The Scarring Effect of Recessions

Min Ouyang*

March 2008

Abstract

According to the conventional Schumpeterian view, recessions improve resource allocation by driving out less productive producers. We posit that recessions bring an additional scarring effect by impeding the developments of potentially superior producers, which can be destroyed during their infancy and never realize their potential. A model combining creative destruction with learning is developed to capture both the cleansing and scarring effects. A key ingredient of our model is that idiosyncratic productivity is not directly observable, but can be learned over time. When calibrated with statistics on productivity and on cyclical entry and exit, the model suggests that the scarring effect can dominate the cleansing effect, and can give rise to lower average productivity during recessions.

Keywords: Cleansing Effect, Scarring Effect, Creative Destruction, Learning, Demand Shocks.

JEL: E32, L16, C61

*University of California at Irvine, Department of Economics, 3151 Social Science Plaza B, Irvine, CA, 92697, mouyang@uci.edu. I am indebted to Bob King, an anonymous referee, Richard Rogerson, and John Shea for their detailed comments. I also thank Ricardo Caballero, Gadi Barlevy, Mark Gertler, John Haltiwanger, Hugo Hopenhayn, Micheal Pries, Vincenzo Quadrini, John Rust, Dmitriy Stolyarov and Daniel Vincent for their generous advice. Seminar participants at UMD, Yale, USC, UC-Irvine, Kansas City Fed, San Francisco Fed, Cleveland Fed, McGill, Queen's, and UC-San Diego provided helpful discussions. 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 (1899, p. 368)

1 Introduction

How do recessions affect resource allocation? This question has long attracted the attention of economists. As early as 1934, Schumpeter advanced the concept of "cleansing": recessions are times when outdated techniques and products are driven out, and resources are freed for more productive uses. This view has been revived in an assortment of theoretical work such as Caballero and Hammour (1994, 1996), Hall (2000), and Mortensen and Pissarides (1994). While none of these theories argue that recessions actually increase welfare, some adverse-cleansing views have emerged recently. Barlevy (2002, 2003) posits that the cleansing effect can be dominated or reversed by other effects arising from on-the-job search or credit market frictions. Therefore, the impact of recessions on resource allocation should be more complicated than what the conventional wisdom suggests.

This paper proposes a scarring effect of recessions, motivated by three empirical findings. First, a close examination of the cleansing hypothesis with data often returns confusing results: while producers' deaths do surge during recessions, the exiters are not necessarily the least productive. For example, Baden-full (1989) examines exits in the British steel industry during the 1980s recession and finds that many of the closing plants were more profitable than the surviving ones (p. 950). Second, many producers exit young. According to Dunne et al. (1989), over 75 percent of the exiting plants in the U.S. manufacturing sector were no older than five years (Table 1, p. 676). Most importantly, recessions appear to affect producers' deaths *disproportionally.* Figure 1 plots the exit rate for U.S. manufacturing plants across three age categories. Apparently, it is young producers' deaths that surge the most during recessions. For example, during the second quarter of 1984, the quarterly exit rate jumped from 1.35% to 3.41% for plants aged one year or less; such a rate increased from 0.77% to 0.87% for plants aged between one and nine years; however, for plants aged 10 years or more, it only rose from 0.35% to 0.36%. In this paper, we ask if these findings can be reconciled, and what would add to the conventional cleansing hypothesis once we consider what young producers' deaths imply.

We assert that young producers' deaths suggest the loss of some potentially good producers, so that, in addition to the cleansing effect, recessions bring a scarring effect. While recessions do drive out some of the least productive producers as the cleansing hypothesis suggests, they also kill of some of the potentially good ones — those that have the potential to be proven productive in the future close down because of reduced profitability. The exit of potentially good producers leaves scars when a recession arrives, and such scars deepen as the recession persists. Therefore, with the exit of potentially good producers taken into consideration, the overall impact of recessions on resource allocation should depend on the relative magnitude of two competing effects — cleansing and scarring.

To understand the scarring effect, consider a producer's life cycle. Suppose we call a producer a firm. Often a firm starts without knowing its own quality. Uncertainty may come from the unobservable talent of the manager, the unknown appeal for a product, or the unpredictable profitability of a store location. As the firm operates, realized revenue signals its true quality: high revenue suggests that the firm is more likely productive and encourages its continuing operation; low revenue suggests otherwise. The longer a firm operates, the more it knows its own quality. Therefore, the potentially good firms – those that do not yet know they are good – must be relatively young. During recessions, the profitability declines in general so that a firm cannot bear to learn as *long* as during good times. A potentially good firm that would have survived during good times, might thus exit during recessions *before* it learns. At the industry level, the exit of potentially good firms reduces the proportion of good firms at present times, as well as in the future because fewer potentially good firms are left to learn. The reduced proportion of good firms lowers the average productivity, which is defined in this paper as a scarring effect.

The above story reflects the spirit of learning, theoretically proposed by Jovanovic (1982) and has been promoted by the empirical literature as a powerful tool to understand firm turnover (e.g., Caves, 1998, and Foster et al., 2008). In recent years, a number of authors (e.g., Moscarini, 2003, Pries, 2004, Pries and Rogerson, 2005) have devised models with learning to understand the nature and the impact of reallocation. This paper borrows a simplified learning mechanism from Pries (2004) with the vintage framework of Caballero and Hammour (1994) to capture cleansing and scarring theoretically.

As the model setup, firm-level productivity is decomposed into two components – vintage and unobservable idiosyncratic productivity – so that the industrial average productivity is determined by the distribution of firms across both dimensions. The idiosyncratic productivity is not directly observable, but can be learnt over time. Demand variations serve as the source of economic fluctuations. Lower demand reduces profitability in general so that firms' exit ages become younger. Younger exit-ages directs, on the one hand, resources to younger and more productive vintages, causing a cleansing effect that raises average productivity; on the other hand, they *truncate* the learning process that directs resources toward firms with higher idiosyncratic productivity, creating a scarring effect that pulls down average productivity. Hence, recessions cause two competing effects – cleansing and scarring. The question then becomes, which effect dominates?

We turn to the data to explore the quantitative implications of the scarring effect. We calibrate our model so that the equilibrium exit-rates, entrants' productivity growth, cohort productivity differential, and cyclical entry and exit rates match those observed in the U.S. manufacturing sector. An application of the calibrated model to stochastic demand fluctuations shows that the scarring effect is likely to dominate the cleansing effect, and generate lower average productivity during recessions.

Previous authors have also critiqued the cleansing hypothesis, but the scarring effect differs from their proposed adverse effect in important ways. The focus of Ramey and Watson (1997) and Caballero and Hammour (2005) is on whether cyclical reallocation is socially efficient: in their models, recessions still promote more productive allocation of resources, but are associated with lower social welfare. The sullying effect proposed by Barlevy (2002) arises from reduced entry (creation) rather than concentrated exit (destruction). Barlevy (2003) focuses on credit market imperfections rather than the learning of unobservable qualities; moreover, the scarring effect impacts resource allocation both in present times and in the future – a dynamic effect that is missing in Barlevy (2003). Nevertheless, these various adverse-cleansing effects should be seen as complementary effects that likely amplify each other in reality. For example, during recessions, the credit market frictions can further tighten young businesses' borrowing constraints, so that more potentially good businesses are driven out before they learn; as a result, credit market frictions deepen the scarring effect.

The rest of the paper is organized as follows. Section 2 lays out the model. The cleansing and scarring effects are motivated in Section 3 by analytical comparative static exercises. Section 4 applies the model to stochastic demand fluctuations, confirms that the cleansing and scarring effects carry over, and studies their quantitative implications using data from the U.S. manufacturing sector. We conclude in Section 5.

2 A Renovating Industry with Learning

Consider an industry where labor and capital combine in fixed proportions to produce a homogenous output. Firms that enter at different times coexist. Each age cohort consists of a continuum of firms. Each firm hires one worker, and is characterized by two components:

1. Vintage;
2. idiosyncratic productivity.

A firm's vintage is given by an exogenous technological progress $\{A_t\}_0^\infty$ that grows at a constant rate $\gamma > 0$ so that

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

where A_0 is a constant. A firm enters the industry embodied with the leading technology. We assume that, only entrants have access to the updated technology, incumbents cannot retool. Since technology grows exogenously, young firms are always technologically more advanced than old firms. With a as the firm age, the vintage of a firm of age a in period t is A_{t-a},

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

At the time of entry, a firm is endowed with idiosyncratic productivity θ. It 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. The key assumption regarding θ is that its value, although fixed at the time of entry, is not directly observable.

The period-t output of a firm of age a and idiosyncratic productivity θ is:

$$q_t(a,\theta) = A_{t-a}x_t, \tag{1}$$

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. We set the wage rate to one by normalization, and let P_t denote the output price in period t. Then the period-t profit generated by a firm of age a and idiosyncratic productivity θ is

$$\pi_t\,(a,\theta) = P_t A_0(1+\gamma)^{t-a}(\theta + \varepsilon_t) - 1. \tag{2}$$

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.

2.1 "All-Or-Nothing" Learning

Each firm is a price taker and profit maximizer. It attempts to resolve the uncertainty about θ to decide whether to continue or terminate the production. The random component ε_t represents transitory factors that are independent of the idiosyncratic productivity θ. Assuming that ε_t has mean zero, we have $E_t(x_t) = E_t(\theta) + E_t(\varepsilon_t) = E_t(\theta)$. Hence, given knowledge of the distribution of ε_t, a sequence of observations of x_t allows the firm to learn about its θ. Here we borrow from Pries (2004) an "all-or-nothing" learning process by assuming only two values of θ: 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 $[\theta_b-\omega,\theta_b+w]$. Lastly, θ_g, θ_b and ω satisfy $0<\theta_b-\omega<\theta_g-\omega<\theta_b+\omega<\theta_g+\omega$.

Hence, with an observation of x_t within $(\theta_b+\omega,\theta_g+\omega]$, the firm learns with certainty that it is a good idiosyncratic productivity; conversely, an observation of x_t within $[\theta_b-\omega,\theta_g-\omega)$ indicates that it has bad idiosyncratic productivity. However, an x_t within $[\theta_g - \omega, \theta_b + \omega]$ reveals nothing, 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 our model considerably. Let θ^e represent the expected θ. Since it is θ^e instead of θ that affects firms' decisions, there are three groups of firms corresponding to the three values of θ^e: firms with $\theta^e = \theta_g$, firms with $\theta^e = \theta_b$, and firms with $\theta^e = \theta_u$, the prior mean of θ. We define "unsure firms" as those with $\theta^e = \theta_u$. We further assume that the unconditional probability of $\theta = \theta_g$ is φ, and let $p \equiv \frac{\theta_g-\theta_b}{2\omega}$ denote the probability of true idiosyncratic productivity being revealed every period. A firm enters the market as unsure; thereafter, every period it stays unsure with probability $1-p$, learns it is good with probability $p \cdot \varphi$ and learns it is 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$(0, 1, 0)$, and a transition matrix

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

If firms were to live forever, eventually all uncertainty would be resolved because the market would provide enough information to reveal each firm's true idiosyncratic productivity. The limiting probability distribution as a goes to ∞ is $(\varphi, 0, 1-\varphi)$.

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*, we have the densities of three groups of firms in a cohort of age a as

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

which implies an evolution of cross-section 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 idiosyncratic productivity, whether good or bad, grow as a cohort ages. Moreover, the two "learning curves" (depicting the evolution of densities of good firms and bad firms) are concave. This feature is defined as the decreasing property of marginal learning in Jovanovic (1982): the marginal learning effect decreases with firm age, which in our model is reflected by the decline in the marginal number of learners 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 firm distribution across ages and idiosyncratic productivity of an industry that features constant entry but no exit. In this industry, cohorts continuously enter in the same size and experience the same dynamics afterward, so that at any one time, different life-stages of different birth cohorts overlap, giving rise to the distribution in Figure 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.

2.2 The Recursive Competitive Equilibrium

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

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

Three assumptions characterize this partial industrial equilibrium: firm rationality, free entry and competitive pricing.

2.2.1 Firm Rationality

Firms are assumed to have rational expectations. They need to observe (F, D) to predict the sequence of prices from today onward. Therefore, the relevant state variables for a firm are its vintage, its belief about its true idiosyncratic productivity, and the aggregate state (F, D). We let $V(\theta^e, a; F, D)$ be the expected value, for a firm with belief θ^e and age a, of staying in operation for one more period and optimizing afterward, when the aggregate state is (F, D). Then V satisfies:

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

subject to

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

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

Since firms enter as unsure, firm rationality implies that entry occurs only if $V(\theta_u, 0; F, D) > 0$. Different from Caballero and Hammour (1994), a firm with belief θ^e and age a in our model exit if and only if $V(\theta^e, a; F, D) < 0$.[1]

2.2.2 Entry Cost and Entry Size

New firms can enter at any instant, as long as they bear an entry cost c. With $f(\theta_u, 0; F, D)$ as the size of the entering cohort with aggregate state (F, D), we assume that entry cost is a linear function of entry size:

$$c = c_0 + c_1 f(\theta_u, 0; F, D), \ c > 0 \text{ and } c_1 > 0, \quad (4)$$

Entry cost can be the cost of establishing a particular location, that of purchasing capital stock, or that of finding a qualified manager. With $c_1 > 0$, entry cost increases in entry size. This can arise from a limited amount of land available to build production sites, or from an upward-sloping supply curve for the industry's capital stock. Goolsbee (1998) provides evidence that supports

[1] The demand sequence in Caballero and Hammour (1994) is determinstic and *smooth*. The absense of sudden changes in demand justifies their simplififying firm exit. They assume that a firm exits as long as its *current* profit drops below zero, because, when demand changes are smooth, the exit of a forward-looking firm is similar to that of a myopic firm. Since this does not hold in our model with demand following a two-state markov chain, we model firm exit as when its expected value of staying drops below zero. Under our setup, the expected value of staying includes the value of waiting: when demand is low, firms' expected value of staying would include that by realizing the probability of demand recovery.

this assumption by showing that higher investment demand drives up the equipment prices. His findings suggest that, as more firms enter, capital becomes increasingly more expensive – this is captured in our model by $c_1 > 0$.

Under the free entry condition, new firms keep entering as long as the value of entry exceeds the cost of entry. As more firms enter, entry cost keep increasing until reaching the value of entry. At this point, entry stops, so that

$$V(\theta_u, 0; F, D) = c_0 + c_1 f(\theta_u, 0; F, D). \tag{5}$$

2.2.3 Competitive Pricing and Industry Demand

Firms are price takers. The equilibrium output price is determined endogenously by:

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

D is an exogenous demand parameter. The industry output, Q, is the sum of production over heterogeneous firms. The sequence of events implies that:[2]

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

where A is the leading technology. $f(\theta^e, a)'$ measures the number of operating firms with θ^e and a *after entry and exit*. $f(\theta^e, a)'$ belongs to F', the updated firm distribution. Since $F' = H(F, D)$, Q is a function of (F, D). High Q reduces P, according to (6). (7) implies that Q depends not only on the number of firms in operation, but also on their distribution. More firms yield higher output and reduces the price; the more the distribution is skewed toward younger vintages or better idiosyncratic productivity, the higher the output and the lower the price.

We assume an exogenous demand parameter, D, to capture changes in profitability arising from demand side. In reality, such fluctuations can arise from taste shocks on the industry's production goods or from productivity shocks of down-stream industries that demand the industry's output

[2] 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.

as their inputs. Apparently, D equals an industry's total revenue. With wage rate normalized as one, it also captures the industry's total profit. In our model, it is the exogenous fluctuations in D that introduces industrial cycles.

With these three conditions, we have the following:

> Definition: A *recursive competitive (partial) 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.

3 Cleansing and Scarring

This section motivates the cleansing and scarring effects by comparative statics. The industrial dynamics generated by demand fluctuations are difficult to characterize, since, as shown in the previous section, the firm distribution F enters the model as a state variable. Nevertheless, it is general true that, if shocks are sufficiently persistent, the effects of temporary changes in transitory shocks are similar to the effects induced by permanent shocks. Therefore, we engage comparative statics on the steady-state equilibrium to capture how demand affects the labor allocation, and to provide a more rigorous intuition for the scarring and cleansing effects described in the introduction. The next section will turn to a numerical analysis of the model's response to stochastic demand fluctuations and confirm that the results carry over.

3.1 The Steady State

A steady state is a recursive competitive equilibrium with time-invariant aggregate states: D is *and is perceived* as time-invariant: $D' = D$; F is also 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, it can be summarized by $\{f(0), \overline{a}_g, \overline{a}_u\}$, with $f(0)$ as the steady-state entry size, $\overline{a}_g$ as the maximum age for good firms, and $\overline{a}_u$ as the maximum age for unsure firms. Proposition 1 establishes that, for each D, a unique steady-state equilibrium exists.

> Proposition 1: With D constant over time, there exists a unique time-invariant $\{f(0), \overline{a}_g, \overline{a}_u\}$ that satisfies the conditions of firm rationality, free entry and competitive pricing.

Appendix 1 shows that, combining the exit conditions for unsure and good firms, we get:

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

We prove in Appendix 1 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. This suggests that demand does not affect the gap between the exit ages of good and unsure firms. They co-move across steady states with the same magnitude. With $\overline{a}_g - \overline{a}_u$ determined by (8) independently, the competitive pricing and the free entry condition jointly determine $f(0)$ and $\overline{a}_g$.

Figure 2 illustrates the steady-state firm distribution, entry margin, and the exit margins. Like Figure 1, there are two ways to interpret Figure 2. First, with the horizontal axis depicting the cohort age *across time,* it displays the steady-state life-cycle dynamics of a representative cohort. A cohort enters as unsure in a measure of $f0$. As it ages, bad firms learn and exit, so that the cohort size declines; good firms learn and stay, thus the density of good firms grows. At age $\overline{a_u}$, all unsure firms exit, because their vintage is too old to survive by remaining unsure; but good firms can stay. After $\overline{a_u}$, learning stops, the cohort contains good firms only, and its size remains constant. Good firms live until $\overline{a_g}$. The vintage after $\overline{a_g}$ is too old even for good firms to survive.

Second, with the horizontal axis depicting the cohort age *cross section,* Figure 2 displays the steady-state firm distribution across ages and idiosyncratic productivity at any one time. Firms of different ages coexist. Because older cohorts have lived longer and learned more, their sizes are smaller and their densities of good firms are higher. Cohorts older than $\overline{a_u}$ are of the same size and contain good firms only. No cohort is older than $\overline{a_g}$.

Moreover, despite its time-invariant structure, the industry experiences continuous entry and exit at a steady state. From a pure accounting point of view, resource reallocation takes place through three margins: the entry margin, the exit margins of good firms and unsure firms, and the learning margin. At the entry margin, new vintages enter; at the exit margins, old vintages leave. This introduces a force of creative destruction that replaces old vintages with new vintages. At the learning margin, bad firms leave, giving rise to a learning force that keeps good firms and drives out bad firms. Because of creative destruction, average labor productivity grows at technological pace γ. Because of learning, the proportion of good firms is higher for older cohorts. The two forces – learning and creative destruction – together drive the entry, the exit, and the related productivity dynamics.

3.2 Comparative Statics: Cleansing and Scarring

This subsection establishes that, across steady states, the model delivers the conventional cleansing effect, and an additional scarring effect. The two effects are formalized in Propositions 2 and 3.

> Proposition 2: In a steady-state equilibrium, the exit age for firms with a given belief is weakly increasing in the demand level

Detailed proof is provided in Appendix 1. Put intuitively, in a high-demand steady state, both unsure firms and good firms live longer. Lower demand drives down price, so that some firms that are viable when demand is high become not viable when demand is low.

If this story carries over when D fluctuates stochastically, then our model delivers a conventional "cleansing" effect, in which average firm age falls so that the average vintage becomes younger and more productive. However, once learning is allowed, we also need to take into account the allocation of labor across the other dimension – idiosyncratic productivity. With only two true idiosyncratic productivity, good and bad, this allocation can be summarized by the fraction of labor at good firms. A higher fraction suggests a more productive allocation of labor. The next proposition establishes how demand affects this ratio at a steady state.

> Proposition 3: In a steady state equilibrium, the fraction of labor at good firms is weakly increasing in demand.

The steady-state fraction of labor at good firms (including both known and yet unknown) is:

$$l_g^{ss} = 1 - \frac{(1-\varphi)}{\frac{p\varphi(\overline{a_u}+1)}{1-(1-p)^{\overline{a_u}+1}} + (1-\varphi) + p\varphi\left(\overline{a_g} - \overline{a_u}\right)}. \tag{9}$$

Because $(\overline{a_g} - \overline{a_u})$ is independent of D according to (8), $\frac{d\left(l_g^{ss}\right)}{d(D)} = \frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})}\frac{d(\overline{a_u})}{d(D)}$. We prove in Appendix 1 that $\frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})} \geq 0$, which, together with $\frac{d(\overline{a_u})}{d(D)} \geq 0$ (Proposition 2), implies $\frac{d\left(l_g^{ss}\right)}{d(D)} \geq 0$. Put intuitively, the impact of demand on the fraction of labor at good firms ($\frac{d\left(l_g^{ss}\right)}{d(D)}$) comes from its impact on unsure firms' exit age ($\frac{d(\overline{a_u})}{d(D)}$). To further understand this result, consider Figure 3.

Figure 3 displays the firm distribution across vintages and idiosyncratic productivity at a high-demand steady state and that at a low-demand steady state, with the entry size normalized as one. The two exit margins in Figure 3 shift to the left corresponding to a lower demand, due to the cleansing effect. This clears out old firms that could be either good or unsure. However, the leftward shift of the *unsure exit margin* also reduces the number of *older good firms*. The latter effect, shown as the shaded area in Figure 3, is the scarring effect of recessions.

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

Hence, the industry suffers from uncertainty: firms that exit at age $\overline{a_u}$ include some that are truly good and should have stayed. How many potentially good firms exit at $\overline{a_u}$ depends on the size of the unsure exit margin, which, in turn, depends on the value of $\overline{a_u}$. A drop in demand truncates learning by reducing $\overline{a}_u$; consequently, more potentially good firms exit, fewer good firms become old, and the proportion of resources at good firms declines.

To summarize from Propositions 2 and 3, a low-demand steady state features a better average vintage, yet a lower proportion of good firms. If the comparative static results carry over when demand fluctuates stochastically, then recessions will have both a conventional cleansing effect that raises average vintage, and a scarring effect that lowers average idiosyncratic productivity. As suggested by $\frac{d\left(l_g^{ss}\right)}{d(D)} = \frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})}\frac{d(\overline{a_u})}{d(D)}$, the two effects are directly related to each other: it is the cleansing effect ($\frac{d(\overline{a_u})}{d(D)}$) that truncates learning and prevents more firms from realizing their potential ($\frac{d\left(l_g^{ss}\right)}{d(\overline{a_u})}$).

When we move beyond steady states to allow for cyclical fluctuations, the intuition behind "cleansing and scarring" still carries over. Again, consider Figure 3. Both exit margins shift as soon as demand drops, so that the cleansing effect takes place immediately. However, the scarring effect takes place both instantaneously and gradually. At the onset of a recession, the composition of idiosyncratic productivity worsens immediately, because the shift of the exit margins clear out old cohorts that contain more good firms – we call this "instantaneous scarring". This effect also comes from learning and therefore is not captured by the conventional cleansing models. Moreover, another "lasting scarring effect" will follow. Notice that, when demand drops, despite the shift in exit margins, the group of firms already in the shaded area in Figure 3 will *stay* by knowing their true idiosyncratic productivity to be good. They leave gradually as the recession persists, creating a "lasting scarring effect": the reduced $\overline{a_u}$ allows fewer good firms to survive past $\overline{a_u}$, so that the shaded area would eventually be left blank. Hence, as described in the introduction, the loss of good firms leaves "scars" when a recession arrives, and the "scars" deepen as the recession persists. Instantaneous and lasting scarring effects together capture the impact of recessions on the composition of idiosyncratic productivity.

Moreover, while the cleansing and scarring effects stem from shift of the exit margins, they are also impacted by changes at the entry margin. During recessions, entry size declines due to lower expected value of entry. *Less* entry lows average vintage on the one hand, because entrants carry the leading technology, and raises average idiosyncratic productivity on the other hand, because entrants contain the highest proportion of bad firms. Therefore, the shift of exit margins create cleansing and scarring, but the changes in entry size dampen both.

3.3 More Discussions

Three modeling assumptions are to be discussed before we move to the quantitative evaluation of cleansing and scarring.

3.3.1 Entry Cost and Entry Size

In our model, entry cost is assumed to increase in entry size: $c_1 > 0$. What if entry cost is independent of entry size ($c_1 = 0$) instead? Then the three conditions that jointly determine $\{f(0), \overline{a}_g, \overline{a}_u\}$ become fully recursive: $\overline{a}_g - \overline{a}_u$ is given by (8) independently; With $\overline{a}_u = \overline{a}_g - (\overline{a}_g - \overline{a}_u)$, the free entry condition determines $\overline{a}_g$; then the competitive pricing condition, where D enters, determines $f(0)$.[3] Hence, demand impacts entry size only. This extreme case is described as "full insulation" in Caballero and Hammour (1994): demand fluctuations are fully accommodated at the entry margin. Intuitively, when *fast* entry is costless ($c_1 = 0$), entry size adjusts proportionally to changes in demand to such an extent that price remains unchanged. As a result, exit margins do not respond; and there would be neither cleansing nor scarring effects.

On the contrary, with $c_1 > 0$, $f(0)$ and $\overline{a}_g$ are jointly determined by the free entry condition and the competitive pricing condition, so that demand impacts both. Therefore, some of the demand variations are accommodated at the entry margin, while the rest are taken as shifts of the exit margins. As a result, both entry and exit fluctuate over the cycle. Apparently, data are consistent with $c_1 > 0$. For example, according to the Business Employment Dynamics (BED), exit and entry display similar volatility:.from the third quarter of 1992 to the second quarter of 2007, the ratio of the standard deviation of the quarterly exit rate over that of the entry rate is 1.00 for the U.S. manufacturing sector, and 1.02 for the entire private sector.[4]

3.3.2 Productivity Composition of Entrants

Moreover, we assume that demand cannot impact the proportion of good firms among entrants: it is fixed at φ. But the literature has pointed out that the a producers' quality can be both an inspection and an experience good (e.g., Pries and Rogerson, 2005). While some information is acquired as firms operate, some other is acquired immediately at the entry stage. If we allow for inspection of θ prior to entry, recessions can change the productivity composition of entrants. If, during recessions, only more promising firms can afford to enter, then the average idiosyncratic productivity of entrants would be countercyclical, and would play against the scarring effect.

[3] Details are provided in Appendix 1 when proving Proposition 1.

[4] The Bureau of Labor Statistics provides the data. The series we examined are the seasonally adjusted non-weighted entry and exit rates. There standard deviations are calculated using their detrended values with the HP filter.

However, it is hard to tell whether recessions affect entrants' productivity composition positively or negatively. Firms enter, expecting their entry value high enough to cover the entry cost. In the current model, the expected entry value depends solely on aggregate states (F, D). With inspection of θ prior to entry, it would depend additionally on θ^e. In that case, lower D implies a higher θ^e threshold, *holding entry cost constant*. But, if lower D reduces entry cost, the θ^e threshold can become lower instead. Intuitively, during recessions, while firms need to feel more optimistic about themselves to enter due to the poor aspect of the industry, it also becomes cheaper to rent land or easier to find a qualified manager. Therefore, the impact of recessions on entrants' productivity can be ambiguous.

Jensen et al.(2001) provides related evidence. They list the average labor productivity of U.S. manufacturing entrants every five years from 1963 to1992 (p. 327, Table 1). According to their estimates: entrants' average productivity goes up in the recovery phase of 1972, but declines sharply in 1977 as in the middle of an expansion; it declines again during the 1982 recession, and spikes up in 1987, which was, again, in the middle of an expansion.[5] While their number of observations are limited and the NBER business cycle dates can have some potential inaccuracy, this evidence suggests, by no means, that recessions *improve* entrants' productivity composition. Similar evidence has been provided by Davis et. al. (1996), who show that job created during recessions tend to live shorter, and by Bowlus (1995), who estimates that job created during recessions are from lower part of the wage distribution. Therefore, it is likely that recessions worsen entrants' productivity composition instead. In that case, productivity dynamics at the entry margin would complement the scarring effect.

3.3.3 More Complicated Learning

We are able to motivate the scarring effect analytically, because the all-or-nothing learning simplifies the analysis to a great extent. With uniformly distributed noises, the expected idiosyncratic productivity, θ^e, takes on only two values: θ_u and θ_g. Would the scarring effect carry over with a more complicated learning? Suppose that the noises covering the true idiosyncratic productivity,

[5] This can be seen in levels as well as in detrened values. Jensen et al. (2001) also provide such estimates after controlling for time effect, industry effect, and input effect. Apparently, those were inapproriate for examining the impact of cycles on entrants' productivity composition, because those controlled effects can be the channels for such impact to take place. Instead, we use the trend growth of those controlled estiamtes to calibrate the technological pace (γ) in the next section.

is distributed normally with mean zero and variance σ: $\omega \sim N(0,\sigma)$. Then, every period a good firm would receive a draw from $N(\theta_g,\sigma)$ while a bad firm receive one from $N(\theta_b,\sigma)$. In this case, almost every x would change the perceived likelihood of a firm being actually good or bad, so that θ^e can take any value between θ_b and θ_g. Then a firm's belief would depend on the entire sequence of x's it receives in the past. To make the matter more complicated, this x sequence needs to have the property that it has never driven the firm pessimistic enough to exit. Accordingly, examining the scarring effect requires keeping track of the distribution of such sequences of x's, and would not be analytically feasible.

Here we take a different approach. Again, we use the steady-state distribution of labor to motivate the scarring effect. In a vintage world with any type of learning, the industrial proportion of good firms at a steady state of demand D, denoted as $l_g(D)$, is

$$l_g(D) = \frac{\varphi + \sum_{a=1}^{\overline{a}(D)} \varphi_a(D) h_a(D)}{1 + \sum_{a=1}^{\overline{a}(D)} h_a(D)}$$

The entrants' average idiosyncratic productivity entrants (φ) is exogenous. The the proportion of good firms of an age-a cohort (φ_a), the maximum firm age ($\overline{a}$), and the size of an age-a cohort *relative to the entry size* (h_a) are affected by D. Lower demand allows firms to live shorter: $\overline{a}'(D) > 0$; meanwhile, more firms exit as a cohort ages, so that any incumbent cohort would be of smaller sizes: $h_a'(D) > 0$. The impact of D on φ_a, however, is negative. Lower demand drives stronger selection, causing a higher proportion of good firms for each incumbent cohort: $\varphi_a'(D) < 0$. Apparently, demand affects θ through its impact on θ_a, on h_a, and on $\overline{a}$. On the one hand, lower D drives down $\overline{a}$ and h_a's, so that older cohorts with higher proportion of good firms weight *less* in determining l_g; this tends to lower l_g. On the other hand, lower D raises φ_a (for any a) so that, within each incumbent cohort, proportion of good firms are higher; this tends to raise l_g. In summary, changes in D cause two competing effects on l_g through three channels.

Does the scarring effect established through all-or-nothing learning capture all these three channels? It *fully* captures the change in $\overline{a}$: as shown in Figure 3, firms's maximum life span shortens at a low-demand steady state. However, it only *partially* captures the changes in h_a and φ_a. Again, as shown in Figure 3, only *some* incumbent cohorts become smaller at the low-demand steady state

– those that contain good firms only. Meanwhile, the cohort proportions of good firms (φ_a) are higher only for some cohorts – those that contain some unsure firms at the high-demand steady state but have good firms only at the low-demand steady state. This is because, with all-or-nothing learning, θ^e takes on only two values. Now the question becomes, with continuous values for θ^e, would further increases in φ_a make the scarring effect disappear, or, would further decreases in h_a amplify the scarring effect?

An important remark should be made. Lower demand causes the exit of potentially good and bad firms both. But it is the exit of potentially bad firms that causes increases in φ_a. In that sense, increases in φ_a adds to the conventional cleansing effect. On the contrary, h_a declines mainly because of the exit of potentially good firms. It not only makes the cohort smaller in present times, but also reduces it size in the future because, over time, good firms would have the chances of growing more optimistic and stayed, while bad firms would learn they are bad and choose to leave on their own anyway – this is the spirit of the scarring effect.

We highlight our point again by comparing three worlds: one with vintage only, one with learning only, and one with vintage-specific learning as in our current model. Apparently, there is only cleansing effect in a world with vintage only, as shown by Caballero and Hammour (1994). With learning only, $\overline{a}$ converges to infinity with some good firms staying for ever. In that case, recessions clear out potentially bad firms, contributing to the conventional cleansing effect, and potentially good firms, contributing to the scarring effect. With vintage and learning, changes in $\overline{a}$ create *additional* cleansing effect by driving out oldest vintages, as well as making the scarring effect stronger – decrease in $\overline{a}$ further lower the weight of older cohorts with higher proportion of good firms in determining l_g for the entire industry. Therefore, cleansing and scarring are present as long as learning takes place, with or without vintage.

In summary, the all-or-nothing learning provides a convenient framework to motivate the scarring effect analytically. With more complicated learning, both the cleansing and the scarring may become stronger. The question again becomes, cleansing and scarring, which effect dominates?

4 Quantitative Implications with Stochastic Demand Fluctuations

This section analyze numerically a stochastic version of our model to show that the scarring effect is likely to dominate the cleansing effect. Throughout this section, demand follows a two-state Markov process with values $[D_h, D_l]$ and transition probability μ; accordingly, firms expect the current demand to persist for another period with probability μ, and to change with probability $1 - \mu$.

4.1 Calibration

We calibrate our model using statistics on plant-level productivity differential, plant entry rate, and plant exit rate from the U.S. manufacturing sector. Although a production unit is called a firm in our model, we apply plant-level measure for calibration because, in reality, firms often retool by closing old plants and opening new plants, so that plant provides a better approximation for technology adoption (Campbell, 1998). Since we apply the manufacturing productivity measure to calibrate the productivity differential, we also restrict our attention to manufacturing entry and exit at the purpose of keeping the calibration consistent.

Table 1 summarizes the calibrated parameter values. With a period representing a quarter, β is set to equal 0.99. μ is chosen to equal 0.95 so that demand switches between a high level and a low level with a constant probability 0.05 per quarter. Bad firms' idiosyncratic productivity, θ_b, is normalized as one. c_1 determines the elasticity of entry cost with respected to entry size. We choose its value according to Goolsbee (1998), who estimates that a 10% increase in demand for capital equipment raises equipment price by 7.284% (p. 143, Table VII). We set $c_1 = 0.7284$.

The rest of the parameters are calibrated as follows. First, we use the cohort cumulative exit rates to calibrate the learning pace (p) and the probability of being a good firm (φ). Second, the technological pace (γ) and the idiosyncratic productivity differential (θ_g) are calibrated to the observed entrants' productivity growth and the cohort productivity differentials. These parameter values jointly determine the relative strengths of the forces of learning and creative destruction. Then, it is the changes in demand together with fixed entry cost that generate cyclical entry and exit, which, in turn, determine the magnitudes of the cleansing and scaring effects. Therefore, we use the observed moments in entry and exit rates to calibrate the values of D_h, D_l, and c_0.

4.1.1 Learning and the cumulative cohort exit rates.

Dunne et al. (1989) report that 57.5% of the 1972-entering cohort in the U.S. manufacturing sector had exited by 1977, and 78.2% of it had exited by 1982. This imposes two conditions on p and φ:

$$\begin{aligned}\varphi + (1-\varphi)(1-p)^{19} &= 1 - 0.575;\\ \varphi + (1-\varphi)(1-p)^{39} &= 1 - 0.782.\end{aligned}$$

They jointly determine that $p = 0.0538$ and $\varphi = 0.1157$.

4.1.2 Technological pace and entrants' productivity growth.

In our model, only entrants adopt the most advanced technology, which allows us to calibrate γ using growth in entrants' productivity. Since such pace can vary over time in reality, we take the historical average of the longest period possible. Jensen et al. (2001) estimate that, after controlling for industry and time effects, the U.S. manufacturing entrants' productivity grew by 46.8% from 1963 to 1992. The 46.8% increase in entrants' productivity over a 29-year horizon suggests a quarterly technological pace of 0.004.

As a comparison, Basu et al. (2001) estimate annual TFP growth by controlling for employment growth, factor utilization, capital adjustment costs, quality of inputs and deviations from constant returns and perfect competition. Their estimate from 1965 to 1990 implies a quarterly rate of 0.0049 for durable manufacturing and one of 0.0054 for non-durable manufacturing, slightly above our calibrated γ. However, not all technological improvements are embodied in entrants, some are achieved through the accumulation of experiences or the within-plant retooling, for example. We control for these effect in calibrating θ_g.

4.1.3 Idiosyncratic productivity differential and productivity differential among birth cohorts.

We calibrate θ_g to cohort productivity differential *in a given year*. In our model, vintage effect makes younger cohort more productive, while the learning effect drives up older cohorts' productivity with higher proportion of good firms. In reality, however, older cohorts' being more productive can be additionally driven by managers' accumulating experiences, workers' learning by doing, the achieving of economies of scale, or the within-plant retooling. Davis and Haltiwanger (1992) define such additional effects as "active learning", and the learning of unobservable predetermined features as "passive learning". Therefore, a careful calibration of θ_g requires matching the observed productivity differential by controlling for the vintage effect, present in our model, as well as the active learning effect, missing from our model.

We control for active learning using surviving plants' productivity growth since they enter, considering that active learning displays diminishing returns (Jovanovic and Nyarko, 1996). Jensen et. al. (2001) track the plants that entered in 1967 and survived through 1992, and estimate that, after controlling for industry effect, their productivity grew by 14.8% over this period of 25 years.[6] They further report that, while these plants are 15.3% more productive than the new entrants in 1992, their productivity back in 1967 were 60.4% lower than that of 1992 entrants.

In summary, the 1967-entering cohort is observed as 15.3% *more* productive in 1992 than the 1992-entering cohort, although their vintage is 60.4% *less* productive than the 1992 vintage. Controlling for the difference of 14.8% driven by active learning leaves a difference of 60.9% to be taken by passive learning. This imposes the following restriction on the value of θ_g, according to all-or-nothing learning.

$$\frac{\left(\varphi\theta_g + (1-\varphi)(1-p)^{100}\theta_b\right)}{\left(\varphi\theta_g + (1-\varphi)\theta_b\right)\left(\varphi + (1-\varphi)(1-p)^{100}\right)} = 1.609$$

This suggests $\theta_g = 1.75$, a idiosyncratic productivity differential of 75%. As a comparison, Bartelsman and Doms (2000) document the productivity of U.S. manufacturing plants in the ninth docile to be 2.75 times of that of plants in the second docile. Based on their results, Davis and Haltiwanger (1999) assumes a high-to-low productivity ratio of 2.4. Our calibrated θ_g is lower, because θ_g is supposed to capture only the passive learning effect, as one of the many factors driving plant productivity differences.

[6] Jensen et al. (2001) also estimate active learning by further controlling for wage and capital intensity. However, they report "Neither input variable is statistically significant, however, and the capital variable is surprisingly negative". By contrast, they mention in Footenote 17 that "the industry-wide variables remain jointly significant at the 99% level". Accordingly, we adopt their estimates controlling for industry effect only.

4.1.4 Calibrating D_h, D_l, and c_0 with average plant age and cyclical entry and exit rates.

We calibrate D_h, D_l, and c_0 by approximating firm distributions at a high-demand equilibrium and a low-demand equilibrium. According to our numerical simulations, the dynamic system eventually settles down with constant entry and exit along any sample path with unchanging demand. The firm distribution at these stable points are similar to those at the steady states, which allows us to use steady-state conditions for approximation.

As established in Section 3, a steady state is characterized by good firms' maximum age, unsure firms' maximum age, and the entry size. We let $\overline{a_g}h$, $\overline{a_u}h$ and fh to represent those when demand is high, and $\overline{a_g}l$, $\overline{a_u}l$, and fl as those when demand is low. Calibrations for p, φ, γ, θ_g determine the value of $\overline{a_g}h - \overline{a_u}h$ and $\overline{a_g}l - \overline{a_u}l$ using (8). Hence, we need to determine only the values of $\overline{a_g}h$, $\overline{a_g}l$, fh, and fl.

As the first step, $\overline{a_g}h$ is chosen to pin down the industry's age structure, and to match the mean entry and exit rates. With $\overline{a_u}h = \overline{a_g}h - (\overline{a_g}h - \overline{a_u}h)$, $\overline{a_g}h$ determines the average firm age and the relative size of the entry and exit margins to the entire industry. Faberman (2003) reports the age of U.S. manufacturing plants to average around 58 quarters using data compiled from the unemployment insurance records at the Bureau of Labor Statistics.[7] Moreover, the quarterly entry rate averages at 3.11% and the quarterly exit rate averages at 3.45% in the manufacturing sector from 1992 to 2007, according to BED. We find that $\overline{a_g}h = 165$ can closely match these conditions.

Then, we use statistics on cyclical entry and exit to calibrate $\overline{a_g}l$, fh, and fl. Figure 5 plots, in the top two panels, the BED quarterly entry and exit rates for the U.S. manufacturing sector from 1992 to 2007. As it shows, both series (especially the entry rate) display a declining trend. To avoid over estimating the magnitude of their cyclicality, we detrend the two series using the HP filter, and present their detrended variations in the bottom two panels of Figure 5. As Figure 5 shows, the entry rate fluctuates between 0.29% below and 0.37% above its trend value, and the exit rate varies between 0.36% below and 0.34% above its trend value. These give us the following restrictions.

First, our model must match the peak in exit rate and the trough in entry rate. They occur

[7] Note that data used by Faberman (2003) covers only five states: Colorado, Michigan, North Carolina, Ohio, and Pennsylvania.

when a negative demand shock hits a high-demand equilibrium. At this moment, the exit rate rises to 3.79% (3.45% + 0.34%), and the entry drops to 2.82% (3.11% − 0.29%). On the contrary, the trough in exit rate and the peak in entry rate take place when a positive demand shock hits a low-demand equilibrium. At this point, the exit rate drops to 3.09% (3.45% − 0.36%) and the entry rate rises to 3.48% (3.11% + 0.37%).

To generate the best match to these statistics, the following transitory dynamics need to be incorporated into the calibration. Our numerical simulations show that, when demand drops, the exit age shifts to an age *younger than* $\overline{a_g}l$, while the entry size drops to a level lower than fl. This is because, some old good firms stay by knowing their true idiosyncratic productivity (shown as those in the shaded area in Figure 3). Their operation keeps the output high but the price low, so that *more* old firms exit and *less* young firms enter. Over time, the exit age extends to $\overline{a_g}l$, and the entry size recovers to fl. In contrast, when demand rises, the exit margin can extend by only one quarter each time;.but the entry size rises *above* fh, because the *past* low demand leaves less firms in operation, which keeps output low but the price high and causes more new firms to enter. As high demand persists, exit margins and the entry size converge to $\overline{a_g}h$ and fh.

Using a search algorithm that incorporates these transitory dynamics (details are presented in Appendix 2), we find that, at the onset of a recession, the exit ages shift by four quarters and the entry size drops by 6.91%. As the recession persists, the exit margins extend by three quarters and the entry size recovers to 3.66% below its high-demand level. In other words, $\overline{a_g}l = 164$ and $\frac{fh}{fl} = 1.0366$.

With $\overline{a_g}h$ and $\overline{a_g}l$, we calculate the high-demand entry value (Vh) to equal 2.0803 and the low-demand entry value (Vl) to be 2.0124. By the free-entry condition, this suggests that the equilibrium entry cost is equivalent to the operation cost of about two quarters. With $fh = \frac{vh-c_0}{c_1}$ and $fl = \frac{Vl-c_0}{c_1}$, $\frac{fh}{fl} = 1.0366$ gives $c_0 = 0.1587$. We then use $c_1 = 0.7283$ (chosen according to Goolsbee, 1997) to calculate the entry sizes, and get that $D_h = 108.7294$ and $D_l = 103.9819$.

With demand equal to the total revenue, the calibration suggests a 4.57% difference in total revenue between a high-demand equilibrium and a low-demand equilibrium. To check if this is plausible, we detrend the 1992-2007 quarterly series of total value of manufacturing shipments using the HP filter, and find that it varies by about 6% around its trend value.[8] While it is difficult

[8] This data is provided by the Census of Bureau at http://www.census.gov/indicator/www/m3/hist/naicshist.htm.

to compare the difference of two discrete values with the standard deviation of continuous value, this does imply that our calibrated revenue differential is smaller than what data suggests. Subsection 4.3 provides further discussions.

4.2 Response to a Negative Demand Shock

With all of the parameter values assigned, we approximate firms' value functions. The key computational task is to map F, the firm distribution across ages and idiosyncratic productivity, given demand level D, into a set of value functions $V(\theta^e, a; F, D)$. Unfortunately, F is a high-dimensional object and it is well-known that the numerical solution of dynamic programming problems becomes increasingly difficult as the size of the state space increases. We follow Krusell and Smith (1998) by, first, shrinking the state space into a limited set of variables, and then showing that these variables' laws of motion can approximate the equilibrium behavior of firms in the simulated time series. Details are presented in Appendix 3. The approximated value functions and the corresponding decision rules enable us to investigate the dynamics of our model's key variables along any particular path of demand realizations, and study the model's quantitative implications.

4.2.1 Scarring and Cleansing

To assess the effect of a negative demand shock, we start with a random firm distribution and use our approximated value functions to simulate our model's response to a *particular* realization of demand sequence — demand stays at D_h until the key variables converge, then drops to D_l, and persists afterward.

Figure 6 illustrates the simulated dynamics of exit and entry rates, with the quarter labeled 0 denoting the onset of a recession. The exit rate jumps up, declines afterward, and converges above its original value. The entry rate drops initially and rises gradually. Hence, a negative demand shock clears out some firms that would have kept their operation if demand remained high. The cleansing effect motivated by the steady-state analysis carries over.

According to the comparative static exercises, recessions bring an additional scarring effect by worsening the composition of idiosyncratic productivity. Figure 7 presents the dynamics of the fraction of labor at good firms in response to a negative demand shock. As it shows, this measure

The series we examined are seasonally adjusted.

drops at the onset of a recession, consistent with the instantaneous scarring effect; it converges to a lower value while the recession persists, as suggested by the lasting scarring effect. Therefore, the scarring effect motivated by the steady-state analysis also carries over.

Moreover, the transitory dynamics described in Subsection 4.1.4, the initial overshift of the exit margins, the initial overdrop of the entry size, and their recovery afterward, are reflected in Figure 6 as the gradual increase in the entry rate, and as the sudden drop in the exit rate right after the onset of a negative demand shock. Such transitory dynamics also drive l_g to be hump-shaped in Figure 7. With the exit margins extending older and with new, unsure firms entering, more good firms learn their true idiosyncratic productivity, so that l_g rises. This can be called as a temporary "plastic-surgery" effect that partially erases the instantaneous scar. l_g declines again as the lasting scarring effect takes place, and converges to its low-demand equilibrium level eventually.

To summarize, despite some transitory dynamics, Figure 6 and Figure 7 suggest that both the conventional cleansing effect established in Proposition 2, and the scarring effect established in Proposition 3, carry over with an unexpected persistent negative demand shock.

4.2.2 Implications for Productivity

Next, we 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 $\frac{A\theta}{(1+\gamma)^a}$, average productivity is affected by A *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, demand shocks add fluctuations around this trend by affecting the labor distribution across a and θ.

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

$$prod = \frac{\sum\limits_{f}\left(\frac{\theta^e}{(1+\gamma)^a}\right)f(\theta^e,a)}{\sum\limits_{f}f(\theta^e,a)}, \quad vin = \frac{\sum\limits_{f}\left(\frac{1}{(1+\gamma)^a}\right)f(\theta^e,a)}{\sum\limits_{f}f(\theta^e,a)}.$$

Apparently, *prod* is affected by both the cleansing and scarring effects, while *vin* is driven by the cleansing effect alone.

Figure 8 traces the percentage changes in *prod* and in *vin* in response to a negative demand shock, with *their initial levels* normalized *as* one. At the onset of the shock, the cleansing effect alone *raises* the average productivity to 1.0012 while, with both the cleansing and the instantaneous scarring effects, average productivity drops to 0.9995. Afterward, *prod* rises with the temporary "plastic-surgery" effect, and declines again while the lasting scarring takes place. Eventually, *vin* converges to 1.0013, implying a 0.13% growth in average productivity by the cleansing effect alone. But, *prod* converges to 0.9994, suggesting a 0.06% decline in average productivity under both scarring and cleansing effects.

Therefore, a plausible calibration of our model suggests that scarring effect can dominate the cleansing effect, by lowering the average manufacturing productivity with a magnitude of up to 0.19%.

4.3 The average plant age, the wage rate, and the revenue differential

Before we conclude, it is important to discuss the three elements of the calibration. First, a vintage can live up to 165 quarters (41.25 years) in our calibrated model. This result does not necessarily imply that some firms never retool over 40 years. In reality, some of the retooling takes place through entry and exit; some other occurs within plant. The unobservable within-plant retooling can drive part of the active learning, which is controlled for already in calibrating θ_g. Therefore, while our model cannot account for all the technology adoption and learning going on in reality, it aims to examine the part that is associated specifically with entry and exit.

Second, wage rate is fixed at one in our model. By contrast, labor supply is not perfectly elastic in reality, so that wage would decline when demand drops. However, our goal is not to identify how much price or wage change, but to evaluate cleansing and scarring as the consequences of entry and exit. To serve this goal, the calibration exercise aims to replicate the observed entry and exit by generating the related profit margin. With wage fixed at one, revenue represents this margin directly; with lower wage, revenue would have to drop by more to give the same margin. Therefore, allowing wage to vary would not affect any of the results, as long as entry and exit are calibrated to the data.

Third, our calibration suggests that, in the U.S. manufacturing sector from 1992 to 2007, entry and exit dynamics are associated with a revenue differential of 4.57%. But, as pointed out in Subsection 4.1.4, this number may have underestimated the actual revenue differential. The difference can be explained by an upward-slopping labor supply, absent from our model but present in reality, so that total revenue changes by more to generate the observed entry and exit.

5 Conclusion

How do recessions affect resource allocation? This paper suggests that firm learning has important consequences for this question. We posit that, recessions create, in addition to the cleansing effect, a scarring effect by interrupting firms' learning of their idiosyncratic productivity. The scarring effect is evaluated quantitatively using statistics on entry, exit, and productivity differentials. We find that it can dominate the cleansing effect, and can give rise to lower average productivity during recessions.

Firm size can be added to this model, by allowing firms with better vintage or higher expected idiosyncratic productivity to hire more workers. Extensions into general equilibrium can also be made, by modeling demand shocks as from taste shocks or from down-stream industry's productivity shocks, and by incorporating labor supply. Can a model with vintage and learning account for the observed firm dynamics with respect to firm age and size, that have been explored theoretically by Cooley and Quadrini (2001) using financial frictions?[9] Can labor-supply elasticities propagate taste shocks or productivity shocks of plausible sizes into the observed cyclicality in entry and exit? We leave such questions for future research.

[9] According to Cooley and Quadrini (2001), firm dynamics (growth, volatility of growth, job flows) display a "size independence" and an "age independence". That is, they are negatively correlated with firm size holding age constant, and negatively correlated with firm age holding size constant. Our model has the potential of accounting for such patterns. Unsure firms must be smaller among those of the same age (vintage), and younger (better vintage) for those of the same size – it is unsure firms that display stronger dynamics driven by learning.

6 Appendix

6.1 Appendix 1

6.1.1 Proof of Proposition 1

Proof. According to the condition of competitive pricing and the definition of a steady state,

$$D = P_t A_t \cdot \left\{ \sum_{a=0}^{\overline{a_u}} \left[\theta_u f(\theta_u, a)(1+\gamma)^{-a} \right] + \sum_{a=0}^{\overline{a_g}} \left[\theta_g f(\theta_g, a)(1+\gamma)^{-a} \right] \right\} \tag{A1}$$

with D as the time-invariant demand, $f(\theta^e, a)$ the time-invariant number of firms with (θ^e, a), and $\overline{a_g}$, $\overline{a}_u$ the time-invariant exit ages for good and unsure firms. It suggests that $P_t A_t$ must also be time-invariant. We let $P_t A_t = PA$.

$f(0)$ represents the time-invariant entry size at the steady state. Let $V(\theta^e, a)$ be the time-invariant expected value of staying of a firm with belief θ^e and age a. The exit condition for good firms, $V(\theta_g, \overline{a_g}) = 0$, suggests:

$$\theta_g PA (1+\gamma)^{-\overline{a_g}} - 1 = 0. \tag{A2}$$

With $f(\theta^e, a)$ given by all-or-nothing learning, (A1) and (A2) together with the steady-state structure as shown in Figure 2, imply

$$f(0)\frac{(1+\gamma)^{\overline{a}_g}}{\theta_g} \left[\begin{array}{c} (\theta_u - \varphi\theta_g) \sum_{a=0}^{\overline{a}_u} \left(\frac{1-p}{1+\gamma}\right)^a + \varphi\theta_g \sum_{a=0}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a + \\ \varphi\theta_g (1-p)^{\overline{a}_u+1} \sum_{a=\overline{a}_u+1}^{\overline{a}_g} \left(\frac{1}{1+\gamma}\right)^a \end{array} \right] = D. \tag{A3}$$

The free entry condition, $V(\theta_u, 0) = c_0 + c_1 f(0)$, suggests

$$\sum_{a=0}^{\overline{a}_u} \beta^a \left[\frac{PA\theta_u}{(1+\gamma)^a} - 1 \right] \lambda(\theta_u, a) + \sum_{a=0}^{\overline{a}_g} \beta^a \left[\frac{PA\theta_g}{(1+\gamma)^a} - 1 \right] \lambda(\theta_g, a) = V(\theta_u, 0) = c_0 + c_1 f(0). \tag{A4}$$

$\lambda(\theta_u, a)$ and $\lambda(\theta_g, a)$ are the probabilities of staying in operation at age a as an unsure firm and a good firm, and are given by the all-or-nothing learning.

The exit condition for unsure firms, $V(\theta_u, \overline{a_u}) = 0$, gives:

$$\theta_u PA (1+\gamma)^{-\overline{a_u}} - 1 + \beta p \varphi \sum_{a=\overline{a_u}+1}^{\overline{a_g}} \beta^{a-\overline{a_u}-1} \left[\theta_g PA (1+\gamma)^{-a} - 1 \right] = 0 \tag{A5}$$

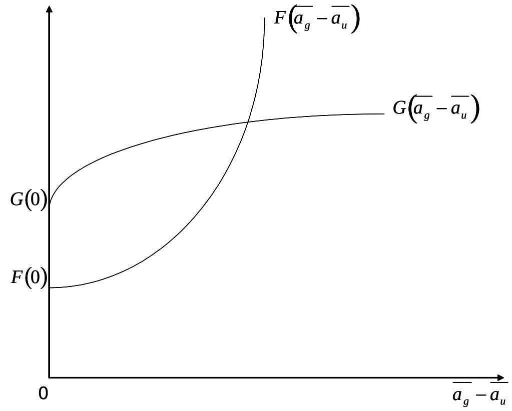
Combining (A2) and (A5) gives

$$\left(\frac{\theta_u}{\theta_g} + \frac{p\varphi\beta}{1+\gamma-\beta} \right) (1+\gamma)^{\overline{a}_g - \overline{a}_u} = 1 + \frac{p\varphi\beta}{1-\beta} - \frac{p\varphi\beta\gamma}{(1-\beta)(1+\gamma-\beta)} \beta^{\overline{a}_g - \overline{a}_u}. \tag{A6}$$

(A6) solves $\overline{a}_g - \overline{a}_u$. To establish the existence and the uniqueness of the solution, let $F(\overline{a}_g - \overline{a}_u)$ represents the left-hand side, and $G(\overline{a}_g - \overline{a}_u)$ be the right-hand side of (A6). It can be shown that $G' > 0$ but $G'' < 0$, $F' > 0$ and $F'' > 0$; moreover,

$$F(0) < G(0) \text{ as long as } \frac{\theta_u}{\theta_g} < 1$$

Since $\theta_u < \theta_g$ holds by definition, F and G must cross *once* at a positive value of $\overline{a}_g - \overline{a}_u$, as shown in the following figure



Hence, (15) determines a unique value for $\overline{a}_g - \overline{a}_u$. With $\overline{a}_u = \overline{a}_g - (\overline{a}_g - \overline{a}_u)$ and A(2), (A3) and (A4) jointly determine $f(0)$ and $\overline{a}_g$ when $c_1 = 0$.

Notice that with entry cost independent of entry size, $c_1 = 0$. (A6), (A3) and (A4) become recursive. (A6) determines $\overline{a}_g - \overline{a}_u$. With $\overline{a}_u = \overline{a}_g - (\overline{a}_g - \overline{a}_u)$, (A4) determines $\overline{a}_g$. Then (A3) determines $f(0)$. Since D is only present in (A3), variations in D would be exclusively accommodated by variations in $f(0)$. ■

6.1.2 Proof of Proposition 2:

Proof. combining (A3) with (A4) and replacing $\overline{a}_u$ by $\overline{a}_g - (\overline{a}_g - \overline{a}_u)$ gives

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

The left-hand monotonically increases 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 (A6), $\frac{d(\overline{a}_u)}{dD} = \frac{d(\overline{a}_g - (\overline{a}_g - \overline{a}_u))}{dD} \geq 0$.

Similarly, with $\overline{a}_g - \overline{a}_u$ independent of D, $\frac{d(jd^{ss})}{dD} = \frac{d(jd^{ss})}{d\overline{a}_u} \frac{d\overline{a}_u}{dD} \leq 0$. ■

6.1.3 Proof of Proposition 3:

Proof.

$$
\begin{aligned}
l_g^{ss} &= \frac{\sum_{a=0}^{\overline{a}_g}\left[f\left(\theta_g,a\right)+\varphi f\left(\theta_u,a\right)\right]}{\sum_{a=0}^{\overline{a}_g}\left[f\left(\theta_g,a\right)+f\left(\theta_u,a\right)\right]} \\
&= \frac{\varphi(\overline{a}_u+1)+\varphi\left[1-(1-p)^{\overline{a}_u+1}\right]\left(\overline{a}_g-\overline{a}_u\right)}{\sum_{a=0}^{\overline{a}_u}\left[\varphi+(1-\varphi)\left(1-p\right)^a\right]+\varphi\left[1-(1-p)^{\overline{a}_u+1}\right]\left(\overline{a}_g-\overline{a}_u\right)} \\
&= 1-\frac{\sum_{a=0}^{\overline{a}_u}\left(1-\varphi\right)\left(1-p\right)^a}{\sum_{a=0}^{\overline{a}_u}\left[\varphi+(1-\varphi)\left(1-p\right)^a\right]+\varphi\left[1-(1-p)^{\overline{a}_u+1}\right]\left(\overline{a}_g-\overline{a}_u\right)} \\
&= 1-\frac{(1-\varphi)}{\frac{p\varphi(\overline{a_u}+1)}{1-(1-p)^{\overline{a_u}+1}}+(1-\varphi)+p\varphi\left(\overline{a_g}-\overline{a_u}\right)}
\end{aligned}
$$

(15) implies that $\overline{a}_g-\overline{a_u}$ is independent of D, so that

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

Proposition 3 has established that $\frac{d(\overline{a_u})}{d(D)}\geq 0$. Therefore, $\frac{d(l_g)}{d(D)}\geq 0$ if and only if $\frac{d(l_g)}{d(\overline{a_u})}\geq 0$. With $\frac{\overline{a_u}+1}{1-(1-p)^{\overline{a_u}+1}}=x$, $\frac{d(l_g)}{d(\overline{a_u})}=\frac{d(l_g)}{d(x)}\frac{d(x)}{d(\overline{a_u})}$. Since $\frac{d(l_g)}{d(x)}>0$, $\frac{d(l_g)}{d(\overline{a_u})}\geq 0$ if and only if $\frac{d(x)}{d(\overline{a_u})}\geq 0$. Hence, we need to prove that $\frac{d(x)}{d(\overline{a_u})}\geq 0$.

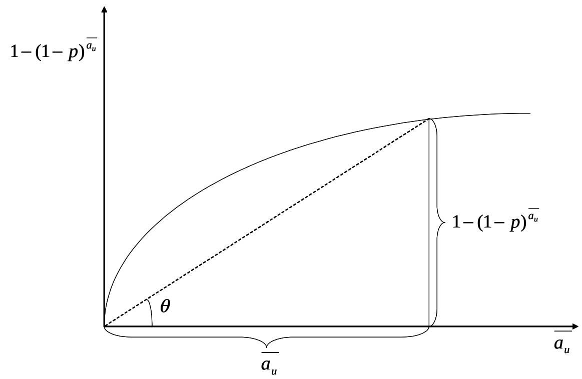
$1-(1-p)^{\overline{a_u}+1}$ is plotted in the following graph as a function of $\overline{a_u}+1$. Since

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

but

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

the curve is concave.



Clearly, it indicates that $x = \frac{\overline{a_u}+1}{1-(1-p)^{\overline{a_u}+1}} = \cot(\theta)$ .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 $\overline{a_u}$. ■

6.2 Appendix 2: Calibrating demand differential and entry cost.

The algorithm includes the following steps:

1. Loop around three conditions to find $\overline{a_g}h$. It generates an average firm age of around 58 quarters, a mean entry rate of around 3.11%, and a mean exit rate of around 3.44%.

2. Let the exit margin to shift from $\overline{a_g}h$ to younger age quarter by quarter, until it generates the observed peak in exit rate. We then use the observed trough in entry rate to find the *proportional* drop in entry size at this moment. This is when low demand hits the high-demand equilibrium.

3. Use the size of the shift and the proportional drop in entry size to calculate the output at this moment normalized by the high-demand entry size, which, combined with (A2) from Appendix 1, gives us $\frac{Dl}{fh}$.

4. Assume that demand stays low. Let the exit margin to move to older age quarter by quarter; meanwhile, calculate the exit rate arising from the learning margin. Stops when it reaches the observed minimum exit rate. Move the exit margin back by one quarter – this is where $\overline{a_g}l$ positions.

5. Calculate $\frac{Dl}{fl}$ using (A3), which, together with $\frac{Dl}{fh}$ from Step 3, determines $\frac{fh}{fl}$.

6. Use (A2) and (A4) from Appendix 1 to find the entry values: Vh and Vl. Calculate c_0 by equating $\frac{fh}{fl}$ to $\frac{Vh-c_0}{Vl-c_0}$.

7. Calculate fh and fl using $Vh = c_0 + c_1 fh$ and $Vl = c_0 + c_1 fl$.

8. With fh and fl, we calculate Dh and Dl using (A3).

6.3 Appendix 3: Approximating Value Functions with Krusell & Smith (1998) Approach.

The key computational task is to map F, the firm distribution across ages and idiosyncratic productivity, given demand level D, into a set of value functions $V(\theta^e, a; F, D)$. To make the state space tractable, we define a variable X such that:

$$X(F) = \sum_a \sum_{\theta^e} (1+\gamma)^{-a} \theta^e f(\theta^e, a). \tag{A7}$$

Combining (A7) with (6) and (7) in the text gives

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

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

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

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

$$\pi(a, \theta; F, D) = \frac{D}{X(H(F, D))}(1+\gamma)^{-a}(\theta + \varepsilon) - 1. \tag{A8}$$

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 we need to find an information set Ω that delivers a good approximation of firms' equilibrium behavior, yet is small enough to reduce the computational difficulty.

I look for an Ω through the following procedure. In step 1, we choose a candidate Ω. In step 2, we postulate perceived laws of motion for all members of Ω, denoted H_Ω, such that $\Omega' = H_\Omega(\Omega, D)$. In step 3, given H_Ω, we calculate firms' value functions on a grid of points in the state space of Ω applying value function iteration, and obtain the corresponding industry-level decision rules – entry sizes and exit ages across aggregate states. In step 4, given such decision rules and an initial firm distribution. We simulate the behavior of a continuum of firms along a random path of demand realizations, and derive the implied aggregate behavior — a time series of Ω. In step 5, we 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, we update H_Ω, return to step 3 and continue until convergence. In step 6, once H_Ω converges, we evaluate the fit of H_Ω in terms of tracking the aggregate behavior. If the fit is satisfactory, we stop; if not, we 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. We further assume that firms perceive the sequence of future coming $X's$ as depending on nothing more than the current observed X and the state of demand. The perceived law of motion for X is denoted H_x so that $X' = H_x(X, D)$. We then apply the procedure described above and simulate the behavior of a continuum of firms over 5000 periods. The results are presented in the following Table.

Ω	$\{X\}$
H_Ω	$H_x(X, D_h)$: $\log X' = 0.8429 + 0.8880 \log X$ $H_x(X, D_l)$: $\log X' = 1.2588 + 0.8321 \log X$
R^2	for D_h: 0.9959 for D_l: 0.9472
standard forecast error	for D_h: 0.00029 for D_l: 0.00079
maximum forecast error	for D_h: 0.0001 for D_l: 0.0007
Den Haan & Marcet test statistic (χ^2_7)	0.7343

The estimated H_x is log-linear. The fit of H_x is quite good, as suggested by the high R^2, the low standard forecast error, 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, we implement the Den Haan and Marcet (1994) test using instruments $[1, X, \mu_a, \sigma_a, \gamma_a, \kappa_a, r_u]$, where μ_a, σ_a, γ_a, κ_a,r_u are the mean, standard deviation, skewness, and kurtosis of the age distribution of firms, and the fraction of unsure firms, respectively. The test statistic is 0.7343, well below the critical value at the 1% level. This suggests that given the estimated laws of motion, we do not find much additional forecasting power contained in other variables.

Figure A1 displays the value of staying for heterogeneous firms as a function of a, θ^e, D and X ($\log X$). Figure A2 displays the corresponding optimal exit ages and entry sizes. These tables and figures suggest that our solution using X to approximate the aggregate state closely replicates optimal firm behavior at the equilibrium. These results were robust when experimented with different parameterization of the model. Therefore, we use the solution based on $\Omega = \{X\}$ to generate all the relevant series.

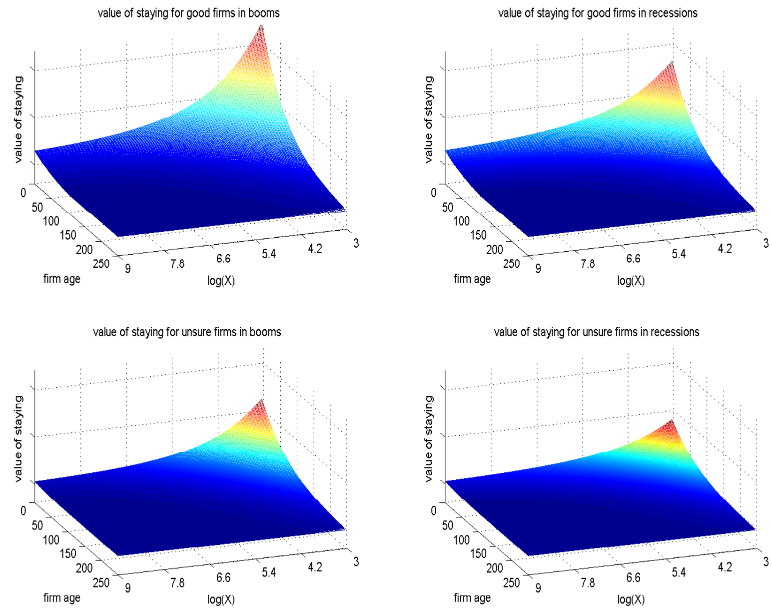


Figure A1: Expected Value of Staying: aggregate state variables are D and $\log X$ (the log of detrended output), firm-level state variables are firm age and belief (good or unsure); the parameter choices underlying these figures are summarized in Table 2.

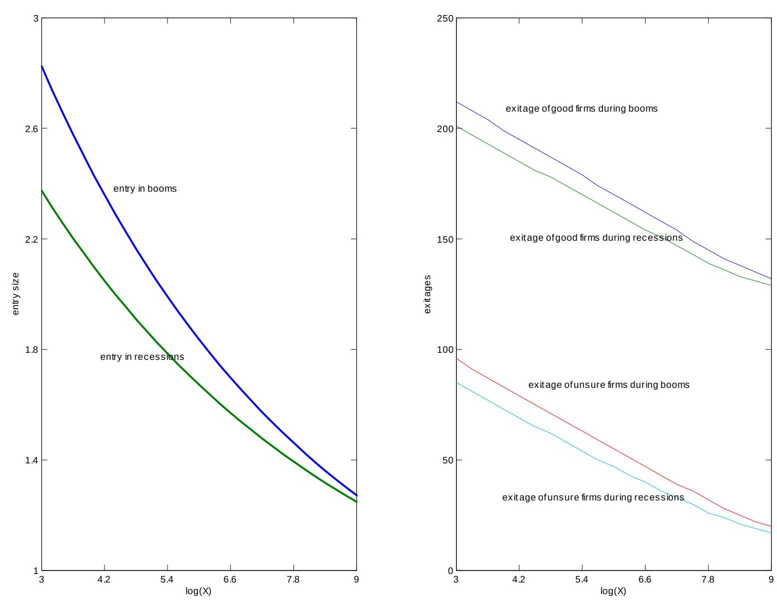


Figure A2: Industry-level Policy Functions: Entry Size and Exit Ages. Aggregate states are D (booms or recessions) and $\log X$ (the log of detrended output).

References

[1] Campbell, J. 1998. Entry, Exit, Embodied Technology, and Business Cycles. Review of Economic Dynamics 1, 371-408.

[2] Baden-Fuller, C., 1989. Exits From Declining Industries and the Case of Steel Castings. Economic Journal 99, 949 – 961.

[3] Barlevy, G., 2002. The Sullying Effect of Recessions. Review of Economic Studies 69(1), 65-96.

[4] Barlevy, G., 2003. Credit Markt Frictions and the Allocation of Resources Over the Business Cycle. Journal of Monetary Economics 50, 1795 – 1818.

[5] Bartelsman, E. J., Doms, M., 2000. Understanding Productivity: Lessons from Longitudinal Microdata. Journal of Economic Literature 38(3), 569-594.

[6] Basu, S., Fernald, J., Shapiro, M., 2001. Productivity Growth in the 1990s: Technology, Utilization, or Adjustment? Carnegie-Rochester Conference Series on Public Policy 55(1), 117-165.

[7] Bowlus, A., 1995. Matching Workers and Jobs: Cyclical Fluctuations in Match Quality. Journal of Labor Economics 13(2), 335-350.

[8] Caballero, R., Hammour, M., 1994. The Cleansing Effect of Recessions. American Economic Review 84(5), 1350-1368.

[9] Caballero, R., Hammour, M., 1996. On the Timing and Efficiency of Creative Destruction. Quarterly Journal of Economics 111(3), 805-852.

[10] Carlyle, T., 1899. Critical and Miscellaneous Essays (5). Charles Scribner's Sons, New York.

[11] Caves, R., 1998. Industry Organization and New Findings on the Turnover and Mobility of Firms. Journal of Economic Literature, 1947-1982.

[12] Cooley, T., Quadrini, V., 2001. Financial Markets and Firm Dynamics. American Economic Review 91(5), 1286-1310.

[13] Davis, S., Haltiwanger, J., 1992. Gross Job Creation, Gross Job Destruction, and Employment Reallocation. Quarterly Journal of Economics 7(3), August 1992, 819-863.

[14] Davis, S., Haltiwanger, J., 1999. Gross Job Flows. In: Ashenfelter, O., Card, D. (Eds), Handbook of Labor Economics 41. Elsevier, Amsterdam, pp. 2711-2805.

[15] Davis, S., Haltiwanger, J., Schuh, S., 1996. Job Creation and Destruction. MIT Press, Cambridge.

[16] Den Haan, W.J., Marcet, A., 1994. Accuracy in Simulations. Review of Economic Studies 61, 3-18.

[17] Dunne, T., Roberts, M.J., Samuelson, L., 1989. The Growth and Failure of U.S. Manufacturing Plants. Quarterly Journal of Economics 104(4), 671-98.

[18] Faberman, R. J., 2003. Job Flows and Establishment Characteristics: Variations across Metropolitan Areas. William Davidson Institute Working Papers 609.

[19] Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? American Economic Review, forthcoming.

[20] Goolsbee, A.,1998. Investment Tax Incentives, Prices, and the Supply of Capital Goods. Quarterly Journal of Economics 113(1), 121-148.

[21] Hall, R., 2000. Reorganization. Carnegie-Rochester Conference Series on Public Policy 52, 1-22.

[22] Lucas, R. E., 1978. On the Size Distribution of Business Firms. The Bell Journal of Economics 9(2), 508-523.

[23] Jensen, J. B., McGuckin, R. H., Stiroh, K.J., 2001. The Impact of Vintage and Survival on Productivity: Evidence from Cohorts of U.S. Manufacturing Plants. Review of Economics and Statistics 83(2), 323-332.

[24] Jovanovic, B., 1982. Selection and the Evolution of Industry. Econometrica 50(3), 649-670.

[25] Jovanovic, B., Nyarko, Y. 1996. Learning by Doing and the Choice of Technology. Econometrica 64(6), 1299-1310.

[26] Krusell, P., Smith, A. Jr., 1998. Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy 106(5), 867-895.

[27] Mortensen, D., Pissarides, C., 1994. Job Creation and Job Destruction in the Theory of Unemployment. Review of Economic Studies 61(3), 397-415.

[28] Moscarini, G., 2003. Skill and Luck in the Theory of Turnover. working paper, Yale economics department.

[29] Pries, M. J., 2004. Persistence of Employment Fluctuations: a Model of Recurring Firm Loss. Review of Economic Studies 71(1), 193-215.

[30] Pries, M.J., Rogerson, R., 2005. Hiring Policies, Labor Market Institutions, and Labor Market Flows. Journal of Political Economy 113(4), 811-839.

[31] Ramey, G., Watson, J., 1997. Contractual Fragility, Job Destruction, and Business Cycles. Quarterly Journal of Economics 112(3), 873-911.

[32] Schumpeter, J. A., 1934. Depressions. In: Douglass V. B., Edward C., Seymour E. H. (Eds), Economics of the Recovery Program. Whittlesey House, New York, pp. 3-12.

7 Figures and Tables

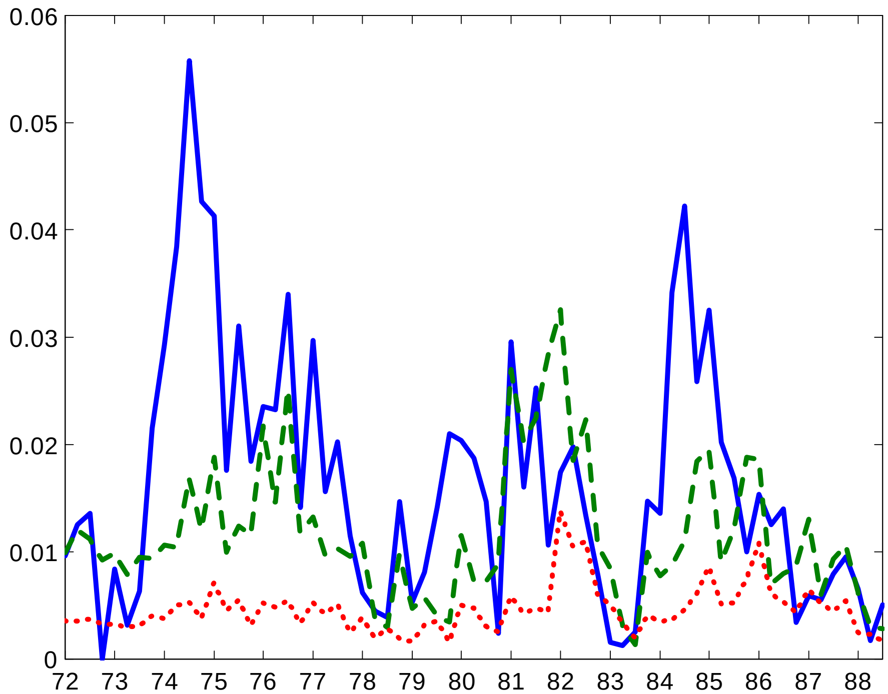


Figure 1: Quarterly Exit Rates of the U.S. Manufacturing Sector. Solid line represents that exit rate of plants aged ten years or older; dashed line that of plants aged between one and ten years; dotted line that of plants aged younger than a year. All rates are employment-weighted. Data sources: series of gross job flows compiled by Davis and Haltiwanger.

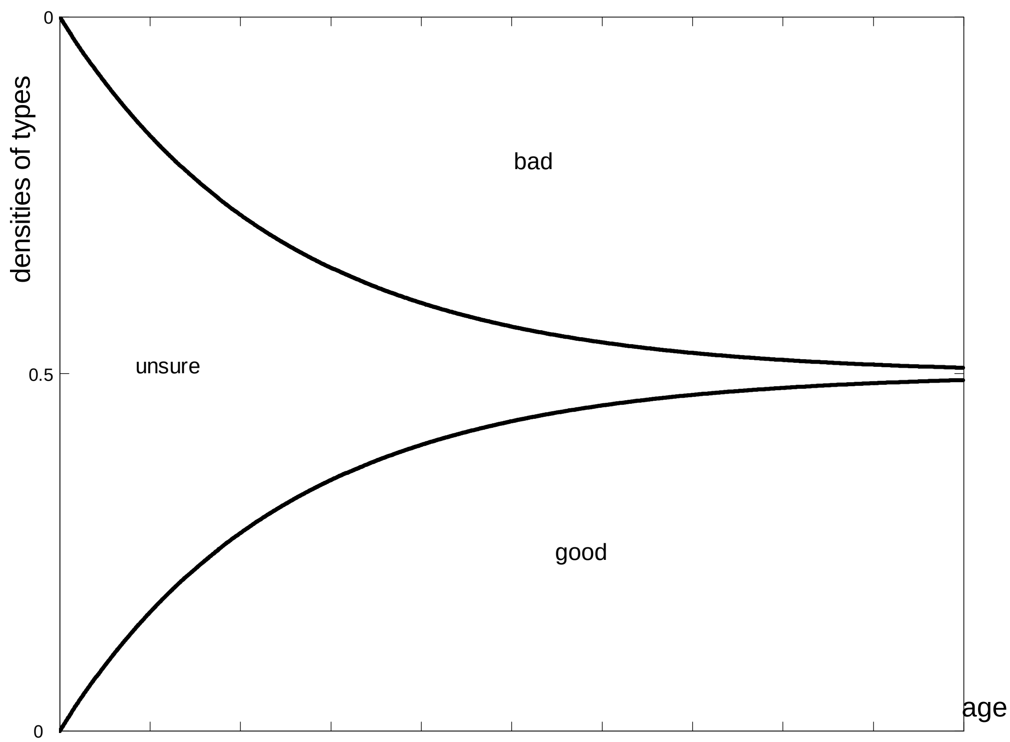


Figure 2: Dynamics of a Birth Cohort: the distance between the concave curve and the bottom axis measures the density of firms with $\theta^e = \theta_g$; the distance between the convex curve and the top axis measures the firms with $\theta^e = \theta_b$; the distance between the two curves measures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

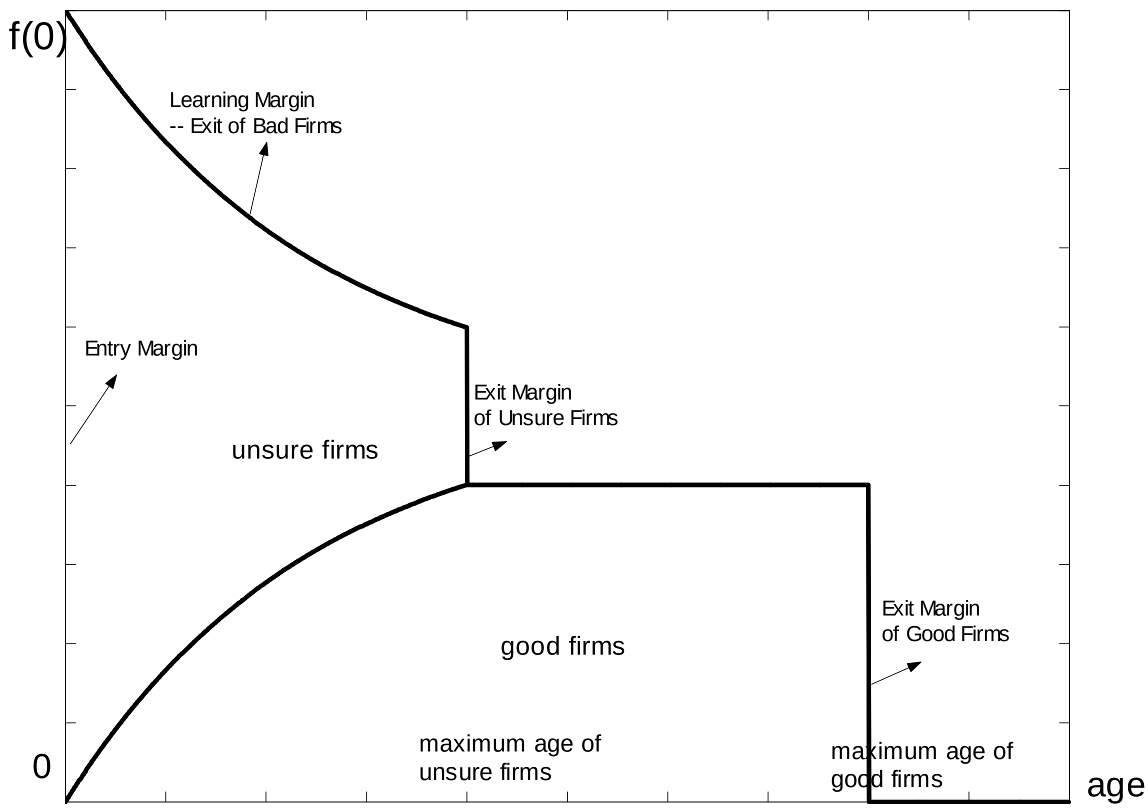


Figure 3: The Steady-state Labor Distribution and Job Flows: the distance between the lower curve (extended as the horizontal line) and the bottom axis measures the density of good firms; the distance between the two curves measures the density of unsure firms.

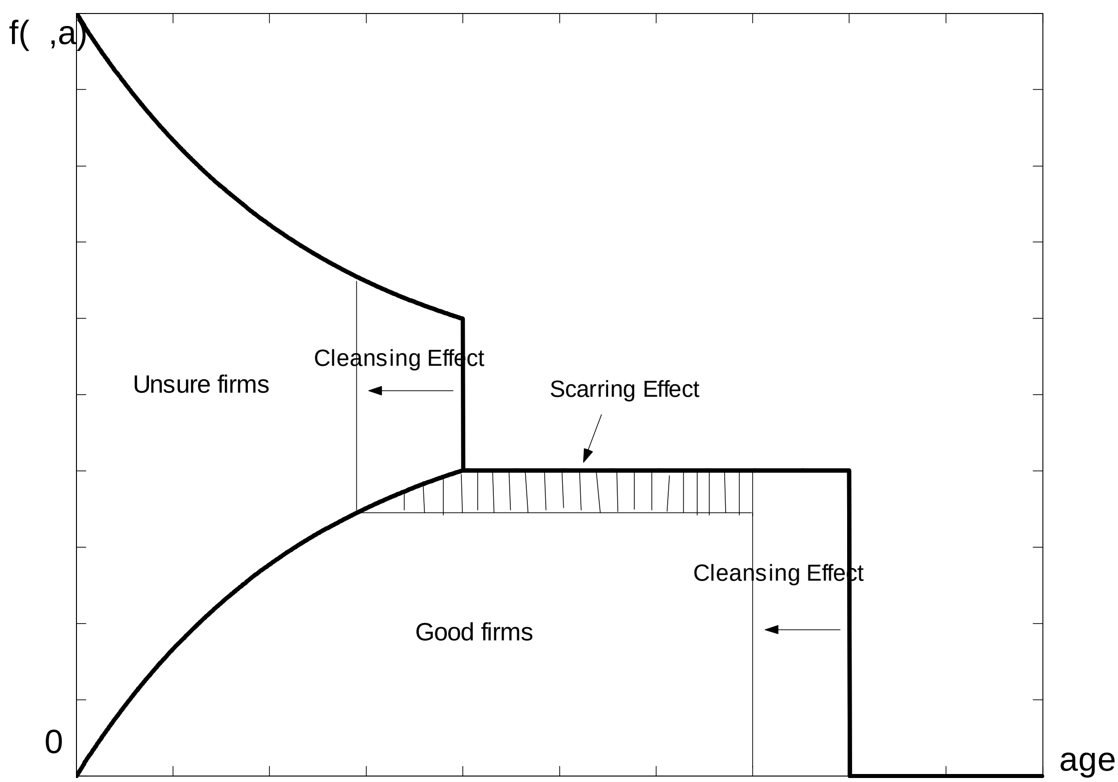


Figure 4: Cleansing and Scarring

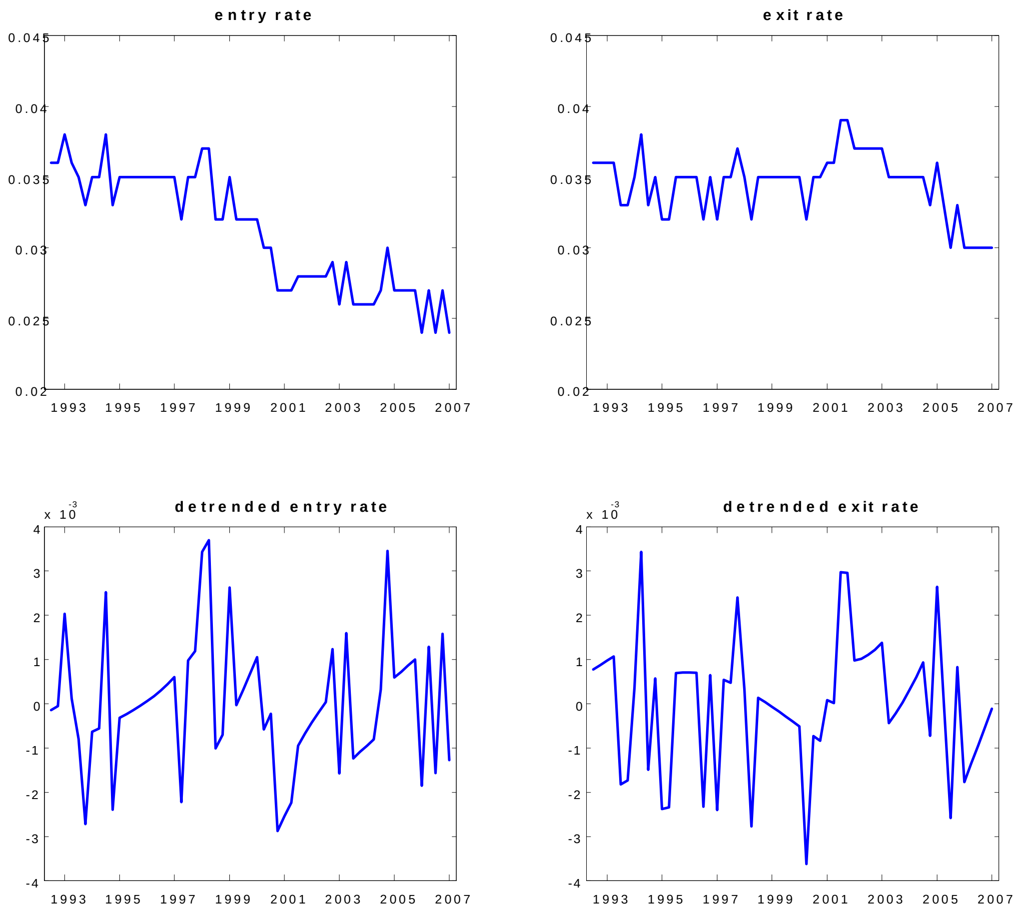


Figure 5: U.S. Manufacturing Quarterly Entry and Exit Rates (1992 to 2007). The top two panels display the actual data; the bottom two panels display the detrended data using the HP filter. Data source: the Business Employment Dynamics provided by the Bureau of Labor Statistics.

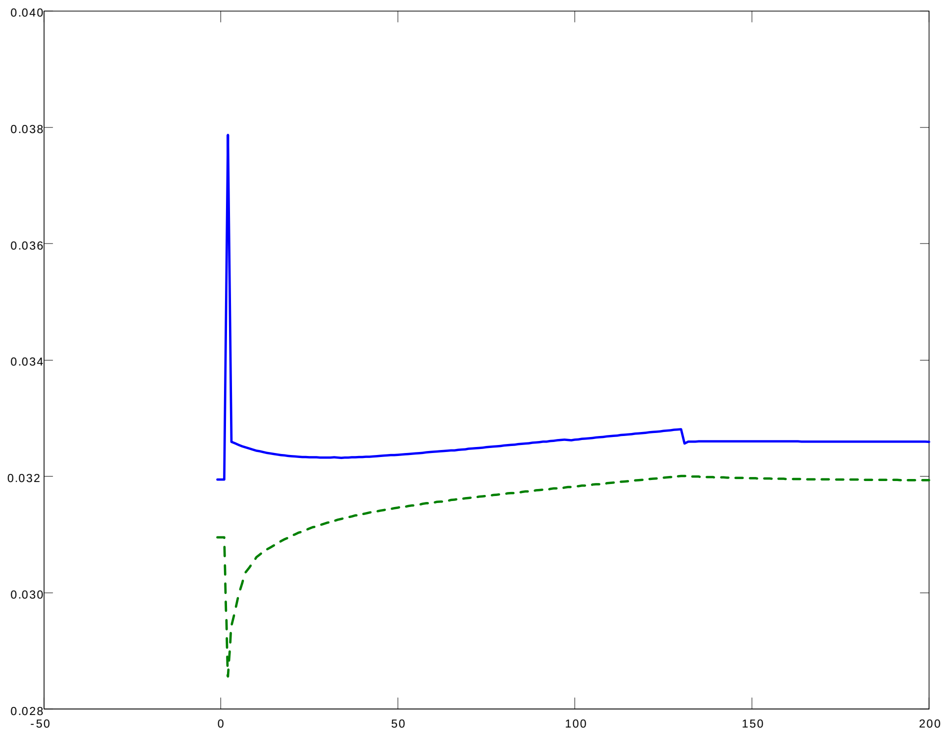


Figure 6: Response of Exit and Entry Rates To a Negative Demand Shock: the horizontal axis denotes quarters, with the quarter labeled 0 denoting the onset of a recession; solid lines denotes the dynamics of exit, dashed line that of entry.

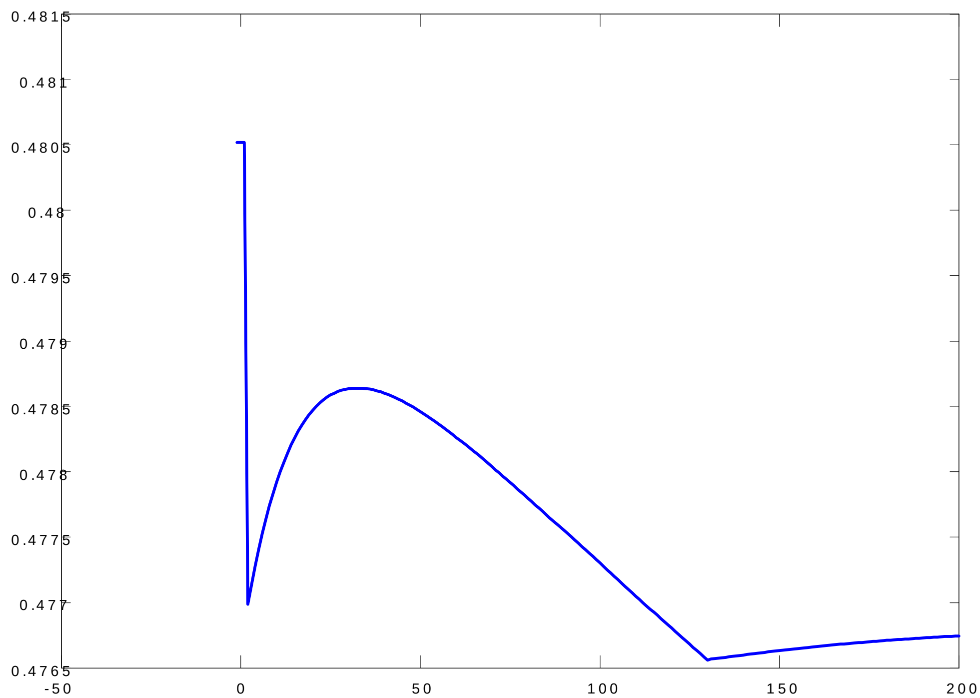


Figure 7: Response of the Ratio of Labor at Good Firms to A Negative Demand Shock: the horizontal axis denotes quarters, with the quarter labeled 0 denoting the onset of a recession.

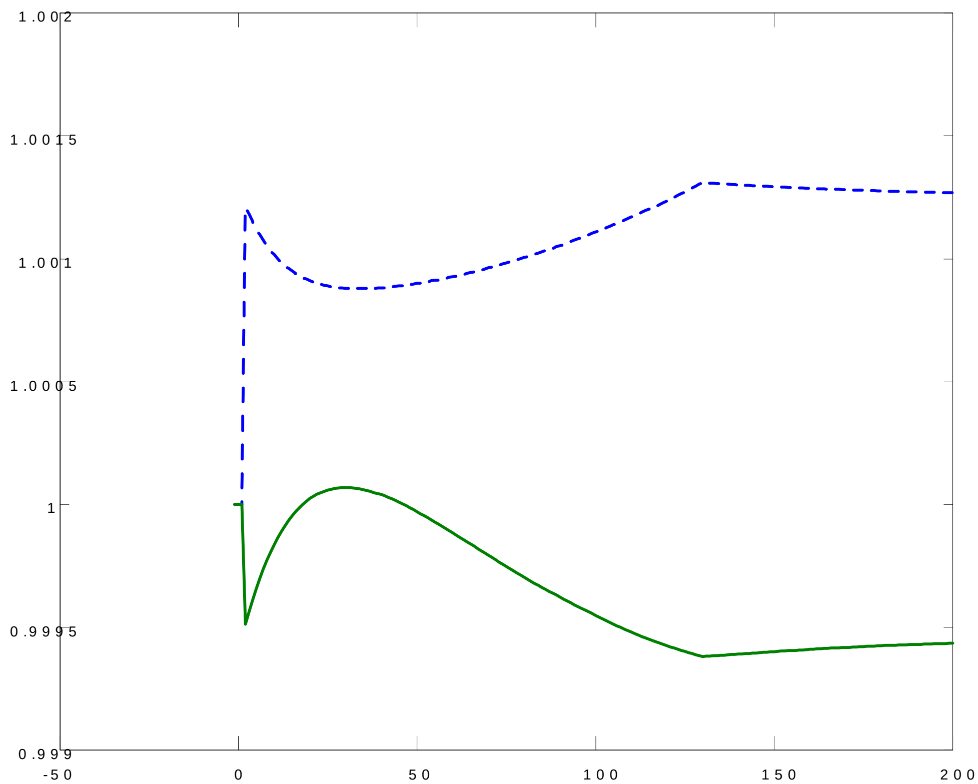


Figure 8: Response of Productivity to a Negative Demand Shock. Dashed line represents productivity driven by the cleansing effect only; solid line denotes that driven by cleansing and scarring. Period zero denotes the onset of the shock.

parameters	**value**
quarterly discount factor: β	0.9900
persistence rate of demand: μ	0.9500
prior probability of being a good firm: φ	0.1157
quarterly pace of learning: p	0.0538
quarterly technological pace: γ	0.0040
productivity of bad firms: θ_b	1
productivity of good firms: θ_g	1.7500
entry cost parameter: c_1	0.7284
entry cost parameter: c_0	0.1587
High demand: D_h	108.7294
Low demand:D_l	103.9819

Table 1: Calibration