However, once learning is allowed, we also need to take into account the allocation of labor across idiosyncratic productivity. With only two true idiosyncratic productivity, good and bad, the idiosyncratic productivity distribution of labor can be summarized by the fraction of labor at good firms. A higher fraction suggests a more efficient cross-idiosyncratic productivity allocation of labor. The next proposition establishes how the level of demand affects this ratio in a steady state.

> Proposition 3.4: In a steady state equilibrium, the fraction of labor at good firms is weakly increasing in the demand level.

It can be shown that the steady-state fraction of labor at good firms, denoted l_g^{ss}, equals:

$$l_g^{ss} = 1 - \frac{(1-\varphi)}{\frac{p\varphi\overline{a_u}}{1-(1-p)^{\overline{a_u}}} + (1-\varphi) + p\varphi\left(\overline{a_g} - \overline{a_u}\right)}.$$

Again, since $(\overline{a_g} - \overline{a_u})$ is independent of D, demand affects l_g^{ss} only through its impact on $\overline{a_u}$: $\frac{d\left(l_g^{ss}\right)}{d(D)} = \frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})} \cdot \frac{d(\overline{a_u})}{d(D)}$. I prove $\frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})} \geq 0$ in the appendix, which, together with $\frac{d(\overline{a_u})}{d(D)} \geq 0$, implies $\frac{d\left(l_g^{ss}\right)}{d(D)} \geq 0$.

My analysis suggests that the impact of demand on the fraction of labor at good firms comes from its impact on the exit age of unsure firms. To understand this result intuitively, consider Figure 10.

Figure 10 displays the steady-state industry structures corresponding to two demand levels.[34] The cleansing effect formalized in Proposition 3.3 is shown as the

[34] The entry sizes of the two steady states, although different, are normalized as 1. Since the steady state features time-invariant entry and all cohorts are the same size, entry size matters only as a scale.

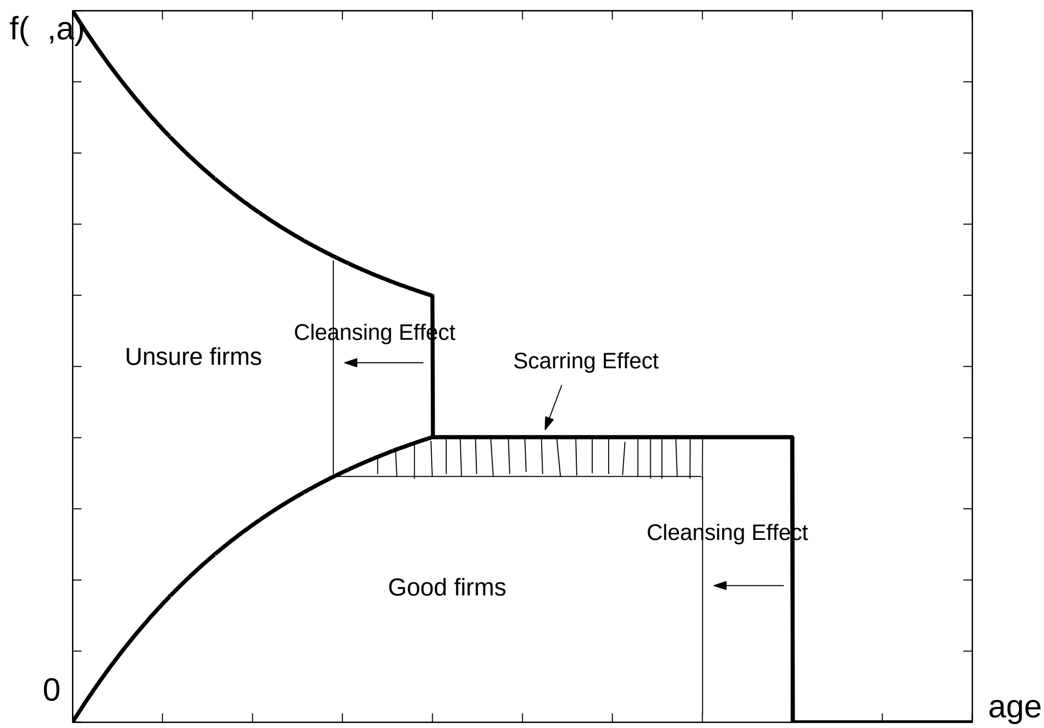


Figure 10: Cleansing and Scarring

leftward shift of the two exit margins. The shifted margins clear out old firms that could be either good or unsure. However, the leftward shift of the *unsure exit margin* also reduces the number of *older good firms*. The latter effect, shown as the shaded area in Figure 10, is the scarring effect of recessions.

The scarring effect stems from learning. New entrants begin unsure of their idiosyncratic productivity, although a proportion φ are truly good. Over time, more and more bad firms leave while good firms stay. Since learning takes time, the number of "potentially good firms" that realize their true idiosyncratic productivity depends on how many learning chances they have. If firms could live forever, eventually all the potentially good firms would get to realize their true idiosyncratic productivity. But a finite life span of unsure firms implies that if potentially good firms do not learn before age $\overline{a}_u$, they exit and thus forever lose the chance to learn. Therefore, $\overline{a}_u$ represents not only the exit age of unsure firms, but also the number of learning opportunities. A low $\overline{a_u}$ allows potentially good firms fewer chances to realize their true idiosyncratic productivity, so that the number of *old good firms* in operation after age $\overline{a_u}$ is also reduced.

Hence, the industry suffers from uncertainty; it tries to select out bad firms but the group of firms it clears at age $\overline{a_u}$ includes some firms that are truly good. The number of clearing mistakes the industry makes at $\overline{a_u}$ depends on the size of the unsure exit margin, which in turn depends on the value of $\overline{a_u}$.[35] When a drop in demand reduces the value of $\overline{a}_u$, this reduces the number of learning opportunities, allows fewer good firms to become old and thus shifts the labor distribution toward bad firms.

[35] The all-or-nothing learning suggests that the number of truly good firms cleared out at $\overline{a_u}$ equals $f(0)(1-p)^{\overline{a_u}}\varphi$.

To summarize from Propositions 3.3 and 3.4, a low-demand steady state features a better average vintage, yet a less efficient cross-idiosyncratic productivity distribution of labor. If the comparative static results carry over when demand fluctuates stochastically, then recessions will have both a conventional cleansing effect, shifting resources to better vintages, and a scarring effect, shifting resources to bad idiosyncratic productivity. The two effects are directly related to each other: it is the cleansing effect that significantly reduces learning opportunities and hence prevents more firms from realizing their potential.

When we move beyond steady states to allow for cyclical fluctuations, the intuition behind "cleansing and scarring" still carries over. Consider Figure 6. Both exit margins shift as soon as demand drops so that the cleansing effect takes place immediately.[36] However, the scarring effect takes place gradually. When a recession first arrives, the group of firms already in the shaded area in Figure 6 will not leave despite the shift in exit margins, since they know their true idiosyncratic productivity to be good. They leave gradually as the recession persists. At this point, the scarring effect starts to take place: the reduced $\overline{a_u}$ allows fewer good firms to survive past $\overline{a_u}$. The shaded area would eventually be left blank, and the "scar" left by recessions would surface.

3.3.3 Sensitivity Analysis

Two modifications are examined in this subsection to check the robustness of my results from the comparative static exercises: first, I allow the entry cost to be inde-

[36] My numerical exercises imply that when demand falls, these margins initially shift more than suggested by the comparative static exercises. The margins shift back partially as the recession persists. A detailed discussion of this phenomenon is contained in Section 4.

pendent of entry size; second, I allow the process of learning to be more complicated than "all-or-nothing".

Entry Cost Independent of Entry Size The previous subsection has argued that the shift of the exit margins creates both a cleansing effect and a scarring effect. Now, focus on the entry side. How does demand affect entry, and how would alternative assumptions on entry affect my results?

To address these questions, recall that the free entry condition requires $V(\theta_u, 0) = C(f(\theta_u, 0))$, and C is assumed to depend positively on entry size. Since low demand reduces the value of entry by driving down profitability, $C'(f(\theta_u, 0)) > 0$ implies less entry (smaller $f(\theta_u, 0)$) for the low-demand steady state. Hence, an industry in my model has two margins along which it can accommodate low demand. It can either reduce entry, or increase exit by shifting the exit margins. The issue is which of these two margins will respond when demand falls, and by how much. If the drop in demand level can be fully incorporated as a decrease in entry size, the exit margins might not respond.

The extreme case that the entry margin *exclusively* accommodates demand fluctuations is defined as the "full-insulation" case in Caballero and Hammour (1994). They argue that creation (entry) "insulates" destruction (exit), and the extent of the insulation effect depends on the cost of *fast* entry, that is, $C'(f(\theta_u, 0))$. The full-insulation case occurs when $C'(f(\theta_u, 0)) = 0$. The intuition is as follows. If entry cost is independent of entry size, then *fast* entry is costless and the adjustment on the entry margin becomes instantaneous. When demand falls, entry will adjust to such a level that aggregate output falls by the same proportion, which keeps price at the same level. Then the value of staying remain unaffected, and the exit margins do not respond. Hence, with entry cost independent of entry size, *there is neither a cleansing effect nor a scarring effect.*

Two remarks can be made. First, in reality, an industry may not be able to create all the necessary production units instantaneously. Goolsbee (1998) shows empirically that higher investment demand drives up both the equipment prices and the wage of workers producing the capital goods. His findings suggest that as more firms enter and increase the demand for capital, it becomes increasingly costly to purchase capital. As another intuitive example, when more new stores are built, land prices and rentals usually rise. Therefore, $C'(f(\theta_u, 0)) > 0$ seems more reasonable. Second, data does not support the assumption that $C'(f(\theta_u, 0)) = 0$. In the full-insulation case, job creation fully accommodates demand fluctuations and job destruction does not respond. This contradicts the large and robust evidence that job destruction is *more* responsive than job creation to the business cycle.[37]

More Complicated Learning As I have argued in subsection 2.2, the all-or-nothing learning with a uniform distribution of random noise simplifies the analysis considerably. But how restrictive is it? Would the scarring effect carry over with a more complicated process of learning?

In general, we can define the scarring effect as a drop in the fraction of labor at good firms. To look at the scarring effect from a different angle, suppose we divide firms into two groups, young and old.[38] With l_g^o denoting the fraction of labor at good firms among the old, l_g^y as the fraction among the young, f^y as the density of young firms and f^o as the density of old firms, the fraction of labor at good firms for

[37] See footnote 6.

[38] The cut-off age to define "young" and "old" is arbitrarily chosen. Changing this cut-off age does not affect the analysis that follows.

the industry as a whole, l_g, can be written as:

$$l_g = \frac{f^y l_g^y + f^o l_g^o}{f^y + f^o} = \frac{l_g^y + l_g^o \frac{f^o}{f^y}}{1 + \frac{f^o}{f^y}}.$$

The first order derivative of l_g with respect to $\frac{f^o}{f^y}$ equals:

$$\frac{d\left(l_g\right)}{d\left(\frac{f^o}{f^y}\right)} = \frac{l_g^o - l_g^y}{1 + \frac{f^o}{f^y}}.$$

which is greater than or equal to zero as long as $l_g^o - l_g^y \geq 0$, which should hold for *any* learning process, since old firms have experienced more learning. Hence, the scarring effect of recessions should occur under any idiosyncratic productivity of learning as long as recessions reduce the ratio of old to young firms ($\frac{f^o}{f^y}$), which by definition will be true in any model in which recessions cleanse the economy of older vintages. Intuitively, the scarring effect suggests that recessions shift resources toward younger firms, so that there cannot be as much learning taking place as in booms.

Now suppose we assume a more complicated learning process with normally distributed random noise, so that the signals received by good firms are normally distributed around θ_g and the signals received by bad firms are normally distributed around θ_b. In that case, a firm can never know *for certain* that it is good or bad, and posterior beliefs are distributed continuously between θ_b and θ_g. The expected value of staying would still depend positively on θ^e and negatively on age. Thus, given the aggregate state, there would be a cut-off age for each belief, $\overline{a}\left(\theta^e; F, D\right)$, such that firms with belief θ^e do not live beyond $\overline{a}\left(\theta^e; F, D\right)$.

With a recession, the value of staying across all ages and idiosyncratic productivity falls, so that for each belief θ^e, the cut-off age $\overline{a}(\theta^e; F, D)$ becomes younger. Hence, the firm distribution tilts toward younger ages and $\frac{f^o}{f^y}$ falls. Since $\frac{d(l_g)}{d\left(\frac{f^o}{f^y}\right)} \geq 0$, a fall in $\frac{f^o}{f^y}$ drives down the ratio of good firms and creates the scarring effect. Although this analysis is preliminary,[39] we can still argue that recessions would allow for less firm learning, so the scarring effect would carry over even with a more complicated process of learning.

3.4 Quantitative Implications with Stochastic Demand Fluctuations

I establish in Section 3 that across steady states, variations in demand induce competing cleansing and scarring effects on productivity. In this section, I address whether the two effects carry over when demand fluctuates stochastically, and which effect dominates quantitatively.

This section turns to numerical techniques to analyze a stochastic version of my model in which the demand level follows a two-state Markov process with values $[D_h, D_l]$ and transition probability μ. Throughout this section, firms expect the current demand level to persist for the next period with probability μ, and to change with probability $1 - \mu$.

I first describe my computational strategy, which follows Krusell and Smith (1998) by shrinking the state space into a limited set of variables and showing that these variables' laws of motion can approximate the equilibrium behavior of firms in the simulated time series. Later in this section, I confirm that the basic insights from the comparative static exercises carry over with probabilistic business cycles. Then I

[39] For instance, the analysis cannot address the relative sizes of the cleansing effect on young firms versus old firms. Whether cleansing affects primarily young or old firms depends on the specifics of the learning process.

examine whether the scarring effect is likely to be empirically relevant. Specifically, I calibrate my model so that its equilibrium job destruction rate mimics the observed pattern in U.S. manufacturing. As I have argued, recessions clear out old firms, including some good firms that have not yet learned their idiosyncratic productivity. Therefore, the model allows us to use the job destruction rate to make inferences on the size of the cleansing and scarring effects.

3.4.1 Computational Strategy

The definition of the *recursive competitive equilibrium* in Section 2 implies that individual decision rules can be generated from the value functions V; by summing up the corresponding individual decision rules, we can get the laws of motion H, then trace out the evolution of industry structure. Therefore, the key computational task is to map F, the firm distribution across ages and idiosyncratic productivity, given demand level D, into a set of value functions $V(\theta^e, a; F, D)$. Unfortunately, the endogenous state variable F is a high-dimensional object. The numerical solution of dynamic programming problems becomes increasingly difficult as the size of the state space increases. To make the state space tractable, I define a variable X such that[40]

$$X(F) = \sum_a \sum_{\theta^e} (1+\gamma)^{-a} \cdot \theta^e \cdot f(\theta^e, a). \tag{3.10}$$

Combining (3.9) with (3.6) and (3.7), I get

$$P(F, D) \cdot A = \frac{D}{X(F')}.$$

[40] X can be interpreted as detrended output.

A is the leading technology; F' is the updated firm distribution after entry and exit; X' corresponds to F'; $P(F, D)$ is the equilibrium price in a period with initial aggregate state (F, D). Since $F' = H(F, D)$, the above equation can be re-written as

$$P(F, D) \cdot A = \frac{D}{X(H(F, D))}$$

Given these definitions, the single-period profitability of a firm of idiosyncratic productivity θ^e and age a, given aggregate state (F, D), equals

$$\pi(a, \theta; F, D) = \frac{D}{X(H(F, D))} \cdot (1+\gamma)^{-a} \cdot (\theta + \varepsilon) - 1. \tag{3.11}$$

Thus, the aggregate state (F, D) and its law of motion help firms to predict future profitability by suggesting sequences of X's from today onward under different paths of demand realizations. The question then is: what is the firm's critical level of knowledge of F that allows it to predict the sequence of $X's$ over time? Although firms would ideally have full information about F, this is not computationally feasible. Therefore I need to find an information set Ω that delivers a good approximation of firms' equilibrium behavior, yet is small enough to reduce the computational difficulty.

I look for an Ω through the following procedure. In step 1, I choose a candidate Ω. In step 2, I postulate perceived laws of motion for all members of Ω, denoted H_Ω, such that $\Omega' = H_\Omega(\Omega, D)$. In step 3, given H_Ω, I calculate firms' value functions on a grid of points in the state space of Ω applying value function iteration, and obtain the corresponding industry-level decision rules – entry sizes and exit ages across aggregate states. In step 4, given such decision rules and an initial firm distribution,[41] I simulate

[41] I start with a uniform firm distribution across types and ages. My numerical exercises suggest that the dynamic system of my model is stable and that the initial firm distribution does not affect

Ω	$\{X\}$
H_Ω	$H_x(X, D_h)$: $\log X' = 1.2631 + 0.8536 \log X$ $H_x(X, D_l)$: $\log X' = 2.4261 + 0.7172 \log X$
R^2	for D_h: 0.9876 for D_l: 0.9421
standard forecast error	for D_h: 0.0000036073% for D_l: 0.000030068%
maximum forecast error	for D_h: 0.000049895% for D_l: 0.00074675%
Den Haan & Marcet test statistic (χ^2_7)	0.8007

Table 5: The Estimated Laws of Motion and Measures of Fit

the behavior of a continuum of firms along a random path of demand realizations, and derive the implied aggregate behavior — a time series of Ω. In step 5, I use the stationary region of the simulated series to estimate the *implied* laws of motion and compare them with the *perceived* H_Ω; if different, I update H_Ω, return to step 3 and continue until convergence. In step 6, once H_Ω converges, I evaluate the fit of H_Ω in terms of tracking the aggregate behavior. If the fit is satisfactory, I stop; if not, I return to step 1, make firms more knowledgeable by expanding Ω, and repeat the procedure.

I start with $\Omega = \{X\}$ — firms observe X instead of F. I further assume that firms perceive the sequence of future coming $X's$ as depending on nothing more than the current observed X and the state of demand. The perceived law of motion for X is denoted H_x so that $X' = H_x(X, D)$. I then apply the procedure described above and simulate the behavior of a continuum of firms over 5000 periods. The results are presented in Table 5 As shown in Table 5, the estimated H_x is log-linear. The fit of H_x is quite good, as suggested by the high R^2, the low standard forecast error,

the result.

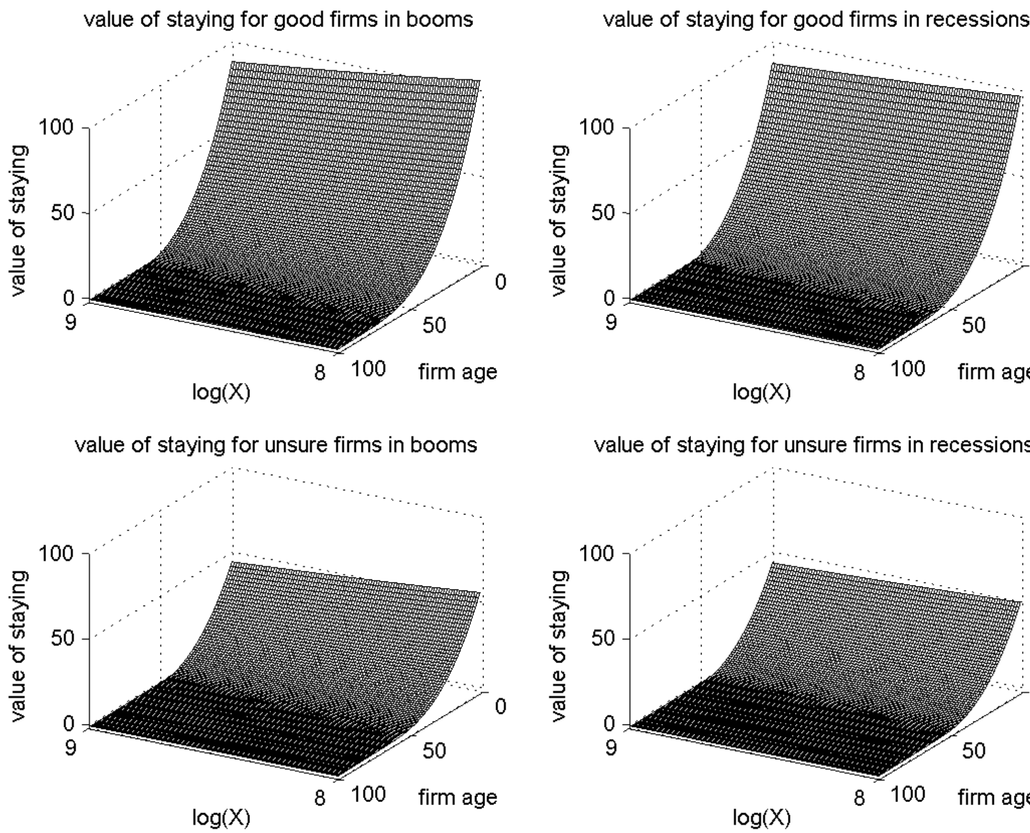


Figure 11: Expected Value of Staying: aggregate state variables are D and $\log X$ (the log of detrended output), firm-level state variables are firm age and belief (good or unsure); the parameter choices underlying these figures are summarized in Table 2 and discussed in Subsection 4.2.

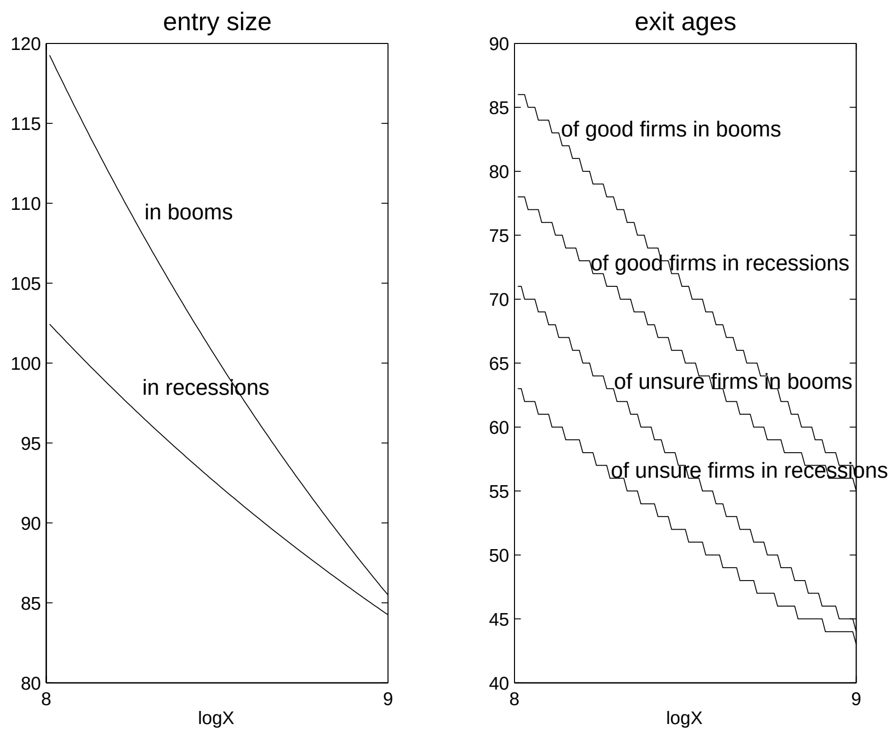


Figure 12: Industry-level Policy Functions: Entry Size and Exit Ages. Aggregate states are D (booms or recessions) and $\log X$ (the log of detrended output).

and the low maximum forecast error. The good fit when $\Omega = \{X\}$ implies that firms perceiving these simple laws of motion make only small mistakes in forecasting future prices. To explore the extent to which the forecast error can be explained by variables other than X, I implement the Den Haan and Marcet (1994) test using instruments $[1, X, \mu_a, \sigma_a, \gamma_a, \kappa_a, r_u]$, where μ_a, σ_a, γ_a, κ_a,r_u are the mean, standard deviation, skewness, and kurtosis of the age distribution of firms, and the fraction of unsure firms, respectively.[42] The test statistic is 0.8007, well below the critical value at the 1% level. This suggests that given the estimated laws of motion, I do not find much additional forecasting power contained in other variables. Nevertheless, I expand Ω further to include σ_a, the standard deviation of the age distribution of firms. The results when $\Omega = \{X, \sigma_a\}$ are presented in the appendix. The measures of fit do not change much.[43] Furthermore, the impact of changes in σ_a on the approximated value function is very small (less than 0.5%). This confirms that the inclusion of information other than X improves the forecast accuracy by only a very small amount.

Figure 11 displays the value of staying for heterogeneous firms as a function of a, θ^e, D and X ($\log X$). Figure 12 displays the corresponding optimal exit ages and entry sizes. The properties of value functions and exit ages stated in Proposition 3.2 are satisfied in both figures: given the aggregate state, the value of staying is increasing in the perceived idiosyncratic productivity θ^e and decreasing in firm age;

[42] Den Haan and Marcet (1994) offer a statistic for computing the accuracy of a simulation. It has an asymptotic χ^2 distribution under the null that the simulation is accurate. The statistic for my industry is given by $TB_T^{'}A_T^{-1}B_T$, where $B_T = \frac{1}{T}\sum u_{t+1} \otimes h(G_t)$, $A_T = \frac{1}{T}\sum u_{t+1}^2 \otimes h(G_t)\, h(G_t)'$, u_{t+1} is the expectation error for X_{t+1}(or $\log X_{t+1}$), and $h(G_t)$ is some function of variables dated t. I choose $h(G_t) = [1, X, \mu_a, \sigma_a, \gamma_a, \kappa_a, r_u]$, which gives my test statistic 7 degrees of freedom.

[43] Actually the fit during recessions becomes worse to some extent. Young (2002) adds an additional moment to the original Krusell & Smith approach, and also gets a worse measure of fit for the bad state (recessions). He attributes this result to numerical error.

parameters (pre-chosen)	value
productivity of bad firms: θ_b	1
productivity of good firms: θ_g	3.5
quarterly technological pace: γ	0.007
quarterly discount factor: β	0.99
parameters (calibrated)	**value**
high demand: D_h	2899
low demand: D_l	2464
prior probability of being a good firm: φ	0.14
quarterly pace of learning: p	0.08
persistence rate of demand: μ	0.58
entry cost function	$0.405 + 0.52 * f(0, \theta_u)$

Table 6: Base-line Parameterization of the Model

and good firms exit at an older age than unsure firms.

To conclude, Table 5 and Figures 11 and 12 suggest that my solution using X to approximate the aggregate state closely replicates optimal firm behavior at the equilibrium.[44] Therefore, I use the solution based on $\Omega = \{X\}$ to generate all the series in the subsequent analysis.

3.4.2 Calibration

Table 6 presents the assigned parameter values. Some of the parameter values are pre-chosen. The most significant in this group are the relative productivity of good and bad firms. I follow Davis and Haltiwanger (1999), who assume a ratio of high-to-low productivity of 2.4 for total factor productivity and 3.5 for labor productivity based on the between-plant productivity differentials reported by Bartelsman and Doms

[44] These results were robust when I experimented with different parameterizations of the model. Although they suggest that my approximation is good, one could say that these are self-fulfilling equilibria: because everyone perceives a simple law of motion, they behave correspondingly so that the aggregate states turn out as predicted. However, it has been difficult to prove theoretically the existence of such self-fulfilling equilibria in my model.

(1997). Since labor is the only input in my model, I normalize productivity of bad firms as 1 and set productivity of good firms as 3.5. I allow a period to represent one quarter and set the quarterly discount factor $\beta = 0.99$. Next, I need to choose γ, the quarterly pace of technological progress. In a model with only creative destruction, Caballero and Hammour (1994) choose the quarterly technological growth rate as 0.007 by attributing all output growth of US manufacturing from 1972 (II) to 1983 (IV) to technical progress. To make comparison with their results convenient in the coming subsections, I also choose $\gamma = 0.007$. Caballero and Hammour (1994) assume a linear entry cost function $c_0 + c_1\ f(0, \theta_u)$ with $f(0, \theta_u)$ denoting the size of entry, which is also applied in my calibration exercises.

The remaining undetermined parameters are: p, the pace of learning; φ, the probability of being a good firm; D_h and D_l, the demand levels; μ, the probability with which demand persists; and c_0 and c_1, the entry cost parameters. The values of these parameters are chosen so that the job destruction series in the calibrated model matches properties of the historical series from the U.S. manufacturing sector. Their values are calibrated in the following manner.

First, I match the long-run behavior of job destruction. My numerical simulations suggest that the dynamic system eventually settles down with constant entry and exit along any sample path where the demand level is unchanging. The industry structures at these stable points are similar to those at the steady states, which allows me to use steady state conditions for approximation.[45] I let $\overline{a_g}$ and $\overline{a_u}$ represent the maximum ages of good firms and unsure firms at the high-demand steady state and $\overline{a_g}'$ and

[45]However, a stable point is different from a steady state. In a steady state, firms perceive demand as constant, while in a stable point, firms perceive demand to persist with probability μ, and to change with probability $1 - \mu$.

$\overline{a_u}'$ represent the exit ages at the low-demand steady state. The steady-state job destruction rate, denoted jd^{ss}, is given by (3.9).

Second, I match the peak in job destruction that occurs at the onset of a recession. My model suggests that the jump in the job destruction rate at the beginning of a recession comes from the shift of exit margins to younger ages. I assume that when demand drops, the exit margins shift from $\overline{a_g}$ and $\overline{a_u}$ to $\overline{a_g}'$ and $\overline{a_u}'$ immediately, so that the job destruction rate at the beginning of a recession, denoted as $jd_{\max}$, is approximately:[46]

$$jd_{\max} = \frac{\begin{array}{c} \varphi\left[1-(1-p)^{\overline{a_u}+1}\right](\overline{a_g}-\overline{a_g}')+ \\ 2\cdot\ \left[\frac{1}{p}+\varphi-1-\frac{1}{p}(1-p)^{\overline{a_u}-\overline{a_u}'}\right](1-p)^{\overline{a_u}'+1}+ \\ (1-\varphi) \end{array}}{\begin{array}{c} \varphi\left(\overline{a_u}+\overline{a_u}'\right)+\frac{(1-\varphi)}{p}\left[2-(1-p)^{\overline{a_u}+1}-(1-p)^{\overline{a_u}'+1}\right]+ \\ \varphi\left[1-(1-p)^{\overline{a_u}+1}\right](\overline{a_g}-\overline{a_u})+\varphi\left[1-(1-p)^{\overline{a_u}'+1}\right](\overline{a_g}'-\overline{a_u}') \end{array}} \tag{3.12}$$

Third, I match the trough in job destruction that occurs at the onset of a boom. My model suggests that when demand goes up, the exit margins extend to older ages, so that for several subsequent periods job destruction comes only from the learning margin, implying a trough in the job destruction rate. The job destruction rate at this moment, denoted as $jd_{\min}$, is approximately:

$$jd_{\min} = \frac{(1-\varphi)\left[1-(1-p)^{\overline{a_u}'+1}\right]}{\overline{a_u}'\cdot\varphi+[\frac{1-\varphi}{p}+(\overline{a_g}'-\overline{a_u}')\cdot\varphi]\cdot[1-(1-p)^{\overline{a_u}'+1}]} \tag{3.13}$$

[46] As I have noted earlier, the calibration exercises suggest that when a negative aggregate demand shock strikes, the exit margins shift more than $\overline{a_g}'$ and $\overline{a_u}'$. The bigger shift implies a bigger jump in job destruction, This is why I require $neg_{\max}$ to lie below 11.60%. I experiment with different demand levels to find those that generate the closest fit.

Descriptive Statistics	Mean	Min.	Max.	Std.
Value	5.6%	2.96%	11.60%	1.66%

Table 7: Descriptive Statistics of Quarterly Job Destruction in U.S. Manufacturing (1972:2-1993:4), constructed by Davis and Haltiwanger.

Now I turn to data for conditions on jd^{ss}, $jd_{\max}$, and $jd_{\min}$. Table 7 lists descriptive statistics for the job destruction series of the U.S. manufacturing sector from 1972:2 to 1993:4 compiled by Davis and Haltiwanger. This data places three restrictions on the values of p, φ, $\overline{a_g}$, $\overline{a_u}$, $\overline{a_g}'$ and $\overline{a_u}'$. First, the implied jd^{ss} with either $(\overline{a_g}, \overline{a_u})$ or $(\overline{a_g}', \overline{a_u}')$ must be around 5.6%.[47] Second, the implied $jd_{\max}$ must not exceed 11.6%. Third, the implied $jd_{\min}$ must be above 3%. Additionally, $(\overline{a_g}, \overline{a_u})$ and $(\overline{a_g}', \overline{a_u}')$ must satisfy (3.8), the gap between the exit ages of good and unsure firms suggested by the steady state. There are six equations in total to pin down the values of these six parameters. Using a search algorithm, I find that these conditions are satisfied for the following combination of parameter values: $p = 0.06$, $\varphi = 0.18$, $\overline{a_g} = 78$, $\overline{a_u} = 62$, $\overline{a_g}' = 73$, $\overline{a_u}' = 57$. By applying these $\overline{a_g}$, $\overline{a_u}$, $\overline{a_g}'$ and $\overline{a_u}'$ to the steady state industry structure, I find $D_h = 2899$ and $D_l = 2464$.

The value of μ is calibrated to match the observed standard deviation of the job destruction rate. In my model, the job destruction rate jumps above its mean when demand drops and falls below when demand rises. Thus, the frequency of demand switches between D_h and D_l determines the frequency with which the job destruction rate fluctuates between 11.6% and 3%, which in turn affects the standard deviation of the simulated job destruction series. My calibration exercises suggest $\mu = 0.58$. Finally, the entry cost parameters are adjusted to match the observed mean job creation rate of 5.19%.

[47] The job destruction rate implied by $(\overline{a_g}', \overline{a_u}')$ is slightly higher since $\overline{a_g}' < \overline{a_g}$ and $\overline{a_u}' < \overline{a_u}$.

3.4.3 Response to a Negative Demand Shock and Simulations of U.S. Manufacturing Job Flows

With all of the parameter values assigned, I approximate firms' value functions applying the computational strategy described in subsection 4.1. With the approximated value functions, the corresponding decision rules and an initial firm distribution, I can investigate the dynamics of my model's key variables along any particular path of demand realizations, and study the model's quantitative implications.

Scarring and Cleansing over the Cycle To assess the effect of a negative demand shock, I start with a random firm distribution and simulate my model with demand level equal to D_h for the first 200 quarters. Regardless of the initial firm distribution, I find that the exit age of good firms settles down to 76, the exit age of unsure firms settles down to 62, the job destruction rate converges to 5.38%, and the fraction of good firms converges to 49.8%. This suggests that my model is globally stable. Once the key variables converge, I simulate the effects of a negative demand shock that persists for the next 87 quarters.

The dynamics of the job destruction rate and the job creation rate are illustrated in Panel 1 of Figure 13, with the quarter labeled 0 denoting the onset of a recession. The job destruction rate goes up from 5.38% to 10.84% on impact. Thus, the immediate effect of a negative demand shock is to clear out some firms that would have stayed in had demand remained high. After 70 quarters, the job destruction rate converges to 5.63%, still above its original value. Hence, the conventional cleansing effect of demand on job destruction that I establish analytically in steady state carries over with probabilistic cycles.

Unlike the job destruction rate, the job creation rate drops from 4.69% to 4.32% when a recession strikes, rises gradually and converges later. This matches the finding of Davis and Haltiwanger (1992) that the job creation rate falls during recessions and co-moves negatively with the job destruction rate over the cycle.[48]

The analysis of the steady state also suggests that recessions will bring a scarring effect by shifting labor resources toward bad firms. As shown in Panel 2 of Figure 13, the fraction of labor at good firms drops from 49.8% to 48.07% when the negative demand shock strikes and converges to 47.87% after 70 quarters. This implies that the negative demand shock shifts the cross-idiosyncratic productivity firm distribution toward bad firms. Hence, the scarring effect suggested by the steady-state analysis also carries over with probabilistic business cycles.

Two remarks are in order regarding the response of the fraction of labor at good firms to a negative demand shock. First, the initial drop in l_g at the onset of a recession contradicts my argument in Section 3.2 that the scarring effect takes time to work. My calibration exercises suggest that this feature is robust and can be understood as follows. Recessions shift both exit margins to younger ages. While the shift of the exit margin for unsure firms clears out *both* bad firms *and* good firms, the shift of the exit margin for good firms clears out *only* good firms, so that in total more good firms are cleared out than bad firms initially and l_g drops at the onset of a recession. Since l_g eventually converges to a value below the initial drop, and the initial drop in l_g also stems from learning, this result does not hurt my argument that in a model with learning, recessions create a scarring effect by shifting resources toward bad firms.

[48] Davis and Haltiwanger (1999) report a correlation coefficient of -0.17 of job destruction and job creation for the U.S. Manufacturing from 1947:1-1993:4.

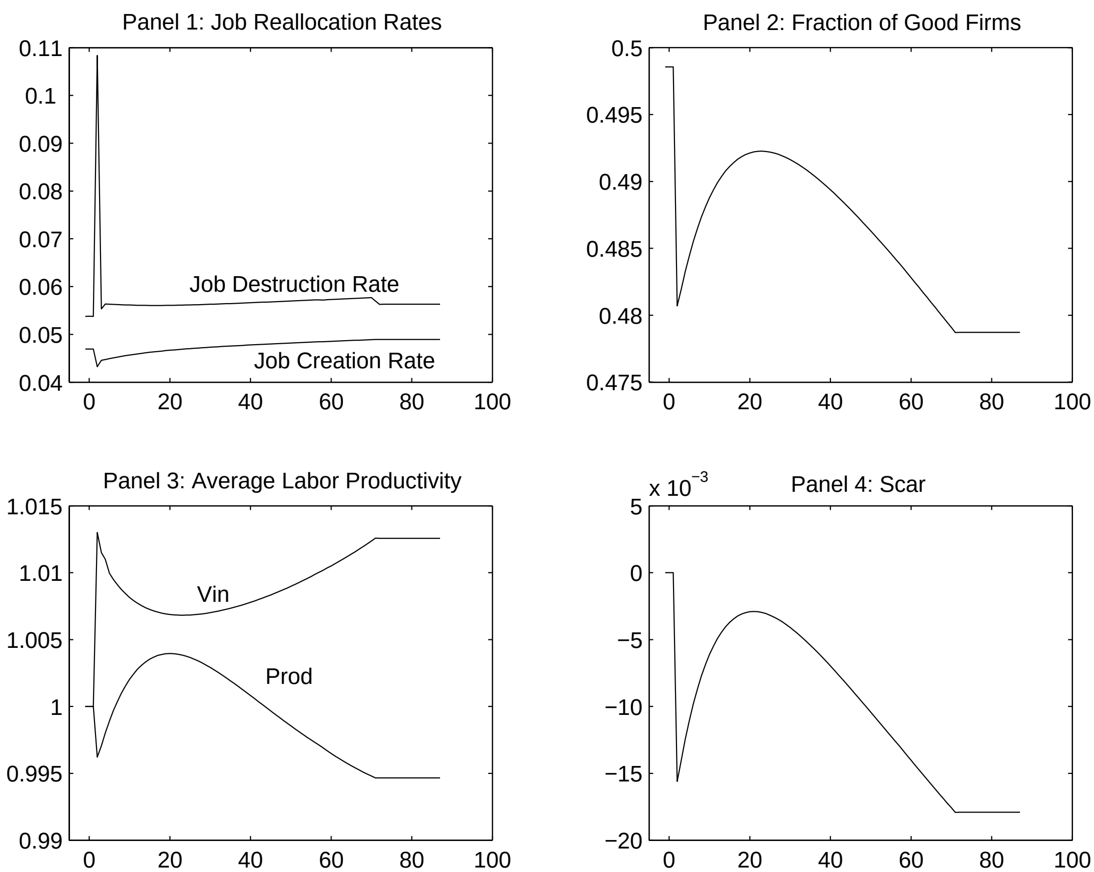


Figure 13: Response to a Negative Demand Shock: vin is the detrended average labor productivity driven only by the cleansing effect, $prod$ is the detrended average labor productivity driven by both the cleansing effect and the scarring effect. $Scar = prod - vin$. The horizontal axis denotes quarters, with the quarter labeled 0 denoting the onset of a recession.

Second, the response of l_g shown in Panel 2 is hump-shaped: it drops initially, increases gradually, then declines again. This feature is mainly due to the response of the exit margins over the cycle. When a recession first strikes, the exit margins overshift to the left, and shift back gradually as the recession persists. As the exit margin for unsure firms shifts back, more good firms are allowed to reach their potential; meanwhile, as the exit margin for good firms shifts back, no old good firms exit for several quarters. Hence, l_g increases after the initial drop. The exit margins reach their stable points after about 20 quarters. From then on, l_g starts to fall, with old good firms gradually being cleared out but not enough new good firms being realized. Another part of this hump-shaped response comes from the entry margin. Because they have had no time to learn, newly entered cohorts have the least efficient cross-idiosyncratic productivity firm distribution in the industry, so that entry tends to drive down l_g. When entry falls in a recession, the negative impact of entry on l_g is also reduced, which contributes to part of the increase in l_g after the initial drop.

To summarize, despite some transitory dynamics, Panel 1 and Panel 2 of Figure 13 suggest that both the conventional cleansing effect established in Proposition 3.2, and the scarring effect established in Proposition 3.3, carry over with probabilistic business cycles.

Implications for Productivity Next, I turn to the quantitative implications of the model for the cyclical behavior of average labor productivity. With one worker per firm setup and firm-level productivity given by $\frac{A \cdot \theta}{(1+\gamma)^a}$, average labor productivity is affected by A, the level of the leading technology, and the firm distribution across a and θ. While technological progress drives A, and thus average labor productivity, to grow at a trend rate γ (the technological pace), demand shocks add fluctuations around this trend by affecting the labor distribution across a and θ.

To analyze the fluctuations of average labor productivity over the cycle, I define *de-trended average labor productivity* as the average of $\frac{\theta}{(1+\gamma)^a}$ over heterogeneous firms. In evaluating this measure, recall that there are two competing effects. On the one hand, the cleansing effect drives down the average a by lowering the cut-off ages for each idiosyncratic productivity, causing average labor productivity to rise. On the other hand, the scarring effect drives down the average θ by shifting resources away from good firms, causing average labor productivity to fall. To separate the two effects, I generate two indexes for average labor productivity. The first index is the average of $\frac{\theta}{(1+\gamma)^a}$ across all firms in operation, defined as the following:

$$prod = \frac{\sum_f \left(\frac{\theta^e}{(1+\gamma)^a}\right) \cdot f\left(\theta^e, a\right)}{\sum_f f\left(\theta^e, a\right)}.$$

This measure is affected by both cleansing and scarring effects. The other index is the average of $\frac{1}{(1+\gamma)^a}$ across all existing firms, defined as:

$$vin = \frac{\sum_f \left(\frac{1}{(1+\gamma)^a}\right) \cdot f\left(\theta^e, a\right)}{\sum_f f\left(\theta^e, a\right)}.$$

This measure is affected only by the cleansing effect. To compare the relative magnitude of these two effects, *their initial levels are both normalized as* 1. Since only the cleansing effect drives the dynamics of vin but both cleansing and scarring effects drive the dynamics of $prod$, the gap between vin and $prod$ reflects the magnitude of the scarring effect. A scarring index measures this gap. It is defined as:

$$scar = prod - vin.$$

Panel 3 in Figure 13 traces the evolution of vin and $prod$ in response to a negative demand shock. As the negative demand shock strikes, the cleansing effect *alone* raises the average labor productivity to 1.013 while the scarring effect brings the average labor productivity down to 0.9974. After 70 quarters, $prod$ converges to 0.9947 while vin converges to 1.0126. The dynamics of the scarring index in response to a negative demand shock is plotted in Panel 4 of Figure 13. The scarring index remains negative following a negative demand shock and eventually converges to -0.0179. This matches the predictions of my model that the scarring effect plays against the conventional cleansing effect during recessions by shifting resources away from good firms, driving down the average labor productivity.

3.4.4 Simulation of U.S. Manufacturing Job Flows

To gauge whether the scarring effect is likely to be relevant at business cycle frequencies, I simulate my model's response to random demand realizations generated by the model's Markov chain. I perform 1000 simulations of 87 quarters each. Results are presented in Table 8. The reported statistics are means (standard deviations) based on 1000 simulated samples. Sample statistics for U.S. Manufacturing data for the 87 quarters from 1972 (II) to1993(IV) are included for comparison. In the table, jd and jc represent the job destruction and job creation rate; $prod$ and q represent de-trended average labor productivity and de-trended output.

Table 8 suggests that my calibrated model can replicate the observed patterns of job flows; moreover, the positive correlation coefficient of 0.1675 between $prod$ and q implies that my model generates procyclical average labor productivity for the U.S.

	simulation statistics	**data**
jd_{mean}	5.29%(0.0100%)	5.6%
jd_{std}	1.65%(0.3100%)	1.66%
jc_{mean}	4.72%(0.0581%)	5.19%
jc_{std}	0.72%(0.0595%)	0.95%
$corr(prod, q)$	0.1675(0.7504)	0.5537*

Table 8: Means (std errors) of 1000 Simulated 87-quarter Samples: *jd* is the job destruction rate, *jc* is the job creation rate, *prod* is detrended average labor productivity, q is detrended aggregate output. Data comes from the U.S. Manufacturing job flow series for 1972:2-1993:4, compiled by Davis and Haltiwanger. *Detrended average labor productivity is calculated as output per production worker, with output measured by industrial production index. The quarterly series of industrial production index of U.S. manufacturing sector for 1972:2-1993:4 comes from the Federal Reserve and the series of total production workers comes from the Bureau of Labor Statistics.

manufacturing sector in the relevant period. Put differently, under my benchmark calibration the scarring effect on cyclical productivity dominates the cleansing effect.

3.5 Sensitivity Analysis of the Dominance of Scarring over Cleansing

In the baseline parameterization of subsection 4.2, I followed Caballero and Hammour (1994) in setting the quarterly technological pace γ equal to 0.007. The value was estimated by attributing *all* output growth of the U.S. manufacturing sector to technological progress, which may exaggerate the technological pace in the relevant period. An alternative estimate of γ, has been provided by Basu, Fernald and Shapiro (2001), who estimate TFP growth for different industries in the U.S. from 1965 to 1996 after controlling for employment growth, factor utilization, capital adjustment costs, quality of inputs and deviations from constant returns and perfect competition. They estimate a quarterly technological pace of 0.0037 for durable manufacturing, a pace

Calibration Results	$\gamma = \mathbf{0.003}$	$\gamma = \mathbf{0.007}$
calibrated p	0.0830	0.0800
calibrated φ	0.1200	0.1420
Response to a Negative Demand Shock		
vin (when a recession strikes)	1.0052	1.0130
vin (70 quarters after a recession strikes)	1.0029	1.0126
prod (when a recession strikes)	0.9866	0.9974
prod (70 quarters after a recession strikes)	0.9820	0.9947
scar (when a recession strikes)	−0.0186	−0.0156
scar (70 quarters after a recession strikes)	−0.0209	−0.0179

Table 9: Sensitivity Analysis to a Slower Technological Pace (I): *prod* is detrended average labor productivity, driven by both the cleansing and the scarring effects, *vin* is the component of detrended average labor productivity driven only by the cleansing effect, *scar* = *prod* - *vin*. Other parameter values are as shown in Table 2.

of 0.0027 for non-durable manufacturing and an even slower pace for other sectors.

How would a slow pace of technological progress affect the magnitudes of the scarring and cleansing effects? To address this question, I re-calibrate my model assuming $\gamma = 0.003$, matching the same moments of job creation and destruction as before, and simulate responses to a negative demand shock. The results are presented in Table 9 together with results from the baseline parameterization.

The calibration results in Table 9 suggest that the model with $\gamma = 0.003$ needs a faster learning pace ($p = 0.083$ compared to 0.08) and a smaller prior probability of firms' being good ($\varphi = 0.120$ compared to 0.142) to match the observed moments of

	simulation statistics with $\gamma = 0.003$	**simulation statistics with $\gamma = 0.007$**	**data**
jd_{mean}	5.73%(0.0799%)	5.29%(0.0100%)	5.6%
jd_{std}	1.42%(0.2800%)	1.65%(0.3100%)	1.66%
jc_{mean}	5.14%(0.0565%)	4.72%(0.0581%)	5.19%
jc_{std}	0.34%(0.0059%)	0.37%(0.0535%)	0.95%
$corr(prod, q)$	0.4819(0.5212)	0.1675(0.7504)	0.5537

Table 10: Sensitivity to A Slower Technological Pace (II): Means (std errors) of 1000 Simulated 87-quarter Samples. Definitions, measures and data sources are the same as Table 4.

job flows.[49] The simulated responses suggest that slower technological progress magnifies the scarring effect, weakens the cleansing effect, and magnifies the procyclical behavior of productivity.

This result can be explained as follows. First, slower technological progress implies that the force of creative destruction is weak. A lower γ weakens the technical disadvantage of old firms and allows both good firms and unsure firms to live longer, so that less job destruction occurs at the exit margins. A lower γ also implies a smaller cleansing effect on average labor productivity. A recession clears out marginal firms by shifting the exit margins toward younger ages. The size of the shift is pinned down in my calibration exercises by matching $jd_{\max} \approx 11.6\%$. Given the shift of exit margins, a slower technological pace *shrinks* the productivity difference between the vintages that have been killed and the ones that have survived, so that the impact of

[49] Consider (9), the expression of jd^{ss}, for intuition. My calibration exercises look for parameter values that satisfy three moment conditions on job flows, one of which is that $jd^{ss} \approx 5.6\%$. Proposition 3 establishes that jd^{ss} decreases with the exit ages ($\overline{a_g}$ and $\overline{a_u}$). It can be further shown that it increases in p but decreases in φ. A slower technological pace weakens the technical disadvantage of old firms and extends their life span so that both $\overline{a_g}$ and $\overline{a_u}$ tend to increase. Hence, the job destruction rate would decrease if p and φ remain the same. A faster learning pace and a lower prior probability of being good are thus needed to match the observed mean job destruction. Thus, the paramerization of my model with $\gamma = 0.003$ suggests that more job destruction comes from learning rather than creative destruction.

the cleansing effect on average labor productivity declines.

Second, when I assume a lower γ, I must also assume a higher p and a lower φ to match the moments of job destruction. This re-calibration implies a larger role for learning in job destruction: firms not only learn faster, but are more likely to learn that they are bad. This also gives a larger scarring effect on average labor productivity: a faster learning pace implies a higher *opportunity cost* of not allowing unsure firms to survive; a smaller prior probability of being good suggests that learning has a greater *marginal* impact on cross-idiosyncratic productivity efficiency.

Table 10 reports the simulation statistics of 1000 simulated 87-quarter samples when $\gamma = 0.003$. Results when $\gamma = 0.007$ and sample statistics from data are included for comparison. My model with $\gamma = 0.003$ generates a correlation coefficient of 0.4819 between detrended average labor productivity and detrended output. Productivity is strongly procyclical, almost as much as in the data.

3.6 Cost of Business Cycles with Heterogeneous Firms

This sub-section explores possible welfare cost of business cycles with the scarring effect's presence. Suppose my modeled industry's output is consumed by a representative consumer, whose utility depends positively on the level of consumption. Then lower output implies lower consumption and consequently lower welfare. Since the scarring effect drives down average labor productivity during recessions, it can lead to lower equilibrium output, and hence imply a welfare cost.

To proceed, I compare the output series of two industries, a cyclical industry whose demand follows a Markov chain and a steady-state industry with time-invariant demand. I let the steady-state industry's demand equal the cyclical industry's average

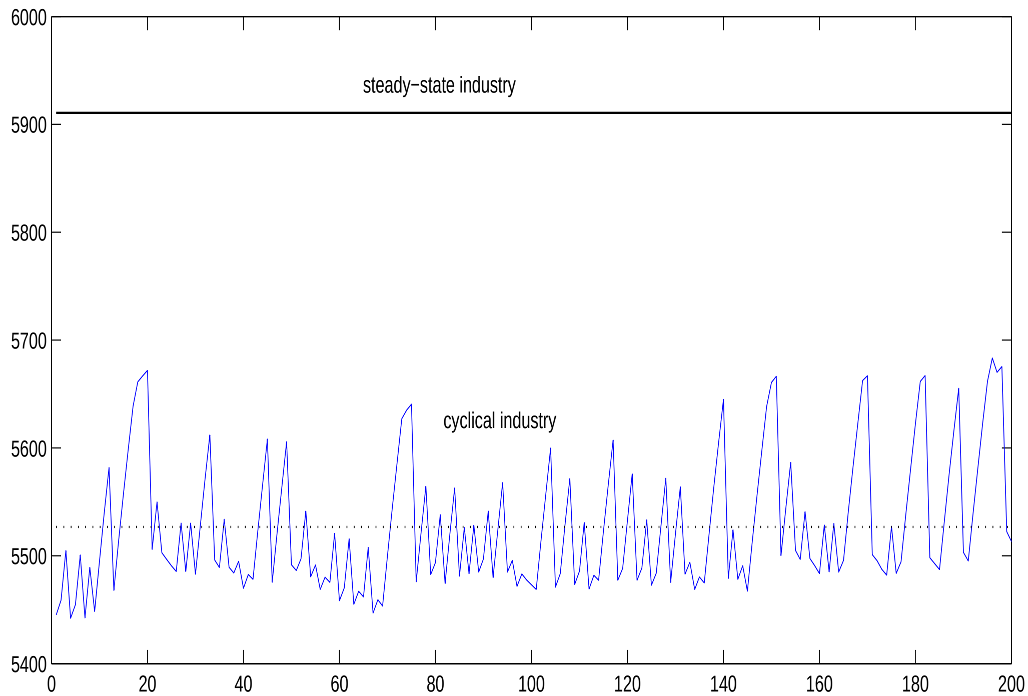


Figure 14: Time Series of Detrended Output of a Cyclical industry and a Steady-state Industry. Dashed line represents the average of cyclical output. This figure is generated using the base-line calibration, with the steady-state industry's demand level assumed equal to the cyclical industry's average demand level.

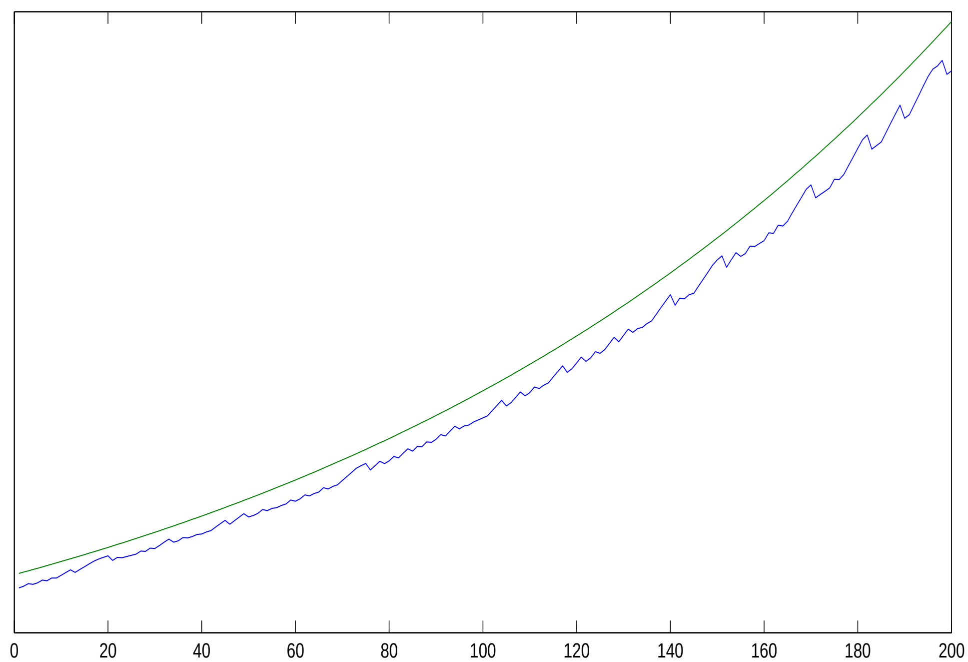


Figure 15: Time Series of Output with Trend of a Cyclical Industry and a Steady-state Industry. This figure is generated with the base-line calibration, assuming the steady-state industry's demand equal to the cyclical industry's average demand.

demand. The results are shown in Figure 14 and Figure 15. Both figures are generated with baseline calibrations. The cyclical industry's demand switches between 2899 and 2464 with probability 0.42 $(1-0.58)$. This implies an average demand of 2657, which is applied as the steady-state industry's demand.

Figure 14 presents the time series of the two industries' *de-trended* output. The cyclical industry's de-trended output fluctuates around a mean of 5515.8, below the steady-state industry's de-trended output of 5910.6. Figure 15 shows the two industries' output series with technological progress added. Both industries' outputs grow. But only the cyclical industry's output fluctuates around the growth trend. Moreover, the cyclical output series stay strictly below the steady-state output series.

Figure 14 and Figure 15 suggest that, in my model's framework, more output would be produced if business cycles could be eliminated. Hence, business cycles may possibly bring a welfare cost of a representative consumer who consumes the industry's output.

The discussion of welfare cost of business cycles traces back to Lucas (1987), who put forth an argument that the welfare gains from reducing the volatility of aggregate consumption is negligible. Subsequent work that revisited Lucas calculation continued to find only small benefits from reducing the consumption volatility, reinforcing the perception that business cycles do not matter. However, with a heterogeneous-firm setup, my model argues from the supply side that the business cycles reduces the average output by affecting production efficiency. In Lucas' argument, eliminating business cycles only eliminates the consumption volatility, but does not affect the mean of consumption. My model suggests that, eliminating business cycles may raise up the mean by providing more output.[50] Although my analysis here is very

[50] Since consumption grows over time, this mean refers to the mean of detrended consumption.

preliminary,[51] it does point out an interesting research direction.

3.7 Conclusion

How do recessions affect resource allocation? My theory suggests learning has important consequences for this question. I posit that in addition to the cleansing effect proposed by previous authors, recessions create a scarring effect by interrupting the learning process. Recessions kill off potentially good firms, shift resources toward bad firms and exacerbate the allocative inefficiency in an industry. The empirical relevance of the scarring effect is examined in Section 4. Using data on U.S. manufacturing job flows, I find that the scarring effect dominates the cleansing effect in the U.S. manufacturing sector from 1972 to 1993, and can account for the observed degree of procyclical productivity.

The scarring effect stems from learning. Recessions bring a scarring effect by limiting the learning scope. Figure 3 of the paper provides intuition. Recessions force firms to exit at earlier ages. The shortened firm life allows less learning time, so that fewer truly good firms get to realize their potential and the shaded area in Figure 3 would disappear. The decrease in the fraction of labor at good firms implies a less efficient allocation of labor during recessions.

My theory highlights a firm' age as an indicator for its number of learning opportunities. The existing empirical literature documents that firm age has important

A higher mean of detrended cosumption (output) is shown in Figure 15: although both growing over time, the steady-state series with a higher detrended mean stays strictly above the other series. Another interesting exploration of the welfare cost of business cycles is Barlevy (2003), who posits that eliminating cycles may give rise to a higher growth rate of consumption.

[51] A more careful exploration of this question should study cyclical labor supply and cyclical equilibrium labor input. Seperating the cleansing effect from the scarring effect is also important.

explanatory power for micro-level job flow patterns.[52] My model predicts that the mean and the dispersion of firm age both decline during recessions, while the productivity dispersion within an age cohort goes up on average. These are testable hypotheses with detailed data on the age distribution of firms over the cycle.

The empirical relevance of the scarring effect remains to be explored in a wider framework. My calibration exercises have focused on the U.S. manufacturing sector, where job destruction is more responsive to business cycles than job creation. However, Foote (1997) documents that in services, fire, transportation and communications, retail trade, and wholesale trade, job creation is more volatile than job destruction. Would relatively more responsive job creation hurt the dominance of the scarring effect? It could, since recessions leave "scars" by killing off potentially good firms on the destruction side. It may not, because a larger decline in job creation also introduces fewer potentially good firms on the creation side. Whether "scarring" dominates "cleansing" in sectors other than manufacturing remains an interesting question.

[52] See Caves (1998) for an extensive review of recent findings on firm turnover and industrial dynamics.

4 Appendix

Proof of Proposition 2.1 and 3.1(three steps):

Step1: to prove that $\frac{V(\theta^e,a;F,D)}{\partial a} < 0$:

Proof. Compare two firms with same belief θ^e, but different ages $a_1 > a_2$. To prove $\frac{V(\theta^e,a;F,D)}{\partial a} < 0$, I need to show that

$$V\left(\theta^e, a_1; F, D\right) < V\left(\theta^e, a_2; F, D\right).$$

Suppose that the aggregate state is (F, D) at the beginning of period t_0. I assume there are n different possible paths of demand realizations from t_0 onward, each with probability p^i, where $i = 1, ..., n$. I also assume that under the i'th path of demand realizations, the firm with a_1 expects itself to exit at the end of period $t_1^i \geq t_0$ and the firm with a_2 expects itself to exit at the end of period $t_2^i \geq t_0$, then:

$$V\left(\theta^e, a_1; F, D\right) = \sum_{i=1}^{n} \sum_{t=t_0}^{t_1^i} \left\{\beta^{t-t_0} E\left[\pi_t^i\left(\theta^e, a_1 + t - t_0\right)|F, D\right]\right\} \cdot p^i,$$

and

$$V\left(\theta^e, a_2; F, D\right) = \sum_{i=1}^{n} \sum_{t=t_0}^{t_2^i} \left\{\beta^{t-t_0} E\left[\pi_t^i\left(\theta^e, a_2 + t - t_0\right)|F, D\right]\right\} \cdot p^i,$$

where $\pi_t^i\left(\theta^e, a_1 + t - t_0\right)$ is the expected profit (of a firm with current age a_1 and current belief θ^e) at period $t \geq t_0$ under demand path i. Firms have rational expectations and expect a price sequence $\left\{P_t^i\left(F, D\right)\right\}_{t \geq t_0}$ conditional on the realization of path i. Since price is competitive and firms are price takers, I must have:

$$V\left(\theta^{e}, a_{1}; F, D\right)=\sum_{i=1}^{n} \sum_{t=t_{0}}^{t_{1}^{i}}\left\{\beta^{t-t_{0}}\left[A\left(t_{0}-a_{1}\right) \theta^{e} P_{t}^{i}(F, D)-1\right]\right\} \cdot p^{i}$$

and

$$V\left(\theta^{e}, a_{2}; F, D\right)=\sum_{i=1}^{n} \sum_{t=t_{0}}^{t_{2}^{i}}\left\{\beta^{t-t_{0}} \cdot\left[A\left(t_{0}-a_{2}\right) \theta^{e} P_{t}^{i}(F, D)-1\right]\right\} \cdot p^{i}.$$

There are three possibilities for any i.

Possibility 1, if $t_{1}^{i}=t_{2}^{i}=t^{i}$:

since $A\left(t_{0}-a_{1}\right)<A\left(t_{0}-a_{2}\right)$,

$$\left(t_{0}-a_{1}\right) \theta^{e} P_{t}^{i}(F, D)-1<A\left(t_{0}-a_{2}\right) \theta^{e} P_{t}^{i}(F, D)-1$$

holds for any t. Hence,

$$\begin{aligned} & \sum_{t=t_{0}}^{t^{i}}\left\{\beta^{t-t_{0}}\left[A\left(t_{0}-a_{1}\right) \theta^{e} P_{t}^{i}(F, D)-1\right]\right\} \\ < & \sum_{t=t_{0}}^{t^{i}}\left\{\beta^{t-t_{0}} \cdot\left[A\left(t_{0}-a_{2}\right) \theta^{e} P_{t}^{i}(F, D)-1\right]\right\} \end{aligned}$$

Possibility 2, if $t_{1}^{i}<t_{2}^{i}$:

then it must be true that,

$$\begin{aligned}
&\sum_{t=t_0}^{t_2^i}\left\{\beta^{t-t_0}\cdot\left[A\left(t_0-a_2\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\}\\
=&\sum_{t=t_0}^{t_1^i}\left\{\beta^{t-t_0}\cdot\left[A\left(t_0-a_2\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\}+\\
&\sum_{t=t_1^i+1}^{t_2^i}\left\{\beta^{t-t_0}\cdot\left[A\left(t_0-a_2\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\},
\end{aligned}$$

and hence,

$$\begin{aligned}
&\sum_{t=t_0}^{t_1^i}\left\{\beta^{t-t_0}\left[A\left(t_0-a_1\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\}\\
<&\sum_{t=t_0}^{t_2^i}\left\{\beta^{t-t_0}\cdot\left[A\left(t_0-a_2\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\},
\end{aligned}$$

Possibility 3, if $t_1^i > t_2^i$:

when it comes to period t_2^i under path i, the firm aged $a_1+t_2^i-t_0$ chooses to stay and the firm aged $a_2+t_2^i-t_0$ decides to leave. Based on the exit condition, it must be true that,

$$V\left(\theta^e, a_1+t_2^i-t_0; F', D'\right) > 0 and V\left(\theta^e, a_2+t_2^i-t_0; F', D'\right) < 0.$$

The firm aged $a_1+t_2^i-t_0$ chooses to stay to capture the potential profit

$$\sum_{t=t_2^i+1}^{t_1^i}\left\{\beta^{t-t_2^i}\cdot\left[A\left(t_0-a_1\right)\theta^e P_t^i\left(F,D\right)-1\right]\right\}$$

and he expects those future profits can cover any possible cost if demand path does not goes as expected. Since

$$\sum_{t=t_2^i+1}^{t_1^i} \left\{ \beta^{t-t_2^i} \cdot \left[A\left(t_0 - a_1\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\}$$
$$< \sum_{t=t_2^i+1}^{t_1^i} \left\{ \beta^{t-t_2^i} \cdot \left[A\left(t_0 - a_2\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\},$$

the firm aged $a_2 + t_2^i - t_0$ should have expected even higher potential profits in the future which is worth waiting for. Hence, it must not choose to leave at period t_2^i. Therefore, $t_1^i > t_2^i$ cannot be true.

1), 2) and 3) help me conclude that:

$$\sum_{t=t_0}^{t_1^i} \left\{ \beta^{t-t_0} \left[A\left(t_0 - a_1\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\}$$
$$< \sum_{t=t_0}^{t_2^i} \left\{ \beta^{t-t_0} \cdot \left[A\left(t_0 - a_2\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\}$$

holds for any i. Then it must be true that,

$$\sum_{i=1}^{n} \sum_{t=t_0}^{t_1^i} \left\{ \beta^{t-t_0} \left[A\left(t_0 - a_1\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\} p^i$$
$$< \sum_{i=1}^{n} \sum_{t=t_0}^{t_2^i} \left\{ \beta^{t-t_0} \cdot \left[A\left(t_0 - a_2\right) \theta^e P_t^i \left(F, D\right) - 1 \right] \right\} p^i$$

or

$$V\left(\theta^e, a_1; F, D\right) < V\left(\theta^e, a_2; F, D\right).$$

■

Step 2: to prove $\frac{V(\theta^e,a;F,D)}{\partial\theta^e} > 0$.

Proof. It is similar to the proof of $\frac{V(\theta^e,a;F,D)}{\partial a} > 0$. ■

Step 3: to prove the existence of cut-off age $\overline{a}(\theta^e;F,D)$ and $\overline{a}(\theta^{e\prime};F,D) \geq \overline{a}(\theta^e;F,D)$, for $\theta^{e\prime} > \theta^e$.

Proof. The existence of $\overline{a}(\theta^e;F,D)$ is straightforward. Holding θ^e constant, $V(\theta^e,a;F,D)$ is monotonically decreasing in a, then there must be $\overline{a}(\theta^e;F,D)$ such that

$$V(\theta^e,\overline{a}(\theta^e;F,D);F,D) > 0$$

but

$$V(\theta^e,\overline{a}(\theta^e;F,D)+1;F,D) \leq 0.$$

And since $\frac{V(\theta^e,a;F,D)}{\partial\theta^e} > 0$, I have:

$$V\left(\theta^{e^\prime},\overline{a}(\theta^e;F,D);F,D\right) > V(\theta^e,\overline{a}(\theta^e;F,D);F,D) = 0 \text{ holds for any } \theta^{e\prime} > \theta^e.$$

Therefore, it must be true that $\overline{a}(\theta^{e\prime};F,D) \geq \overline{a}(\theta^e;F,D)$. ■

PROOF OF PROPOSITION 3.2 (three steps):

Proof. *Step 1: to show that a steady state features time-invariant P_tA_t, such that $P_tA_t = PA$, $\forall\ t$, where P_t represents the equilibrium price and A_t represents the leading technology in period t.*

The condition of competitive pricing tells that:

$$D_t = P_t \cdot Q_t.$$

Q_t is the aggregate output over heterogeneous firms.

$$Q_t = \sum_a \sum_{\theta^e} A_t \theta^e f_t(\theta^e, a)(1+\gamma)^{-a}.$$

so that:

$$D_t = P_t A_t \cdot \sum_a \sum_{\theta^e} \theta^e f_t(\theta^e, a)(1+\gamma)^{-a}. \tag{1}$$

By definition, a steady state features constant level of demand, $D_t = D$ ($\forall\ t$). and time-invariant firm distribution. Let $f(\theta^e, a)$ denote the number of firms with (θ^e, a) and $\overline{a_g}$, $\overline{a_u}$ denote the maximum ages for good firms and unsure firms in operation, respectively. The above equation can be rewritten as:

$$D = P_t A_t \cdot \left\{ \sum_{a=0}^{\overline{a_u}} \left[\theta_u f(\theta_u, a)(1+\gamma)^{-a}\right] + \sum_{a=1}^{\overline{a_g}} \left[\theta_g f(\theta_g, a)(1+\gamma)^{-a}\right] \right\}$$

so that

$$P_t A_t = \frac{D}{\left\{ \sum_{a=0}^{\overline{a_u}} \left[\theta_u f(\theta_u, a)(1+\gamma)^{-a}\right] + \sum_{a=1}^{\overline{a_g}} \left[\theta_g f(\theta_g, a)(1+\gamma)^{-a}\right] \right\}}.$$

Hence, $P_t A_t$ must be time-invariant. I let $P_t A_t = PA$.

Step 2: solve for $\overline{a}_g - \overline{a}_u$ by firms' exit conditions.

At a steady state, the aggregate state $\{D, F\}$ is perceived to be time-invariant. Thus, good firms know they will live until $\overline{a_g}$, and unsure firms know they will live until $\overline{a_u}$. The time-invariant decision rules at the steady state imply time-invariant value functions. Let $V(\theta^e, a)$ represent the steady-state expected value of staying of a firm with belief θ^e and age a.

Since $\overline{a_g}$ denote the maximum age of good firms in operation, and $V(\theta_g, a)$ decreases in a monotonically, the condition of firm rationality suggests it must be true for $\overline{a_g}$ that:

$$\begin{aligned} V(\theta_g, \overline{a_g}) &= 0 \\ \theta_g PA (1+\gamma)^{-\overline{a_g}} - 1 &= 0 \end{aligned}$$

so that

$$PA = \frac{(1+\gamma)^{\overline{a_g}}}{\theta_g}. \tag{2}$$

Similarly, exit condition for unsure firms suggest:

$$\begin{aligned} V(\theta_u, \overline{a_u}) &= 0 \\ \theta_u PA (1+\gamma)^{-\overline{a_u}} - 1 + \beta p \varphi V(\theta_g, \overline{a_u}+1) &= 0 \\ \theta_u PA (1+\gamma)^{-\overline{a_u}} - 1 + \beta p \varphi \sum_{a=\overline{a_u}+1}^{\overline{a_g}} \beta^{a-\overline{a_u}-1} \left[\theta_g PA (1+\gamma)^{-a} - 1\right] &= 0 \end{aligned}$$

With (15) plugged in, I have (8):

$$\left(\frac{\theta_u}{\theta_g} + \frac{p\varphi\beta}{1+\gamma-\beta}\right)(1+\gamma)^{\bar{a}_g-\bar{a}_u} = 1 + \frac{p\varphi\beta}{1-\beta} - \frac{p\varphi\beta\gamma}{(1-\beta)(1+\gamma-\beta)}\beta^{\bar{a}_g-\bar{a}_u} \tag{8}$$

which can be re-written as:

$$F(\bar{a}_g - \bar{a}_u) = G(\bar{a}_g - \bar{a}_u)$$

Proposition 1 suggests that $\bar{a}_g - \bar{a}_u \geq 0$. To establish the existence of $\bar{a}_g - \bar{a}_u \geq 0$ that satisfies the above equation, I need to show that F and G cross each other at a positive value of $\overline{a}_g - \overline{a}_u$.

$$G' = -\frac{p\varphi\beta\gamma}{(1-\beta)(1+\gamma-\beta)}\beta^{\overline{a}_g-\overline{a}_u}\ln\beta > 0, \text{ but}$$
$$G'' = -\frac{p\varphi\beta\gamma}{(1-\beta)(1+\gamma-\beta)}\beta^{\overline{a}_g-\overline{a}_u}(\ln\beta)^2 < 0$$

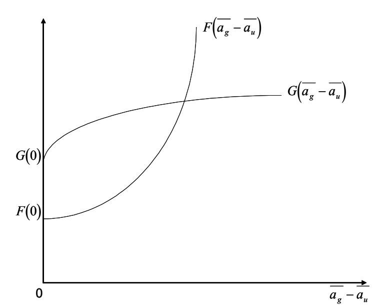
moreover,

$$F(0) = \frac{\theta_u}{\theta_g} + \frac{p\varphi\beta}{1+\gamma-\beta}, \text{ and}$$
$$G(0) = 1 + \frac{p\varphi\beta}{1+\gamma-\beta}.$$

and:

$$F(0) < G(0)$$

because $\frac{\theta_u}{\theta_g} < 1$ by definition ($\theta_u = \varphi\theta_g + (1-\varphi)\theta_b$ and $\theta_g > \theta_b$). $F(0) < G(0)$ suggests that the curve of F starts at $\overline{a}_g - \overline{a}_u = 0$ below the curve of G. $F' > 0$ and $G' > 0$ imply that both of F and G increase monotonically in $\overline{a}_g - \overline{a}_u$. $F'' > 0$ suggests that F is convex but $G'' < 0$ suggests that G is concave. Hence, F and G must cross *once* at a positive value of $\overline{a}_g - \overline{a}_u$, as shown in the following figure:



Therefore, (8) alone determines a unique value for $\overline{a}_g - \overline{a}_u$.

Step 3, solve for $f(0)$ and $\overline{a}_g$ *by combining the free entry condition and the competitive pricing condition:*

$$V(\theta_u, 0) = C(f(0))$$

where $f(0)$ represents the size of the entering cohort. With time-invariant life-cycle dynamics for each cohort shown in Figure 2, I have:

$$V(\theta_u, 0) = \sum_{a=1}^{\overline{a}_u} \beta^a \left[\frac{PA\theta_u}{(1+\gamma)^a} - 1 \right] \lambda(\theta_u, a) + \sum_{a=1}^{\overline{a}_g} \beta^a \left[\frac{PA\theta_g}{(1+\gamma)^a} - 1 \right] \lambda(\theta_g, a)$$

where $\lambda(\theta_u, a)$ denotes the probability of staying in operation at age a as an unsure firm, and $\pi(\theta_g, a)$ denotes the probability of staying in operation at age a as a good firm. All-or-nothing learning suggests that:

$$\begin{aligned} \lambda(\theta_u, a) &= (1-p)^a \text{ for } 0 \leq a \leq \overline{a}_u, \\ \lambda(\theta_g, a) &= \varphi\left[1 - (1-p)^a\right] \text{ for } 0 \leq a \leq \overline{a}_u, \\ \lambda(\theta_g, a) &= \varphi\left[1 - (1-p)^{\overline{a}_u + 1}\right] \text{ for } \overline{a}_u + 1 \leq a \leq \overline{a}_g \end{aligned}$$

Plugging $\lambda(\theta_u, a)$, $\lambda(\theta_g, a)$ and $PA = \frac{(1+\gamma)^{\overline{a}_g}}{\theta_g}$ into $V(\theta_u, 0)$, I have:

$$\frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left\{ \begin{array}{c} \sum_{a=1}^{\overline{a}_u} \beta^a \left[\begin{array}{c} (1-p)^a \left(\frac{\theta_u}{(1+\gamma)^a} - 1 \right) + \\ \varphi\left(1 - (1-p)^a\right) \left(\frac{\theta_g}{(1+\gamma)^a} - 1 \right) \end{array} \right] + \\ \varphi\left(1 - (1-p)^{\overline{a}_u + 1}\right) \sum_{a=\overline{a}_u + 1}^{\overline{a}_g} \beta^a \left(\frac{\theta_g}{(1+\gamma)^a} - 1 \right) + \\ \theta_u - 1 \end{array} \right\} = C(f(0)) \qquad (3)$$

Plugging $PA = \frac{(1+\gamma)^{\overline{a}_g}}{\theta_g}$ back into (14) and applying the steady state industry structure suggested by all-or-nothing learning and exit ages, I have:

$$f(0) \cdot \frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left[\begin{array}{c} (\theta_u - \varphi\theta_g) \sum_{a=1}^{\overline{a}_u} \left(\frac{1-p}{1+\gamma}\right)^a + \varphi\theta_g \sum_{a=1}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a + \\ \varphi\theta_g (1-p)^{\overline{a}_u+1} \sum_{a=\overline{a}_u+1}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a \end{array} \right] = D \qquad (4)$$

$\overline{a}_g - \overline{a}_u$ has been given by (8). The left-hand sides of (16) and (17) are both monotonically increasing in $\overline{a}_g$; The left-hand side and the right-hand side of (16) are both monotonically increasing in $f(0)$. Hence, with $\overline{a}_u$ replaced by $\overline{a}_g - (\overline{a}_g - \overline{a}_u)$, (16) and (17) jointly determine $\overline{a}_g$ and $f(0)$.

Therefore, for any D, there exists a steady state that can be captured by $\{f(0), \overline{a}_g, \overline{a}_u\}$.

■

PROOF OF PROPOSITION 3.3:

Proof. To prove that $\frac{d(\overline{a}_g)}{dD} \geq 0$ and $\frac{d(\overline{a}_u)}{dD} \geq 0$ at the steady state, combining (16) with (17) and replacing $\overline{a}_u$ by $\overline{a}_g - (\overline{a}_g - \overline{a}_u)$ gives the following:

$$\begin{aligned} & \frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left[\begin{array}{c} (\theta_u - \varphi\theta_g) \sum_{a=1}^{\overline{a}_u} \left(\frac{1-p}{1+\gamma}\right)^a + \varphi\theta_g \sum_{a=1}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a + \\ \varphi\theta_g (1-p)^{\overline{a}_u+1} \sum_{a=\overline{a}_u+1}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a \end{array} \right] \cdot \\ & c^{-1} \left(\frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left\{ \begin{array}{c} \sum_{a=1}^{\overline{a}_u} \beta^a \left[\begin{array}{c} (1-p)^a \left(\frac{\theta_u}{(1+\gamma)^a} - 1\right) + \\ \varphi (1-(1-p)^a) \left(\frac{\theta_g}{(1+\gamma)^a} - 1\right) \end{array} \right] + \\ \varphi \left(1 - (1-p)^{\overline{a}_u+1}\right) \sum_{a=\overline{a}_u+1}^{\overline{a}_g} \beta^a \left(\frac{\theta_g}{(1+\gamma)^a} - 1\right) + \\ \theta_u - 1 \end{array} \right\} \right) \\ = \; & D \end{aligned}$$

The left-hand is monotonically increasing in $\overline{a}_g$. Hence, $\frac{d(\overline{a}_g)}{dD} \geq 0$. With $\overline{a}_g - \overline{a}_u$ independent of D as suggested by (8), $\frac{d(\overline{a}_u)}{dD} = \frac{d(\overline{a}_g - (\overline{a}_g - \overline{a}_u))}{dD} \geq 0$. ■

PROOF OF PROPOSITION 3.4:

Proof. Since $r_g = 1 - \frac{(1-\varphi)}{\frac{p\varphi\overline{a_u}}{1-(1-p)^{\overline{a_u}}} + (1-\varphi) + p\varphi(\overline{a_g} - \overline{a_u})}$ and $\overline{a_g} - \overline{a_u}$ is independent of D,

$$\frac{d(r_g)}{d(D)} = \frac{d(r_g)}{d(\overline{a_u})} \cdot \frac{d(\overline{a_u})}{d(D)}$$

Proposition 2 has established that $\frac{d(\overline{a_u})}{d(D)} \geq 0$. Therefore, $\frac{d(r_g)}{d(D)} \geq 0$ if and only if $\frac{d(r_g)}{d(\overline{a_u})} \geq 0$.

With $\frac{\overline{a_u}}{1-(1-p)^{\overline{a_u}}} = x$, $\frac{d(r_g)}{d(\overline{a_u})} = \frac{d(r_g)}{d(x)} \cdot \frac{d(x)}{d(\overline{a_u})}$. Since $\frac{d(r_g)}{d(x)} > 0$, $\frac{d(r_g)}{d(\overline{a_u})} \geq 0$ if and only if $\frac{d(x)}{d(\overline{a_u})} \geq 0$.

Hence, I need to prove that $\frac{d(x)}{d(\overline{a_u})} \geq 0$.

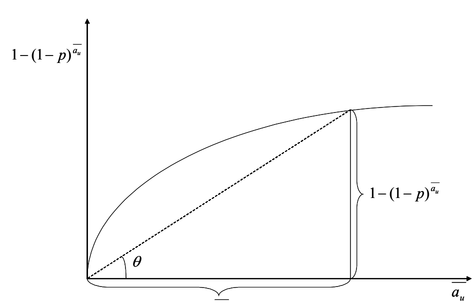
$1 - (1-p)^{\overline{a_u}}$ is plotted in the following graph as a function of $\overline{a_u}$. Since

$$\frac{d\left(1 - (1-p)^{\overline{a_u}}\right)}{d(\overline{a_u})} = -(1-p)^{\overline{a_u}} \cdot \ln(1-p) > 0$$

but

$$\frac{d^2\left(1 - (1-p)^{\overline{a_u}}\right)}{d(\overline{a_u})^2} = -(1-p)^{\overline{a_u}} \cdot (\ln(1-p))^2 < 0,$$

the curve is concave.



Clearly, it indicates that $x = \frac{\overline{a_u}}{1-(1-p)^{\overline{a_u}}} = \cot(\theta)$.The concavity of the curve suggests that as $\overline{a_u}$ increases, the angle of θ shrinks and $\cot(\theta)$ increases. Therefore, x increases in $\overline{a_u}$. ■

Results from two-moment Krusell-Smith approach:

Ω	$\{X, \sigma_a\}$
H_Ω	booms ($\log X$): $\log X' = 0.1261 + 0.9653 \log X + 0.3246\sigma_a$ recessions($\log X$): $\sigma_a' = 0.0079 + 0.0076 \log X + 0.8988\sigma_a$ booms (σ_a): $\log X' = -0.1485 + 0.9291 \log X + 1.0317\sigma_a$ recessions(σ_a): $\sigma_a' = 0.0789 + 0.0166 \log X + 0.6924\sigma_a$
R^2	booms ($\log X$): 0.9940 recessions($\log X$): 0.9287 booms (σ_a): 0.9571 recessions(σ_a): 0.5812
standard forecast error	booms ($\log X$): 0.0000069741% recessions($\log X$): 0.000068307% booms (σ_a): 0.00012513% recessions(σ_a):0.00097406%
maximum forecast error	booms ($\log X$): 0.000087730% recessions($\log X$):0.0016626% booms (σ_a):0.0014396% recessions(σ_a):0.028074%
Den Haan & Marcet test statistic (χ^2_7)	0.9216

REFERENCES

Aghion, Philippe and Howitt, Peter. "Growth and Unemployment." *Review of Economic Studies*, July 1994, 61(3), pp. 477-494.

Aghion, Philippe and Saint-Paul, Gilles. "Virtues of Bad Times." *Macroeconomic Dynamics*, September 1998, 2(3), pp. 322-44.

Aruoba, S. Boragan; Rubio-Ramirez, Juan F. and Fernandez-Villaverde, Jesus. "Comparing Solution Methods for Dynamic Equilibrium Economies." Working Paper 2003-27, Federal Reserve Bank of Atlanta, 2003.

Aw, Bee Yan; Chen, Xiaomin and Roberts, Mark J. "Firm-level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing." *Journal of Development Economics*, October 2001, 66(1), pp. 51-86.

Baily, Martin Neil; Bartelsman, Eric J. and Haltiwanger, John. "Labor Productivity: Structural Change and Cyclical Dynamics." *Review of Economics and Statistics*, August 2001, 83(3), pp. 420-433.

Baldwin, John R. *The Dynamics of Industrial Competition.* Cambridge University Press, 1995.

Barlevy, Gadi. "The Sullying Effect of Recessions." *Review of Economic Studies*, January 2002, 69(1), p65-96.

Basu, Sustanto. "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" *Quarterly Journal of Economics*, August 1996, 111(3), pp. 719-51.

Basu, Susanto and Fernald, John G. "Aggregate productivity and aggregate technology," European Economic Review, Elsevier, 2002, vol. 46(6), pages 963-991.

Basu, Susanto; Fernald, John G. and Shapiro, Matthew D. "Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?" NBER Working paper 8359, 2001.

Basu, Susanto and Miles S. Kimball, "Cyclical Productivity with Unobserved Input Variation," NBER Working Papers 5915, 1997, National Bureau of Economic Research, Inc.

Bowlus, Audra J. "Job Match Quality over the Business Cycle." *Panel Data and Labour Market Dynamics*, Amsterdam: North Holland, 1993, pp. 21-41.

Brsnahan, Timothy F. and Raff, Daniel M.G. "Intra-industry Heterogeneity and the Great Depression: the American Motor Vehicles Industry, 1929-1935." *Journal of Economic History*, June 1991, 51(2), pp. 317-31.

Caballero, Ricardo J, "A Fallacy of Composition," American Economic Review, American Economic Association, 1992, vol. 82(5), pages 1279-92.

Caballero, Ricardo J. and Hammour, Mohamad L. "The Cleansing Effect of Recessions." *American Economic Review*, December 1994, 84(5), pp. 1350-68.

Caballero, Ricardo J. and Hammour, Mohamad L. "On the timing and Efficiency of Creative Destruction." *Quarterly Journal of Economics*, August 1996, 111(3), pp. 805-52.

Caballero, Ricardo J. and Hammour, Mohamad L. "The Cost of Recessions Revisited: a Reverse-Liquidationist View." NBER Working Paper #7355, 1999.

Campbell, Jeffrey R. and Fisher, Jonas D.M. "Technology Choice and Employment Dynamics at Young and Old Plants." Working Paper Series-98-24, Federal Reserve Bank of Chicago, 1998.

Campbell, Jeffrey R. and Fisher, Jonas D.M. "Aggregate Employment Fluctuations with Microeconomic Asymmetries," *American Economic Review*, Volume 90, Number 5, December 2000. pp. 1323-1345

Campbell, Jeffrey R. and Fisher, Jonas D.M. "Idiosyncratic Risk and Aggregate Employment Dynamics." *Review of Economic Dynamics*, 2004.

Carlyle, Thomas. *Critical and Miscellaneous Essays.* New York, 1904, vol. 4, pp. 368-369.

Cooley, Thomas and Prescott, Edward. "Economic Growth and Business Cycles" (pp. 1-38), in Thomas Cooley, (Ed.), *Frontiers of Business Cycle Research*, Princeton, Princeton University Press, 1994.

Cooper, Russell; Haltiwanger, John; and Power, Laura. "Machine Replacement and Business Cycles: Lumps and Dumps." *American Economic Review,* 1999.

Davis, Steven J. and Haltiwanger, John. "Gross Job Flows." *Handbook of Labor Economics*, Amsterdam: North-Holland, 1999.

Davis, Steven J.; Haltiwanger, John and Schuh, Scott. *Job Creation and Job Destruction*, Cambridge, MIT Press,1996.

Davis, Steven J. and Haltiwanger, John. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." *Quarterly Journal of Economics*, August 1992, 107(3), pp. 818-63.

Den Haan, Wouter J. and Marcet, Albert. "Accuracy in Simulations." *Review of Economic Studies*, 1994, 61, pp. 3-18.

Dickens, William T. "The Productivity Crisis: Secular or Cyclical?" *Economic Letters*, 1982, 9(1), pp. 37-42.

Dunne, Timothy; Roberts, Mark J. and Samuelson, Larry. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics*, November 1989, 104(4), pp. 671-98.

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

Foote, Christopher L. "Trend Employment Growth and the Bunching of firm Creation and Destruction." *Quarterly Journal of Economics.* August 1998, 113(3), pp. 809-834.

Foster, Lucia; Haltiwanger, John and Krizan, C. J. "The Link Between Aggregate and Micro Productivity Growth: Evidence from Retail Trade." NBER working paper 9120, 2002.

Gomes, Joao, Jeremy Greenwood, and Sergio Rebelo. "Equilibrium Unemployment." *Journal of Monetary Economics*, August 2001, 48(1), pp. 109-52.

Goolsbee, Austan. "Investment Tax Incentives, Prices, and the Supply of Capital Goods." *Quarterly Journal of Economics,* February 1998, 113(1), pp. 121-148.

Hall, Robert E. "Labor Demand, Labor Supply, and Employment Volatility." in

Olivier J. Blanchard and Stanley Fischer, *NBER Macroeconomics Annual.* Cambridge, MA: MIT Press, 1991.

Hall, Robert E. "Lost Jobs." *Brookings Papers on Economic Activity*, 1995:1, pp. 221-256.

Hall, Robert E. "Reorganization." *Carnegie-Rochester Conference Series on Public Policy*, June 2000, pp. 1-22.

Lucas, Robert, E. "On the Size Distribution of Business Firms." *The Bell Journal of Economics*, Autumn1978, 9(2), pp. 508-523.

Lucas, Robert, E. *Models of Business Cycles,* Oxford: Basil Blackwell, 1987.

Jensen, J. Bradford; McGuckin, Robert H. and Stiroh, Kevin J. "The Impact of Vintage and Survival on Productivity: Evidence from Cohorts of U.S. Manufacturing Plants." Economic Studies Series Working paper 00-06, Census of Bureau, May 2000.

Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica*, May 1982, 50(3), pp. 649-70.

Krusell, Per. and Smith, Anthony A. Jr. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 1998, 106(5), pp. 867-895.

Kydland, Finn E. and Prescott, Edward C, 1982. "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50(6), pages 1345-70.

Mortensen, Dale and Pissarides,Christopher. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies,* July 1994, 61(3), pp. 397-415.

Montgomery, Edward and Wascher, William. "Creative Destruction and the Behavior of Productivity over the Business Cycle." *Review of Economics and Statistics,* February 1998, 70(1), pp. 168-172.

Pries, Michael J. "Persistence of Employment Fluctuations: a Model of Recurring Firm Loss." *Review of Economic Studies*, January 2004, 71(1), pp. 193-215.

Ramey, Garey and Watson, Joel. "Contractual Fragility, Job Destruction, and Business Cycles." *Quarterly Journal of Economics,* August 1997, 112(3), pp. 873-911.

Rust, John. "Using Randomization to Break the Curse of Dimensionality." *Econometrica*, May 1997, 65(3), pp. 487-516.

Rust, John. "Structural Estimation of Markov Decision Processes." *Handbook of Econometrics*, 1994, V IV.

Schumpeter, Joseph A. "Depressions." in Douglas Brown et al., *Economics of the Recovery Program*, New York, 1934, pp. 3-12.

Young, Eric R. "Approximate Aggregation: an Obstacle Course for the Krusell-Smith Algorithm." Florida University, 2002.