Working Paper Series

Firm-Specific Learning and the Investment Behavior of Large and Small Firms

WP 99-03

Wenli Li
Federal Reserve Bank of Richmond

John A. Weinberg
Federal Reserve Bank of Richmond

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Firm-Specific Learning and the Investment Behavior of Large and Small Firms

Wenli Li and John A. Weinberg*
Federal Reserve Bank of Richmond

April 2, 1999

Abstract

We examine a model of the size distribution and growth of firms whereby firms learn about idiosyncratic productivity parameters. Aggregate shocks, by adding noise to learning at the firm level, can produce differentiated response across firms with their reactions depending on the position of the firms in their individual life cycle. In particular, young firms, which are smaller on average than older firms, can "overreact" to aggregate shocks. Such differences across firm sizes and ages, which arise here in a model with perfect financial markets, are often attributed to financial frictions that hit small and large firms differently.

*Correspondence: Research Department, Federal Reserve Bank of Richmond, 701 E. Byrd Street, Richmond, VA 23261. E-mail: wenli.li@rich.frb.org, john.weinberg@rich.frb.org. We would like to thank seminar participants at the Federal Reserve Bank of Richmond, the Federal Reserve Bank of Minneapolis, and Indiana University.

1. Introduction

In recent years, macroeconomists have given considerable attention to the implications of heterogeneity among firms. The focus has been on differences in behavior across the size distribution of firms. While many empirical studies have established numerous differences, a growing body of theoretical work has examined the equilibrium behavior of economies with heterogeneous firms, the key focus being on firm-level investment and growth dynamics. The traditional view on this subject was that the growth rates of firms followed Gibrat's Law, which holds the rate of growth to be independent of firm size. This view, however, has been largely rejected by more recent evidence suggesting a negative correlation between size and growth (although the correlation exists mainly in subpopulations of smaller firms).[1] Not only did the studies show the correlation, but they also tended to find an inverse relationship between size and the variability of growth rates.

Differences between large and small firms appear to be related to differences among firms that vary by age. For example, younger firms are generally smaller and have faster average growth rates. They also have greater variance in their growth rates and face higher probability of exiting (through either failure or acquisition).[2]

A number of authors have offered equilibrium models that capture some or all of these empirical facts. Most of the models derive the size distribution of firms from an underlying distribution of firm-specific managerial abilities (or other specific factors affecting productivity).[3] The stochastic process governing the firm-specific productivity factor (or firm's beliefs about that factor) drives the firm-level dynamics. In Jovanovic (1982), firms' abilities are fixed, and firms learn about their abilities through their own (noisy) production experience. In Hopenhayn (1992), firms' abilities change over time, causing firms to grow or decline in size.

Most recently, macroeconomists have focused on specific aspects or implications of firm-level heterogeneity. Davis, Haltiwanger, and Schuh (1996), for instance, have studied employment dynamics of manufacturing firms. A number of other authors have focused on the investment behavior of heterogeneous firms. Using a general equilibrium model of employment dynamics that is consistent with the findings of the latter focus, Hopenhayn and Rogerson (1993) study the welfare implications of various employment and unemployment policies. Perhaps the most widely noted strand of literature on differences between large

[1] Important contributions on this subject include Evans (1987a and b), and Hall (1987).
[2] Exit rates across the size and age distribution are studied by Dunne, Roberts and Samuelson (1988).
[3] Most such models are dynamic versions of Lucas (1978).

and small firms, however, deals with investment behavior. Fazzari, Hubbard, and Petersen (1988) noted that the investment of small, fast growing firms does not fit traditional models of investment behavior. In particular, investment by these firms is sensitive to current period cash flow. This finding, which has given rise to a number of similar and related results in the literature, has been interpreted as evidence of financial constraints facing some firms. Since access to external funding is more costly or difficult for smaller firms to obtain, these firms' investment must be more closely tied to the availability of internal funds. A related finding, presented by Gertler and Gilchrist (1994), is the greater volatility of small firm sales in response to some aggregate shocks, including monetary shocks. Such findings form the basis for the view that monetary policy influences the economy through a credit channel that mainly affects the investment behavior of small firms.

Many authors have examined models with heterogeneous firms and financial frictions that are broadly consistent with the empirical findings on the investment behavior of large and small firms.[4] These models differ from the standard growth model in two ways: they deal with heterogeneous firms and there exists financial frictions. What is the relative importance of the frictions compared with the fundamental forces driving the growth of firms of varying sizes? For one thing, it is well known that at least some of the sensitivity of firm-level investment to cash flow is consistent with a model in which cash flow is informative about investment opportunities. The Jovanovic (1982) learning model is just such a model.

In this paper, we offer a benchmark against which to measure models of financial frictions. Our model looks at heterogeneous firms that learn about their own attributes and make entry, exit, and investment decisions in an environment with aggregate uncertainty and perfect capital markets. We calibrate a steady-state, general equilibrium version of the model to certain characteristics of the size distribution of U.S. manufacturing firms. We then examine the dynamic behavior of a partial equilibrium version of the model with aggregate shocks. In particular, we consider the average responses to aggregate shocks of broadly defined classes of large and small firms and find that small-firm behavior is more responsive than that of large firms.

A key to the effects of aggregate shocks in the model is an assumed inability to fully and immediately distinguish aggregate shocks from firm-specific factors. Firms seeking to learn about their own abilities must filter out the effects that changes in aggregate conditions have on their production experience. We assume that firms know their own current period output when they make their investment decisions, but that they do not observe current

[4] Fisher (1996), Gomes (1997), Cooley and Quadrini (1998), and Li (1998).

aggregate results until they have committed to an investment level for the next period's production. Firms' ability to filter out aggregate factors differs according to firm age, since beliefs about ability become increasingly precise with age. Hence, younger firms are more likely to overreact to an aggregate shock than are older firms. This difference, together with the equilibrium relationship between firm size and age, drives the differences in behavior between different classes of size.

2. Firm Behavior and Learning

We examine an environment in which the production behavior of firms is driven both by firm-specific productive capabilities and by specific as well as aggregate shocks. Firm-specific productivity is a permanent characteristic that the firm is unable to observe directly. Instead, a firm must learn about its ability through its own productive experience. Hence, firm behavior is as modeled by Jovanovic (1982). The firm's learning process is complicated by the presence of aggregate shocks and by the firm's inability to distinguish aggregate from idiosyncratic shocks when making its investment decisions.

There is a continuum of potential firms. A potential firm makes its entry decision before observing any information on the firm-specific productivity. Hence, the entry decision compares the expected present value of profits for the average firm to the cost of entry. Firms make decisions that maximize the value delivered to shareholders. This premise assumes that there is free entry into the (costless) activity of firm management, and that households can hold perfectly diversified portfolios (managers earn no rents, and there are no differences of opinion among shareholders).

After a firm incurs the entry cost, it observes a signal of its productivity. Production requires acquiring capital that becomes productive with a one-period lag. An existing firm must make an exit decision each period before making its investment choice.

A firm that has acquired k_t units of capital produces

$$y_t = e^{(\theta + z_t + \varepsilon_t)} f(k_t) \equiv e^{m_t} f(k_t), \tag{2.1}$$

where θ is the firm's productivity, z_t is an aggregate shock, and ε_t is a firm-specific shock. The production function, f, is strictly concave and increasing (with $\lim_{k \to 0} f\,'(k) = \infty$). From this output, a firm pays $r_{t-1} k_t$ for the capital it has used. The rental rate is dated $t-1$ to indicate that the capital was acquired in the previous period. The firm's profit is denoted $\pi_t \equiv y_t - r_{t-1}k_t + (1-\delta)k_t$, where δ is the rate of depreciation of capital. The residual π_t is paid to shareholders. Note that there may be states of the world in which some firms' realized profits are negative. We assume that the residual claimant shareholders are obligated to make up any shortfall. In a large economy with diversified shareholdings, aggregate profits and average return to shareholders will always be positive. A firm that chooses to exit makes a final distribution of w to its shareholders. In addition to voluntary exit, we assume that a fraction ξ of all firms active in period t die prior to making their period $t+1$ exit decision.

A firm makes its investment decision (k_{t+1}) knowing its own current output (y_t), but not aggregate output (Y_t). Consequently, it does not know the current value of the aggregate shock (z_t). That is, at the time of its time t investment decision, the firm's information set includes the histories $m^t$and z^{t-1}, as well as the current interest rate r_t.[5] From this information, the firm constructs an expectation of m_{t+1}. Information on aggregate output, and therefore on the current value of the aggregate shock, z_t, is not observed until after the investment decision has been made.

We make the following assumptions about the probability structure governing θ, ε, and z: $\theta \sim N(\mu, \sigma^2)$; $\varepsilon_t \sim N(0, \tau^2)$; and $z_t \sim N(\rho z_{t-1}, \eta^2), 0 < \rho < 1$. Let θ_t denote the expected value of θ, conditional on the history (m^t, z^t), and let σ_t^2 denote the conditional variance. These beliefs evolve according to

$$\theta_t = \theta_{t-1} + \frac{\sigma_{t-1}^2}{\sigma_{t-1}^2 + \tau^2}(m_t - z_t - \theta_{t-1}) \tag{2.2}$$

$$\sigma_t^2 = \frac{\sigma_{t-1}^2 \tau^2}{\sigma_{t-1}^2 + \tau^2}. \tag{2.3}$$

The updating rule is based on the observation of the current aggregate shock. When making its investment decision, however, the firm has not observed this shock. At that stage, the firm holds an interim belief about its productivity, summarized by the pair, $(\widetilde{\theta}_t, \widetilde{\sigma}_t^2)$, where

$$\widetilde{\theta}_t = \theta_{t-1} + \frac{\sigma_{t-1}^2}{\sigma_{t-1}^2 + \tau^2 + \eta^2}(m_t - \rho z_{t-1} - \theta_{t-1}), \tag{2.4}$$

[5] We assume that the price r_t is not informative about the aggregate state. This assumption will be discussed more below.

$$\widetilde{\sigma}_t^2 = \frac{\sigma_{t-1}^2(\tau^2 + \eta^2)}{\sigma_{t-1}^2 + \tau^2 + \eta^2}. \tag{2.5}$$

The updating rule indicates how much of the observed sum $m_t = \theta + z_t + \varepsilon_t$ the firm attributes to news about its own productivity. At the same time, the firm updates its beliefs about the current values of the aggregate and idiosyncratic shocks:

$$\widetilde{z}_t = \rho z_{t-1} + \frac{\eta^2}{\sigma_{t-1}^2 + \tau^2 + \eta^2}(m_t - \rho z_{t-1} - \theta_{t-1}), \tag{2.6}$$

$$\eta_t^2 = \frac{\eta^2(\sigma_{t-1}^2 + \tau^2)}{\sigma_{t-1}^2 + \tau^2 + \eta^2}, \tag{2.7}$$

$$\widetilde{\varepsilon}_t = \frac{\tau^2}{\sigma_{t-1}^2 + \tau^2 + \eta^2}(m_t - \rho z_{t-1} - \theta_{t-1}), \tag{2.8}$$

$$\tau_t^2 = \frac{\tau^2(\sigma_{t-1}^2 + \eta^2)}{\sigma_{t-1}^2 + \tau^2 + \eta^2}. \tag{2.9}$$

To summarize, the evolution of the firm's beliefs about its own productivity is a two-stage process in each period. The firm comes into period t with beliefs $(\theta_{t-1}, \sigma_{t-1}^2)$. Upon observing m_t, it adjusts its beliefs to $(\widetilde{\theta}_t, \widetilde{\sigma}_t^2)$. Based on this belief, the firm makes its exit and investment decisions. Once aggregate output is observed, beliefs adjust again to (θ_t, σ_t^2). The beliefs with which the firm exits period t are consistent both with updating from $(\theta_{t-1}, \sigma_{t-1}^2)$ upon observing $\theta + \varepsilon_t$, and with updating from $(\widetilde{\theta}_t, \widetilde{\sigma}_t^2)$ upon observing z_t. For computational purposes, we will also use an alternative formulation of the stochastic process governing the aggregate shock. Under this alternative, we will assume that z_t follows a two-state Markov process with states z^h and z^l and transition probabilities p_{ij}. This process for z clearly changes the evolution of beliefs about θ. For a given θ_{t-1} and $m_t = \theta + z_t + \varepsilon_t$, θ_t will take one of two values. Specifically, if $z_t = z^j$ $(j = l, h)$ then

$$\theta_t = \theta_t^j \equiv \theta_{t-1} + \frac{\sigma_t^2}{\sigma_t^2 + \tau^2}(m_t - z^j - \theta_{t-1}). \tag{2.10}$$

When the firm makes its time t investment decision (when m_t is known but z_t is not), its belief, which is conditioned on $z_{t-1} = z^i$, is given by

$$\widetilde{\theta}_t^i = \frac{p_{ih} g_{t-1}(m_t - z^h)}{p_{ih} g_{t-1}(m_t - z^h) + p_{il} g_{t-1}(m_t - z^l)} \theta_t^h + \frac{p_{il} g_{t-1}(m_t - z^l)}{p_{ih} g_{t-1}(m_t - z^h) + p_{il} g_{t-1}(m_t - z^l)} \theta_t^l, \tag{2.11}$$

where g_{t-1} is the normal density with mean θ_{t-1} and variance σ_{t-1}^2. Except where noted, all discussion in what follows refers to the continuous process for z.

A firm makes its choices to maximize the expected present value of its profits. The firm's behavior can be characterized by the Bellman equation for a firm that has chosen not to exit in period t:

$$\begin{aligned} v(z_{t-1}, \theta_{t-1}, \sigma_{t-1}^2, m_t; r_t) \;\; &= \;\; \beta \max E\left[e^{m_{t+1}} f(k_{t+1}) - r_t k_{t+1} + (1-\delta) k_{t+1} \mid z^{t-1}, m^t\right] + \\ & \quad (1-\xi)\beta E\{\max[w; v(z_t, \theta_t, \sigma_t^2, m_{t+1}; r_{t+1})] | z^{t-1}, m^t\} \end{aligned} \tag{2.12}$$

The investment decision is a static choice problem, since k_{t+1} has no effect beyond π_{t+1}. Hence, investment is determined by the first-order condition:

$$E\left[e^{m_{t+1}} \mid z_{t-1}, \theta_{t-1}, \sigma_{t-1}^2, m_t\right] f'(k_{t+1}) = r_t + \delta - 1, \tag{2.13}$$

or

$$\exp[\widetilde{\theta}_t + \rho \widetilde{z}_t + \frac{1}{2}(\widetilde{\sigma}_t^2 + \rho^2 \eta_t^2 + \eta^2 + \tau^2)] f'(k_{t+1}) = r_t + \delta - 1. \tag{2.14}$$

This condition determines the firm's investment as a function of the rental rate and its beliefs conditional on the history (z^{t-1}, m^t). The second term on the right-hand side of equation (2.12) reflects the firm's choice between exiting and continuing, if it survives to that decision point. The exit choice (in period $t+1$) is conditioned on the history (z^t, m^{t+1}).

The evolution of beliefs about θ is stochastic, depending on the firm's realized output experience. However, the variance of this perceived distribution, σ_t^2, follows a deterministic path. That is, σ_t^2 depends only on a firm's age. Since it will be useful to consider the behavior of firms in a given age cohort, we can accordingly characterize the value function v for a given σ_t^2.

Lemma 2.1. *The value function $v(z,\theta,\sigma^2,m;r)$ is increasing in z,θ, m, and decreasing in r. For any (z,σ^2,m,r), there is a $\widehat{\theta}$ such that $v(z,\theta,\sigma^2,m;r) \geq w$ for all $\theta \geq \widehat{\theta}$, and $v(z,\theta,\sigma^2,m;r) \leq w$ for all $\theta \leq \widehat{\theta}$.*

One can gain some intuition regarding the differences in behavior of large and small firms by examining how the investment decision varies with m_t. Since a firm cannot distinguish aggregate from idiosyncratic shocks when it makes its investment decision, its reaction to its observed m_t in part could be an "overreaction" to the amount of m_t the firm attributes to the permanent components. An important determinant of this reaction is the firm's uncertainty regarding its own permanent productivity parameter, θ. The less precise its estimate of θ (the greater σ^2_{t-1}), the greater will be its response to variations in m_t. One can understand this result by noting that the first order condition for investment (equation (2.14)), after substituting for $\widetilde{\theta}_t$, $\widetilde{\sigma}^2_t$, $\widetilde{z}_t$, and η^2_t, can be written as follows:

$$\begin{aligned} r_t + \delta - 1 &= \exp[\theta_{t-1} + \rho^2 z_{t-1} + \frac{\sigma^2_{t-1} + \rho\eta^2}{\sigma^2_{t-1} + \eta^2 + \tau^2}(m_t - \rho z_{t-1} - \theta_{t-1}) \\ &+ \frac{\sigma^2_{t-1}(\tau^2 + \eta^2) + \rho^2\eta^2(\sigma^2_{t-1} + \tau^2)}{2(\sigma^2_{t-1} + \eta^2 + \tau^2)} \\ &+ \frac{1}{2}(\eta^2 + \tau^2)] f\,'(k_{t+1}) \end{aligned} \qquad (2.15)$$

Consider two firms with the same θ_{t-1} and m_t but with different σ^2_{t-1}. From the above equation, one can see that the firm with the less precise estimate (greater σ^2_{t-1}) will choose a greater investment level if $m_t > \rho z_{t-1} + \theta_{t-1}$ and a smaller investment if $m_t < \rho z_{t-1} + \theta_{t-1}$. This difference across variances (σ^2_{t-1}) is one across firms of different ages, since σ^2_{t-1} varies deterministically with age. However, the difference is also the key to differences in behavior across firm sizes.

The lemma states how the value function varies with θ_{t-1}, m_t, and z_{t-1}. The value function's variation with σ^2_{t-1} (and hence with age) is more complex. Here there are two opposing effects. On the one hand, uncertainty about ability causes a firm to make the "wrong" investment decision relative to the choice it would make if θ were known. In this respect, uncertainty is costly to the firm's value. On the other hand, the exit option introduces a nonconvexity in the value of remaining in business. If a firm has a low current belief (θ_{t-1}), there is a lower bound on how much worse off it can make itself by taking one more observation on its ability. The option value of an incremental observation is increasing in σ^2_{t-1}. This effect, however, is greatest in the neighborhood of $v = w$ and falls (ultimately vanishing) as v rises above w. To be more precise, if a firm's current expectation of its next period's interim belief, $E_t\left[\widetilde{\theta}_{t+1}\right]$, is close to a given exit threshold, then its next period's expected value is strictly increasing in the variance of $\widetilde{\theta}_{t+1}$. This variance is a strictly increasing function of σ_{t-1}^2. Hence, in the neighborhood of the exit threshold, the value function is increasing in σ_{t-1}^2, while beyond that neighborhood, it is decreasing. This creates a tendency for average firm size to increase with age. In particular, if one were to segment the population of firms into large and small firms, the average age of the large group would be greater than that of the small group. As a result, differences in behavior that are driven by age will show up as differences between large and small firms.

3. A Stationary Model

In this section, we present a general equilibrium version of the model without aggregate shocks. We calibrate this model so that the steady state industry structure matches the structure of the U.S. manufacturing sector along some key dimensions. The calibrated parameters are used for simulations of the model with aggregate shocks.

Existing firms behave as described above with one key exception; in the absence of an aggregate shock, a firm is able to update its beliefs from $(\theta_{t-1}, \sigma_{t-1}^2)$ to (θ_t, σ_t^2) upon observation of $m_t = \theta + \varepsilon_t$. At any point in time there is a continuum of potential firms, with management responsibilities randomly allocated among the agents in the economy and with each firm (and potential firm) having one manager.[6] Ownership is also allocated among the agents. Even ownership in potential firms that never enter is distributed. All households own identical portfolios of claims to the cash flows of all firms and potential firms. Hence, these portfolios are perfectly diversified and riskless.

Each period, a potential entrant chooses whether to incur the entry cost C_e. If the cost is incurred, the firm draws a θ from the population distribution, which is $N(\mu, \sigma^2)$. The firm does not observe its θ directly. Instead, it observes a signal S from a normal distribution with mean θ and variance σ_s^2. One can think of the entry cost as the cost of producing the initial signal of firm productivity. Conditional on this signal, the entrant's beliefs are summarized by the pair (θ_e, σ_e^2), where

[6]The act of managing a firm imposes no costs on the individual agent. Hence, there are no incentive problems, and we assume that firms are operated to maximize shareholder returns.

$$\theta_e = \mu + \frac{\sigma_s^2}{\sigma^2 + \sigma_s^2}(S - \mu), \tag{3.1}$$

$$\sigma_e^2 = \frac{\sigma_s^2 \sigma^2}{\sigma^2 + \sigma_s^2}. \tag{3.2}$$

Therefore, (θ_e, σ_e^2) plays the same role for entrants that $(\theta_{t-1}, \sigma_{t-1}^2)$ plays for existing firms. Based on these beliefs, the entrant makes an investment choice.

3.1. The Household's Problem

A continuum of identical households (with total measure one) makes consumption and savings decisions in each period. Savings out of current income (s_t) takes the form of selling some current consumption good to firms for use as capital in the next period. At time zero, each household has an endowment of goods that can be consumed or sold to firms as period-one capital. At each subsequent period $(t \geq 1)$, the household earns $r_{t-1}s_{t-1}$ on its previous period's savings and also receives its share of current period profits. These profits include the w (units of consumption good) produced by each exiting firm and the $\pi_t = y_t - r_{t-1}k_t + (1 - \delta)k_t$ for each active firm. Given that each household holds an equal share of the market portfolio, a household's income in period t will simply be equal to the period's aggregate, per capita output. These households' behavior, therefore, can be treated as though it arises from a representative household solving:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$s.t. \;\; c_0 + s_0 \leq Y_0 \tag{3.3}$$

$$c_t + s_t \leq X_t w + r_{t-1}s_{t-1} + \Pi_t, \tag{3.4}$$

$$t \geq 1,$$

where Y_0 is the aggregate initial endowment, X_t is the measure of exiting firms in period t, and Π_t is the aggregate profit of operating firms. The household chooses to make s_t units of current income available to firms for use as next period's capital. In the next period, the household receives its income in the form of rent on capital and firms' distributed profits. With no aggregate uncertainty and perfectly diversified portfolios, future income and prices are certain. The household's problem, then, is quite standard.

3.2. Industry Structure

The industrial structure of this economy at time t is determined by the distribution of firms over the $(\theta_{t-1}, \sigma_{t-1}^2)$ characteristic space. This distribution, in turn, is described by a set of measures γ_t^i $(i \geq 0)$ giving the measure of firms with i periods of production experience over θ_{t-1}, with $i = 0$ denoting the measure of entrants over θ_e. To see how these measures evolve over time, denote a firm's belief about θ coming into a period by θ', and denote the updated belief upon observing $m_t = \theta + \varepsilon_t$ by θ''. We can then write the evolution of the measures γ_t^i as

$$\gamma_{t+1}^{i+1}(\theta'') = (1-\xi) \int \gamma_t^i(\theta') g_{\theta'}^i[(\theta'' - \theta') \frac{(\sigma^i)^2 + \tau^2}{(\sigma^i)^2} + \theta'] d\theta', \tag{3.5}$$

$$\gamma_{t+1}^0(\theta'') = N_{t+1} g_\mu^0[(\theta'' - \mu) \frac{\sigma^2 + \tau^2}{\sigma^2} + \mu]. \tag{3.6}$$

where N_{t+1} is the mass of entrants at $t+1$, $(\sigma^i)^2$ is the variance of an age-i firm's beliefs about its θ, and $g_{\theta'}^i$ is the probability density function of a normal distribution with mean θ' and variance $(\sigma^i)^2 + \tau^2$.[7]

The size distribution of firms is a transformation of the distribution over θ_{t-1}. To construct this distribution, note that for firms of a given age (i) and with a given belief θ', there is a one-to-one correspondence between their output experience (m_t) and their updated belief θ'', which in turn determines their investment choice k_{t+1}. Let $h^i(k)$ denote the belief θ'' that induces an age-i firm to choose investment k. The measure of firms with a given age and prior belief θ' choosing size k is[8]

[7] Given our notation for the age-dependent variance $(\sigma^i)^2$, we will also use $(\sigma^0)^2$, for σ_e^2.

[8] In this notation, the time subscript on the measures γ^i has been surpressed. In a stationary equilibrium, these measures will be time-invariant.

$$\gamma^i(\theta')g^i_{\theta'}[(h^i(k)-\theta')\frac{(\sigma^i)^2+\tau^2}{(\sigma^i)^2}+\theta']. \tag{3.7}$$

The measure of firms at size k, then, is obtained by integrating over θ' and summing over i. Denoting this measure by $\lambda(k)$, we have

$$\lambda(k)=\sum_{i=0}^{\infty}\int_{\theta'}\gamma^i(\theta')g^i_{\theta'}[(h^i(k)-\theta')\frac{(\sigma^i)^2+\tau^2}{(\sigma^i)^2}+\theta']d\theta'. \tag{3.8}$$

Another important characteristic of the industrial structure is the rate of exit. As defined in (3.3) in the household's problem above, X_t is the mass of exiting firms. Exit comes in two forms. All firms face a constant, exogenous probability ξ of disappearing. In addition, some fraction of the firms coming into period t will have period t output so low as to drive their θ_t below the exit threshold. This latter, endogenous exit ultimately decreases in frequency with firm age as firms' beliefs about their abilities become more precise. As age increases, a cohort's exit rate approaches the exogenous probability ξ. Denoting the age-i cohort's exit threshold by $\widehat{\theta}^i$, we can write the mass of exiting firms as

$$\begin{aligned} X &= \sum_{i=1}^{\infty}\int_{\theta'}\gamma^i(\theta')\{\xi+(1-\xi)\int_{\theta''\leq\widehat{\theta}^i}g^i_{\theta'}[(\theta''-\theta')\frac{(\sigma^i)^2+\tau^2}{(\sigma^i)^2}+\theta']d\theta''\}d\theta'. \\ &\equiv \sum_{i=1}^{\infty}X^i \end{aligned} \tag{3.9}$$

Hence, the measure of endogenously exiting firms is the measure of firms whose production experience causes their beliefs to fall below the exit threshold for their age.

The total number (measure) of firms of a given age is $\Gamma^i=\int_{\theta'}\gamma^i(\theta')$, and the total number of firms in the population is $\Gamma=\sum_i\Gamma^i$. This aggregate measure evolves according to $\Gamma'=\Gamma-X+N$.

3.3. Equilibrium

A stationary equilibrium consists of exit rules given by $\widehat{\theta}^i$, for $i\geq 0$; investment decisions given by $k^{i*}(\theta'';r)$; an aggregate savings choice $S(r)$; a set of measures γ^i for $i\geq 0$; a mass of entrants N; and a price r. The decision rules $\widehat{\theta}^i$ and k^{i*} must solve the firm's problem conditional on the observation of current period output. Savings must solve the household problem. Market clearing requires that

$$S(r) = (1-\xi)\sum_{i=0}^{\infty}\int_{\theta'}\int_{\theta''\geq\widehat{\theta}^i}\gamma^i(\theta')g^i_{\theta'}[(\theta''-\theta')\frac{(\sigma^i)^2+\tau^2}{(\sigma^i)^2}+\theta']k^{i*}(\theta'';r)d\theta''d\theta'. \tag{3.10}$$

The right-hand side of (3.10) gives the total investment demand of surviving firms. Consistency conditions for the measures γ and N include

$$N = X; \tag{3.11}$$

$$\int_{\theta'}\gamma^0(\theta')d\theta' = N; \text{ and} \tag{3.12}$$

$$\gamma^i(\theta') = 0 \text{ if } \theta' < \widehat{\theta}^{i-1}. \tag{3.13}$$

That is, entry equals exit, the measure across all sizes of the youngest cohort is equal to the measure of entrants, and there are no firms below the exit threshold for their generation.

Finally, entrants must be indifferent between entering and staying out:

$$C_e = \int_{\theta_e} v(\theta_e, \sigma_e^2, r)g_\mu[(\theta_e-\mu)\frac{\sigma_s^2}{\sigma_s^2+\sigma^2}+\mu]d\theta_e, \tag{3.14}$$

where g_μ is the density for a normal distribution with mean μ and variance $\sigma_s^2+\sigma^2$ (here, μ and σ^2 are the parameters of the population distribution of θ, and σ_s^2 is the variance of the noise in an entrant's initial signal of θ).

3.4. Calibration

We would like the stationary equilibrium size distribution of firms to match certain characteristics of the U.S. manufacturing sector. We will choose certain parameters relating to the firm growth process so that aggregate exit rates and the proportion of firms that are young agree with statistics on exit rates of manufacturing firms. These growth process parameters are the entry cost C_e; the opportunity cost of continuing in operation w; the parameters of the population distribution of productive ability, μ and σ^2; the variance of the noise in an entrant's initial observation on θ, σ_s^2; the variance of the noise in individual output (ε_t), τ^2; and the exogenous exit rate ξ.

We normalize the mean of the population distribution to 0. Other parameters will affect entry and exit behavior by their magnitudes relative to μ, and choices of values for these parameters will be driven by how they jointly effect average entry and exit choices. The exogenous exit rate ξ is an exception, since this form of exit is not a matter of choice. One can note from equation (3.9), however, that the exit rate of a given age cohort, X^i, approaches ξ as i grows. Therefore we set ξ at the estimated exit rate for mature firms from Dunne, Roberts and Samuelson (1988). They estimate the five-year exit rate for large and mature firms to be about 20 percent. Accordingly, we set the quarterly rate ξ at 0.01.

For the remaining growth-process parameters, we seek to match the aggregate exit rate and the fraction of firms in the economy that are "young." The aggregate exit rate in U.S. manufacturing is about 1.9 percent per quarter (or about 40 percent per five years) as documented by Hopenhayn and Rogerson (1993). Evans (1987b) gives statistics on the age distribution of U.S. manufacturing firms. From his data we seek to match the fact that 30 percent of all firms are 6 years old or less, 33 percent of the firms are between the age 7 and 20. Finally, we also seek an age distribution in which the average age of firms is 12 years. These characteristics are achieved by the following parameter settings: $\sigma^2 = 0.04$; $\sigma_s^2 = \tau^2 = 0.01$; $w = 150$.

The entry cost C_e is something of a technicality, since the "supply" of entrants in this economy is perfectly elastic with respect to this cost. That is, in order for exactly the right number (measure) of firms to be willing to enter each period, C_e must be exactly equal to the expected value of being an entrant. Hence, after calculating the stationary distribution under the assumption that entry just balances exit, we set the cost equal to the expected value of being an entrant so that all potential entrants are indifferent at that point.

The remaining parameters in the model are more standard. We set the discount rate, β, at 0.989 and the depreciation rate, δ, at 0.02. We use the production function $f(k) = k^\alpha$ and set α at 0.33. Table 1 summarizes the parameterization.

Table 1. Calibrated Parameters

symbol	variable	value
μ	population mean of firms' productivity	0
σ^2	population variance of firms' productivity	0.04
σ_s^2	variance of the noise in entrants' initial productivity	0.01
τ^2	variance of the noise in firms' productivity	0.01
ξ	exogenous exit rate	0.01
w	exit value	150
α	capital income share of total income	0.33
β	discount rate	0.989
δ	capital depreciation rate	0.02

With these parameters, the stationary equilibrium industry structure is depicted in Figures 1 through 4. Figure 1 depicts the joint distribution of size and age. Figures 2 and 3 give the age and size distributions separately, while Figures 4 gives the distributions of size at various ages. Notice first how the size distribution shifts to the right with increasing age, reflecting the fact that $\widehat{\theta}^i$ is increasing with age (i). This, together with the declining density of the age distribution, contributes to the overall skewness of the size distribution. All of these characteristics are present in the data for U.S. manufacturing. The model produces an industrial structure that is qualitatively similar to that of U.S. manufacturing.

4. Aggregate Shocks and Exogenous Prices

The differential sensitivity of firms of different sizes or ages to aggregate fluctuations rests on the inability to fully filter out aggregate shocks in updating their beliefs about their own ability parameters. The description of a firm's decision problem assumes that the firm chooses its investment after observing its current output but before knowing aggregate output for the current period. If the firm knew aggregate output, it could update its beliefs solely on the observation of $\theta + \varepsilon_t$, and the learning rule would be identical to that in the previous section's stationary case.

Since firms buy their investment goods at a competitively determined market price, the price will reveal the aggregate state if there are no other sources of price fluctuation. We think the presence of a signal extraction problem when firms do not fully know the aggregate state is a desirable and reasonable feature for the type of learning-based model we are examining. While the model is one with a single price and a single source of aggregate uncertainty, the notion that firms face a nontrivial filtering problem is consistent with a more complicated world. One can imagine, for instance, an environment in which a firm faces some shocks that are common to a relatively narrow group (say, its industry) and other shocks that are common to a broader group. If industry-level shocks affect industry demand for inputs that are purchased in markets that cross industry boundaries, then the prices of those inputs will be functions of multiple shocks, some of which are not relevant to the individual firm's forecast of its industry's future state. In such an environment, a firm would face an inference problem much like the filtering problem presented above (complicated somewhat by the presence of additional shocks).

One way to proceed, then, would be to add both the technology shock z and an aggregate shock to the supply of investment goods (for instance, through a preference shock) to the model of the previous section. Here, we take the initial step in that direction by examining the behavior of the model with exogenous prices by examining an industry equilibrium for an industry facing a perfectly elastic supply curve (that is, for a small industry). While the U.S. manufacturing sector most likely does not satisfy this condition, we think the results of this exercise are likely to be robust along the key dimension: differences in behavior between large and small firms.

While the focus of the Gertler and Gilchrist's discussion is the differential responses of small and large firms to monetary shocks, the evidence they examine also suggests a more general difference in responses to aggregate shocks. This difference is demonstrated in Figures 5 and 6. Figure 5 illustrates the relative growth rates of sales over business cycles. For the most part, it appears that small firms decline sharply relative to large firms during recessions. Figure 6 reconfirms the pattern by showing the average sales growth experience of small and large manufacturing firms after the onset of a recession. The financial frictions interpretation of this difference is that the growth of small firms is constrained by their limited access to external funds. The limitation becomes all the more severe when internal funds are squeezed by a negative aggregate shock. Hence, the difference between large and small firms should be more pronounced during economic downturns, a feature that Gertler and Gilchrist find to be true in the data.

Figures 7 through 10 show the behavior of large and small firms in the model with an aggregate shock and a constant interest rate (equal to $1/\beta$). The aggregate shock follows a two-state Markov process with $z^l = -0.007$, $z^h = 0.007$, $p_{hh} = p_{ll} = 0.975$, and $p_{lh} = p_{hl} = 0.025$. Figures 7 and 8 show the average responses to a transition from z^h to z^l, while figures 9 and 10 show the average responses to a transition from z^l to z^h. The classification of firms into the large and small groups is achieved by setting a threshold size such that firms below the threshold account for about 38 percent of output in the steady-state equilibrium.[9]

Figures 7 and 8 indicate a greater response to the downturn among small firms. The difference across firm sizes is related to the difference across firm ages. The response in output works through the response in investment. At the realization of z^l(following z^h in the previous period), firms experience unexpectedly low (firm-level) output. The firm's forecast error is represented by $m_t - E\left[z_t \mid z_{t-1} = z^h\right] - \theta_{t-1}$. Given the persistence in the z process, this forecast error is large in the transition period. A mature firm (with small σ^2_{t-1}) attributes relatively little of the error to news about its firm-specific productivity and, accordingly, makes a relatively large adjustment to its expectation for next period's z. Therefore, its current period investment falls because of adjusted expectations about the aggregate state. On average (across many realizations of such a transition), z remains at z^l for a number of periods. After the transition, then, the firm will be expecting an increase in z, so that investment rises after the transition period.

For a young firm, the response to the transition in the aggregate state is compounded by the fact that the firm attributes more of its forecast error to bad news about its own ability. While such attribution diminishes the firm's adjustment of expectations about the aggregate state, the overall effect is a greater reduction in investment (since less of the error is also attributed to the pure error term ε_t). Since the small firm group has a greater share of young firms than does the large firm group, the small firm response will look more like the young firm response. After aggregate output is observed, all firms' expectations about z are the same, so there is an investment rebound for all firms as they realize that they are at a low point in the z process.

Differences in the aggregate behavior of large and small firms are also affected by exit behavior. Recall that small, young firms (with greater σ^2_{t-1}) are more inclined to stay in business in hopes that they will receive good news about their abilities, as compared to small, old firms. This difference shows up in exit thresholds that rise with age and reflects the option value of staying in business. The option value varies with the aggregate state. In particular, a low z tends to flatten the value function at any given $(\theta_{t-1}, \sigma^2_{t-1}, m_t)$, which smooths out the nonconvexity in the firm's decision problem, reducing the importance of the exit option and the rate at which the exit threshold grows with age. The result is that relative number of small firms in the population of firms will tend to be smaller in periods

[9] The cutoff size for small firms in Gertler and Gilchrist (1994) is determined so that small firms output account for one-third of the total output.

of low z.

5. Concluding Remarks

This paper has examined a model of firm-level dynamics absent of financial market imperfections. The model predicts differences in behavior across size and age classes that are qualitatively similar to observed cross-sectional differences often attributed to financial frictions. There is obviously other empirical evidence that the model is not equipped to address, such as the asymmetries of the cross-sectional behavior over business cycles. In other words, adverse aggregate shocks affect small firms more than positive ones do. This asymmetry could be due to the nature of the underlying shocks, or the presence of the financial constraints.

This paper contributes to the discussion of the importance of financial factors for aggregate behavior of the economy. It provides a model that might serve as a useful benchmark against which models that incorporate financial frictions can be measured. The results clearly indicate that any assessment of the aggregate importance of financial frictions should focus on identifying implications of such frictions beyond those of a nonfinancial theory of the size distribution and growth of firms.

Appendix: Computation Procedure

We describe first the computation procedure for solving the stationary equilibrium and then the procedure used to compute the equilibrium with aggregate uncertainty.

1. Solving for the stationary equilibrium

We use a method of successive approximation to solve numerically for the stationary equilibrium for this economy. The iterative procedure consists of several steps. From the first order conditions of the household problem, we know that the equilibrium price for capital, i.e. the interest rate, is equal to one over the discount rate. Given the price, we then use value iteration to solve the functional equations defined in the firm problem. Next, the invariant distribution corresponding to these decision rules is found by iterating on equation (3.5).

The method used to solve for the decision rules involves discretizing the state space by choosing the grids for firms belief θ and their age σ^2. We also need a grid for m, the sum of firms individual technology θ and the idiosyncratic shock ε in order to decide the state of a firm in the next period. We set the upper bound for firms' age at 140 quarters, at which level the variance is 0.00017.[10] The determination of the bounds for θ and m is more involved. We start by finding the ranges of θ and m for firms at each age. For entrants, we choose $[\underline{\theta}_0, \overline{\theta}_0]$ so that the probability of getting a value outside of this range is 0.0001. Similarly, we choose such a range for the idiosyncratic shock ε $[\underline{\varepsilon}, \overline{\varepsilon}]$. Then the range of m for the entrants is $[\underline{m}_0, \overline{m}_0]$, where $\underline{m}_0 = \underline{\theta}_0 + \underline{\varepsilon}$ and $\overline{m}_0 = \overline{\theta}_0 + \overline{\varepsilon}$. Denote the range of the belief θ for firms at age 1 as $[\underline{\theta}_1, \overline{\theta}_1]$, we have $\underline{\theta}_1 = \underline{\theta}_0 + \frac{(\sigma^0)^2}{(\sigma^0)^2+\tau^2}(\underline{m}_0 - \underline{\theta}_0)$, and $\overline{\theta}_1 = \overline{\theta}_0 + \frac{(\sigma^0)^2}{(\sigma^0)^2+\tau^2}(\overline{m}_0 - \overline{\theta}_0)$. The range of m for firms of age 1 is denoted as $[\underline{m}_1, \overline{m}_1]$, where $\underline{m}_1 = \underline{\theta}_1 + \underline{\varepsilon}$, and $\overline{m}_1 = \overline{\theta}_1 + \overline{\varepsilon}$. We continue this process until we get the ranges for firms of age 140. Obviously, the minium level for θ is $\underline{\theta}_{140}$, and the maximum level for θ is $\overline{\theta}_{140}$; the minimum level for m is $\underline{m}_{140}$ and the maximum level for m is $\overline{m}_{140}$. We choose 20 grid points for both θ and m. All the continuous normal distributions are substituted by their discrete counterparts. For example, if we use n discrete points within a given range to approximate a normally distributed variable x and denote the value of these discrete points as $x_1, x_2, ..x_n$,

[10] Firms' variances at each age is determined as follows:
$(\sigma^0)^2 = \sigma_e^2$,
$(\sigma^1)^2 = \frac{(\sigma^0)^2\tau^2}{(\sigma^0)^2+\tau^2}$,
...
$(\sigma^{i+1})^2 = \frac{(\sigma^i)^2\tau^2}{(\sigma^i)^2+\tau^2}$,
...
$(\sigma^{140})^2 = \frac{(\sigma^{139})^2\tau^2}{(\sigma^{139})^2+\tau^2}$.

then the probability of x having each value x_i would be $\text{Prob}(x = x_i) = prob(x_{i-1} \leq x \leq x_i)$ ($x_{-1} = -\infty$ and $x_{n+1} = +\infty$).

The optimal value functions and decision rules for this finite state discounted dynamic programming problem are obtained by successive approximations. This approach involves starting with initial approximations for the value functions and using them to obtain a subsequent approximation by computing the right side of the firm value function. This process continues until the sequence of value functions so obtained converges.

Given the state transition function implied by the equilibrium decision rules is ergodic, there exists a unique invariant distribution. To compute the invariant distribution, we begin with an initial approximation and evaluate equation (3.5) using decision rules we obtained. Note that this only gives us the measure of firms aged 1 to 140. The measure of entrants is 1 subtracted by the total measure of firms aged 1 and 140, and their distribution in the state space is determined by their initial believes. The result is used as the next candidate, and the process is repeated until successive approximations are sufficiently close. At this point a steady-state equilibrium is computed.

2. Solving for the equilibrium with aggregate uncertainty

We only solve for a partial equilibrium with aggregate uncertainty, i.e., we set the equilibrium interest rate to be the steady-state interest rate.

The main difference between this environment and one without aggregate uncertainty is that both z, the aggregate shock, and m, the sum of the individual technology, the idiosyncratic shock, and the aggregate shock, are now part of the state variables. As stated in the main text, for computational simplicity the aggregate shock z is assumed to take only two values. The grid point for variance is chosen exactly as in our computation of the stationary equilibrium. The determination for the range of θ and m is also similar to the one used when solving the stationary equilibrium, except that $\underline{m}_i = \underline{\theta}_i + \underline{\varepsilon} + z_l$, $\overline{m}_i = \overline{\theta}_i + \overline{\varepsilon} + z_h$ and $\underline{\theta}_{i+1} = \underline{\theta}_i + \frac{(\sigma^i)^2}{(\sigma^i)^2+\tau^2}(\underline{m}_i - \underline{\theta}_i - z_h)$, $\overline{\theta}_{i+1} = \overline{\theta}_i + \frac{(\sigma^i)^2}{(\sigma^i)^2+\tau^2}(\overline{m}_i - \overline{\theta}_i - z_l)$. Again the minimum level of θ is $\underline{\theta}_{140}$ and the maximum level of θ is $\overline{\theta}_{140}$; the minimum level of m is $\underline{m}_{140}$ and the maximum level of m is $\overline{m}_{140}$. Also we approximate all the continuous distributions with their discrete counterparts.

Except for the calculations of the period profit function, the optimal value functions are solved exactly as those in the stationary equilibrium.

To compute our impulse responses, we simulate the model according to the optimal decision rules obtained for a sufficiently long period, 2000 periods as in our experiment, starting with an arbitrary distribution. During the simulation, at each time period t, we need to keep track of the aggregate shock that is known to the firms z_{t-1} and the one z_t, which is not yet known to the individual firm, according to which the output of each firm is realized. The transition of a firm from one state to another over time is similar to that in the above section. We then simulate the model for another 2000 periods, setting the sequence of the aggregate shock to be a series of low values (or high values) with a high (low) value every 20 periods (depending on how long we want the sequence). We discard the first 2000 simulations, then compute the average response of the system in which the aggregate shock transited from its low value to its high value across all instances in our simulation, or vice versa. These average responses are analogous to impulse response functions.

References

[1] Cooley, T.F. and V. Quadrini (1998). Monetary Policy and the Financial Decisions of Firms. Manuscript.

[2] Davis, S. J., C. C. Haltiwanger and S. Schuh (1996). *Job Creation and Destruction.* Cambridge, Mass. MIT Press.

[3] Dunne T., M.J. Roberts and L. Samuelson (1988). Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *Rand Journal of Economics*, 19, 495-515.

[4] Evans, D.S. (1987a). The Relationship between Firm Growth, Size and Age; Estimates for 100 Manufacturing Industries. *Journal of Industrial Economics*, 35, 567-582.

[5] Evans, D.S. (1987b). Tests of Alternative Theories of Firm Growth. *Journal of Political Economy*, 95, 657-674.

[6] Fazzari, S.M., R.G. Hubbard and B.C. Petersen (1988). Financial Constraints and Corporate Investment. *Brookings Papers on Economic Activity*, 88, 141-195.

[7] Fisher, J.D.M. (1996). Credit Market Imperfections and the Heterogeneous Responses of Firms to Monetary Shocks. Working Paper 96-23, Federal Reserve Bank of Chicago.

[8] Gertler, M. and S. Gilchrist (1994). Monetary Policy, Business Cycles and the Behavior of Small Manufacturing Firms. *Quarterly Journal of Economics*, 109, 309-340.

[9] Gomes, J. (1997). *Financing Investment.* Ph.D. thesis, University of Rochester.

[10] Hall, B. (1987). The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector. *Journal of Industrial Economics*, 35, 583-599.

[11] Hopenhayn, H. A. (1992). Entry, Exit and Firm Dynamics in Long Run Equilibrium. *Econometrica*, 60, 1127-1150.

[12] Hopenhayn, H. A. and R. Rogerson (1983). Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy*, 101, 915-938.

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

[14] Li, W. (1998). Capital Constraints and Firm Heterogeneity over Business Cycles. Manuscript.

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

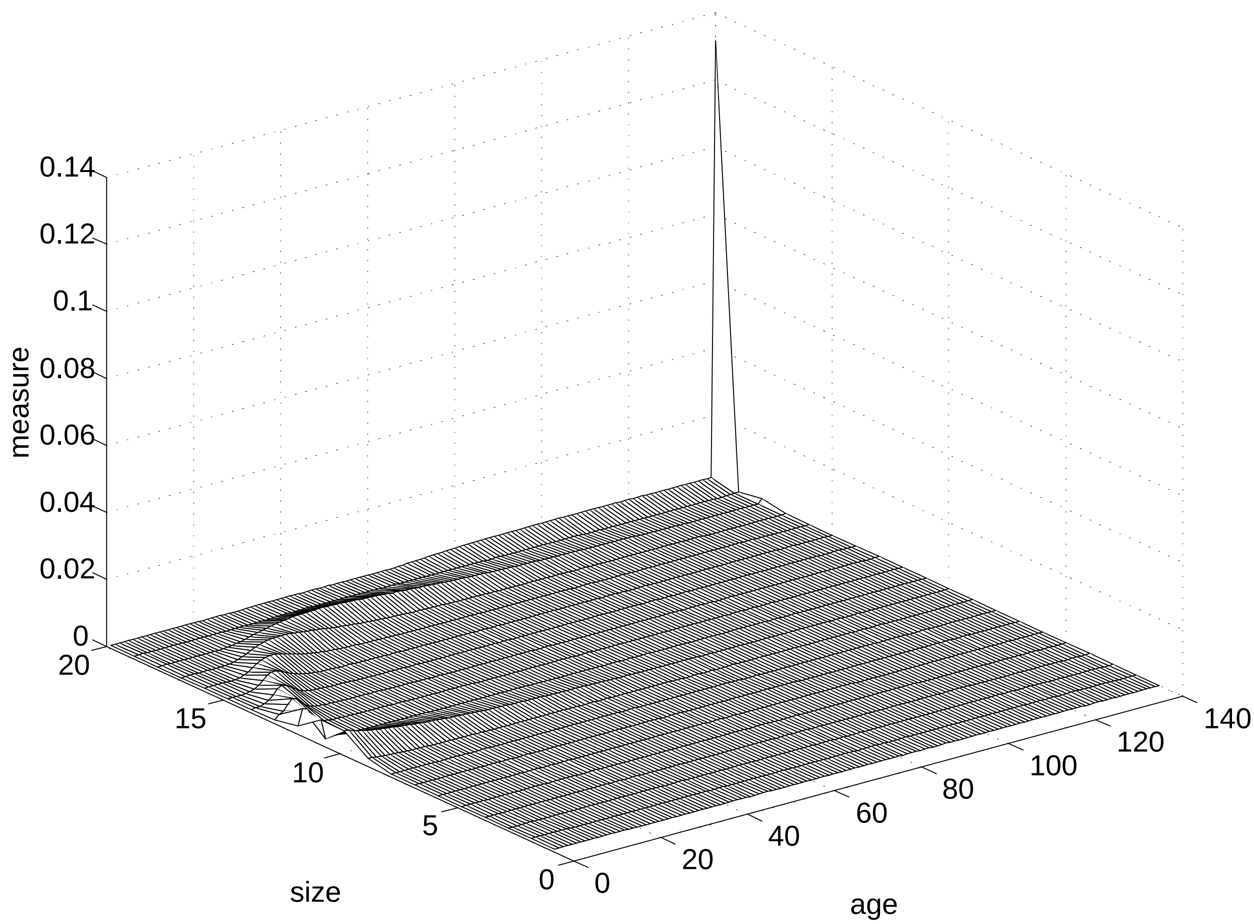

Figure 1: Joint Distribution of Size and Age

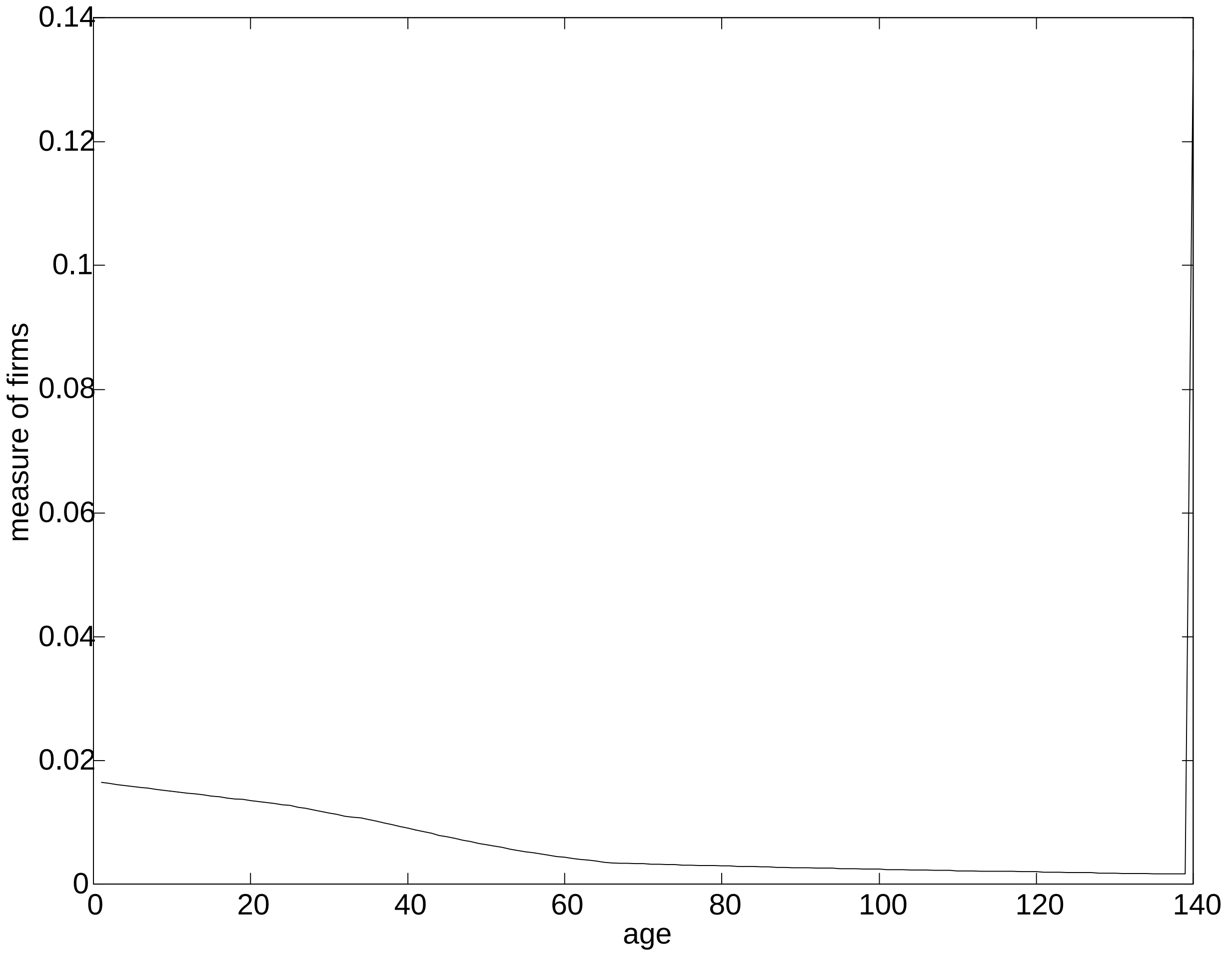

Figure 2: Age Distribution

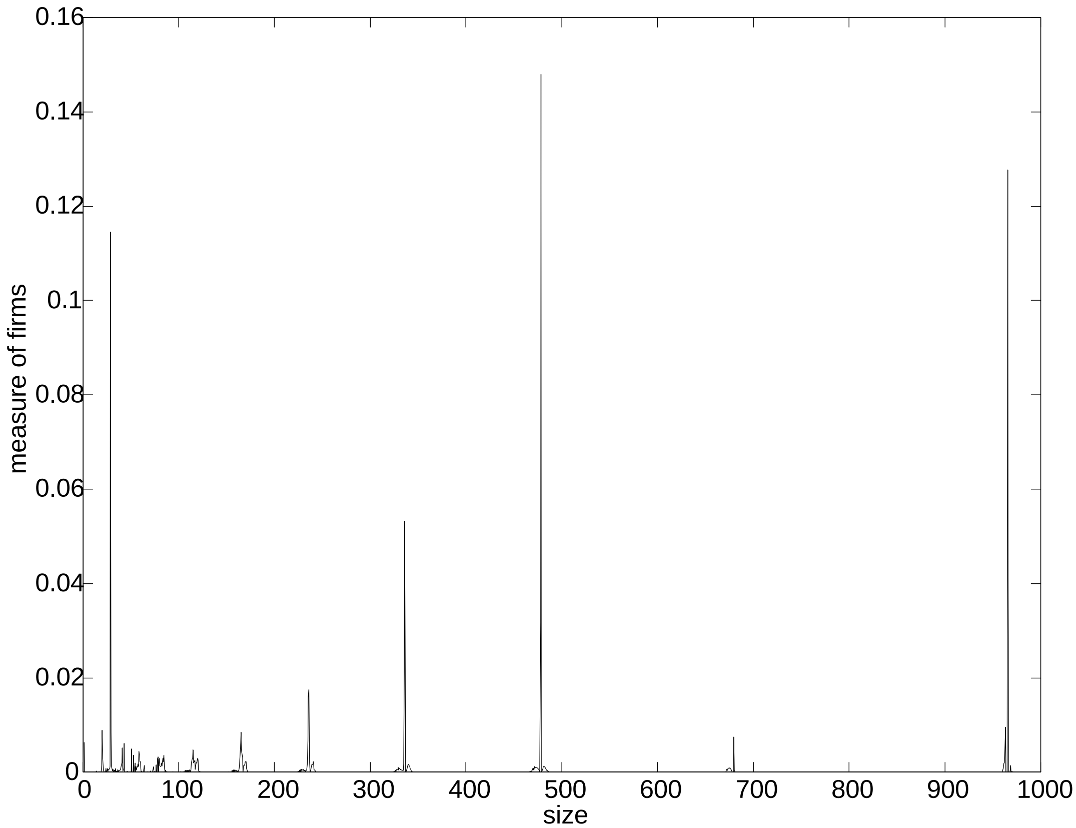

Figure 3: Size Distribution

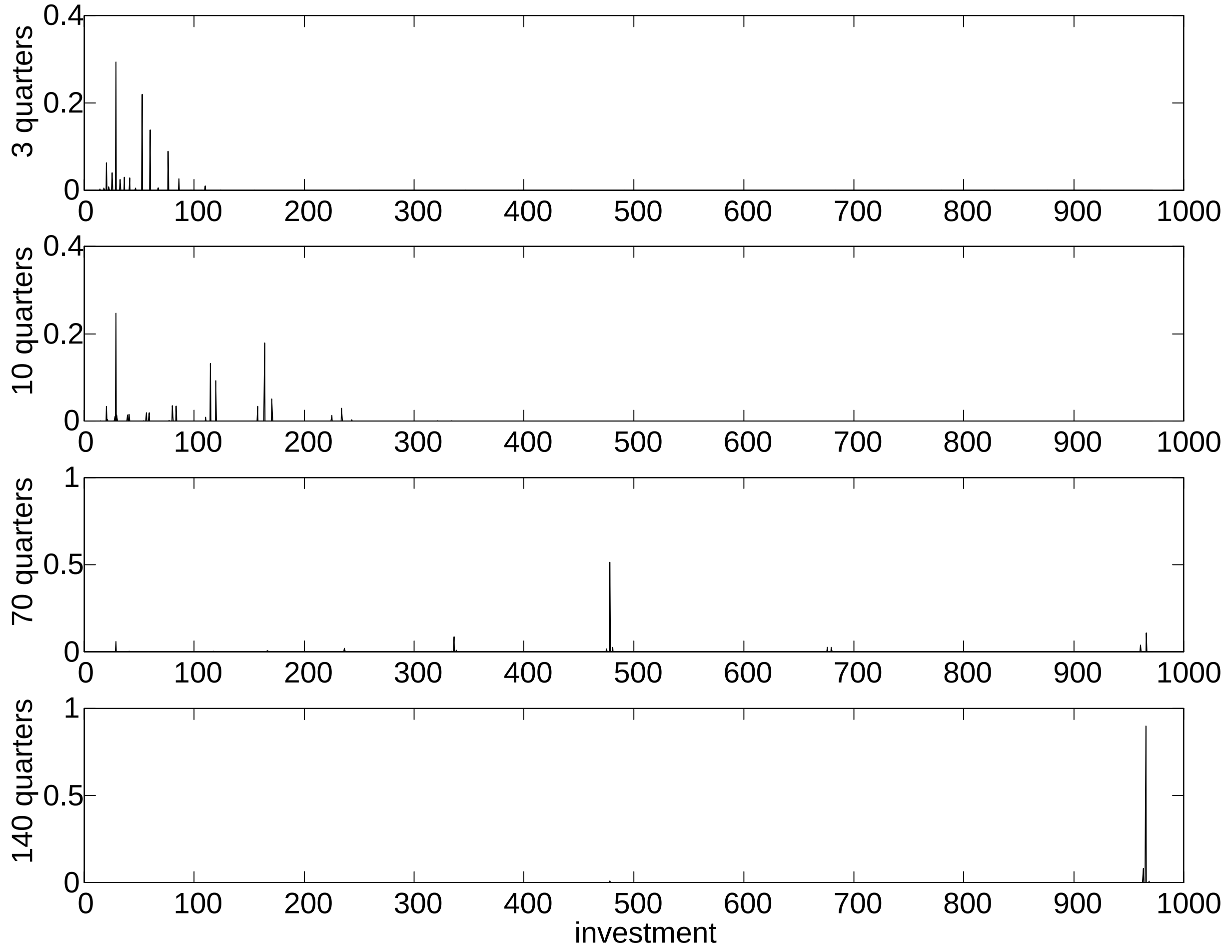

Figure 4: Size Distribution at Various Ages

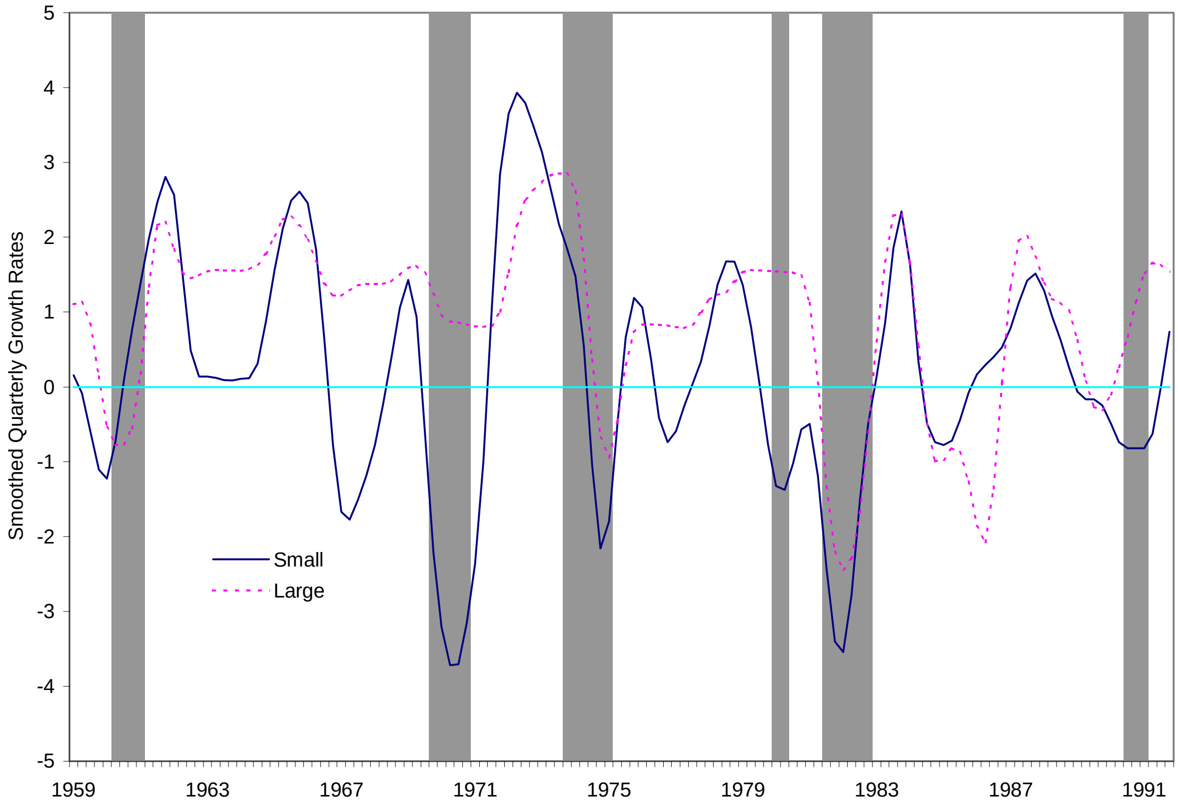

Figure 5: Small and Large Firm Sales Over Business Cycles (the growth rate series are smoothed using an S-plus program; the shaded regions indicate the NBER recessions)

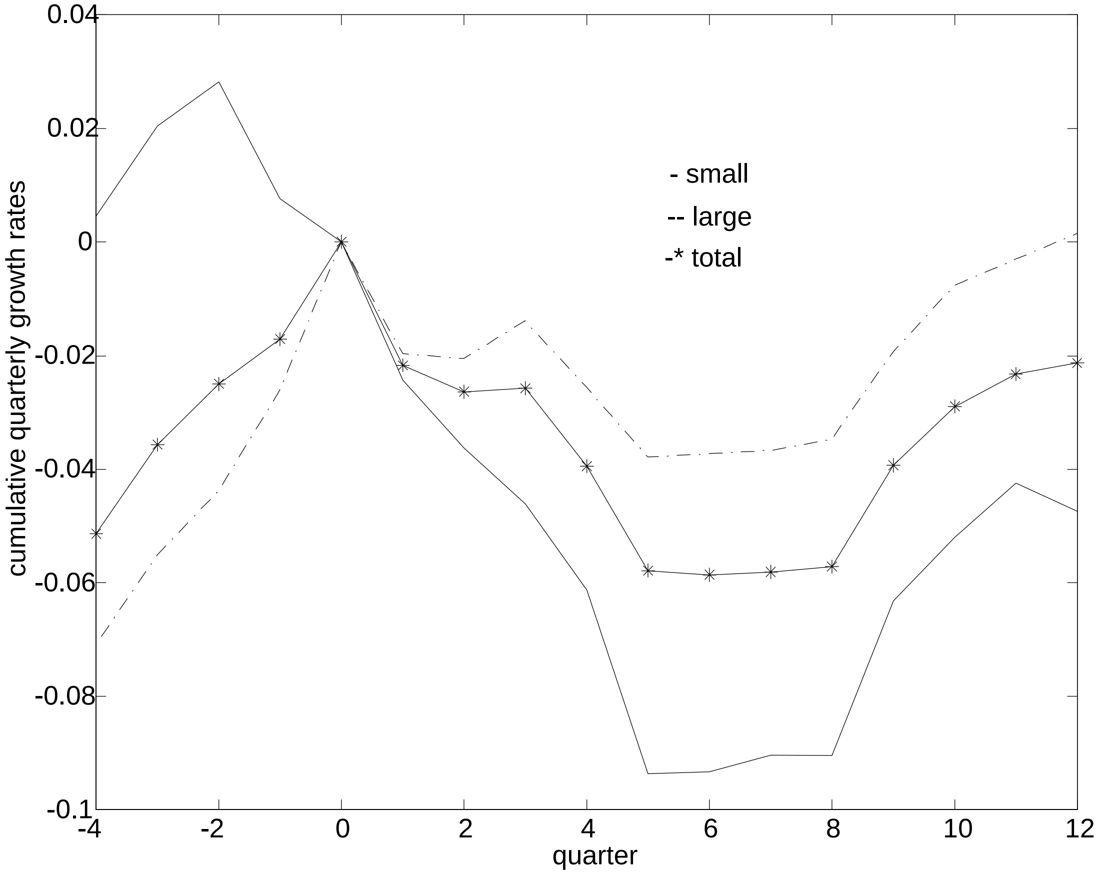

Figure 6: Response of Large and Small U.S. Manufacturing Firms to Economic Downturns

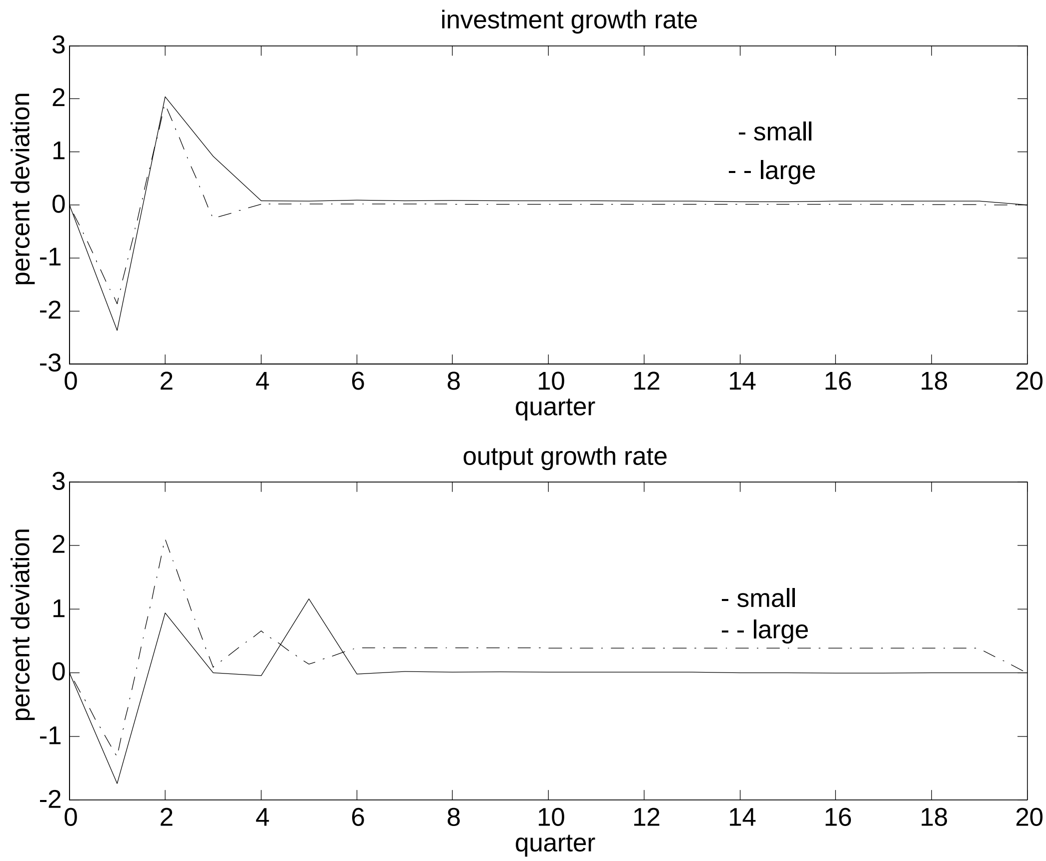

Figure 7. Average Response of Investment and Output Growth to an Economic Downturn

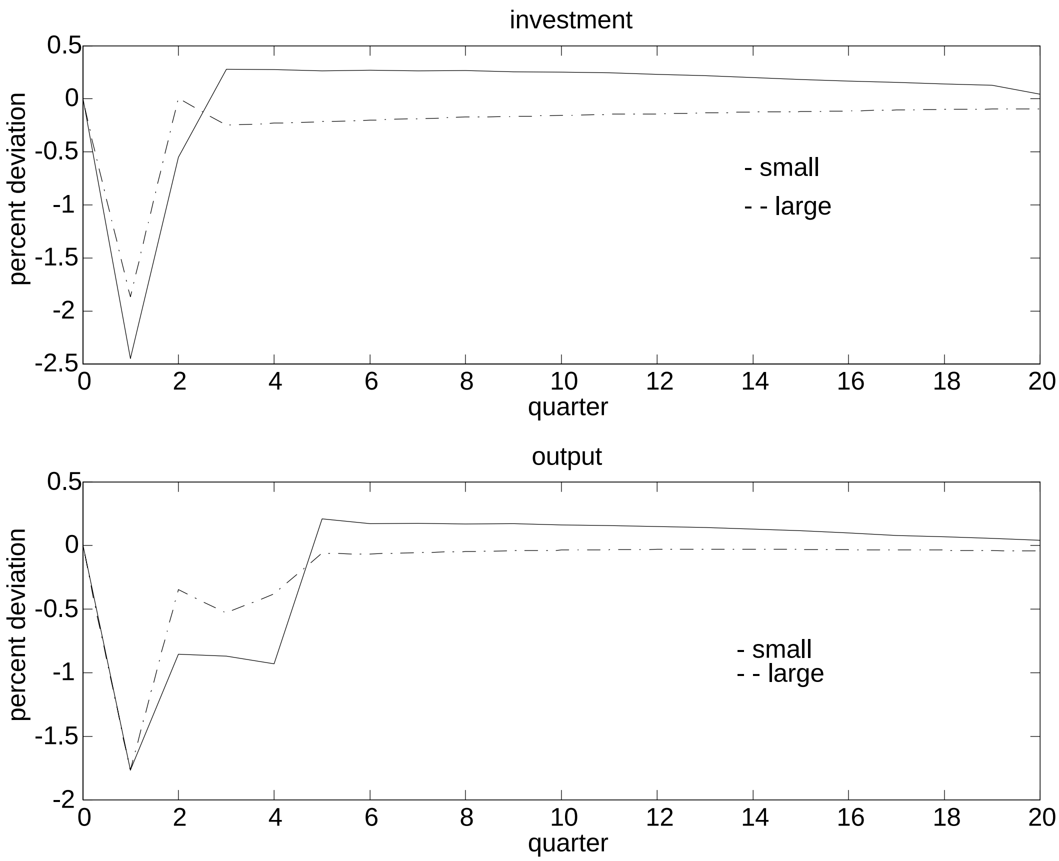

Figure 8. Average Response of Investment and Output to an Economic Downturn

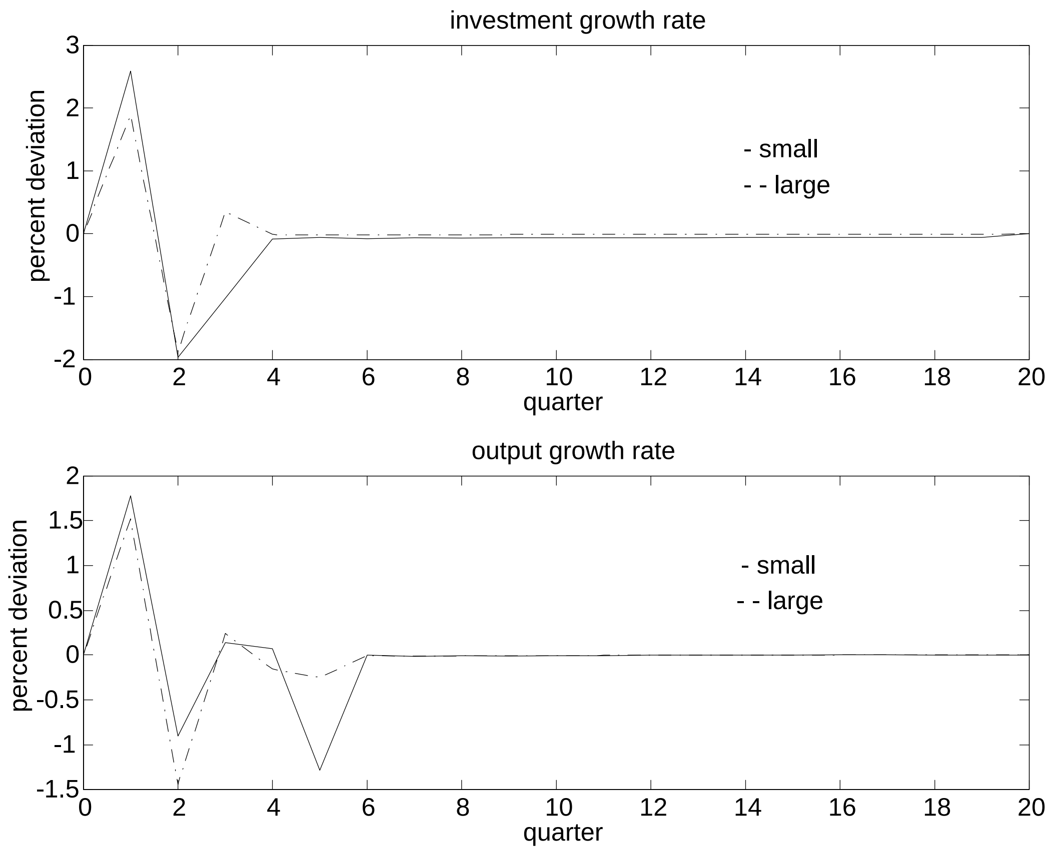

Figure 9. Average Response of Investment and Output Growth to an Economic Boom

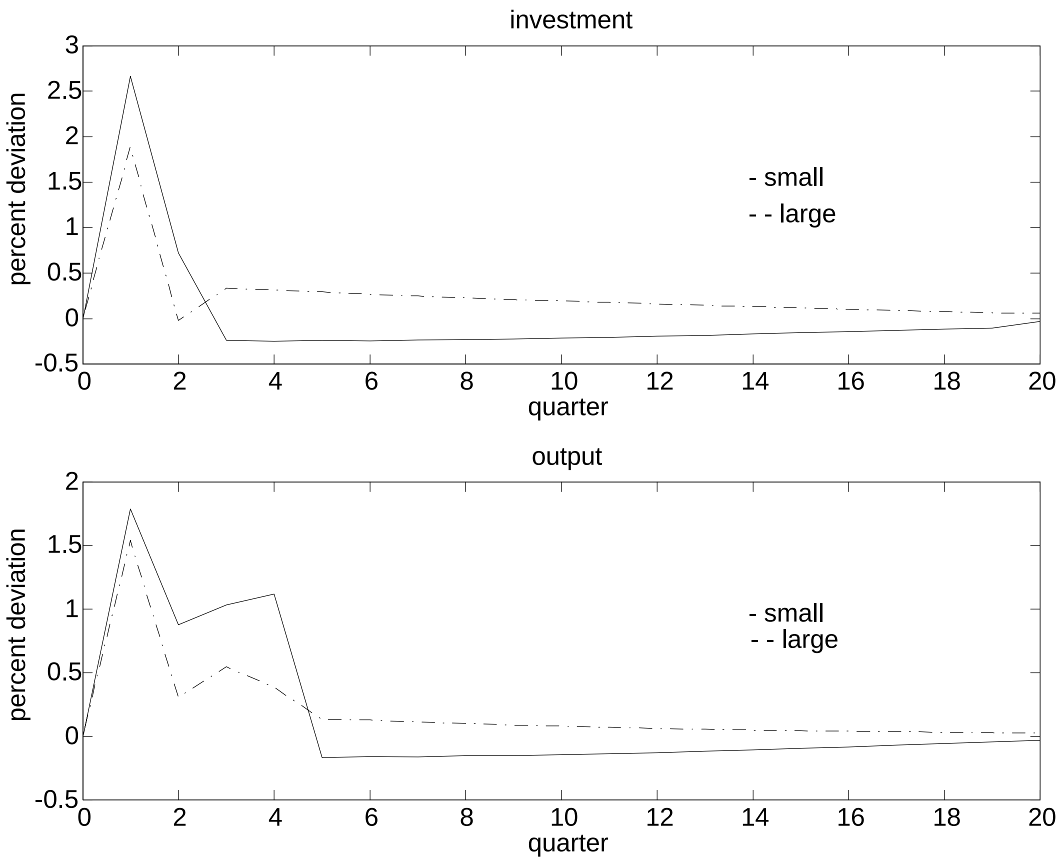

Figure 10. Impulse Response of Investment and Output Levels to an Economic Boom