Learning By Investing
Embodied Technology and Business Cycles

Geng Li

2007-15

Learning By Investing

Embodied Technology and Business Cycles

Geng Li [*]

Federal Reserve Board

December 2006

KEYWORDS: Embodied Technology, Learning, Overinvestment

JEL Classification: E22, E32

Abstract

In the last decade of the 20th century, the U.S. economy witnessed a persistent and substantial increase in private investment. The boom was sharply reversed in 2001, and a great deal of evidence suggests that the capital stock had become excessive. Standard equilibrium business cycle models have difficulties in predicting the investment boom and overshooting. An embodied technology model is constructed to replicate the pattern of investment boom and collapse. Unlike previous models of embodiment, the present model assumes that new technology increases the productivity of capital of all vintages, but only new capital can facilitate the adoption of the new technology. Further, although agents in this model know about the advent of a new technology, they have imperfect information about its magnitude. Agents learn the magnitude by investing in new capital. I present a sufficient condition for having a persistent investment boom and overshooting. I also solve the model numerically in a dynamic general equilibrium (DGE) setup. The model presented in this paper extends the standard DGE business cycle models in two ways: First, it presents a strong internal propagation mechanism with respect to technology shocks; second, it generates endogenous recessions without invoking technological regress. The model also offers a possible explanation on why consumption growth was strong during the last recession.

1

1 Introduction

The most recent business cycle demonstrated many unique characteristics that are at variance with previous cycles. One of the most remarkable features of the economic expansion in the 1990s was the persistent and substantial increase in private investment, in terms of both its level and its share in GDP. Investment growth stopped abruptly in the second half of the year 2000. In the subsequent years, investment decreased sharply and at the same time the economy as a whole experienced a recession.

One of the distinguishing characteristics of this investment boom is that the economy may have accumulated too much capital and overinvested, especially in certain sectors. For instance, more than 90 percent of the optical fiber cables installed during the 1990s were left unused in the years that followed, resulting in thousands of miles of "dark fiber."[1] At the moment when the investment boom was about to crash, the overinvestment phenomenon had already drawn much attention from entrepreneurs and policy makers. On July 14, 2000, close to the peak of the boom, Microsoft President and CEO Steve Ballmer commented that, "A lot of people are overinvesting in dot-com start-ups ... There has been a hysteria. There is too much money chasing Internet ideas in the short run."[2] Similarly, one year later, after the investment spending boom collapsed and the economy was in a recession, the vice chairman of the Federal Reserve Board, Roger W. Ferguson, commented on July 18, 2001, that " ... for a variety of reasons ... firms may be holding considerably more capital now than they would prefer . . . although it is difficult to determine how large the overhangs of capital might be at present, they seem likely to exert at least a modest amount of drag on the economy over the near term, even as growth picks up."[3]

Another aspect of the last cycle is that consumption did not show as much weakness as it

---

[1]Such a level of "dark fiber" is also partially due to the advent of a new signal transmitting technology, which dramatically increased the per-cable transmitting capacity.

[2]Atlanta Business Chronicle. http://atlanta.bizjournals.com/atlanta/stories/2000/07/10/daily19.html.

[3]Speech to the Charlotte Economics Club. http://www.federalreserve.gov/boarddocs/speeches/2001/20010718/default.htm.

had (on average) in prior downturns. Consumption growth, especially for durable goods, was much stronger than GDP growth during the years of 2001 and 2002.

In this paper I construct an embodied technology model to explain why the last business cycle exhibited the unique characteristics described above. In my model, new technology has to be embodied into new capital goods before it can increase total factor productivity (TFP). Different from the traditional embodiment models, *a la* Solow (1959), my model assumes that new technology increases the productivity of capital of all vintages, instead of only the new vintage. The more the economy has invested in new capital, the higher is the TFP, up to a point. In addition, I assume that although the agents know about the advent of a new technology, they have imperfect information about the magnitude of the new technology shock. Agents have to learn the magnitude of the new technology by investing in new capital goods.[4] They observe the output and evaluate whether they have invested beyond the optimal amount. If they have not, they revise up their beliefs about the magnitude of the technological innovation and invest more in subsequent periods, hence an investment boom follows. If they find that investment has overshot, they will sharply reduce subsequent investment, and thereby trigger a recession. During the recession, resources are reallocated from investment to consumption. Therefore, even though GDP does not increase, we may still observe healthy growth in consumption.

I first set up the firm's capital demand problem under the proposed embodiment and learning mechanisms and contrast the capital demand in this model with that in the standard models. Then I provide a sufficient condition on the prior belief over the technology shock distribution that generates persistent investment booms and overshooting. I also solve the model numerically in a stochastic dynamic general equilibrium setup to study the business cycle dynamics. The model can qualitatively replicate what happened during the last business cycle. I find that after a reasonably large permanent technology shock, investment can keep increasing for as long as nine years. Naturally, the model also predicts an output boom as persistent as the investment boom.

---

[4]The learning mechanism adopted in this paper is similar to that in Zeira (1987), Rob (1991) and Barbarino and Jovanovic (forthcoming). In later sections I will show, however, that their models differ in important ways from mine.

Investment overshoots at its peak and is sharply reduced subsequently. When investment is cut back, output stops growing. Under certain calibrations, output decreases in absolute terms. However, during the recession, consumption rises to a higher level, even relative to the level in the boom.

Standard dynamic general equilibrium (DGE) models have often been criticized on two grounds. First, in many DGE models, if the shock process is not very persistent, the output, investment, and consumption dynamics typically are not very persistent either. Put differently, these models lack strong internal propagation mechanisms.[5] Second, DGE models usually demand certain levels of technological regress to generate significant recessions.[6] This paper addresses these two problems. In my model, output does not jump in response to a technology shock, as the economy has to invest in new capital to take advantage of the new technology. Because the agents have to learn the magnitude of the technology shock, they are cautious in making investment decisions before they have learned much about the underlying technology. Consequently, GDP growth is gradual, which stretches out the length of booms. In addition, the model generates endogenous recessions. Under certain assumptions, the investment boom caused by the new technology will eventually overshoot the optimal level. It is the favorable technology shock that leads the economy into a recession. In this sense, *a favorable technology shock already contains the seed of a future recession.* There is no need for any negative technology shocks to trigger recessions.

Finally, it is worthwhile to point out that if one technology shock can generate a persistent investment and output boom as well as sizable overinvestment, this shock should be sufficiently large. It should have a significant contribution to TFP. Such technology shocks can be mo-

---

[5]An important critique of the real business cycle (RBC) models, an important version of the DGE models, on this ground is Cogley and Nason (1995). They find that most RBC models have very weak internal propagation mechanisms to produce both positively correlated GDP growth and a mean-reverting component. For permanent shocks, they find that, for most RBC models, output will jump to the maximum response and then begin to decline slowly. However, empirically derived impulse response functions show quite a different picture. After a technology shock arrives, output jumps only mildly, and in the subsequent quarters, output keeps increasing gradually.

[6]As Beaudry and Portier (2004) noted, "it is well known that the standard real business cycles models have difficulties explaining recessions–at least the size observed in postwar U.S. data — without invoking technological regress."

tivated as General Purpose Technology (GPT)-shocks, which have been found empirically to impact most industries of the economy.[7] The point of view that the recent boom was at least partially driven by a GPT shock finds support in many recent empirical studies.[8]

The paper is organized as follows. Section 2 presents a model of embodied technology with imperfect information and learning. I provide a sufficient condition under which a new technology leads to a persistent acceleration of investment. Section 3 calibrates the functional form of embodiment and learning and solves the model numerically in a DGE setup. Section 4 discusses the related literature. Section 5 documents more carefully the investment boom and collapse and other characteristics of the 1990s. I also briefly discuss other historical episodes that share similar characteristics. Section 6 provides some concluding remarks and directions for future research.

2 The Model of Embodiment and Learning

2.1 Production Function with Embodiment and Learning

I study investment dynamics in a model in which new technology has to be embodied in new capital before it can increase TFP, and the agents have only imperfect information about the new technology. First I summarize the assumptions I make and introduce the production function under these assumptions.

Assumption 1 *After a new technology arrives, firms need to invest in new capital to use the new technology. Before new capital is accumulated, TFP does not rise.*

Assumption 2 *After new capital is accumulated, TFP increases for all capital, including the capital that existed before the technological innovation.*

Assumption 3 *Bounded by the underlying technological innovation, TFP increases with the amount of new capital that has been accumulated relative to the stock of old capital.*

---

[7]See Jovanovic and Rousseau (2005) for a thorough treatment of GPT.
[8]See, for instance, Oliner and Sichel (2000).

Assumption 4 *Firms know when a technological innovation has arrived, but do not know the magnitude of the innovation. However, firms know the distribution the magnitude has been drawn from.*

Assumption 5 *The technology shock affects all firms in the economy uniformly. However, firms are able to learn the magnitude of the shock only from their own activities. There is no information externality and spillover. One firm can not learn the magnitude of the shock by analyzing another firm's information.*

A technological innovation arrives at $t = 0$. Before then, the production function of a typical firm is

$$Y_0 = A_0 \times K_0^\alpha. \tag{1}$$

where Y_0 is output, K_0 is the pre-shock level of capital, and A_0 is the pre-shock level of technology. The post-shock level of technology is denoted by, A. The shock is multiplicative,

$$A = (1 + \epsilon) \times A_0. \tag{2}$$

We normalize A_0 to 1, and hence $A = 1 + \epsilon$. After the technology shock arrives, the production function becomes

$$Y_t = \widetilde{A}_t \times K_t^\alpha \qquad \forall \quad t > 0 , \tag{3}$$

$$\widetilde{A}_t = \min \left[A, \ \Psi \left(\frac{K_{new,t}}{K_{old,t}} \right) \right]. \tag{4}$$

Capital accumulations follow

$$
\begin{aligned}
K_t &= K_{new,t} + K_{old,t} \\
K_{new,t} &= I_{new,t-1} + (1 - \delta)K_{new,t-1} \\
K_{old,t} &= I_{old,t-1} + (1 - \delta)K_{old,t-1} \qquad \forall \quad t > 0,
\end{aligned} \tag{5}
$$

6

and

$$K_{new,0} = 0,$$

$$K_{old,0} = K_0. \tag{6}$$

In production function (3), $\widetilde{A}_t$ is the effective TFP level at period t. It is equal to the minimum of A, the underlying post-shock level of technology, and $\Psi\left(\dfrac{K_{new,t}}{K_{old,t}}\right)$, an embodiment function. K_t is the capital stock combining old and new capital. Old capital is available before and after the arrival of the new technology, whereas new capital is available only after the technological innovation, that is, when $t \geq 0$. We assume that there is a transformation technology that can transform one unit of output into one unit of either old or new capital. The capital accumulation equations and their initial levels are given by equations (5) and (6). I impose two very general restrictions on Ψ

$$\Psi(0) = 1, \tag{7}$$

and

$$\Psi' > 0 \quad and \quad \Psi'' < 0. \tag{8}$$

Restriction (7) satisfies *Assumption 1.* To see this, notice that if there is no new capital invested after the innovation, $K_{new,t} = 0$, and $\min[A, \Psi(0)] = \min[A, 1] = 1$. Restriction (8) establishes that the Ψ function is strictly monotonically increasing and concave.

In each period, the firm may convert its output into investment in old capital, which embodies no new technology, or it may convert its output into investment in new capital, which embodies the technology and potentially increases TFP. Therefore the firm has to simultaneously determine the amount of old and new investment. To simplify the analysis, I assume that there is a minimum level of investment in old capital that equals its the maintenance cost

$$I_{old,t} \geq \delta \times K_{old,t}. \tag{9}$$

It is straightforward to show that the marginal product of new capital is always at least weakly larger than that of the old capital. We then have the following lemma:

Lemma 1 *Any net investment made after the technology shock is only in new capital, and, in each period, firms replenish their depreciated old capital.*[9]

Proof: Refer to Appendix A.

Subsequently, we have

$$K_{old, t} \equiv K_0 \qquad \forall \quad t > 0. \tag{10}$$

Therefore, we have to focus only on the investment decisions on new capital. Note that although the firms have to replenish their depreciated old capital, they may choose to sell their old capital if a market for it exists. However, in our analysis we do not have to model this directly because at equilibrium, the price of old capital will adjust to such a level that makes the firms indifferent between selling and buying a marginal unit of capital. Because all firms are the same in our model, there will be no transactions in the market. It is useful to keep in mind, however, that the shadow price of old capital is pinned down by the marginal product of old capital relative to new capital and by the adjustment cost associated with the resale of capital.

Because of the monotonicity of Ψ and because the level of $K_{old, t}$ is fixed along the optimal investment path, there exists a level of $K_{new, t}$ such that $\Psi\left(\frac{K_{new, t}}{K_0}\right) = A$. Denote this level of $K_{new, t}$ by K_{new}^*, and introduce the notation $Z_t = \Psi\left(\frac{K_{new, t}}{K_0}\right)$. We can rewrite the production function (3) as

$$Y_t = \begin{cases} Z_t \times K_t^\alpha & if \quad K_{new, t} < K_{new}^*, \\ A K_t^\alpha & if \quad otherwise. \end{cases} \tag{11}$$

---

[9]Without the minimum investment assumption, the firm may let old capital depreciate without fully replenishing it and invest faster in new capital. In an appendix available from the author, I show that in a generic multi-vintage model where technology of each vintage is embodied in the capital of only the same vintage, the stock of old capital may have a U-shape dynamic. Firms will reduce the holding of old capital by letting it depreciate while they are focusing on investing in new capital. After a sufficient amount of new capital is built, the firms replenish the old capital. The steady state holdings of old capital are not changed by the advent of technological innovation. How deep and fast they allow the old capital to depreciate depends on the empirical adjustment cost related to negative net investment. Section 3 provides a discussion of asymmetric adjustment cost.

and

$$K_t = K_{new,\,t} + K_0. \tag{12}$$

By *Assumption 4*, the firm does not know how big A is after the new technology arrives. It does know the distribution of A. The firm learns how big the shock is by continuously investing in new capital and evaluating whether adding new capital keeps contributing to increasing TFP. Denote the probability density function (PDF) and cumulative density function (CDF) of A by $\phi(A)$ and $\Phi(A)$ respectively. Mechanically, learning is carried out as follows: The firm observes the output and the levels of both old and new capital and computes the value of Z_t. By equation (11) the firm can infer the conditional distribution of A by applying

$$\begin{cases} Y_t = Z_t \times K_t^\alpha & \implies & A > Z_t\,, \\ Y_t \neq Z_t \times K_t^\alpha & \implies & A = \dfrac{Y_t}{K_t^\alpha}\,. \end{cases} \tag{13}$$

The idea is simple: If the firm has not invested enough to completely embody the new technology, $\min\,[A,\ Z_t] = Z_t$, and in turn, $Y_t = Z_t \times K_t^\alpha$. The firm therefore is able to infer that A is greater than Z_t and that the conditional distribution of A is $\dfrac{\phi(A)}{1 - \Phi(Z_t)}$. Conversely, if the invested new capital is sufficient to embody the new technology in the sense that $\min\,[A,\ Z_t] = A$, the exact magnitude of A can be inferred as $A = \dfrac{Y_t}{K_t^\alpha}$.

2.2 The Firm's Problem and Capital Demand

In this section, I set up the firm's profit maximization problem and derive a sufficient condition under which investment will accelerate after the technological innovation. As we have shown, investment in old capital is a constant and equal to the amount of old capital depreciation. However, investment in new capital includes the part that replenishes depreciated new capital, which grows with the stock of new capital. Therefore, if, under some conditions, investment accelerates from t to $t+1$ in an economy of zero depreciation, the same set of conditions should also lead to accelerating investment in an economy of positive depreciation.[10] Hence, without

---

[10]That is, investment acceleration in an economy of zero depreciation demands stronger conditions than that in an economy of positive depreciation.

loss of generality, I assume that $\delta = 0$. In this case, all investment along the optimal path will be the net increase of new capital.

Consider a firm that maximizes the discounted sum of dividend flows under the constant interest rate, r. Immediately after the arrival of the technology shock, the *cum dividend* value function of the firm is

$$V(K_0) = \max_{[I_t]_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \left(\frac{1}{1+r}\right)^t (Y_t - I_t), \tag{14}$$

subject to the production function (3), the information update function (13), and the modified capital accumulation function

$$\begin{cases} K_t & = K_{new,\,t} + K_{old,\,t} \\ K_{new,\,t} & = K_{new,\,t-1} + I_{t-1} \\ K_{old,\,t} & \equiv K_0 \, . \end{cases} \tag{15}$$

To capture the unique characteristics of capital demand in the underlying model, consider an auxiliary problem in which the firm has a standard production function but has to invest before the magnitude of a technological innovation is revealed. To fix the idea, consider a firm that anticipates a shock of unknown magnitude will arrive at $t = 1$. The firm has to make an investment at $t = 0$. In this case, the firm's optimality condition is simply to invest up to such a level that the expected marginal product of capital is equal to the borrowing cost, r. That is

$$r = \int_1^{\infty} \alpha \, A \, K_1^{\alpha-1} \, \phi(A) \, dA, \tag{16}$$

or

$$r = E_0 \, \alpha \bar{A} K_1^{\alpha-1} = \int_{A \in sup(A)} \alpha A K_1^{\alpha-1} \phi(A) \, dA.$$

where $\bar{A}$ is the unconditional mean of A. In the next period, after the magnitude of A is revealed, the firm will make investment decisions upon complete information.

In contrast, in our model with Ψ-type embodiment and learning, the first-order condition for capital is more complex, and, in turn, the investment dynamics are very different. In such a

model, the *cum dividend* value of the firm is the sum of current period payoffs and the discounted weighted average of future value in two regimes, $Z_1 > A$ and $Z_1 < A$. In the first regime, with capital K_1, the true magnitude of A has not been revealed, and the learning process will continue. Let $V(K_1)$ be the value of the firm in this regime. In the second regime, the firm finds that K_1 is sufficient to entirely pick up the new technology, and A is learned. The firm value is given by $W(K_1, A)$.

Note that V is a function of K_1 only, whereas W is a function of both K_1 and A (figure 1). Suppose a firm in the middle of the learning process with capital K_1 finds that A has not been revealed yet. In such a scenario, the firm learns that $A > Z_1$. The current capital stock K_1 contains all the information available about the conditional distribution of A. Therefore, K_1 contains all information available about the value of the firm. However, if the same firm realizes that A is completely revealed with capital K_1, then all levels of A that are between 1 and Z_1 can induce this. Z_1 can be either very close to the true level of A or much bigger than A, and the corresponding values of the firms are very different. Therefore, the value of a firm that has completed the learning process depends on both the capital level K_1 as well as the value of A. *Ex ante*, the firm can compute the probability of being in each regime when it increases capital stock from K_1 to K_2, but it does not know the exact magnitude of the shock. The firm has to compute the expectation of A conditioned on A being between 1 and Z_1.

Define $\pi(K_t, K_{t+1})$ to be the probability of the firm remaining in the learning process after it increases its capital stock from K_t to K_{t+1}, then $\pi(K_t, K_{t+1})$ can be computed as

$$\pi(K_t, K_{t+1}) = \frac{1 - \Phi(Z_{t+1})}{1 - \Phi(Z_t)} \ . \tag{17}$$

The firm's problem can be rewritten as the following Bellman equation system

$$V(K_0) = \max_{I_0} \ Y_0 - I_0 + \frac{1}{1+r}\{\pi(K_0, K_1) \times V(K_1) + \cdots$$
$$\cdots + [1 - \pi(K_0, K_1)] \times E_A \left[\, W(K_1, \ A) \mid 1 < A < Z_1 \, \right]\}, \tag{18}$$

where $V(K_1)$ is recursively defined as

$$V(K_1) = \max_{I_1} Y_1 - I_1 + \frac{1}{1+r}\{\pi(K_1, K_2) \times V(K_2) + \cdots$$

$$\cdots + [1 - \pi(K_1, K_2)] \times E_A [\, W(K_2, A) \mid Z_1 < A < Z_2 \,]\}, \qquad (19)$$

and

$$W(K_1, \, A) = \max_{[I_t]} \sum_{t=1}^{\infty} \left(\frac{1}{1+r}\right)^{t-1} [AK_t^{\alpha} - I_t]. \qquad (20)$$

In equation (18) the term $E_A[\, W(K_1, A) \mid 1 < A \leq Z_1 \,]$ is the expected value of a firm that has learned the value of A. The expectation is taken with respect to A conditional on $1 < A < Z_1$ because only when A is within this range can a firm with capital level of K_1 have completed the learning process.

I now introduce a useful lemma:

Lemma 2 *Given the same K_0 and $\phi(A)$, the optimal level of investment I_0 in the model with Ψ-embodiment and learning is larger than in the auxiliary model, in which the technology is not required to be embodied and the uncertainty about A is completely revealed at $t = 1$.*

Proof: Refer to Appendix A. Appendix A also discusses in detail the economic interpretation of the first-order condition for capital demand in the model with learning and embodiment.

The intuition of *Lemma 2* is simple. In a model with the $\Psi-$type embodiment and learning, new capital has additional roles besides production: First, new capital embodies new technology and increases TFP; second, new capital helps learning. Conditional on the value of A not being completely learned, more investment leads to a higher value of Z and in turn leads to a tighter conditional distribution of A. Hence, the marginal value of investment is higher than in a conventional model, and the firm will invest more.

Now we discuss the conditions under which investment accelerates persistently as we observed in the 1990s. Heuristically, this requires some properties related to the distribution of A. In particular, conditional on the current capital stock not completely revealing the level of A, the posterior expectation of A should increase at a sufficiently fast rate. I will provide a sufficient

12

condition on the distribution of A so that investment will be increasing over time, that is, $K_{t+1} - K_t > K_t - K_{t-1}, \, \forall \, t$.

I focus on the condition under which the optimal I_1 is greater than the optimal I_0, or $K_2 - K_1 > K_1 - K_0$.[11] The result can be extended to $I_t > I_{t-1}$. As shown in *Lemma 1*, the stock of old capital is a constant; therefore, in order to have $I_1 > I_0$, it is sufficient to show that the increment of the effective TFP associated with capital increasing from K_1 to K_2 is bigger than the increment of the effective TFP associated with capital increasing from K_0 to K_1. This is due to the concavity of the Ψ function. Specifically, we have

$$(21)$$

$$\Psi\left(\frac{K_{new,2}}{K_0}\right) - \Psi\left(\frac{K_{new,1}}{K_0}\right) > \Psi\left(\frac{K_{new,1}}{K_0}\right) - \Psi\left(\frac{K_{new,0}}{K_0}\right) \Rightarrow K_2 - K_1 > K_1 - K_0,$$

and, of course, $K_{new,0} = 0$.

Let $\tilde{K}_2$ be the level of capital that satisfies

$$r = \int_{Z_1}^{\infty} \alpha \, A \, \tilde{K}_2^{\alpha-1} \, \frac{\phi(A)}{1 - \Phi(Z_1)} \, dA. \tag{22}$$

We recognize that $\widetilde{K}_2$ is the optimal capital level of in our auxiliary problem if A follows the truncated distribution $\frac{\phi(A)}{1 - \Phi(Z_1)}$.[12] By *Lemma 2*, we know that the optimal capital level under the Ψ-type embodiment and learning, K_2, should be greater than $\widetilde{K}_2$. Therefore, we have

$$\widetilde{K}_2 - K_1 > K_1 - K_0 \Longrightarrow K_2 - K_1 > K_1 - K_0. \tag{23}$$

Finally, combining the relationship (21) and (23) and letting $\widetilde{Z}_2 = \Psi\left(\frac{\widetilde{K}_2}{K_0}\right)$, we have

$$\widetilde{Z}_2 - Z_1 > Z_1 - Z_0 \Longrightarrow K_2 - K_1 > K_1 - K_0, \tag{24}$$

where $\widetilde{Z}_2 - Z_1$ can be interpreted as the "lower bound" of the increment of effective TFP between $t = 1$ and $t = 2$ under the Ψ-type embodiment and learning. If even the lower bound is greater

---

[11]Recall that we assume that $\delta = 0$ to simplify notation.
[12]I slightly abuse notation here.

than the increment of the effective TFP between $t = 0$ and $t = 1$, it must be true that the optimal increment is even bigger. Therefore the increase of investment that induces this TFP increase is larger.

We have a lemma to characterize $\widetilde{Z}_2$:

Lemma 3

$$\widetilde{Z}_2 = \Psi\left[\left(\int_{Z_1}^{\infty} A\, \frac{\phi(A)}{1 - \Phi(Z_1)}\, dA\right)^{\frac{1}{1-\alpha}} - 1\right]. \tag{25}$$

Proof: By applying the FOC of the auxiliary problem and the steady state relationship, it is straightforward to have (25).

Now we reach the following proposition:

Proposition 1 *To have $I_1 > I_0$, it is sufficient if we have*

$$\Psi\left[\left(\int_{Z_1}^{\infty} A\, \frac{\phi(A)}{1 - \Phi(Z_1)}\, dA\right)^{\frac{1}{1-\alpha}} - 1\right] > 2Z_1 \qquad \forall\ Z_1\,. \tag{26}$$

Because Ψ is an increasing concave function, *Proposition 1* basically argues that if an increasing and concave transformation, which is pinned down by the function Ψ, of the conditional expectation of the technology shock is at least twice as big as the lower bound of the support of the conditional distribution, investment will accelerate.

An Example

A particularly interesting example is

$$\Psi\left(\frac{K_{new}}{K_{old}}\right) = \left(\frac{K_{new}}{K_{old}} + 1\right)^{1-\alpha}. \tag{27}$$

If Ψ has such a functional form, the model will have some desirable properties. We will defer more careful discussion of the properties to the next section. For now, notice that if Ψ is chosen to be as in (27), the sufficient condition given in (26) will be reduced to

$$\int_{Z_1}^{\infty} A\, \frac{\phi(A)}{1 - \Phi(Z_1)}\, dA > 2Z_1 \qquad \forall\ Z_1, \tag{28}$$

which is simply that the conditional expectation is at least twice as big as the lower bound. It is not difficult to locate a distribution with such properties. For example, let A follow the Pareto distribution $\Phi(A) = 1 - \dfrac{1}{A^\omega}$, $\phi(A) = \dfrac{\omega}{A^{\omega+1}}$, where $\omega > 1$ is the distribution parameter. For the conditional expectation, if $A > Z_1$, we have

$$\int_{Z_1}^{\infty} A \, \frac{\phi(A)}{1 - \Phi(Z_1)} \, dA = \frac{\omega}{\omega - 1} Z_1. \tag{29}$$

For $1 < \omega < 2$, $\dfrac{\omega}{\omega - 1} Z_1 > 2Z_1$, $\forall \, Z_1$.

This example gives us some hints about the shape of the distributions that are likely to induce increasing investment. The Pareto distribution is heavily tailed. The fat-tail property is particularly pronounced for the cases in which $\omega \in (1, 2)$. In addition, the Pareto distribution has a decreasing hazard rate. Let $H(A)$ denote the hazard rate of the distribution. For the Pareto distribution we have

$$H(A) = \frac{\omega}{A}, \tag{30}$$

which is decreasing with A. Why does the hazard rate affect how much a firm should invest? The hazard rate is positively related to the probability of ending learning. If the probability is high, the firm should be more cautious about investment. A decreasing hazard rate implies a decreasing probability of being able to complete the learning process. In this scenario, the firm will choose to engage in a more aggressive investment plan.

It is worthwhile to point out that what is required in Proposition 1 is a very strong condition. Indeed, because the concavity of the Ψ function can be quite powerful and the difference between K_2 and $\widetilde{K}_2$ can be quite large, a distribution that does not have a tail as fat as that of the Pareto distribution can generate the investment acceleration as well. I have not found an analytical expression for the necessary condition for accelerating investment.

2.3 Discussion

In this section, I explain more carefully the assumptions I made and contrast my model with other familiar models in the literature.

Technology is not like fertilizer in most cases. Better fertilizer increases the harvest on the same land with the same farmers and tractors being used. Better technology typically requires producers to invest in new capital before they may enjoy the higher productivity brought by the new technology. In the polar case, new technology and new capital can be complements in a Leontief sense. Output does not increase after a new technology arrives if no new capital investment has been made. The empirical relevance of embodied technology has been established in many studies (for example, Hercowitz 1998).

Roughly speaking, there are two types of technological innovation. One is the small, incremental innovation that happens probably every day. This type of progress might increase the productivity of only one sector or even one production unit. The other type is the large, revolutionary innovation, such as the invention of electricity and the introduction of information technology (IT). This type of progress is what is usually referred to as GPT, and occurs infrequently. In this paper, I focus on the second type of technological progress, the GPT.

The model introduced in this paper differs from the standard embodied technology models in many aspects. Solow (1959) is one of the classical contributions to the idea of embodied technology. In his model, capital of a particular vintage embodies the technology of the same vintage, and total output is the sum of the output produced by capital of various vintages. Letting $Y_{\nu, t}$ denote the output produced by capital and technology of vintage ν in time t, we have

$$Y_t^\nu = f(A_t^\nu, K_t^\nu). \tag{31}$$

The aggregated output at time t is given by

$$Y_t = \sum_\nu f(A_t^\nu, K_t^\nu). \tag{32}$$

One important property of the Solow-type embodiment is that a new technology makes capital of only the same vintage more productive. The productivity of capital of older vintages does not change. This can be shown by observing that $\dfrac{\partial MP^\nu}{\partial A^{\nu'}} = 0$ for $\nu \neq \nu'$, where MP^ν is the marginal product of capital of vintage ν. Put differently, old capital is not able to benefit from

16

the innovations.[13] In contrast, in a Ψ-embodiment model, it is the new capital that facilitates the adoption of the innovations. However, once adopted, the new technology increases the productivity of all capital, old and new.

The contrast between Ψ-embodied and Solow-embodied technology reflects two categories of technological innovation. One type of innovation does require substantial replacement of old capital with new capital, whereas the other type of innovation requires adding new capital but does not require replacing old capital. Now consider the following hypothetical example. An automobile company produces cars using assembly lines, controlled by a mechanical system. A new technology that controls assembly lines with computers is then introduced in the market. If the old assembly lines cannot be adapted to the computer system and have to be replaced with new assembly lines, the scenario is best described by a Solow model. If the computer system can be used to operate the old assembly lines, the scenario is consistent with the Ψ-embodied technology. In this setup, it is the productivity of the entire firm (both the assembly lines and the controlling computer system) that will increase. I will provide more examples in Section 5 to show that the Ψ-model captures important characteristics of the 1990s expansion.

In addition, our model also differs from the Solow embodiment because the latter does not distinguish between the concepts of *effective* TFP and *potential* TFP introduced in this paper. The Solow model assumes that new technology is entirely embodied in the first slice of new capital. The marginal productivity level is irrelevant to the amount of new capital invested. In contrast, in *Assumption 2*, our model explicitly assumes that, subject to a certain limit, A, a monotonically increasing functional relationship exists between the amount of new investment and the extent to which new technology is transformed into TFP.

In the Solow-type embodied technology model, the firm will replace old capital with new

---

[13]Consequently, after a new technology arrives, the market value of old capital decreases because it becomes less productive than new capital. Hobijn and Jovanovic (2001) exploit this idea to explain the stock market crash in 1974, and Laitner and Stolyarov (2003) try to explain why Tobin's q was persistently below 1 from 1974 to 1984. Both studies argue that the market value of the old firms that were incumbents before 1974 was destroyed by the arrival of the IT revolution in 1974. Although the IT revolution did not arrive at its full power until years later, these authors argued that this new technology was heralded and well understood by the middle of the 1970s.

capital because only capital of the latest vintage has higher productivity. When there are convex investment adjustment costs, or capital is not completely reversible, this replacement will be carried out over time. The observed TFP averaging across capital of all vintages converges to the state-of-the-art technology level at $t = \infty$. At the other extreme, if technology is disembodied as in the fertilizer example, a technology innovation would imply an immediate jump of TFP to the post-shock level. The Ψ-type embodiment introduced in this paper implies a time series path of effective TFP lying between the above two polar cases.

Figure 2 illustrates this contrast. Curve A is the TFP path after a technology shock in a disembodied model. It jumps to A_1 from A_0 right after the shock arrives at time t. Curve C is the path of the Solow-type embodiment with partially irreversible capital. The TFP only asymptotically converges to A_1 because old capital is gradually replaced by the capital that embodies the new technology. Curve B represents the TFP path implied by the Ψ-type embodiment. At time t, the firm learns the news about a technology innovation and starts to invest. It accumulates a sufficient amount of capital at t' to pick up all of the new technology, and from then on curve B coincides with curve A. However, the firm does not realize that it has invested enough in new capital until it observes that learning is completed at t''. Therefore the expected TFP level between t' and t'' exceeds the true level of the underlying technology.

The model assumes – simply for convenience – a standard lag of one period between the time of the investment and the time when the new capital becomes productive. One consequence is that the firm can make an investment that leads to a suboptimal level of capital for at most one period. However, the model can be generalized to allow for a T-period gestation delay, which can be either time-to-build or time-to-plan. Then the firm may make what are *ex post* suboptimal investments for multiple periods. As long as the investment decisions are irreversible, the extent to which the firm may accumulate capital beyond the first best level is largely irrelevant to the choice of investment-decision-making frequency (including continuous time modelling). Consider a T-period time-to-build setup as an example. Suppose that, at time t, the capital invested T periods before becomes productive and the firm realizes that the capital stock is above the

18

optimal level. The firm wishes that it had not invested any new capital after $t - T$. However, since it cannot reverse the capital that has already been invested, the capital of the firm will keep rising for the next $T - 1$ periods. To sum up, in a high frequency model, the investment per period can be infinitesimally small, but the total amount of capital overinvested can still be substantial if the time-to-build is sufficiently long.

Finally, *Assumption 5* assumes away the possibility of learning from other firms. Should this assumption be relaxed, firms can strategically choose the timing for investment. In particular, firms can delay their own investment until other firms have invested enough and the true magnitude of A is completely revealed. This will further complicate the analysis. I leave the discussion of this to a separate appendix available from the author.[14] On the other hand, this assumption is not wildly unrealistic. It captures the notion that the technology shock may have hit firms in various industries in very different ways even if it is a GPT. An auto producer may learn very little about how much IT would increase its productivity by observing by how much IT has increased the productivity in the food industry. In addition, even within the same industry, the integration of the new capital and new technology with existing capital may require firm-specific knowledge.

3   A Numerical Example: Dynamic General Equilibrium

This section gives an example to show quantitatively how Ψ-type embodiment and learning lead to overinvestment and affect the investment and output impulse response functions (IRF) with respect to the technology innovation. The computational details are in Appendix B. In the example, the Ψ function is explicitly calibrated and the model is set up in a dynamic general equilibrium (DGE) environment.

---

[14]Basically I found that in a monopolistic competition economy where information spill-over and learning-from-others are allowed for, there will be an endogenously determined fraction of firms that want to invest independently and remaining firms choose to wait until the uncertainty is resolved. However, it is difficult computationally to have the results similar to what we will present in the paper.

3.1 Embodiment Function Calibration

In the previous section, we require only that Ψ be monotonically increasing and concave. In this section, we calibrate the functional form of Ψ explicitly. We assume that pre-shock old capital, K_0, is the steady state level vis-à-vis A_0, and we impose only two conditions that Ψ should satisfy: First, the minimum amount of new capital required to completely embody the new technology is the steady state level of capital. Second, the Ψ-embodied technology production function has the same long-run steady state as in a standard neoclassical production function with disembodied technology for any level of A. The first condition requires that if K_{new}^* is the steady state level of new capital, we have

$$\min\left[A,\ \Psi\left(\frac{K_{new}^*}{K_0}\right)\right] = A = \Psi\left(\frac{K_{new}^*}{K_0}\right). \tag{33}$$

In a standard Cobb-Douglas production function of disembodied technology, $Y_t = AK_t^\alpha$, let $K^*(A')$ and $K^*(A'')$ be the steady state level of capital associated with two levels of technology, A' and A''. Then we have the well-known relationship

$$\frac{A''}{A'} = \left[\frac{K^*(A'')}{K^*(A')}\right]^{1-\alpha} \qquad \forall\ A'\ \text{and}\ A''. \tag{34}$$

Because the above relation holds regarding all pairs of A' and A'', we can let $A' = A_0$ and $A'' = A$ as the pre- and post-shock level of technology specified in (2). By (33) and (34), and keeping in mind $A_0 = 1$ by normalization, we have

$$A = \Psi\left(\frac{K_{new}^*}{K_0}\right) = \left(\frac{K_{new}^* + K_{old}^*}{K_{old}^*}\right)^{1-\alpha}. \tag{35}$$

If we rearrange and notice that the above relationship holds for any A, then for any $K_{new,t}$, the embodiment function reads

$$\Psi\left(\frac{K_{new,t}}{K_0}\right) = \left(\frac{K_{new,t}}{K_0} + 1\right)^{1-\alpha}. \tag{36}$$

This is exactly what we used as the example to illustrate the sufficient condition for accelerating investment. Indeed, this surprisingly concise functional relationship requires only the two

sensible conditions we introduced in the beginning paragraph of this section. Finally, plugging equation (36) into the production function (11), we reach

$$Y_t = \begin{cases} K_0^{\alpha-1} \times K_t & if \quad K_t < K^*(A), \\ AK_t^\alpha & if \quad otherwise. \end{cases} \tag{37}$$

Figure 3 contrasts production function (37) with a standard disembodied production function. In the latter, output jumps from Y_0 to $A \times Y_0$ when technology increases from 1 to A, even without any investment in new capital. The entire production function shifts up. For the Ψ-type embodied technology, the production function is linear before the capital stock approaches the new steady state $K^*(A)$. Beyond $K^*(A)$, the Ψ-type production function coincides with the disembodied production function.

One key feature that this Ψ function delivers is that sufficient learning is closely related to overinvestment. On the one hand, the firm wants to have a sufficient amount of capital to make the new technology completely embodied. On the other hand, with such an embodiment function, the firm will not learn the true value of A until it has accumulated capital above the steady state level. This pattern very much resembles the last business cycle. The economy heavily invested in IT-related capital to test the boundary of this great new technology. After each wave of investment, it seemed that adding more new capital would further advance productivity. However by the time the economy understood the potential of IT, it had invested too much. Investment then subsequently reversed, and a recession occured.

3.2   A Dynamic General Equilibrium Model

The Household

A representative agent maximizes the discounted sum of future utility over an infinite time horizon. The agent owns the shares of the firm, receives dividend payments each period, and can trade the shares to smooth consumption. The representative household's problem is

$$\max_{[C_t, S_t]} E_0 \sum_{t=0}^{\infty} \beta^t U(C_t), \tag{38}$$

subject to

$$C_t + P_t S_t = (P_t + D_t)S_{t-1}. \tag{39}$$

where

C_t is the consumption level in period t,

S_t is the share holdings at the end of period t,

D_t is the per-share dividend payment in period t,

P_t is the ex-dividend share price of the firm at the end of period t.

The Firm

A fixed number of firms maximize the sum of discounted dividend flows

$$\max_{I_t} \ E_0 \sum_{t=0}^{\infty} \beta^t \frac{U'(C_t)}{U'(C_0)} D_t \ , \tag{40}$$

subject to

$$D_t = Y_t - I_t - \Theta_t, \tag{41}$$

where output Y_t is defined as in (37). The new ingredient is

$$\Theta_t = c\left(\frac{K_{t+1} - K_t}{K_t}\right) \times K_t \tag{42}$$

It is the investment adjustment cost, and $c(\cdot)$ is a convex $\mathcal{C}^2$ function. In this numerical example, we assume that the depreciation rate is positive. The information structure is the same as in Section 2.2 such that the post-shock level of the underlying technology is A, the CDF and PDF of which are $\Phi(A)$ and $\phi(A)$ respectively.

Equilibrium

Given the initial state of the economy, A_0, K_0, and the unconditional distribution of A, the equilibrium conditions are familiar to us and are shared by many DGE models. An equilibrium is given by a sequence of quantities $\{C_t, I_t, K_t, Y_t\}_{t=0}^{\infty}$ and prices $\{P_t\}_{t=0}^{\infty}$ such that given prices $\{P_t\}_{t=0}^{\infty}$, the representative household solves (38); the firm solves (40) subject to underlying constraints. The markets for equity shares and goods clear.

22

The equity market clearing condition is

$$S_t \equiv 1 \qquad \forall \;\; t. \tag{43}$$

The optimal consumption in this model is therefore trivially determined as

$$C_t = D_t. \tag{44}$$

Let the marginal value of capital be MPK_t; then the optimal Euler equation for investment is the familiar

$$E_t \left[\beta \frac{U'(C_{t+1})}{U'(C_t)}(MPK_{t+1} + 1 - \delta) \right] = 1. \tag{45}$$

An Equivalent Social Planner's Problem

It is usually hard to compute the numerical results in a decentralized model. The equilibrium path of the quantities that are interesting, $\{C_t, I_t, K_t, Y_t, \}_{t=0}^{\infty}$, in the decentralized model can be replicated by a planner's problem. The social planner will solve the following problem

$$\max_{[C_t]} E_0 \sum_{t=0}^{\infty} \beta^t U(C_t), \tag{46}$$

subject to

$$Y_t = C_t + I_t + \Theta_t. \tag{47}$$

as well as the same capital accumulation and production functions.

Proposition 2 *The social planner's problem has the same first-order condition with respect to K_t as in the decentralized model, and hence has the same quantity path of $\{C_t, I_t, K_t, Y_t, \}_{t=0}^{\infty}$.*

In a manner similar to the Bellman equations (18)–(20), the social planner's problem can be written as

$$\begin{aligned} V(K_0) &= \max_{I_0} U(C_0) + \beta\{\pi(K_0, K_1) \times V(K_1) + \cdots \\ &\cdots + [1 - \pi(K_0, K_1)] \times E_A[\, W(K_1, \; A) \mid 1 < A < Z_1 \,]\}, \end{aligned} \tag{48}$$

23

where

$$V(K_1) = \max_{I_1} U(C_0) + \beta\{\pi(K_1,\ K_2) \times V(K_2) + \cdots$$

$$\cdots + [1 - \pi(K_1,\ K_2)] \times E_A[\ W(K_2,\ A)|Z_1 < A < Z_2\ ]\}, \tag{49}$$

and

$$W(K_1,\ A) = \max_{[I_t]} \sum_{t=1}^{\infty} \beta^{t-1} U(A\ K_t^\alpha - I_t - \Theta_t)\ . \tag{50}$$

The analytical solution to this problem is not tractable. This Bellman equation system, however, can be solved numerically by iterating both the value functions V and W. Appendix B provides the computational details.

3.3 Parameters Choices and Numerical Results

The model is calibrated annually, and the values of most parameters are set to the standard values in the literature. First, preferences are assumed to be CRRA,

$$U(C_t) = \frac{C_t^{1-\sigma}}{1 - \sigma}, \tag{51}$$

where σ is the coefficient of relative risk aversion. Following Hall (2001), the investment adjustment cost function is assumed to be

$$c\left(\frac{K_{t+1} - K_t}{K_t}\right) = \quad \frac{\theta^+}{2}\left(\frac{K_{t+1} - K_t}{K_t}\right)^2 \times \mathbf{P}(K_{t+1} - K_t) + \tag{52}$$
$$\frac{\theta^-}{2}\left(\frac{K_{t+1} - K_t}{K_t}\right)^2 \times [1 - \mathbf{P}(K_{t+1} - K_t)],$$

where $\mathbf{P}$ is an indicator function such that $\mathbf{P}(K_{t+1} - K_t) = 1$ if $K_{t+1} - K_t \geq 0$ and $\mathbf{P}(K_{t+1} - K_t) = 0$ if $K_{t+1} - K_t < 0$. θ^+ and θ^- are the parameters of adjustment costs for positive and negative investments, respectively. The value of the parameters α, β, δ, σ, θ^+, and θ^- are chosen as in table 1.

The calibration of θ^+ and θ^- follows Hall (2001). As Hall (2001) points out, θ^+, the adjustment cost coefficient for positive capital stock changes, can be related to the time needed to double the capital stock. Summers (1981) and Shapiro (1986) provide some empirical evidence

24

on the size of θ^+. Shapiro reports θ^+ is equal to eight quarters, or two years. Summers reports a much larger number for θ^+. His findings suggest that θ^+ is equal to thirty-two years! Summers's finding has been viewed as unrealistic by many authors (for example, see Tobin 1981). Therefore, Hall (2001) adopts two values for this parameter. He uses the value that Shapiro (1986) reports (two years) as the lower bound and eight years as the upper bound. I synthesize Hall's calibration by using the geometric average of the lower and upper value he used, which leads to $\theta^+ = 4$. The parameter θ^- is then chosen to be equal to 40, which is ten times θ^+, to capture the irreversibility of installed capital. The level of the adjustment cost plays an important role in determining the length of the period of increasing investment. Because it is possible to have increasing investment only before the true magnitude of A is learned, then faster investment can accelerate the learning process. A higher adjustment cost keeps the firm from investing too fast and prolongs the learning process. At the pre-shock steady state, the investment share is 21 percent, which is higher than the investment share in the data. The reason for this discrepancy is that I chose $\delta = 0.1$, whereas the economywide depreciation rate can be considerably lower.

What is still left to be "calibrated" is the distribution of the technological innovation. One restriction we impose is that the support of A is $(1, \infty)$ to reflect the notion that there is no technological regress. To begin, I constrain that the difference between A and A_0, denoted by ϵ, to follow an exponential distribution. The PDF for ϵ is $\lambda e^{-\lambda \epsilon}$, whereas the CDF is $1 - e^{-\lambda \epsilon}$. One reason for choosing an exponential distribution is that it has an analytical closed-form CDF, which is convenient when computing the *ex ante* probability of overshooting. Another consideration is that the exponential distribution has a constant hazard rate, which is the border case between an increasing and decreasing hazard rate. If an exponential distribution generates the desired dynamics, namely, accelerating investment, then for the families of distributions that have a decreasing hazard rate, the acceleration should be more pronounced.

The exponential distribution has only one parameter, λ, and $E(\epsilon) = \dfrac{1}{\lambda}$. Because the new technology in this paper is set to be a GPT, the innovation is not expected to take place every single year. Rather, it arrives once in a long period. The postwar data suggest that,

measured by the Solow residual, the average annual percentage growth of TFP is 0.79 percent. After controlling for the variations in capital utilization and nonconstant returns to scale, Basu, Fernald, and Kimball (2004) report that the "purified" residual increases on average 0.35 percent annually. Suppose that a GPT innovation takes place once every decade, and that all the productivity increase is due to the GPT progress; the mean of ϵ should then be between 8.19 percent and 3.56 percent.[15] If we postulate that only a fraction of the productivity growth is due to the GPT progress, then the mean of ϵ should be correspondingly lower. As a benchmark, I choose the mean of ϵ to be 4 percent and $\lambda = 25$.

Figure 4 illustrates the impulse-response dynamics of the economy after a new technology arrives. The magnitude of this technology shock is equal to 0.2. This is close to the cumulative productivity growth during the 1990s. The upper-left panel shows the conditional expectation of the magnitude of the technology shock. We see that right after the shock arrives, the expectation is simply equal to the unconditional mean, 0.04. After the agent invests in new capital and some of the new technology has been embodied, the agent learns that the technology shock is at least as large as what has already been embodied and revises the conditional belief about the magnitude of the shock. Therefore, the conditional expectation of ϵ keeps increasing until the investment overshoots. In the period before overshooting, the conditional expectation of ϵ is equal to 0.24, which is higher than the true magnitude of the shock. After overshooting, the true magnitude of the shock is completely learned and the conditional expectation goes back to 0.2, which is the true value of the shock. In this setup, we find that investment finally overshoots in the tenth year after the arrival of the new technology.

The investment dynamics, in the middle-left panel, show that in the nine consecutive years, in which investment has been increasing since the innovation, the annual growth rate is 3 percent and the cumulative growth is 28.5 percent. In the tenth year, when the agent learns that capital stock has overshot, investment is reduced dramatically, to 1.31, a decrease of 22 percent.

The output dynamics, in the bottom-left panel, indicate a long-lasting output boom and

---

[15]The cumulative non-purified Solow residual is 0.0819, and the cumulative purified residual is 0.0356.

an endogenous recession. Before the magnitude of the shock has been learned completely, investment keeps increasing and fuels the output boom. Output climbs from 1.00 to 1.30, with an average annual growth rate of about 3.0 percent. The collapse of the investment boom also ends the output boom. We see a sharp decrease in the output growth, though the level of output does not decrease in absolute terms.

One way in which the benchmark result is at variance with the data is that although investment increases quite persistently, the growth rate of investment is not as high as that of output. Consequently, the investment share is flat. In contrast, the data indicate that the share of private fixed investment (PFI) increases from 13.4 percent to 17.1 percent during the 1990s.

The consumption dynamics, in the middle-right panel, largely replicate the pattern of the output IRF before overshooting. At the moment of overshooting, consumption actually increases substantially because investment is sharply reduced while output is largely unchanged. This pattern is consistent with the most recent recession, during which investment crashed but consumption was sustained. This is one of the characteristics that makes the most recent business cycle unique. Typically, consumption is significantly procyclical. However, in the two years following the cyclical peak, real GDP grew a cumulative 2.6 percent, whereas real consumption grew almost 5.7 percent. I will discuss the dynamics of the implied (shadow) risk free interest rate, r_t^f (bottom-right panel) in the next section.

The post-overshooting dynamics of consumption and output are almost flat. The reason is that the post-overshooting level of capital is very close to the post-shock steady state. In other words, although the economy overshot, it has not overshot by much. Therefore, the economy undergoes only limited capital decumulation, and both output and consumption remain flat.

The substantial adjustment cost assumed in this model contributed to the long-lasting boom through two channels. First, high θ^+ slows new investment. Second, high θ^- leads to a large cost of overshooting and makes the firm invest in new capital more cautiously. If we use the lower value adopted by Hall (2001), namely $\theta^+ = 2$ and $\theta^- = 20$, the investment path will be

almost 10 percent higher than that in figure 4. Consequently, the boom period is shorter, only seven years. In addition, in the low adjustment cost model, the economy overshoots by a large amount of new capital, which makes the investment crash and recession more pronounced.

To summarize, the model generates persistent investment and output booms and endogenous overinvesting and recession. However, the investment share of GDP does not increase as much as we observe in the data. Meanwhile the investment growth rate is lower than that in the data. One potential reason for this is that the revision of the firm's conditional expectation of ϵ is not climbing very fast. Therefore, one round of successful investment does not prompt a round of sharply more aggressive investment. A distribution that induces a larger leap of conditional expectations also induces a larger leap of investment. I introduce an alternative distribution of technology shocks to allow for a fatter tail. Assume that the PDF of ϵ is $\zeta \lambda_1 \, e^{-\lambda_1 \epsilon} + (1-\zeta) \lambda_2 \, e^{-\lambda_2 \epsilon}$ and $\lambda_1 < \lambda_2$. This is a weighted average of two exponential distributions. Unlike the standard exponential distribution that has a constant hazard rate, this distribution has a decreasing hazard rate. We choose the parameters λ_1, λ_2, and ζ to make this distribution have the same mean as the previous one. Figure 5 illustrates the IRF after a shock of $\epsilon = 0.2$ in this setup.

In contrast to the previous case, the conditional expectation of ϵ increases much more rapidly. Within eight years, it jumps from 0.04, the unconditional mean, to near 0.4 before the economy overshoots. This rapidly revised belief about the magnitude of the technology shock induces a similarly rapid increase in investment. It grows from 1.28, a level lower than but close to the previous case in the first year, to 1.86, which reflects a 5.5 percent average annual increase. The initial investment in this setup is similar to the previous one because the two distributions have the same mean, whereas the investment grows faster in this setup because of the decreasing hazard rate.

During the same period, output grows at a slower rate of 3.3 percent. Therefore, the investment share increases from about 27.15 percent to nearly 31.44 percent over seven periods. Because investment grows faster, the boom lasts a shorter time. The economy overshoots after

eight years. An important distinction is that the recession is more pronounced in this setup, and we are able to visualize the decrease in output because the economy has overshot by a large amount and has to decumulate excess capital in a more aggressive way. For the same reason the output and investment dynamics in this model are not linear in ϵ.

Now I contrast the results presented with two alternative cases. First I present the impulse response of investment with respect to a technology shock in a standard DGE-growth model in figure 6 (the solid curve). Investment is normalized as a quantity index relative to the level in the first period after the technology shock. Two points are notable: (1) In reacting to the technology shocks, the level of investment does not change much over time; within twenty periods, the change is less than 4 percent. (2) The investment level is decreasing over time instead of increasing.

Second, to highlight how the hazard rate and infinite support of the distribution of ϵ matter, I slightly change the model. I assume that A follows a uniform distribution, which has an increasing hazard rate and a well defined upper bound. The dashed curve in figure 6 shows that in this scenario, if a shock of same magnitude arrives, investment is high at the early stage of the boom but diminishes over time even if the investment has not overshot. The intuition is simple: Because of the known upper bound as well as the increasing probability of overshooting, the firm invests at an increasingly cautious pace.

3.4 Discussion on Interest Rate and Firm Value Dynamics

The bottom-right panel of figures 4 and 5 show the dynamics of the implied (shadow) risk-free interest rate, r_t^f. Recall that the following familiar Euler equation should hold in the decentralized model, if the consumer can borrow and lend at a risk-free interest rate to smooth consumption[16]

$$U'(C_t) = E_t[\beta \, (1 + r_t^f) \, U'(C_{t+1})], \tag{53}$$

---

[16]In this paper, it is not necessary to include a risk-free asset in the model. Because in a homogeneous-agents economy nobody is lending or borrowing, such an asset is a redundant variable. However, the consumption Euler equation still prices the risk free asset, and that is why I label it the shadow risk-free rate.

which implies that

$$r_t^f = \frac{U'(C_t)}{\beta E_t\, U'(C_{t+1})} - 1 = \frac{1}{\beta}\frac{C_t^{-2}}{E_t\, [C_{t+1}^{-2}]} - 1, \tag{54}$$

where the last equality follows because $U(C) = -\dfrac{1}{C}$. The interest rate before the economy overshoots is very high in this model because of high expected consumption growth. High consumption growth is expected because before investment overshoots, some consumption has been sacrificed to increase investment. In the next period, if the economy remains undershot, consumption will remain low. When the economy overshoots, consumption will bounce back dramatically. Expected consumption growth is the average of the two scenarios weighted by the probability of overshooting. Because the probability of overshooting is not trivially small, expected consumption growth is quite high. After overshooting, the interest rate jumps down to the ordinary level, which is close to the steady state level, $\dfrac{1}{\beta} - 1$.

One implication of the high discount rate is the substantial decline of firm value. Indeed, during the learning phase of the transition, firm value decreases upon the arrival of the GPT innovation and remains low until the economy overshoots (figure 7). When the economy overshoots, the interest rate is adjusted to its normal level, as does firm value.

This scenario is apparently at odds with what happened in the 1990s. We did not see interest rates as high, or firm value as weak, as the present model predicts. One important reason for this gap is the present model is a closed economy model. Every dollar of increase in investment is at the cost of contemporaneous consumption, which will lift expected consumption growth and the interest rate. However, during the 1990s, especially during the later part of the decade, large amounts of foreign capital flowed in and helped keep the interest rate low. The interest rate and firm value puzzle can thus be reconciled in an open economy setup. The caveat of that approach is that if the interest rate is kept low because of the sufficient foreign asset supplies, investment will jump faster because capital is cheaper, and the investment boom may be shorter. *Ceteris paribus*, to keep interest rates low and the investment boom long-lasting, we need either faster depreciation of new capital or a larger adjustment cost or a combination of the two forces.

4 Related Literature

The learning mechanism employed in this paper is closely related to the literature often referred to as "Bayesian learning by doing." Aghion, Bolton, Harris, and Jullien (1991) study the problem of optimal learning through sequential experimentation by a single decision maker. In an application of the theory, Aghion, Bolton, and Jullien (1987) investigate experimental price setting by a monopolist facing uncertainty about demand.

Zeira (1987), Rob (1991) and Barbarino and Jovanovic (forthcoming) study uncertainty about some types of underlying capacity limits. For example, Rob (1991) proposes a mechanism for learning about market capacity through sequential entry of firms. Barbarino and Jovanovic (forthcoming) study a related learning problem in which the market capacity is learned through capacity expansion at firm level. One restrictive assumption in their models is that overshooting happens at an industry level. Within an industry, any firm level expansions on the margin do not alter the probability of overshooting of the industry because firms are infinitesimally small. Hence, when a firm decides to expand, it does not have to take into account the increased likelihood of overshooting that such an expansion would cause.

In addition, my model differs from those in previous studies in that it is a DGE model and can be used to study aggregated fluctuations, whereas Zeira (1987), Rob (1991) and Barbarino and Jovanovic (forthcoming) are partial equilibrium models that focus on firm or industry-level dynamics.

As for the origin of the uncertainty, my model is similar in spirit to that of Zeira (1987). Zeira sets forth the idea of productive capital and assumes that output is not a function of existing capital but a function of the minimum value between existing capital and productive capital. The level of the productive capital is unknown to the producers. In my model, the Ψ-type embodiment suggests that the effective TFP is equal to the minimum between the level of underlying technology and the TFP that can be embodied in the current amount of net new investment. In a certain sense, my model can be viewed as an extension and articulation of Zeira

(1987). However, because of complexities elsewhere in my model, my analytical results are not as elegant as those in Zeira (1987).

In recent studies of business cycles, various authors have proposed mechanisms to make DGE business cycle models exhibit impulse response functions that are more consistent with the data. Much of this effort has focused on imperfect information and expectations. Beaudry and Portier (2004) revitalized the idea of Pigou (1926) and argued that an upward-biased expectation of future productivity growth will lead to a current expansion. Similar to their work, Jaimovich and Rebelo (2006) show that "recessions are caused not by contemporaneous negative shocks but by lackluster news about future TFP or investment-specific technical change." Unlike my model, Beaudry and Portier (2004) and Jaimovich and Rebelo (2006) do not explicitly model learning. As time progresses, the agents should be able to collect additional information about future productivity growth. If the agents are allowed to revise their expectations about future productivity growth conditional on the evolving information set, as long as new information arrives at a sufficiently slow rate, it is less likely that recessions of the scale in their papers will happen.

Learning plays a central role in Van Nieuwerburgh and Veldkamp (2006). They focus on the asymmetry of business cycles, that is, why booms usually last a long period while recessions are short and sharp. In their model, the production function is characterized by exogenous noise, which makes it hard to tell if high output is due to better technology or lucky output shocks. In contrast to my model, their learning mechanism is more like a signal-extraction device. When productivity is expected to be high, production is higher and this will thereby lead to a flow of more precise information. When the economy passes the peak of a productivity boom, precise estimates of the slowdown prompt investment and labor to fall sharply. Conversely, at the end of a slump, low production impedes learning, slows the recovery, and makes booms more gradual than crashes. Thus booms and recessions in their model are still exogenous, and boom does not have to induce a recession. Learning in their model affects the duration of booms and recessions only and does not lead to any overshooting.

5 A Closer Look at the Last Business Cycle

The model in this paper generates a persistent investment boom as well as a surge in observed TFP. In this section, I document the pattern of the cycle in the 1990s in greater detail. I present several stylized examples to highlight the relevance of the Ψ-embodied technology in that era. I also provide examples of other historical episodes that share characteristics similar to the most recent one.

The upper panel of figure 8 plots the logarithm of real PFI and its three major components. From 1991 to 2000, real GDP increased at an annual average of 3.3 percent, whereas real PFI increased at an annual average of 7.3 percent. When PFI is decomposed into equipment and software (E&S), nonresidential structures, and residential structures, we find that over that decade, real E&S investment increased at an annual average of about 10.3 percent. Investment growth was even stronger over certain years of the decade and with respect to certain investment goods. For instance, from 1997 to 2000, investment in information-processing equipment increased 25 percent per year! In contrast to the rapid growth of E&S investment, other components such as non-residential and residential structure did not increase nearly so fast. After the economic expansion reached its peak, we saw a sharp decrease in investment in the subsequent years.

The lower panel of figure 8 plots investment as a share of GDP. Starting in 1991, the share of PFI largely increased in parallel to the share of E&S investment. The share of real PFI increased from 13.4 percent to 17.1 percent by 2000, whereas the share of E&S investment increased from 6.9 percent to 9.5 percent.[17] In respect to real PFI, only four other years in history saw higher investment share, [18] whereas for E&S investment reached new historical highs every year from 1995 to 2000.

Another well-documented characteristic of the 1991-2000 period is a rapid increase in pro-

---

[17]I calculated the investment share using the nominal level of investment divided by nominal GDP in order to avoid the potential bias due to the use of an ideal chain index. The bias is particularly large in the E&S sectors. For more detailed discussion on this problem, see Whelan (2000).

[18]During the period of 1978-1981, the PFI share of GDP reached a historical high of 18.5 percent. The increase in investment shares in this era was more concentrated in structures. E&S investment shares rose to a lesser extent relative to the rise in the 1990s.

ductivity. Productivity growth was particularly strong during the second part of the decade. Oliner and Sichel (2000) report that from 1995 to 1999, output per hour increased 2.5 percent annually, which is almost double the average growth rate in the previous twenty-five years. They estimate that about two thirds of this increase was due to the contribution of information technology. Furthermore, they estimate that about 1 percentage point of the annual increase in labor productivity was due to computer hardware and software.

Indeed, a substantial part of the increase in productivity in the 1990s did not require replacing old capital with new capital. On the contrary, old capital benefitted from being operated with new capital that embodied new technology. One celebrated example comes from Wal-Mart (See, for example, Tsao 2003). The retailer did not replace its truck fleet but equipped itself with better inventory control and logistics technology, which are rooted in the IT revolution. Similarly, businesses that expand their e-commerce operation only have to install better information processing and telecom equipment but the productivity of all its buildings and physical structures benefit. If we interpret capital broadly and include human capital, we can see that the IT-related knowledge is most valuable when it is combined with the existing base of knowledge and the old human capital has not been completely substituted with IT-based human capital.

Returning to the most recent recession, although real GDP increased only 0.8 percent in 2001, real personal consumption increased 2.5 percent. The dispersion between GDP and consumption growth was quite persistent in the next year as well. Real GDP increased 1.9 percent in 2002, whereas real consumption grew 3.1 percent. Consumption growth in that year was particularly strong for durable goods – 6.6 percent.

Several important episodes in history share similar patterns of investment. Jovanovic and Rousseau (2005) document that the share of total power generated by electricity increased sharply between 1900 and 1929. This suggests an investment boom in electricity generating and related industries. The growth of electric power share slowed down abruptly in 1929, the first year of the Great Depression, and electricity expansion never returned to its previous strength,

even after the end of the depression. This pattern indicates a certain degree of capital overhang. Similar anecdotes can be found in the studies of canal and railroad expansions in the 19th century. Each boom started with the advent of a revolutionary technology and ended with overinvestment. Canal building boomed after the invention of the steamboat, and by the year 1860 more than 4,000 miles of canal had been completed. However, many of these canals did not live up to the expectations of their promoters. Many of these projects eventually turned out to be financial failures.[19] Later in the same century, the railroad expansion shared a similar fate. Thousands of miles of railroad were built and left unused or under used, a phenomenon described by Schumpeter (1949) as construction "ahead of demand." Fogel's investigation (1964) found that the *ex post* return in the railroad industry in the early 1870s turned out to be too low to attract capital. Overall, the steamboat, the railroad, and electricity all contributed to productivity increases to a great extent, but investors eventually overshot.

6 Conclusion and Future Research

In this paper, I propose a new model in which, once new technology has been embodied into new capital, it increases the productivity of capital of all vintages. Subject to the limits of the available new technology, the more new capital that has been installed, the higher is the productivity level. Agents in the economy do not know exactly the magnitude of the new technology shock and therefore have to learn it by investing in new capital. The true magnitude is learned after their investments overshoot the optimal level. I provide a sufficient condition under which investment increases persistently and overshoots at the end. After overshooting, investment is sharply reduced and a recession is triggered. On the other hand, consumption can increase during recessions because output is reallocated from investment to consumption.

The investment, consumption, and output dynamics derived from the model match qualitatively what we observed in the data of the last business cycle. I argue that the model adds a strong internal propagation mechanism to the standard DGE models. In addition, it generates

---

[19]Atack and Passell (1994) provide a detailed treatment of the canal-building boom and its economic consequences.

endogenous recessions without invoking technological regress.

Some extensions to the model seem promising. One is to allow for a more flexible and generalized learning structure. In particular, a "learning from others" framework where firms can learn from other firms might be more realistic. As Bolton and Harris (1999) point out, if we extended the single-agent problem of optimal learning by experimentation into a multi-agent setup, the free-rider problem would arise. In the context of my model, this implies that if one firm invests more, not only this firm but all other firms can potentially acquire better information about the underlying new technology. Therefore, a typical firm will strategically delay its investment in the absence of additional incentives. One such incentive for investment would be in a monopolistic environment. Because a firm's profit is a function of its productivity level, a firm that invests quickly acquires a higher level of productivity relative to those firms investing slowly. The tradeoff in such a model would be between investing quickly and acquiring higher productivity sooner versus investing slowly and avoiding the potential costs due to overshooting. This might result in a Nash equilibrium at which a fraction of the firms choose to be leaders and invest quickly and others choose to be followers.

References

[1] Aghion, Philippe, Patrick Bolton, and Bruno Jullien (1987), "Learning Through Price Experimentation by a Monopolist Facing Unknown Demand", MIT Working Paper.

[2] Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien (1991), "Optimal Learning by Experimentation", *Review of Economic Studies*, vol. 58 (June), pp. 621-54.

[3] Atack, Jeremy and Peter Passell (1994) *"A New Economic View of American History from Colonial Times to 1940"*, New York: W.W. Norton & Company.

[4] Barbarino, Alessandro and Boyan Jovanovic (forthcoming), "Shakeouts and Market Crashes", *International Economic Review*.

[5] Basu, Susanto, John Fernald and Miles Kimball (forthcoming). "Are Technology Improvements Contractionary?", *American Economic Review*.

[6] Beaudry, Paul and Franck Portier (2004), "An Exploration into Pigou's Theory of Cycles", *Journal of Monetary Economics*, vol. 51 (September), pp. 1183-216.

[7] Bolton, Patrick and Christopher Harris (1999), "Strategic Experimentation", *Econometrica*, vol. 67 (March), pp. 349-74.

[8] Cogley, Timothy and James M. Nason (1995) "Output Dynamics in Real-Business-Cycle Models", *American Economic Review*, vol. 85 (June), pp. 492-511.

[9] Fogel, Robert (1964), *"The Union Pacific Railroad: A Case in Premature Enterprise"*, Baltimore: Johns Hopkins University Press.

[10] Hall, Robert (2001), "The Stock Market and Capital Accumulation", *American Economic Review*, vol. 91 December, pp. 1185-202.

[11] Hercowitz, Zvi (1998), "The Embodiment Controversy: A Review Essay", *Journal of Monetary Economics*, vol 41 (February), pp. 217-24.

[12] Hobijn, Bart and Boyan Jovanovic (2001), "The Information Technology Revolution and the Stock Market: Evidence", *American Economic Review*, vol. 91 (December), pp. 1203-220.

[13] Jaimovich, Nir and Sergio Rebelo (2006), "Can News About the Future Drive the Business Cycle?", working paper.

[14] Jovanovic, Boyan and Peter Rousseau (2005), "General Purpose Technology", chapter in the *Handbook of Economic Growth*, vol.1, part 2, pp. 1181-224.

[15] Laitner, John and Dmitriy Stolyarov (2003), "Technological Change and the Stock Market", *American Economic Review* vol. 93 (September), pp. 1240-67.

[16] Van Nieuwerburgh, Stijn and Laura Veldkamp (2006), "Learning Asymmetries in Real Business Cycles", *Journal of Monetary Economics*, vol. 53 (May), pp. 753-72.

[17] Oliner, Stephen and Daniel Sichel (2000), "The Resurgence of Productivity Growth in the Late 1990s: Is Information Technology the Story?", *Journal of Economics Perspectives*, vol. 14 (Autumn), pp. 3-22.

[18] Pigou, Arthur Cecil (1926), *"Industrial Fluctuations"*, New York: A.M. Kelly.

[19] Rob, Rafael (1991), "Learning and Capacity Expansion under Demand Uncertainty", *Review of Economic Studies*, vol. 58 (June), pp. 655-75.

[20] Shapiro, Matthew D. (1986), "The Dynamic Demand for Capital and Labor", *Quarterly Journal of Economics*, vol. 101 (August) pp. 513-42.

[21] Schumpeter, Joseph (1949), *"The Theory of Economic Development"*. Cambridge: Harvard University Press.

[22] Solow, Robert Solow (1959), "Investment and Technological Progress", In Kenneth Arrow, Samuel Karlin and Patrick Suppes, eds., *Mathematical Methods in the Social Sciences*, Stanford, CA: Stanford University Press, pp. 89-104.

[23] Summers, Lawrence (1981), "Taxation and Corporate Investment: A q-Theory Approach", *Brooking Papers on Economic Activity*, vol. 1981, pp. 67-127.

[24] Tobin, James (1981), "Discussion of Summers", *Brookings Papers on Economic Activity*, vol. 1981, pp. 132-9.

[25] Tsao, A. (2002) , "Will Wal-Mart Take Over the World?", *BusinessWeek Magazine*, Nov. 27, 2002.

[26] Zeira, Joseph (1987), "Investment as a Process of Search", *Journal of Political Economy*, vol. 95 (February), pp. 204-10.

A Proof of Lemmas

Proof of Lemma 1: By the production function (3), if $A > \Psi\left(\frac{K_{new,t}}{K_{old,t}}\right)$, then $Y = \Psi\left(\frac{K_{new,t}}{K_{old,t}}\right) \times$
$(K_{new,t} + K_{old,t})^{\alpha}$. The marginal product of K_{new} and K_{old} are

$$MP_{new} = \frac{1}{K_{old,t}} \times \Psi' + \alpha\Psi \times (K_{new,t} + K_{old,t})^{\alpha-1}, \tag{A.1}$$

and

$$MP_{old} = -\frac{K_{new,t}}{K_{old,t}^2} \times \Psi' + \alpha\Psi \times (K_{new,t} + K_{old,t})^{\alpha-1}. \tag{A.2}$$

If $A < \Psi\left(\frac{K_{new,t}}{K_{old,t}}\right)$, then $Y = A \times (K_{new,t} + K_{old,t})^{\alpha}$. The marginal product of K_{new} and
K_{old} are equal. Remember $\Psi' > 0$, therefore,

$$MP_{new,t} \geq MP_{old,t} \qquad \forall \quad t. \tag{A.3}$$

■

Proof of Lemma 2 The first order condition of Bellman equations (18) - (20) is

$$\int_{Z_1}^{\infty} \frac{\partial V}{\partial K_1} \phi(A) \, dA + \int_1^{Z_1} \frac{\partial W}{\partial K_1} \phi(A) \, dA +$$

$$\{W[K_1, Z_1] - V(K_1)\} \times \phi(Z_1) = 1 + r. \tag{A.4}$$

By the envelope theorem and repeated use of the Newton-Leibnitz Theorem we have

$$\frac{\partial V}{\partial K_1} = K_0^{\alpha-1} + \cdots$$

$$+\frac{1}{1+r} \times \frac{\phi(Z_1)\,\Psi'(K_1)}{[1-\Phi(Z_1)]^2} \times \left[\int_{Z_2}^{\infty} V(K_2)\,\phi(A)\,dA + \int_{Z_1}^{Z_2} W(K_2, A)\,\phi(A)\,dA\right]$$

$$+\frac{1}{1+r} \times \frac{1}{1-\Phi(Z_1)} \times \left[\int_{Z_2}^{\infty} \frac{\partial V}{\partial K_2}\,\phi(A)\,dA + \int_{Z_1}^{Z_2} \frac{\partial W}{\partial K_2}\,\phi(A)\,dA\right]$$

$$+\frac{1}{1+r} \times \frac{1}{1-\Phi(Z_1)}$$

$$\times\{[\,W(K_2, Z_2) - V(K_2)\,]\phi(Z_2)\,\Psi'(K_2) - W(K_2,\,Z_1)\phi(Z_1)\,\Psi'(K_1)\}$$

$$. \tag{A.5}$$

The first order condition for (19) provides that

$$
1 + r = \frac{1}{1 - \Phi(Z_1)} \times \quad \text{(A.6)}
$$

$$
\left\{ \int_{Z_2}^{\infty} \frac{\partial V}{\partial K_2} \phi(A) \, dA \; + \; \int_{Z_1}^{Z_2} \frac{\partial W}{\partial K_2} \phi(A) \, dA + [W(K_2, Z_2) - V(K_2)]\phi(Z_2) \, \Psi'(K_2) \right\}.
$$

Plug (A.6) into (A.5) we have

$$
\frac{\partial V}{\partial K_1} = 1 + K_0^{\alpha - 1} + \frac{1}{1+r} \times \frac{\phi(Z_1) \, \Psi'(K_1)}{1 - \Phi(Z_1)} \times \quad \text{(A.7)}
$$

$$
\left[\frac{\displaystyle\int_{Z_2}^{\infty} V(K_2)\phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A)\phi(A) \, dA}{1 - \Phi(Z_1)} - W(K_2, Z_1) \right].
$$

However, we still have

$$
\frac{\partial W(K_1, A)}{\partial K_1} = 1 \; + \; \alpha A \, K_1^{\alpha - 1}. \quad \text{(A.8)}
$$

Finally, plug (A.7) and (A.8) into (A.4) we reach

$$
r = \left[K_0^{\alpha - 1}[1 - \Phi(Z_1)] + \alpha K_1^{\alpha - 1} \int_1^{Z_1} A \, \phi(A) \, dA \right] \; +
$$

$$
\phi(Z_1)\Psi'(K_1)[W(K_1, Z_1) - V(K_1)] + \quad \text{(A.9)}
$$

$$
\frac{1}{1+r} \times \phi(Z_1) \, \Psi'(K_1) \times
$$

$$
\left[\frac{\displaystyle\int_{Z_2}^{\infty} V(K_2) \, \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, \, Z_2) \, \phi(A) \, dA}{1 - \Phi(Z_1)} - W(K_2, Z_1) \right].
$$

Keep in mind that the optimality condition of the auxiliary problem is

$$
r = \int_1^{\infty} \alpha \, A \, K_1^{\alpha - 1} \, \phi(A) \, dA
$$

$$
= \int_{Z_1}^{\infty} \alpha \, A \, K_1^{\alpha - 1} \, \phi(A) \, dA + \int_1^{Z_1} \alpha \, A \, K_1^{\alpha - 1} \, \phi(A) \, dA. \quad \text{(A.10)}
$$

Notice that in (A.9) and (A.10), the common part is $\alpha K_1^{\alpha - 1} \int_1^{Z_1} A \, \phi(A) \, dA$. We can show that $\forall \, K_1$, the value of the RHS of (A.9) excluding the common part has a larger value than the RHS of (A.10) excluding the common part. In order to make the RHS of (A.9) equal to the RHS of (A.10), we must have the optimal K in (A.9) greater than the optimal K in (A.10) ∎

The Economic Interpretation of (A.9)

The RHS of equation (A.9) has three components, and here is an intuitive interpretation of them.

1. $K_0^{\alpha-1}[1 - \Phi(Z_1)] + \alpha K_1^{\alpha-1} \int_1^{Z_1} A \, \phi(A) \, dA$ is the expected marginal productivity that incorporates the Ψ-type embodiment.

2. As illustrated in figure 4, the extent to which investment overshoots varies, hence the *ex post* loss of overshooting varies. The marginal gain due to the reduction of the loss of overshooting in the subsequent periods is $\phi(Z_1)\Psi'(K_1) \times [W(K_1, Z_1) - V(K_1)]$. To see this, assume that the firm has capital K_1 and that the true size of A is equal to $Z_1 + \eta$, where η is a sufficiently small positive number. Because $Z_1 < A$; the firm would have found undershooting and the continuation value would be $V(K_1)$. However, because A is so close to Z_1 and the firm has not realized this, the cost of overshooting will be quite big. Suppose that the firm increases investment on the margin and pushes K_1 up so that $Z_1 > A$, investment will overshoot and learn the true size of A. The continuation value will be $W(K_1, \ A)$. Because η is sufficiently small, we then should have $W(K_1, \ A) \approx W(K_1, \ Z_1)$. Because K_1 is the steady state level of capital corresponding to the technology level Z_1, we have $W(K_1, \ Z_1) > V(K_1)$. The difference measures the loss reduced by avoiding large overshooting in the subsequent period. The gain is weighted by the change of the probability of overshooting on the margin, $\phi(Z_1)\Psi'(K_1)$.

3. Finally, suppose the firm has undershot with K_1, the term

$$\frac{1}{1+r} \times \phi(Z_1) \, \Psi'(K_1) \times \left[\frac{\int_{Z_2}^{\infty} V(K_2) \, \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, \ A) \, \phi(A) \, dA}{1 - \Phi(Z_1)} - W(K_2, \ Z_1) \right]$$

captures the value of the additional information provided by the marginal investment. Recall that the firm updates the belief about A conditional on $A > Z_1$ after observing undershooting. The firm should be better off if instead it learns that $A > Z_1'$, where $Z_1' > Z_1$, because tighter conditional distribution provides more precise information about A. Since this helps to increase the firm's value only after the next period, the value of information has to be discounted by $1 + r$. The fraction term in the bracket is the expected value of the firm conditional on $A > Z_1$, and the second term in the bracket, $W(K_2, \ Z_1)$, is the value of the firm if the true size of the shock is Z_1 but the firm overshoots its capital to K_2. The difference between these two terms is the gain from pushing up the lower bound of the conditional distribution of A.

B Computation Method

I use value function iteration over a discrete grid to compute the value functions. I discretize the A distribution with N points, and each point has weight equal to $\frac{1}{N}$. Depending on the shape of the distribution, N may have to be rather large to capture the tail properties. Meanwhile I discretize the state variable, K.

I begin by computing the value function $W(K, A)$ because it does not involve $V(K)$. This can be computed with the standard value function iteration technique. I iterate the value functions until the average gap between two consecutive iterations is smaller than 0.0001. Because the computation process is time consuming and N is large, I do not compute $W(\cdot, A)$ for each A I simulated in the discretized distribution. Instead, I compute $W(\cdot, A)$ on a rougher grid of ϵ and interpolate the value function for all A grids.

I then compute $\pi(K_t, K_{t+1})$ for all possible pairs $K_{t+1} > K_t$ on the K grid. With W and π in hand, I iterate $V(K)$ until the gap is smaller than 0.0001. Thus, we have a value function of the economy with capital level K without having overshot.

α	β	δ	σ	θ^+	θ^-
0.3	0.96	0.10	2	4	40

Table 1: Parameters

φ(A)

If $1 < A < Z_1$, then learning
is completed after one round
of investment

If $A > Z_1$ then learning
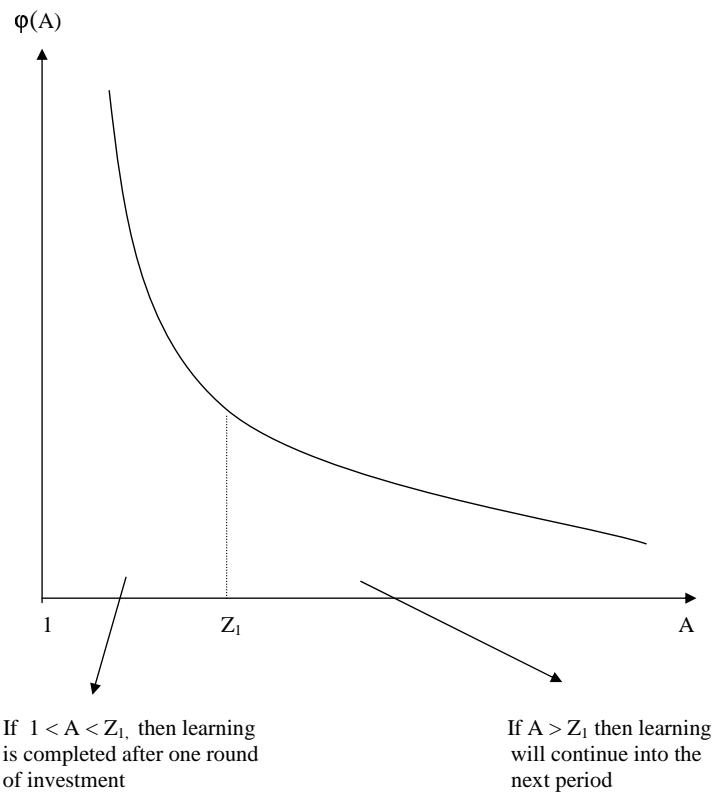will continue into the
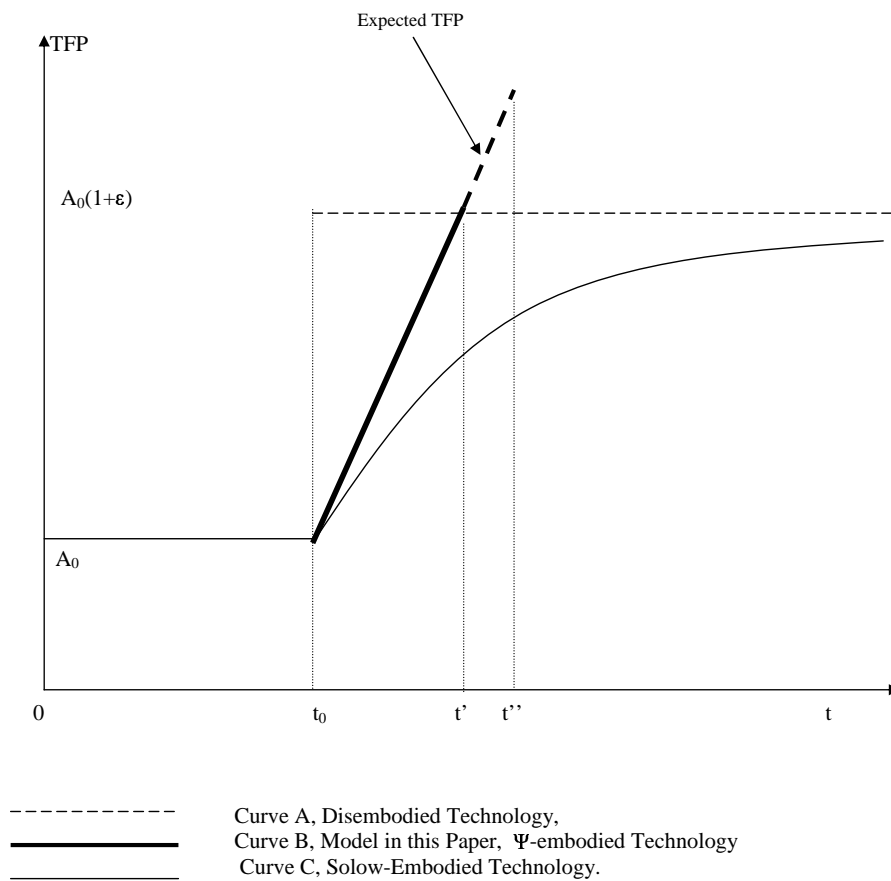next period

Figure 1: Illustration of Learning

44

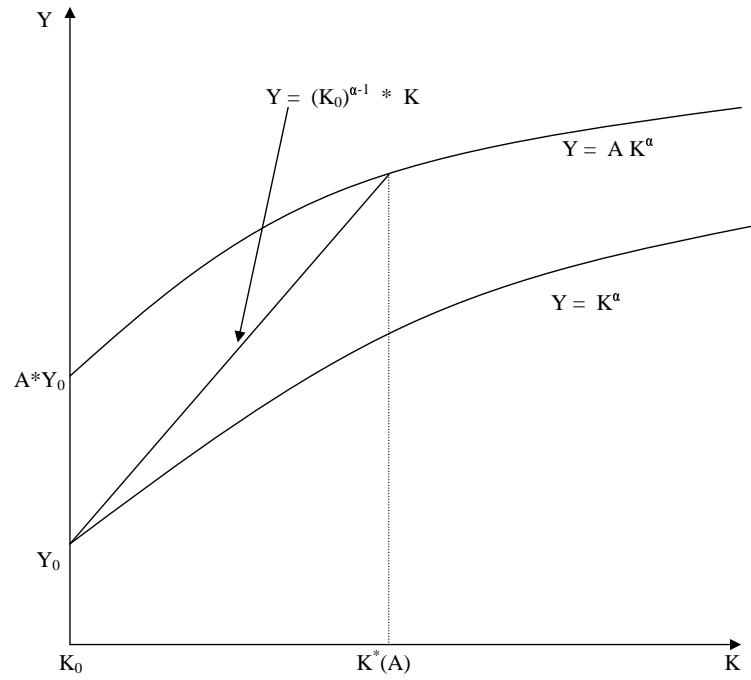Figure 2: TFP in Disembodied, Solow-Embodied, and Ψ-Embodied Technology Models

45

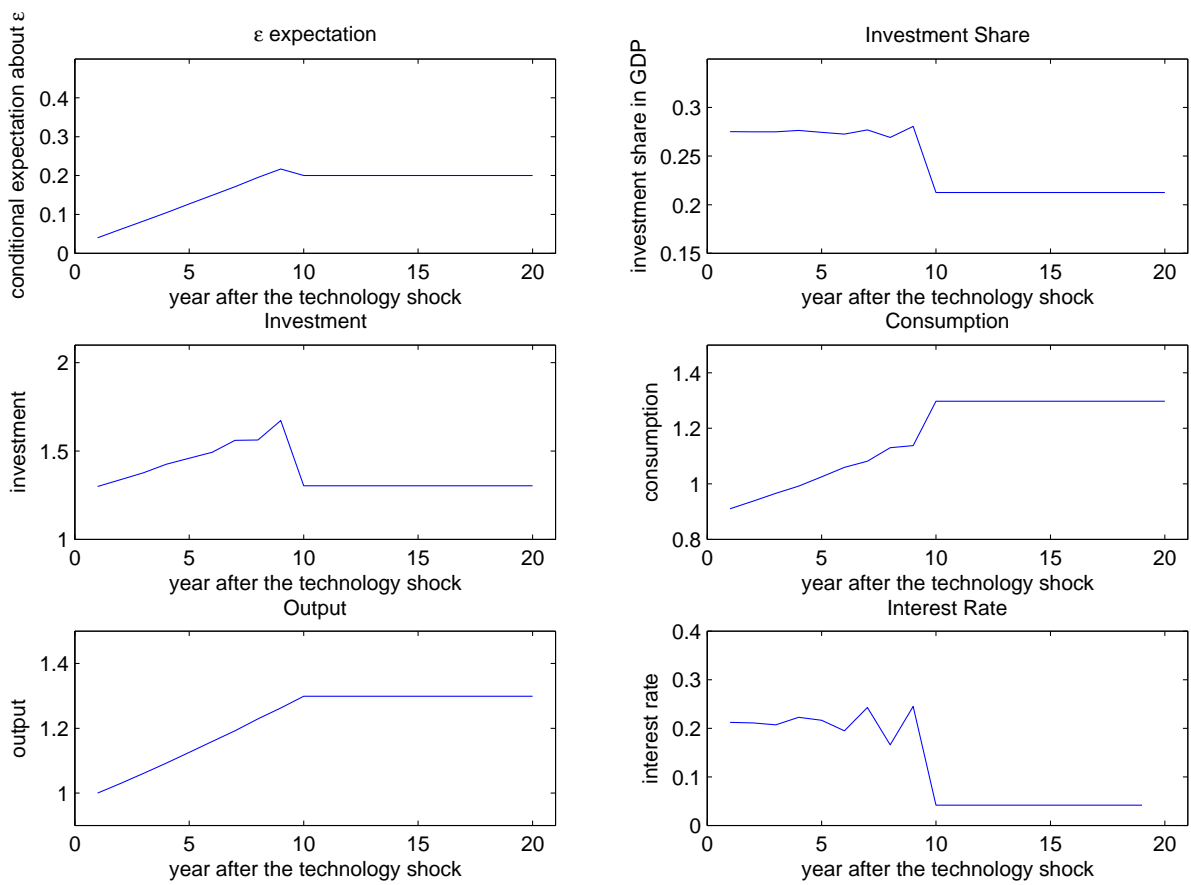Figure 3: Comparison of Output Changes between the Standard Cobb-Douglas and the Ψ Embodied Technology

Figure 4: Impulse Response of a Constant Hazard Rate ϵ Distribution
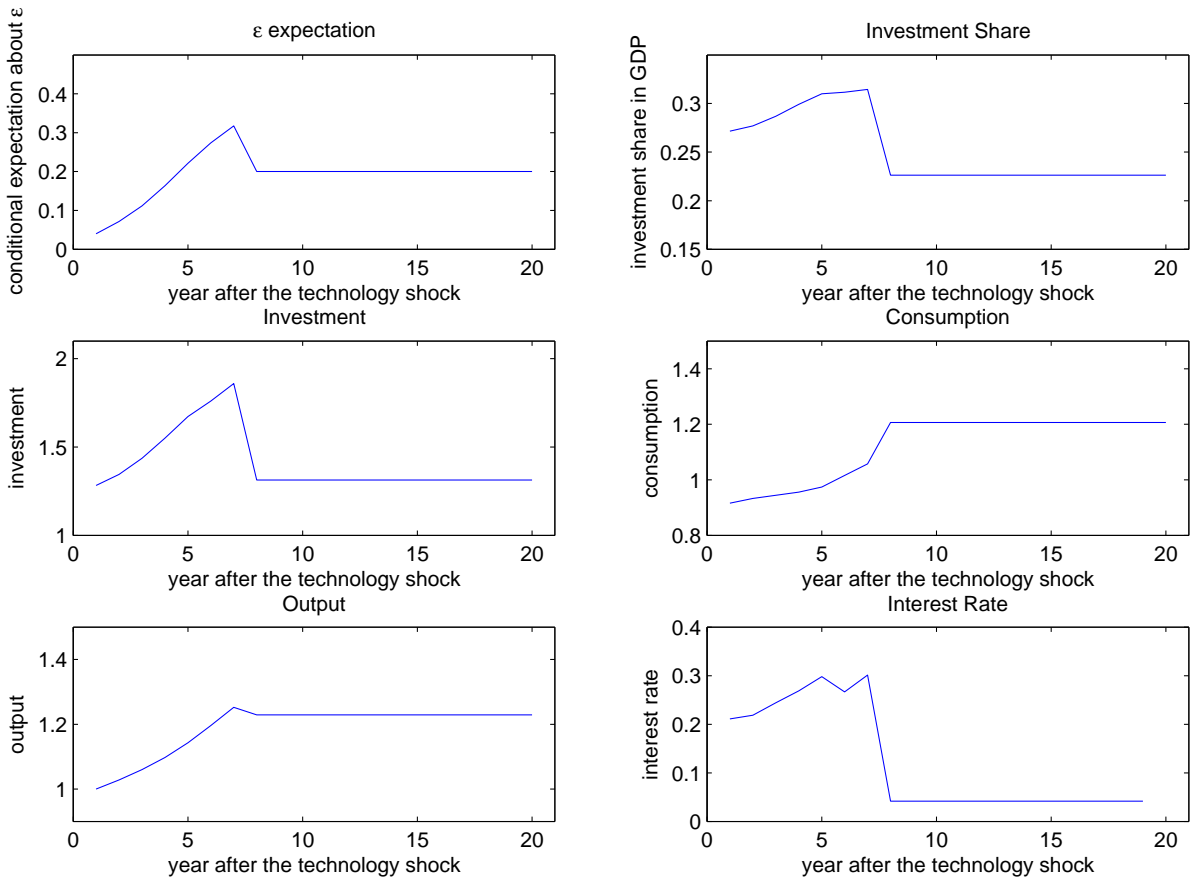
Figure 5: Impulse Response of a Decreasing Hazard Rate ϵ Distribution
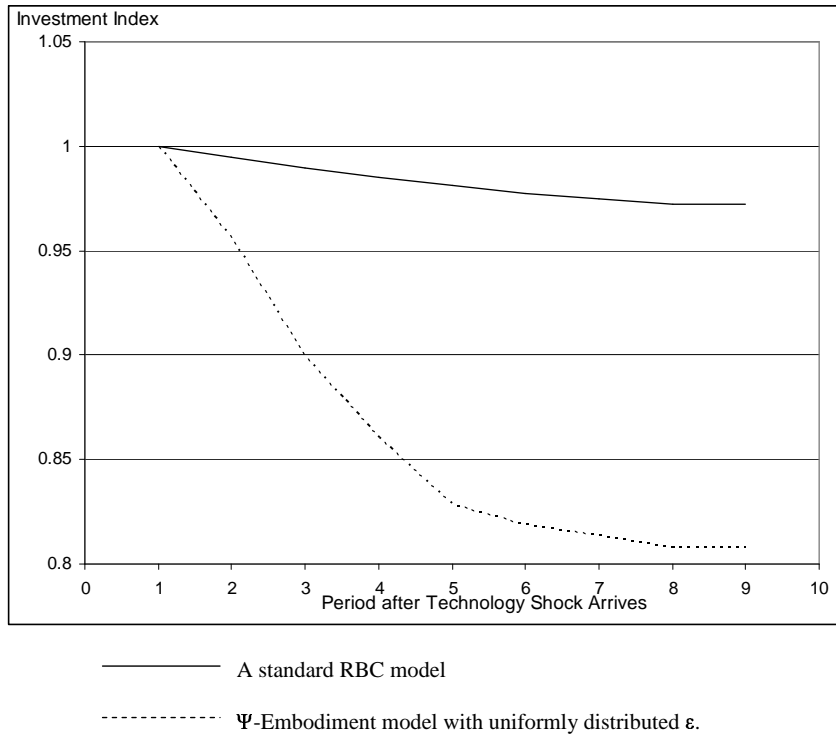
48

Investment Index

Period after Technology Shock Arrives

———————— A standard RBC model

------------ Ψ-Embodiment model with uniformly distributed ε.

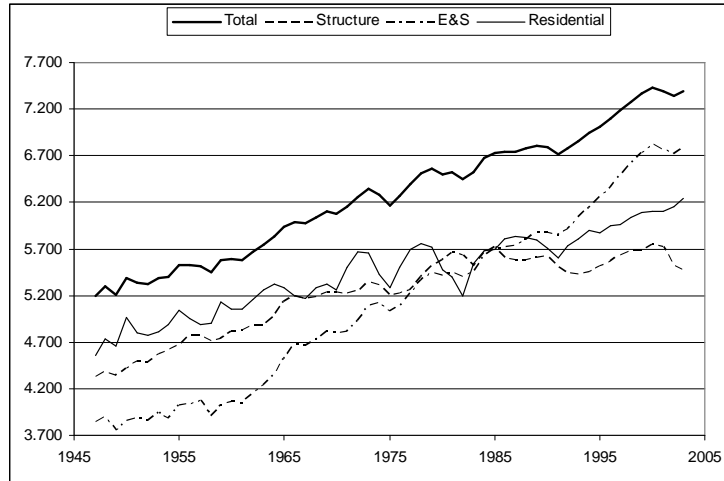Figure 6: Impulse Response in Two Alternative Environments
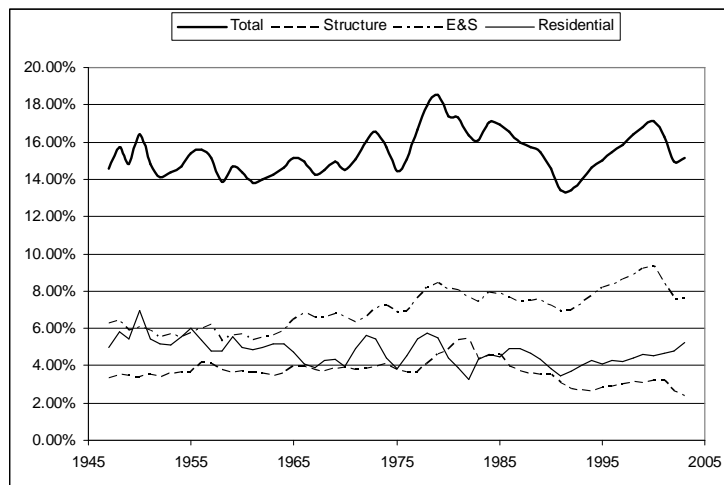
49
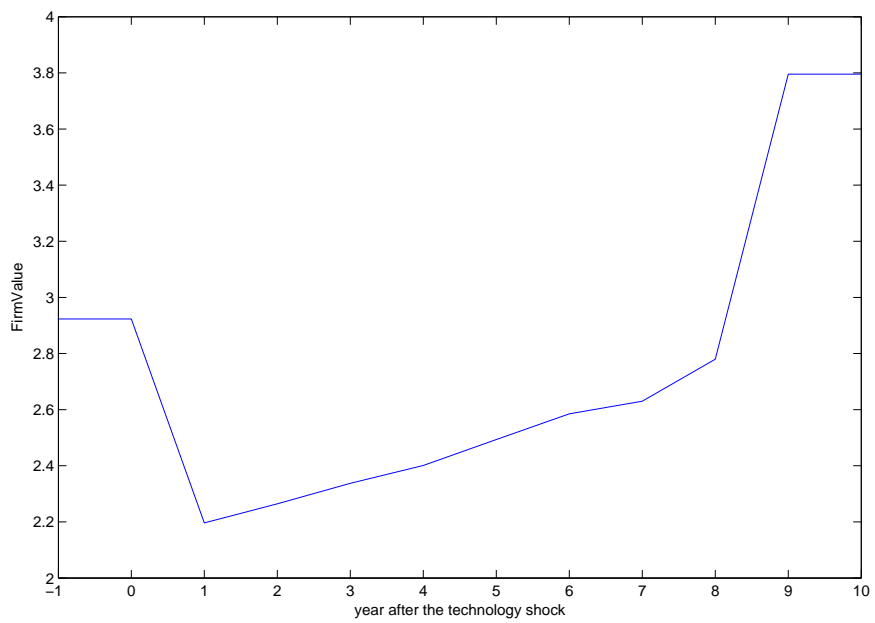
Figure 7: Firm Value Dynamics

Log of Real Private Investment and Its Components

GDP Share of Investment

Data Source: National Income and Product Accounts Tables

Figure 8: Investment Levels and Shares

51