Title
Essays on Macroeconomics and Cyclical Fluctuations

Permalink
https://escholarship.org/uc/item/58g5k9rv

Author
SHI, WEI

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Macroeconomics and Cyclical Fluctuations

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Wei Shi

2013

<div align="center">

ABSTRACT OF THE DISSERTATION

Essays on Macroeconomics and Cyclical Fluctuations

by

Wei Shi

Doctor of Philosophy in Economics

University of California, Los Angeles, 2013

Professor Lee E. Ohanian, Chair

</div>

This series of essays studies the observed fluctuations in the aggregate economy and the factors behind. I first examine the cyclical behaviors of the aggregate productivity shocks, as measured by the aggregate Solow residual and how it relates to the technology adoption decision done by individual firms. Then I divert my attention to the labor market, and enquire into (i) why workers with different skills show such significant differences as observed in the U.S. in terms of their unemployment rates and wages; and (ii) what the so-called labor wedge might reflect.

The first chapter of my dissertation formalizes Schumpeter's idea that the firm level technological changes are what cause changes in the aggregate Solow residual. The analysis starts with a characterization about how new firms make their technology adoption decision, taking into account both the average productivity of the candidate technology and the risk associated with its adoption. Then through the creative destruction process, these newly adopted technologies gradually prevail in the market, and eventually manifest themselves in the aggregate Solow residual. The quantitative experiments confirm that the Schumpeterian story told in this chapter is able to amplify the traditional aggregate productivity shock, as well as to transform other shocks to the economy into variations in the Solow residual, and thus generating significant business cycle fluctuations. The model also has a few reasonable firm-level implications.

The second chapter develops a framework for the study of the labor market dynamics when workers differ in their production efficiency, which I call skills in the chapter, and when there are search frictions. Compared to the standard business cycle model with frictional labor market, skill heterogeneity in my model creates dispersion in the match

surpluses between the workers and the firms, and thus necessitates a screening process that results in the termination of the unprofitable matches in equilibrium. This endogenous separation mechanism disproportionately influences the employment status of the less-skilled workers and not only exposes them to larger layoff risks, but also inflicts on them greater difficulties in terms of reestablishing their employment relationship with the firms. Quantitatively, the model has cross-sectional implications for the unemployment rates and the wages that are consistent with the observed differences across skill groups in the U.S. labor market.

The last chapter carefully studies the hypothesis that the empirical labor wedge as defined by Robert Shimer may reflect the existence of a household production sector that is largely uncounted by the standard macroeconomic framework. By enriching an otherwise standard real business cycle model with a household production sector, I find that if the hours worked at home and the utility obtained from the home-produced goods are not included in the calculation, the model generates a wedge between the marginal product of labor (MPL) and the household's marginal rate of substitution between consumption and leisure (MRS) that assembles the empirical labor wedge. With reasonable parameter values, the quantitative properties of the model-predicted labor wedge also match those of their empirical counterpart.

The dissertation of Wei Shi is approved.

Hanno Lustig

Mark L. J. Wright

Pierre-Olivier Weill

Lee E. Ohanian, Committee Chair

University of California, Los Angeles

2013

To my father, SHI Benzhi, and to my step-mother, CHEN Huihui, whose support and care are among my most valuable treasures.

To my late mother, QIAN Hong, whom I still miss so much even after these many years.

To my grandmother, QIAN Shuhua, who brought me up and taught me right versus wrong.

To all my friends, who help make the best part of my life.

No words are ever enough to express my gratitude to them.

List of Figures

List of Tables

Vita

2003–2007   B.S. (Mathematics), Peking University, Beijing, China.

2008-2011   Teaching Fellow, Department of Economics, University of California, Los Angeles.

2009   M.A. (Economics), University of California, Los Angeles.

2009   C.Phil. (Economics), University of California, Los Angeles.

2011-2012   Special Reader, Department of Economics, University of California, Los Angeles.

CHAPTER 1

Heterogeneous Firms, Technology Adoption and Aggregate Solow Residual

1.1 Introduction

Over the decades, the real business cycle (RBC) literature mainly works with an aggregate production function, and assigns great importance to the technological shocks as the driving force behind the cyclical behavior observed in the aggregate economy. This approach can be justified in many ways. For instance, in Chari, Kehoe and McGrattan (2007, [2]), the technology wedge and the labor wedge are found to have accounted for almost all the aggregate fluctuations during the Great Depression and the 1982 recession. What they call the technology wedge is the gap between the per capita output and the per capita labor-capital composite formed in the Cobb-Douglas fashion, i.e., the well-known Solow residual. To some extent, such specification of the technology wedge demonstrates the widely held belief in the RBC literature that the Solow residual reflects the technological changes experienced by the economy.

However, ever since the early days of the RBC theory, economists, such as Lawrence Summers (1986, [22]) and Gregory Mankiw (1989, [16]) etc., have expressed doubt about the Solow residual as a measure of the technological shocks and the central role played by these shocks. The great importance attributed to the technological shocks is also challenged by a few recent empirical studies which assert that other factors, like the world interest rate shocks in Neumeyer and Perri (2005, [18]), as well as the preference shocks and country spread shocks in García-Cicco, Pancrazi and Uribe (2010, [11]) etc., may also contribute to the observed cyclical movements of the aggregate output and other aggregate variables of interest. Though we still consider the RBC framework as a reasonable and fruitful abstraction of the real world economy, given above concerns, we

think it is worthwhile for us to move away from the aggregation assumption typically maintained in the RBC models, and to perform a more careful examination about the relationship between the Solow residual and the true technological changes taking place among the economical agents.

Ever since Joseph Schumpeter, it has been understood that the technological progress is made possible by the firm level innovative and technology adopting activities. Moreover, the creative destruction process, i.e., the old firms being gradually replaced by the new ones equipped with better technologies, as well as the efficient risk allocation among the whole economy guaranteed by a well-functioning financial system, are both considered to have facilitated the invention and the use of new technologies. Motivated by this view, instead of designating a representative firm to do all the production, we treat the production sector as consisting of a large number of heterogeneous firms, each one of which makes technology installation choice and then gets the draw of the idiosyncratic productivity accordingly.

Hopenhayn (1992, [13]) provides a nice framework to study the productivity evolution for heterogeneous firms and the resulted long run effects on the aggregate economy. The part of our model concerning the firm-level dynamics rests on his framework, but we make several simplifying assumptions. For instance, we abstract from the post-entry learning or technology upgrading for incumbent firms. We also put aside the endogenous exit decision of the incumbent firms. The cost of such simplification is that we lose a few maybe important dynamic properties at the firm level, but our gain is a nicely shaped aggregate production function with an explicit form for the Solow residual, as well as the ability to move from the stationary equilibrium analysis to the dynamic stochastic general equilibrium analysis.

The technology adoption problem we are going to propose in the paper emphasizes the risk of adopting new technologies, which, as suggested by Schumpeter in his discussion of the role of the financial market, can be thought of as including at least the following two aspects. The first concerns the risk-pooling among individual firms in the technology adoption process. The other more general one is about the economy-wide risk-return tradeoff. The resources invested in the risky technologies chosen by the firms have alternative uses, thus the return of the target technologies has to be high enough

2

to compete for these scarce resources. Put in another way, a less risk-averse economy delivers a more friendly environment for the firms to install the risky technologies by asking for a relatively lower compensation for the risks it is about to bear. In the paper, we try to capture both aspects by assuming that technology adoption is subject to a prepaid set-up cost, which is financed by the firms through entering collectively into a borrowing contract with the household, which guarantees the latter a minimum state-price adjusted expected return. This is a reduced form that enables us to incorporate the fundamental characteristics of a well-functioning financial market in our model without being distracted too much from focusing on the firm-side issue. As a partial equilibrium outcome, the effective cost of adopting new technologies is lower when the profitability per unit of firm-specific productivity is higher and/or the economy is less risk-averse.

The creative destruction process also presents in our model. However, rather than making the replacement a one-shot event where the new comer immediately takes the whole market from the incumbent, as in Aghion and Howitt (1992, [1]), we model it as a more gradual process. The feasible technologies in the model have an exogenously growing frontier. The fact that the new firms always install the most advanced technologies bids up the real wage and effectively increases the operational cost for the old firms which can only stick to their outdated technologies. In the long run, the old firms continue downsizing and are driven out of the market asymptotically.

Our model takes into account the cyclical properties of the firm entry and exit observed in the real world as well. Motivated by the empirical findings in Lee and Mukoyama (2011, [15]), we impose a selection mechanism at the entry that screens out the entrants with idiosyncratic productivity much lower than the average within their own age cohort. It turns out that we are able to replicate the business cycle properties for the entering firms documented in their paper without assuming a special dynamic pattern for the entry cost.

In spirit, our model is connected to a relatively new strand of the macroeconomic studies that try to provide mechanisms transforming other disturbances to the economy into shocks to the Solow residual. To offer an example, in a recent paper, Bai, Ríos-Rull and Storesletten illustrate that in a competitive search environment, demand shocks behave as shocks to the aggregate total factor productivity (TFP). Similarly, we provide

an example in the paper that turns the shocks to the world interest rate into variations in the Solow residual. With calibrated parameter values, a 1% increase in the world interest rate generates a 0.6% decline in the Solow residual, which changes would be absent in the standard RBC model with a representative firm and exogenous productivity shocks. Moreover, in our framework, any disturbances that will affect the endogenous firm level technological changes show up as innovations to the Solow residual. Therefore, we think that our model contributes to this literature by providing a more general transformation mechanism.

The rest of the paper is organized as follows. Section 1.2 details the set-up of the model. We characterize the firms' technology adoption decision and the general equilibrium in section 1.3. Section 1.4 provides quantitative analysis. The concluding remarks and our plan for the future research are given in Section 1.5. We suppress all proofs into the appendix.

1.2 Model with Endogenous Technology Adoption

Time is discrete, starting from a base year named by period 0. There is a single perishable output good in the economy. The good is either consumed or invested to make new capital usable next period by the representative household. It is produced from capital and labor by a continuum of heterogeneous firms. These firms are grouped into two categories: the incumbents and the entrants. We shall go over the optimization problems facing these three types of agents in this section and finish with the definition of the general equilibrium.

1.2.1 The Representative Household

The representative household behaves as in the standard RBC literature. It accumulates capital, supplies the capital and labor services to the firms and spends its income on consumption and investment. In short, the representative household solves the following

utility maximization problem:

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t u(c_t, 1-h_t)\right], \tag{1.1}$$

$$s.t. \quad c_t + i_t + \sum Q_{t,t+1}b_{t+1} \le w_t h_t + r_t k_t + \text{Profit}_t + b_t, \tag{1.2}$$

$$i_t = k_{t+1} - (1-\delta)k_t, \tag{1.3}$$

where c_t is the time t consumption and h_t is the time spent on the market activity. The total time endowment for the representative household in each period is normalized to 1. The instantaneous utility function $u(\cdot, \cdot)$ is continuous, increasing and concave in both arguments. Since our model economy is growing, we also require that $u(\cdot, \cdot)$ be consistent with a balanced growth path. The household is able to trade a set of state-contingent bonds $\{b_{t+1}\}$ which can be conditional on every aggregate state.

The representative household gets income from the labor it supplies to the firms at the wage rate w_t, and the rental income from capital it owns for the rental rate r_t. As the ultimate owner of all firms in the economy, the representative household also collects profits, if any, from the firms.

1.2.2 The Incumbent Firms

A typical incumbent firm produces the output good in period t according to a decreasing-return-to-scale production function

$$Y_t(\tau, z) = (g^\tau z)^{1-\nu} A_t[K_t(\tau, z)^{1-\alpha}L_t(\tau, z)^\alpha]^\nu, \tag{1.4}$$

where $\nu, \alpha < 1$. The term $g^\tau z$ indicates how efficient a firm with individual characteristic (τ, z) is in terms of transforming capital and labor into the output good, where τ is the time when the firm was created, or equivalently, $t-\tau$ is the age of the firm. We assume that the available technology has an exogenously growing trend, so a newer generation of the firms are on average g ($g \ge 1$) times more productive than their immediate predecessors. The variable z is the idiosyncratic component of the firm-specific productivity within each age cohort. To simplify the general equilibrium analysis done later, we assume that once drawn, z remains fixed over time until the firm exits the market. Moreover, the

firm's production is subject to an aggregate technology shock A_t that has no growing trend and equally applies to all the firms in the economy.

In each period t, the incumbent firm (τ, z) will allocate capital and labor in order to maximize its profit

$$\Pi_t(\tau, z) = Y_t(\tau, z) - w_t L_t(\tau, z) - r_t K_t(\tau, z), \qquad (1.5)$$

subject to the production function (1.4), where the wage rate w_t and the rental rate of capital r_t are taken as given. Solving this constrained optimization problem leaves us with the optimal factor demands $L_{dt}(\tau, z)$ and $K_{dt}(\tau, z)$ for this firm, as well as its optimal output $Y_t(\tau, z)$ and the maximized profit $\Pi_t(\tau, z)$.

At the end of the period, an incumbent firm may be hit by an exogenous exit shock with constant probability x and is forced to leave the market. We assume that such exit shocks are independently distributed across firms and over time. As a result, the value of the firm (τ, z) at time t can be defined as the expected present value of its future profits, conditional on survival:

$$V_t(\tau, z) = \mathbb{E}_t \left[\sum_{j=0}^{\infty} (1 - x)^j Q_{t,t+j} \Pi_{t+j}(\tau, z) \right], \qquad (1.6)$$

where

$$Q_{t,t+j} = \beta^j \frac{u'(c_{t+j}, 1 - h_{t+j})}{u'(c_t, 1 - h_t)} \qquad (1.7)$$

is the household stochastic discount factor.

1.2.3 The Entrants

In this paper, technologies are considered as risky projects which could turn out to be much less productive than the ex ante expectation. As a result, the entrants face idiosyncratic risks in the technology installation process. The following formulation of the entrant's problem allows the entrants to share, to some extent, such idiosyncratic risks if they target the same technology.

From the whole economy's perspective, installing a certain technology means that

the resource required in the installation process can not be used for other maybe less profitable but safer purposes. As a result, when making such decision, the economy implicitly weighs the riskiness associated with the candidate technologies against the potential benefits they are able to generate on average.

Let us turn to the modeling details. The time t entrants do not do any production in period t. They come to the market and are shown to a continuum of candidate technologies among which they have to choose one to install if they want to become active next period. The installed technology determines how efficient its users will be in their future production. To be more precise, we suppose that an entrant at time t has productivity $g^t Z$. The term g^t illustrates the trend of the exogenous technological progress, while Z is the comparable part of the z component for an incumbent firm. The only difference is that the incumbent firm knows its z, while the entrant, since it has never participated in production before and thus has no way to observe the realization of its idiosyncratic productivity, views Z as a random variable. A technology specifies a particular distribution for the random variable Z, which we denote by its cumulative distribution function $F(z; \epsilon)$. The following are a few assumptions about F we maintain throughout the paper:

Assumption 1 *The c.d.f $F(z; \epsilon)$ characterizes the distribution of a continuous random variable Z such that:*
[1] Z has a non-negative support;
[2] F can be fully characterized by a scaler parameter ϵ, and thus from now on, a candidate technology is identified with the parameter ϵ and its time of installation t. Without loss of generality, we assume that the mean of Z, $\mathbb{E}[Z]$, is increasing in ϵ.[1]

Assumption [1] reasonably restricts the support of the possible idiosyncratic productivity realizations to be non-negative. We will mention the implication of Assumption [2] soon.

The resources spent in the technology adoption process is modeled as a setup cost $g_Y^t C(\epsilon)$ the entrants have to pay before installing technology ϵ at time t, where g_Y is the output growth rate on the balanced growth path. We will derive an expression for g_Y as

---

[1]As can be seen later, the formulation of the technology adoption problem refers to general distributions. Therefore, though we assume that $\mathbb{E}[Z]$ is increasing in ϵ to simplify description, we do not put more weights on the mean productivity of the target technology than the higher moments.

a function of the technology progress rate g when we characterize the balanced growth of our model.[2] This specification, along with Assumption [2] above, highlights the social risk-return tradeoff, i.e., the tradeoff between adopting a more productive technology and the greater opportunity cost of the required initial investment.

In the paper, we incorporate this tradeoff into the entrants' problem by endowing the entrants with zero net worth and requiring the setup cost to be financed by the household. Since the household has other ways to invest or save for the future, the entrants have to offer a high enough rate of return in order to get the financial support from the household. In this way, risks associated with adopting new technologies appear as costs imposed on the entrants.

To keep things tractable in an environment with aggregate uncertainty, we model the borrowing and lending relationship between the entrants and the representative household by a one-period defaultable debt contract. A typical contract specifies the amount B the entrants borrow from the household in the current period, the borrowing interest rate R^e according to which the entrants will repay next period, and an implicit cutoff of the idiosyncratic productivity $\bar{z}$ below which the entrants have to claim default on their debt in the following period.[3] This contract can be conditional on the aggregate state of the date when it is signed. It also depends on the target technology ϵ because different technologies, by assigning different distributions to Z, deliver different levels of the average productivity and risk.

Given the amount of borrowing B_t, the target technology at time t is constrained by

$$g_Y^t C(\epsilon_t) \leq B_t. \tag{1.8}$$

Once a technology ϵ_t is adopted, the realization of the random Z at the beginning of the next period will be drawn according to the c.d.f $F(z; \epsilon_t)$.

The borrowing rate R_t^e should be such that the representative household is willing to

---

[2]It will be shown that $g_Y = g^{(1-\nu)/(1-\nu+\alpha\nu)}$.

[3]Standard literature on the competitive intermediary also includes a monitoring cost. There, the realization of z is the borrower's private information, and the outsiders, such as the household or the intermediary, have to pay the monitoring cost for such information. Though meaningful and interesting, the informational friction is not the main purpose of this project, so we simply assume zero monitoring cost and do not emphasize it in the paper.

supply funds to the entrants,[4] i.e.,

$$\mathbb{E}_t \left[Q_{t,t+1} \left\{ R_t^e B_t [1 - F(\bar{z}_{t+1}; \epsilon_t)] + \int_0^{\bar{z}_{t+1}} \Pi_{t+1}(t, z) dF(z; \epsilon_t) \right\} \right] \geq B_t, \qquad (1.9)$$

where $\bar{z}_{t+1}$ is determined implicitly by

$$\Pi_{t+1}(t, \bar{z}_{t+1}) = R_t^e B_t. \qquad (1.10)$$

$\bar{z}_{t+1}$ is the cutoff of the realization z so that the entrants with an even lower z will default. In the paper, we assume that such firms exit the market immediately. In this sense, $\bar{z}_{t+1}$ serves a second purpose in our model, i.e., the cutoff specified by the selection mechanism at the entry margin, below which the entrants are screened out. Theoretically, $\bar{z}_{t+1}$ could be exogenously given and fixed. Our way of determining $\bar{z}_{t+1}$ by comparing the profit of this particular entrant and the repayment it should be making is motivated by the consideration that the whole economy's tolerance of the bad firms might not be stable, but rather change with the other aggregate conditions.

The left-hand side of (1.9) is the expected present value of all the debt repayments the representative household is able to collect from the entrants with technology ϵ_t. These repayments could vary across different states of nature at time t. The fact that the household has access to a complete bond market based on the aggregate states guarantees that this state-price adjusted expected present value is enough to characterize the household's budget. If it does not weakly exceed the size of the loan itself (right-hand side of (1.9)), the household would not support the entrants' decision of adopting the target technology ϵ_t.

The value of an entrant is thus

$$V_t^E = \max_{\epsilon_t, R_t^e, B_t} \mathbb{E}_t \left[Q_{t,t+1} \int_{\bar{z}_{t+1}}^{\infty} \{V_{t+1}(t, z) - R_t^e B_t\} dF(z; \epsilon_t) \right], \qquad (1.11)$$

where $V_{t+1}(t, z) - R_t^e g_Y^t C(\epsilon)$ gives the value of the entrant with respect to a specific

---

[4]The informational friction studied by the competitive intermediation literature manifests as a factor $(1 - \mu)$ pre-multiplied by the profit function Π on the left-hand side. Parameter μ is the proportional monitoring cost. Firms with $z < \bar{z}$ can not fully repay the debt and would claim default. They get monitored and thus a fraction μ of their profits are lost. Not very concerned with such informational imperfections, we assume that $\mu = 0$ throughout the paper and do not show it explicitly.

realization z if $z \geq \bar{z}_{t+1}$. If $z < \bar{z}_{t+1}$, all the profits are collected by the representative household as the debt repayment. The entrant is left with zero profits and discontinues its production, and thus has a value of zero. The optimization is subject to constraints (1.8) and (1.9).

It can be seen that, since they are ex ante identical, all time t entrants will optimally install the same technology and enter into the same contract with the representative household. More properties concerning the technology adoption problem will be postponed to the next section. We will now turn to the aggregation of the production sector and eventually to the definition of the general equilibrium.

1.2.4 Aggregation

The endogenous technology adoption done by different generations of the firms gives rise to a non-degenerate distribution over the firm age $t - \tau$ and the idiosyncratic productivity z. Let N_t denote the mass of the firms producing at time t. Also let $G_t(\tau, z)$ be the c.d.f. of the time t joint distribution over the firm age $t - \tau$ and the idiosyncratic productivity z among the producing firms. Then the aggregate output is

$$Y_t = N_t \int_{(\tau, z)} Y_t(\tau, z) dG_t(\tau, z),$$

and the aggregate factor demands are

$$L_t = N_t \int_{(\tau, z)} L_{dt}(\tau, z) dG_t(\tau, z),$$

$$K_t = N_t \int_{(\tau, z)} K_{dt}(\tau, z) dG_t(\tau, z).$$

The Cobb-Douglas production function (1.4) implies that the individual labor and capital demands satisfy:

$$L_{dt}(\tau, z) = \alpha \nu \frac{Y_t(\tau, z)}{w_t},$$

$$K_{dt}(\tau, z) = (1 - \alpha) \nu \frac{Y_t(\tau, z)}{r_t}.$$

After some algebra, we reach the following aggregation result:

$$Y_t = \left\{ A_t N_t^{1-\nu} \left[\int_{(\tau,z)} g^\tau z \, dG_t(\tau, z) \right]^{1-\nu} \right\} [K_t^{1-\alpha} L_t^\alpha]^\nu. \tag{1.12}$$

The theoretical Solow residual (the aggregate TFP) is thus

$$\text{TFP}_t = A_t N_t^{1-\nu} \left[\int_{(\tau,z)} g^\tau z \, dG_t(\tau, z) \right]^{1-\nu}. \tag{1.13}$$

As in the standard RBC model with a representative firm, shocks to the aggregate productivity A_t show up as shocks to the Solow residual. However, in our model, changes in the total number of operating firms and the firm-level distributional changes also appear in the Solow residual. Expression (1.13) allows us to study in more details how other disturbances to the economy affect the Solow residual by influencing the characteristics of individual firms, such as the their number or the age-size distribution. Several quantitative examples are offered in Section 1.4.

1.2.5   The General Equilibrium

We will first describe how the firm age-size distribution evolves. Our model economy features endogenous firm entry, where the number M_t os the time t entrant is determined by the following free-entry condition,

$$g_Y^t C_e = V_t^E. \tag{1.14}$$

Parameter C_e is the fixed entry cost. It is made to grow at rate g_Y to be consistent with the balanced growth experienced by the economy.

With respect to the exit margin, we need to differentiate the newly entered firms from all other incumbent firms. Period t is the first time when the time $t-1$ entrants M_{t-1} actually produce. Among them, a fraction $F(\bar{z}_t; \epsilon_{t-1})$ are screened out by the selection mechanism we impose at the entry margin, the remaining $t-1$ entrants are subject to the exogenous exit shocks, as all other incumbent firms. In a word, the mass of $t+1$

operating firms is

$$N_{t+1} = [N_t - M_{t-1}](1 - x) + M_{t-1}[1 - F(\bar{z}_t; \epsilon_{t-1})](1 - x) + M_t, \qquad (1.15)$$

where the first term refers to the evolution of the incumbent firms except for those created last period. The second term states that the newly established firms exit both because of too bad a draw of z and the exogenous exit shocks, while the third term adds the mass of the current period entrants.

Similarly, the distribution evolves according to the law of motion

$$G_{t+1}(\tau, z) = \begin{cases} \frac{1}{N_{t+1}} \{N_t G_t(\tau, z)(1 - x)\}, \text{ if } \tau < t - 1 \\ \frac{1}{N_{t+1}} \{M_{t-1}[F(z; \epsilon_{t-1}) - F(\bar{z}_t; \epsilon_{t-1})](1 - x)\}, \text{ if } \tau = t - 1, z \geq \bar{z}_t \\ 0, \text{ if } \tau = t - 1, z < \bar{z}_t \\ \frac{1}{N_{t+1}} \{M_t F(z; \epsilon_t)\}, \text{ if } \tau = t \end{cases}$$

where $\tau \leq t$ and $z \geq 0$. This law of motion is derived by first considering the mass of each specified type of firms at the end of the current period and then dividing it by the total mass of producing firms in the following period. The largest value of τ in distribution $G_t(\tau, z)$ with a positive value is $t - 1$, due to the assumption that time t entrants have to postpone their production to period $t + 1$.

Let us turn to the definition of the general equilibrium.

There are four active markets in the economy: the market for output good, the labor market, the rental market for capital service and the bond market between the household and the entrants. The bond market will clear if

$$\sum Q_{t,t+1} b_{t+1} = B_t.$$

The factor market clearing conditions are

$$L_t = h_t, \quad K_t = k_t.$$

Since the entry cost and set-up cost are both covered by the output good, the output

good market clears if

$$c_t + i_t + M_t[g_Y^t C(\epsilon_t) + g_Y^t C_e] = Y_t.$$

In equilibrium, the representative household chooses consumption, investment, hours worked and funds to supply to the entrants in order to maximize the expected lifetime utility. The incumbent firms schedule their production optimally according to the competitive factor prices w_t and r_t. Given the expected profit and value functions Π_{t+1} and V_{t+1}, the entrants' choices over ϵ_t, R_t^e, B_t and $\bar{z}_{t+1}$ solve their technology adoption problem. At last, the four markets clear.

1.3 Characterizing the Equilibrium

This section discusses the qualitative properties of the model. We will begin with a partial equilibrium analysis focusing on the entrants' technology adoption problem, then turn to the characterization of the balanced growth path for the whole economy. Major properties are stated as propositions whose proofs are contained in the attached appendix.

1.3.1 The Entrant's Problem

The profit maximization problem of the incumbent firm (τ, z) produces the following optimal profit function:

$$\Pi_t(\tau, z) = \left[(1-\nu)A_t^{\frac{1}{1-\nu}}\nu^{\frac{\nu}{1-\nu}}\left(\frac{1-\alpha}{r_t}\right)^{\frac{(1-\alpha)\nu}{1-\nu}}\left(\frac{\alpha}{w_t}\right)^{\frac{\alpha\nu}{1-\nu}}\right]g^\tau z. \tag{1.16}$$

To ease exposition, define π_t to be the coefficient appearing inside the brackets in (1.16). As can be seen from the definition of the value function (1.6), the value function $V_t(\tau, z)$ also permits the same decomposition

$$V_t(\tau, z) = v_t g^\tau z$$

with the coefficient v_t satisfying

$$v_t = \pi_t + (1-x)\mathbb{E}_t[Q_{t,t+1}v_{t+1}]. \tag{1.17}$$

Fix technology ϵ_t. The relationship between the borrowing rate R_t^e and the cutoff $\bar{z}_{t+1}$ can be rewritten as

$$\pi_{t+1} g^t \bar{z}_{t+1} = R_t^e B_t. \tag{1.18}$$

Substituting R_t^e using (1.18), the entrant's problem becomes

$$\max_{\epsilon_t, \bar{z}_{t+1}} \quad \mathbb{E}_t \left[Q_{t,t+1} g^t \left\{ v_{t+1} \int_{\bar{z}_{t+1}}^{\infty} z dF(z; \epsilon_t) - \pi_{t+1} \int_{\bar{z}_{t+1}}^{\infty} \bar{z}_{t+1} dF(z; \epsilon_t) \right\} \right],$$

$$s.t. \quad \mathbb{E}_t \left[Q_{t,t+1} \pi_{t+1} g^t \left\{ \int_0^{\infty} \min(z, \bar{z}_{t+1}) dF(z; \epsilon_t) \right\} \right] \geq B_t.$$

Let

$$\Gamma(\bar{z}; \epsilon) = \int_0^{\infty} \min(z, \bar{z}) dF(z; \epsilon).$$

It is clear that the function of the expected repayment to the household $\Gamma(\cdot; \cdot)$ is increasing in $\bar{z}$ for all ϵ. This implies that, if the entrant wants to borrow more (by Assumption [2], it will need a larger loan if it targets a more productive technology), it has to choose a higher cutoff $\bar{z}$, or equivalently to promise a higher borrowing rate R^e in order to satisfy the household's participation constraint.

The next proposition shows a few interesting comparative statics.

Proposition 1 *Suppose there is no aggregate uncertainty, and the entrant has perfect foresight of the household discount factor $Q_{t,t+1}$, the profit function π_{t+1} and the value function v_{t+1}. Then the optimal technology and the corresponding contract are such that:*
(a) The shadow value of an additional unit of the resource to the entrants is the same as that to the household, i.e., the multipliers with respect to constraints (1.8) and (1.9) are the same;
(b) When the household participation constraint (1.9) binds, the optimal borrowing rate is such that

$$\frac{\partial R_t^e}{\partial B_t} > 0, \quad \frac{\partial R_t^e}{\partial \pi_{t+1}} < 0$$

Conclusion (a) makes clear our purpose of requiring the set-up cost to be financed by the household. At the margin, the entrants and the household assign the same value to the resources spent in the technology installation process.

Conclusion (b) states the comparative statics with respect to the effective borrowing

rate R^e. As expected, more borrowing leads to higher R^e, while greater profitability reduces it. The optimal technology choice ϵ_t is beyond the scope of the partial equilibrium analysis performed in this subsection. According to the free-entry condition (1.14), the optimal technology ϵ_t interacts with the mass of entrants M_t to maintain the equality between the option value of the entrants V_t^E and the fixed entry cost $g_Y^t C_e$. We will numerically characterize ϵ_t in the next section.

1.3.2 The Balanced Growth Path

Suppose that the aggregate shock $A_t \equiv \bar{A}$. The following proposition characterizes the balanced growth path (BGP) of our economy.

Proposition 2 *Suppose that the technology progress rate g is not too high[5], and the mass of the entrants is strictly positive in every period, then on the BGP,*
(i) The aggregate output Y_t grows at rate $g_Y = g^{(1-\nu)/(1-\nu+\alpha\nu)}$;
(ii) The representative household's consumption c_t, investment i_t and accumulated capital stock k_{t+1} grow at rate g_Y while the hours worked h_t remains constant;
(iii) For a given incumbent firm (τ, z), its output $Y_t(\tau, z)$, labor demand $L_{dt}(\tau, z)$, the demand for capital services $K_{dt}(\tau, z)$ and the profit $\Pi_t(\tau, z)$ all shrink at a constant rate $g^{\alpha\nu/(1-\nu+\alpha\nu)}$ (or grow at rate $g^{-\alpha\nu/(1-\nu+\alpha\nu)}$);
(iv) For the entrants, the optimal technology ϵ_t, the corresponding borrowing rate R_t^e and the cutoff $\bar{z}_{t+1}$ are all constant;
(v) The wage rate w_t grows at rate g_Y while the rental rate of capital r_t stays unchanged.

The first condition assumed in the proposition is to ensure that the household lifetime utility remains bounded on the BGP. The assumption of the positive mass for the entrants guarantees that the actual technology frontier grows at rate g, which facilitates our characterization of the BGP. In the numerical experiments we do later, we calibrate the model so that the annual exit rate matches its empirical counterpart at the non-stochastic steady state, which guarantees the positive mass of entrants.

---

[5]A more specific characterization is possible if we explicitly parameterize the household instantaneous utility function $u(\cdot, \cdot)$.

Conclusion (i) makes connection between the technology progress rate g and the output balanced growth rate g_Y. (ii) and (v) show that the aggregate quantities and prices in our model economy exhibit the same BGP pattern as in the standard real business cycle model with a representative firm. (iii) and (iv) concern the behaviors of individual firms along the BGP which differentiate our model from the standard literature and thus deserve a few more words, which are given in the next subsection.

1.3.3 The Long Run Evolution of Individual Firms

This subsection studies the long run behavior of individual firms in our model. Most of the analysis is done by contrasting two steady states, one with a moderate 2% annual output growth rate. The other has a 4% annual growth rate and represents a fast growing economy. The calibration treats the first economy as the benchmark. All parameters except for g_Y bear the same values across these two economies. More details about the parameterization and calibration are included in the next section.

The main result concerning the long run evolution of individual firms is shown in the proposition characterizing the balanced growth path of the economy. To keep the tractability, we do not consider post-entry firm-level improvement in the model. Thus, a firm starts with the highest relative efficiency. Then since its idiosyncratic productivity $g^\tau z$ is fixed over time, its relative idiosyncratic productivity compared to the technology frontier keeps going down stairs. Eventually, its market share, no matter measured by output or profits generated, or the inputs used, converges monotonically to zero. Such dynamics may be at odd for the short run evolution of the individual firms observed in the real world. However, they are consistent with the long run tendency, the so-called creative destruction process, first addressed by Schumpeter, then rigorously formulated by Aghion and Howitt (1992, [1]).

Figure (1.1) gives a numerical example of the creative destruction process. The blue solid line refers to the benchmark with expected annual growth rate for the per capita real output equal to 2%. The red dash line is with respect to a higher growth rate, 4% per year. The vertical axis in the upper left panel is the actual quantity of output produced by a typical firm which was created in period 0. We normalize the technology frontier at time 0 to have production efficiency 1. Its demands for capital and labor, as well as its
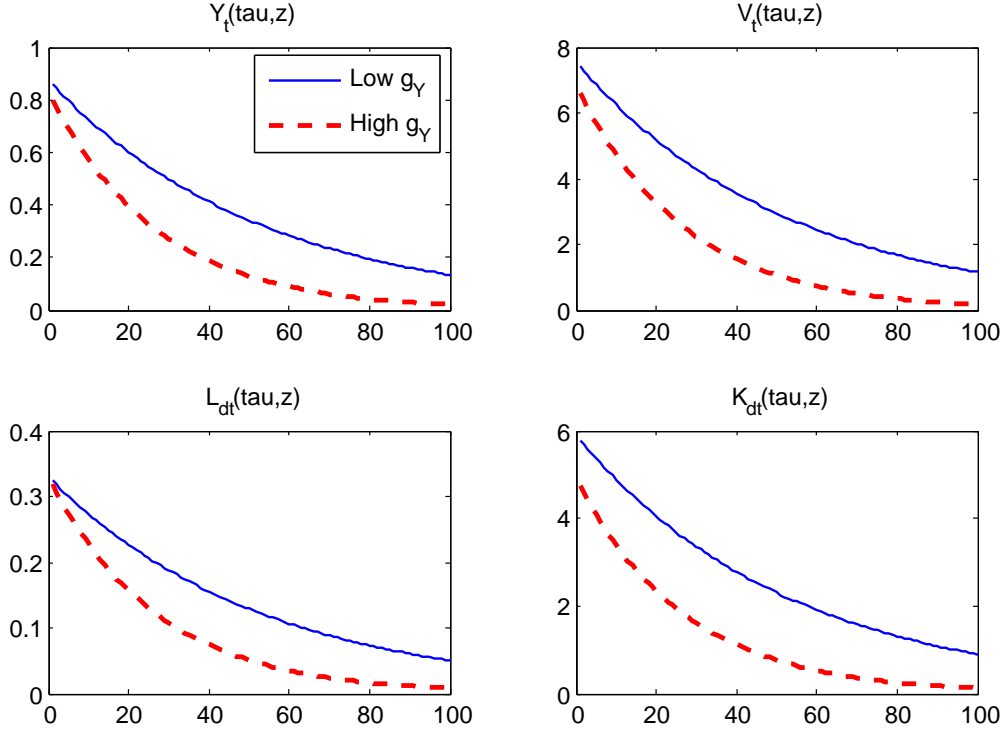
Figure 1.1: Evolution of Individual Firm Characteristics

This figure illustrates the evolution of a typical firm's output, labor demand, capital demand, profit and value. We normalize the technology frontier at period 0 to have efficiency 1. The number on the vertical axis indicates the level of the corresponding variable. Time is placed on the horizontal axis. The blue solid curves refer to a slow growing economy while the green dash curves are for a fast growing economy.

profit function and value function, display qualitatively similar pattern, as can be seen from the other panels of Figure (1.1). Hence, we can conclude that in our model, the creative destruction process means that not only the relative market share, but also the absolute scale of the firms, shrink over time.

The balanced growth trend of the real wage is what underlies such long run dynamics. Since the technology frontier is moving forward over time while the total labor supply is fixed, the labor market clearing condition requires that the real wage appreciates with the technology progress. As a result, for the incumbent firm (τ, z) whose idiosyncratic productivity $g^\tau z$ stays constant after its creation, it has to reduce the scale of its production because it is facing more and more expensive inputs. Eventually, its scale converges to zero, so it as if exits the market when it is becoming far behind the technology frontier.

By comparing the blue and red lines, it is easy to see that the downsize of the old firms is accelerated if people expect faster growth. This result is intuitive because higher
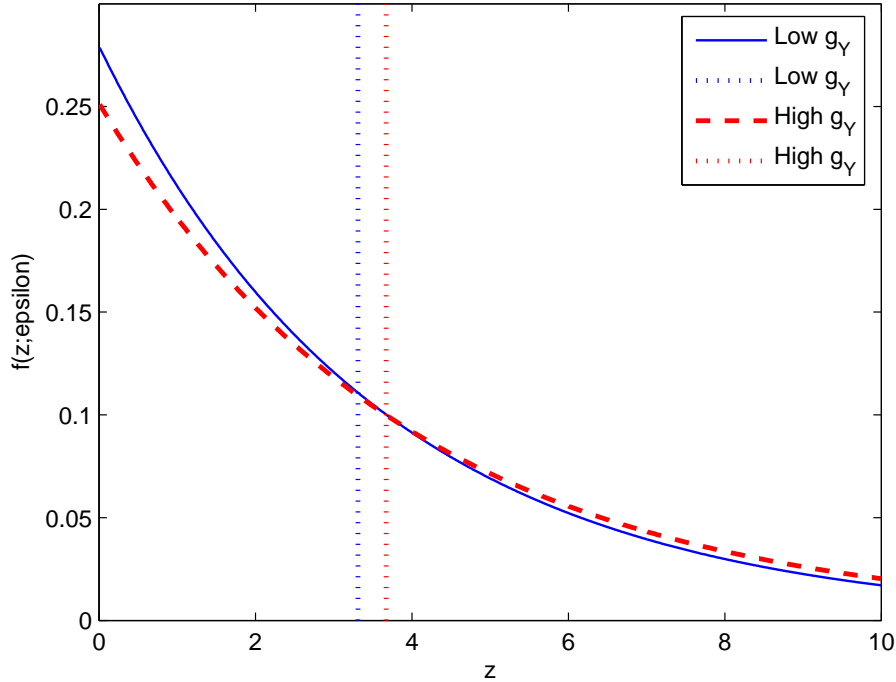
Figure 1.2: Stationary Distribution of the Firm Level Productivity

This graph shows the p.d.f. and the cutoff values associated with idiosyncratic productivity z. The blue solid curve illustrates the stationary distribution of the idiosyncratic z among the entrants given a low expected growth rate g_Y. The green dash curve is the corresponding p.d.f. related to a high g_Y. The vertical straight lines show the steady state cutoff $\bar{z}$ in these two cases, with blue for the smaller g_Y and green for the greater g_Y. The graph is truncated at $z = 10$, while in the numerical exercises performed in the paper, we do not do any truncation.

expected growth rate gives rise to more quickly appreciating wages, and thus makes it more difficult for the old firms to compete with their more productive successors.

With respect to the new entrants, if there are no exogenous shocks present in the economy, each generation of entrants would optimally adopt the same technology $\epsilon_t \equiv \epsilon_{SS}$, and enter into the same contract with the representative household, i.e., $R_t^e \equiv R_{SS}^e$, and are screened out according to the same criterion $\bar{z}_{t+1} \equiv \bar{z}_{SS}$. These implications are consistent with the way we model the general trade-off facing the entrants, i.e., the one between the household's willingness to bear risk and to save, and the average efficiency associated with each technology. Since the representative household's preference stays stable along the balanced growth path, it will favor a constant technology adoption plan.

Figure (1.2) compares the stationary firm-size distribution among the entrants in the two economies with different expected output growth rate g_Y. The faster growing

economy is featured by a more productive technology and a more strict rule to determine which entrants would discontinue their production after the first period.

In the literature, Schumpeter's ideas are usually associated with long term dynamics. In the next section, we will present how the Schumpeterian economy as modeled in our paper would behave at the business cycle frequency.

1.4 Quantitative Studies

This section is devoted to the short run dynamic properties of the model. We first consider a closed economy which is subject to an aggregate productivity shock in the first case, and a shock to the labor wedge in the second case. Then we turn to a small open economy and take shocks to the household non-contingent savings interest rate as the source of uncertainty in the economy. While presenting results in these numerical experiments, we also provide comments on a few puzzles in the RBC or the international RBC literature and show how our model helps us understand them. However, before jumping into details of these experiments, let us first discuss how we parameterize and calibrate the model.

1.4.1 Parameterization and Calibration

The period in our model is defined as one quarter in the data.

In order to obtain an analytical form of the upper-sided expectation, we assume that the random productivity Z follows an exponential distribution:

$$F(z; \epsilon) = 1 - \exp\left(-\frac{z}{\epsilon}\right). \tag{1.19}$$

Then ϵ coincides with the mean of Z.

We parameterize the set-up cost associated with technology ϵ as

$$C(\epsilon) = \kappa \epsilon^{\psi} \tag{1.20}$$

where $\kappa > 0$ is a scaling parameter, $\psi > 1$ shows the convexity of the cost function. We calibrate κ and ψ so that the steady state total mass of the operating firms is 1, so is the

steady state Solow residual.

The return-to-scale parameter ν in the production function is set to 0.9, somewhat higher than the point estimate 0.85 in Atkeson, Khan and Ohanian (1996, [2]). It is easy to see from (1.12) that small value of ν gives more room for the firm level factors, i.e., the mass of producing firms and changes in the firm age-size distribution, to affect the Solow residual. Thus, our baseline choice should be viewed as the lower bound of what our model is able to achieve.

We use the observed labor income share in U.S. to calibrate α, so $\alpha = 0.67$. The technology progress rate g is chosen so that the implied balanced growth rate of the output matches the average growth rate of per capita real GDP in the data. The capital depreciation rate δ is consistent with the average investment-output ratio.

Lee and Mukoyama (2011, [15]) find that in U.S. from 1972 to 1997, the average annual entry rate is 6.2% while the average annual exit rate is 5.5%. Concerning the post-entry evolution of individual firms, Mata, Portugal and Guimarães (1995, [17]) document that for a sample of Portugal plants, more than 20% of new plants died during their first years, and only 30% of the initial population survived for seven years. Motivated by these empirical evidence, we calibrate the exogenous exit rate x and the fixed cost C_e so that in the non-stochastic steady state of our model, the mass of entrants in each period is around 1.5%, and around 60% of plants are screened out by the selection mechanism imposed at the entry margin.[6]

With respect to the representative household, we adopt different functional forms for the utility u, depending on whether the economy is closed or open. In the closed economy version, u is assumed to be

$$u^C(c_t, 1 - h_t) = \frac{1}{1 - \sigma} \left[c_t^a (1 - h_t)^{1-a} \right]^{1-\sigma} \qquad (1.21)$$

The utility is derived from a composite of consumption and leisure with consumption share parameter $a \in [0, 1]$. Parameter σ determines the intertemporal substitutability

---

[6]Literally speaking, such calibration strategy means that 60% of the entrants operate for only one period. However, in reality, the less productive firms are driven out of the market in a much more gradual manner. Therefore, we tend to interpret the selection mechanism in our model as a reduced form and the result it produces summarizes what would happen in the real world over a much longer time horizon.

with respect to the consumption-leisure composite. In the following numerical experiments, we fix it as 2, as commonly assumed in the RBC literature.

We also consider an open economy version of the model where the utility function u is parameterized as

$$u^O(c_t, h_t) = \frac{1}{1-\sigma} \left[c_t - \theta g_Y^t h_t^\varepsilon \right]^{1-\sigma} \tag{1.22}$$

Parameter ε determines the Frisch elasticity of the household's labor supply and is assigned to 1.6. We give a value of 5 to σ. Both are taken from Neumeyer and Perri, (2005, [18]). In both the close economy and the open economy versions, the remaining parameters a and θ are chosen to match the target of the steady state hours worked h.

In all cases, the subjective discount factor β is calibrated so that the household demands a 4% annual interest rate in the steady state absent of any growth.

Table 1.1: Parameter Values (Closed Economy)

Common parameters			
α	0.67		observed labor share
ν	0.9		return to scale
g_w	$1.02^{0.25}$		2% annual growth rate of real output
δ	0.025		10% annual depreciation rate
β	0.99		4% annual interest rate
σ	2		curvature in utility function
ε	1.6		curvature of labor
x	0.6%		exogenous exit probability
	A shock	τ shock	
a	0.2980	0.4637	consumption share
κ	0.0092	0.0091	scale parameter in cost $C(\epsilon)$
ψ	2.353	2.3524	curvature parameter in cost $C(\epsilon)$
C_e	13.2702	12.5599	fixed entry cost
ρ_A & ρ_τ	0.95	0.9645	autocorrelation
σ_A & σ_τ	0.007	0.009	std. of innovations
$\bar{A}$ & $\bar{\tau}$	1	0.4	steady state values

Table 1.1 and Table 1.2 show the calibrated values for all these parameters, where the closed economy version uses U.S. as a benchmark, and Argentina guides our calibration for the open economy version. Table 1.2 also includes a capital adjustment cost and a bond-holding cost that usually appear in the small open economy models.

Table 1.2: Parameter Values (Open Economy)

α	0.67	labor share
ν	0.9	return to scale
g_w	$1.025^{0.25}$	2.5% annual growth rate of real output
δ	0.057	investment-output ratio 0.2
β	0.99	4% annual interest rate
σ	5	curvature with respect to the consumption-leisure composite
θ	1.9820	steady state hours worked 0.2
ε	1.6	curvature of labor
c_b	0.003	bond adjustment cost
$\bar{b}$	-0.4	steady state bond-output ratio
c_k	8	capital adjustment cost
κ	0.0032	scale of the set-up cost
ψ	2.3358	curvature of the set-up cost
x	0.006	exogenous exit rate
C_e	3.0304	entry cost
ρ_R	0.81	persistence of the international interest rate
σ_R	0.0063	standard deviation of the international interest rate

1.4.2 Solow Residual and Aggregate Technology Shock

Our model interprets A_t in the individual firm production function (1.4) as the aggregate technology shock. To make a comparison, we also include in the experiment a standard RBC model with a representative firm operating the following production function:

$$Y_t = g_Y^t A_t \left[K_t^{1-\alpha} L_t^{\alpha} \right]^{\nu}. \tag{1.23}$$

The stochastic process governing A_t is assumed to be a log-AR(1) process:

$$\log A_t = \rho_A \log A_{t-1} + e_t^A \tag{1.24}$$

where the persistence $\rho_A = 0.95$ and the volatility of innovation e_t^A, $\sigma_A = 0.7\%$. Both are values commonly used in the RBC literature.

Table 1.3 displays the business cycle statistics of the main variables implied by our baseline model with heterogeneous firms and those by a standard RBC model.[7] Our

---

[7]The real investment is not shown in the table because it has different meanings in these two models. The representative household invests only in physical capital in the standard RBC model while in our framework, it also invests in making new firms.

Table 1.3: Predicted Moments (Productivity Shock)

	Volatility (%)		$Corr(\cdot, A)$		$Corr(\cdot, Y)$	
	Heter.	Rep.	Heter.	Rep.	Heter.	Rep.
A	2.24	2.24	1.00	1.00	0.96	0.98
Y	4.69	4.36	0.96	0.98	1.00	1.00
c	3.34	3.52	0.82	0.89	0.95	0.97
h	2.25	1.89	1.00	0.99	0.96	0.95
TFP	2.65	2.24	0.99	1.00	0.99	0.98
Y/H	2.62	2.63	0.86	0.91	0.97	0.97
mass of firms	2.83	–	0.36	–	0.52	–
ϵ	1.15	–	0.12	–	-0.16	–
$\bar{z}$	1.62	–	-0.14	–	-0.29	–

model generates similar correlations among the variables as the standard RBC model. However, due to the amplified changes in the Solow residual, the implied volatility of our model is significantly higher. Table 1.4 shows that with the relative conservative value for ν ($\nu = 0.9$), our model generates the Solow residual that is 18% more volatile than the original technology shocks. If $\nu = 0.8$, a value not too small to be unreasonable in the literature, the amplification becomes as big as around one third relative to the underlying A shock.

Table 1.4: Volatility of Implied Solow Residual

std(TFP)	$\nu = 0.9$	$\nu = 0.85$	$\nu = 0.8$
Heter.(%)	2.65	2.81	2.97
Rep.(%)	2.24	2.24	2.24
Change	18%	26%	33%

To some extent, this experiment helps reconcile the real business cycle theory and a few critical opinions about its approach expressed by economists like Lowrence Summers and Gregory Mankiw. These economists point out that there is little evidence illustrating that the technological changes, understood as the shifts in the economy-wide production possibility frontier, can be as dramatic as indicated by the observed Solow residual over such short time horizons as the business cycles. We do not intend to reject other explanations of the Solow residual suggested by these economists, such as the labor hoarding, etc. Our goal is to show that the discrepancy in terms of the duration and the magnitude between the Solow residual and the aggregate technology shock can be partially resolved

by the consideration of the firm level changes.

In our model, individual firms stick to their installed technologies for a relatively long period of time. The frequent fluctuation in the implied Solow residual is due to the constantly entering and exiting activities, as well as the creative destruction process, at the micro level. Moreover, the amplification mentioned above suggests that the underlying exogenous technology shocks can be made less volatile than the observed Solow residual if we take into account the firms' endogenous reactions with respect to their entry decision and their optimal choices over technologies.

1.4.3 Solow Residual and Labor Wedge

Before starting formal analysis, let us first modify the model to incorporate the exogenous shocks to the labor wedge. To exclusively focus on the labor wedge shocks, the term A_t in firm (τ, z)'s production function (1.4) is fixed at its steady state value 1.

Our paper does not aim at providing an explanation for the labor wedge. As a result, in this experiment, we simply adopt its reduced form, i.e., we model the labor wedge as a proportional tax on the labor income received by the representative household $\tau_t w_t h_t$. Then it follows that the representative household equates its marginal rate of substitution between consumption and leisure (MRS) to $(1 - \tau_t)w_t$ in the optimality condition concerning its time allocation. On the other hand, the firms' first order condition states that the equilibrium wage w_t is equal to the marginal product of labor (MPL) for the equivalent representative firm. Therefore, τ_t in our model matches the definition of the labor wedge in the labor market literature.

Again, the contrast is made relative to the RBC case, where the production is undertaken by a single firm according to

$$Y_t = g_Y^t A_t \left[K_t^{1-\alpha} L_t^{\alpha} \right]^{\nu}, \tag{1.25}$$

and $A_t \equiv 1$. The Solow residual, coinciding with the aggregate productivity shock A_t, thus remains constant at its steady state value in the RBC case.

Following Shimer (2010,[9]), we parameterize the labor wedge τ_t as obeying the fol-

lowing law of motion:

$$\log \tau_t = (1 - \rho_\tau) \log \bar{\tau} + \rho_\tau \log \tau_{t-1} + e_t^\tau \qquad (1.26)$$

We repeat Shimer's calculation to obtain the empirical labor wedge and calibrate the unconditional mean $\bar{\tau}$, the persistence ρ_τ and the volatility σ_τ according to our calculated sample statistics. As shown in Table 1.5, shocks to the labor wedge does little to differentiate these two models. It does generate movements in the measured Solow residual in our model, but such movements are too small to have significant impacts on the aggregate economy. We interpret this result as consistent with the finding in Chari, Kehoe and McGrattan (2007, [2]), where they identify both the technology wedge and the labor wedge as responsible for most fluctuations in the real economy. Our model focuses exclusively on the technology wedge. The fact that shocks to the labor wedge under our mechanism contribute to only a tiny part of the technology wedge speaks out the distinctive effects of these two wedges. That helps explain why they are picked separately by Chari, Kehoe and McGrattan's business cycle accounting approach.

Table 1.5: Predicted Moments (Shock to Labor Wedge)

	Volatility (%)		$Corr(\cdot, \tau)$		$Corr(\cdot, Y)$	
	Heter.	Rep.	Heter.	Rep.	Heter.	Rep.
τ	3.41	3.41	1.00	1.00	-0.94	-0.97
Y	2.44	2.38	-0.94	-0.97	1.00	1.00
c	2.05	2.12	-0.87	-0.92	0.99	0.99
h	2.95	3.03	-1.00	-1.00	0.97	0.98
TFP	0.25	0.00	-0.76	0.00	0.93	0.00
Y/H	0.87	0.85	0.74	0.85	-0.48	-0.70
mass of firms	1.52	—	-0.36	—	0.54	—
ϵ	0.56	—	-0.30	—	-0.03	—
$\bar{z}$	0.71	—	-0.06	—	-0.14	—

1.4.4 Solow Residual and Shocks to Interest Rate

The example we are about to analyze concerns the effect of the world interest rate variations on the predicted Solow residual, and moreover, on other aggregate variables of interest in the business cycle studies. There is a large strand of literature analyzing

the large drop in the measured Solow residual in emerging market economies. There are also papers emphasizing the negative correlation between the world interest rates and the emerging market business cycles. In this experiment, we try to connect these two fields of research. As a result of our endogenous technology adoption theory, higher world interest rates tend to reduce the Solow residual, which in turn leads to a temporary recession. Intuitively, changes in the world interest rate alter the domestic household's consumption and savings decision and thus the social opportunity cost of the resources used to create new firms in the domestic market. By suppressing the potential entrants' incentive to start business, as well as to invest in productive projects, the implied Solow residual at the aggregate level falls as a response to a positive innovation to the world interest rate.

Before starting the analysis, it helps to first clarify the environment in which we are going to perform our experiment. The economy is assumed to satisfy the small open economy assumption with fully mobile output good and immobile capital and labor. The representative household trades a non-contingent bond b_t^W at the world interest rate R_t with the rest of the world. From the perspective of this small economy, R_t is exogenously given and its stochastic movements are the only source of uncertainty. This set-up gives us an easy way to shock the state prices the representative household assigns to goods in the future market. Since we are more interested in how the social opportunity costs of the resources required by the technology adoption process impact the entrants' choices, we assume that the local firms do not have access to the international financial market and rely completely on the household for any external funds they need. To stationarize the model, we include the bond-holding cost

$$\Phi_b(b_{t+1}) = \frac{c_b}{2} Y_t \left(\frac{b_{t+1}}{Y_t} - \bar{b} \right)^2 \tag{1.27}$$

as standard in the international literature. We also add a capital adjustment cost to the capital accumulation constraint (1.3) in order to avoid the excess volatility of the real investment in the equilibrium:

$$\Phi_k(k_{t+1}, k_t) = \frac{c_k}{2} k_t \left(\frac{k_{t+1} - g_Y k_t}{k_t} \right)^2 \tag{1.28}$$

The two parameters c_b and c_k are chosen so that the volatility of real output and invest-
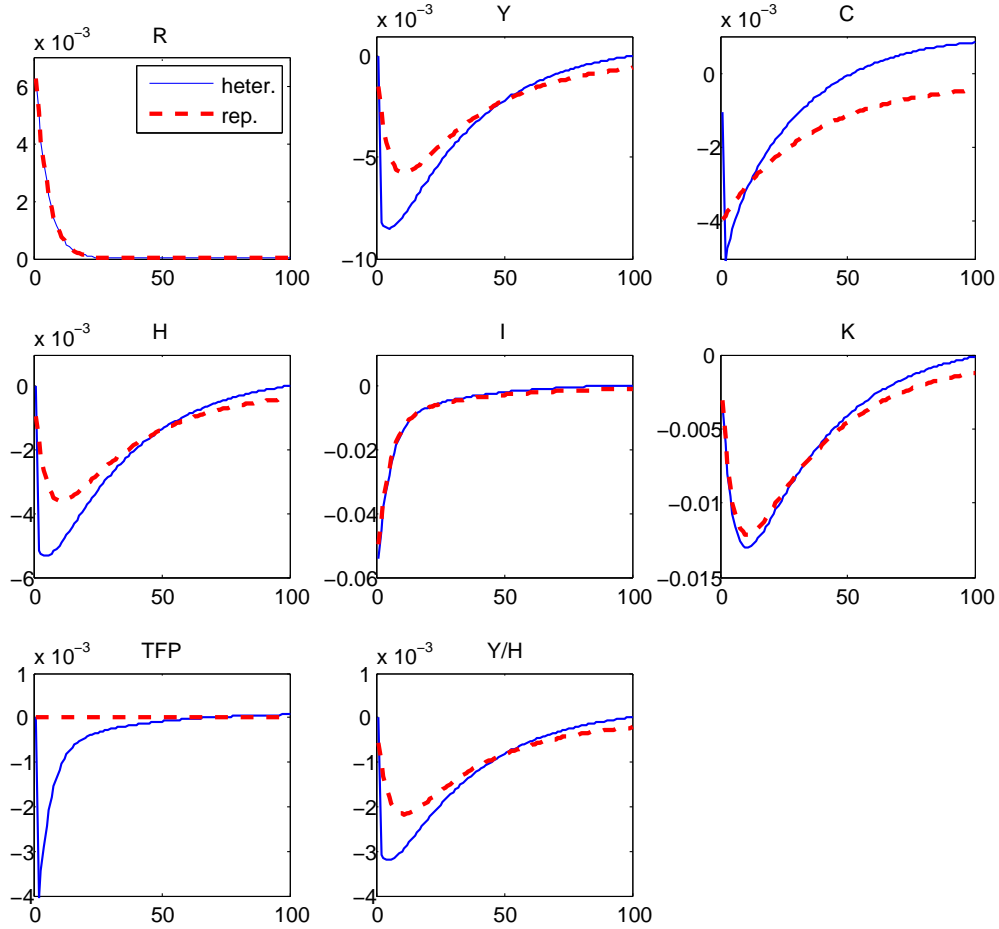
Figure 1.3: Impulse Responses to A One-Standard-Deviation Interest Rate Shock
The impulse responses generated by our model with heterogenous firms and endogenous technology adoption decision are given by the blue solid lines. As a comparison, the responses if the production is done by a representative firm with constant TFP are shown by the red dash lines. All variables are log-deviations from the corresponding steady state, following the exogenous innovation to the world interest rate shock R_t (upper left panel).

ment are in the reasonable range. (See Table 1.2.)

As a comparison, the implication of an international real business cycle model with a representative firm and the following aggregate production function

$$Y_t = g_Y^t A_t [K_t^{1-\alpha} L_t^{\alpha}]^{\nu} \tag{1.29}$$

is also provided in this subsection. We fix A_t at its steady state value, 1, in both cases.

From the impulse responses of the aggregate variables shown in Figure (1.3), it is easy to see that the positive innovation to the international interest rate generates a temporary but relatively more severe recession in our model. Through trading of the non-contingent

bond in the international financial market, the interest rate fluctuations mainly map into variations in the household stochastic discount factor $Q_{t,t+1}$. In the standard set-up with a representative firm and exogenously given production efficiency, $Q_{t,t+1}$ impacts the aggregate production by affecting the Euler equation with respect to new capital stock. This channel also presents in our model, which gives rise to the similar patterns of the aggregate investment in both models. However, in general, the responses produced by our model are more pronounced, and such amplification is due to the declining Solow residual caused by the interest rate shock, which is absent in the standard model. To quantify the result, the implied Solow residual immediately drops by approximately 0.6% after a one-percent increase in the world interest rate.

Table 1.6: Predicted Moments (Interest Rate Shock)

	Volatility (%)		$Corr(\cdot, R)$		$Corr(\cdot, Y)$	
	Heter.	Rep.	Heter.	Rep.	Heter.	Rep.
R	1.07	1.07	1.00	1.00	-0.50	-0.38
Y	4.12	3.11	-0.50	-0.38	1.00	1.00
c	2.06	1.78	-0.52	-0.61	0.79	0.96
h	2.58	1.94	-0.50	-0.38	1.00	1.00
TFP	0.81	0.00	-0.78	0.00	0.79	0.00
Y/H	1.55	1.17	-0.50	-0.38	1.00	1.00
mass of firms	2.39	–	-0.63	–	0.75	–
ϵ	3.30	–	0.95	–	-0.25	–
$\bar{z}$	3.22	–	0.77	–	-0.38	–

Table 1.6 illustrates the same dynamic patterns by the business cycle statistics. As desired, the world interest rate is found to be negatively related with the real output where the correlation is -0.5, comparable to the average correlation between the interest rate and the real output for a set of emerging market economies given in Neumeyer and Perri (2005, [18]). The world interest rate is also negatively associated with the implied Solow residual and there is one period time difference between the changes of these two variables, both of which are qualitatively consistent with the empirical findings.

In our model, the entrants have to compete with investment opportunities abroad, of which the internationally traded non-contingent household bond is a proxy, for the scarce resources necessary to their technology adoption process. To show more explicitly the presence of such tension, Figure (1.4) plots the output volatility against the bond-holding
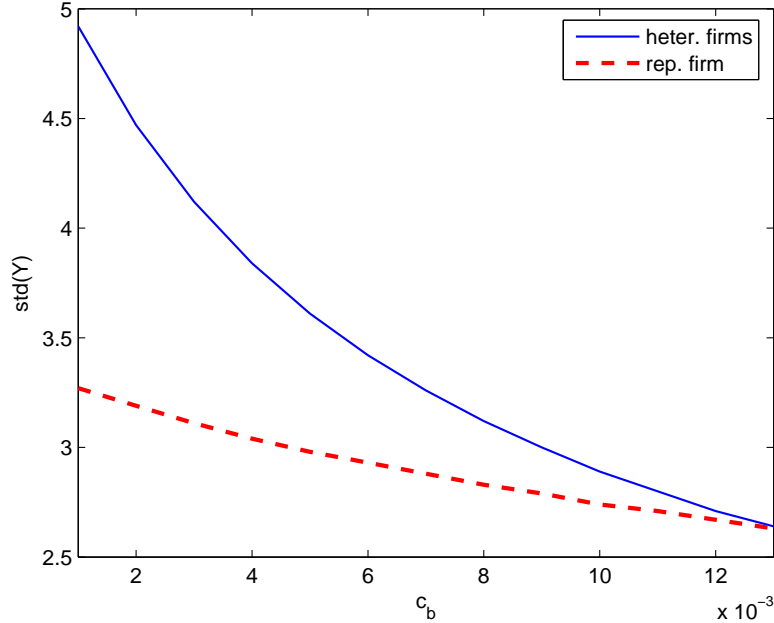
Figure 1.4: Bond-Holding Cost v.s. Output Volatility

This figure shows how the bond-holding cost (given by parameter c_b) affects the standard deviation of the aggregate output Y. In our model, the bond-holding cost determines how effectively the representative household diversifies the domestic risk with the rest of the world, which in turn determines its value of the production-enhancing risky technologies. As a result, when c_b is small, the technology adoption channel transforms the household's willingness in bearing risks into greater volatility of the production sector. The standard deviation of Y is expressed in percentages.

cost. Small bond-holding costs make it easier for the household to shift unspent income in or out of the country. As a result, the new firms' entry and technology adoption decision becomes more sensitive to interest rate movements. In other words, when the firms' productivity is endogenously determined by a mechanism as studied in this paper, temporarily higher world interest rate may have severely adverse impact on the real economy if the household enjoys a lot of flexibility in choosing whether to invest in domestic firms or to invest in the foreign safe assets.

To sum up, this subsection delivers a quantitatively significant example showing how our technology adoption mechanism is able to transform other disturbances to the economy, such as the world interest rate fluctuations discussed above, into shocks to the measured Solow residual, and therefore generates sizable fluctuations at the business cycle frequency. The next subsection will close the discussion by a more detailed description about what is underlying such changing Solow residual.

1.4.5 The Cleansing Effect of Recessions

The cleansing effect sometimes attributed to recessions is a closely related idea to the Schumpeterian creative destruction. It is argued in the literature that, by forcing the firms equipped with outdated technologies to permanently shut up, recessions help to disseminate the newly developed better technologies in the production sector. However, as shown in Caballero and Hammour (1994, [6]), theoretically, it is also possible that these outdated firms are insulated during recessions since too few new firms are created. The empirical findings in a recent paper of Lee and Mukoyama (2011, [15]) are consistent with such insulation effect. The permanently exiting firms in their sample do not show significant differences in their size or productivity over the business cycles, which may be attributed to the observed strong procyclicality of the number of the entrants, thus, insulation does seem to occur in times of the economic downturn.

On the other hand, by looking at the cyclical properties of the entering firms, Lee and Mukoyama find that the average productivity of the entrants is significantly counter-cyclical, which they interpret as indicating that the cleansing effect actually takes place at the entry margin, rather than the traditionally thought exit margin.

In the appendix, we also present the implied micro level moments (Lower panels of Tables 1.3, 1.5 and 1.6) for the previously studied three experiments. It can be seen that for all three cases, the total number of the firms displays strong positive correlation with the aggregate output. Recall expression (1.13), the procyclicality of the total number of the firms is what underlies the procyclicality of the model generated Solow residual. Our model does not have much to say about the exit margin since a large portion of the firms quit randomly and exogenously. However, it generates desirable implications at the entry margin. These panels also show that, the average productivity of the entrants (labeled by ϵ), as well as the cutoff below which the newly entered firms are screened out (represented by $\bar{z}$), move countercyclically in all experiments. This result may be counter-intuitive at the first glance, but with the constant de-trended entry cost C_e and the free-entry condition, the entrants during bad times have to install a more productive technology in order to have the same expected future value as the entrants that enter in good times. In the paper, Lee and Mukoyama obtain such countercyclical average

productivity of the entrants by including a countercyclical entry cost. The installation cost $C(\epsilon)$ exhibits the required dynamic patterns in the general equilibrium of our model, so we do not have to attach special cyclical properties to it in order to replicate these plant-level entering behaviors.

1.5   Conclusion and Future Research Agenda

In this paper, we aim at providing a theory behind the idea of the aggregate production function and the Solow residual based on the Schumpeterian story of the firm level technological changes. Our theory is featured by the heterogeneous firms which, at the entry, endogenously adopt production-enhancing technologies, and over the long run, are subject to the creative destruction process. Our model permits easy aggregation of the production sector, and provides an expression for the Solow residual that makes clear the impact of the micro level factors, such as the number of the firms and changes in the firm age-size distribution. The quantitative experiments performed in Section IV illustrate the significance of our model and its usefulness as a mechanism transforming other exogenous shocks into variations to the Solow residual.

We plan to extend our research in several directions.

One possible improvement for our model is related to the exit margin. Due to the tractability issue, the current version of the model leaves out the firms' endogenous exit decision. It would be an interesting next step to go. We also consider the generalization of incorporating the post-entry technology upgrading or learning problems for incumbent firms. In the data, both channels significantly contribute to the productivity gain of individual firms. We are expecting to see more significant effects if these channels are included.

The role of the financial market is another interesting question to ask in our framework. As the link between the firms that demand external funds and the household that has unspent income to invest, the disfunction of banks or other financial intermediaries will definitely have great influence on the technology adoption process at the firm level, which in turn influences the growth path undertaken the whole economy. We would like to have a better understanding towards this direction in the future.

31

The last thing on our recent research agenda concerns the empirical strategy we would use to test our model's explanatory power. In the paper, we treat each modeling period as one quarter in the data and look at cyclical properties at the business cycle frequency. However, economists, such as Comin and Gertler (2006,[10]), have suggested that cycles induced by the technological changes are usually of longer duration. They propose a way of de-trending the data series that preserves the statistical properties of what they call the medium-term cycles. We plan to follow their empirical approach and evaluate the performance of our model on a more solid empirical foundation.

1.6 Appendix

1.6.1 Appendix A: Proof of Proposition 1

For ease of exposition, define

$$g_{1t+1}(\bar{z}_{t+1}, \epsilon_t) = \int_{\bar{z}_{t+1}}^{\infty} z dF(z; \epsilon_t)$$

$$g_{2t+1}(\bar{z}_{t+1}, \epsilon_t) = 1 - F(\bar{z}_{t+1}; \epsilon_t)$$

Attach to constraints (1.8) and (1.9) multipliers ξ_t and γ_t, respectively. The first order conditions with respect to the three choices variables ϵ_t, R_t^e and d_{t+1} are

$$\epsilon_t: \ \mathbb{E}_t \left[Q_{t,t+1} \left\{ v_{t+1} g^t \frac{\partial g_{1t+1}}{\partial \epsilon_t} - R_t^e B_t \frac{\partial g_{2t+1}}{\partial \epsilon_t} \right\} \right]$$
$$+ \gamma_t \mathbb{E}_t \left[Q_{t,t+1} \left\{ R_t^e B_t \frac{\partial g_{2t+1}}{\partial \epsilon_t} + \pi_{t+1} g^t \frac{\partial g_{1t+1}}{\epsilon_t} \right\} \right] - \xi_t g_Y^t C'(\epsilon_t) = 0, \qquad (1.30)$$

$$R_t^e: \ \mathbb{E}_t \left[Q_{t,t+1} \left\{ v_{t+1} g^t \frac{\partial g_{1t+1}}{\partial \bar{z}_{t+1}} - B_t g_{2t+1} - R_t^e B_t \frac{\partial g_{2t+1}}{\partial \bar{z}_{t+1}} \right\} \frac{\partial \bar{z}_{t+1}}{\partial R_t^e} \right]$$
$$+ \gamma_t \mathbb{E}_t \left[Q_{t,t+1} \left\{ B_t g_{2t+1} + R_t^e B_t \frac{\partial g_{2t+1}}{\partial R_t^e} + \pi_{t+1} g^t \frac{\partial g_{1t+1}}{\partial \bar{z}_{t+1}} \right\} \frac{\partial \bar{z}_{t+1}}{\partial R_t^e} \right] = 0, \qquad (1.31)$$

$$B_t: \ \mathbb{E}_t \left[Q_{t,t+1} \left\{ v_{t+1} g^t \frac{\partial g_{1t+1}}{\partial \bar{z}_{t+1}} - R_t^e g_{2t+1} - R_t^e B_t \frac{\partial g_{2t+1}}{\partial \bar{z}_{t+1}} \right\} \frac{\partial \bar{z}_{t+1}}{\partial B_t} \right] + \xi_t$$
$$+ \gamma_t \mathbb{E}_t \left[Q_{t,t+1} \left\{ R_t^e g_{2t+1} + R_t^e B_t \frac{\partial g_{2t+1}}{\partial \bar{z}_{t+1}} + \pi_{t+1} g^t \frac{\partial g_{1t+1}}{\partial \bar{z}_{t+1}} \right\} \frac{\partial \bar{z}_{t+1}}{\partial B_t} \right] = \gamma_t. \qquad (1.32)$$

Expression (1.18) with respect to the cutoff $\bar{z}_{t+1}$ implies that

$$\frac{\partial \bar{z}_{t+1}}{\partial R_t^e} = \frac{R_t^e}{B_t} \frac{\partial \bar{z}_{t+1}}{\partial B_t}$$

Substituting above equation into (1.31) and subtracting (1.31) from (1.32) yield

$$\xi_t = \gamma_t.$$

Thus, (a) is established.

For any given technology ϵ, the binding household participation constraint implies that the optimal cutoff satisfies the following equation:

$$Q_{t,t+1}\pi_{t+1}g^t \frac{\Gamma(\bar{z}_{t+1}(\epsilon),\epsilon)}{B_t} = 1 \tag{1.33}$$

The first derivative of Γ is

$$\frac{\partial}{\partial z}\Gamma(z,\epsilon)) = [1 - F(z;\epsilon)]$$

which is clearly non-negative.

If

$$\lim_{z \to \infty} Q_{t+1}\frac{\pi_{t+1}g^t}{B_t}\Gamma(\epsilon,z) < 1$$

the intermediary will not offer any loans greater than or equal to B_t to the firms targeting technology ϵ because ex post, the opportunity cost of such a loan is not fully covered by its gain. Otherwise, the borrowing rate R_t^e is defined by the cutoff $\bar{z}_{t+1}$ that makes above inequality hold as an equality.

Given that function Γ is increasing in $\bar{z}_{t+1}$, it is trivial to see that the optimal cutoff $\bar{z}_{t+1}$ is increasing in B_t and decreasing in $\pi_{t+1}g^t$. Next, I will show that the comparative statics of the borrowing rate R_t^e follows the same pattern.

Take the total derivative of the equality (1.33) with respect to B_t and $\bar{z}_{t+1}$,

$$-\frac{Q\pi_{t+1}g^t\Gamma(\epsilon,\bar{z}_{t+1})}{B_t^2}dB_t + \frac{Q\pi_{t+1}g^t}{B_t}\Gamma'(\epsilon,\bar{z}_{t+1})d\bar{z}_{t+1} = 0$$

where Γ' denotes the partial derivative of function Γ with respect to $\bar{z}_{t+1}$. Then,

$$\frac{d\bar{z}_{t+1}}{dB_t} = \frac{\Gamma(\epsilon, \bar{z}_{t+1})}{B_t \Gamma'(\epsilon, \bar{z}_{t+1})}$$

Expression (1.18) implies that $R_t^e = \pi_{t+1} g^t \bar{z}_{t+1}/B_t$, so

$$\frac{\partial R_t^e}{\partial B_t} = -\frac{\pi_{t+1} g^t \bar{z}_{t+1}}{B_t^2} + \frac{\pi_{t+1}^t}{B_t} \frac{d\bar{z}_{t+1}}{dB_t}$$

$$= \frac{\pi_{t+1} g^t}{B_t^2} \left[\frac{\Gamma(\epsilon, \bar{z}_{t+1})}{\Gamma'(\epsilon, \bar{z}_{t+1})} - \bar{z}_{t+1} \right]$$

$$> 0,$$

since

$$\Gamma(\epsilon, \bar{z}_{t+1}) - \bar{z}_{t+1} \Gamma'(\epsilon, \bar{z}_{t+1}) = \int_0^{\bar{z}_{t+1}} z \, dF(z; \epsilon) > 0$$

Following a similar logic, it can be shown that the optimal borrowing rate R_t^e is decreasing in π_{t+1}, so (b) is proved.

1.6.2   Appendix B: Proof of Proposition 2

(i) The balanced growth rate g_Y:

The profit maximization problem for a typical firm (τ, z) implies that the optimal output supplied is

$$Y_t(\tau, z) = \left[A_t^{\frac{1}{1-\nu}} \nu^{\frac{\nu}{1-\nu}} \left(\frac{1-\alpha}{r_t} \right)^{\frac{(1-\alpha)\nu}{1-\nu}} \left(\frac{\alpha}{w_t} \right)^{\frac{\alpha\nu}{1-\nu}} \right] g^\tau z. \qquad (1.34)$$

Again, to simplify notation, define y_t to be the coefficient inside the brackets. Similarly, its optimal labor demand is

$$L_{dt}(\tau, z) = \left[A_t^{\frac{1}{1-\nu}} \nu^{\frac{\nu}{1-\nu}} \left(\frac{\alpha\nu}{w_t} \right) \left(\frac{1-\alpha}{r_t} \right)^{\frac{(1-\alpha)\nu}{1-\nu}} \left(\frac{\alpha}{w_t} \right)^{\frac{\alpha\nu}{1-\nu}} \right] g^\tau z. \qquad (1.35)$$

Given that along the balanced growth path, both total hours worked h_t and the rental rate of capital r_t stay constant[8], condition (1.35) suggests the wage rate w_t should grow

---

[8]The stationarity of r_t can be seen from the household's Euler equation evaluated along the balanced

34

at rate

$$g_w = g^{\frac{1-\nu}{1-\nu+\alpha\nu}}$$

in order to offset the impact of the growing efficiency of the technology frontier on the labor demand. g_w is enough to guarantee a constant aggregate labor demand because the aggregation over all operating firms results in the summation $\int_{\tau<t,z} g^\tau dG_t(\tau,z)$ which is of the same order as g^t. Then, condition (1.34) implies that the growth rate of the output produced by firms on the technology frontier is

$$g \cdot g_w^{-\frac{\alpha\nu}{1-\nu}} = g^{\frac{1-\nu}{1-\nu+\alpha\nu}}.$$

Since the output produced by other firms is of a lower order, the growth of the aggregate output is dominated by the output growth for firms that are on the frontier, thus

$$g_Y = g^{\frac{1-\nu}{1-\nu+\alpha\nu}} = g_w$$

(ii) The budget constraint of the representative household suggests that on the balanced growth path, consumption c_t, investment i_t and the capital stock k_t should have the same growth rate as the wage rate w_t, which is g_Y as computed in the previous part of Appendix B.

(iii) To discuss the individual firms' behavior on the balanced growth path, it is easy to start with the coefficients y_t and π_t as defined in (1.34) and (1.16), which, once multiplied by the firm-specific $g^\tau z$, yield the firm-specific output $Y_t(\tau,z)$ and profit $\Pi_t(\tau,z)$ in period t, respectively.

Expression (1.34) implies that y_t is growing at rate

$$g_y = g_w^{-\frac{\alpha\nu}{1-\nu}} = g^{-\frac{\alpha\nu}{1-\nu+\alpha\nu}}$$

The negative power suggests that y_t is actually shrinking over time because the technology progress bids up the equilibrium wage w_t. For a given firm (τ,z),

$$Y_t(\tau,z) = y_t g^\tau z$$

_____

growth path.

where $g^\tau z$ stays constant, so its output $Y_t(\tau, z)$ is decreasing at rate g_y^{-1}. This argument makes no use of z, thus it applies to all firms that were created in period τ.

The Cobb-Douglas structure of the production function suggests that

$$L_{dt}(\tau, z) = \frac{\nu\alpha}{w_t}Y_t(\tau, z), \quad \text{and } K_{dt}(\tau, z) = \frac{\nu(1-\alpha)}{r_t}Y_t(\tau, z)$$

Hence, the individual demand for capital $K_{dt}(\tau, z)$ decreases at the same rate as the individual output $Y_t(\tau, z)$, while the individual labor demand decreases at rate g_w/g_y.

It can be shown in the same way using expression (1.16) that the individual profit $\Pi(\tau, z)$ is growing at rate g_y.

(iv) When constraint (1.8) is binding, it can be expected that on the balanced growth path, the borrowing of a new firm B_t grows at rate g_Y. Moreover, due to (1.16), coefficient π_t grows at rate $g_y^{-\alpha\nu/(1-\nu)}$. So does v_t. Perform the following transformation of variables:

$$\hat{B}_t = B_t g_y^{-t}, \quad \hat{\pi}_t = \pi_t g_y^{\frac{\alpha\nu}{1-\nu}t}, \quad \hat{v}_t = v_t g_y^{\frac{\alpha\nu}{1-\nu}t}$$

Notice that

$$g \cdot g_y^{-\frac{\alpha\nu}{1-\nu}} = g$$

The technology adoption problem can be rephrased as

$$\max \ \mathbb{E}_t \left[Q_{t,t+1} \int_{\bar{z}_{t+1}}^{\infty} \left\{ \frac{\hat{v}_{t+1}}{g_y}z - R_t^e\hat{B}_t \right\} dF(z; \epsilon_t) \right]$$

$$s.t. \ \hat{C}(\epsilon_t) \leq \hat{B}_t$$

$$\bar{z}_{t+1} = \frac{R_t^e\hat{B}_t}{\hat{\pi}_{t+1}g_y}$$

$$\mathbb{E}_t \left[Q_{t,t+1} \left\{ R_t^e\hat{B}_t \left(1 - F(\bar{z}_{t+1}; \epsilon_t)\right) + (1-\mu)\frac{\hat{\pi}_{t+1}}{g_y}\int_0^{\bar{z}_{t+1}} z dF(z; \epsilon_t) \right\} \right] \geq \hat{B}_t$$

Since all the hat variables are stationary, the aforementioned rephrase shows that the optimal project ϵ_t and the corresponding borrowing interest rate R_t^e are stationary. Without exogenous disturbance to the aggregate economy, $\hat{v}_{t+1}$ and $\hat{\pi}_{t+1}$ are constant, which results in a constant project ϵ_t and a constant borrowing rate R_t^e on the balanced growth path.

36

(v) The balanced grow property for the wage rate w_t has already been shown in (i). The constancy of the rental rate for capital r_t results from the household Euler equation.

References

[1] Philippe Aghion, and Peter Howitt, "A Model of Growth through Creative Destruction", *Econometrica*, Vol. 60, No. 2 (Mar., 1992), pp. 323-351.

[2] Andrew Atkeson, Aubhik Khan, and Lee Ohanian, "Are data on industry evolution and gross job turnover relevant for macroeconomics?", Carnegie-Rochester Conference Seires on Public Policy, 44 (1996) 216-250.

[3] Yan Bai, José-Víctor Ríos-Rull, and Kjetil Storesletten, "Demand Shocks that Look Like Productivity Shocks", working paper, April 2011.

[4] Torbjörn Becker, and Paolo Mauro, "Output Drops and the Shocks That Matter", IMF working paper, July 2006.

[5] Ben S. Bernanke, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative real business cycle framework", *Handbook of Macroeconomics*, Volumn 1, Edited by J. B Taylor and M. Woodford, 1999.

[6] Ricardo J. Caballero, and Mohamad L. Hammour, "The Cleansing Effect of Recessions", *The American Economic Review*, Vol. 84, No. 5(Dec., 1994), pp. 1350-1368.

[7] Ricardo J. Caballero, and Adam B. Jaffe, "How High Are the Giant's Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth", *NBER Macroeconomics Annual*, Vol. 8 (1993), pp. 15-74.

[8] Murillo Campello, John R. Graham, and Campbell R. Harvey, "The real effects of financial constraints: Evidence from a financial crisis", *Journal of Financial Economics*, 97 (2010) 470-487.

[9] V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan, "Business cycle accounting", *Econometrica*, Vol. 75, No. 3(May, 2007), 781-836.

[10] Diego Comin, and Mark Gertler, "Medium-Term Business Cycles", *The American Economic Review*, June 2006, 523-551.

[11] Javier García-Cicco, Roberto Pancrazi, and Martín Uribe, "Real Business Cycles in Emerging Countries?", *American Economic Review*, 100 (December 2010): 2510-2531.

[12] Jeremy Greenwood, and Gokce Uysal, "New Goods and the Transition to a New Economy", *Journal of Economic Growth*, 10, 99-134, 2005.

[13] Hugo A. Hopenhayn, "Entry, exit, and firm dynamics in long run equilibrium", *Econometrica*, Vol. 60, No. 5(Sep., 1992), pp. 1127-1150.

[14] Timothy J. Kehoe, and Kim J. Ruhl, "Are Shocks to the Terms of Trade Shocks to Productivity?", Federal Reserve Bank of Minneapolis, Research Department Staff Report 391, April 2008

[15] Yoonsoo Lee, and Toshihiko Mukoyama, "Entry, Exit, and Plant-level Dynamics over The Business Cycle", working paper, July 2011.

[16] N. Gregory Mankiw, "Real Business Cycles: A New Keyesian Perspective", *Journal of Economic Perspectives*, Volumn 3, Number 3, Summer 1989, Pages 79-90.

[17] José Mata, Pedro Portugal, and Paulo Guimarães, "The survival of new plants: Start-up conditions and post-entry evolution", *International Journal of Industrial Organization*, 13 (1995) 459-481.

[18] Pablo A. Neumeyer, and Fabrizio Perri, "Business cycles in emerging economies: the role of interest rates", *Journal of Monetary Economics*, 52 (2005) 345-480.

[19] Diego Restuccia, and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments", *Review of Economic Dynamics*, 11 (2008) 707-720.

[20] Kjell G. Salvanes, and Ragnar Tveteras, "Plant exit, vintage capital and the business cycle", *The Journal of Industrial Economics*, Vol. 52, No. 2 (June 2004), pp. 255-276.

[21] Robert Shimer, *Labor Markets and Business Cycles*, Princeton University Press, 2010.

[22] Lawrence H. Summers, "Some skeptical observations on real business cycle theory", *Federal Reserve Bank of Minneapolis Quarterly Review*, 1986 Fall, pp. 23-7.

CHAPTER 2

Skill Heterogeneity, Search Frictions and Labor Market Dynamics

2.1 Introduction

The issue of unemployment has been an important topic in economic research for a long time, and it is relevant both for understanding the basic principles of the labor market and for analyzing the wellbeing of individual workers. In recent years, the search and matching models as formulated in Mortensen and Pissarides (1994, [13]) or in Merz (1995, [11]) emerge as the main device in macroeconomics to study the unemployment dynamics and have been proven quite fruitful. However, due to its consideration of only the homogeneous workers, this type of models is totally silent about the differences in the unemployment spells and frequencies for workers with heterogeneous individual characteristics in the data, which, as discussed by several empirical studies, are widely observed and of nontrivial significance.

To list some of these empirical studies, Ravenna and Walsh (2012, [17]) find that in both the U.S. and the Europe, young workers who can be thought of as having both less education and less working experience face on average higher and more volatile unemployment risks over the business cycles. Focusing on the recent Great Recession, Elsby et al. (2010, [6]) also document that the demographic groups which can be classified as relatively low-skilled experience constantly higher inflow into unemployment than the rest of the population, and their unemployment rate exhibits steeper rise at the onset of the recession. Moreover, it is a long-established fact that there exists positive skill premium, meaning that the more skilled workers are also constantly paid more than the relatively low-skilled ones. In our opinion, all these rich heterogeneities resulted from skill differences shed light on the functioning of the labor market, to be able to capture which

is the main motivation of our extending the standard search and matching framework to incorporate skill heterogeneity.

There are also theoretical concerns justifying the inclusion of the skill heterogeneity in the search and matching models. One gain from allowing skill heterogeneity among the potential employees is that it permits the separate consideration of the endogenous job destruction, as oppose to the exogenous job destruction that is usually assumed in the literature to ensure the existence of the stationary equilibrium. This idea is closed related to the screening and the employer search emphasized in a few papers, such as Ravenna and Walsh (2012, [17]), Villena-Roldán (2010,[20]), etc., and is supported by empirical evidence reported in Barron, et al. (1985, [3]) and van Ours and Ridder (1992, [21]). The basic logic goes as follows. The firms have limited information in targeting their vacancy-posting activities to specific skill requirements. As a result, their recruiting process first finds a pool of interested job-seekers and then only the more productive ones will be actually hired. Moreover, the firms and their existing employees may also constantly re-evaluate the already formed employment relations, and both would agree to dissolve the ones where too few surpluses are generated to be shared with each other. The separation between the firms and those job-seekers who are contacted by the firms but do not end up with an job offer, or between the firms and those workers who were employed previously but turn out to be no more productive given the current aggregate state, is what we call the endogenous job destruction in this paper.

Our model is mainly built on Merz (1995, [11]). There are homogeneous firms producing and posting costly vacancies in the labor market in order to attract potential employees. Infinitely-living identical households consist of workers who differ in their skills but have the same preference over consumption and leisure. For tractability, we simplify the household's decision making problem by assuming the existence of an intra-household perfect financial market, thanks to which all workers within the same household enjoy the same level of consumption, no matter what income they make. We also abstract from the labor market participation concern for the households by assuming that job searching imposes no cost, both in terms of resource and in terms of utility, on the households. As a result, all unemployed workers are job-seekers in our model. As in Merz (1995, [11]), matches between the vacancies and the unemployed workers are formed according to a

random matching technology, and equilibrium wage is determined as the Nash bargaining outcome where the Hosios efficiency condition holds. However, on top of the wage determination, our paper adds another stage during which the firms and the workers together assess the profitability of the existing matches, no matter whether these matches were newly formed or formed long ago, and immediately destroy those which do not pass the test. This stage is not necessary in Merz's original setup because with homogeneous workers, all matches generate identical surpluses which should be positive under reasonable calibration. Skill heterogeneity creates dispersion among the surplus delivered by each match, and thus necessitates such an examination for the firms and the workers before actually going to the bargaining table. We will show in the quantitative part of this paper that through such endogenous separation mechanism, skill heterogeneity generates significant differences in the labor market outcomes across skill groups.

Several papers in the literature have already touched on the topic of the skill heterogeneity, such as Pries (2008, [16]) and Ravenna and Walsh (2012, [17]), etc. In spirit, our model tells the same story as these studies. However, in modeling details, our model considers a continuous skill distribution, rather than artificially dividing the working population into the high-skill and the low-skill groups. There are also papers studying heterogeneity in the match-specific productivity, like Merz (1999,[12]). We think our view that skills are characteristics of the workers, rather than that of the matches, may be more relevant for assessing the welfare costs of unemployment, which we have not got the chance to analyze in this paper but is on our future research agenda.

The general equilibrium characterization of our model requires to track the changing skill distribution among the employed workers and to forecast the endogenous separation criteria in the following period. The first issue is largely simplified by our financial market arrangements within the household which make the household-level consumption and hours worked sufficient for the household's decision making. Since the households are identical, they can be represented by a single typical household. Considering the workers' idiosyncratic income shocks would lead to heterogeneous decision makers and greatly increase the difficulty of the general equilibrium analysis. The second issue is approximated by a forecasting rule that depends on a few moments of the skill distribution, the detailed algorithm of which is proposed by Krusel and Smith (1998, [8]), and we iterate until the

resulted policy functions show satisfactorily small changes.

The rest of the paper is organized as follows. Section 2.2 is fully devoted to the description of the equilibrium with search frictions and heterogeneous workers. Section 2.3 characterizes the equilibrium allocation by analyzing an equivalent constrained social planning problem. We also briefly discuss the efficiency issue of the equilibrium in Section 2.3. The quantitative analysis is provided in Section 2.4. Lastly, we conclude with a few plans for the future extension of this project. All proofs are suppressed in the appendix.

2.2 Equilibrium

Our model is built on the business cycle framework with the labor market search frictions developed in Merz (1995, [11]) and Andolfatto (1996, [1]). The economy consists of two types of agents, the workers and the firms. Different from the tradition in the search and matching literature, the workers are endowed with heterogeneous skills which provide them with distinguished identities when they come in front of the recruiting firms. This section details the decision-making problems considered by the workers and the firms in this context and concludes with a full definition of the general equilibrium. Next section will compare the equilibrium allocation with that favored by a social planner.

2.2.1 General Environment

There are a continuum of workers in the economy, each is indexed by his or her skill s which is restricted to a set S. With no population growth, we normalize the total mass of the workers to one. As a result, the population skill distribution can be characterized by a probability mass function $g(s)$ defined for every $s \in S$. Each worker has one unit of discretionary time per period that will be supplied entirely to the firms if he or she is currently employed. We abstract from the consideration of labor market participation in this paper.

Switching from employment to unemployment exposes the individual workers to idiosyncratic income shocks, which, if treated as uninsurable, add much more difficulties in keep the model tractable. As a result, in this paper, we make the simplification assumption of perfect insurance against these idiosyncratic income shocks. To be more

precise, we assume that the workers are randomly distributed in a continuum of mass one identical households, within which all income are shared among the household members. As a decision-making unit, a typical household has the following utility function defined over the total consumption and the total mass of employed workers:

$$u(c, n) = \log c - \frac{B}{1 + \phi} n^{1+\phi}, \tag{2.1}$$

where $n(s)$ is the mass of skill s employed workers and

$$n = \sum_{s \in S} n(s).$$

Notice that the household does not make distinctions among the different skilled workers when calculating the utility loss from working. In other words, skills, as the workers' innate abilities, do not affect their tastes with respect to consumption and leisure.

The economy is also resided by a continuum of mass one firms which operate the same production technology to create output from the physical capital and the skill-adjusted labor. Formally, a firm with $n(s)$ skill s workers and $k(s)$ capital will at most produce

$$y(s) = Ak(s)^\alpha [sn(s)]^{1-\alpha},$$

where A is the economy-wide productivity and $sn(s)$ is the efficient units of labor the $n(s)$ skill s workers supply. Since all production results in the same good, the total production of a firm with employment profile $\{n(s)\}$ and the corresponding capital allocation $\{k(s)\}$ is

$$y = \sum_{s \in S} y(s) = \sum_{s \in S} Ak(s)^\alpha [sn(s)]^{1-\alpha}. \tag{2.2}$$

We assume that the capital rental market is subject to no frictions and is competitive. As a result, taking as given the market rental rate of capital r, a typical firm with employment profile $\{n(s)\}$ faces the following cost-minimization problem if it targets the

44

total output y:

$$\min_{k(s)} \quad r \sum_{s \in S} k(s),$$

$$s.t. \quad \sum_{s \in S} A[k(s)]^{\alpha}[sn(s)]^{1-\alpha} \geq y.$$

The solution $\{k^*(s)\}$ to this problem provides the optimal capital allocation rule the firm will follow. Define $k = \sum_{s \in S} k^*(s)$, we show in Appendix A that the firm will behave in the capital rental market and in the good market as if it has the total production function

$$y = Ak^{\alpha}\left[\sum_{s \in S} sn(s)\right]^{1-\alpha}. \tag{2.3}$$

Hence, the firm views these different types of labor as perfect substitutes, and cares only about a skill-weighted sum of the total labor service delivered by its employees.

The unemployed workers and the recruiting firms form matches according to a random matching technology. Let U denote the mass of unemployed workers. The assumption that job searching is costless makes sure that the households will send all their unemployed workers to hunt for future employment, and therefore causes u to be the mass of the job-seekers in the labor market. Also let V be the mass of vacancies posted by the recruiting firms. The matching function is

$$M = m_0 V^{1-\lambda} U^{\lambda}, \tag{2.4}$$

where m_0 captures the impact of the unemployed worker's constant search effort on the number of matching rendered in the random search and matching process. Our adoption of a constant search effort is to guarantee a negative relationship between the equilibrium unemployment rate and vacancies, or the so-called Beverage curve, which is consistent with the quantitative findings in Merz (1995, [11]). Parameter λ measures the elasticity of the matching function (2.4) with respect to the mass of the unemployed workers. In search models with homogeneous workers and Nash bargaining over the match surplus, the Hosios condition which sets the worker's bargaining power to λ delivers efficiency of the equilibrium. However, the skill heterogeneity adds the compositional externality

that can not be fully internalized in the equilibrium. As a result, our equilibrium fails to achieve the first best allocation. We will return to this issue in more details in the next section.

The undirect feature of the searching and matching technology is supported by several empirical evidence. For instance, Villena-Roldan (2010, [20]) documents that U.S. firms interview a median of five applicants for each filled vacancy and spend around 2.5% of their total labor cost in recruiting activities. Moreover, it is also consistent with our view of the skills as something that influences the worker's labor productivity and may include a few unobserved factors other than education, experience, etc., which the firms could not target effectively in their job posting announcements.

The matching technology (2.4) gives rise to the following aggregate job finding rate and vacancy posting rate that do not distinguish skills:

$$JF = \frac{M}{U}, \quad VF = \frac{M}{V}.$$

However, since the skills are not distributed uniformly among the unemployed workers, the ex post probability that a skill s unemployed worker finds meets a vacancy and the ex post probability that a vacancy is filled by a skill s worker are dependent on s. Let $U(s)$ be the number of skill s unemployed workers, then $U = \sum_{s \in S} U(s)$ and the ex post number of matches that involve a skill s agent is

$$m \cdot \frac{U(s)}{\sum_{s' \in S} U(s')}.$$

We can define the ex post job finding rate for a skill s unemployed worker to be

$$p(s) = \frac{M}{U} \frac{U(s)}{U}. \tag{2.5}$$

Similarly, the ex post vacancy filling rate with respect to hiring a skill s worker is

$$q(s) = \frac{M}{V} \frac{U(s)}{U}. \tag{2.6}$$

Both $p(s)$ and $q(s)$ will be taken as given by the households and the firms in equilibrium.

2.2.2 Timing of Events

The economy starts period t with a set of matched pairs between the workers and the firms. Let $\tilde{N}_t(s)$ be the aggregate mass of skill s workers that are paired with the firms. Period t is divided into two subperiods, the first one of which is devoted to the re-evaluation of the existing matches, while the households and the firms undertake job-hunting and recruiting activities in the second subperiod.

After observing the realization of the aggregate productivity shock A_t in the first subperiod, the matched firms and workers come together to reevaluate the existing employment relations. The unprofitable matches are immediately ended with the involved workers instantly returning to the unemployed pool who are eligible for the job searching in the second subperiod. The remaining workers re-bargain their wages with the firms using Nash bargaining. The first subperiod ends when the bargaining is over.

The firms rent capital and produce output in the second subperiod. Workers get paid afterwards according to the compensation previously bargained, and a fraction ψ of these workers are hit by an exogenous separation shock and become unemployed in next period. At the same time, the unemployed workers are hunting jobs in the labor market and will be matched with the vacancies posted by the recruiting firms as indicated by the matching technology (2.4). These newly matched workers can not produce within the current period, and will be included in the re-evaluation procedure in the following period.

The rest of this section will follow a backward-induction logic by first considering the second subperiod's production and recruitment decisions of the representative household and the representative firm, and then analyzing how they evaluate the existing matches in the first subperiod. This section will close by the definition of the equilibrium. More detailed analysis will be postponed to the next section.

2.2.3 The Representative Household in Subperiod II

As mentioned in the previous subsection, period t starts with a set of already matched workers and firms, which are denoted by $\{\tilde{N}_t(s)\}$. In addition, the households and the firms also know the beginning-period capital stock K_t and the aggregate productivity A_t before making the time t decisions. As a result, the aggregate state Γ_t consists of A_t, K_t and $\{\tilde{N}_t(s)\}$.

From the perspective of the representative household, its second subperiod at time t begins with the employment profile $\{n_t(s)\}$ it mutually determines with the firms. Therefore, its second subperiod value function is defined as

$$W^H(k_t, \{n_t(s)\}, \Gamma_t) = \max \left\{ u(c_t, n_t) + \beta \mathbb{E}_t \left[W^H(k_{t+1}, \{n_{t+1}(s)\}, \Gamma_{t+1}) \right] \right\}, \qquad (2.7)$$

where c_t is the good consumption in period t, $n_t = \sum_{s \in S} n_t(s)$ is the total mass of currently employed workers within the representative household. Utility function $u(\cdot, \cdot)$ is given in (2.1).

Since there is no cost associated with allocating the unemployed workers to job hunting, the total mass of the job-seekers within the representative household will be equal to the mass of the unemployed, i.e., $1 - n_t$. The representative household takes as given the aggregate job finding rate, as well as the aggregate conditional probability of the matched worker being of skill s. In a word, the representative household treats the following ex post probability of a skill s match as given

$$p_t(s) = \frac{M_t}{F_t(1 - N_t)} \cdot \frac{g(s) - N_t(s)}{1 - N_t},$$

where the upper-case variables are the aggregate counterparts of the corresponding lower-case variables. After the job-hunting activity, the representative household expects to have

$$\tilde{n}_{t+1}(s) = (1 - \psi)n_t(s) + p_t(s)(1 - n_t) \qquad (2.8)$$

skill s workers matched with the firms at the beginning of period $t + 1$. Whether or not these $\{\tilde{n}_{t+1}(s)\}$ workers will be kept by the firms depends on the productivity realization in period $t+1$. To assess the continuation value of the time t job searching effort, we endow

the representative household with a perceived rule for the next period's employment decision:

$$n_{t+1}(s) = \chi^H(s, \tilde{n}_{t+1}(s), \Gamma_{t+1}). \tag{2.9}$$

We will follow the rational expectation tradition to assume that the perceived χ^H is consistent with the optimal separation decisions of the households and the firms in equilibrium.

The representative household also takes as given the wages $w_t(s)$ bargained in the first subperiod, so its time t budget constraint is

$$c_t + \left[k_{t+1} - e^{-\mu}(1-\delta)k_t\right] = \sum_s w_t(s)n_t(s) + r_t k_t + \text{Profit}_t. \tag{2.10}$$

Parameter μ accounts for the deterministic technological growth, δ is the depreciation rate of the physical capital. Profits are collected from the firms in the economy, whose shares are equally divided by the households. The representative household optimizes over consumption c_t, future capital stock k_{t+1}, and future matched workers $\{n_{t+1}(s)\}$ in the second subperiod, subject to the budget constraint (2.10), the law of motion for next period's matched skill s agents (2.8), and the perceived employment decision (2.9).

2.2.4 The Representative Firm in Subperiod II

The firms post vacancies that are open to all types of the unemployed agents in the second subperiod. For each vacancy posted by the representative firm, the chance that it will meet a skill s job-seeker is

$$q_t(s) = \frac{M_t}{V_t} \cdot \frac{g(s) - N_t(s)}{1 - N_t},$$

which is the product of the aggregate vacancy filling rate common to all skills, and a skill-specific conditional probability restricted to the current unemployed pool. As a result, the mass of skill s matches the representative firm will obtain at the beginning of time $t+1$ if it posts v_t vacancies is

$$\tilde{n}_{t+1}(s) = (1-\psi)n_t(s) + q_t(s)v_t. \tag{2.11}$$

Similar to the representative household, to determine the continuation value of the vacancy posting, the representative firm also has to be endowed with the perceived rule for next period's hiring decision:

$$n_{t+1}(s) = \chi^F(s, \tilde{n}_{t+1}(s), \Gamma_{t+1}). \tag{2.12}$$

As in the household's problem, we impose the rational expectation requirement to this perceived hiring rule.

Taking as given the bargained wages $\{w_t(s)\}$ and ex post probability of forming a skill s match $\{q_t(s)\}$, the representative firm chooses k_t and v_t to maximize its value in the second subperiod of time t:

$$W^F(\{n_t(s)\}, \Gamma_t) = \max \left\{ y_t - r_t k_t - \sum_{s \in S} w_t(s) n_t(s) + \mathbb{E}_t \left[Q_{t,t+1} W^F(\{n_{t+1}(s)\}, \Gamma_{t+1}) \right] \right\}, \tag{2.13}$$

where $Q_{t,t+1} = \beta C_t / C_{t+1}$ is the stochastic discount factor of its owner, the representative household. The optimization is subject to the total production function (2.3), the law of motion for $\tilde{n}_{t+1}(s)$ (2.11) and the perceived hiring rule next period (2.12).

2.2.5 Screening and Bargaining in Subperiod I

From the household's envelope condition with respect to $n_t(s)$, the marginal value of an additional employed skill s worker with respect to equilibrium employment profile $\{n_t(s)\}$ and the equilibrium wages $\{w_t(s)\}$ is[1]

$$MV_t^H(s) = -B n_t^\phi + \frac{w_t(s)}{c_t} + \mu_t(s)(1-\psi) - \sum_{s' \in S} \mu_t(s') p_t(s'). \tag{2.14}$$

This expression has clear interpretation. The first term measures the direct utility loss at the margin associated with the lost leisure time of the marginal skill s worker. The second term is the income gain from this additional employment, which, adjusted by the marginal utility of consumption, is expressed in utilities rather than in goods. The last two terms relate to the continuation value of a marginal increase in $n_t(s)$, in which $\mu_t(s)$

---

[1]Its formal derivation is included in Appendix B and Appendix C.

is the multiplier to the law of motion for $\tilde{n}_{t+1}(s)$ (2.8). Intuitively, $\mu_t(s)$ measures the impact of the time t actual employment of skill s agents $n_t(s)$ on the time $t+1$ potential employment of the same skill. The additional hired skill s agent positively influences $\tilde{n}_{t+1}(s)$ since it continues into period $t+1$ with probability $1 - \psi$. At the same time, under random matching, increasing current skill s employment means that there will be fewer job-seekers searching for vacancies in the second subperiod, which results in less matches formed, and its total negative effect on the potential employment of all skill levels is captured by the last term above. In Appendix C, we show that this term is actually independent of s.

Similarly, from the firm's envelope condition with respect to $n_t(s)$, we can get how it values hiring an additional skill s agent at the margin:

$$MV_t^F(s) = MPL_t s - w_t(s) + \gamma_t(s)(1 - \psi), \tag{2.15}$$

where

$$MPL_t = (1 - \alpha)A_t k_t^\alpha \left[\sum_{s' \in S} s' n_t(s') \right]^{-\alpha}, \tag{2.16}$$

and $\gamma_t(s)$ is the multiplier to the law of motion of $\tilde{n}_{t+1}(s)$ (2.11). The firm's immediate gain from this extra skill s employment is the difference between its marginal product, which is equal to the average marginal product of labor MPL_t scaled up by s, and the wage $w_t(s)$ it has to pay. The match survives the exogenous separation shock with probability $1 - \psi$, and thus benefits the future skill s employment by $\gamma_t(s)(1 - \psi)$. All the firm-revelent payoffs are measured in units of the output good.

As in Merz (1995), the total surplus generated by a match is the sum of the marginal values it brings to both the household and the firm, where the latter is multiplied by the marginal utility of consumption in order to change its units from good to utility:

$$TS_t(s) = -Bn_t^\phi - \sum_{s' \in S} \mu_t(s')p_t(s') + \frac{MPL_t}{c_t}s + \left[\mu_t(s) + \frac{\gamma_t(s)}{c_t} \right](1 - \psi). \tag{2.17}$$

To get a brief idea of what the above expression of the total surplus illustrates, let us temporarily put aside the multipliers in above expression, which, at the steady state, will be multiples of the other terms in the total surplus expression. The total surplus is

determined by two factors, the negative of the household's utility if this worker remains in idle, adjusted by the costs associated with the searching effort, and the marginal product this worker is able to produce if he or she is employed. Therefore, the total surplus is simply the gap between the benefit of the specific match to the firm and the utility loss suffered by the household.

It can be imagined that if the firm's benefit from the match exceeds the worker's loss, i.e., if the total surplus is positive, it is possible for the firm to offer wage payment to the worker so that the latter is willing to work for it. Otherwise, both the firm and the worker would be better off if they stay unmatched. Therefore, the outcome of the first period's screening process is as follows. The representative household and the representative firm will first review all the existing matches $\{\tilde{n}_t(s)\}$, keeping those $n_t(s) \leq \tilde{n}_t(s)$ if the corresponding total surplus $TS_t(s) \geq 0$ and destroying the rest. Only after finishing this re-evaluation process, will they start bargaining wages for the remaining matches.

The bargaining process is exactly the same as in Merz (1995, [11]) and Andolfatto (1996, [1]). We assume that the representative household and the firms split the non-negative total surplus so that the share going to the household is equal to the elasticity of the matching technology with respect to the total search effort exerted by the household, i.e., let $\tilde{\beta}$ denote the bargaining power of the household, then the following Hosios condition holds:

$$\tilde{\beta} = \lambda.$$

The Nash solution implies that the bargained wage $w_t(s)$ is set so that

$$MV_t^H(s) = \lambda TS_t(s), \quad MV_t^F(s) = (1-\lambda)TS_t(s).$$

The bargained wage will be provided in the next section after introducing the constrained social planning problem that facilitates our characterization of the equilibrium. Let us finish this section by the formal definition of the equilibrium we focus in this paper.

2.2.6 Equilibrium Definition

The equilibrium consists of: (1) A value function W^H and a set of decision rules $\{c_t, k_{t+1}, n_t(s)\}$ for the representative household; (2) A value function W^F and a set of policy functions $\{k_t, n_t(s), v_t\}$ for the representative firm; (3) The market rental rate of capital r_t and a skill-specific wage schedule $\{w_t(s)\}$; (4) The perceived employment decisions χ^H and χ^F for the representative household and the representative firm, respectively; and (5) The perceived law of motion for the aggregate state Γ_t, so that

(i) The representative household's decision rules solve its utility maximization problem given (3), (4) and (5);

(ii) The representative firm's policy rules solve its profit maximization problem given (3), (4) and (5);

(iii) The perceived future hiring rules χ^H and χ^F are consistent with the actual hiring rules mutually determined by the representative household and the representative firm;

(iv) The representative household and the representative firm also correctly perceive the law of motion for the aggregate state Γ_t;

(v) The markets clear given prices (3), meaning that in the good market, the resource constraint holds,

$$c_t + \left[k_{t+1} - e^{-\mu}(1-\delta)k_t\right] + av_t = A_t k_t^\alpha l_t^{1-\alpha}. \tag{2.18}$$

In the rental market for capital service, the demand of the firms is equal to the supply of the households,

$$k_t^F = k_t^H. \tag{2.19}$$

In the labor market, no matches with negative surpluses survive the screening process in the first subperiod of each period, and the wages $\{w_t(s)\}$ are consistent with Nash bargaining.

2.3 Constrained Social Planning Problem

Our quantitative approach follows the tradition in the search and matching literature. We first construct a version of the constrained social planning problem the efficient al-

location implied by which coincides with the equilibrium allocation, then we provide a brief explanation about the equivalence between the equilibrium and the constrained social planning problem. The last subsection supplies a tentative discussion concerning the optimality of the our search equilibrium with heterogeneous workers. Contrast to the findings with single skill, the equilibrium in our paper brings up a compositional externality that can not be fully internalized by the households and the firms even though the bargaining powers are correctly chosen. In our words, in our model, the Hosios condition is not enough to deliver the efficiency of the equilibrium with search frictions.

2.3.1 Constrained Social Planning Problem

As the representative household and the firms in the equilibrium setup, the constrained social planner starts period t with a stock of matches $\{\tilde{n}_t(s)\}$ that were formed in the previous periods, from which the planner will dissolve the unprofitable ones and keep the others. Then the planner will determine how much effort the household is to exert in job-searching, as well as how many vacancies the firms are to post. Matches are generated by the same matching technology as in the equilibrium, and all the producing workers are subject to the same exogenous separation shocks which force them to leave the firms when the current production is done.

Conditional on the matches $\{n_t(s)\}$ that the constrained social planner decides to keep in period t, let us define the total employment and the total efficient units of labor as

$$n_t = \sum_{s \in S} n_t(s), \quad l_t = \sum_{s \in S} s n_t(s).$$

Suppose we continue using the same division of each period into two subperiods, in the first one of which the planner evaluates the profitability of the existing matches, and in the second one of which the production and the recruitment take place. The planner's choice problem in the second subperiod can be written in terms of the above-defined

aggregate variables:

$$W^S(k_t, n_t, l_t, \Gamma_t) = \max \left\{ \log c_t - \frac{B}{1+\phi} n_t^{1+\phi} + \beta \mathbb{E}_t \left[W^S(k_{t+1}, n_{t+1}, l_{t+1}, \Gamma_{t+1}) \right] \right\}, \quad (2.20)$$

$$s.t. \quad c_t + \left[k_{t+1} - e^{-\mu}(1-\delta)k_t \right] + \kappa(f_t)(1-n_t) + av_t = A_t k_t^\alpha l_t^{1-\alpha}, \quad (2.18)$$

$$\tilde{n}_{t+1} = (1-\psi)n_t + v_t^{1-\lambda}[f_t(1-n_t)]^\lambda, \quad (2.21)$$

$$\tilde{l}_{t+1} = (1-\psi)l_t + v_t^{1-\lambda}[f_t(1-n_t)]^\lambda \frac{\sum_{s\in S} sg(s) - L_t}{1 - N_t}, \quad (2.22)$$

$$n_{t+1} = \chi^S(\tilde{n}_{t+1}, \Gamma_{t+1}), \quad l_{t+1} = \zeta^S(\tilde{l}_{t+1}, \Gamma_{t+1}). \quad (2.23)$$

The constrained social planner aims at maximizing the expected discounted utility of the representative household (2.20), subject to the resource constraint (2.18), the laws of motions for $\tilde{n}_{t+1}$ (2.21) and $\tilde{l}_{t+1}$ (2.22), and the perceived hiring decision for the next period (2.23).

In the first subperiod, the constrained social planner faces the following problem:

$$\max_{n_t(s)} \quad W^S(k_t, \sum_s n_t(s), \sum_s sn_t(s), \Gamma_t),$$

$$s.t. \quad 0 \leq n_t(s) \leq \tilde{n}_t(s).$$

The choice of $n_t(s)$ depends on the first order condition with respect to $n_t(s)$

$$\frac{\partial W^S}{\partial n_t} + s \frac{\partial W^S}{\partial l_t},$$

i.e., no matches with above expression being negative will survive the examination of the constrained social planner. In the next subsection, we will study more closely the relationship between the criterion of the constrained social planner and the total surplus $TS_t(s)$ we defined in the equilibrium in Section 2.2, and show that under the Hosios condition, our equilibrium achieves constrained efficiency.

2.3.2 Equilibrium v.s. Constrained Social Planning Problem

Let us start developing the equivalence between the equilibrium allocation and the constrained social planning solution with comparing the first-subperiod's screening process

in these two contexts. For the constrained social planner,

$$\frac{\partial W^S}{\partial n_t} + s\frac{\partial W^S}{\partial l_t} = -Bn_t^\phi - \frac{\lambda}{1-\lambda}\frac{a}{c_t}\frac{v_t}{1-n_t} + \frac{MPL_t}{c_t}s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s], \quad (2.24)$$

where λ_{2t} is the multiplier of constraint (2.21), which measures the effect of the time t's total mass of workers on the the time $t+1$'s total mass of workers. Similarly, λ_{3t} is the multiplier of constraint (2.22), which can be viewed as the impact of the time t's total efficient units of labor on that of time $t+1$.

Equation (2.24) has clear interpretation. Let us temporarily put aside the term containing the multipliers. The constrained social planner's first order condition with respect to $n_t(s)$ consists of three major parts. The first part is the utility loss of the skill s worker at the margin which negatively contributes to the first order condition. The second part is a negative impact coming from the changing vacancy posting cost since the additional employment in the first subperiod of time t reduces the size of job-seekers in the second subperiod. The last part is the gain in production that the marginal skill s worker is able to create. Recall our interpretation of the total surplus (2.17) in Section II. The total surplus with which the representative household and the representative firm decide whether or not to separate an existing match also consists of utility loss to the representative household, production gain to the representative firm, as well as a term related to the reduction of the search and recruiting costs. The only difference is that in equilibrium, the search costs fall on the representative household while in the constrained social planning problem, the planner measures them using the vacancy-post costs incurred by the firms. These two approaches are connected by the planner's first order condition with respect to v_t:

$$a\lambda_{1t} = m_0(1-\lambda)v_t^{-\lambda}(1-n_t)^\lambda \left[\lambda_{2t} + \lambda_{3t}\frac{\sum_{s\in S} sg(s) - L_t}{1-N_t}\right].$$

We show in Appendix C that these two ways of measuring the search and recruiting costs are equivalent, and thus the constrained social planner agrees with the equilibrium households and firms in the screening process.

That the equilibrium supports the constrained social planner's decision in the second subperiod is relatively more standard compared with other studies in the literature. The

Hosios condition, i.e., the assumption that the worker's bargaining power is equal to the elasticity of the matching function with respect to the worker's search effort, λ, implies that after the bargaining, a fraction λ of the total surplus goes to the households, while the rest $1 - \lambda$ is retained by the firms. As already established in the search and matching literature, this arrangement aligns the households' and the firms' incentive of job searching and recruiting in the labor market to that of the constrained social planner. We will leave the technical details to the appendix.

The resulted wage schedule has the following form expressed in utilities[2]:

$$\frac{w_t(s)}{c_t} = \lambda \frac{MPL_t}{c_t} s + (1 - \lambda) \left[Bn_t^\phi - \frac{\kappa(f_t) - \kappa'(f_t)}{c_t} \right]. \tag{2.25}$$

It turns out that the bargained wage is a weighted average of two objects. The first object, $(MPL_t/c_t)s$ is the maximum marginal gain of the firm from this skill s employment. The firm would not want to make the hiring if the household asks for anything above. The second object, the terms in the square brackets, illustrates the utility gain of the household if this skill s worker stays idle instead of working for the firm. Similarly, the household would not agree to providing this skill s labor if the firm offers anything below. These two objects indicate the boundaries of the bargaining set, and for matches with non-negative total surpluses, it is always well-defined and delivers well-behaved non-negative wage. Moreover, (2.25) also suggests that in our model, the bargained wage includes a base pay that is the same for all employees, and a skill-specific pay that gives the relative high skilled a premium due to their greater contribution to the marginal product of labor. In other words, our model supplies an intuitive story for the skill premium, that the high-skill workers get paid more because they are more efficient in the production process.

2.3.3 The Unconstrained Social Planning Problem

To complete the comparison between the equilibrium and the social planning problem, this subsection provides a tentative discussion of the choices that would be favored by the unconstrained social planner. Since it is not the main point of our paper, rather

---

[2]Refer to Appendix C for details about how it is derived.

than delivering a full characteristics of the unconstrained social planning problem, our emphasis will be placed on its differences from the constrained one, and how it improves upon the equilibrium allocation which achieves the constrained efficiency as defined in the previous subsection.

The unconstrained social planner faces the same problem as the constrained social planner in the second subperiod of time t, i.e., maximizing the life-time utility of the households (2.20), subject to the resource constraint (2.18), the laws of motion for next period's potential employment (2.21) and efficient units of labor (2.22), and the perceived hiring decision (2.23). However, when determining $n_t(s)$ in the first subperiod, the constrained social planner views its impact on the law of motion for $\tilde{l}_{t+1}$ as exogenous,

$$\tilde{l}_{t+1} = (1-\psi)l_t + m_0 v_t^{1-\lambda}(1-n_t)^\lambda \frac{\sum_{s \in S} sg(s) - L_t}{1 - N_t},$$

where upper-case letters N_t and L_t indicate that they are not choice variables for the constrained social planner, the unconstrained social planner takes into account such quality change in the unemployed pool:

$$\tilde{l}_{t+1} = (1-\psi)l_t + m_0 v_t^{1-\lambda}(1-n_t)^\lambda \frac{\sum_{s \in S} sg(s) - l_t}{1 - n_t}.$$

As a result, the first order condition of the unconstrained social planner with respect to $n_t(s)$ in the first subperiod becomes:

$$\frac{\partial W^{US}}{\partial n_t(s)} = -Bn_t^\phi - \frac{\lambda}{1-\lambda}\frac{a}{c_t}\frac{v_t}{1-n_t} + \frac{MPL_t}{c_t}s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s]$$
$$+ m_0 v_t^{1-\lambda}(1-n_t)^\lambda \left[\frac{\sum_{s' \in S} s'g(s') - l_t}{1 - n_t} - s\right].$$

Compared to the corresponding first order condition (2.24), the unconstrained social planner's first order condition contains an additional term, which measures the impact of $n_t(s)$ on the unemployed pool that both the households and the firms will face in the labor market opening in the second subperiod. Intuitively, $[\sum_{s' \in S} s'g(s') - l_t]/[1 - n_t]$ is the average skill among the unemployed workers if the first subperiod's decision leads to $\{n_t(s)\}$. If the planner decides to have more workers with s higher than this average employed, the quality of the unemployed pool will deteriorate and eventually translate

58

into a negative effect in the labor market through the law of motion for $\tilde{l}_{t+1}$ (2.22). The compositional effect goes the opposite way if the workers being hired are of low-than-average-unemployed skill. By taking as given the share of skill s worker in the unemployed pool, $[g(s) - N_t(s)]/[1 - N_t]$, both the equilibrium agents and the constrained social planner fail to internalize this externality. That is why the equilibrium does not achieve the first best, even though the Hosios condition is assumed to hold.[3]

2.4 Quantitative Analysis

The major question we explore in this paper is whether or not the observed differences in the labor market outcomes among workers can be explained by their heterogeneous skills as shown in the good production. As a result, the quantitative analysis performed in this section will mainly focus on the cross-sectional implications of our model. Moreover, to highlight the impact of the skill heterogeneity, we will suppress other possible types of heterogeneity in our model, and base our numerical experiments on a calibrated economy that mostly targets on the aggregate properties of the U.S. economy. We will start our discussion with parameterizing and calibrating the parameters in our model, and then turn to its cross-sectional implications and how they are compared to the data.

2.4.1 Parameterization and Calibration

We need to specify the stochastic process of the aggregate productivity before the calibration. As in the standard real business cycle literature, we assume the following AR(1) process:

$$A_t = \exp[(1 - \alpha)(\mu t + z_t)], \text{ where } z_t = \rho_Z z_{t-1} + e_{Zt}, \tag{2.26}$$

and the innovations e_{Zt} are i.i.d. $N(0, \sigma_Z^2)$. Parameter μ is the rate of the deterministic technological growth, which is taken as 0.4% per quarter. The persistence ρ_Z and the volatility σ_Z are assigned their typical values.

Table 2.1 shows our baseline calibration, which treats each model period as one quarter in the data. As can be easily seen, most of the parameters concerning the firms' produc-

---

[3]We can also illustrate the suboptimality of the equilibrium allocation by finding a Pareto improvement. Appendix D contains such an example.

tion or the household preference are given their widely used values in the macroeconomic literature. For instance, the capital share in the Cobb-Douglas production function α is 0.36. Capital depreciates by 2.2% each quarter. The household's discount factor β is 0.99, which implies a 4% annual interest rate at the steady state. The Frisch elasticity of the labor supply is chosen to be 1, suggesting a value for ϕ equal to 1 as well. Under the Hosios condition, the elasticity of the matching technology with respect to the household searching effort λ is calibrated to be the average share of the total surplus that goes to the household in the Nash bargaining, 40%.

Table 2.1: Calibration

Production & Productivity		
α	0.36	capital share
δ	0.022	capital depreciation rate
μ	0.4%	balanced growth rate
ρ_Z	0.95	persistence
σ_Z	$0.007/(1-\alpha)$	volatility
Preference		
B	0.90	disutility
β	0.99	discount factor
ϕ	1	Frisch elasticity of labor supply 1
Search & Matching		
a	1.00	linear vacancy posting cost
λ	0.4	worker's share in the matching function
m_0	1.50	scale parameter of the matching function
ψ	0.0218	exogenous separation rate
Skill Distribution		
σ_S	0.0082	standard deviation of the log(skill)

The remaining parameters, i.e., the scale of the household utility loss from working B, the linear vacancy posting cost a, the scale parameter in the matching function m_0, as well as the transitional rate from employment to unemployment ψ, are calibrated using the U.S. labor market related data from BLS.[4] We look at several targets, all of which but one concern the aggregate labor market. The first target is the average unemployment rate which is around 5.7% per quarter from 1948 Q1 to 2010 Q1. We also target the average job finding rate, which is 0.6134 according to the series constructed by Robert

---

[4]To be more precise, the unemployment series are from CPS. The total number of separations and the number of layoffs are from JOLTS, starting 2002. We also check the job finding rate series and the separation rate series available in Robert Shimer's website (http://home.uchicago.edu/ shimer/data/).

Shimer. The last target is the average layoff rate, defined as the ratio of the number of layoffs to the number of the total employment. Its quarterly average over 2002 and 2011 is 1.5%, as calculated from data collected by the Job Opening and Labor Turnover Survey (JOLTS), which begins in the year 2002. It worths mentioning that in this paper we use exclusively the layoff rate as a proxy for the endogenous separation rate. This is mainly because that the endogenous separation mechanism in our model works by screening out the less efficient workers in the production process, which is more likely to be paralleled by the layoff procedure in the real world. It should be admitted that in the data, a large portion of separations happen when workers willingly leave the job. However, given the simplified household structure we impose to our model to keep it tractable, we can do nothing but to reserve the workers' incentive to quit for future scrutiny.

The last thing to specify is the population skill distribution. In our quantitative experiments, we assume that the logarithmic skill is shaped by a normal distribution with mean zero, whose spread is chosen with other parameters mentioned above to ensure that the steady state prediction of the endogenous separation rate is in line with the observed layoff rate. To be more precise, at the de-trended steady state, there exists a cutoff of skill, call it $\underline{s}$, so that any individuals with $s < \underline{s}$ would not be hired even if they are paired with vacancies posted by the firms. Once the steady state mass of employed workers n is fixed, the cutoff $\underline{s}$ can be recovered from the zero total surplus condition and it does not depend on the shape of the skill distribution. Therefore, we can alter the spread of the skill distribution so that the resulted fraction of skill realization that lies below $\underline{s}$ is consistent with the one implied by the layoff rate. It turns out that the standard deviation of the logarithmic skill is 0.82%, or equivalently, over 99% workers have skills within the interval $[0.95, 1.05]$, and a fraction of around 2.34% of the workers at the bottom of the skill distribution will never be employed at the steady state.[5]

Our calibration implies a total separation rate equal to 3.65%, fairly close to its quarterly average 3.53% suggested by JOLTS. The total resource spent in the job searching and recruiting process takes up only a small fraction of the aggregate output (less than 0.5%), which complies well with the empirical findings that the resource costs of the search frictions are by themselves insignificant.

---

[5]Figure (2.1) in the appendix includes a graphic representation of the population skill distribution and the mass of skill s workers that are employed at the steady state.

2.4.2   Model Implication across Skill Groups

Figure (2.1) in the appendix depicts the skill distribution in the whole population (the blue dash curve) and how it spreads among the employed workers at the steady state (the red solid curve). The black dotted vertical lines indicates the critical value of the skill which generates zero surplus at the steady state. As a result, no workers with skills to the left of this cutoff value are employed. To the right of the cutoff, unemployment arises due to the fact that the matches are subject to exogenous separation and that search frictions prevent those who get separated from instantly finding new employers. We can also see from Figure (2.1) how skill heterogeneity contributes to the observed differences across skill groups in our model. Since the endogenous separation exclusively concentrates on the bottom of the skill distribution, it creates much more difficulties for the less skilled workers to establish a long-term employment relationship with the firms. When the economy is hit by exogenous shocks, the zero-surplus cutoff will shift accordingly, influencing dramatically the employment status of workers whose skills are in its neighborhood. As a result, these workers are also subject to more volatile unemployment risks. Next we will show that these statements are consistent with the quantitative implications of our model and compare their magnitude to what is observed in the data.

Before presenting the results, we will offer a few more words concerning how we perform the analysis. As in the data, our results focus on only a finite number of skill groups rather than a continuous skill distribution. To be more specific, we split the skill distribution into two skill groups. The low skill group in the data corresponds to workers between age 16 and 24, who are both less educated and less experienced relative to workers over age 25 who constitute the high skill group. All series in the calculation are taken from the Current Population Survey (CPS) published by the Bureau of Labor Statistics. The sample period is from 2001 Q1 to 2011 Q4, which is restricted by the availability of the unemployment rates and the wages across these two age groups.[6] Under this specification, the average share of the low skill group in the total labor force is around 15%. Accordingly, we call the workers at the bottom 15% of the skill distribution in our

---

[6]Due to the Great Recession, the average total unemployment within this sample period, 6.31%, is above the one we use in the calibration, 5.7%. The 5.7% unemployment rate is the average between 1951 Q1 and 2004 Q4, which is covered by Robert Shimer's estimates for the job finding rate and the separation rate. Our calibration chooses to target this longer time period since it is our source of the average job finding rate.
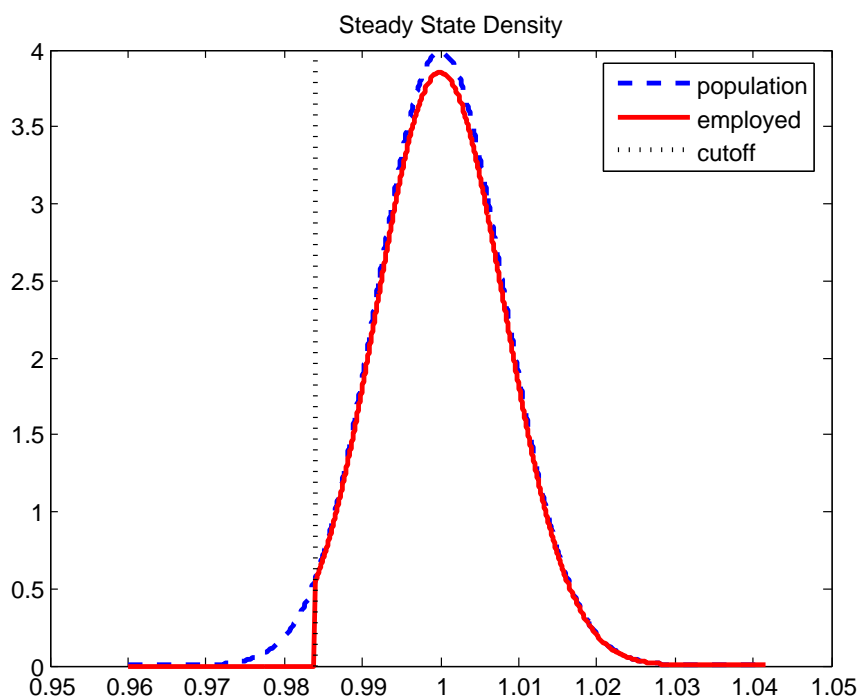
Figure 2.1: Steady State Skill Distribution

The graph shows the steady state density function of the skills in the whole population (blue dashed curve) and within the employed workers (red solid curve). The black dotted vertical line shows the cutoff below which the corresponding workers are so unproductive that they will never be hired by the firms even though the two parties do meet each other. Under the baseline calibration, the fraction of these extremely unskilled workers is tiny, around 2.34%.

model the low skill workers, and define the rest to be the high skill workers.

The steady state implications of our model is shown in Table 2.2, and Table 2.3 offers their empirical counterparts. In the data, quite intuitively, the low skill group has a much higher average unemployment rate, 13.12%, than the high skill group, 5.15%, while the whole economy's unemployment rate is 6.31%. The low skilled are also less paid: The median low skill worker's wage is around 60% that of the median wage in the whole economy, while the median high skill worker's wage is 6% higher than the economy-wide median wage. Our model correctly captures these differences across skill groups, as indicated by Table 2.2. The low skill group experiences an average unemployment rate equal to 17.57%, and the average wage within the low skill group is 87% of the average wage in the whole economy. On the other hand, the unemployment rate for the high skill group is only 3.44% and their average wage is 2% above the average wage over all workers. It should be admitted that our model underestimates the cost of being unemployed for the less skilled workers and thus overstates their unemployment rate. This

63

Table 2.2: Steady State (Model)

Average	Low	High	Total
U	17.57%	3.44%	5.70%‡
W (mean)	0.87	1.02	1†
Labor force	0.16	0.84	1†
Job finding	0.5056	0.6134	0.6134‡
Total separation	——	——	3.65%
Layoff rate	——	——	1.50%‡

†: normalization

‡: calibration targets

Table 2.3: Long Run Averages (Data)

Average	Low	High	Total
U	13.12%	5.15%	6.31%
W (median)	0.61	1.06	1†
Labor force	0.15	0.85	1†
Job finding	——	——	0.6134
Total separation	——	——	3.53%
Layoff rate	——	——	1.43%

†: normalization

might be caused by two factors. The first is the arbitrariness of our selection of the skill distribution. The log-normal skill distribution is chosen for convenience, and we plan to do more research with respect to the empirical distribution of the skills in the data. The second and maybe more crucial factor is the perfect insurance market we assume to exist within each household. The possibility for the low skill workers to diversify their income risks with their more productive fellowmen in the same household raises their outside option as oppose to being employed, and thus leads to an exaggerated unemployment rate prediction of the model. Our model could be improved along this dimension if we have a more realistic household structure that takes seriously the idiosyncratic income risks into consideration.

The dynamic properties of our model are first hinted by the impulse response functions shown in Figure (2.2). After a negative productivity shock, the typical aggregate variables, such as output Y, consumption C, investment I and total employment N, all shown signs of getting into a recession followed by a gradual recovery. The total number of vacancies also fall at the onset of the negative productivity shock, even though it
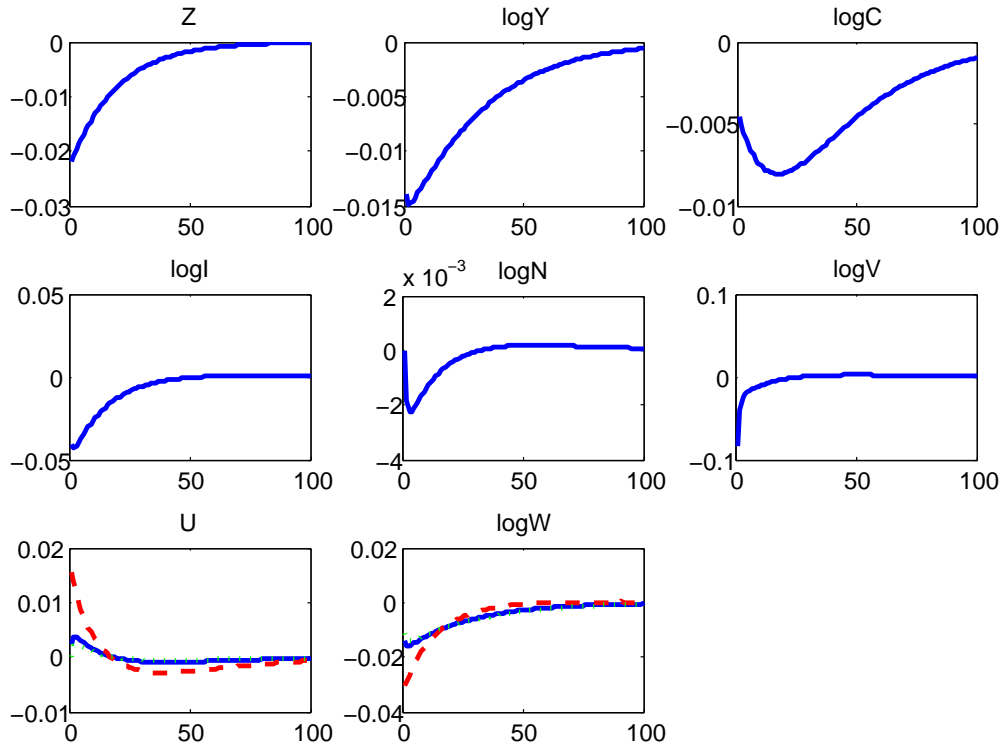
Figure 2.2: Impulse Response to Adverse Productivity Shock
The graph provides the impulse response functions of the major variables to a sizable adverse productivity shock (two-standard-deviation shock). The responses of the aggregate variables are given by the blue solid curves. The red dashed curves are responses for the low-skill group (age 16 to 24), while the green dotted curves are for the high-skill group (over age 25). All variables except the unemployment rate are deviations from the deterministic steady state, while the unemployment rate is simply the deviation from its steady state value.

recovers much more quickly. The two panels at the bottom show the responses of the unemployment rate and the wage. The blue solid curve refers to the aggregate economy, while the one for the low skill group is in red, and the one for the high skill group is in green. It can be seen that the high skill group shows similar dynamics as the whole economy, because the high skill workers make up the majority of the economy and our endogenous separation mechanism mainly targets the low skill group. The sharp increase of the low skill unemployment rate and the moderate increase of the high skill unemployment rate resulted from the negative productivity shock are consistent with what people observe in the real world. For instance, Elsby et al. (2011, [6]) find using data from the Great Recession that recessions are the times that all workers are subject to higher unemployment risks, yet the less skilled are particularly vulnerable. Our model

also predicts that the average wage in the low skill group falls more during a recession, which would make their situation even worse. However, again, since our simplified household structure that abstracts from any form of idiosyncratic income shocks, we have to postpone the welfare analysis to the future.

The business cycle statistics calculated from the real data and our model simulations are given in Table 2.4 and Table 2.5, respectively. Our main comparison will be done for the unemployment rate, for with respect to the wage, our model has different definition from the data. Our data series is the median wage in the whole population or within the specific age group, while in the model, it is more convenient for us to calculate the corresponding mean wages. This discrepancy of the standards makes it difficult to assess our model in the light of the data, and thus we decide to put more emphasis on the unemployment rate which can be treated in a uniform way.

Table 2.4: Business Cycle Properties (Data)

$std(\cdot)/std(Y)$	Low	High	Total
$std(Y)$	1.32%		
U	1.01	0.63	0.67
W	1.77	0.52	0.72
$corr(\cdot,Y)$	Low	High	Total
U	-0.7772	-0.8186	-0.8160
W	-0.1545	-0.1387	-0.3643
$corr(\cdot,N)$	Low	High	Total
U	-0.9836	-0.9979	-0.9994
W	0.0891	0.0408	-0.1664
$corr(\cdot,V)$	Low	High	Total
U	-0.9297	-0.9308	-0.9364
W	-0.0486	-0.1251	-0.2847
Skill premium	Y	N	V
$corr(sp,\cdot)$	0.1247	-0.0844	0.0133

As in other search models of the labor market, our model underestimates the volatility of the unemployment rate relative to output. We can improve upon this issue by allowing variable search effort exerted by the unemployed workers. However, as Merz (1995, [11]) indicates, varying search cost gives rise to a positive correlation between the number of vacancies and the unemployment rate, which contradicts to the empirically identified downward sloping Beverage Curve. Since our goal is not to amplify the predicted volatility

Table 2.5: Business Cycle Properties (Model)

$std(\cdot)/std(Y)$	Low	High	Total
$std(Y)$	0.91%		
U	0.85	0.15	0.24
W	1.81	0.92	1.04
$corr(\cdot, Y)$	Low	High	Total
U	-0.9505	-0.7077	-0.9241
W	0.9859	0.9949	0.9978
$corr(\cdot, N)$	Low	High	Total
U	-0.6040	-0.9944	-0.8755
W	0.6507	0.7695	0.7445
$corr(\cdot, V)$	Low	High	Total
U	-0.8120	-0.0903	-0.5139
W	0.7872	0.6551	0.6904
Skill premium	Y	N	V
$corr(sp, \cdot)$	-0.9241	-0.5006	-0.8745

of the unemployment rate, we will keep the assumption of an exogenously given search effort. In the data, the unemployment rate for the low skill workers is around twice of that for the high skill workers. Our model captures this qualitatively, but with a low skill unemployment rate is over 5 times more volatile than the high skill unemployment rate, our model again overstates the magnitude due to the perfect insurance assumption we adopt to keep the model tractable. As in the data, our model predicted unemployment rate is negatively correlated with output Y, employment N and vacancy V, at both the aggregate level or across different skill groups.

2.5   Conclusion

In this paper, we extend the search and matching model developed in Merz (1995, [11]) to the case where the workers are heterogeneous in their skills. Skill heterogeneity creates dispersion in the total surpluses generated by the matches between the firms and the workers and thus gives rise to an endogenous separation mechanism which screens out the least productive workers from the workforce. With respect to wage determination, the workers and the firms in our model still engage in Nash bargaining to split the surpluses for the matches that survive the screening process. However, the traditional Hosios condition, i.e., the condition that the workers' bargaining power is equal to the

elasticity of the random matching function with respect to the workers' search effort, is no longer sufficient for the efficiency of the equilibrium due to the so-called compositional externality. Quantitatively, our model has cross-sectional implications consistent with the observed labor market experience for workers in different skill groups.

Our next step is to relax the perfect intra-household financial market assumption we currently maintain for tractability concern. Perfectly insurable income risks lower the cost of being unemployed for the less skill workers, and thus make them too likely to reject the wages proposed by the firms. Directly consider partly uninsurable income risks should help mitigate our model's overstatement about the unemployment rate for the low skill workers, both in terms of levels and in terms of volatility. Moreover, we also would like to perform a welfare analysis for individual workers to assess the adverse influence of unemployment risks, as well as the effectiveness of the policies that are usually adopted by the government to combat against unemployment. All these are saved for future scrutiny.

2.6 Appendix

2.6.1 Appendix A: The Total Production Function with Heterogeneous Labor

In this appendix, we will characterize the firm's optimal capital allocation and derive the profit-maximizing output

$$y_t = A_t k_t^\alpha \left[\sum_s s n_t(s) \right]^{1-\alpha}$$

for a firm that has employment profile $\{n_t(s)\}$ and optimally uses in total k_t units of capital in the production. As a price-taker in the rental market for capital service, the firms solves

$$\max_{k_t(s)} \sum_s A_t k_t(s)^\alpha [s n_t(s)]^{1-\alpha} - r_t \sum_s k_t(s) - \sum_s w_t(s) n_t(s),$$

The first order condition with respect to $k_t(s)$ is

$$A_t \alpha \left[\frac{s n_t(s)}{k_t(s)} \right]^{1-\alpha} = r_t, \quad \forall s.$$

Therefore,

$$k_t^*(s) = \left[\frac{A_t \alpha}{r_t} \right]^{\frac{1}{1-\alpha}} s n_t(s),$$

and the optimal output produced by skill s workers is

$$y_t^*(s) = A_t \left[\frac{A_t \alpha}{r_t} \right]^{\frac{\alpha}{1-\alpha}} s n_t(s)$$

Summing over all skill levels, the total output produced by this firm is

$$y_t = \sum_s y_t^*(s) = A_t \left[\frac{A_t \alpha}{r_t} \right]^{\frac{\alpha}{1-\alpha}} \sum_s s n_t(s).$$

Write k_t as the total amount of capital used,

$$k_t = \sum_s k_t^*(s) = \left[\frac{A_t \alpha}{r_t} \right]^{\frac{1}{1-\alpha}} \sum_s s n_t(s),$$

it is straight-forward to see that

$$y_t = A_t k_t^\alpha \left[\sum_s s n_t(s) \right]^{1-\alpha}$$

2.6.2 Appendix B: Marginal Value of Skill s Employment

Relative to the equilibrium mass of skill s workers $n_t(s)$ chosen in Subperiod I of time t and the corresponding bargained wages $w_t(s)$, suppose the representative household now has an additional ϵ skill s workers employed at wage w, while for all other skill $\hat{s} \neq s$, $n_t(s')$ and $w_t(s')$ are as in equilibrium. Moreover, also assume that the representative household follows the equilibrium decision rules from time $t+1$ on. We can define the

representative household's value function with above deviation from equilibrium to be

$$\tilde{W}^H(k_t, \{n_t(s)\}, s, \epsilon, \Gamma_t) = \max \left\{ \log c_t - \frac{B}{1+\phi} \left[\sum_{s \in S} n_t(s) + \epsilon \right]^{1+\phi} \right.$$

$$+ \beta \mathbb{E}_t \left[W^H(k_{t+1}, \{n_{t+1}(s)\}, \Gamma_{t+1}) \right] \Big\},$$

$$s.t. \quad c_t + k_{t+1} - e^{-\mu}(1-\delta)k_t = \sum_{s \in S} w_t(s)n_t(s) + w\epsilon + r_t k_t + \text{Profit}_t, \quad \tilde{\mu}_{1t},$$

$$\tilde{n}_{t+1}(s) = (1-\psi)[n_t(s) + \epsilon] + p_t(s) \left[1 - \sum_{s' \in S} n_t(s') - \epsilon \right], \quad \tilde{\mu}_{2t}(s),$$

$$\tilde{n}_{t+1}(\hat{s}) = (1-\psi)n_t(\hat{s}) + p_t(\hat{s}) \left[1 - \sum_{s' \in S} n_t(s') - \epsilon \right], \quad \tilde{\mu}_{2t}(\hat{s}),$$

$$n_{t+1} = \chi^H(s, \tilde{n}_{t+1}(s), \Gamma_{t+1}).$$

Its partial derivative respect to ϵ is

$$\frac{\partial \tilde{W}^H}{\partial \epsilon} = -B \left[\sum_{s \in S} n_t(s) + \epsilon \right]^\phi + \tilde{\mu}_{1t} w + (1-\psi)\tilde{\mu}_{2t}(s) - \sum_{s' \in S} \tilde{\mu}_{2t}(s)p_t(s').$$

The marginal value of skill s employment is thus the limit of above expression as $\epsilon \to 0$:

$$MV_t^H(s) = -Bn_t^\phi + \frac{w}{c_t} + (1-\psi)\mu_{2t}(s) - \sum_{s' \in S} \mu_{2t}(s')p_t(s'),$$

where multipliers $\mu_{1t} = 1/c_t$ and $\tilde{\mu}_{2t}(s) \to \mu_{2t}(s)$ as $\epsilon \to 0$. Treating w as the equilibrium wage $w_t(s)$, we obtain the marginal value of a skill s employment to the representative household (Equation (2.14)).

Similar, consider the situation that if the representative firm decides to hire an additional skill s workers at wage w, i.e., the employed mass of skill s workers is $n_t(s) + \epsilon$, while keeping all other $n_t(s')$ and $w_t(s')$ as in the equilibrium. The firm's new value

function becomes

$$\tilde{W}^F(\{n_t(s)\}, s, \epsilon, \Gamma_t) = \max \left\{ y_t - r_t k_t - \sum_{s \in S} w_t(s) n_t(s) - w\epsilon - a v_t \right.$$

$$\left. + \beta \mathbb{E}_t \left[Q_{t,t+1} W^F(\{n_{t+1}(s)\}, \Gamma_{t+1}) \right] \right\},$$

$$s.t. \quad y_t = A_t k_t^\alpha \left[\sum_{s \in S} s n_t(s) + s\epsilon \right]^{1-\alpha},$$

$$\tilde{n}_{t+1}(s) = (1 - \psi)[n_t(s) + \epsilon] + q_t(s) v_t, \quad \tilde{\gamma}_t(s),$$

$$\tilde{n}_{t+1}(\hat{s}) = (1 - \psi) n_t(\hat{s}) + q_t(\hat{s}) v_t, \quad \tilde{\gamma}_t(\hat{s}),$$

$$n_{t+1}(s) = \chi^F(s, \tilde{n}_{t+1}, \Gamma_{t+1}).$$

The derivative of $\tilde{W}^F$ with respect to ϵ is

$$\frac{\partial \tilde{W}^F}{\partial \epsilon} = MPL_t s - w + (1 - \psi) \tilde{\gamma}_t(s).$$

Taking the limit as $\epsilon \to 0$ and substituting w for $w_t(s)$ yields the marginal value of a skill s employment to the representative firm as in (2.15).

2.6.3 Appendix C: Equilibrium v.s. Constrained Social Planning Problem

Equilibrium

The representative household solves the following problem in the second subperiod of time t:

$$W^H(k_t, \{n_t(s)\}, \Gamma_t) = \max \left\{ \log c_t - \frac{B}{1+\phi} \left[\sum_{s \in S} n_t(s) \right]^{1+\phi} + \beta \mathbb{E}_t \left[W^H(k_{t+1}, \{n_{t+1}(s)\}, \Gamma_{t+1}) \right] \right\},$$

$$s.t. \quad c_t + k_{t+1} - e^{-\mu}(1 - \delta) k_t = \sum_{s \in S} w_t(s) n_t(s) + r_t k_t + \text{Profit}_t, \quad \mu_{1t},$$

$$\tilde{n}_{t+1}(s) = (1 - \psi) n_t(s) + p_t \frac{g(s) - N_t(s)}{1 - N_t} \left[1 - \sum_{s' \in S} n_t(s') \right], \quad \mu_{2t}(s),$$

$$n_{t+1} = \chi^H(s, \tilde{n}_{t+1}(s), \Gamma_{t+1}).$$

Let μ_{1t} and $\mu_{2t}(s)$ be multipliers for the budget constraint and the law of motion

for the number of skill s match at the beginning of the next period.[7][8] The first order conditions are

$$c_t : \frac{1}{c_t} - \mu_{1t} = 0;$$

$$k_{t+1} : \beta\mathbb{E}_t\left[W^H_{1t+1}\right] - \mu_{1t} = 0;$$

$$n_{t+1}(s) : \beta\mathbb{E}_t\left[W^H_{2t+1}(s)\right] - \mu_{2t}(s) = 0,$$

where

$$W^H_{1t+1} \triangleq \frac{\partial}{\partial k_{t+1}}W^H(k_{t+1}, \{n_{t+1}(s)\}, \Gamma_{t+1}),$$

$$W^H_{2t+1}(s) \triangleq \frac{\partial}{\partial n_{t+1}(s)}W^H(k_{t+1}, n_{t+1}(s), \{n_{t+1}(s')\}_{s'\neq s}, \Gamma_{t+1}).$$

The envelope conditions with respect to k_t and $n_t(s)$ are

$$W^H_{1t} = \mu_{1t}\left[r_t + e^{-\mu}(1 - \delta)\right],$$

$$W^H_{2t}(s) = -B\left[\sum_{s'\in S}n_t(s')\right]^\phi + \mu_{1t}w_t(s) + \mu_{2t}(s)(1 - \psi) - p_t\sum_{s'\in S}\mu_{2t}(s')\frac{g(s') - N_t(s')}{1 - N_t}.$$

Conditional on the employment $\{n_t(s)\}$ and the bargained wage $\{w_t(s)\}$, the representative firm faces the following problem in the second subperiod:

$$W^F(\{n_t(s)\}, \Gamma_t) = \max\left\{y_t - r_tk_t - \sum_{s\in S}w_t(s)n_t(s) - av_t + \beta\mathbb{E}_t\left[Q_{t,t+1}W^F(\{n_{t+1}(s)\}, \Gamma_{t+1})\right]\right\},$$

$$s.t. \quad y_t = A_tk_t^\alpha\left[\sum_{s\in S}sn_t(s)\right]^{1-\alpha},$$

$$\tilde{n}_{t+1}(s) = (1 - \psi)n_t(s) + q_t\frac{g(s) - N_t(s)}{1 - N_t}v_t, \gamma_t(s),$$

$$n_{t+1}(s) = \chi^F(s, \tilde{n}_{t+1}, \Gamma_{t+1}).$$

---

[7]The difference between n_{t+1} and $\tilde{n}_{t+1}$ is due to the matches that turn out to be unproductive next period and thus get destroyed by the households and the firms. As a result, we consider the following form for the perceived relationship:

$$n_{t+1}(s) = \tilde{n}_{t+1}(s) - x^N_t(s, \Gamma_{t+1}).$$

[8]Multiplier $\mu_{2t}(s)$ is the $\mu_t(s)$ in the main body of the paper, especially in Section II and Section III.

where $Q_{t,t+1} = \beta C_t / C_{t+1}$.

Attach $\gamma_t(s)$ to the law of motion for $\tilde{n}_{t+1}(s)$, the first order conditions of the firm are

$$k_t : A_t \alpha k_t^{\alpha-1} \left[\sum_{s' \in S} s' n_t(s') \right]^{1-\alpha} - r_t = 0;$$

$$n_{t+1}(s) : \beta \mathbb{E}_t \left[Q_{t,t+1} W_{1t+1}^F(s) \right] - \gamma_t(s) = 0;$$

$$v_t : -a + q_t \sum_{s' \in S} \gamma_t(s') \frac{g(s) - N_t(s)}{1 - N_t} = 0,$$

with the notation

$$W_{1t+1}^F(s) =\triangleq \frac{\partial}{\partial n_{t+1}(s)} W^F(k_{t+1}, n_{t+1}(s), \{n_{t+1}(s')\}_{s' \neq s}, \Gamma_{t+1}).$$

The envelope condition is

$$W_{1t}^F(s) = A_t k_t^\alpha (1-\alpha) \left[\sum_{s' \in S} s' n_t(s') \right]^{-\alpha} s - w_t(s) + \gamma_t(s)(1-\psi).$$

Define

$$MPK_t \triangleq A_t \alpha k_t^{\alpha-1} \left[\sum_{s \in S} s n_t(s) \right]^{1-\alpha}, \quad MPL_t \triangleq A_t k_t^\alpha (1-\alpha) \left[\sum_{s \in S} s n_t(s) \right]^{-\alpha}.$$

Both notations will be used throughout the rest of Appendix B.

Back to the screening process in the first subperiod, the total surplus generated by a skill s match is

$$TS_t(s) = W_{2t}^H(s) + \frac{1}{c_t} W_{1t}^F(s)$$

$$= -B \left[\sum_{s' \in S} n_t(s') \right]^\phi - p_t \sum_{s' \in S} \mu_{2t}(s') \frac{g(s') - N_t(s')}{1 - N_t} + \frac{MPL_t}{c_t} s + \left[\mu_{2t}(s) + \frac{\gamma_t(s)}{c_t} \right] (1-\psi)$$

The employment decision will be made jointly by the households and the firms so that in equilibrium, no matches with negative surpluses will be allowed to survive the screening process. Next we will turn to the discussion of the constrained social planning problem.

Constrained Social Planning Problem:

The social planner is constrained in the sense that in the second subperiod matching process, the composition of the unemployed labor force is taken as given. Formally, the employment decision $\{n_t(s)\}$ at the beginning of time t leads to the aggregate number of workers

$$n_t = \sum_{s \in S} n_t(s),$$

as well as the aggregate efficient units of labor

$$l_t = \sum_{s \in S} s n_t(s).$$

The second subperiod's social planning problem can be written in terms of these aggregate variables:

$$W^S(k_t, \{n_t(s)\}, \Gamma_t) = \max \left\{ \log c_t - \frac{B n_t^{1+\phi}}{1+\phi} + \beta \mathbb{E}_t \left[W^S(k_{t+1}, n_{t+1}, l_{t+1}, \Gamma_{t+1}) \right] \right\}$$

s.t. $c_t + k_{t+1} - e^{-\mu}(1-\delta)k_t + av_t = A_t k_t^\alpha l_t^{1-\alpha}, \quad \lambda_{1t}$

$\tilde{n}_{t+1} = (1-\psi)n_t + m_0 v_t^{1-\lambda}(1-n_t)^\lambda, \quad \lambda_{2t}$

$\tilde{l}_{t+1} = (1-\psi)l_t + m_0 v_t^{1-\lambda}(1-n_t)^\lambda \frac{\sum_{s \in S} sg(s) - L_t}{1 - N_t}, \quad \lambda_{3t}$

Laws of motion for n_{t+1}, l_{t+1} and Γ_{t+1},

where

$$\frac{\sum_{s \in S} sg(s) - L_t}{1 - N_t} = \sum_{s \in S} s \frac{g(s) - N_t(s)}{1 - N_t},$$

the conditional probability that the newly formed match involves a skill s worker $[g(s) - N_t(s)]/[1 - N_t]$ is treated as exogenous by the constrained social planner.

The corresponding first order conditions are

$$c_t : \frac{1}{c_t} - \lambda_{1t} = 0,$$

$$k_{t+1} : \beta \mathbb{E}_t \left[W_{1t+1}^S \right] - \lambda_{1t} = 0,$$

$$n_{t+1} : \beta \mathbb{E}_t \left[W_{2t+1}^S \right] - \lambda_{2t} = 0,$$

$$l_{t+1} : \beta \mathbb{E}_t \left[W_{3t+1}^S \right] - \lambda_{3t} = 0,$$

$$v_t : -a\lambda_{1t} + m_0(1-\lambda)v_t^{-\lambda}(1-n_t)^{\lambda} \left[\lambda_{2t} + \lambda_{3t} \frac{\sum_{s \in S} sg(s) - L_t}{1 - N_t} \right] = 0.$$

The notations are defined in a similar way as the equilibrium first order conditions, i.e.,

$$W_{1t+1}^S \triangleq \frac{\partial}{\partial k_{t+1}} W^S(k_{t+1}, n_{t+1}, l_{t+1}, \Gamma_{t+1}),$$

$$W_{2t+1}^S \triangleq \frac{\partial}{\partial n_{t+1}} W^S(k_{t+1}, n_{t+1}, l_{t+1}, \Gamma_{t+1}),$$

$$W_{3t+1}^S \triangleq \frac{\partial}{\partial l_{t+1}} W^S(k_{t+1}, n_{t+1}, l_{t+1}, \Gamma_{t+1}).$$

The envelope conditions are

$$W_{1t}^S = \lambda_{1t} \left[MPK_t + e^{-\mu}(1-\delta) \right],$$

$$W_{2t}^S = -Bn_t^{\phi} + (1-\psi)\lambda_{2t} - m_0 v_t^{1-\lambda}\lambda(1-n_t)^{\lambda-1} \left[\lambda_{2t} + \lambda_{3t} \frac{\sum_{s \in S} sg(s) - L_t}{1 - N_t} \right],$$

$$W_{3t}^S = \lambda_{1t} MPL_t + \lambda_{3t}(1-\psi).$$

We can simplify W_{2t}^S by the first order condition with respect to v_t, which implies

$$\lambda_{2t} + \lambda_{3t} \frac{sg(s) - L_t}{1 - N_t} = \frac{1}{m_0(1-\lambda)} \frac{a}{c_t} \frac{v_t^{\lambda}}{(1-n_t)^{\lambda}},$$

then

$$W_{2t}^S = -Bn_t^{\phi} + (1-\psi)\lambda_{2t} - \frac{\lambda}{1-\lambda} \frac{a}{c_t} \frac{v_t}{1-n_t}.$$

In the first subperiod, the constrained social planner chooses $n_t(s) \le \tilde{n}_t(s)$ in order to

$$\max \quad W^S(k_t, \sum_{s \in S} n_t(s), \sum_{s \in S} sn_t(s), \Gamma_t).$$

The partial derivative of above objective function with respect to $n_t(s)$ is

$$W_{2t}^S + W_{3t}^S s = -Bn_t^\phi - \frac{\lambda}{1-\lambda}\frac{a}{c_t}\frac{v_t}{1-n_t} + \frac{MPL_t}{c_t}s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s].$$

Equilibrium v.s. Constrained Social Planning Problem

Our goal in this section is to show that the following wage schedule

$$w_t(s) = \lambda\left[MPL_t s + \frac{av_t}{1-n_t}\right] + (1-\lambda)Bn_t^\phi c_t \tag{2.27}$$

decentralizes the constrained social planning solution, and is consistent with the Nash bargaining if the workers are assigned bargaining power λ.

Start with the constrained social planning solution (as indicated by the asterisk sign) and the multipliers λ_{2t} and λ_{3t}. Construct the equilibrium multipliers as follows:

$$\mu_{2t}(s) = \lambda[\lambda_{2t} + \lambda_{3t}s], \quad \frac{\gamma_t(s)}{c_t^*} = (1-\lambda)[\lambda_{2t} + \lambda_{3t}s].$$

Since

$$p_t = m_0 v_t^{*1-\lambda}(1-n_t^*)^{\lambda-1}, \quad q_t = m_0 v_t^{*-\lambda}(1-n_t^*)^\lambda,$$

respectively,

$$p_t \sum_{s \in S} \mu_{2t}(s)\frac{sg(s) - N_t^*(s)}{1-N_t^*} = \frac{\lambda}{1-\lambda}\frac{a}{c_t^*}\frac{v_t^*}{1-n_t^*},$$

$$q_t \sum_{s \in S} \gamma_t(s)\frac{sg(s) - N_t^*(s)}{1-N_t^*} = a.$$

The second equality indicates that the social planner's choice of v_t^* is optimal from the perspective of the representative firm. From the first equality, we can derive the envelope condition of the household

$$W_{2t}^H(s) = -Bn_t^{*\phi} - \frac{\lambda}{1-\lambda}\frac{a}{c_t^*}\frac{v_t^*}{1-n_t^*} + \frac{w_t(s)}{c_t^*} + (1-\psi)\lambda[\lambda_{2t} + \lambda_{3t}s],$$

as well as the total surplus

$$TS_t(s) = -Bn_t^{*\phi} - \frac{\lambda}{1-\lambda}\frac{a}{c_t^*}\frac{v_t^*}{1-n_t^*} + \frac{MPL_t^*}{c_t^*}s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s].$$

Since the above expression of the total surplus coincides with $W_{2t}^S + W_{3t}^S s$, the constrained social planner's decision over $n_t(s)$ is also optimal to the households and the firms in the equilibrium. Next we will show that allocations implied by the Euler equations of the constrained social planner are also consistent with the Euler equations for the households and the firms.

Since in equilibrium, the representative firm will equate its marginal product of capital with the market rental rate of capital,

$$MPK_t^* = r_t,$$

the representative household's Euler equation with respect to k_{t+1} holds if the corresponding Euler equation for the social planner is satisfied as an equality.

Combining the constrained social planner's Euler equations for n_{t+1} and l_{t+1}, we obtain

$$\lambda_{2t} + \lambda_{3t}s = \beta\mathbb{E}_t\left[-Bn_{t+1}^{*\phi} - \frac{\lambda}{1-\lambda}\frac{a}{c_{t+1}^*}\frac{v_{t+1}^*}{1-n_{t+1}^*} + \frac{MPL_{t+1}^*}{c_{t+1}^*}s + (1-\psi)(\lambda_{2t} + \lambda_{3t}s)\right].$$

Given the wage (2.27), the household's first order condition with respect to $n_{t+1}(s)$ evaluated at the social planning solution becomes

$$\beta\mathbb{E}_t\left[W_{2t+1}^H(s)\right]$$

$$= \beta\mathbb{E}_t\left[-Bn_{t+1}^{*\phi} - \frac{\lambda}{1-\lambda}\frac{a}{c_{t+1}^*}\frac{v_{t+1}^*}{1-n_{t+1}^*} + (1-\psi)\lambda(\lambda_{2t+1} + \lambda_{3t+1}s)\right.$$

$$\left. + \lambda\left\{\frac{MPL_{t+1}^*}{c_{t+1}^*}s + \frac{av_{t+1}^*}{c_{t+1}^*(1-n_{t+1}^*)}\right\} + (1-\lambda)Bn_{t+1}^{*\phi}\right]$$

$$= \lambda\beta\mathbb{E}_t\left[-Bn_{t+1}^{*\phi} - \frac{\lambda}{1-\lambda}\frac{a}{c_{t+1}^*}\frac{v_{t+1}^*}{1-n_{t+1}^*} + \frac{MPL_{t+1}^*}{c_{t+1}^*} + (1-\psi)(\lambda_{2t+1} + \lambda_{3t+1}s)\right]$$

$$= \lambda(\lambda_{2t} + \lambda_{3t}s)$$

$$= \mu_{2t}(s).$$

Thus, the social planning solution is consistent with the household's Euler equation for next period's skill s employment.

Similarly, the firm's first order condition with respect to $n_{t+1}(s)$ at the constrained optimum implies

$$
\begin{aligned}
\mathbb{E}_t & \left[\beta \frac{c_t^*}{c_{t+1}^*} W_{1t+1}^F(s) \right] \\
= & c_t^* \beta \mathbb{E}_t \left[\frac{MPL_{t+1}^*}{c_{t+1}^*} s + (1-\psi)(1-\lambda)(\lambda_{2t+1} + \lambda_{3t+1} s) \right. \\
& \left. -\lambda \left\{ \frac{MPL_{t+1}^*}{c_{t+1}^*} s + \frac{a v_{t+1}^*}{c_{t+1}^*(1 - n_{t+1}^*)} \right\} - (1-\lambda) B n_{t+1}^{*\phi} \right] \\
= & c_t^* (1-\lambda) \beta \mathbb{E}_t \left[-B n_{t+1}^{*\phi} - \frac{\lambda}{1-\lambda} \frac{a}{c_{t+1}^*} \frac{v_{t+1}^*}{1 - n_{t+1}^*} + \frac{MPL_{t+1}^*}{c_{t+1}^*} s + (1-\psi)(\lambda_{2t+1} + \lambda_{3t+1} s) \right] \\
= & c_t^* (1-\lambda)(\lambda_{2t} + \lambda_{3t} s) \\
= & \gamma_t(s).
\end{aligned}
$$

The constrained social planning solution also satisfies the firm's Euler equation for $n_{t+1}(s)$.

Above derivation also shows that the wage (2.27) splits the total surplus so that

$$
W_{2t}^H(s) = \lambda TS_t(s), \quad \frac{W_{1t}^F(s)}{c_t^*} = (1-\lambda) TS_t(s),
$$

thus, it is indeed resulted from Nash bargaining when the workers are given bargaining power λ.

2.6.4 Appendix D: A Pareto Improvement Relative to the Equilibrium Allocation

Suppose the social planner decides to hire an additional ϵ skill s workers relative to the equilibrium employment level $n_t(s)$ in the first subperiod of time t, and then let the households and the firms behave as in equilibrium from time $t+1$ on, i.e., the social planner changes and only changes the total mass of employed workers and their efficient units of labor to

$$
\hat{n}_t = \sum_{s'} n_t(s') + \epsilon = n_t + \epsilon, \quad \hat{l}_t = \sum_{s'} s' n_t(s') + s\epsilon = l_t + s\epsilon.
$$

Let SV be the social value function with respect to $\hat{n}_t$ and $\hat{l}_t$ specified above,

$$SV(\epsilon, s, k_t, \{n_t(s')\}, \Gamma_t) = W^S(k_t, n_t + \epsilon, l_t + s\epsilon, \Gamma_t).$$

Then it can be shown that

$$\lim_{\epsilon \to 0} \frac{\partial SV}{\partial \epsilon} = -Bn_t^\phi - \frac{\lambda}{1-\lambda} \frac{a}{c_t} \frac{v_t}{1-n_t} + \frac{MPL_t}{c_t} s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s]$$
$$+ \lambda_{3t} m_0 v_t^{1-\lambda}(1-n_t)^{\lambda-1} \left[\frac{\sum_{s' \in S} s'g(s') - L_t}{1 - N_t} - s \right]$$
$$> -Bn_t^\phi - \frac{\lambda}{1-\lambda} \frac{a}{c_t} \frac{v_t}{1-n_t} + \frac{MPL_t}{c_t} s + (1-\psi)[\lambda_{2t} + \lambda_{3t}s] = TS_t(s),$$

if

$$s < \bar{s}_t \triangleq \frac{\sum_{s' \in S} s'g(s') - L_t}{1 - N_t}.$$

Since $\{n_t(s)\}$ is the equilibrium employment at time t, $TS_t(s) \geq 0$ for all s such that $n_t(s) > 0$. Therefore, if in the equilibrium, $n_t(\bar{s}_t) > 0$, we can find $s < \bar{s}_t$ so that the social planner can enlarge the economy-wide welfare by hiring more skill s workers at the margin. This constitutes a Pareto improvement because the social planner can give all the additional values generated by this change to these skill s workers while keeping all the other workers' payoffs as in the equilibrium.

The last thing in need of justification is that $n_t(\bar{s}_t)$ is indeed positive. Unfortunately, we fail to establish this statement analytically because it lends too much dependence on the shape of the skill distribution. However, when we calibrate our model to the data, whose details are contained in Section IV, the steady state $n(\bar{s})$ is in fact above zero and it is robust to small changes of other parameter values.

References

[1] David Andolfatto, "Business Cycles and Labor-Market Search", *The American Economic Review*, Vol. 86, No. 1 (Mar., 1996), 112-132.

[2] Almut Balleer, "New Evidence, Old Puzzles: Technology Shocks and Labor Market Dynamics", working paper, October 2009.

[3] John M. Barron, John Bishop, and William C. Dunkelberg, "Employer Search: The Interviewing and Hiring of New Employees", *The Review of Economics and Statistics*, Vol. 67, No. 1 (Feb., 1985), pp. 43-52.

[4] Almut Balleer, and Thijs van Rens, "Skill-Biased Technological Change and the Business Cycle", working paper, May 2011.

[5] Andri Chassamboulli, "Job Competition Over the Business Cycle", working paper, June 20, 2007.

[6] Michael Elsby, Bart Hobijn, Ayşegűl Şahin, and Robert G. Valletta, "The Labor Market in the Great Recession", working paper, April 15th 2010.

[7] Marcus Hagedorn, and Iourh Manovskii, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited", *American Economic Review* 2008, 98:4, 1692-1706.

[8] Per Krusell, and Anthony A. Smith, Jr., "Income and Wealth Heterogeneity in the Macroeconomy", *Journal of Political Economy*, 1998, vol. 106, no. 5.

[9] Finn E. Kydland, "Labor-Force Heterogeneity and the Business Cycle", *Carnegie-Rochester Conference series on Public Policy*, 21 (1984) 173-208.

[10] Finn E. Kydland, and Edward Prescott, "Time to Build and Aggregate Fluctuations", *Econometrica*, Vol. 50, No. 6 (Nov., 1982), pp. 1345-1370.

[11] Monika Merz, "Search in the labor market and the real business cycle", *Journal of Monetary Economics*, 36 (1995) 269-300.

[12] Monika Merz, "Heterogeneous job-matches and the cyclical behavior of labor turnover", *Journal of Monetary Economics*, 43 (1999) 91-124.

[13] Dale T. Mortensen, and Christopher A. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment", *The Review of Economic Studies*, Vol. 61, No. 3 (Jul., 1994), pp. 397-415.

[14] Barbara Petrongolo, and Christopher A. Pissarides, "Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, Vol. 39, No. 2 (Jun., 2001), pp. 390-431.

[15] Eswar Prasad, "Skill Heterogeneity and the Business Cycle", *The Canadian Journal of Economics*, Vol. 29, No. 4 (Nov., 1996), pp. 910-929.

[16] Michael J. Pries, "Worker heterogeneity and labor market volatility in matching models", *Review of Economic Dynamics*, 11 (2008) 664-678.

[17] Federico Ravenna, and Carl E. Walsh, "Screening and Labor Market Flows in a Model with Heterogeneous Workers", working paper, May 2012.

[18] Robert Shimer, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies", *The American Economic Review*, March 2005.

[19] Robert Shimer, *Labor Markets and Business Cycles*, Princeton University Express, 2010.

[20] Benjamín Villena-Roldán, "Aggregate Implications of Employer Search and Recruiting Selection", working paper, April 2010.

[21] Jan van Ours, and Geert Ridder, "Vacancies of the Recruitment of New Employees", *Journal of Labor Economics*, Vol. 10, No. 2(Apr., 1992), pp. 138-155.

CHAPTER 3

Household Production and Labor Wedge

3.1 Introduction

The labor wedge is defined as the gap between the marginal product of labor (MPL) and the marginal rate of substitution between consumption and leisure (MRS) in the data:

$$\text{MPL}(1 - \tau) = \text{MRS}, \tag{3.1}$$

where τ is the labor wedge.

Ever since Chari, Kehoe and McGrattan propose the business cycle accounting method, i.e., the method of analyzing data by looking at the implied deviations of the first order conditions from those of a neo-classical model, the labor wedge has drawn much attention from the macroeconomists because it, along with the efficiency wedge (the Solow residual), accounts for almost all fluctuations during the Great Depression and 1982 recession in the U.S. economy (Chari, Kehoe and McGrattan 2007, [2]). In his 2010 book [9], Shimer documents that the observed labor wedge is volatile and co-moves negatively with the market hours worked, which is in contrast to the predictions of both the neo-classical growth model and the search model that are usually used to study the labor market by macroeconomists.

There have been several explanations of the labor wedge in the literature. To list some of them, the labor wedge has been interpreted as the distortionary taxes on the labor income, the search and matching frictions of the labor market, or the time-varying bargaining power of the workers due to unionization, etc. All these explanations treat the labor wedge as a market distortion that prevents the agents in the model from optimally allocating their time endowment between leisure and work. However, the labor wedge is not necessarily a sign of distortion. Literally speaking, it says that the time spent

working for the firms has a higher marginal value than if it is spent on leisure, but the latter may not be the opportunity cost of the market hours for the households. In reality, an alternative use of time for the household members could be to work at home, and they also receive utilities from consuming the goods and services they produce at home. This paper explores this alternative in detail and shows that it is able to account for a large part of the fluctuations for the observed labor wedge.[1]

The idea is intuitive. What a household production sector adds to the model is an additional option towards which the households can allocate their time and an additional source from which the households can get consumption goods. If the labor wedge is calculated using only the market hours and the market consumption, the marginal utility of leisure will be underestimated because the working hours of the households are underestimated. On the other hand, since the consumption of the home-produced goods is not included in the analysis, the marginal utility of consumption would probably be exaggerated. The two factors jointly yield an underestimated MRS which gives rise to a non-zero misspecified labor wedge even though the economy is absent of any distortions with respect to time allocation.

Since expansions are usually times when the market technology is relatively more productive, the households may be more willing to supply labor to the market sector and substitute some home goods by their counterparts in the market. This results in a relatively small measured labor wedge based on the market hours and the market consumption. On the other hand, during recessions when the market technology is less productive, the households would choose to work more at home and consume more home-produced good. In this situation, the increase in the hours worked at home is recorded as an increase in the leisure the households enjoy, while the recorded market consumption is lower than the total consumption of the households. They together translate into a large measured labor wedge during the recessions. In a word, the model with a home production sector generates counter-cyclical labor wedge naturally. Moreover, since the two mis-measurements always push the labor wedge toward the same direction, the model also has the potential to produce sizable volatility in the misspecified labor wedge, almost comparable to what we observe in the data.

---

[1]The relative volatility of the implied labor wedge, relative to the market output, is over 90% of its observed value in the data.

The impact of the household production on economic agents has long been analyzed. Reuben Gronau (1997, [4]) provides a detailed study on the allocation of time among working in the market, working at home and leisure, and concludes that including the household production sector has important implications on the market labor supply. Benhabib, Rogerson and Wright (1991, [1]) find that adding a household production sector to a standard real business cycle model significantly improves its quantitative performance in several aspects, especially the statistical properties of the market hours worked, final output, consumption, investment and labor productivity. The idea expressed in our paper is mostly related to a recent paper of Karabarbounis (2011,[7]) which also interprets the observed labor wedge as reflecting the omitted household production sector. However, Karabarbounis' main goal is to address a few puzzles in the international business cycle theories, and the relationship between the household production sector and the labor wedge is used as a guidance when estimating the unknown parameters in his model. Our paper focuses more on the labor wedge itself, both in the long run and at the business cycle frequency. Moreover, since the household production sector can affect both the marginal utility of consumption and the marginal utility of leisure for the representative household, we also study their individual effects on the misspecified labor wedge.

The rest of the paper is organized as follows. Section 3.2 lays out the theoretical model with a household production sector that is to be used in all the following analysis. Section 3.3 provides a brief summary of the empirical evidence concerning the labor wedge. The calibration and quantitative implications of the model are reported in Section 3.4. A robust check is also included in the same section. Finally, Section 3.5 delivers the concluding remarks. Most technical details and the data details are placed in the appendix.

3.2 Model with A Household Production Sector

The economy is resided by an infinitely living representative household which consumes both a market-produced good and a home-produced good. The market good is produced

by a competitive representative firm using capital (k_m) and labor (l_m) as inputs:

$$y_{mt} = A_{mt} k_{mt}^{\alpha_m} l_{mt}^{1-\alpha_m}, \tag{3.2}$$

where A_{mt} is an aggregate shock that affects the efficiency of producing the market good.

The production of the home good utilizes the same inputs as the market production:

$$y_{nt} = A_{nt} k_{nt}^{\alpha_n} l_{nt}^{1-\alpha_n}. \tag{3.3}$$

However, it can only be done within the household, subject to the efficiency shock A_{nt}.

In addition to household consumption, the market good can also be invested to get new capital usable next period. The home good is neither tradable nor storable, so within the household, its production is equal to its consumption in every period and under every circumstance.

In the following exposition, the market good is taken as the numeraire.

3.2.1 The Firm's Problem

The representative firm is competitive in both the output market and the input markets. Given the wage rate w_t and the rental rate for capital r_t, the firm faces the following profit-maximization problem in period t:

$$\max_{k_{mt}, l_{mt}} \ y_{mt} - r_t k_{mt} - w_t l_{mt}, \tag{3.4}$$

$$s.t. \quad y_{mt} = A_{mt} k_{mt}^{\alpha_m} l_{mt}^{1-\alpha_m}, \tag{3.5}$$

whose first order conditions imply that the inputs are priced by their marginal products:

$$r_t = \alpha_m \frac{y_{mt}}{k_{mt}}, \tag{3.6}$$

$$w_t = (1 - \alpha_m) \frac{y_{mt}}{l_{mt}}. \tag{3.7}$$

3.2.2 The Representative Household's Problem

The representative household is endowed with the instantaneous utility function

$$U(C, H) = \frac{1}{1-\sigma} \left\{ C^{1-\sigma} \left[1 + (\sigma - 1)\frac{\theta\varepsilon}{1+\varepsilon} H^{\frac{1+\varepsilon}{\varepsilon}} \right]^{\sigma} - 1 \right\} \qquad (3.8)$$

where $\varepsilon > 0$ is the Frisch elasticity of labor supply, $\sigma > 0$ affects the household's willingness to substitute between consumption and leisure. The consumption aggregate C is made from the market good c_m and the home good c_n according to

$$C = [ac_m^e + (1-a)c_n^e]^{\frac{1}{e}}, \qquad (3.9)$$

where parameter e determines the substitutability between the market good and the home good. The term H denotes the total hours worked of the household, i.e.,

$$H = h_m + h_n. \qquad (3.10)$$

Physical capital is accumulated by the household and can be moved freely between the market sector and the household sector. In a word, the representative household solves the following life-time utility maximization problem:

$$\max \ \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{1}{1-\sigma} \left\{ C_t^{1-\sigma} \left[1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right]^{\sigma} - 1 \right\} \right],$$

$$s.t. \quad C_t = [ac_{mt}^e + (1-a)c_{nt}^e]^{\frac{1}{e}},$$

$$H_t = h_{mt} + h_{nt}$$

$$c_{nt} = A_{nt}k_{nt}^{\alpha_n} h_{nt}^{1-\alpha_n} \quad \text{multiplier } \beta^t \mu_t,$$

$$c_{mt} + [K_{t+1} - (1-\delta)K_t] = w_t h_{mt} + r_t k_{mt} \quad \text{multiplier } \beta^t \lambda_t$$

where the choice is over consumption (c_{mt}, c_{nt} and C_t), hours worked (h_{mt}, h_{nt} and H_t) and capital stock (k_{mt}, k_{nt} and K_t).

As in the standard real business cycle literature, the representative household optimally allocates its time between working and leisure so that at the margin, the disutility due to less idle time is fully compensated by the utility increase resulted from enjoying a

little more consumption good:

$$-\frac{\partial U_t}{\partial H_t} = \lambda_t w_t. \tag{3.11}$$

The utility gain in equation (3.11) is expressed as if the extra working time is spent in the market sector. It creates no loss of generality because the optimal allocation of time between the market sector and the household sector requires that the representative household is indifferent between the marginal values of these two alternative uses of time:

$$\lambda_t w_t = \mu_t (1 - \alpha_n) \frac{c_{nt}}{h_{nt}}. \tag{3.12}$$

Similarly, the total capital stock next period is determined by the usual Euler equation

$$\lambda_t = \beta \mathbb{E}_t [\lambda_{t+1}(r_{t+1} + 1 - \delta)] \tag{3.13}$$

where the future value of the capital stock is measured by its return in the market sector. At the optimum, another indifference condition equates the return of the capital in the market sector to that in the household sector

$$\lambda_t r_t = \mu_t \alpha_n \frac{c_{nt}}{k_{nt}}. \tag{3.14}$$

3.2.3 The Market Clearing Conditions

There are three markets in this economy: the market for the market good, the market for labor and the rental market for the capital service. Therefore, the following three market clearing conditions should be satisfied in the equilibrium:

$$c_{mt} + [K_{t+1} - (1 - \delta)K_t] = y_{mt}, \tag{3.15}$$

$$l_{mt} = h_{mt}, \tag{3.16}$$

$$k_{mt} = K_t - k_{nt}. \tag{3.17}$$

The equilibrium in the economy is a set of decision rules for the representative household $\{c_{mt}, c_{nt}, C_t, h_{mt}, h_{nt}, H_t, k_{mt}, k_{nt}, K_t\}$, a set of policy functions for the firm $\{l_{mt}, k_{mt}\}$, a set of prices $\{w_t, r_t\}$ and a set of perceived laws of motion for the aggregate productivity

shocks $\{A_{mt}, A_{nt}\}$ so that (1) the household maximizes its life-time utility; (2) the firm maximizes its profits every period and (3) the markets for the market good, labor and capital service all clear.

3.3   The Labor Wedge

3.3.1   Empirical Evidence

As in Shimer (2010,[9]), the empirical labor wedge is calculated from a one-sector real business cycle model without any consideration of the household production:

$$\hat{\tau}_t = 1 - \frac{\sigma\theta}{1-\alpha}\frac{\hat{C}_t}{\hat{Y}_t}\hat{H}_t^{\frac{1+\varepsilon}{\varepsilon}}\left\{1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon}\hat{H}_t^{\frac{1+\varepsilon}{\varepsilon}}\right\}^{-1} \tag{3.18}$$

where α is the capital income share in the market sector (i.e., α_m in our model). Given the preference of the representative household, the empirical labor wedge $\hat{\tau}_t$ can be recovered from the observed market hours worked $\hat{H}_t$ and the empirical consumption-output ratio $\hat{C}_t/\hat{Y}_t$ in the market sector, where the hat variables denote deviations from the hp-trend with smoothing parameter 1600.

Table 3.1: Empirical Moments

	Y	C	H	C/Y	τ
std.(%)	1.55	1.18	1.39	1.00	1.19
corr. Y	1.00	0.81	0.85	-0.40	-0.54
C	——	1.00	0.70	-0.02	-0.65
H	——	——	1.00	-0.28	-0.61
C/Y	——	——	——	1.00	-0.08
τ	——	——	——	——	1.00

Table 3.1 contains the empirical moments we obtain where the labor wedge is calculated using the parameter values provided in table 3.2 in Section 3.4. Two features can be seen that are consistent with Shimer's observations. First, the empirical labor wedge is relatively volatile: the volatility of τ is over 3 quarters of that of the real GDP. Second, the empirical labor wedge is strongly negatively correlated with the market hours worked (correlation -0.61). Both observations raise challenge to the standard productivity driven real business cycle models. However, as will be shown in the next section, the RBC
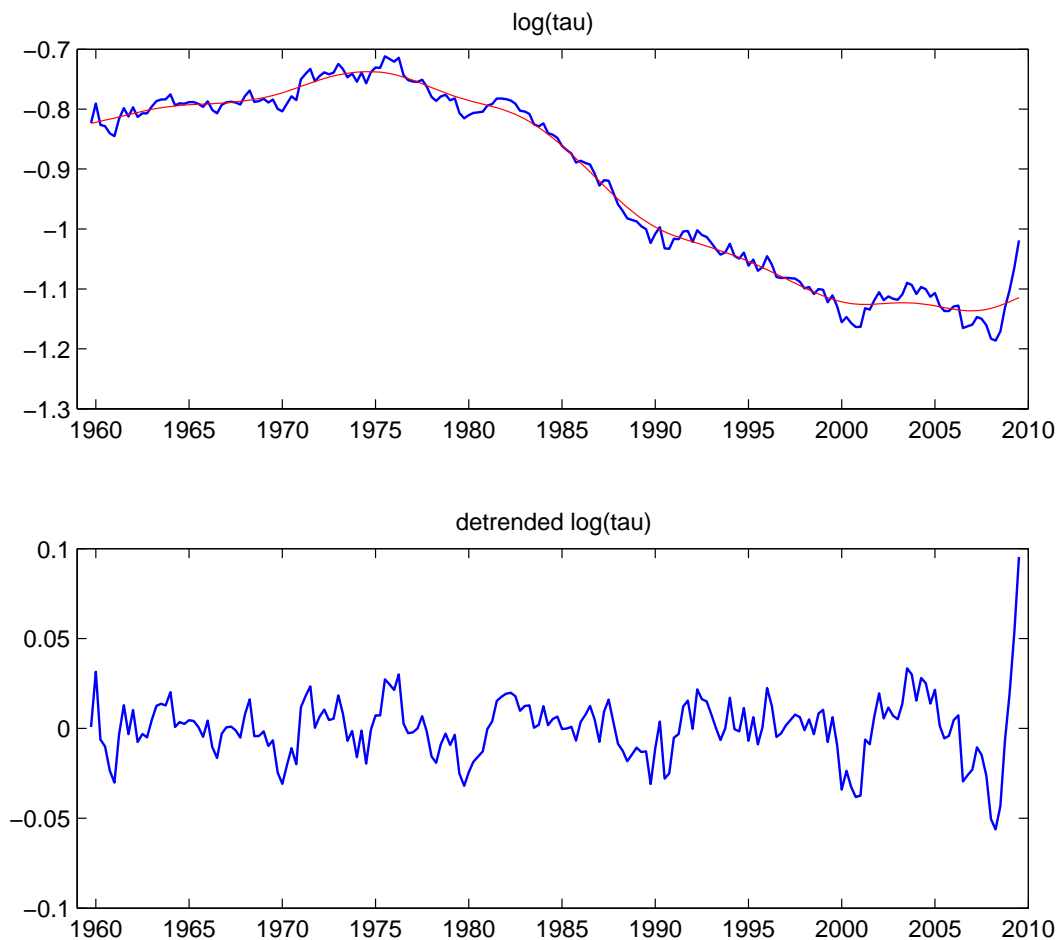
Figure 3.1: Empirical Labor Wedge

The empirical labor wedge (1959Q4-2009Q3). Parameters $\varepsilon = 3$, $\sigma = 3$, $\alpha = 0.36$. θ is chosen so that the average labor wedge is around 0.4. The red-dotted line in the upper panel is the HP-filtered trend with the smoothing parameter 1600 (quarterly data).

framework enriched by a household production sector is able to do the work.

Before starting the quantitative analysis, we will end this section by introducing the labor wedge in our model and explicitly specify how it would be calculated.

3.3.2 The Model Implied Labor Wedge

By definition, the labor wedge is the gap between the marginal product of labor in the market sector and the household's marginal rate of substitution between consumption and leisure

$$MPL_t(1 - \tau_t) = MRS_t.$$

Since there is no distortion in the firm's problem, the optimal labor demand is chosen so that the marginal product of labor is equal to the market wage (equation (3.7)):

$$MPL_t = (1 - \alpha_m) \frac{y_{mt}}{h_{mt}} = w_t.$$

On the other hand,

$$MRS_t = \frac{-\partial U_t/\partial h_{mt}}{\partial U_t/\partial c_{mt}} = \frac{C_t^e}{ac_{mt}^{e-1}} \sigma \theta H_t^{\frac{1}{\varepsilon}} \left\{ 1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} = w_t$$

according to equation (3.11). This implies that the correctly measured labor wedge

$$\tau_t^0 = 1 - \frac{C_t^e}{ac_{mt}^{e-1}} \frac{\sigma \theta}{1 - \alpha_m} \frac{h_{mt}}{y_{mt}} H_t^{\frac{1}{\varepsilon}} \left\{ 1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \qquad (3.19)$$

is constantly equal to zero, and there is no mis-allocation of time in this economy if the household production sector is fully taken into account.

However, information of the household production sector may not be as easy to collect and study as that of the market sector. As a result, it is possible that data on the hours worked at home h_{nt} and the home-produced consumption c_{nt} are measured with large errors or even unavailable in certain cases. To highlight the problem generated by our limited ability of acquiring information inside the household, we make the extreme assumption that the whole household production sector is ignored by the researchers and see how much we can get out of the misspecified labor wedge. A robust check with other cases is provided at the end of Section IV.

This assumption modifies the instantaneous utility function to be

$$U(c_m, h_m) = \frac{1}{1 - \sigma} \left\{ c_m^{1-\sigma} \left[1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} h_m^{\frac{1+\varepsilon}{\varepsilon}} \right]^{\sigma} - 1 \right\}. \qquad (3.20)$$

The corresponding marginal rate of substitution between consumption and leisure is

$$MRS_t = \sigma \theta \, c_{mt} \, h_{mt}^{\frac{1}{\varepsilon}} \left[1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} h_{mt}^{\frac{1+\varepsilon}{\varepsilon}} \right]^{-1}. \qquad (3.21)$$

Let τ_t^1 denote the implied labor wedge in this case,

$$\tau_t^1 = 1 - \frac{\sigma\theta}{1 - \alpha_m}\frac{c_{mt}}{y_{mt}}\ h_{mt}^{\frac{1+\varepsilon}{\varepsilon}}\left[1 + (\sigma - 1)\frac{\theta\varepsilon}{1 + \varepsilon}h_{mt}^{\frac{1+\varepsilon}{\varepsilon}}\right]^{-1}. \tag{3.22}$$

In the next section, it will be shown that quantitatively, such mis-measured labor wedge τ_t^1 to a large extent well resembles the dynamic properties of the empirical labor wedge discussed in the first subsection.

3.4 The Quantitative Properties of Labor Wedge

3.4.1 Calibration

The period in the model is interpreted as one quarter in the real world.

Table 3.2 contains the calibrated values for all parameters used in our model. Concerning the technology, the capital share in the market production function α_m is set to 0.36 as in many real business cycle models. Following Benhabib, Rogerson and Wright (1991, [1]), we choose α_n, the capital share in the household production function, so that the steady-state c_n/Y_m is around 0.26. The productivity process for the market technology A_m is modeled as a log-AR(1) process with the usual persistence parameter $\rho_A = 0.95$. The volatility σ_A is chosen so that the model generated output series has the same standard deviation as the observed output series. The household productivity A_n is assumed to be governed by an identical process and the correlation between the two productivity processes is set to 2/3, which also follows Benhabib, Rogerson and Wright (1991, [1]).

The representative household combines the market good and the home-produced good according to the aggregator:

$$C = [ac_m^e + (1 - a)c_n^e]^{\frac{1}{e}}$$

where parameter e determines the substitutability between the two goods. We adopt the same value for e (=0.8) as in Benhabib, Rogerson and Wright (1991, [1]). A robust check is conducted later in this section to see how changes in this parameter affect the labor

Table 3.2: Calibration

Production functions		
α_m	0.36	capital share in the market production function
α_n	0.08	capital share in the household production function
Household preference		
a	0.35	share of market good in consumption composite
β	0.99	discount factor
e	0.8	elasticity of substitution between home and market goods is 5
σ	2	determines the substitution between consumption and leisure
θ	1.33	disutility from working
ε	4	Frisch elasticity of labor supply
Depreciation		
δ	0.025	capital depreciation rate
Shocks		
ρ_A	0.95	persistence of both market and home productivity shocks
σ_A	0.0057	volatility of the innovations to productivity shocks
γ	2/3	correlation between productivity shocks

wedge τ^1 in our model.

The share of the market good a is chosen together with the disutility parameter θ so that at the steady state, the representative household spends approximately 33% of its time working in the market, and 28% of the time working at home. The parameter σ determines the household's willingness of substitution between consumption and leisure. Shimer argues in his 2010 book ([9]) that values for σ greater than 2 would not be reasonable because it implies too large a consumption gap between employed workers and unemployed agents. In our work, we take σ as 2. Given σ, we consider a value of 4 for the Frisch elasticity of substitution ε in order to get a steady-state labor wedge τ_1 of around 40%.[2]

Lastly, the household discount factor β is 0.99 and the capital depreciation rate δ is 0.025. Both are widely used in the macroeconomic studies.

3.4.2 The Quantitative Properties of Labor Wedge

It has been mentioned in section 3.3 that the correctly measured labor wedge τ^0 carries a constant value of zero in our model, indicating that no inefficiency associated with

---

[2]Robust checks are performed for other reasonable values of these parameters.

the time allocation presents in the economy. However, if the household sector is partially ignored, the representative household will appear to enjoy more leisure and less consumption than it actually does, both of which tend to reduce the marginal rate of substitution between consumption and leisure for the representative household, and thus result in a seemingly non-trivial distortionary term in the household's optimal labor supply condition.

Table 3.3: Model Predicted Moments

	y_m	c_m	h_m	k_m	c_m/y_m	τ_1
std.(%)	1.54	0.91	1.21	0.45	1.10	1.12
corr. y_m	1.00	0.71	0.99	0.41	-0.81	-0.33
c_m	——	1.00	0.79	0.72	-0.16	-0.89
h_m	——	——	1.00	0.40	-0.73	-0.46
k_m	——	——	——	1.00	0.02	-0.60
c_m/y_m	——	——	——	——	1.00	-0.27
τ_1	——	——	——	——	——	1.00

The steady state mis-measured labor wedge τ^1 is approximately 0.4, which is our target in the calibration. Its dynamic properties are displayed in Table 3.3. As can be easily seen, our model generates sizable volatility in τ^1, comparable to its empirical counterpart. The implied correlation between τ^1 and the total market hours worked h_m is -0.46, which also matches qualitatively the observed negative association between the labor wedge and the market hour quite well.

The impulse response functions shown in Figure (3.2) and Figure (3.3) reveal more dynamic properties of the modeling economy. In general, favorable shocks to the market productivity increase resources allocated to the market sector, while positive innovations to the home productivity have the opposite effects. The mis-measured labor wedge τ^1 falls below its steady state level in the first case because, as the representative household works more time in the market sector and consumes more the market good, the ignorance of the home sector becomes less a problem. To the contrary, when it is the time that the home technology is relatively productive, the understatement of the household MRS gets worse since neglecting the household production sector means a greater gap between the true hours worked/consumption and their counterparts in the market sector. As a result, we observe an increase in τ^1 as the home productivity A_n is hit by a positive shock.
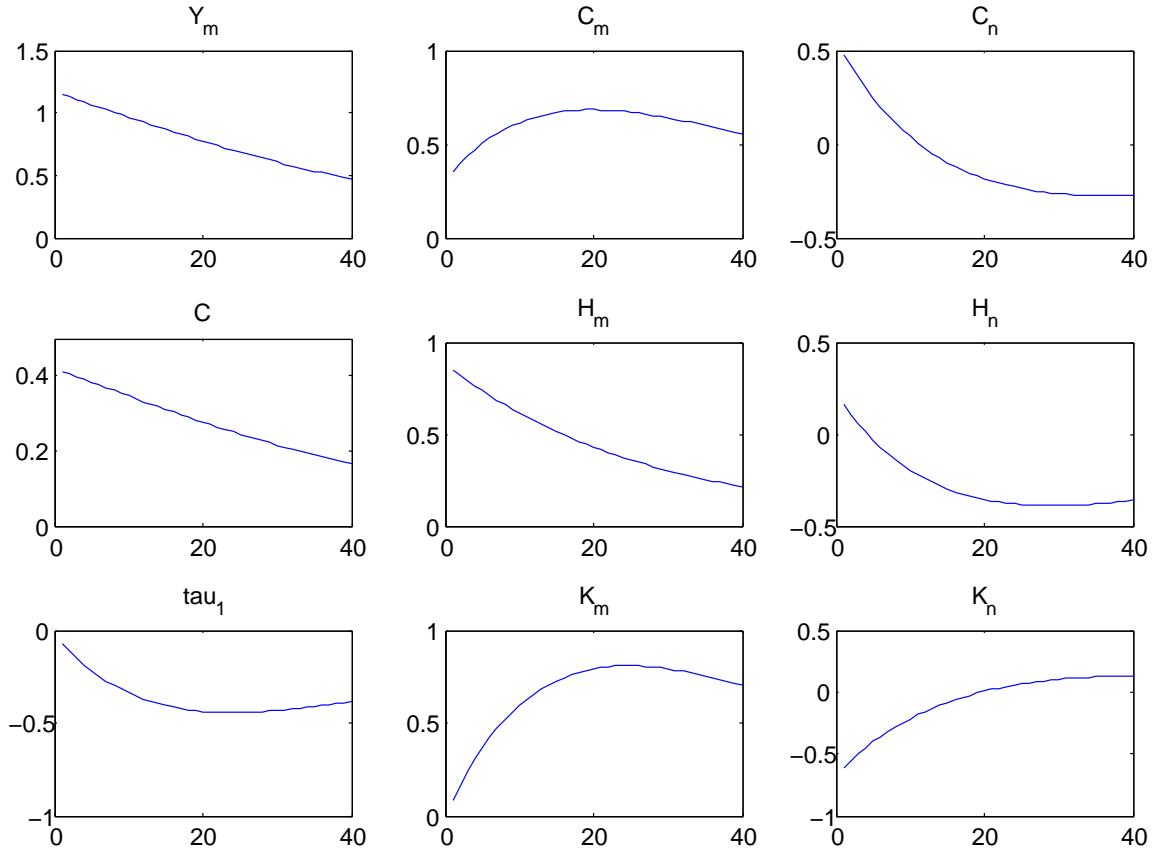
Figure 3.2: Impulse Response: Shocks to Market Productivity
The impulse response functions with respect to a one-standard-deviation positive shock to A_m. All variables are percentage deviations from the steady state.

3.4.3 Robust Checks

This subsection is devoted to the robust checks. We will first analyze the impact of a few key parameters in the household utility function on the dynamic properties of the mis-measured labor wedge τ^1, then we will turn to study how the mis-measured total hours worked and consumption, respectively, contribute to such dynamics.

There are three important parameters in the household's problem, e determining the elasticity of substitution between the market and the home-produced goods, σ governing the elasticity of substitution between the total consumption and leisure, and ε relevant to the Frisch elasticity of labor supply. The robust check is done by fixing any two out of the three, as well as other parameters except share of the market good a and the disutility θ, at their baseline values given in Table 2 and altering the remaining one in a reasonable
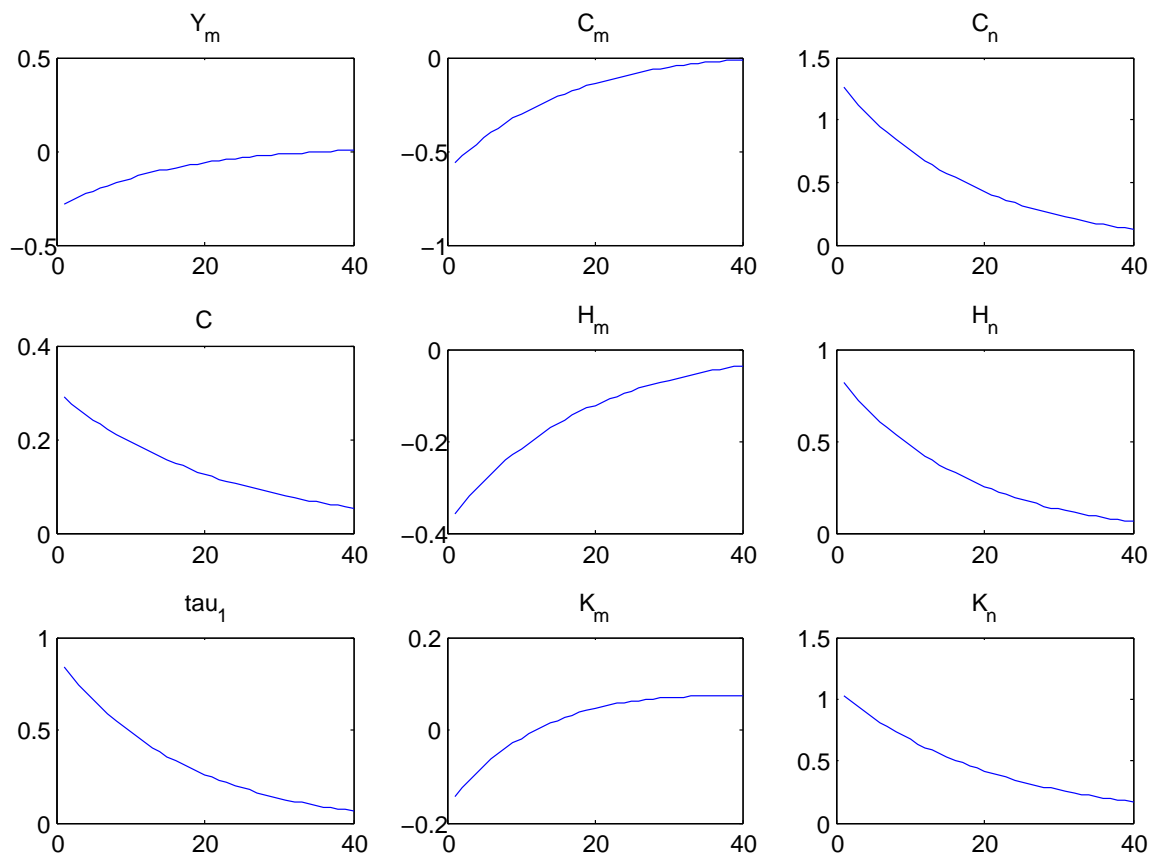
Figure 3.3: Impulse Response: Shocks to Home Productivity
The impulse response functions with respect to a one-standard-deviation positive shock to A_n.
All variables are percentage deviations from the steady state.

range. Each time, a and θ are re-calibrated so that the steady state hours worked in the market and at home are $h_m = 0.33$ and $h_n = 0.28$.

From Table 3.4, it is easy to see that one robust feature of the mis-measured labor wedge τ^1 is its negative association with the market hours worked h_m. In almost all the parameter combinations considered, the simulated correlation coefficient is more negative than the one with baseline calibration, bringing the model's prediction closer to its empirical counterpart.

When it comes to the implied volatility of τ^1, ε has a small effect while σ has a moderate effect. The substitutability between the market good and the home good, described by parameter e, changes substantially the volatility of τ^1, which bears a big value when the two goods are close to be perfect substitutes. It turns out that with the

Table 3.4: Robust Check

ε	std($\cdot$) (%)			corr($h_m, \cdot$)		
	1	2	3	1	2	3
τ_1	0.81	1.00	1.09	-0.87	-0.70	-0.56
σ	0.5	1	1.5	0.5	1	1.5
τ_1	0.59	0.86	1.03	-0.59	-0.54	-0.50
e	0.5	0.7	0.9	0.5	0.7	0.9
τ_1	0.35	0.73	2.10	-0.44	-0.39	-0.61

current calibration, the value of e chosen by Benhabib, Rogerson and Wright (1991, [1]) produces the labor wedge with standard deviation comparable to the data.

So far, we have been focusing on a complete ignorance of the household production sector in order to highlight its role in determining the measured labor wedge dynamics. Next we will turn to studying the cases where only part of the household sector is absent in the calculation of the researchers. To be more specific, we assume that the researchers use the following utility function for the representative household:

$$U(\hat{C}, \hat{H}) = \frac{1}{1 - \sigma} \left\{ \tilde{C}^{1-\sigma} \left[1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} \tilde{H}^{\frac{1+\varepsilon}{\varepsilon}} \right]^{\sigma} - 1 \right\}.$$

where $\tilde{C}$ and $\tilde{H}$ are measured consumption and hours worked. In the specification of τ^1, both are taken as only from the market sector, i.e., $\tilde{C} = c_m$ and $\tilde{H} = h_m$. However, in reality, data for these two variables may come from different sources and thus are not necessarily treating the household sector in a uniform way. Motivated by these thoughts, we consider the following two specifications of the measured labor wedge:

$$\tau_t^2 = 1 - \frac{\sigma \theta c_{mt} H_t^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.23}$$

$$\tau_t^3 = 1 - \frac{\sigma \theta C_t h_{mt}^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma - 1) \frac{\theta \varepsilon}{1 + \varepsilon} h_{mt}^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.24}$$

The idea behind these specifications is that (3.23) looks at the case where the total hours worked are correctly measured $\tilde{H} = H$ while data on consumption only cover the market sector $\tilde{C} = c_m$, while (3.24) studies the opposite case where $\tilde{C} = C$ but $\tilde{H} = h_m$.

The dynamic properties of τ^2 and τ^3 as opposed to those of τ^1 are presented in Table

Table 3.5: Different Specifications of Labor Wedge

	τ_1	τ_2	τ_3
$std(\cdot)$ (%)	1.12	1.03	0.16
$corr(h_m, \cdot)$	-0.46	$-0.32^{\dagger}$	-0.09

†: $corr(H, \tau_2) = 0.43$
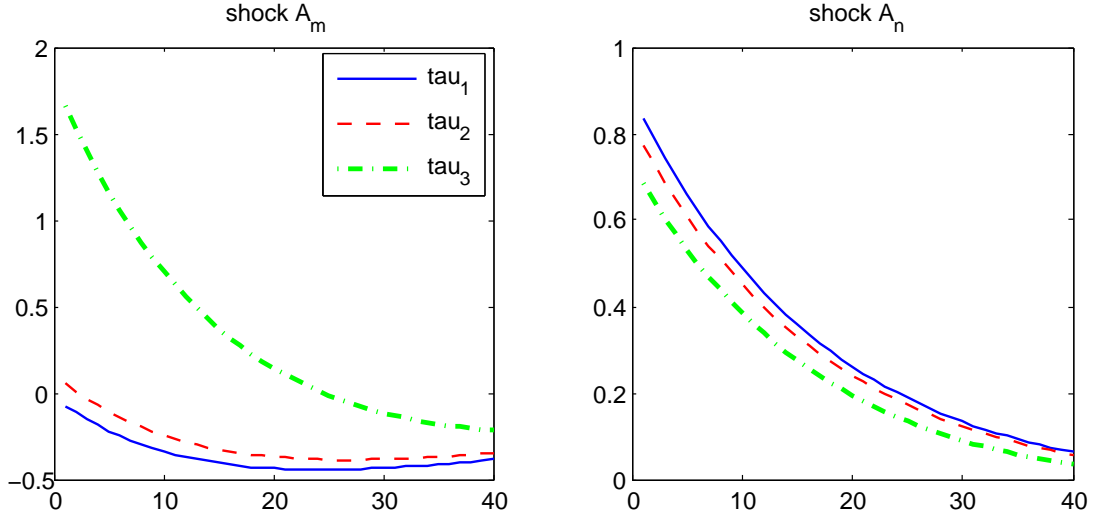
Figure 3.4: Dynamics of τ^1, τ^2 and τ^3

Three mis-specifications of the labor wedge. τ_1 is when the household sector is completely ignored. In τ_2, hours worked are correctly measured while only the consumption of the market good is taken into account. On the other hand, τ_3 considers only hours worked in the market sector but the total consumption is correctly measured. All variables are percentage deviations from the steady state.

3.5 and Figure (3.4). Specification τ^2 resembles τ^1 both qualitatively and quantitatively,[3] though if calculations are done consistently, we would observe a positive association between τ^2 and the measured hours worked H.

The mis-measurement of the total hours worked has relatively small quantitative impact on the implied labor wedge. Volatility of τ^3 is around one-ninth of that of τ^1, and it shows a weakly negative correlation with the market hours h_m. The observation that τ^3 falls below its steady state level when the non-market productivity A_n is temporarily high may seem counter-intuitive (see Figure (3.4)) because h_m declines as a response to

---

[3]The rise of τ^2 above its steady state value right after the positive A_m shock is due to the fact that the initial increase in the market consumption is relatively small since it is a good time to invest in the physical capital. As a result, the measured marginal utility of consumption does not decline quickly enough.

an increase in A_n, and thus the marginal utility of leisure measured from h_m would decline as well, which tends to reduce MRS and raise τ^3. However, the positive innovation to A_n also brings an increase in the total consumption for the representative household, which leads to a drop in the marginal value of consumption. With the current calibration, the net effect of these two forces is that MRS rises and the labor wedge τ^3 falls temporarily. In a word, τ^3 and h_m co-move with each other for a while after shocks to the non-market productivity, and this is what mitigates the negative relationship between the two variables.

To summarize, if the dynamics of the empirical labor wedge does result from the omission of a household production sector, the mis-measurement of the true hours worked has a much less role than the mis-measured total consumption for the representative household.

3.5 Conclusion

In this project, we formally analyze the hypothesis that the observed labor wedge in the data reflects the existence of an under accounted household production sector. We show that after being extended to include the household production, as Benhabib, Rogerson and Wright (1991, [1]), an otherwise standard real business model generates the wedge in the household's optimal labor allocation condition that displays similar business cycle properties as the empirical labor wedge. The results are robust to a large range of parameter values. We also find that, compared to the mis-measurement of the total hours worked, the mis-measured total consumption due to the omission of the household production sector is much more relevant in terms of counting for the dynamic properties of the empirical labor wedge. To sum up, the model with exogenous labor wedge can be well viewed as a reduce form of the full model where both the market sector and the household sector present.

3.6 Appendix: Equilibrium Conditions

[1] The market-good producing firms:

$$y_{mt} = A_{mt}k_{mt}^{\alpha_m}l_{mt}^{1-\alpha_m} \tag{3.25}$$

$$\alpha_m y_{mt} = r_t k_{mt} \tag{3.26}$$

$$(1 - \alpha_m)y_{mt} = w_t l_{mt} \tag{3.27}$$

[2] The representative household:

[2.1] The marginal utilities of consumption and leisure:

$$\frac{\partial U_t}{\partial C_t} = C_t^{-\sigma} \left\{ 1 + (\sigma - 1)\frac{\theta\varepsilon}{1+\varepsilon}H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{\sigma} \tag{3.28}$$

$$-\frac{\partial U_t}{\partial H_t} = \sigma\theta C_t^{1-\sigma} \left\{ 1 + (\sigma - 1)\frac{\theta\varepsilon}{1+\varepsilon}H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{\sigma-1} H_t^{\frac{1}{\varepsilon}} \tag{3.29}$$

[2.2] The marginal utility of the market consumption λ and the marginal utility of the non-market consumption μ:

$$\lambda_t = \frac{ac_{mt}^{e-1}}{C_t^{e-1}}\frac{\partial U_t}{\partial C_t} \tag{3.30}$$

$$\mu_t = \frac{(1-a)c_{nt}^{e-1}}{C_t^{e-1}}\frac{\partial U_t}{\partial C_t} \tag{3.31}$$

[2.3] The first order conditions:

$$-\frac{\partial U_t}{\partial H_t} = \lambda_t w_t \tag{3.32}$$

$$\lambda_t w_t = \mu_t(1 - \alpha_n)\frac{c_{nt}}{h_{nt}} \tag{3.33}$$

$$\lambda_t = \beta\mathbb{E}_t[\lambda_{t+1}(r_{t+1} + 1 - \delta)] \tag{3.34}$$

$$\lambda_t r_t = \mu_t \alpha_n \frac{c_{nt}}{k_{nt}} \tag{3.35}$$

[2.4] Production function of the non-market good:

$$c_{nt} = A_{nt}k_{nt}^{\alpha_n}h_{nt}^{1-\alpha_n} \tag{3.36}$$

[2.5] Total consumption and hours worked:

$$C_t^e = ac_{mt}^e + (1-a)c_{nt}^e \tag{3.37}$$

$$H_t = h_{mt} + h_{nt} \tag{3.38}$$

[3] The market clearing conditions:

$$y_{mt} = c_{mt} + K_{t+1} - (1-\delta)K_t \tag{3.39}$$

$$h_{mt} = l_{mt} \tag{3.40}$$

$$K_t - k_{nt} = k_{mt} \tag{3.41}$$

[4] Laws of motion for the aggregate productivity shocks:

$$\log A_{mt} = \rho_A \log A_{mt-1} + u_{mt} \tag{3.42}$$

$$\log A_{nt} = \rho_A \log A_{nt-1} + u_{nt} \tag{3.43}$$

with

$$\begin{bmatrix} u_{mt} \\ u_{nt} \end{bmatrix} \sim^{\text{i.i.d}} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_A^2 \end{bmatrix} \right) \tag{3.44}$$

[5] Specifications of the labor wedge:

$$\tau_t^0 = 1 - \frac{C_t^e}{ac_{mt}^{e-1}} \frac{\sigma\theta H_t^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.45}$$

$$\tau_t^1 = 1 - \frac{\sigma\theta c_{mt} h_{mt}^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon} h_{mt}^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.46}$$

$$\tau_t^2 = 1 - \frac{\sigma\theta c_{mt} H_t^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.47}$$

$$\tau_t^3 = 1 - \frac{\sigma\theta C_t h_{mt}^{1/\varepsilon}}{w_t} \left\{ 1 + (\sigma-1)\frac{\theta\varepsilon}{1+\varepsilon} h_{mt}^{\frac{1+\varepsilon}{\varepsilon}} \right\}^{-1} \tag{3.48}$$

REFERENCES

[1] Jess Benhabib, Richard Rogerson, and Randall Wright, "Homework in Macroeconomics: Household Production and Aggregate Fluctuations", *The Journal of Political Economy*, Vol. 99, No. 6(Dec., 1991), pp. 1166-1187.

[2] V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting", *Econometrica*, Vol. 75, No. 3 (May 2007), 781-836.

[3] Simona E. Cociuba, Edward C. Prescott, and Alexander Ueberfeldt, "U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2009-III", working paper, November 2009.

[4] Reuben Gronau, "Leisure, Home Production, and Work–the Theory of the Allocation of Time Revisited", *The Journal of Political Economy*, Vol. 85, No. 6(Dec., 1977), pp1099-1123.

[5] Reuben Gronau, "Home Production–A Survey", *Handbook of Labor Economics*, Volumn 1, Chapter 4, Elsevier Science Publishers BV 1986.

[6] Reuben Gronau, "The Theory of Home Production: The Past Ten Years", *Journal of Labor Economics*, Vol. 15, No. 2(April, 1997), pp.197-205.

[7] Loukas Karabarbounis, "Labor Wedges and Open Economy Puzzles", working paper, June 2011.

[8] Stephen L. Parante, Richard Rogerson, and Randall Wright, "Homework in Development Economics: Household Production and the Wealth of Nations", *The Journal of Political Economy*, Vol. 108, No. 4(August, 2000), pp. 680-687.

[9] Robert Shimer, *Labor Markets and Business Cycles*, Princeton University Press, 2010.