The Scarring Effect of Recessions

Min Ouyang*

January 2005

Abstract

This paper explores the role that recessions play in resource allocation. The conventional cleansing view, advanced by Schumpeter in 1934, argues that recessions promote more efficient resource allocation by driving out less productive units and freeing up resources for better uses. However, empirical evidence is at odds with this view: average labor productivity is procyclical, and jobs created during recessions tend to be short-lived. This paper posits an additional "scarring" effect: recessions "scar" the economy by killing off "potentially good firms". By adding learning to a vintage model, I show that as a recession arrives and persists, the reduced profitability limits the scope of learning, makes labor less concentrated on good firms, and thus pulls down average productivity. Calibrating my model using data on job flows from the U.S. manufacturing sector, I find that the scarring effect is likely to dominate the conventional cleansing effect, and can account for the observed pro-cyclical average labor productivity.

Keywords: Business Cycles, Cleansing Effect, Scarring Effect, Creative Destruction, Learning, Job Flows.

JEL: E32, L16, C61

*University of Maryland at College Park, Department of Economics, Tydings Hall, College Park, MD, 20742, ouyang@econ.umd.edu. I am grateful to John Haltiwanger, Micheal Pries, John Rust, John Shea and Daniel Vincent for their generous advice and encouragements. I am especially indebted to John Shea for his particularly detailed comments, and to John Rust for providing C codes for multiple-dimentional interpolation. I also thank seminar participants at the University of Maryland, the 2003 Inter-University Students' Conference and the 2004 Midwest Macroeconomics Meeting. S. Boragan Aruoba made helpful suggestions regarding this draft. All errors are mine.

"[Depressions] are the means to reconstruct each time the economic system on a more efficient plan. But they inflict losses while they last, drive firms into the bankruptcy court...before the ground is clear and the way paved for new achievement..." Joseph A. Schumpeter (1934, p. 8)

"You must empty-out the bathing-tub, but not the baby along with it." Thomas Carlyle (1904, p. 368)

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.[1] Attempting to explain these cyclical patterns, an assortment of theoretical work has arisen returning to the Schumpeterian cleansing view.[2] 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,[3] 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.[4]

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 in Caballero and Hammour (1994), where they 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,

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

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

[3] 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).

[4] 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.

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,[5] 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.[6]

In this paper, 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"; the 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 Hammour (1994) with *learning* in the spirit of Jovanovic (1982). As in their model, exogenous technological progress introduces a force of creative destruction that drives in technologically sophisticated entrants to displace older, outmoded firms.[7] However, in my model, firms of the same vintage also differ in "type": some are good and others are bad. Firm type 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, these types 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

[5] 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..

[6] 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.

[7] 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.

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, the limited learning shifts labor toward bad firms, creating a scarring effect that pulls down average labor productivity. The question then becomes, which one 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 type together can explain the observed heterogeneous firm-level productivity. The vintage component suggests that entering cohorts are more productive than incumbents.[8] The type component implies that each vintage cohort is itself a heterogeneous group. Vintage and type 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.[9]

The existing empirical literature has advanced learning and creative destruction as powerful tools to understand the patterns of firm turnover and industrial dynamics.[10] The significance of their interaction has also been suggested. Davis and Haltiwanger (1999) note, "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

[8] 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.

[9] 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).

[10] 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).

the scarring effect of recessions.

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

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 types. As time passes by, good firms survive and bad firms leave. Allocative inefficiency comes from costly entry and time-consuming learning.

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 exit. Each firm is characterized by two components:

1. Vintage $A(t-a)$.

2. Type θ.

$A(t)$ is the exogenous technological progress that grows at a constant rate $\gamma > 0$. A firm that enters the industry in period t embodies the leading technology $A(t)$, which becomes its vintage and will affect its production afterward. Only entrants have access to the leading technology. Incumbents cannot retool. With a as the firm age, $A(t-a)$ is the vintage of a firm of age a in period t. Since $A(t)$ grows exogenously, young firms are always technologically more advanced than old firms.

At the time of entry, a firm is endowed with a type θ. Hence, firms of the same vintage differ in type. 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.[11] I call θ "the technology adoption type".

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.

A firm of age a and type θ produces output in period t, given by

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

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 type θ in period t is

$$\pi_t(a, \theta) = P_t \cdot A(t) \cdot (1+\gamma)^{-a} \cdot (\theta + \varepsilon_t) - 1. \tag{2}$$

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.

2.2 "All-Or-Nothing" Learning

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

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

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

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 type; conversely, an observation of x_t within $[\theta_b - \omega, \theta_g - \omega)$ indicates that it is a bad type. 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 types 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 $\theta^e = \theta_u \equiv$ 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 types 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 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 initial probability distribution:

$$\left(\begin{array}{ccc} 0, & 1, & 0 \end{array} \right),$$

and a transition matrix

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ p \cdot \varphi \quad , & 1 - p \quad , & p \cdot (1 - \varphi) \\ 0 & 0 & 1 \end{array} \right).$$

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

$$\left(\begin{array}{ccc} \varphi, & 0, & (1 - \varphi) \end{array} \right).$$

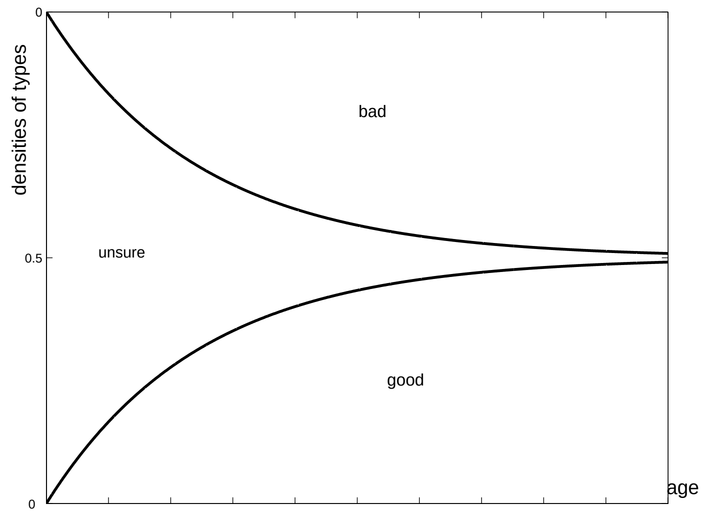


Figure 1: Dynamics of a Birth Cohort: the distance between the concave curve and the bottom axis measures the density of firms with $\theta^e = \theta_g$; the distance between the convex curve and the top axis measures the firms with $\theta^e = \theta_b$; the distance between the two curves measures the density of unsure firms (firms with $\theta^e = \theta_u$).

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 cross-type firm distribution within a birth cohort as shown in Figure 1, with the horizontal axis depicting the age of a cohort *across time*. The densities of firms that are certain about their types, 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 1 can tell. If we let the horizontal axis depict the cross-sectional distribution of firm ages at any instant, then Figure 1 can be interpreted as the cross-age and cross-type firm distribution 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 1. Under this interpretation, Figure 1 indicates that at any instant older cohorts contain fewer unsure firms, because they have lived longer and learned more.

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 operation 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 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 types. The part of F that measures the number of firms with θ^e and a is denoted $f(\theta^e, a)$. D is an exogenous demand parameter; it captures aggregate conditions and is fully observable. The part of the law of motion for D is exogenous, described by D's transition matrix. The part of 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, 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:

$$V\left(\theta^{e}, a; F, D\right) = E\left[\pi\left(\theta^{e}, a\right)|F, D\right] + \beta E\left[\max\left(0, V\left(\theta^{e\prime}, a+1; F', D'\right)\right)|F, D\right] \tag{3}$$

subject to

$$F' = H\left(F, D\right)$$

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

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

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 $f\left(\theta_{u}, 0; F, D\right)$ represents the size of the entering cohort when the aggregate state is (F, D), and letting c represent the entry cost, I have

$$c = C\left(f\left(\theta_{u}, 0; F, D\right)\right), c > 0 \text{ and } C' \geq 0. \tag{4}$$

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.[12] The free entry condition equates a firm's entry cost to its value of entry, and can be written as

$$V\left(\theta_{u}, 0; F, D\right) = C\left(f\left(\theta_{u}, 0; F, D\right)\right). \tag{5}$$

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

$$P\left(F, D\right) = \frac{D}{Q\left(F, D\right)} \tag{6}$$

[12] See Subsection 3.3.1 for further discussion.

Q represents aggregate output; it equals the the sum of production over heterogeneous firms. Given (1), the sequence of events implies that:[13]

$$Q(F,D) = Q(F') = \sum_a \sum_{\theta^e} A \cdot (1+\gamma)^{-a} \cdot \theta^e \cdot f'(\theta^e, a), \tag{7}$$

where $f'(\theta^e, a)$ measure the number of in-operation 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).

(6) implies that high output drives down the price. (7) implies that Q depends on not only the number of firms in operation, but also their distribution. More firms yield higher output and drive down the price; the more the distribution is skewed toward younger vintages and better types, 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 (6) and (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 types of firms in operation – unsure and good.

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

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

[13] 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.

The proof for Proposition 1 presented in the appendix is not restricted to all-or-nothing learning. Hence, Proposition 1 holds for *any* learning process. It follows from the fact that firms with smaller a and higher θ^e have higher expected value of staying. As V is strictly decreasing in a, firms with belief θ^e 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 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.[14] 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.

3.1 The Steady State

I consider a steady state a recursive competitive equilibrium with time-invariant aggregate states.[15] 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 2: With D constant over time, there exists a unique

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

[15] I call it "steady state" following Caballero and Hammour (1994). Although it is called "steady", the steady-state price decreases but the steady-state average labor productivity increases over time driven by technological progress.

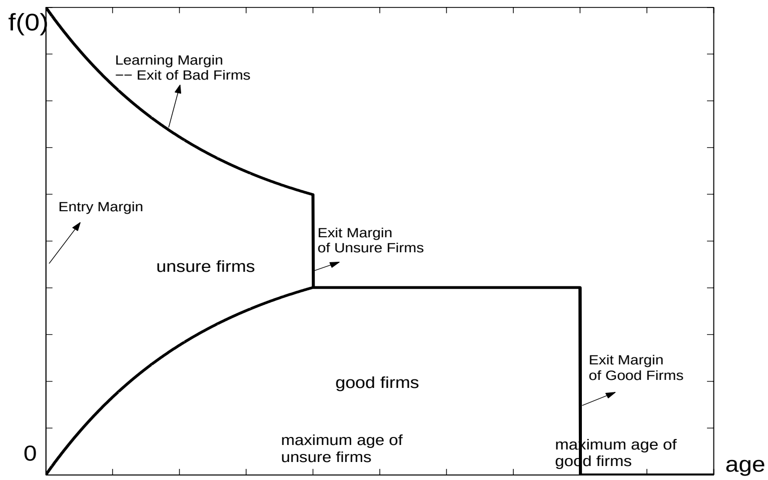


Figure 2: The Steady-state Labor Distribution and Job Flows: the distance between the lower curve (extended as the horizontal line) and the bottom axis measures the density of good firms; the distance between the two curves measures the density of unsure firms.

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 2. Like Figure 1, there are two ways to interpret Figure 2. 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. Since then, the cohort contains only good firms and the number of good firms remains constant because learning has stopped. Good firms live until $\overline{a_g}$. The vintage after $\overline{a_g}$ is too old even for good firms to survive.

Second, Figure 2 also displays the firm distribution across ages and types at any one time, with the horizontal axis depicting the cohort age *across section*. At the steady state, firms of different ages coexist. Since older cohorts have lived longer and learned more, cohort size declines and the density of good firms increases with age. 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: they are 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 the 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 hired only by good firms.

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 will illustrate 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 and 4.

> Proposition 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, compare two steady states with different demand levels, $D_h > D_l$. For any time t, (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 type 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:

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

I prove in the appendix that (8) gives an unique solution for $\overline{a}_g - \overline{a}_u$ as long as $\theta_g > \theta_u$. Since D does not enter (8), $\overline{a}_g - \overline{a}_u$ is independent of demand: $\frac{d(\overline{a}_g - \overline{a}_u)}{dD} = 0$. (8) suggests that the demand level does not affect the gap between the exit ages of good and unsure firms. They tend to co-move across steady states with the same magnitude.

The steady-state job destruction rate, denoted jd^{ss}, equals the following:[16]

$$jd^{ss} = \frac{1}{\overline{a_u}\cdot\varphi + [\frac{1-\varphi}{p} + (\overline{a_g} - \overline{a_u})\cdot\varphi]\cdot[1-(1-p)^{\overline{a_u}+1}]}. \tag{9}$$

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 less jobs are destroyed at the exit margins.

To summarize, Proposition 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.

[16] 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}.

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

> Proposition 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 3.

Figure 3 displays the steady-state industry structures corresponding to two demand levels.[17] The cleansing effect formalized in Proposition 3 is shown as the 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 amount of *older good firms*. The latter, shown as the shaded area in Figure 3, is the scarring effect of recessions.

The scarring effect stems from learning. New entrants begin unsure of their type, 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 types depends on

[17] 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.

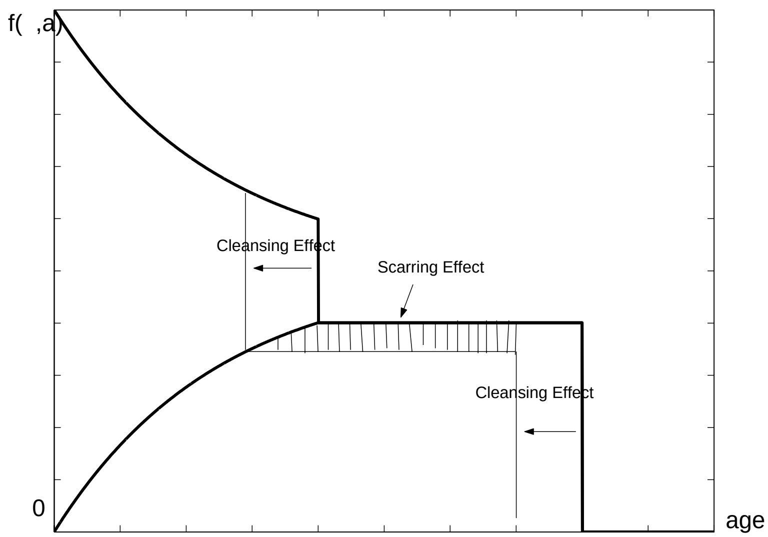


Figure 3: Cleansing and Scarring

how many learning chances they have. If firms could live forever, eventually all the potentially good firms would get to realize their true types. 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 types, so that the number of *old good firms* in operation after age $\overline{a_u}$ are 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 $\overline{a_u}$.[18] 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.

To summarize from Proposition 3 and Proposition 4, a low-demand steady state features a better average vintage, yet a less efficient cross-type distribution of labor. If the comparative static results carry over when

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

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 types. 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 3. Both exit margins shift as soon as demand drops so that the cleansing effect takes place immediately.[19] However, the scarring effect takes place gradually. When a recession first arrives, the group of firms already in the shaded area in Figure 3 will not leave despite the shift in exit margins, since they know their true types 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 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 independent of entry size; second, I allow the process of learning to be more complicated than "all-or-nothing".

3.3.1 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 this 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

[19] My numerical exercises with demand fluctuations 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 discussed in Section 4.

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 coming in with rising demand for capital, it becomes increasingly costly to adjust the capital stock. 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.[20]

3.3.2 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, I divide firms into two groups, young and old.[21] With l_g^o denoting the fraction

[20] See footnote 6.

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

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 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 type 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 types 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,[22] 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.

[22] 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.

4 Quantitative Implications with Stochastic Demand Fluctuations

I establishe in Section 3 that across steady states, my model delivers two competing effects – cleansing and scarring. Now the questions are, whether the two effects carry over when demand fluctuates stochastically, and which one 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 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 the U.S. manufacturing industry. As I have argued, recessions clear out old firms, including some good firms that have not yet learned their type. Therefore, the model allows us to use the job destruction rate to make inferences on the size of the cleansing and scarring effects.

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 types, 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[23]

$$X(F) = \sum_{a}\sum_{\theta^e}(1+\gamma)^{-a} \cdot \theta^e \cdot f(\theta^e, a). \tag{10}$$

Combining (9) with (6) and (7), I get

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

A is the leading technology; F' is the updated firm distribution after the 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 type θ^e and age a, given aggregate state (F, D), equals

$$\pi(a, \theta; F, D) = \frac{D}{X(H(F, D))} \cdot (1+\gamma)^{-a} \cdot (\theta + \varepsilon) - 1. \tag{11}$$

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 a Ω through the following procedure. In step 1, I choose a candidate Ω. In step 2, I choose the 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 approach, 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,[24] I simulate the behavior of a continuum of firms along a random path of demand

[23] X can be interpreted as detrended output.

[24] 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 the result.

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 1: The Estimated Laws of Motion and Measures of Fit

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. 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 1. As shown in Table 1, the estimated H_x is log-linear. And the fit of H_x is quite good, as suggested by the high R^2, the low standard forecast error, 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 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.[25] The test statistic is 0.8007,

[25] 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

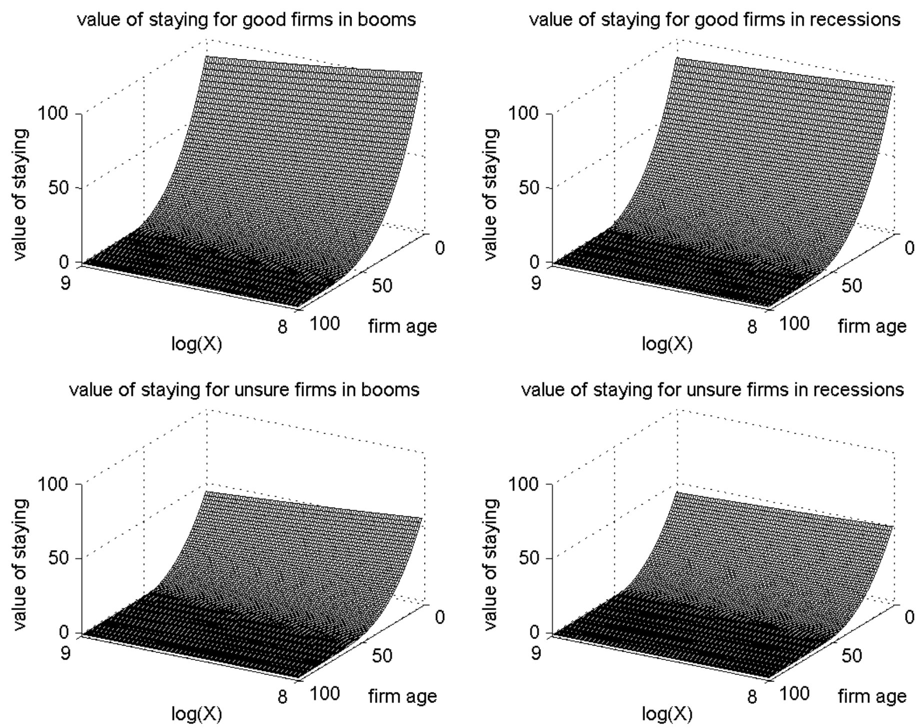


Figure 4: Expected Value of Staying: aggregate state variables are D and $\log X$ (the log of detrended output), firm-level state variables are firm age and belief (good or unsure); applied calibration is summarized in Table 2 and discussed in Subsection 4.2.

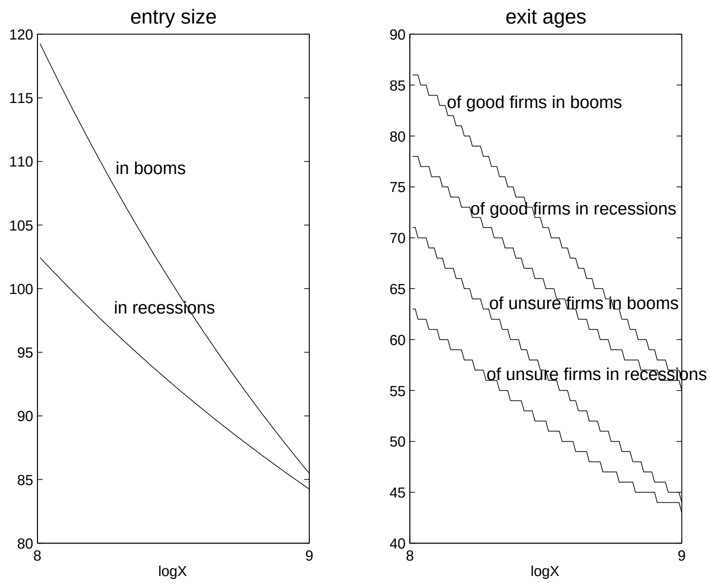


Figure 5: Industry-level Policy Functions: Entry Size and Exit Ages. Aggregate states are D (booms or recessions) and $\log X$ (the log of detrended output).

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.[26] 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 4 displays the value of staying for heterogeneous firms as a function of a, θ^e, D and X ($\log X$). Figure 5 displays the corresponding optimal exit ages and entry sizes. The properties of value functions and exit ages stated in Proposition 2 are satisfied in both figures: given the aggregate

$h(G_t)$ 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.

[26] 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 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 2: Base-line Parameterization of the Model

state, the value of staying is increasing in the expected type θ^e and decreasing in firm age; and good firms exit at an older age than unsure firms.

To conclude, Table 1, Figures 4 and 5 suggest that my solution using X to approximate the aggregate state closely replicates optimal firm behavior at the equilibrium.[27] Therefore, I use the solution based on $\Omega = \{X\}$ to generate all the series in the subsequent analysis.

4.2 Calibration

Table 2 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 (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

[27] Those results were robust when I experimented with different parameterizations of the model. Although it suggests that the simulation 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.

growth rate as 0.007 by attributing all output growth of US manufacturing from 1972:2 to 1983:4 to technical progress. To make a convenient comparison with their result 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 path where demand level is unchanging. The industry structures at the stable points are similar to those at the steady states, which allows me to use steady state conditions for approximation.[28] 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 $\overline{a_u}'$ represent the exit ages at the low-demand steady state. The steady-state job destruction rate, denoted jd^{ss}, is given by (9).

Secondly, 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 $\overline{a_u}'$ at once, so that the job destruction rate at the beginning

[28] 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 $1 - \mu$.

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

Table 3: Descriptive Statistics of Quarterly Job Destruction in U.S. Manufacturing (1972:2-1993:4), constructed by Davis and Haltiwanger.

of a recession, denoted as $jd_{\max}$, is approximately:[29]

$$jd_{\max} = \frac{\begin{array}{c}\varphi\left[1-(1-p)^{\overline{a_u}+1}\right](\overline{a_g}-\overline{a_g}') + \\ 2\cdot \quad \left[\frac{1}{p}+\varphi-1-\frac{1}{p}(1-p)^{\overline{a_u}-\overline{a_u}'}\right](1-p)^{\overline{a_u}'+1} + \\ (1-\varphi)\end{array}}{\begin{array}{c}\varphi(\overline{a_u}+\overline{a_u}') + \frac{(1-\varphi)}{p}\left[2-(1-p)^{\overline{a_u}+1}-(1-p)^{\overline{a_u}'+1}\right] + \\ \varphi\left[1-(1-p)^{\overline{a_u}+1}\right](\overline{a_g}-\overline{a_u}) + \varphi\left[1-(1-p)^{\overline{a_u}'+1}\right](\overline{a_g}'-\overline{a_u}')\end{array}} \tag{12}$$

Thirdly, 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:

$$jd_{\min} = \frac{(1-\varphi)\left[1-(1-p)^{\overline{a_u}'+1}\right]}{\overline{a_u}'\cdot\varphi + [\frac{1-\varphi}{p} + (\overline{a_g}'-\overline{a_u}')\cdot\varphi]\cdot[1-(1-p)^{\overline{a_u}'+1}]} \tag{13}$$

Now I turn to data for conditions on jd^{ss}, $jd_{\max}$, and $jd_{\min}$. Table 3 lists descriptive statistics for the job destruction series of the U.S. manufacturing sector from 1972:2 to 1993:4 compliled 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%.[30] 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 (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

[29] 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.

[30] The one implied by $(\overline{a_g}', \overline{a_u}')$ is slightly higher since I assume $\overline{a_g}' < \overline{a_g}$ and $\overline{a_u}' < \overline{a_u}$.

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 1.66. 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 5.19%.

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.

4.3.1 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 distributions, 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 6, 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 if demand had remained high. After 70 quarters, the job destruction rate converges to 5.63%, still above its original value. Hence, the conventional cleansing effect on job destruction from the comparative static exercises 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.[31]

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 6, 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-type 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 2.2 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.

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 over-shift 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

[31] 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.

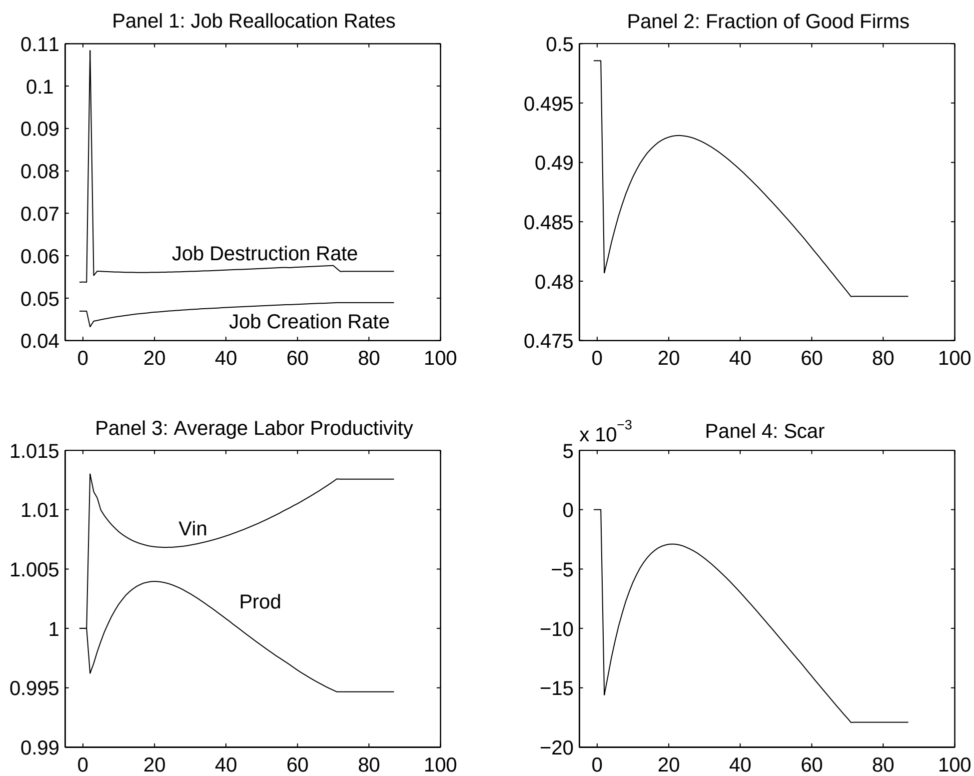


Figure 6: Response to a Negative Demand Shock: *vin* is the detrended average labor productivity driven only by the cleansing effect, *prod* is the detrended average labor productivity driven by both the cleansing effect and the scarring effect. $Scar = prod - vin$. The horizontal axis denotes quarters, with the quarter labeled 0 denoting the onset of a recession.

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-type 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 6 suggest that both the conventional cleansing effect established in Proposition 2, and the scarring effect established in Proposition 3, carry over with probabilistic business cycles.

4.3.2 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 type, 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:

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

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 6 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 6. 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.

4.3.3 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 4. The reported statistics are means (standard deviations) based on 1000 simulated samples. Sample statistics for U.S. Manufacturing data for the 87 quarters from 1972:2 to1993:4 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 4 suggests that my calibrated model can replicate the observed patterns of job flows; moreover, the positive correlation coefficient of 0.1675

	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}	0.37%(0.0535%)	0.95%
$corr(prod, q)$	0.1675(0.7504)	0.5537*

Table 4: 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.

between *prod* and *q* implies that my model generates procyclical average labor productivity for the U.S. manufacturing sector in the relevant period. Put differently, under my benchmark calibration the scarring effect on cyclical productivity dominates the cleansing effect.

4.4 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. Table 5 presents their results for the period 1979-1990: a quarterly technological pace of 0.0037 for durable manufacturing, a pace 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 6 together with results from the

	1979-1990 (*quarterly*)
Durable Manufacturing	0.0037
Non-durable Manufacturing	0.0027
Non-manufacturing	−0.0005
Private Sector	0.0005

Table 5: Estimated Growth in TFP (Basu, Fernald and Shapiro 2001)

Calibration Results	$\boldsymbol{\gamma = 0.003}$	$\boldsymbol{\gamma = 0.007}$
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 6: 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.

baseline parameterization.

The calibration results in Table 6 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 job flows.[32] The simulated responses suggest

[32] 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

	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 7: 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.

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 difference between the vintages that have been killed and the ones that have survived, so that the impact of 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 also every period they learn, more learners exit as bad firms. It also gives a larger scarring effect on average labor productivity: a faster learning pace implies a higher *opportunity cost* of not learning; a smaller prior probability of being good suggests that learning has a greater *marginal* impact on cross-type efficiency.

Table 7 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. This is a strong procyclical behavior of

my model with $\gamma = 0.003$ suggests that more job destruction comes from learning rather than creative destruction.

productivity and is close to the one suggested by data.

5 Conclusion

How do recessions affect resource allocation? My paper suggests learning has important consequences for this question. I posit that in addition to the cleansing effect argued by previous authors, recessions create a scarring effect by interrupting the learning process. They 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 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 paper highlights firm age as an indicator for the number of learning opportunities. The existing empirical literature documents that firm age has important explanatory power for micro-level job flow patterns.[33] 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 sectors of 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

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

introduces fewer potentially good firms on the creation side. Whether "scarring" dominates "cleansing" in sectors other than manufacturing remains an interesting question.

APPENDIX

Proof of Proposition 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

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

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:

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

and

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

where $\pi_t^i(\theta^e, a_1 + t - t_0)$ 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 $\left\{ P_t^i (F, D) \right\}_{t \geq t_0}$ conditional on the realization of path i. Since price is competitive and firms are price takers, I must have:

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

and

$$V\left(\theta^{e},a_{2};F,D\right)=\sum_{i=1}^{n}\sum_{t=t_0}^{t_2^i}\left\{\beta^{t-t_0}\cdot\left[A\left(t_0-a_2\right)\theta^{e}P_t^i\left(F,D\right)-1\right]\right\}\cdot p^{i}.$$

There are three possibilities for any i.
Possibility 1, if $t_1^i=t_2^i=t^i$:
since $A\left(t_0-a_1\right)<A\left(t_0-a_2\right)$,

$$\left(t_0-a_1\right)\theta^{e}P_t^i\left(F,D\right)-1<A\left(t_0-a_2\right)\theta^{e}P_t^i\left(F,D\right)-1$$

holds for any t. Hence,

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

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

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

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:

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

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

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

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 $\overline{a}\left(\theta^e; F, D\right)$ and $\overline{a}\left(\theta^{e\prime}; F, D\right) \geq \overline{a}\left(\theta^e; F, D\right)$, for $\theta^{e\prime} > \theta^e$.

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

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

but

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

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

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

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

PROOF OF PROPOSITION 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.

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

so that:

$$D_t = P_t A_t \cdot \sum_a \sum_{\theta^e} \theta^e f_t\left(\theta^e, a\right)(1+\gamma)^{-a}. \tag{14}$$

By definition, a steady state features constant level of demand, $D_t = D$ ($\forall$ t). and time-invariant firm distribution. Let $f\left(\theta^e, a\right)$ 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

$$PA = \frac{(1+\gamma)^{\overline{a_g}}}{\theta_g}. \tag{15}$$

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):

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

which can be re-written as:

$$F\left(\overline{a}_g - \overline{a}_u\right) = G\left(\overline{a}_g - \overline{a}_u\right)$$

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

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

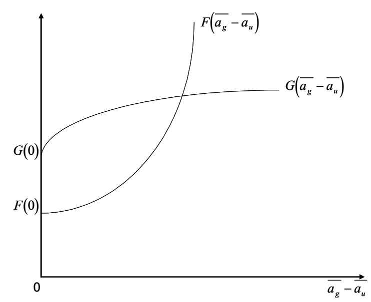
moreover,

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

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:

$$\frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left\{ \begin{array}{c} \sum\limits_{a=1}^{\overline{a}_u} \beta^a \left[\begin{array}{c} (1-p)^a \left(\frac{\theta_u}{(1+\gamma)^a} - 1 \right) + \\ \varphi\left(1-(1-p)^a\right)\left(\frac{\theta_g}{(1+\gamma)^a} - 1 \right) \end{array} \right] + \\ \varphi\left(1-(1-p)^{\overline{a}_u+1}\right) \sum\limits_{a=\overline{a}_u+1}^{\overline{a}_g} \beta^a \left(\frac{\theta_g}{(1+\gamma)^a} - 1 \right) + \\ \theta_u - 1 \end{array} \right\} = C(f(0)) \quad (16)$$

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:

$$f(0) \cdot \frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left[\begin{array}{c} (\theta_u - \varphi\theta_g) \sum\limits_{a=1}^{\overline{a}_u} \left(\frac{1-p}{1+\gamma} \right)^a + \varphi\theta_g \sum\limits_{a=1}^{\overline{a}_g} \left(\frac{1}{1+\gamma} \right)^a + \\ \varphi\theta_g (1-p)^{\overline{a}_u+1} \sum\limits_{a=\overline{a}_u+1}^{\overline{a}_g} \left(\frac{1}{1+\gamma} \right)^a \end{array} \right] = D \quad (17)$$

$\overline{a}_g - \overline{a}_u$ 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:

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:

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

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

$$= D$$

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 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,

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

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 $\frac{d(x)}{d(\overline{a_u})} \geq 0$.

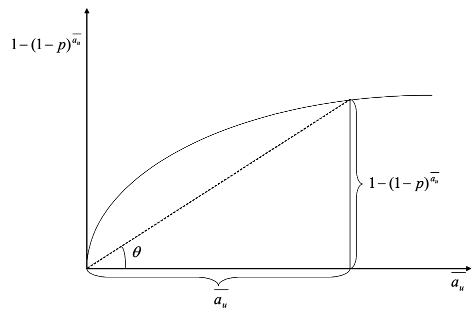
$1-(1-p)^{\overline{a_u}}$ is plotted in the following graph as a function of $\overline{a_u}$. Since

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

but

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

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

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. 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. "Idiosyncratic Risk and Aggregate Employment Dynamics." NBER Working Paper #7936, 2001.

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.

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.

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.

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.