Learning By Investing

Embodied Technology and Business Cycles

Geng Li *

November 10, 2004

Department of Economics
University of Michigan

KEYWORDS: Embodied Technology, Learning, Overinvestment

JEL Classification: E22, E32

*I thank my dissertation committee members Bob Barsky, Matthew Shapiro, Dmitriy Stolyarov and Tyler Shumway for their superb guidance and priceless encouragement. Part of the research of this paper has been done when the author was a Ph.D. Dissertation Intern at the Federal Reserve Board. Their hospitality and support are gratefully acknowledged. I thank the seminar participants at the Economics Department and Business School of the University of Michigan and Federal Reserve Board for helpful comments they gave. I am also very much grateful for the discussions with Susanto Basu, Jason Cummins, Rochelle Edge, Alexander Field, Chris House, Mike Kiley, Miles Kimball, John Laitner, Andreas Lehnert, Steve Oliner, Mike Palumbo, Dave Reifschneider, Peter Rousseau and Dan Sichel. All errors are my own. Email: gli@umich.edu
Abstract

In the last decade of the 20th century, the U.S. economy witnessed a persistent and substantial increase in private investment. This investment boom was sharply reversed in 2001. Standard equilibrium business cycle models have difficulties in predicting the investment boom and overshooting without assuming ad hoc exogenous technology shocks. I construct an embodied technology model to replicate the pattern of investment boom and collapse. Different from previous models of embodiment, I assume that new technology increases the productivity of capital of all vintages, but only new capital can embody the new technology. In addition, while agents 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 technology regress. The model also sheds some light on such important issues as: why consumption growth was strong during the last recession; why Tobin’s q can be well above 1 persistently; and why new technology did not destroy the value of old capital but instead has raised its value during the 1990s.
1 Introduction

The most recent business cycle demonstrated many interesting characteristics that are at variance with previous cycles. One of the most remarkable features of the recent economic expansion is the persistent and substantial increase in private investment, both in terms of the level and share of GDP. This investment growth stopped abruptly in the second half of year 2000. In the subsequent years, there was a sharp decrease of investment and at the same time the whole economy was dragged into a recession.

One important characteristic of this investment boom is that the economy may have accumulated too much capital and has overinvested, especially in certain sectors. For instance, it was reported that over 90 percent of the optical fiber cables installed during the 1990s were left unused, resulting in thousands of miles of "dark fiber". At the moment when the investment boom was to crash, the overinvestment phenomenon had already drawn considerable amount of attention from entrepreneurs and policy makers. On July 14, 2000, close to the very 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". Similarly, one year later after the investment boom collapsed and the economy was deep into the 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".

1Such a level of "dark fiber" is partially due to the advent of a new signal transmitting technology, which dramatically increased the per-cable transmitting capacity. However, even after taking into account the effect of this new development, the over-capacity of optical fiber cable is still very substantial. I thank Alexander Field for pointing this out to me
Another interesting aspect of the last recession is that it seemed to be a relatively painless recession in the sense that consumption was not hurt much. Consumption growth, especially for durable goods, was much stronger than the GDP growth during the recession years, 2001 and 2002.

In this paper I construct an embodied technology model to explain why the last business cycle exhibited the interesting characteristics described above. In my model, new technology has to be embodied into new capital goods before it can increase the total factor productivity (TFP). Different from the traditional embodiment models, a la Solow(1960), I assume that new technology increases the productivity of capital of all vintages, instead of the new vintage capital only. Subject to the limits of the available new technology, the more the economy has invested in new capital, the higher the TFP is. On the other hand, I assume that while the agents know the advent of a new technology, they have imperfect information about the magnitude of the new technology. Agents have to learn the magnitude of the new technology by investing in new capital goods. They observe the output and assess whether they have invested beyond the optimal amount. If they have not, they revise up their beliefs about the magnitude and invest more in subsequent periods, hence an investment boom follows. If they find that investment has overshot, subsequent investment will be sharply reduced and a recession will be triggered. During the recession, resources are reallocated from investment to consumption. Therefore, even the GDP does not increase, we may still observe a healthy growth in consumption.

I first set up the firm’s capital demand problem under the proposed embodiment and learning mechanisms. I contrast the capital demand in this model with that in standard models. Then I provide a sufficient condition about the prior belief over the technology shocks distribution that generates persistent investment booms and overshooting. I also solve the model numerically under 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

---

4The learning mechanism adopted in this paper is similar to that in Zeira(1987), Rob(1991) and Barbarino and Jovanovic(2004). In later sections I contrast my model with theirs. It will become clearer that there are important difference between my model and theirs.
after a reasonably large permanent technology shock arrives, investment can keep increasing for as many as 8 years. Naturally, the model also predicts an output boom as persistent as the investment boom. 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 from its previous level when the economy was booming.

It is well known that the standard dynamic general equilibrium (DGE) models have often been criticized on two grounds. First, the DGE models typically lack of strong internal propagation mechanisms\(^5\). Second, the DGE models usually demand certain levels of technology regress to generate significant recessions\(^6\).

My model contributes to the DGE literature in the above two directions. First, embodiment and learning provide the model with a strong internal propagation mechanism. Because new technology has to be embodied into new capital goods and it takes time to accumulate new capital, TFP cannot jump to the level of the underlying technology. It only increases gradually, so does the output. On the other hand, because the agents have to learn the magnitude of the technology shock, they can be cautious in making investment decisions before they have learned much about the underlying technology. This further stretches the length of booms. Second, 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 has already contained the seed of a future recession. There is no need for any negative technology shocks to trigger recessions.

---

\(^5\)An important critique on the Real Business Cycle (RBC) models on this ground is Cogley and Nason(1995). They find that most RBC models have very weak internal propagation mechanism to produce both positively correlated GDP growth and mean reverting component. For permanent shocks, they find that for most RBC models output will jump to the maximum response and then begins to decline slowly. The empirical impulse response functions show quite different pictures. After the 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.”
This paper also contributes to the literature of investment under uncertainty. Due to the information and learning structure of this model, the uncertainty faced by agents evolves over time. Finally, the paper also sheds light on important financial issues such as why Tobin’s $q$ can be well above 1 persistently; and why new technology did not destroy the value of old capital but instead has raised its value during the 1990s.

Finally it is worthwhile to point out that if one technology shock can generate the persistent investment and output boom as well as sizable overinvestment, this shock should be sufficiently large. It should have a significant contribution to the TFP. Such technology shocks can be motivated as the General Purpose Technology (GPT) that have impacts on most industries of the economy. The point of view that the recent boom was at least partially driven by a GPT finds support in many recent empirical studies.

The paper will be organized as follows: Section 2 presents a model of embodied technology with imperfect information and learning. Then I show how embodiment and learning affect firm’s demand for capital. I provide a sufficient condition under which a new technology leads to a persistent investment boom and overshooting. Section 3 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 facts during the 1990s. I also briefly discuss other historical episodes that share the similar characteristics. Section 6 provides some concluding remarks and directions of future research.

2 The Model of Embodiment and Learning

2.1 Embodiment

Technology is not like fertilizer in most cases. Better fertilizer increases 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

---

7 See Jovanovic and Rousseau (forthcoming) for a thorough treatment on GPT
8 For instance, among others Cummins and Violante (2002) find that due to the IT revolution, there was a 6 percent annual growth in the quality of E&S investment in the 1990s, which is 2 percent higher than the postwar average. This acceleration occurred virtually in every industry, suggesting that IT should be viewed as a GPT.
technology. In the polar case, new technology and new capital can be complements in a Leontief way. Output does not increase after a new technology arrives if no new capital goods have been invested. The empirical relevance of embodied technology has been forcefully established in many studies, see for instance, Hercowitz (1998).

In this paper, I propose a novel model of embodied technology. First I introduce the following assumptions:

**Assumption 1**  *Only capital invested after the advent of the new technology can embody this new technology. The TFP does not increase without new capital being invested after a technological progress.*

**Assumption 2**  *Once it has been embodied into new capital goods, the new technology improves the productivity of capital of all vintages, including the capital existing before the new technology arrives.*

**Assumption 3**  *Subject to the limits of the available new technology, the more new capital has been invested, the higher the TFP is.*

**Assumption 4**  *At the new steady state after a technology shock, the net accumulated new investment is exactly enough to embody the available new technology.*

In this model, there is a distinction between the potential level of TFP and the effective level of TFP after the advent of a new technology. The potential TFP is pinned down by the underlying new technology while the effective TFP is determined by both the amount of new investment and the potential TFP. Furthermore, let \( \{A^0, A^1, \ldots, A^t\} \) be the technology increments due to the arrival of new technologies at period \( \{0, 1, \ldots, t\} \). (\( A^\tau = 0 \) if there is no new technology arrival at period \( \tau \).) Then the period \( t \) potential TFP level is

\[
A_t = A_0 + \sum_{\tau=0}^{t} A^\tau
\]  

(2.1)
Assumption 4 implies that suppose a technology shock $A^t$ increases the underlying technology level from $A_{t-1}$ up to $A_t$, and the economy have invested such amount of new capital since the arrival of $A^t$ that $A^t$ is exactly entirely embodied, i.e., the effective TFP is equal to the potential TFP. Then the aggregated capital stock should increase from the steady state level corresponding to $A_{t-1}$ to that corresponding to $A_t$.\(^9\) To fix the idea, I focus on the dynamics after only one technology shock without explicitly modelling recurrent shocks. The following assumption describes the pre-shock state and the technology shock:

**Assumption 5** suppose that the pre-shock level of technology is $A_0$, which is normalized to be equal to 1, and the economy is at the steady state before a new technology arrives. This new technology increases the underlying TFP level from 1 to $1 + \epsilon$. The post-shock level of technology is denoted by the generic letter $A$.

Under this assumption, before reaching the new steady state, the effective TFP level can be represented as a function of the net new investment, which is equivalent to a function of the current capital stock. In particular, we have the following functional relationship:

$$\tilde{A}_t = \min [A, Z_t] \quad (2.2)$$

$$Z_t = \Psi(K_t) \quad (2.3)$$

where $\tilde{A}_t$ is the effective TFP level at period $t$; and $Z_t = \Psi(K_t)$ is the level of technology that can be embodied into $K_t$ given that $K_0$ has already embodied technology level $A_0$.\(^{10}\) $\Psi(\cdot)$ is increasing and concave. $\Psi'(\cdot) > 0$ and $\Psi''(\cdot) < 0$

Equation 2.2 requires that the level of effective TFP is equal to the minimum between the potential TFP and the level of technology that can be embodied into the current level of $K_t$.\(^9\)Assumption 4 has two implications: First, it stresses out the importance of embodiment. Under this assumption, the process of embodiment does not end until the economy gets to the new steady state associated with the higher technology level. Second, this assumption makes the calibration of the embodiment function easier, as will become clear later in this section.

\(^9\)Put differently, because the net accumulated new investment is equal to $K_t - K_0$, $\Psi(K_t)$ can be equivalently interpreted as the pre-shock level of technology plus the portion of the underlying new technology that has been embodied into the net accumulated new investment up to period $t$, $K_t - K_0$.\(^{10}\)
Because that $\Psi(\cdot)$ is increasing, equations (2.2) and (2.3) imply that before the effective TFP hits the bound of the potential TFP, the effective TFP is increasing with $K_t$. Hence Assumption 3 holds.

Assuming that the only input is capital, the production function is set up as:

$$Y_t = \tilde{A}_t \times K_t^\alpha$$

For reference purpose, I will call $\Psi(\cdot)$ the embodiment function and the above embodiment model the $\Psi$-type embodiment. To further carry out the analysis, we need to nail down a specific functional form of $\Psi(K_t)$. There are a few a priori criteria I impose when choosing the functional form. First it has to be consistent with Assumption 1, which requires that

$$\Psi(K_0) = 1$$

Second, $\Psi(\cdot)$ should preserve the neoclassical (disembodied) quantity relationships at the steady states when these relationships are derived from production function (2.4). This is because that the neoclassical growth model performs remarkably well in explaining long run economic growth and therefore it is desirable to make my model be consistent with the neoclassical model in the long run. In addition to that, this paper focuses on business cycle frequency fluctuations, hence preserving the steady state neoclassical relationships does not affect the core results. More specifically I require that adding the $\Psi$-type embodiment into the otherwise standard production function only affects the path through which the economy converges from the pre-shock steady state to the post-shock steady state, whereas at the steady states, the neoclassical production function and (2.4) imply the same results.

Therefore, let $K^*(A)$ denote the steady state level of capital corresponding to technology level $A$ that is derived from a neoclassical production function, $Y = AK^\alpha$. The above restriction requires that $\Psi(\cdot)$ should satisfy 11:

$$\Psi[ K^*(A) ] = A \quad \forall A$$

11We may think $\Psi(\cdot)$ as the inverse function of $K^*(A)$. In addition, due to Assumption 4 and 5, $\Psi(\cdot)$ has
For the standard Cobb-Douglas production function \( Y = AK^\alpha, \forall A_0, A_1 \), we have:

\[
\frac{A_1}{A_0} = \left[ \frac{K^*(A_1)}{K^*(A_0)} \right]^{1-\alpha}
\]  

Plug \( A_0 = 1, K_0 = K^*(A_0) \) from Assumption 5 into the above equation and use restriction (2.6), we get

\[
\Psi(K_t) = \left[ \frac{K_t}{K^*(A_0)} \right]^{1-\alpha} = \left( \frac{K_t}{K_0} \right)^{1-\alpha}
\]  

This relationship should be interpreted as follows: a post-shock level of capital embodies such amount of new technology that this level of capital is at the steady state corresponding the level of the technology it embodies. Here the assumption that \( K_0 \) is at the steady state is important. Without this assumption the relationship (2.6) is not appropriate 12.

Plug (2.3) into (2.4) we can rewrite the production function as

\[
Y_t = \begin{cases} 
K_0^{\alpha-1} \times K_t & \text{if } K_t \leq K^*(A) \\
AK_t^\alpha & \text{if } K_t > K^*(A)
\end{cases}
\]  

(2.10)

Figure 1 is an illustration of the comparison between the standard neoclassical production function and the production function with the \( \Psi \)-type embodied technology. Curve \( \hat{OD} \) is the production function before new technology arrives, \( Y = A_0K^\alpha \). Curve \( \hat{EBC} \) is the post-shock production function under neoclassical disembodied technology. In the model, output jumps from \( Y_0 \) to \( Y'_0 \) after the shock, even without any investment in new capital. The post-shock production function is simply the pre-shock function scaled up by the magnitude of the technology shock. Curve \( \hat{OBC} \) is the production function with the \( \Psi \)-type embodied technology. Before the an additional desirable property. \( K^*(A) \) is a function of the interest rate, \( r \) or the discount factor, \( \beta \). However, the embodiment function \( \Psi(\cdot) \) should not be a function of either \( r \) or \( \beta \), conceptually. Indeed, because how much new technology the economy has embodied depends on the amount of new capital invested relative to the previous steady state capital level (which is a function of \( r \) or \( \beta \)), as long as \( r \) and \( \beta \) are constants, \( \Psi(\cdot) \) does not explicitly involve \( r \) and \( \beta \).

12Remember only capital invested after the arrival of a new technology can embody it. Suppose otherwise that \( K_0 \) is higher than \( K^*(A_0) \) hence the economy has more capital than needed to embody \( A_0 \) before the new technology arrives. When technology level jumps to \( A \) after the shock, those inefficient old capital, the amount of which is \( K_0 - K^*(A_0) \), cannot embody the new technology. However, if one simply plug \( K_0 \) into

\[
\tilde{A} = \min \left\{ A, \Psi(K_0) \right\} = \min \left\{ A, \left[ \frac{K_0}{K^*(A_0)} \right]^{1-\alpha} \right\}
\]  

(2.9)

\( \tilde{A} \) would have increased, which is contradicting to Assumption 1. In this case, if there is inefficient capital before the arrival of a technology shock, we need to subtract the residual of these inefficient capital from \( K_t \).
capital stock approaches to the new steady state $K^*[(1+\epsilon)A_0]$, the production function is linear, as indicated in equation (2.10). Beyond $K^*[(1+\epsilon)A_0]$, the $\Psi$-type production function coincides with the neoclassical production function. The existing capital has sufficiently embodied the available new technology and additional capital only bears normal and decreasing return.

To summarize, the $\Psi$-type embodiment production function (2.4) demonstrates three properties of embodiment. First, new technology, if not embodied in new capital, cannot increase TFP. Second, productivity of new capital and old capital are both increased by the new technology. Third, up to a certain limit, the more new capital is invested, the higher is the effective TFP.

2.2 Imperfect Information and Learning

Most of the DGE models assume at least ex post perfect information about technology progress, i.e., after the technology shock hit the economy, agents know about how big the shock is. In many cases, this does not necessarily hold. It is not unrealistic to assume that agents only have imperfect information about the technology progress. More often than not, agents know of the advent of a new technology but know little about how productive this technology is at an early stage, even after firms start to adopt this new technology in production. This is because that the agents do not know for sure if there is still any further potential to exploit this new technology. They learn how good this new technology is gradually and possibly from various ways including learning by doing, human capital accumulation and etc.. In this paper, I will focus on one of the channels of learning, namely, learning by investing. I first introduce the following assumptions:

**Assumption 6** After a technology progress takes place, the firm knows that there is a new technology but does not know exactly the new level of the underlying technology, $A$. However, the firm does know the distribution of the shock. The cumulative density function (CDF) of $A$ is $\Phi(A)$ and the probability density function (PDF) of $A$ is $\phi (A)$

**Assumption 7** The capital a firm invested in period $t$ becomes productive in period $t+1$. 

10
Assumption 8 The technology shock affects all firms in the economy uniformly. However, the firms are able to learn the magnitude of the new technology only from their own activities. There is no information externality and spill over. Firm A can not learn about the magnitude of the shock by analyzing Firm B’s information.

Mechanically learning is carried out as follows: The firm observes capital level $K_t$ and output level $Y_t$. By (2.10), as long as $K_t \neq K_0 \times A^{1-\alpha}$, $K_0^{\alpha-1} \times K_t \neq A K_0^{\alpha}$. Therefore firm can unambiguously tell if investment has overshot or not by applying the following rules\textsuperscript{13}:

$$Y_t = K_0^{\alpha-1} \times K_t \quad \Rightarrow \quad \text{Undershooting}$$

$$Y_t \neq K_0^{\alpha-1} \times K_t \quad \Rightarrow \quad \text{Overshooting} \quad (2.11)$$

If investment has overshot, $Y_t = AK_t^{\alpha}$. In the case where the firm finds that investment has overshot, the level of the underlying technology $A$ can be discovered as:

$$A = \frac{Y_t}{K_t^{\alpha}} \quad (2.12)$$

since the firm does observe $Y_t$ and $K_t$. Otherwise if undershooting is observed, the firm is able to infer that the underlying technology is at least as large as the level that the current level of net new investment has embodied. Hence, we have

$$A \geq Z_t = \left( \frac{K_t}{K_0} \right)^{1-\alpha} \quad (2.13)$$

and will make decisions about future investment based on the updated distribution of $A$ conditional on (2.13), whereas the conditional PDF function is $\frac{\phi(A)}{1 - \Phi(Z_t)}$.

So far, I have discussed the investment dynamics in response to only one wave of technology shock. This is not a too bad set up if we view the technology shock as a GPT innovation. Such technology progress does not happen every quarter or every year. Indeed, there are only

\textsuperscript{13}For the knife edge case where $K_t = K_0 \times A^{1-\alpha}$, I assume that the firm infers that investment has not overshot.
a handful identifiable GPT innovations, such as electricity, automobiles, nuclear power and IT, in the last century. Arguably after one wave of GPT innovation there is sufficient time to allow the economy to move quite close to the steady state before the next wave of GPT innovation. However, it is straightforward to extend the model to allow for compound technology shocks, though this requires that more state variables have to be added.

2.3 Demand for Capital

Different from the standard models with disembodied technology, investment in new capital plays multiple roles in this paper. First, the newly invested enters the production function in the $K^\alpha_t$ part and plays the conventional production role; second, new capital enters the embodiment function, $\Psi(K_t)$, and has the role to embody new technology; and finally new capital facilitates the learning process about the magnitude of the underlying technology shock. Because this model is the first effort (as far as the author understands) to synthesis embodiment and learning, it is important to understand how these new ingredients affect a firm’s demand for capital comparing with the standard models. I will compare the firm’s capital demand function of the following three cases:

Case I. Investment under technological uncertainty with no embodiment and learning;

Case II. Investment under uncertainty and the $\Psi^-$ type embodiment but no learning and

Case III. Investment under uncertainty, $\Psi^-$ type embodiment and learning.

Consider that a firm maximizes discounted sum of profit flow under the constant interest rate, $r$

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

subject to$^{14}$:

$$K_t = K_{t-1} + I_{t-1}$$

$^{14}$For the time being, assume that the depreciation rate is zero
Case I

To begin with, consider at time zero the firm expect that a *disembodied* technology shock arrives next period. The firm has to invest before the post-shock level of technology, $A$, is observed. The capital invested will become productive at $t = 1$ when the shock arrives. The magnitude of the shock will be completely revealed at that time as well. In this case, the firm’s optimality condition only requires that the expected marginal product of capital is equal to the borrowing interest rate, $r$:

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

or,

$$r = \bar{A} \alpha K_1^{\alpha-1}$$

where $\bar{A}$ is the unconditional mean of $A$. At $t = 1$, $A$ is observed, the optimality condition for investment in this period is simply the familiar

$$r = \alpha A K_2^{\alpha-1}$$

Case II

Suppose that the technology progress is the $\Psi$-type embodied technology as specified in equation (2.4) and (2.3). I keep the assumption that $A$ will be revealed at $t = 1$. We have the following optimality condition for $K_1$:

**Lemma 1**

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

**Proof:** See Appendix A.

The first term on the right hand side (RHS) in equation (2.16) is the expected marginal product of capital if the investment of $t = 0$ has not overshot whereas the second term is the expected marginal product of capital if $I_0$ has overshot. Because $A$ is assumed to be learned at
$t = 1$, regardless whether investment has overshot, (2.17) still holds as the optimality condition for $K_2$.

**Case III**

Finally, consider that the firm has to make investment decisions after an $\Psi$-type embodied technology shock and has to learn the magnitude of the shock through the mechanism described in (2.11). In this case, with the same amount of capital, the value of the firm is contingent on if investment has overshot. Let $V(K)$ be the value of the firm that has not overshot and let $W(K, A)$ be the value of the firm if the firm has overshot with capital stock $K$ and has learned that the true magnitude of the post-shock level of the underlying technology is $A$. Note that $V$ is a function of $K$ only whereas $W$ is a function of both $K$ and $A$. This can be seen from Figure 2. The graph shows a typical PDF of $A$. Suppose a firm that has not overshot with capital $K_1$ invests up to capital level $K_2$ and again finds that investment has not overshot yet. In such a scenario, the firm learns that $A > Z_2$. The current capital stock $K_2$ contains all the information available about the conditional distribution of $A$. However, if the same firm realizes that investment has overshot with capital $K_2$, then all levels of $A$ that are between point $A$ and $B$ can induce this observed overshotting. Investment can either have overshot by a lot if the true level of $A$ is close to point $A$, or have overshot by only a little bit if the true level of $A$ is close to $B$. The value of a firm that has overshot depends on both capital level $K_2$ as well as the level of the true $A$. *Ex ante*, the firm can compute the probability of overshotting when increase capital from $K_1$ to $K_2$, but even conditional on overshotting, the firm still does not know exactly about the magnitude of the shock. It has to take expectation with respect to the $A$s between $A$ and $B$.

Define $\pi(K_t, K_{t+1})$ to be the probability of remaining undershotting when the firm increases its capital stock from $K_t$ to $K_{t+1}$, and $\pi(K_t, K_{t+1})$ can be computed as

15 The numerator is the unconditional probability of having not overshot with capital stock $K_{t+1}$. This is the double shaded area in Figure 4. Likewise, the denominator is the unconditional probability of having not overshot with capital $K_t$. This is the horizontally shaded area in Figure 4. Because when a firm increases its capital from $K_t$ to $K_{t+1}$, it has already known that $A_t > Z_t$, therefore, $\pi(K_t, K_{t+1})$ should be computed as the ratio between the double shaded area and the horizontally shaded area.
\[ \pi(K_t, K_{t+1}) = \frac{1 - \Phi(Z_{t+1})}{1 - \Phi(Z_t)} \] (2.19)

The firm’s problem can be rewritten as the following Bellman equations system:

\[ V(K_0) = \max_{I_0} Y_0 - I_0 + \frac{1}{1+r} \left\{ \pi(K_0, K_1) \times V(K_1) + \cdots + [1 - \pi(K_0, K_1)] \times E_A \left[ W(K_1, A) \mid 1 < A < Z_1 \right] \right\} \] (2.20)

where \( V(K_1) \) is recursively defined as

\[ V(K_1) = \max_{I_1} K_0^{\alpha-1} K_1 - I_1 + \frac{1}{1+r} \left\{ \pi(K_1, K_2) \times V(K_2) + \cdots + [1 - \pi(K_1, K_2)] \times E_A \left[ W(K_2, A) \mid Z_1 < A < Z_2 \right] \right\} \] (2.21)

and

\[ W(K_1, A) = \max_{[I_i]} \sum_{t=1}^{\infty} \left( \frac{1}{1+r} \right)^{t-1} [AK_1^{\alpha} - I_t] \] (2.22)

In equation (2.20) the term \( E_A \left[ W(K_1, A) \mid 1 < A < Z_1 \right] \) is the expected value of an overshoot firm. The expectation is taken with respect to \( A \) conditional on \( 1 < A < Z_1 \) because only a level of \( A \) within this range can make \( K_1 \) overshoot. For state variable \( K_1 \) we have the following optimality condition:

**Lemma 2**

\[ r = \left[ K_0^{\alpha-1}[1 - \Phi(Z_1)] + \alpha K_1^{\alpha-1} \int_1^{Z_1} A \phi(A) dA \right] + \phi(Z_1)\Psi'(K_1)[W(K_1, Z_1) - V(K_1)] + \frac{1}{1+r} \times \phi(Z_1) \Psi'(K_1) \times \left[ \int_{Z_1}^{\infty} V(K_2) \phi(A) dA + \int_{Z_1}^{Z_2} W(K_2, Z_2) \phi(A) dA \right] \]

\[ 1 - \Phi(Z_1) \]

\[ - W(K_2, Z_1) \]

\[ (2.23) \]

**Proof:** See Appendix A

In the subsequent period, if the firm sees that \( K_1 \) has overshot, the size of \( A \) will be learned and the optimality condition for \( K_2 \) will be the same as (2.17). Otherwise, the size of \( A_1 \) remains
unknown, but the conditional PDF of $A_1$ becomes $\frac{\phi(A)}{1 - \Phi(Z_1)}$. The optimality condition for $K_2$ is simply (2.23) forwarded by one period with all PDFs adjusted by the factor $1 - \Phi(Z_1)$ to convert the unconditional probability to conditional probability.

The RHS of equation (2.23) has three components, the intuition of which are discussed as follows:

1. $K_0^{\alpha-1}[1 - \Phi(Z_1)] + \alpha K_1^{\alpha-1} \int_1^{Z_1} A \phi(A) dA$ is the $\Psi$-type embodiment affected expected marginal productivity. This is identical to the RHS of (2.18)

2. As illustrated in Figure 4, the extent to which investment overshoots varies, hence the ex post loss of overshooting varies. $\phi(Z_1)\Psi'(K_1) \times [W(K_1, Z_1) - V(K_1)]$ is the marginal gain due to the reduction of the loss of overshooting in the subsequent periods. To see this, assuming that the firm has capital $K_1$ and the true size of $A$ is equal to $Z_1 + \eta$, where $\eta$ 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 $\eta$ is sufficiently small, we then should have $W(K_1, A) \approx W(K_1, Z_1)$. Notice that $K_1$ 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 due to 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 does have undershot with $K_1$, the term

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

$$\frac{1}{1 - \Phi(Z_1)} - W(K_2, Z_1)$$

captures the value of the addition information provide 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 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 $A > Z_1$, 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 $A$.

Compare the optimality conditions (2.16) and (2.18) with (2.23) we find that with the learning mechanism, the optimality condition is no longer static. It involves the state variable in the current period and in the subsequent period.

With respect to the level of capital demand, we have the following result:

**Proposition 1** *Ceteris paribus, the optimal investment for Case 2 is higher than that for Case 1, whereas the optimal investment for Case 3 is higher than that for Case 2.*

**Proof** See Appendix A.

The intuition is simple. Because in a model with the $\Psi$–type embodiment and learning, investment has new roles besides production, the marginal value of investment is higher, therefore the firm will invest more.

### 2.4 Investment Dynamics

Now we discuss the condition under which investment increases persistently like what we have observed in the 1990s. Heuristically, this requires some properties related with the distribution of $A$. For instance, conditional on current capital stock has undershot, the posterior expectation of $A$ should increase at a sufficiently fast rate. In this section I will provide a sufficient condition on the distribution of $A$ so that investment will be persistently increasing, i.e., $K_{t+1} - K_t > K_t - K_{t-1} \forall t$.

Without loss of generality, I focus on the condition under which the optimal $I_1$ is greater than the optimal $I_0$, or $K_2 - K_1 > K_1 - K_0$. To have $I_1 > I_0$, it is sufficient to show that the
increment of the technology embodied from capital stock $K_1$ to $K_2$ is bigger than the increment from $K_0$ to $K_1$. This is due to the concavity of $\Psi(K_1)$. Specifically, we have:

$$\left(\frac{K_2}{K_0}\right)^{1-\alpha} - \left(\frac{K_1}{K_0}\right)^{1-\alpha} > \left(\frac{K_0}{K_0}\right)^{1-\alpha}$$

\[\Rightarrow K_2^{1-\alpha} - K_1^{1-\alpha} > K_1^{1-\alpha} - K_0^{1-\alpha}\]

\[\Rightarrow K_2 - K_1 > K_1 - K_0 \quad (2.24)\]

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

$$r = \int_{Z_1}^{\infty} A \frac{\phi(A)}{1 - \Phi(Z_1)} dA \quad (2.25)$$

$\tilde{K}_2$ is the optimal capital level of the Case I in Section 2.3 conditional on $A > Z_1$. By Proposition 1, we know that the optimal capital level under the $\Psi$-type embodiment and learning should be greater than $\tilde{K}_2$. Therefore, we have

$$\tilde{K}_2 - K_1 > K_1 - K_0 \Rightarrow K_2 - K_1 > K_1 - K_0 \quad (2.26)$$

Finally, combine the relationship (2.24) and (2.26) and let $\tilde{Z}_2$ denote the level of technology embodied by $\tilde{K}_2$, we have

$$\tilde{Z}_2 - Z_1 > Z_1 - Z_0 \Rightarrow K_2 - K_1 > K_1 - K_0$$

(2.27)

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

**Lemma 3**

$$\tilde{Z}_2 = \int_{Z_1}^{\infty} A \frac{\phi(A)}{1 - \Phi(Z_1)} dA \quad (2.28)$$

**Proof:** See Appendix A

Now we reach the following proposition:
Proposition 2  In order to have $I_1 > I_0$, it is sufficient if we have

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

(2.29)

An Example

Suppose that $A$ follows the Pareto Distribution, $\Phi(A) = 1 - \frac{1}{A^\omega}$, $\phi(A) = \frac{\omega}{A^{\omega+1}}$, where $\omega > 1$ is the distribution parameter. For the expectation of $A_1$ conditional on $A > Z_1$, we have

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

(2.30)

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

This example gives us some insightful hints on the shape of the distributions that are prone to induce the increasing investment. The Pareto Distribution is heavily tailed. The fat tail property is particularly pronounced for the cases where $\omega \in (1, 2)$. 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

$$H(A) = \frac{\omega}{A}$$

(2.31)

which is decreasing with $A$. Why the hazard rate matters on how much a firm should invest? The hazard rate is positively related with the probability of overshooting. If the overshooting probability is high, the firm should be more cautious about investment. A decreasing hazard rate implies a decreasing probability of overshooting. Because for the same amount of targeted increment of TFP it is less likely to be overshot, the firm should engage in more aggressive investment plan.

It is worthwhile to point out that Proposition 2 is a very strong condition. Indeed, because the concavity of the $\Psi(\cdot)$ function and the difference between $K_2$ and $\tilde{K}_2$ can be quite big, a distribution that does not have a tail as fat as the Pareto Distribution can generate the increasing investment as well. However, the analytical necessary condition for increasing investment is not available.
2.5 Discussions

My model differs from the standard models of embodied technology in many aspects. In this subsection, I will discuss the assumption I made and contrast my model with the conventional models.

Solow (1960) is one of the classical contributions on formalizing the idea of embodied technology. In his model, capital of a particular vintage embodies the technology of the same vintage. New technology has to be embodied into new capital goods before it increases the level of TFP. This is consistent with the Assumption 1.

In the Solow-type embodiment model, total output is the sum of the output produced by capital of various vintages. Let $Y_{\nu,t}$ denote the output produced by equipment of vintage $\nu$ in time $t$, we have:

$$Y_{\nu,t} = f(A_{\nu,t}, K_{\nu,t}) \quad (2.32)$$

where $A$ and $K$ are technology level and capital input, respectively and both are indexed by its vintage, $\nu$ and time, $t$. The aggregated output at time $t$ is given by

$$Y_t = \sum_{\nu=0}^{t} f(A_{\nu,t}, K_{\nu,t}) \quad (2.33)$$

One important property of the Solow-type embodiment is that a technology progress only makes the capital invested afterward more productive. The productivity of the capital of older vintages does not change. In particular, we have $\frac{\partial MPK_{\nu}}{\partial A_{\nu'}} = 0$ for $\nu < \nu'$ where $MPK_{\nu}$ is the marginal product of capital of vintage $\nu$. Consequently, after a new technology arrives, the market value of old capital decreases because they become 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 reconcile 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\(^{16}\).

\(^{16}\)The IT revolution did not arrive at its full power until years later. However, it is argued that this new
This approach has difficulties to explain the stock market dynamics in the 1990s. It is well known that the vigorous boom of the NASDAQ index was due to the success of the IT technology. On one hand, the market value of the new entrants that embraced this new technology surged. On the other hand we have not seen that the value of the incumbent firms was destroyed as what happened in 1974. Rather, the S&P 500 index, which heavily consists of traditional industry firms, rose dramatically as well, although the magnitude of the S&P run-up was not as large as the NASDAQ. This fact hints that the capital of both old and new vintages had benefitted from the spreading of the IT technology.

Assumption 3 assumes that once has been embodied, the new technology increases the productivity of capital of all vintages, including those exist before the arrival of the new technology.

Another aspect on which my model is that the Solow-type embodiment does not have the distinction between effective TFP and potential TFP. The Solow model assumes that new technology is indivisible and can be entirely embodied in the first slice of the new capital. The productivity level is irrelevant with the amount of new investment. In contrast, in Assumption 2 I explicitly assume that subject to a certain limit, there exists a monotonic increasing functional relationship between the amount of investment in new capital and the extent to which new technology helps to increase the TFP.

To motivate Assumption 2 and 3, think about the following imaginary example: An automobile company is producing cars using its assembly lines, which are controlled by a computer system. Suppose that there is a new technology that doubles the computer processing speed. A faster computer system can operate the assembly lines more efficiently. If the firm invests in new computers, it is the productivity of the whole firm (both the assembly lines and the controlling computer system) that will increase.

Furthermore if the firm has only replaced ten percent of its old controlling system with faster computers, presumably its overall productivity should not be as high as if twenty percent has been replaced. On the other hand, investment in new computers alone can not beef up the technology was heralded and well understood by the middle of the 1970’s.
productivity without a limit. The firm can exploit the new technology by investing in new computer only up to a certain limit. This limit is intrinsically determined by the magnitude of the new technology.

To summarize, in the Solow-type embodied technology model, because only capital of the latest vintage has higher productivity due to the arrival of the new technology, the firm will replace old capital with new capital. 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 polar, if technology is disembodied, a technology progress would imply a jump of TFP to the post-shock level immediately. The $\Psi$-type embodiment introduced in this paper implies a time series path of effective TFP lying between the above two polar cases. Figure 3 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 irreversible capital. the TFP only asymptotically converges to $A_1$ since old capital is gradually replaced by the capital that embodies the new technology. Curve B represents the TFP path implied by the $\Psi$-type embodiment. At time $t$, the firm learns the news about a technology progress and start to invest. It accumulates sufficient amount of capital at $t'$ to pick up all of the new technology and from then on it coincides with Curve A. However, the firm yet has not realized this until observing overshooting at $t''$. Therefore the expected TFP level between $t'$ and $t''$ exceeds the true level of the underlying technology.

Assumption 7 assumes the length of time-to-build is equal to one period. This is simply for convenience. One consequence related with this assumption is that the firm can make investment that leads to overshooting for at most one period. However, the model can be generalized to allow for a $T$—period gesture delay, which can be either time-to-build or time-to-plan. Under such extension, the firm may observe overshooting for multiple periods. As long as the investment on pipeline and the investment decisions are irreversible, the extent to which the firm may overinvest is largely irrelevant to the choice of investment-decision-making frequency(including
continuous time modelling). Take a $T$-period time-to-build setup as an example. Suppose at time $t$ the capital invested at $t - T$ becomes productive and the firm realizes that the investment has overshot. The firm would hope that it had not invested after $t - T$. However, since it cannot reverse the capital that has already been invested, its capital will keep increasing for the next $T - 1$ periods. Therefore, in a high frequency model, although the investment per period can be infinitesimally small, the total amount of capital overinvested can still be substantial if the time-to-build is sufficiently long.

Assumption 8 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 investment until other firms have overshot and the true magnitude of the shock is completely revealed. This will further complicate the analysis. I will postpone the discussion on this till a later section. 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 very differently 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, it may require firm-specific knowledge to incorporate the new capital and new technology with the existing capital.

Finally, I want to briefly discuss how my model is related with the dynamics of Tobin’s $q$. It is well known that Tobin’s $q$ has been persistently above 1 for most of the post war era\(^\text{17}\). Suppose the Hayashi condition holds, the marginal $q$ should be persistently above 1 as well. On the other hand, the textbook $q$ Theory taught us that, at the presence of investment adjustment cost, marginal $q$ should decrease and converge to 1 as capital converges to the optimal level. The uncertainty structure and learning mechanism proposed in this paper contribute to reconcile this phenomenon with the theory.

In the standard model with time invariant information about the level of the underlying technology, marginal $q$ is equal to the sum of the discounted stream of future marginal products plus \(^\text{17}\)The exception is the period between 1974 and 1984, as Laitner and Stolyarov(2003) point out.
the sum of the discounted stream of its marginal contributions to the reduction in future capital installation costs. When more capital is invested overtime, the marginal products diminish and therefore the marginal $q$ decreases. This is not the case when learning with respect to the level of technology is involved. After each round of investment, the firm sees if it has overinvested. As long as the firm yet has not overshot, it will revise up its expectation about the underlying technology. Therefore, even after more capital is in place, the expected marginal products do not have to diminish. This mechanism supports a level of marginal $q$ that is persistently above one. Depending on how the conditional expectation about the underlying technology evolves, the marginal $q$ can even increase over time, as the data from the 1990s seem to suggest.

3 A Dynamics General Equilibrium Model

To understand quantitatively how the $\Psi$-type embodiment and learning affect the investment and output impulse response functions (IRF) with respect to technology shocks, I construct a dynamic general equilibrium (DGE) model as follows:

3.1 The Household

A representative agent maximizes the discounted sum of future utility over an infinite time horizon. The agent supplies one unit of labor inelastically every period and earns competitive wages. He owns the shares of the firm and receives dividend payment each period. He can trade the shares to smooth his consumption. The representative household’s problem is

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

subject to:

$$C_t + P_t S_t = (P_t + D_t) S_{t-1} + W_t L_t^s$$

$$L_t^s \equiv 1$$

The notations used are listed below

24
$C_t$ 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 firm share price at the end of period $t$, it is the ex-dividend share price

$W_t$ is the wage rate

$L_t^s$ is the amount of labor supplies in period $t$ and $L_t^s$ is a constant equal to one for all $t$.

### 3.2 The Firm

The representative competitive firm maximizes sum of the discounted dividend flow:

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

subject to:

$$
D_t = Y_t - I_t - \Theta_t - W_t
$$

$$
Y_t = \tilde{A}_t \ K_t^\alpha \ L_t^{1-\alpha}
$$

$$
\tilde{A}_t = \min \{ A, Z_t \}
$$

$$
Z_t = \Psi(K_t)
$$

$$
K_{t+1} = (1 - \delta)K_t + I_t
$$

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

Equation (3.6) is almost the same production function as introduced in Section (2.1) apart from that now the firm uses both capital and labor as input\(^{18}\). $\Theta_t$ is the investment adjustment cost and $c(\cdot)$ is a convex $C^2$ function.

The information structure is similar to that in Section 2.2, the post-shock level of the underlying technology is $A$, the CDF and PDF of which are $\Phi(A)$, and $\phi(A)$, respectively.

\(^{18}\)Adding labor makes the production function has constant returns to scale at steady states
3.3 Equilibrium

The good market clear condition is

\[ Y_t = C_t + I_t + \Theta_t \]  \hspace{1cm} (3.11)

The labor market clear condition is

\[ L_t^s = L_t^d \equiv 1 \]  \hspace{1cm} (3.12)

Finally the equity market clear condition is

\[ S_t \equiv 1 \]  \hspace{1cm} (3.13)

The firm’s labor demand is given by:

\[ W_t = \frac{(1 - \alpha) \cdot Y_t}{L_t} \]  \hspace{1cm} (3.14)

Given initial state of the economy, \( A_0, K_0 \) and the unconditional distribution of \( A, \Phi(A) \), the equilibrium conditions are familiar to us and are shared by many DGE models. An equilibrium is given by a sequence of quantities \{\( C_t, I_t, K_t, Y_t, L_t \)\}\( \infty \) and prices \{\( P_t, W_t \)\}\( \infty \) such that given prices \{\( P_t, W_t \)\}\( \infty \), the representative household solves (3.19) and (3.2); the firm solves (3.4)-(3.10). The markets for labor, share and goods clear.

Because labor is inelastically supplies, labor market equilibrium is simply

\[ L_t^s = L_t^d \equiv 1 \]  \hspace{1cm} (3.15)

The equilibrium wage rate is given by:

\[ W_t = (1 - \alpha) \cdot Y_t \]  \hspace{1cm} (3.16)

The optimal consumption in this model is trivially determined as:

\[ C_t = W_t + D_t \]  \hspace{1cm} (3.17)

Let the marginal value of capital be \( MPK_t \), the optimal condition for investment is the familiar

\[ E_t \left[ \beta \frac{U'(C_{t+1})}{U''(C_t)} (MPK_{t+1} + 1 - \delta) \right] = 1 \]  \hspace{1cm} (3.18)

where \( MPK \) is defined as the LHS of (2.23).
3.4 An Equivalent Social Planner’s Problem

The equilibrium path quantities \( \{ C_t, I_t, K_t, Y_t, L_t \}_{t=0}^{\infty} \) and prices \( \{ P_t, W_t \}_{t=0}^{\infty} \) of the model above can be replicated by a centralized problem. The social planner will solve the following problem:

\[
\max_{E_0} \sum_{t=0}^{\infty} \beta^t U(C_t) \quad (3.19)
\]

subject to:

\[
Y_t = C_t + I_t + \Theta_t \quad (3.20)
\]
\[
Y_t = \tilde{A}_t K_t^\alpha \quad (3.21)
\]

where \( \Theta_t, \tilde{A}_t \) and the capital accumulation equation are defined same as before. Similar to the Bellman equation (2.20) and (2.22), the social planner’s problem can be written as

\[
V(K_0) = \max_{I_0} U(C_0) + \beta \{ \pi(K_0, K_1) \times V(K_1) + \cdots + [1 - \pi(K_0, K_1)] \times E_A [ W(K_1, A) | 1 < A < Z_1 ] \} \quad (3.22)
\]

where

\[
V(K_1) = \max_{I_1} U(C_0) + \beta \{ \pi(K_1, K_2) \times V(K_2) + \cdots + [1 - \pi(K_1, K_2)] \times E_A [ W(K_2, A) | Z_1 < A < Z_2 ] \} \quad (3.23)
\]

and

\[
W(K_1, A) = \max_{|I_t|} \sum_{t=1}^{\infty} \beta^{t-1} U(A K_t^\alpha - I_t - \Theta_t) \quad (3.24)
\]

The analytical solution to this problem is not approachable. This Bellman equations system, however, can be solved numerically by iterating both value function \( V \) and \( W \). Appendix B discusses the computational details.

3.5 Calibration and Numerical Results

The model is calibrated annually and the calibration of most parameters is standard in the literature. First, the preference is assumed to be CRRA,

\[
U(C_t) = \frac{C_t^{1-\sigma}}{1-\sigma} \quad (3.25)
\]
where $\sigma$ is the coefficient of relative risk aversion. Following Hall(2001), the investment adjustment cost function is assumed to be:

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

(3.26)

where $P$ is an indicator function such that $P(K_{t+1} - K_t) = 1$ if $K_{t+1} - K_t \geq 0$ and $P(K_{t+1} - K_t) = 0$ if $K_{t+1} - K_t < 0$. The value of the parameters, $\alpha$, $\beta$, $\delta$, $\theta^+$ and $\theta^-$ are chosen as in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\delta$</th>
<th>$\sigma$</th>
<th>$\theta^+$</th>
<th>$\theta^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.96</td>
<td>0.10</td>
<td>2</td>
<td>4</td>
<td>40</td>
</tr>
</tbody>
</table>

All parameters are calibrated at an annual basis. The calibration of $\theta^+$ and $\theta^-$ follows Hall (2001). As he pointed out, $\theta^+$, the adjustment cost coefficient for positive capital stock changes, can be related with the time needed to double the capital stock. Shapiro(1986) and Summers(1981) provide some empirical evidence on the size of $\theta^+$. Shapiro reported $\theta^+$ is equal to eight calendar quarters, or two calendar years. Summers reported a much larger number for $\theta^+$. His findings suggest that $\theta^+$ is equal to 32 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 Shapiro(1986) reported as the lower bound and 8 years as the upper bound. I synthesis Hall’s calibration by using the geometric average of the lower and upper value he used, which gives me $\theta^+ = 4$. $\theta^-$ is then chosen to be equal to 40, which is ten times $\theta^+$, to capture the irreversibility of installed capital. The level of adjustment cost play 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. Higher adjustment cost keeps the firm from investing too fast and prolongs the learning process. In this setup, the pre-shock steady state is $A_0 = 1$, $K_0 = 2.92$, $Y_0 = 1.38$ and $I_0 = \delta K_0 = 0.292$. At the steady
state, the investment share is 21 percent. This is higher than the investment share in the data. The reason for this discrepancy is that I choose $\delta = 0.1$ whereas the economy wide depreciation rate can be considerably lower.

What still left to be “calibrated” is the distribution of the technology progress. One restriction I imposed is that I that the support of $A$ is $[1, \infty]$ to reflect the notion that there is no technology regress. To begin with, I first set that the magnitude of the shock, i.e., the difference between $A$ and $A_0$ follows an exponential distribution. Let $\epsilon$ denote this shock($A = 1 + \epsilon$), the PDF for $\epsilon$ is $\phi(\epsilon) = \lambda e^{-\lambda \epsilon}$ whereas the CDF is $\Phi(\epsilon) = 1 - e^{-\lambda \epsilon}$. There is only one parameter to pick. One reason for choosing exponential distribution is that it has an analytical closed form CDF, which is convenient when compute the \textit{ex ante} probability of overshooting. Another concern is that the exponential distribution has constant hazard rate, which is the border case between increasing and decreasing hazard rate. If exponential distribution generates the desired dynamics, namely, the persistently increasing investment, then for the families of distributions that have decreasing hazard rate, this dynamics should be more pronounced.

The exponential distribution has only one parameter, $\lambda$, and $E(\epsilon) = 1 / \lambda$. Because the new technology in this paper is set to be a GPT progress, it is not expected to take place every single year. Rather, it arrives once a long period. The postwar data suggest that the average annual percentage change of Solow Residual is 0.79 percent. After controlling for the variation in capital utilization and non-constant returns to scale, Basu, Fernald and Kimball (2004) report that the “purified” residual increases on average 0.35 percent annually. Suppose a GPT progress takes place once every decade, and all productivity increase is due to the GPT progress, the mean of $\epsilon$ should be between 8.19 percent and 3.56 percent \textsuperscript{19}. If we postulate that only a fraction of the productivity growth is due to the GPT progress, then the mean of $\epsilon$ should be lower correspondingly. As a benchmark, I choose the mean of $\epsilon$ equal to 0.04 and $\lambda = 25$.

Figure 4-1 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 consistent with the

\textsuperscript{19}The cumulative “non-purified” Solow Residual is 0.0819, and the cumulative “purified” residual is 0.0356
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 economy invests in new capital and some of the new technology has been embodied, the economy learns that the technology shock is at least as large as what has already been embodied and revise the conditional belief about the magnitude of the shock. Therefore, the conditional expectation of $\epsilon$ keeps increasing until the investment overshoots. In the period before overshooting, the conditional expectation of $\epsilon$ 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. The conditional expectation goes back to 0.2 thereafter. In this setup, we find that it takes 8 years since the arrival of the new technology before the investment finally overshoots.

To contrast my results with the standard DGE models, I present the impulse response of investment for such a model in Figure 5. In this graph, investment is normalized as a quantity index relative to the level in the first period after the technology shock. We see two things from the graph. First, reacting to technology shocks, the level of investment does not change much over time. Within 20 periods, the change is less than 4%; second, investment level is decreasing over time instead of increasing.

Now look at the investment dynamics in the middle-left panel. Before finding overshooting in period 9, for 8 consecutive periods, the investment has been increasing from about 0.405 to 0.513, the annual growth rate is 3 percent and the cumulative growth is 26.7 percent. At $t = 9$, the economy learns that the investment has overshot. The investment in this period is reduced dramatically to 0.36, which marks a decrease of 30 percent.

The output dynamics is shown in the bottom-left panel. The graph indicates a long lasting output boom and an endogenous recession. Before the magnitude of the shock has not been learned completely, investment keeps increasing. This fuels the output boom. Output climb up from 1.38 to 1.8, with an average annual growth rate about 3.3 percent. After the investment
boom collapses, the output boom freezes. We see a sharp decrease of the output growth, though the level of output did not decrease in absolute term. One aspect that this 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 not growing, in contrast to what we have observed in data, the share of PFI increases from 13.4 percent to 17.1 percent during the 1990s.

The consumption dynamics in the middle-right panel largely replicates the pattern of the output IRF before overshooting. At the moment of overshooting, because investment is sharply reduce while output is largely unchanged, consumption actually increased substantially. This pattern is consistent with the most recent recession, during which investment crashed, but the consumption was not hurt. This is another characteristic that distinct the most recent business cycle from others. Typically, consumption is significantly procyclical. However, between year 2000 and 2002, real GDP grew cumulatively 2.6 percent whereas real consumption grew almost 5.7 percent.

Finally, the bottom-right panel shows the dynamics of the implied (shadow) risk free interest rate, \( r^f \). Recall in the decentralized model, if the consumer can borrow and lend at a risk free interest rate to smooth consumption\(^{20}\), the following familiar Euler equation should hold:

\[
U'(C_t) = E_t[\beta (1 + r^f_t) U'(C_{t+1})]
\]  

(3.27)

which implies that

\[
r^f_t = \frac{U'(C_t)}{\beta E_t U'(C_{t+1})} - 1 = \frac{1}{\beta E_t \left[ C_{t+1}^2 \right]} - 1
\]

(3.28)

where the last equality follows since \( U(C) = -\frac{1}{C} \). The interest rate before overshooting is very high in this model due to the high expected consumption growth. This is because before overshooting, some consumption is sacrificed in order to increase investment. In the next period, if the economy remains undershot, consumption will remain to be low. If the economy overshoots,

\(^{20}\)In this paper, it is not necessary to include risk free asset into the model. Because in a homogeneous agents economy nobody is lending or borrowing, such asset is a redundant variable. However, the consumption Euler equation still prices the risk free asset. That is why I label it the shadow risk free rate.
the consumption will bounce back dramatically. The expected consumption growth is the average of the two scenarios weighted by the probability of overshooting. Since the probability of overshooting is not trivially small, the expected consumption growth is quite high. After overshooting, interest jumps down to the ordinary level, which is close to the steady state level, $\frac{1}{\beta}$.

To summarize, the model generates persistent investment and output booms and endogenous overinvesting and recession. However, the share of investment did not increase as what we observe in the data. Meanwhile the investment growth rate is lower than in the data. One potential reason for this is that the revision of the firm’s conditional expectation of $\epsilon$ is not increasing very fast. Therefore, one round of successful investment does not prompt a round of more aggressive investment. A distribution that induces a larger leap of conditional expectation also induces larger leap of investment. Therefore I modify the distribution of technology shocks to allow for a fatter tail. I assume that $\phi(\epsilon) = \zeta \lambda_1 e^{-\lambda_1 \epsilon} + (1 - \zeta) \lambda_2 e^{-\lambda_2 \epsilon}$ and $\lambda_1 < \lambda_2$. This is a weighted average of two exponential distribution. Unlike the standard exponential distribution that has constant hazard rate, this distribution has a decreasing hazard rate. I choose the parameters $\lambda_1$, $\lambda_2$ and $\zeta$ to make this distribution have the same mean as the previous one.

Figure 6-2 illustrated the IRF after a shock $\epsilon = 0.2$ in this setup.

In contrast to the previous case, the conditional expectation of $\epsilon$ increases much more rapidly. Within six periods, it jumps from 0.04, the unconditional mean, up to near 0.5 before the economy realizes overshooting. This rapidly revised belief about the magnitude of the technology shock induces a similarly rapid increase of investment. It grew from 0.41, a level similar to the previous case at $t = 0$, to 0.72, which reflects a 9.8 percent annual increase on average. During the same period, output grew more slowly at an annual rate of 5.5 percent. Therefore, the investment share increases from about 30 percent to nearly 40 percent over five periods. Because the investment grew faster, the boom lasts a shorter length of time. The economy overshoots after 6 years.
To highlight how decreasing hazard rate and infinite support of $\epsilon$ matter, I next show the investment IRF with respect to a technology shock of the same magnitude but drawn from a uniform distribution, which has a monotonically increasing hazard rate and a finite compact support. Figure 6 shows that for uniformly distributed technology shocks with the same mean, 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 although it has not invested over the optima.

4 Related Literature

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

Zeira (1987), Rob (1991) and Barbarino and Jovanovic (2004) study uncertainty about some type of underlying capacity limits. For example, Rob (1991) proposed a mechanism of learning of market capacity through sequential entry of firms. Barbarino and Jovanovic (2004) study a related learning problem where the market capacity is learned through capacity expansion at firm level. One restrictive assumption in their models is that overshooting happens at industry level. Within an industry, since firms are all infinitesimally small, any firm level expansions on the margin does not alter the probability of overshooting of the industry. Hence, when a firm is making expansion decisions, it does not have to take into account the increase of likelihood of overshooting it contributes.

My model is different from the previous studies also on the aspect that my model is a dynamic general equilibrium model and can be used to study aggregated fluctuations while Zeira (1987), Rob (1991) and Barbarino and Jovanovic (2004) are partial equilibrium models that focus on firm
or industry level dynamics.

As for the origin of the uncertainty, my model share similar spirit to that of Zeira(1987). In his study, Zeira sets forth the idea of productive capital and assumes that the output is not a function of existing capital but a function of the minimum between the existing capital and the productive capital. The level of the productive capital is unknown to the producers. In my model, the $\Psi$–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 into current amount of net new investment. In a certain sense, my model can be viewed as an extension and articulation of Zeira(1987). However, due to the complexities introduced into my model on other aspects, I cannot derive the same elegant analytical results as Zeira(1987) gets.

In the recent studies of business cycles, various authors have proposed mechanism to make DGE business cycle models be able to exhibit impulse response functions that are more consistent with the data. Much of such effort has focused on imperfect information and expectation. Beaudry and Portier(2004) revitalized the idea due to Pigou(1926) and argued that upward biased expectation on future productivity growth will lead to a current expansion. Their model can generate endogenous recession without invoking technology regress as well. In particular, an optimistic expectation for future productivity change stimulates current investment and output growth. When the economy realizes that the expected productivity jump has not happened as expected, a recession begins. Beaudry and Portier(2004) did not model learning even though learning may play an important role in their model. As time goes, the agents should be able to collect additional information about future productivity growth. If the agents are allowed to revise their expectation 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 paper would happen.

Learning plays a central role in Van Nieuwerburgh and Veldkamp(2003). They focus on the asymmetry of business cycles, i.e. why booms usually last a long period while recessions are short.
and sharp. In their model, the production function is contaminated by additive exogenous noise, which makes it hard to tell if high output is due to better technology or lucky output shocks. Different from my model, their learning mechanism is more like a signal-extraction device. When the productivity is expected to be high, production activity is higher and this will 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. Booms and recessions in their model are still exogenous. A boom does not have to induce a recession. Learning in their model affects the duration of booms and recessions only.

5 A Closer Look at the Last Business Cycle

The model of this paper generates a persistent investment boom as well as a surge of the observed TFP. In this section, I document the pattern of the cycle in the 1990s in greater details. I also try to provide examples of other historical episodes that share similar characteristics as the most recent one.

The upper panel of Figure 7 plots the logarithm of the real private fixed investment (PFI) and its three major components. From 1991 to 2000, real GDP increased by 38.27 percent, an average of 3.27 percent annually; whereas real PFI increased by 102.78 percent, on average 7.33 percent annually. When the PFI is decomposed into equipment and software investment (E&S), structure investment and residential investment, we find that over that decade real E&S investment increased by 165.65 percent, an average of 10.26 percent each year! Investment growth was even stronger over certain years of the decade and with respect to certain investment goods. For instance, from 1997 to 2000, information-processing equipment investment increased 25 percent each year! In contrast to the rapid growth of E&S investment, other components of the PFI did not increase as fast. The average growth rate of structure and residential investment was 2.54 percent and 5.17 percent, respectively. After the economic expansion reached its peak, we saw a sharp decrease in investment in the years 2001 and 2002. Real PFI decreased 7 percent over
the two years and real E&S decreased more than ten percent.

The lower panel of Figure 7 plots investment as a share of GDP. From 1991, the share of PFI largely had been increasing parallel to the share of the E&S investment. The share of real PFI had increased from 13.40 percent to 17.10 percent by 2000, while the share of E&S investment increased from 6.91 percent to 9.36 percent. In respect to real PFI, only four other years in history saw higher investment share while for E&S investment, it climbed to new historical highs every year from 1995 to 2000. In contrast, the share of structure and residential investment remained rather stable over the same period. To summarize, in the 1990s, there was a remarkable investment boom fuelled by a surge in E&S investment. This boom was sharply reversed in 2000 and the recession followed soon after that.

Another well documented characteristic of the 1991-2000 decade is the rapid increase in productivity. The productivity growth was particularly strong during the second part of the decade. Oliner and Sichel(2000) report that from 1995 to 1999 output per labor hour increased at two and a half percent annually, which almost doubled the average growth rate in the previous 25 years. They estimate that a significant part, about two thirds, of this increase was due to the contribution of information technology. Furthermore, they identify that about one percent of the annual labor productivity increase was contributed by computer hardware and software.

The parallel surge of investment and productivity is not unprecedented. Many episodes in history share similar characteristics. Jovanovic and Rousseau (forthcoming) documented 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 electricity power share slowed down abruptly in 1929, the very year when the Great Depression began. The electricity expansion never returned to its previous strength, even after the economy had recovered from the Great Depression. This does indicate a certain degree of capital overhang.

---

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

22 During the period of 1978-1981, PFI share in GDP reached to the historical high at 18.53 percent.
Similar anecdotal stories can be found in the studies on canal and railroad expansions in the 19th century. Each boom started with the advent of a revolutionary technology and ended up with overinvestment. Stimulated by the invention of the steamboat, by the year 1860 over 4000 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. 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. This is described by Schumpeter as the construction “ahead of demand”. Fogel’s investigation (1960) found that the ex post return in the railroad industry in the early 1870s turned out to be too low to attract capital. Overall, steamboat, railroad and electricity all have contributed to the productivity increase to a great extent, however, these expansions eventually all turned out to have been overinvested.

Although real GDP grew by 0.75 percent only during 2001, real personal consumption increased by 2.5 percent. The dispersion between GDP and consumption growth was quite persistent in the subsequent year. Real GDP increased by 1.87 percent from 2001 to 2002 whereas real consumption grew 3.08 percent. Consumption growth was particularly strong for durable goods, which increased 6.54 percent during the same time. In contrast to the evident deviation from the GDP growth trend during the recession in the years 2001 and 2002, consumption growth largely followed the pace of GDP growth for most of the years during the 1990s when the economy was in the boom.

6 Conclusion and Future Research

In this paper, I propose a new model of embodied technology. In my model, new technology, once has been embodied into new capital, would increase productivity of capital of all vintages. Subject to the limits of the available new technology, the more new capital has been invested, the higher is the productivity level. Agents in the economy do not know exactly the magnitude of the new technology and therefore have to learn it by investing in new capital. The true magnitude

\footnote{Atack and Passell (1994) provides a detailed treatment on the canal expansion and its economic consequences}
is learned after the investment overshoots 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 will be 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 qualitatively with what we observed in the data of the last business cycle. We argue that the model presents a strong internal propagation mechanism to the standard DGE models. In addition, it generates endogenous recessions without invoking technological regress.

There are a number of dimensions along which the paper can be strengthened and extended. One way is to allow for a more flexible and generalized learning structure. In particular, it is desirable to allow for “Learning from Others”, i.e., a firm can learn from other firms instead of from itself only. As Bolton and Harris (1999) point out, extending 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 unless there are additional incentives to make them invest more swiftly. For instance, the government can subsidize this kind of investment. At the equilibrium, the benefit of procrastination on investing is offset by the loss of the subsidy. Alternatively, incentive for investment can be motivated in a monopolistic environment. Because firm’s profit is a function of its productivity level, a firm that invests fast acquires higher level of productivity relative to those firms investing in a slow rate. Therefore the tradeoff in such a model is between investing fast and acquiring higher productivity sooner versus investing slowly and avoiding the potential costs due to overshooting. My conjecture is that there will be a Nash equilibrium where a fraction of the firms choose to be the leaders and invest fast whereas other firms choose to be the followers.

I discuss in the paper verbally how this model provides a possible explanation for the recent
stock market boom. I also contrast my model with Hobijn and Jovanovic (2001) and discuss why IT technology did not destroy the value of the firms in the traditional industries as it seemed had done in the year 1974. The next step is to extend the model to articulate the dynamics of capital good price as well as firm value.

The embodiment function \( \Psi \) plays an important role in the model. It is also the very novel part of the paper. Therefore, it deserves more careful treatment. The functional form assumed in this paper can be viewed as a reduced form representation. It is important to push forward the analysis and provide a more detailed and structural analysis on how the quantity of new capital facilitates the adoption of new technology.

As always, how much truth a model contains can only be examined by looking at how close the model fits the reality. Empirical tests of the model naturally are on the top of the research agenda.
References


Appendix

A Proofs of Lemmas and Propositions

A.1 Proof of Lemma 1:

The firm’s problem is:

$$V_0(K_0) = \max_{I_0} Y_0 - I_0 + \frac{1}{1+r} E_0 V_1(K_1)$$  \hspace{1cm} (A.1.1)

subject to:

$$K_1 = K_0 + I_0$$  \hspace{1cm} (A.1.2)

The first order condition is

$$\frac{\partial E_0 V_1(K_1)}{\partial K_1} = 1 + r$$  \hspace{1cm} (A.1.3)

Let $$V(K_1)$$ be the value of the firm at $$t = 1$$ if it has not yet overshot and $$W(K_1, A)$$ denote the value of the firm if it has overshot and the magnitude of $$A$$ is learned to be $$A$$. We have

$$E_0 V_1(K_1) = \int_{Z_1}^{\infty} V(K_1) \phi(A) \, dA + \int_{Z_1}^{\infty} W(K_1, A) \phi(A) \, dA$$  \hspace{1cm} (A.1.4)

where

$$V(K_1) = \max_{I_1} K_0^{\alpha-1} \times K_1 - I_1 + \frac{1}{1+r} V_2(K_2)$$  \hspace{1cm} (A.1.5)

and

$$W(K_1, \epsilon) = \max_{I_1} (1 + \epsilon) K_1^\alpha - I_1 + \frac{1}{1+r} V_2(K_2)$$  \hspace{1cm} (A.1.6)

subject to

$$K_2 = K_1 + I_1$$  \hspace{1cm} (A.1.7)

where $$V_2(K_2)$$ is given by the familiar

$$V_2(K_2) = \sum_{t=2}^{\infty} \left( \frac{1}{1+r} \right)^{t-2} [(1 + \epsilon)K_t^\alpha - I_t]$$  \hspace{1cm} (A.1.8)

because $$A$$ is learned at $$t = 1$$, by the envelope theorem, we have

$$\frac{\partial V}{\partial K_1} = K_0^{\alpha-1} + \frac{1}{1+r} \left( \frac{\partial V_2}{\partial K_2} \right) |_{K_2 = K_2^*}$$  \hspace{1cm} (A.1.9)

and

$$\frac{\partial W}{\partial K_1} = \alpha A K_1^{\alpha-1} + \frac{1}{1+r} \left( \frac{\partial V_2}{\partial K_2} \right) |_{K_2 = K_2^*}$$  \hspace{1cm} (A.1.10)
where \( K_2^* \) is the optimal level of \( K_2 \) and in this case it is the steady state capital stock level with respect to the technology level \( A \). But evaluated at \( K_2 = K_2^* \), we have

\[
\frac{1}{1 + r} \frac{\partial V_2}{\partial K_2} = 1
\]

By (A.1.3) and (A.1.4) it follows that

\[
\int_{\Psi(K_1) - 1}^{\infty} \frac{\partial V}{\partial K_1} \phi(A) \, dA + \int_{1}^{Z_1} \frac{\partial W}{\partial K_1} \phi(A) \, dA + \{W[K_1, Z_1] - V(K_1)\} \times \phi(Z_1) = 1 + r
\]

and

\[
\int_{Z_1}^{\infty} \frac{\partial V}{\partial K_1} \phi(A) \, dA + \int_{1}^{Z_1} \frac{\partial W}{\partial K_1} \phi(A) \, dA = 1 + r
\]

where (A.1.12) follows by the Newton-Leibnitz Theorem and (A.1.12') is reached by examining equations (A.1.5) and (A.1.6) and noticing that

\[
W[K_1, \Psi(K_1) - 1] - V(K_1) = 0
\]

Plug (A.1.9)-(A.1.11) into (A.1.12') we have

\[
\int_{Z_1}^{\infty} K_0^{\alpha - 1} \phi(A) \, dA + \int_{1}^{Z_1} \alpha A K_1^{\alpha - 1} \phi(A) \, dA = r
\]

**A.2 Proof of Lemma 2:**

Different from the previous problem, it is not necessary that the firm learns the true magnitude of \( A \) when \( t = 1 \) unless it overshoots. Therefore instead of having the convenient presentation (A.1.5) and (A.1.6), the firm’s problem in this case has to be written as:

\[
V(K_0) = \max_{I_0} Y_0 - I_0 + \frac{1}{1 + r} \left[ \int_{Z_1}^{\infty} V(K_1) \phi(A) \, dA + \int_{1}^{Z_1} W(K_1, A) \phi(A) \, dA \right]
\]

(A.2.1)

where

\[
V(K_1) = \max_{I_1} K_0^{\alpha - 1} K_1 - I_1 + \frac{1}{1 + r} \times \frac{1}{1 - \Phi(Z_1)} \times \left[ \int_{Z_2}^{\infty} V(K_2) \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A) \phi(A) \, dA \right]
\]

(A.2.2)
Theorem we have
\[ W(K_1, A) = \max_{I_t} \sum_{t=1}^{\infty} \left( \frac{1}{1 + r} \right)^{t-1} [AK_t^\alpha - I_t] \]  \hspace{1cm} (A.2.3)

The first order condition for (A.2.1) is similar to that for (A.1.1).
\[
\int_{Z_1}^{\infty} \frac{\partial V}{\partial K_1} \phi(A) \, dA + \int_{Z_1}^{Z_2} \frac{\partial W}{\partial K_1} \phi(A) \, dA + \{W[K_1, Z_1] - V(K_1)\} \times \phi(Z_1) = 1 + r 
\] \hspace{1cm} (A.2.4)

However, the terms \[ \frac{\partial V}{\partial K_1} \] and \[ \frac{\partial W}{\partial K_1} \] have different values. Furthermore, \[ W[K_1, Z_1] - V(K_1) \] is no longer equal to zero. By the envelope theorem and using repeatedly the Newton-Leibnitz theorem we have
\[
\frac{\partial V}{\partial K_1} = K_0^{\alpha-1} + \frac{1}{1 + r} \times \frac{\phi(Z_1)}{[1 - \Phi(Z_1)]^2} \times \left[ \int_{Z_2}^{\infty} V(K_2) \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A) \phi(A) \, dA \right] + \frac{1}{1 + r} \times \frac{1}{1 - \Phi(Z_1)} \times \left[ \int_{Z_2}^{\infty} \frac{\partial V}{\partial K_2} \phi(A) \, dA + \int_{Z_1}^{Z_2} \frac{\partial W}{\partial K_2} \phi(A) \, dA \right] + \frac{1}{1 + r} \times \frac{1}{1 - \Phi(Z_1)} \times \left[ \int_{Z_2}^{\infty} \Phi(A) \, dA + \int_{Z_1}^{Z_2} W[K_1, Z_1] \phi(Z_1) \Phi(K_1) \right] \] \hspace{1cm} (A.2.5)

However, the first order condition for (A.2.2) provides that
\[
\frac{1}{1 - \Phi(Z_1)} \times \left[ \int_{Z_2}^{\infty} \frac{\partial V}{\partial K_2} \phi(A) \, dA + \int_{Z_1}^{Z_2} \frac{\partial W}{\partial K_2} \phi(A) \, dA + [W(K_2, Z_2) - V(K_2)] \phi(Z_2) \Phi(K_2) - W(K_2, Z_1) \phi(Z_1) \Phi(K_1) \right] = 1 + r 
\] \hspace{1cm} (A.2.6)

Plug (A.2.6) into (A.2.5) we have
\[
\frac{\partial V}{\partial K_1} = 1 + K_0^{\alpha-1} + \frac{1}{1 + r} \times \frac{\phi(Z_1)}{1 - \Phi(Z_1)} \times \left[ \int_{Z_2}^{\infty} V(K_2) \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A) \phi(A) \, dA \right] \] \hspace{1cm} (A.2.7)

On the other hand, we still have
\[
\frac{\partial W(K_1, A)}{\partial K_1} = 1 + \alpha A K_1^{\alpha-1} 
\] \hspace{1cm} (A.2.8)

Finally plug (A.2.7) and (A.2.8) into (A.2.4) we reach
\[
\begin{align*}
\frac{1}{1 + r} \times \phi(Z_1) \Phi(K_1) & \times \left[ \int_{Z_2}^{\infty} V(K_2) \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A) \phi(A) \, dA \right] \\
& \times \left[ \int_{Z_2}^{\infty} \frac{W(K_2, A) \phi(A) \, dA}{1 - \Phi(Z_1)} - W(K_2, Z_1) \right] \
& = \left[ K_0^{\alpha-1}[1 - \Phi(Z_1)] + \alpha K_1^{\alpha-1} \int_{Z_1}^{Z_2} A \phi(A) \, dA \right] + \phi(Z_1)\Phi(K_1)[W(K_1, Z_1) - V(K_1)] + \frac{1}{1 + r} \times \phi(Z_1) \Phi(K_1) \times \left[ \int_{Z_2}^{\infty} V(K_2) \phi(A) \, dA + \int_{Z_1}^{Z_2} W(K_2, A) \phi(A) \, dA \right] \times \left[ \int_{Z_2}^{\infty} \frac{W(K_2, A) \phi(A) \, dA}{1 - \Phi(Z_1)} - W(K_2, Z_1) \right] \end{align*}
\]
A.3 Proof of Proposition 1:

The optimality condition of Case I is

\[ r = \int_{1}^{\infty} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA \]

\[ = \int_{Z_1}^{\infty} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA + \int_{1}^{Z_1} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA \]  

(A.3.1)

The optimality condition of Case II is

\[ r = \int_{Z_1}^{\infty} K_{0}^{\alpha-1} \phi(A) \, dA + \int_{1}^{Z_1} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA \]  

(A.3.2)

Hence

\[ \int_{\psi(K_{1}')}^{\infty} \alpha A K_{1}^{\alpha-1} \phi(A) \, d\epsilon + \int_{1}^{\psi(K_{1}')} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA = \]

\[ \int_{\psi(K_{1}')}^{\infty} K_{0}^{\alpha-1} \phi(A) \, dA + \int_{1}^{\psi(K_{1}')} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA \]  

(A.3.3)

where \( K_{1}' \) and \( K_{1}'' \) denote the optimal capital level at \( t = 1 \) for Case I and Case II, respectively.

Keep in mind that

\[ \alpha A K_{1}^{\alpha-1} < K_{0}^{\alpha-1} \]  

(A.3.4)

to make (A.3.3) hold, the lower bound of the integral \[ \int_{\psi(K_{1}')}^{\infty} K_{0}^{\alpha-1} \phi(A) \, dA \] has to be higher than the lower bound of the integral \[ \int_{\psi(K_{1}')}^{\infty} \alpha A K_{1}^{\alpha-1} \phi(A) \, dA \]. Hence \( K_{1}' < K_{1}'' \). It is straightforward to show that \( K_{1}'' < K_{1}''' \) by the same method.

A.4 Proof of Lemma 3:

\( \tilde{K}_{2} \) satisfies:

\[ r = \int_{Z_1}^{\infty} \alpha A \tilde{K}_{2}^{\alpha-1} \frac{\phi(A)}{1 - \Phi(Z_1)} \, d\epsilon \]  

(A.4.1)

\[ r \times \tilde{K}_{2}^{1-\alpha} = \int_{Z_1}^{\infty} \alpha A \frac{\phi(A)}{1 - \Phi(Z_1)} \, dA \]  

(A.4.2)

But for the pre-shock steady state capital \( K_{0} \) it is true that

\[ \alpha K_{0}^{\alpha-1} = r \]  

(A.4.3)

Hence

\[ \tilde{Z}_2 = \Psi(\tilde{K}_2) = \int_{Z_1}^{\infty} A \frac{\phi(A)}{1 - \Phi(Z_1)} \, dA \]  

(A.4.4)

■
B Computation Methods

I use grid searching and iterate value functions (4.21)-(4.22) to compute the value functions. I discretize the \( A \) distribution with \( N \) points and each point has weight equal to \( \frac{1}{N} \). Depends 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 with computing the value function \( W(K, A) \) because it does not involve \( V(K) \). This can be computed with the standard value function iterations technique. I iterate the value functions until 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 \( \epsilon \) and interpolate the value function for all \( N A \)s.

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 \( \pi \) in hand, I plug them into (4.21) and iterate \( V(K) \) until the gap is smaller than 0.0001. Thus we have a value function of the economy with capital level \( K \) and having not overshot.
Figure 1  Illustration of the Production Function with the $\psi$-Type Embodied Technology

Notes: $A_0$ is the pre-shock level of technology, $A = (1+\epsilon)\bar{A}_0$ is the post-shock level of technology, $K^*(A_0)$ and $K^*(A)$ are the steady state levels of capital corresponding to technology level $A_0$ and $A$. Curve $OD$ is the pre-shock production function, Curve $EC$ is the post-shock disembodied production function, and Curve $OBC$ is the $\psi$-Type embodied technology production function. Along segment $OB$, output increases linearly with capital, at point $B$, where capital is equal to the steady state level corresponding to $(1+\epsilon)\bar{A}_0$, the new production function coincides with the disembodied production function $EC$. 

47
Figure 2 An Illustration of Learning and Overshooting
Figure 3  TFP Time Series of Three Models

Expected TFP before overshooting

Curve A: Disembodied Technology

Curve B: Model of This Paper

Curve C: Solow Embodied Technology (with irreversible capital)

$A_t = (1 + \delta)A_0$

$E[A_t]$

TFP

0  t  t'  t''  Time
Figure 4.1 Impulse Response Function after a Permanent Technology Shock

- **Expectation**: The expectation shows a gradual increase, reaching a peak around year 10, and then stabilizing around year 20.
- **Investment Share**: There is an initial decrease followed by a stabilization around year 20.
- **Investment**: Investment shows an initial decline, followed by a sharp increase around year 10, and then stabilizes around year 20.
- **Consumption**: Consumption increases steadily over time, showing a smooth upward trend.
- **Output**: Output increases rapidly in the first few years, reaching a peak around year 10, and then stabilizes around year 20.
- **Interest Rate**: The interest rate decreases sharply in the initial years, followed by a stabilization around year 20.
Figure 4.2 Impulse Response Function after a Permanent Technology Shock (an Alternative Distribution)
Figure 5 Investment Response after a Permanent Technology Shock in a Standard DGE Model
Figure 6  Investment Response After a Permanent Technology Shock
Drawn from Uniform Distribution
Figure 7-A  Log of Real Private Investment and Its Components

Figure 7-B GDP Share of Investment