Heterogeneous Costs of Business Cycles with Incomplete Markets*

Toshihiko Mukoyama
Department of Economics
Concordia University
and
CIREQ

Ayşegül Şahin†
Department of Economics
Purdue University

February 12, 2004

Abstract

This paper reconsiders the cost of business cycles under market incompleteness. Primarily, we focus on the heterogeneity in the cost among different groups. In addition to the heterogeneity in asset holdings, this paper considers heterogeneity in earnings and unemployment risk. In particular, we focus on skill heterogeneity. Unskilled workers are subject to a much larger risk of unemployment during recessions than are skilled workers. This serves as an additional source of heterogeneity in the cost of business cycles. This effect is reinforced by the fact that unskilled workers earn less, so that they have less resources for precautionary saving.

Keywords: Cost of Business Cycles; Incomplete Markets; Skill and Unemployment

JEL Classifications: E24, E32; E61

*The discussions with Per Krusell and Tony Smith were very helpful. We thank Sun-Bin Kim for his help in Fortran programming. We also thank Mark Bils, Emanuela Cardia, Rui Castro, Igor Livshits, and seminar participants at Indiana University, University of Illinois at Urbana-Champaign, the Montreal Macro Reading Group, the Rochester Wegmans Conference, the CEA Annual Meetings 2003, and the Society for Economic Dynamics Meetings 2003 for comments and suggestions. We thank Vera Brencic and Roxanne Stanoprud for excellent research assistance. All errors are our own. Mukoyama thanks FQRSC for financial support.

†Corresponding author. *Address:* Purdue University, Krannert School of Management, 403 W. State St. West Lafayette, IN 47907-2056. Tel.: +1-765-494-4419; fax: +1-765-494-9658. *E-mail address:* asahin@mgmt.purdue.edu.

1 Introduction

In everyday discussions of economic policy, it is usually assumed that business cycles are evil and that it is desirable to eliminate them. Many would agree that stabilization is desirable if it comes without cost. However, stabilization policies are often very costly, and it is not obvious whether we should avoid business cycles when the cost of doing so is great. What, in fact, is the cost of having business cycles? How much resource cost can be justified to eliminate business cycles? In an influential study, Lucas (1987) considered these questions. His result was astounding – the cost of having business cycles is almost zero.

Lucas's method is simple. Postulate a standard utility function for a representative agent. Feed a consumption series that resembles the actual aggregate consumption time series into the utility function, and calculate the utility that the representative agent obtains from the consumption stream. Compare this with the utility from a smoothed version of the consumption series (the mean of the original consumption series). How much is the representative agent willing to pay for moving from the fluctuating consumption path to the smooth consumption path? This amount is the measure of the cost of business cycles. Specifically, Lucas calculates λ which satisfies

$$E_0\left[\sum_{t=0}^{\infty}\beta^t U((1+\lambda)c_t^o)\right] = E_0\left[\sum_{t=0}^{\infty}\beta^t U(c_t^s)\right], \tag{1}$$

where $\{c_t^o\}_{t=0}^{\infty}$ is the consumption stream in the original economy (with business cycles), and $\{c_t^s\}_{t=0}^{\infty}$ is the consumption stream in the "smoothed" economy (without business cycles). He finds $\lambda = 0.008\%$ with logarithmic utility.

Some researchers interpreted Lucas's result to indicate that neither the policy nor the research aimed at coping with business cycles is thus relevant. Many other researchers felt that the number Lucas obtained was too small, and challenged Lucas's assumptions. There are three strands of literature that re-examine Lucas's calculation. The first challenges Lucas's assumptions on preferences and the nature of the stochastic process of consumption (e.g. Obstfeld [1994]). If, for example, the consumer's risk aversion is larger, or the stochastic process is not stationary, then the cost of business cycles will be larger. The second asserts that the elimination of business cycles affects the output trend of the economy (e.g. Barlevy [2002]). The third argues that Lucas's use of a representative agent and an aggregate consumption series is not realistic, and considers different model environments. This paper extends this third line of argument.

Lucas's use of a representative agent and an aggregate consumption series are justified in an environment where complete Arrow-Debreu markets exist. Under complete asset markets and a common constant relative risk aversion utility function, the aggregation theorem holds and the consumption series of each individual will parallel the aggregate consumption series. In reality however, it is unlikely that the asset markets are complete.

Recent macroeconomic studies focus increasingly more on an environment where asset markets are not complete. The incompleteness of asset markets can potentially have a large impact on the calculation of the cost of business cycles, since the opportunity for insuring risk is limited in such an environment. İmrohoroğlu (1989) was the first to attempt to measure the welfare cost of business cycles under market incompleteness. She considers an environment where individuals face idiosyncratic unemployment risk, and they cannot insure each other because asset markets are not complete. By comparing environments where aggregate shocks are present and then removed, she obtained a larger cost of business cycles than Lucas did. When the coefficient of relative risk aversion is 1.5, the cost is equivalent to 0.3% of average consumption.

Atkeson and Phelan (1994) re-examined the welfare effect of eliminating business cycles under market incompleteness. When considering the economy without business cycles, instead of removing the aggregate shocks (as İmrohoroğlu did), they removed the correlation of aggregate shocks across individuals. Thus, although aggregate fluctuations are removed, individuals essentially face the same shocks. The resulting cost of business cycles is much

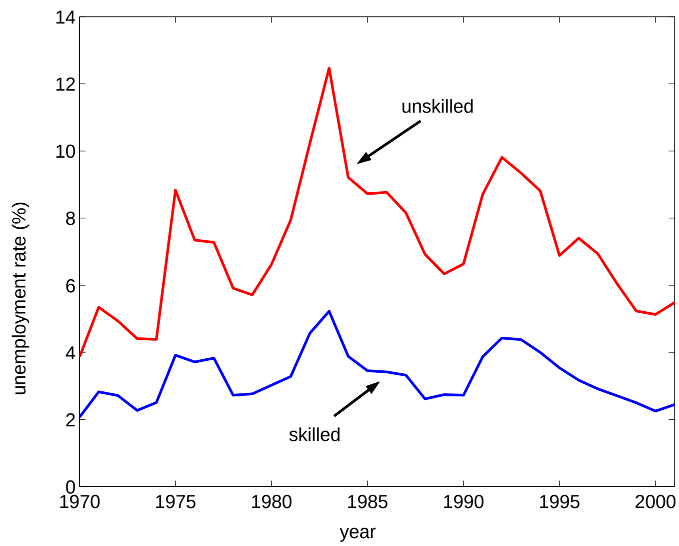

Figure 1: Unemployment Rates by Skill.
Data Source: Current Population Survey

smaller than that in İmrohoroğlu (1989): $\lambda = 0.02\%$ on average, when the coefficient of risk aversion is 1.35.

Krusell and Smith (1999, 2002) analyzed a dynamic general equilibrium model with endogenous prices which matches the U.S. wealth distribution. They proposed a method called the "integration principle" when removing business cycles. They argued that the part of the idiosyncratic shocks which is correlated with the aggregate shocks should be removed when business cycles are eliminated. Their result (Krusell and Smith [2002]) is that $\lambda = 0.09\%$ on average with logarithmic utility.

Perhaps more importantly, Krusell and Smith (1999, 2002) pointed out that the cost of business cycles may differ among different groups of people. For example, agents with larger asset holdings would have a greater opportunity to self-insure against unemployment risk. It can be expected that very poor agents would have a larger cost of business cycles. Krusell and Smith (1999, 2002) reported that there is considerable heterogeneity in the cost of business cycles among agents with different wealth.

We extend Krusell and Smith's analysis to the case where there is an additional source of heterogeneity. Our analysis is motivated by the microeconomic studies of the unemployment process. Figure 1 is drawn from the annual data in the Current Population Survey (1970–2001). It illustrates the unemployment processes for unskilled workers (high school diploma or lower) and for skilled workers (some college or above). The two processes differ dramatically. Unskilled workers are not only subject to higher level of unemployment, but also face a more volatile unemployment process. This implies that unskilled workers are hurt more by recessions.[1] Moreover, unskilled workers earn less income, which limits their ability to self-insure. We examine how this heterogeneity in unemployment risk and income translates into heterogeneity in the cost of business cycles.

This paper is organized as follows. The next section describes our baseline model. In Section 3, we analyze the welfare effects of removing business cycles. Section 4 concludes.

2 Model

2.1 Setup

Our model is a standard Bewley-Aiyagari type dynamic general equilibrium model with incomplete markets (Aiyagari [1994]). In particular, we build upon the model with aggregate shocks developed by Krusell and Smith (1998).

There is a continuum of agents (with measure 1) in the economy. They maximize their discounted utility

$$\mathbf{U} = E_0 \left[\sum_{t=0}^{\infty} \left(\prod_{j=0}^{t} \beta_j \right) \log c_t \right],$$

where c_t is the consumption in period t and $\beta_0 = 1$. We allow the discount factor β_t to differ across agents and to vary over time.

[1] Mincer (1991) documented that unskilled workers are subject to a substantially larger risk of becoming unemployed in recessions than are skilled workers. Topel (1993) shows that the unemployment rate of low-wage men is not only higher, but also much more volatile.

There are two types of agents: skilled ($\eta = s$) and unskilled ($\eta = u$). Each agent's skill status η may change over time (by a stochastic process that is uncorrelated across agents), but the number of skilled (χ_s) and unskilled (χ_u) workers is constant by the law of large numbers. A skilled worker can supply more labor than an unskilled worker. We express this dependence by the function $\phi(\eta)$, where $\phi(s) > \phi(u)$. The value $\phi(s)/\phi(u)$ can be interpreted as the skill premium. We assume that $\phi(\eta)$ is constant over time. Therefore, the skill premium is acyclical.[2]

An agent is either employed ($\epsilon = 1$) or unemployed ($\epsilon = 0$). The employment status is determined by an exogenous random process. When employed, the agent supplies $\phi(\eta)$ units of labor to the market. When unemployed, she engages in household production. Household production utilizes the same technology as the market technology[3], but the agent can supply only a fraction $h < 1$ of her market labor supply (she is less efficient at home than in the market).[4] The probability of becoming unemployed differs between skilled and unskilled agents. The unemployment probability also depends on the aggregate state of the economy.

The market technology is represented by the aggregate production function

$$Y = z\bar{k}^{\alpha}\bar{n}^{1-\alpha},$$

where $\bar{k}$ is the aggregate capital and $\bar{n}$ is the aggregate labor (including the labor supplied for household production). The economy is subject to aggregate shocks. The aggregate state is either good or bad. In a good state, $z = g$ and the unemployment rate for skill level η is μ_η^g. In a bad state, $z = b$ and the unemployment rate for skill level η is μ_η^b. We assume that $g > b$ and $\mu_\eta^g < \mu_\eta^b$ for $\eta = s, u$; i.e., productivity is higher and the unemployment rate is lower in booms than in recessions.

[2]This assumption is motivated by the empirical studies by Raisian (1983) and Keane and Prasad (1993).

[3]It is implicitly assumed that a worker can rent capital for household production without limit. This assumption is not inconsistent with the household borrowing constraint, introduced below, if the household production capital can be collateralized perfectly.

[4]This assumption is made so that an agent can earn some labor income even when she is unemployed. Alternatively, we can introduce an unemployment insurance system à la Hansen and İmrohoroğlu (1992). In this case, a government and its budget constraint would need to be incorporated into our setup.

We assume that there are no insurance markets for idiosyncratic shocks. Agents can hold only one kind of asset – capital. Holding a negative amount of capital (borrowing) is allowed up to an exogenous limit $\underline{k}$.

Aggregate capital and labor are given by

$$\bar{k} = \int_J k_j dj,$$

$$\bar{n} = \sum_{\eta=s,u} \phi(\eta)\chi_\eta\{(1-\mu_\eta^z) + h\mu_\eta^z)\},$$

where k_j is the capital holding of an agent j. The markets are competitive, so the interest rate r and the wage w are determined by their marginal products. Each household's period-by-period budget constraint is

$$c + k' = rk + w\phi(\eta)\theta(\epsilon) + (1-\delta)k,$$

where

$$\theta(\epsilon) = \begin{cases} h & \text{if } \epsilon = 0, \\ 1 & \text{if } \epsilon = 1, \end{cases}$$

k' is the next-period capital, and δ is the depreciation rate of capital.

2.2 Probability Structure

There are two types of exogenous shocks: aggregate and idiosyncratic. The aggregate state evolves stochastically following a Markov process. The probability of moving from state z to state z' is denoted as $p_{zz'}$.

There are three idiosyncratic shocks: ϵ, β, and η. Following the standard Bewley-style model, the individual employment process is treated as an exogenous stochastic process. Idiosyncratic employment shocks $\epsilon \in \{0, 1\}$ follow a Markov process with transition probability $\pi_{\epsilon\epsilon'}^{zz'\eta'}$. Following the tradition of İmrohoroğlu (1989), we assume that the probability of becoming unemployed next period ($\epsilon' = 0$) depends not only on the current employment status (ϵ), but also on the current and next period's aggregate state (z and z'). Additionally, the probability of becoming unemployed depends on the skill level in the next period (η'), to reflect the heterogeneity exhibited in Figure 1.

The discount factor β is assumed to be stochastic. At each point in time some agents are more patient than others. We interpret each agent as an altruistic dynasty. Each agent's patience level may differ across generations. This formulation serves as a device to produce a realistic wealth distribution.[5] We assume that β's process is independent of the other aggregate and idiosyncratic state variables. The Markov transition probability is denoted as $\omega_{\beta\beta'}$.

We assume that the individual skill level $\eta \in \{u, s\}$ follows an exogenous stochastic process. This transition probability is denoted as $q_{\eta\eta'}$. Again, we can interpret each agent as an altruistic dynasty. Within each dynasty, the skill level may differ across generations. The probability $q_{\eta\eta'}$ reflects the intergenerational mobility of skill levels.[6] The skill transition process is assumed to be independent of the other state variables.[7]

2.3 Recursive Competitive Equilibrium

Let Γ denote the measure of agents over $(k, \epsilon, \eta, \beta)$. The state variables relevant to each individual are the aggregate state variables (z, Γ) and the idiosyncratic state variables $(k, \epsilon, \eta, \beta)$. Let $\mathbf{T}$ denote the equilibrium transition function for Γ:

$$\Gamma' = \mathbf{T}(\Gamma, z, z').$$

Definition 1 (Recursive Competitive Equilibrium) *The recursive competitive equilibrium consists of the value function $v(k, \epsilon, \eta, \beta; z, \Gamma)$, a set of decision rules for consumption and asset holdings $\{c(k, \epsilon, \eta, \beta; z, \Gamma), k'(k, \epsilon, \eta, \beta; z, \Gamma)\}$, aggregate capital and labor $\{\bar{k}(z, \Gamma), \bar{n}(z, \Gamma)\}$,*

[5] This method was first developed by Krusell and Smith (1998).

[6] In our formulation, the timing of the switch in β and η may not coincide. Synchronizing these processes is possible, but it complicates the analysis considerably and would not substantially alter our main results.

[7] Belzil and Hansen (2003) estimate a structural model of educational choice and argue that the difference in discount rates has little effect on educational choice.

factor prices $\{w(z,\Gamma), r(z,\Gamma)\}$*, and a law of motion for the distribution,* $\Gamma' = \mathbf{T}(\Gamma, z, z')$*, which satisfy*

1. *Given the aggregate states,* $\{z, \Gamma\}$*, prices* $\{w(z,\Gamma), r(z,\Gamma)\}$*, and the law of motion for the distribution,* $\Gamma' = \mathbf{T}(\Gamma, z, z')$*; the value function* $v(k, \epsilon, \eta, \beta; z, \Gamma)$ *and the individual decision rules* $\{c(k, \epsilon, \eta, \beta; z, \Gamma), k'(k, \epsilon, \eta, \beta; z, \Gamma)\}$ *solve the following dynamic programming problem:*

$$v(k, \epsilon, \eta, \beta; z, \Gamma) = \max_{c,k'} \{\log c + \beta E[v(k', \epsilon', \eta', \beta'; z', \Gamma') | \epsilon, \eta, \beta, z, \Gamma]\}$$

subject to

$$c + k' = r(z,\Gamma)k + w(z,\Gamma)\phi(\eta)\theta(\epsilon) + (1-\delta)k,$$

$$k' \geq \underline{k},$$

and

$$\Gamma' = \mathbf{T}(\Gamma, z, z').$$

2. *Firms optimize:*

$$w(z,\Gamma) = (1-\alpha) z \bar{k}^{\alpha} \bar{n}^{-\alpha},$$

$$r(z,\Gamma) = \alpha z \bar{k}^{\alpha-1} \bar{n}^{1-\alpha}.$$

3. *Markets clear:*

$$\bar{k} = \int k d\Gamma,$$

$$\bar{n} = \int \phi(\eta)\theta(\epsilon) d\Gamma.$$

4. *Consistency:*

$$\Gamma'(K, E, X, B) = \int_{K,E,X,B} \left[\int_{\mathcal{K},\mathcal{E},\mathcal{X},\mathcal{B}} \mathcal{I}_{\{k' = k'(k,\epsilon,\eta,\beta;z,\Gamma)\}} \pi_{\epsilon\epsilon'}^{zz'\eta'} q_{\eta\eta'} \omega_{\beta\beta'} d\Gamma \right] dk' d\epsilon' d\eta' d\beta'$$

for all $K \subset \mathcal{K}$*,* $E \subset \mathcal{E}$*,* $X \subset \mathcal{X}$*, and* $B \subset \mathcal{B}$*.* $\mathcal{K}$*,* $\mathcal{E}$*,* $\mathcal{X}$*, and* $\mathcal{B}$ *are the sets of all possible realizations of* k*,* ϵ*,* η*, and* β*, respectively. The indicator function* $\mathcal{I}_{\{\cdot\}}$ *takes the value*

1 if the statement is true, and 0 if it is false. The transition probabilities $\pi_{\epsilon\epsilon'}^{zz'\eta'}$*,* $q_{\eta\eta'}$*, and* $\omega_{\beta\beta'}$ *are defined in the Appendix.*

2.4 Calibration

We follow a standard calibration for the most part. One period is considered to be six weeks. We choose $\delta = 0.0125$ and the average value of β as 0.995. The capital share $\alpha = 0.36$.

Aggregate shocks take the values $z \in \{0.99, 1.01\}$, and $p_{zz'}$ is chosen so that the average duration of each aggregate state is 2 years. The household production parameter h is assumed to be 0.1. These numbers closely follow the calibration of Krusell and Smith (1999, 2002). The borrowing constraint $\underline{k}$ is set at -13, which is tighter than the "always payback constraint". This number is chosen so that the fraction of the people with negative wealth mimics the actual data.

The idiosyncratic probabilities $\pi_{\epsilon\epsilon'}^{zz'\eta'}$ and $q_{\eta\eta'}$ are chosen so that the processes of unemployment and intergenerational mobility of education mimic the data. The process of β is chosen so that the resulting wealth distribution parallels the real-life wealth distribution. The skill premium, $\phi(s)/\phi(u)$, is set to 1.50. Details on how we calibrated the stochastic processes and the skill premium are found in the Appendix.

2.5 Model Solution

Generally, it is computationally burdensome to solve this type of model. The state variables in the individual optimization include the economy-wide wealth distribution, which is an infinite-dimensional object. In our model, the wealth distribution is in the state variables, since this information is necessary for predicting the next-period prices (for each aggregate state). To predict these prices, the agents have to predict the next period's aggregate capital, which requires knowledge of the current period's wealth distribution. Krusell and Smith (1998) developed a computational method to overcome this obstacle. They found that knowledge of only a few moments of the wealth distribution is often sufficient for predicting the next period's aggregate capital. In fact, they demonstrated that a linear prediction rule based only on the first moment $\bar{k}$ provides very accurate prediction. They call their method "approximate aggregation".

This method works very well in our setting. In our model, we postulate a prediction rule

$$\ln \bar{k}' = \alpha_0 + \alpha_1 \ln \bar{k} + \alpha_2 \ln z.$$

The coefficients $\alpha_0 = 0.1152$, $\alpha_1 = 0.9768$, and $\alpha_2 = 0.0916$ provide a very accurate prediction. The R^2 of this prediction rule is 0.99998.

The aggregate capital fluctuates between the values $\bar{k} = 140.2$ and $\bar{k} = 149.7$. For the most part, $\bar{k}$ is contained in between 142 and 148. On average, skilled agents hold $k = 184.1$ and unskilled agents hold $k = 104.7$. Since the earning ratio is $1 : 1.5$, the difference in asset holdings is more pronounced than the earnings difference.[8] This large difference in asset holding is important, since wealth holdings are the only means of self-insurance for unemployment in our incomplete-market setting.

The wealth distribution is matched to the data. In the data[9] the Gini coefficient of the wealth distribution is 0.80, while in the model, it is 0.79. In the right tail of the distribution, the top 10% of wealth-rich people hold about 69% of the real economy's total wealth, while 68% of wealth is held by the top 10% in the model. In the data, the top 20% (fifth wealth quintile) hold 82% of the total wealth, while the corresponding number is 84% in the model. In the left tail of the distribution, in the model approximately 8% of the agents hold negative wealth. In the data, 7.4% of the population report negative wealth and 2.5% report zero wealth. In the model, less than 2% hold negative wealth within the group of skilled agents, while nearly 15% of unskilled agents fall into this category. The model produces more wealth

[8] This property can also be seen in data. The average wealth-earnings ratio of a college graduate is higher than that of a high school graduate. See Burdía Rodríguez, Díaz-Giménez, Quadrini, and Ríos-Rull (2002).

[9] All the data in this paragraph are drawn from Burdía Rodríguez, Díaz-Giménez, Quadrini, and Ríos-Rull (2002). Wolff (1995) employs a slightly different definition of wealth (most notable difference is that he does not include vehicles and pension plans). In Wolff (1995), top 20% hold 85% of the total wealth and the Gini coefficient is 0.84.

inequality within unskilled agents than within skilled agents: Gini coefficients are 0.86 and 0.71, respectively.

3 Removing Business Cycles

The main question to be answered is: what will happen to people's welfare when business cycles are eliminated? To answer this, we follow Lucas's tradition in not describing specific policies to eliminate cycles. We directly eliminate shocks that are driving the aggregate fluctuations. The elimination is permanent, and this event is unanticipated by the agents.

3.1 Aggregate Shocks

Since the aggregate shocks are the driving force of the business cycles, a natural way to eliminate cycles is to replace the aggregate stochastic process by a deterministic process. In the spirit of Lucas, we replace the aggregate stochastic process by its conditional mean. The aggregate state z starts at $z = 1.01$ or $z = 0.99$ depending on the timing of the removal, and z converges monotonically to the value 1, which is the unconditional mean of z.

3.2 Idiosyncratic Shocks

We assume that when the business cycles are eliminated, the part of the idiosyncratic risk which is correlated with the aggregate shocks is also eliminated. Krusell and Smith (1999, 2002), who first proposed this procedure coined it the "integration principle". Formally, when an idiosyncratic random variable y can be written as a function $\bar{g}(i, z)$ of the aggregate variable z and a random variable i which is independent from z, then the new idiosyncratic variable after the elimination of the cycles is given by

$$\hat{y}(i) = \int \bar{g}(i, z) f_z(z) dz, \tag{2}$$

with density $f_i(i)$ for each i. Here, $f_z(z)$ and $f_i(i)$ are the marginal density functions for z and i, respectively.

As an example[10], consider an individual variable $y = z + i$, where $z \sim N(\mu_z, \sigma_z^2)$ and $i \sim N(\mu_i, \sigma_i^2)$ [thus, $y \sim N(\mu_z + \mu_i, \sigma_z^2 + \sigma_i^2)$]. Our procedure results in $\hat{y}(i) = \mu_z + i$, where since $i \sim N(\mu_i, \sigma_i^2)$, it follows that $\hat{y}(i) \sim N(\mu_z + \mu_i, \sigma_i^2)$. Notice that the mean is the same between y and $\hat{y}(i)$, but the variance is reduced after applying the integration principle.

In our model, the individual random variable ϵ is a two-point Markov process, and the procedure becomes more complex. (Other individual random variables, η and β, are independent from the aggregate shocks, and therefore their processes are unchanged after the elimination of business cycles.) A detailed description of this procedure is found in the Appendix. The resulting employment process $\hat{y}$ has less dispersion than ϵ. The coefficient of variation of $\hat{y}$ is approximately 15% smaller than ϵ for skilled agents, and about 30% smaller for unskilled agents. Therefore, the agents face less earnings risk in an economy without business cycles.

3.3 Competitive Equilibrium

The equilibrium after removing business cycles does not have a recursive structure for two reasons. First, the evolution of aggregate capital $\bar{k}$ and aggregate labor $\bar{n}$ (both are now deterministic) is time-dependent. Second, to calculate $\hat{y}$ at each time period, the entire history of idiosyncratic shocks has to be taken into account.[11]

Definition 2 (Competitive Equilibrium without Aggregate Shocks) *Let time 0 be the moment when the business cycles are removed. Denote the history of idiosyncratic shocks for an agent j by h_j^t. The competitive equilibrium consists of individual decision rules for consumption and the next period capital $\{c_t(h_j^t), k_{t+1}(h_j^t)\}$, factor prices $\{w_t, r_t\}$, and the aggregate variables $\{\bar{k}_t, \bar{n}_t\}$ that satisfy*

[10] This example follows Lucas (2003).

[11] In the Appendix, it is shown that the necessary information about the history of the idiosyncratic shocks can be summarized by two state variables. However, the evolution of those variables is also time-dependent.

1. *Given* $\{w_t, r_t\}$, *consumers optimize:*

$$E_0 \sum_{t=0}^{\infty} (\prod_{j=0}^{t} \beta_j) \log c_t,$$

with $\beta_0 = 1$, *subject to*

$$c_t + k_{t+1} = r_t k_t + w_t \phi(\eta_t) \hat{y}_t + (1-\delta) k_t,$$

$$k_{t+1} \geq \underline{k}.$$

2. *Firms optimize:*

$$w_t = (1-\alpha) \bar{k}_t^{\alpha} \bar{n}_t^{-\alpha},$$

$$r_t = \alpha \bar{k}_t^{\alpha-1} \bar{n}_t^{1-\alpha}.$$

3. *Markets clear:*

$$\bar{k}_t = \int_J k_{t,j} dj,$$

where $k_{t,j}$ *is the asset holding of an agent* j *at time* t,

$$\bar{n} = \int \phi(\eta_t) \hat{y}_t d\Gamma_t,$$

where Γ_t *is the measure over* η_t *and* $\hat{y}_t$.

We employ three approximations to reduce the computational burden. First, since $\bar{n}$ converges to a constant value fairly quickly, we treat $\bar{n}$ as a constant after a sufficient number of periods. Second, we reduce the evolution of $\bar{k}$ to a single law of motion after a sufficient number of periods. Third, we approximate the stochastic process of $\hat{y}$ by a five-point Markov process. These approximations serve to reduce the computational time dramatically, while providing a fairly accurate approximation of the original economy.

3.4 Results

Our experiment is to eliminate business cycles at a specific point from the fluctuating economy.[12] The timing of this "elimination" (call it time 0) is selected according to a specific level of capital stock and aggregate shocks. We select four timings:

1. $\bar{k} = 146$ and $z = g$,

2. $\bar{k} = 146$ and $z = b$,

3. $\bar{k} = 144$ and $z = g$,

4. $\bar{k} = 144$ and $z = b$.

After time 0, the economy experiences the transition to a non-fluctuating steady state. We compare the welfare of each agent at time 0, taking this transition into account.

3.4.1 Case 1: $\bar{k} = 146$ and $z = g$

Consider the case where we remove the business cycles when $\bar{k} = 146$ and $z = g$. The solid line in Figure 2 shows the transition path of capital from the time when the business cycles are removed. The dotted line represents the counterfactual path of capital with business cycle fluctuations. Capital converges to a new steady state with $\bar{k} = 144.4$, which is lower than the average value in the fluctuating economy. The main reason why it is lower is because individuals have less incentive to save for precautionary motives, due to the lower risk in their income.

In the new steady state, the Gini coefficient increases to 0.82. Inequality increases within both unskilled and skilled agents: Gini coefficients rise from 0.86 to 0.90 and from 0.72 to 0.74, respectively. This is largely due to the fact that since there is less precautionary saving, the left tail of the distribution is extended. In fact, the number of agents with negative assets

[12] The algorithm used for the computation can be found in the Appendix.

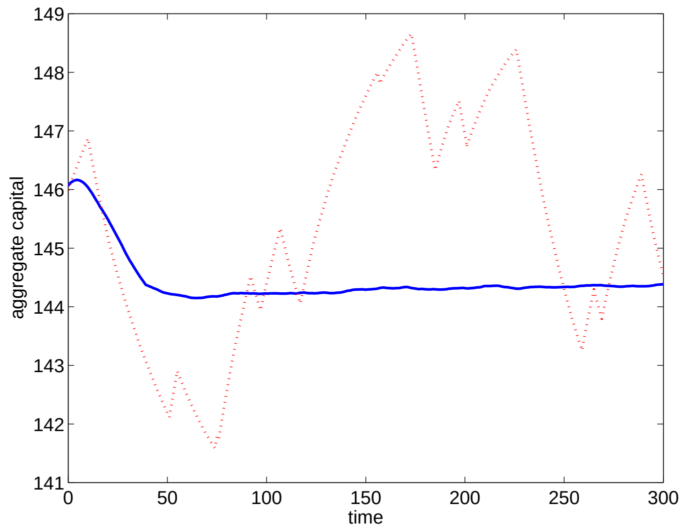

Figure 2: Path of Aggregate Capital after Removing Business Cycles

increases to 15.7% in total (27.4% for unskilled and 4.1% for skilled), which is almost twice compared to the fluctuating economy.

Our main goal is to compare the welfare between the two economies. We follow Lucas's method (1), and calculate the value of λ for each agent. Details of the calculation is delegated to the Appendix. The average value of λ is 0.028%, which is more than three times larger than Lucas's, and comparable to the numbers found in previous studies by Atkeson and Phelan (1994) and Krusell and Smith (2002). There is a large heterogeneity between skill levels: unskilled agents gain 0.069% from the elimination of business cycles, while skilled gain only 0.008%. As a result of stabilization, 93.1% of unskilled workers have increased utility while the fraction of skilled workers with a positive gain is much smaller at 42.9%. In fact, the majority of skilled workers experience lower utility by stabilization. Two effects produce the difference in costs of business cycles between skilled and unskilled workers. First, a majority of the skilled agents already accumulated enough wealth to insure themselves against the idiosyncratic risk. Unskilled workers tend to gain more since their level of wealth is lower on average. Second, at an individual level, the unskilled agents were facing larger unemployment

	asset level (in total wealth distribution)				
	constrained	bottom 1%	50%	99%	99.9%
$\eta = u$, $\epsilon = 0$	0.177%	0.093%	-0.022%	-0.002%	0.119%
$\eta = u$, $\epsilon = 1$	0.145%	0.118%	0.047%	0.017%	0.129%
$\eta = s$, $\epsilon = 0$	0.068%	0.008%	-0.071%	-0.005%	0.090%
$\eta = s$, $\epsilon = 1$	0.043%	0.030%	-0.005%	0.021%	0.108%

Table 1: The values of λ for $\beta = 0.995$

risk under business cycle fluctuations. There is also the general equilibrium effect (discussed below), whose direction is ambiguous.

The costs at an individual level are shown in Table 1. We focus on the agents with $\beta = 0.995$.[13] We observe even larger heterogeneity at an individual level. For example, a wealth-constrained agent with $\eta = u$ (unskilled) and $\epsilon = 0$ (unemployed) realizes more than six times larger cost than on average. (Those are the agents whose cost is the highest.) This figure, 0.177%, is more than twenty-two times larger than the number obtained by Lucas.[14] Tables in the Appendix show that if we eliminate the cycles when the aggregate state is bad, the largest cost is more than 0.75%, which is more than ninety times larger than Lucas's number.

When considering the welfare effect on individuals, two factors have to be taken into account.

1. Direct effect of the risk reduction.

2. General equilibrium effect.

The first is straightforward. When business cycles are eliminated, the aggregate shocks are completely smoothed out (the interest rate and the wage become deterministic variables),

[13]Heterogeneity of the costs across different β is discussed in detail by Krusell and Smith (2002).

[14]It is somewhat smaller than the number that Krusell and Smith (2002) obtained. There are two factors that reduce our numbers compared to theirs. First, in our calibration, the average duration of unemployment is substantially shorter than Krusell and Smith's. Second, our borrowing constraint is tighter, therefore the "constrained agents" here are not as wealth-poor as Krusell and Smith's.

and the idiosyncratic employment risk is also reduced. This benefits all agents, especially the agents who cannot self-insure by their own savings. The second effect calls for a more careful analysis. Since the steady-state capital is lower after eliminating the business cycles (because of less precautionary saving), on average, the interest rate rises and the wage falls. This benefits an agent for whom capital income is more important than wage income.

Keeping the employment status and skill status constant (within each rows of Table 1), the gains from eliminating business cycles exhibit a "U-shape" pattern.[15] Borrowing-constrained agents have a larger gain, reflecting the fact that they cannot self-insure their risk by their own assets. The direct effect of the reduction in idiosyncratic risk is very large for these agents. The "middle class" tends to have small or negative gains. For these agents, the benefit from the reduction in the idiosyncratic risk is small, since they have enough assets to insure themselves. In this case, the general equilibrium effect dominates. The middle-class agents whose income is largely coming from wage may experience lower welfare due to the wage loss. Very rich agents realize welfare gains since their income is largely coming from capital income.

Keeping the wealth level constant (within each columns of Table 1), we cannot always determine whether employed agents gain more than unemployed agents, or if skilled agents gain more than unskilled agents, due to the presence of the general equilibrium effect.

For the borrowing-constrained agents, it is usually possible to make a definite comparison, since the direct effect of the risk reduction dominates the other effects. For the constrained agents, the unemployed agents gain more than the employed agents, and the unskilled agents gain more than the skilled agents. This is due to the simple fact that the unemployed and the unskilled agents have smaller chances of getting out of the borrowing-constrained situation, compared to the the employed and the skilled agents, and therefore the direct benefit from the risk reduction is larger for them.

[15]This pattern is also observed in Krusell and Smith (2002).

	Avg λ	Avg λ for $\eta = u$	Avg λ for $\eta = s$	% of u gaining	% of s gaining
$\bar{k} = 146,\ z = g$	0.028%	0.069%	0.008%	93.1%	42.9%
$\bar{k} = 144,\ z = g$	0.021%	0.048%	0.007%	88.3%	54.5%
$\bar{k} = 146,\ z = b$	0.069%	0.147%	0.048%	100%	100%
$\bar{k} = 144,\ z = b$	0.073%	0.152%	0.050%	99.9%	100%

Table 2: Average gains from different starting points

For the agents who are not constrained, the direct effect is smaller and the general equilibrium effect is more pronounced. The general equilibrium effect works in an ambiguous fashion. Given the wealth level, wage income is more important for the employed agents than for the unemployed agents, simply because the employed agents earn higher wages. Thus unemployed agents should gain more from the general equilibrium effect. In the future however, the wealth level is not given. Given the *current* wealth level, the *future* wealth level is higher for the employed agents. Therefore, employed workers benefit more from the higher interest rate in future. The relative level of this current and future effect determines who will gain more from eliminating business cycles. This same logic is applied to the relationship between the skilled agents and the unskilled agents.

3.4.2 Case 2 – Case 4

For the other cases, we summarize the results for average λ in Table 2. On average, again, the unskilled gain more than the skilled. The individual costs corresponding to Table 1 in Case 1 are delegated to the Appendix. Again, we observe the “U-shaped” pattern in all the cases. For the borrowing constrained agents, it is always the case that the unemployed agents gain more than the employed agents, and the unskilled agents gain more than the skilled agents.

4 Conclusion

In this paper we calculated the costs of business cycles for different groups of people under incomplete markets. We focused primarily on the difference in skills. Unskilled agents face more cyclical unemployment risk and they have less opportunity to self-insure. As a result, the cost of business cycles is much larger for a typical unskilled agent compared to a typical skilled agent.

This difference in costs has an important implication in the political process. It is likely that the majority of unskilled agents favor a stabilization policy (if it comes with a small cost), while many skilled agents may vote against such a policy, if the burden falls evenly on different groups. A policy that directly transfers cyclical risk from unskilled to skilled workers may be politically more agreeable. To analyze such possibilities, incorporating specific policies and political processes into an incomplete markets setting seems to be a promising future research topic.

Appendix

A Constructing the Probability Matrices and Invariant Distributions

A.1 Aggregate Shocks

For aggregate shocks $z \in \{b, g\}$, the transition matrix is:

$$\begin{bmatrix} p_{bb} & p_{bg} \\ p_{gb} & p_{gg} \end{bmatrix},$$

where p_{ij} is the probability of the transition from state i to state j. Following Krusell and Smith (1999, 2002) we set the average business cycle duration to 2 years. Our model period is six weeks, therefore the average duration is 16 periods. From $1/(1-p_{bb}) = 1/(1-p_{gg}) = 16$, $p_{bb} = p_{gg} = 0.9375$, so

$$\begin{bmatrix} p_{bb} & p_{bg} \\ p_{gb} & p_{gg} \end{bmatrix} = \begin{bmatrix} 0.9375 & 0.0625 \\ 0.0625 & 0.9375 \end{bmatrix}.$$

The invariant distribution is [0.5 0.5].

A.2 Skill Transition

For each generation, the transition probability from skilled (college graduate) to skilled is 0.65 and the transition probabilty from unskilled (high school graduate) to skilled is 0.35, the transition matrix satisfies[16]

$$\begin{bmatrix} q_{uu} & q_{us} \\ q_{su} & q_{ss} \end{bmatrix}^{240} = \begin{bmatrix} 0.65 & 0.35 \\ 0.35 & 0.65 \end{bmatrix}.$$

This provides

$$\begin{bmatrix} q_{uu} & q_{us} \\ q_{su} & q_{ss} \end{bmatrix} = \begin{bmatrix} 0.9975 & 0.0025 \\ 0.0025 & 0.9975 \end{bmatrix}.$$

[16]Mayer (2002, Table 1) shows that in PSID data, the intergenerational transition probability (between fathers and sons) from no-college to college is 35%, while college to college is 73%. Statistics Canada (1998, p.37) compares the intergenerational transition of schooling attainment across countries. It shows that in United States, 64.2% of the population attains postsecondary schooling if their parents attained the postsecondary level education. If the parents attain only up to secondary schooling, the percentage drops to 35.7%.

The invariant distribution is

$$\begin{bmatrix} \chi_u & \chi_s \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}.$$

Note that if we start from the invariant distribution, the fraction of skilled workers remains constant by the law of large numbers.

A.3 Individual Shocks

For individual shocks $\epsilon \in \{0, 1\}$, the transition matrix has to be conditioned on last period's aggregate state (z), today's aggregate state (z'), and today's skill level (η'). Denote the matrix as $\Pi^{zz'\eta'}$. Here the case $z = b$, $z' = b$, and $\eta' = u$ is illustrated.

$$\Pi^{bbu} = \begin{bmatrix} \pi_{00}^{bbu} & \pi_{01}^{bbu} \\ \pi_{10}^{bbu} & \pi_{11}^{bbu} \end{bmatrix}.$$

The unemployment rate of skill level η when the aggregate state is z is denoted as μ_η^z. We calibrate μ_η^z from the Current Population Survey. Each year in between 1970 to 2001 is divided into two categories (two equal numbers of good and bad years) by ranking the years according to the total unemployment rate. μ_η^z is given as the average unemployment rate of the skilled and the unskilled for the good and the bad years.[17]

The number of people who were unskilled and unemployed in the last period is $\chi_u \mu_u^b$. They remain unskilled in the current period with probability q_{uu}. Thus, the number of people who were unskilled and unemployed in the last period, and remain unskilled in the current period is $\chi_u \mu_u^b q_{uu}$. The amount of people who were skilled and unemployed in the last period is $\chi_s \mu_s^b$. They become unskilled in the current period with probability q_{su}. Thus, the amount of people who were skilled and unemployed in the last period, and become unskilled in the current period is $\chi_s \mu_s^b q_{su}$. Summing up, the people who were unemployed in the last period

[17] The skilled are defined as some college or college completion, and the unskilled are defined as high school completion or less. Since the population of each group changes over time, we took the weighted average of the unemployment rates in each group using the number of individuals aged 24 and up (the population is taken from the census data) at each year.

and unskilled in the current period is $\chi_u \mu_u^b q_{uu} + \chi_s \mu_s^b q_{su}$. The people who were employed in the last period and unskilled in the current period is $\chi_u(1-\mu_u^b)q_{uu} + \chi_s(1-\mu_s^b)q_{su}$. Thus, as the transition occurs

$$\left[\begin{array}{cc} \chi_u \mu_u^b q_{uu} + \chi_s \mu_s^b q_{su} & \chi_u(1-\mu_u^b)q_{uu} + \chi_s(1-\mu_s^b)q_{su} \end{array}\right] \left[\begin{array}{cc} \pi_{00}^{bbu} & \pi_{01}^{bbu} \\ \pi_{10}^{bbu} & \pi_{11}^{bbu} \end{array}\right]$$

$$= \left[\begin{array}{c} \pi_{00}^{bbu}\chi_u \mu_u^b q_{uu} + \pi_{00}^{bbu}\chi_s \mu_s^b q_{su} + \pi_{10}^{bbu}\chi_u(1-\mu_u^b)q_{uu} + \pi_{10}^{bbu}\chi_s(1-\mu_s^b)q_{su} \\ \pi_{01}^{bbu}\chi_u \mu_u^b q_{uu} + \pi_{01}^{bbu}\chi_s \mu_s^b q_{su} + \pi_{11}^{bbu}\chi_u(1-\mu_u^b)q_{uu} + \pi_{11}^{bbu}\chi_s(1-\mu_s^b)q_{su} \end{array}\right]'.$$

Since the current period is a bad state, the first entry has to be equal to $\chi_u \mu_u^b$. This provides us with the first restriction. (The second entry has to be equal to $\chi_u(1-\mu_u^b)$, but it is easy to show that this is automatically satisfied by the first restriction provided that χ is the invariant distribution.) Now we have two unknowns, π_{00}^{bbu} and π_{10}^{bbu} (the rest are determined by the condition that the probabilities sum up to one: $\pi_{00}^{bbu} + \pi_{01}^{bbu} = 1$ and $\pi_{10}^{bbu} + \pi_{11}^{bbu} = 1$) and one equation. Another restriction is provided from the unemployment duration data. The Current Population Survey provides the average duration of unemployment in each year. We calculated the average duration in the good years and in the bad years (defined by the total unemployment rate), and obtained that the duration is 12.4 weeks for good years and 15.9 weeks for bad years.[18] Mincer (1991) shows (using PSID data) that the average duration of unemployment is not significantly different between skilled and unskilled workers, and therefore we use the same numbers for the skilled and the unskilled. The restriction is $1/(1-\pi_{00}^{bbu}) = 15.9/6$. When selecting $\pi_{00}^{bg\eta'}$ and $\pi_{00}^{gb\eta'}$, there exist two approachs:

1. $\pi_{00}^{bg\eta'} = \pi_{00}^{gg\eta'}$ and $\pi_{00}^{gb\eta'} = \pi_{00}^{bb\eta'}$ (İmrohoroğlu [1989])

2. $\pi_{00}^{bg\eta'} = 0.75 \cdot \pi_{00}^{gg\eta'}$ and $\pi_{00}^{gb\eta'} = 1.25 \cdot \pi_{00}^{bb\eta'}$ (Krusell and Smith [1999, 2002]).

We follow Krusell and Smith.

From the data on unemployment between 1970 and 2001 (described above), we calculate $\mu_u^b = 0.087$, $\mu_u^g = 0.056$, $\mu_s^b = 0.038$, and $\mu_s^g = 0.026$.

[18] İmrohoroğlu (1989) uses the duration of 10 weeks in good times and 14 weeks in bad times. Krusell and Smith (1999, 2002) use 19.5 weeks and 32.5 weeks.

Given above numbers, π_{10}^{bbu} is derived from

$$\pi_{10}^{bbu} = \frac{\chi_u \mu_u^b - \pi_{00}^{bbu}(\chi_u \mu_u^b q_{uu} + \chi_s \mu_s^b q_{su})}{\chi_u(1-\mu_u^b)q_{uu} + \chi_s(1-\mu_s^b)q_{su}}.$$

In general, $\pi_{10}^{zz'\eta'}$ is derived from

$$\pi_{10}^{zz'\eta'} = \frac{\chi_{\eta'} \mu_{\eta'}^{z'} - \pi_{00}^{zz'\eta'}(\chi_u \mu_u^z q_{u\eta'} + \chi_s \mu_s^z q_{s\eta'})}{\chi_u(1-\mu_u^z)q_{u\eta'} + \chi_s(1-\mu_s^z)q_{s\eta'}}.$$

The resulting Π matrices are:

$$\Pi^{bbu} = \begin{bmatrix} 0.6226 & 0.3774 \\ 0.0383 & 0.9617 \end{bmatrix},$$

$$\Pi^{gbu} = \begin{bmatrix} 0.7783 & 0.2217 \\ 0.0483 & 0.9517 \end{bmatrix},$$

$$\Pi^{bgu} = \begin{bmatrix} 0.3871 & 0.6129 \\ 0.0269 & 0.9731 \end{bmatrix},$$

$$\Pi^{ggu} = \begin{bmatrix} 0.5161 & 0.4839 \\ 0.0309 & 0.9691 \end{bmatrix},$$

$$\Pi^{bbs} = \begin{bmatrix} 0.6226 & 0.3774 \\ 0.0152 & 0.9848 \end{bmatrix},$$

$$\Pi^{gbs} = \begin{bmatrix} 0.7783 & 0.2217 \\ 0.0184 & 0.9758 \end{bmatrix},$$

$$\Pi^{bgs} = \begin{bmatrix} 0.3871 & 0.6129 \\ 0.0123 & 0.9877 \end{bmatrix},$$

$$\Pi^{ggs} = \begin{bmatrix} 0.5161 & 0.4839 \\ 0.0134 & 0.9866 \end{bmatrix}.$$

A.4 Stochastic β

Following Krusell and Smith (1998), we assume that the discount factor β_t follows a three-point Markov stochastic process. Let $\beta_t \in \{l, m, h\}$, where $l < m < h$. An agent with the discount factor $\beta = l$ is impatient, and an agent with the discount factor $\beta = h$ is patient. First, we calibrate the Markov transition matrix by imposing the following restrictions.

- 10% of the total population has $\beta = l$, 80% of the population has $\beta = m$, and 10% of the population has $\beta = h$.
- There is no direct transition in between $\beta = l$ and $\beta = h$.
- The average duration of the extreme states, $\beta = l$ and $\beta = h$, is one generation (30 years).

The transition probabilities from state i to state j, ω_{ij} are

$$\begin{bmatrix} \omega_{ll} & \omega_{lm} & \omega_{lh} \\ \omega_{ml} & \omega_{mm} & \omega_{mh} \\ \omega_{hl} & \omega_{hm} & \omega_{hh} \end{bmatrix} = \begin{bmatrix} 239/240 & 1/240 & 0 \\ 1/1920 & 959/960 & 1/1920 \\ 0 & 1/240 & 239/240 \end{bmatrix}.$$

Second, the values of $\tilde{\beta}_l, \tilde{\beta}_m, \tilde{\beta}_h$ are pinned down so that the resulting distribution of the asset holdings mimics the real wealth distribution. In particular, $\tilde{\beta}_l = 0.992$, $\tilde{\beta}_m = 0.995$, and $\tilde{\beta}_h = 0.998$.

A.5 Skill Premium

In the model $\phi(s)/\phi(u)$ is the skill premium. To calibrate $\phi(s)/\phi(u)$ we use the estimates of Murphy and Welch (1992). They compute the ratio of average wage of college graduate workers to high-school graduate workers for different experience groups and for different years. They find that this ratio, which can be interpreted as the skill premium, is between 1.37 to 1.58. To be consistent with their estimates, we set $\phi(s)/\phi(u)$ to 1.50.

One can also calibrate $\phi(s)/\phi(u)$ by using the estimates for return to college education. If one assumes that the return to one year of college education is 10%, which is consistent with the estimates in Card (1995), then $\phi(s)/\phi(u)$ is around 1.50.

B Applying the "Integration Principle" to Idiosyncratic Shocks

Applying the "integration principle" to a two-point process is more difficult than the continuous example in the main text. Here, starting from the static case, we extend the analysis step by step.[19]

B.1 Static Case

Let z take two values, g and b, and ϵ take two values, 0 and 1. The aggregate state z occurs with the probability $\pi_z(z)$, and ϵ's probability depends only on the current z. The conditional probability of ϵ given z is denoted as $\pi(\epsilon|z)$. We define $i \sim U(0,1)$, and

$$\bar{g}(i,z) = \begin{cases} 1 & \text{if } i \leq \pi(1|z), \\ 0 & \text{otherwise.} \end{cases}$$

Thus, when $i \in [0, \pi(1|b)]$,

$$\hat{y} = \sum_z \bar{g}(i,z)\pi_z(z) = \bar{g}(i,g)\pi_z(g) + \bar{g}(i,b)\pi_z(b) = 1 \cdot \pi_z(g) + 1 \cdot \pi_z(b) = 1.$$

When $i \in (\pi(1|b), \pi(1|g)]$,

$$\hat{y} = \bar{g}(i,g)\pi_z(g) + \bar{g}(i,b)\pi_z(b) = 1 \cdot \pi_z(g) + 0 \cdot \pi_z(b) = \pi_z(g).$$

When $i \in (\pi(1|g), 1]$,

$$\hat{y} = \bar{g}(i,g)\pi_z(g) + \bar{g}(i,b)\pi_z(b) = 0 \cdot \pi_z(g) + 0 \cdot \pi_z(b) = 0.$$

In sum, $\hat{y} = 1$ with probability $\pi(1|b)$, $\hat{y} = \pi_z(g)$ with probability $\pi(1|g) - \pi(1|b)$, and $\hat{y} = 0$ with probability $1 - \pi(1|g)$.

B.2 Correlation over Time

When z and ϵ are correlated over time, applying this procedure requires more thought. Suppose that z evolves by a first-order Markov process and that ϵ_t depends on z_t, z_{t-1}, and ϵ_{t-1}. Let i_t be an i.i.d. random variable which follows $U[0,1]$. We must then find a function $\bar{g}(\cdot,\cdot)$ which satisfies

$$\bar{g}_t(\{i_s\}_0^t, \{z_s\}_0^t) = g_t(\{\epsilon_s\}_0^t, \{z_s\}_0^t), \tag{3}$$

[19] For exposition, in this section we utilized slightly different notation from the main text. The correspondence should be clear, however.

and then integrate $\bar{g}_t(\{i_s\}_0^t, \{z_s\}_0^t)$ over $\{z_s\}_0^t$. Note that the right hand side is the employment variable, $g_t(\{\epsilon_s\}_0^t, \{z_s\}_0^t) = \epsilon_t$.

B.2.1 Brute-Force

One way to do this is by brute-force simulation: generate z_t and i_t randomly, and then create ϵ_t from the realization. Iterate the simulation many times – then, for each $\{i_s\}_0^t$, there will be a distribution of ϵ_t (depending on the realizations (history) of z, ϵ can be different for the same $\{i_s\}_0^t$). Average this out and use as the new idiosyncratic shocks at each t. We did not utilize this method here.

B.2.2 Recursive

Instead, we utilized the following method, which exploits the recursive structure of the problem. From the distributional assumptions, to express ϵ_t by an i.i.d. random variable $i_t \sim U(0,1)$, the additional information required is z_{t-1}, z_t, and ϵ_{t-1}. That is, given z_{t-1}, z_t, and ϵ_{t-1}, ϵ_t can be determined by the rule

$$\epsilon_t = \begin{cases} 1 & \text{if } i_t \leq \Omega(z_{t-1}, z_t, \epsilon_{t-1}), \\ 0 & \text{otherwise}, \end{cases} \tag{4}$$

where $\Omega(z_{t-1}, z_t, \epsilon_{t-1})$ is the threshold value calculated from the original Markov transition matrices. However, we can not integrate this yet. $\Omega(z_{t-1}, z_t, \epsilon_{t-1})$ still depends on ϵ_{t-1}. To construct $\bar{g}_t(\cdot,\cdot)$ function, we still require ϵ_{t-1} to be expressed by i and z. By working from $t = 0$ using (4) (ϵ_{-1} is given), we can express ϵ_{t-1} by $\{\{i_s\}_0^{t-1}, \{z_s\}_0^{t-1}\}$. Clearly, this procedure has recursive structure.

Equation (3) can be expressed as $\bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}) = \epsilon_t$, where ϵ_{t-1} on the left hand side is actually a function of $\{\{i_s\}_0^{t-1}, \{z_s\}_0^{t-1}\}$. The integration principle requires us to calculate, for each $\{i_s\}_0^t$,

$$\hat{y}_t = \sum_{z_t} \cdots \sum_{z_0} \bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}) \cdot \pi_z(z_t, \ldots, z_0).$$

This can be rewritten as

$$\begin{aligned}\hat{y}_t &= \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, z_{t-1}, \epsilon_{t-1}) \cdot \pi_z(g, z_{t-1} \ldots, z_0) \\ &\quad + \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, b, z_{t-1}, \epsilon_{t-1}) \cdot \pi_z(b, z_{t-1}, \ldots, z_0).\end{aligned}$$

The first part can be rewritten as

$$\begin{aligned}& \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, z_{t-1}, \epsilon_{t-1}) \cdot \pi_z(g, z_{t-1} \ldots, z_0) \\ = & \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, g, \epsilon_{t-1}) \cdot \pi_z(g, g, z_{t-2} \ldots, z_0) \\ & + \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, b, \epsilon_{t-1}) \cdot \pi_z(g, b, z_{t-2} \ldots, z_0).\end{aligned} \tag{5}$$

The second part can be expressed in the similar way.

Let $P_{t-1}(g, \epsilon)$ be the probability that, for given $\{i_s\}_0^{t-1}$, the realization of $\{z_s\}_0^{t-1}$ induces (i) $z_{t-1} = g$ and (ii) $\epsilon_{t-1} = \epsilon$.

Further, let $\pi_z(z_t | z_{t-1})$ be the conditional probability. Then, the first part of (5) can be rewritten as:

$$\begin{aligned}& \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, g, \epsilon_{t-1}) \cdot \pi_z(g, g, z_{t-2} \ldots, z_0) \\ = \; & \bar{g}_t(i_t, g, g, 1) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 1) + \bar{g}_t(i_t, g, g, 0) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 0).\end{aligned} \tag{6}$$

Here, $\bar{g}_t(i_t, g, g, 1)$ is either 0 or 1, and is easy to calculate using (4). $\pi_z(g|g)$ is given by the Markov transition matrix. Thus, given $P_{t-1}(z_{t-1}, \epsilon_{t-1})$, $\hat{y}_t$ can be calculated only from the information of i_t, using (6) for all possible combinations of z_{t-1} and z_t.

How can we get $P_{t-1}(z_{t-1}, \epsilon_{t-1})$? It can be calculated recursively. First, notice that $P_{t-1}(z_{t-1}, 0) + P_{t-1}(z_{t-1}, 1) = \pi_z(z_{t-1})$, where $\pi_z(z_{t-1})$ can be mechanically calculated from the Markov transition matrix and the initial value z_0. Thus, we only need to keep track of $P_{t-1}(z_{t-1}, 1)$. To obtain $P_t(g, 1)$, we have to calculate the probability of (i) $z_t = g$ and (ii) $\epsilon_t = 1$. This can be done by just picking up the $z_t = g$ part in the $\hat{y}_t$ calculation (sum of (6) for all possible combinations of z_{t-1} and z_t), since $\bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}) = 1$ if $\epsilon_t = 1$ and

$\bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}) = 0$ if $\epsilon_t = 0$. That is,

$$\begin{aligned} P_t(g,1) &= \bar{g}_t(i_t,g,g,1)\cdot\pi_z(g|g)\cdot P_{t-1}(g,1) \\ &+ \bar{g}_t(i_t,g,g,0)\cdot\pi_z(g|g)\cdot P_{t-1}(g,0) \\ &+ \bar{g}_t(i_t,g,b,1)\cdot\pi_z(g|b)\cdot P_{t-1}(b,1) \\ &+ \bar{g}_t(i_t,g,b,0)\cdot\pi_z(g|b)\cdot P_{t-1}(b,0). \end{aligned}$$

In the same way,

$$\begin{aligned} P_t(b,1) &= \bar{g}_t(i_t,b,g,1)\cdot\pi_z(b|g)\cdot P_{t-1}(g,1) \\ &+ \bar{g}_t(i_t,b,g,0)\cdot\pi_z(b|g)\cdot P_{t-1}(g,0) \\ &+ \bar{g}_t(i_t,b,b,1)\cdot\pi_z(b|b)\cdot P_{t-1}(b,1) \\ &+ \bar{g}_t(i_t,b,b,0)\cdot\pi_z(b|b)\cdot P_{t-1}(b,0). \end{aligned}$$

Note that

$$\hat{y}_t = P_t(g,1) + P_t(b,1)$$

by construction. Up to this part, our procedure follows Krusell and Smith (2002).

B.3 Extending to Multiple Skill Levels

Suppose that the employment probability is also dependent on skill level $\eta \in \{u, s\}$, which evolves stochastically. The evolution of η is first-order Markov, and independent of ϵ and z. Specifically, $y_t (= \epsilon_t)$ depends on ϵ_{t-1}, z_{t-1}, z_t, and η_t. Now, (4) has to be modified to[20]

$$\epsilon_t = \begin{cases} 1 & \text{if } i_t \le \Omega(z_{t-1}, z_t, \epsilon_{t-1}, \eta_t), \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Here, ϵ_{t-1} depends on $\{\{i_s\}_0^{t-1}, \{\eta_s\}_0^{t-1}, \{z_s\}_0^{t-1}\}$. Clearly, after integration, $\hat{y}$ is a function of $\{\{i_s\}_0^t, \{\eta_s\}_0^t\}$. Similar steps from the previous case apply. Given $\{\{i_s\}_0^t, \{\eta_s\}_0^t\}$,

$$\hat{y}_t = \sum_{z_t} \cdots \sum_{z_0} \bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}, \eta_t) \cdot \pi_z(z_t, \ldots, z_0).$$

This can be rewritten as

$$\begin{aligned} \hat{y}_t = & \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, z_{t-1}, \epsilon_{t-1}, \eta_t) \cdot \pi_z(g, z_{t-1} \ldots, z_0) \\ & + \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, b, z_{t-1}, \epsilon_{t-1}, \eta_t) \cdot \pi_z(b, z_{t-1}, \ldots, z_0). \end{aligned}$$

[20]Clearly, $\Omega(z_{t-1}, z_t, \epsilon_{t-1}, \eta_t)$ here should be set to $\pi_{\epsilon_{t-1}1}^{z_{t-1}z_t\eta_t}$ in the main text.

The first part can be rewritten as

$$
\begin{aligned}
& \sum_{z_{t-1}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, z_{t-1}, \epsilon_{t-1}, \eta_t) \cdot \pi_z(g, z_{t-1} \ldots, z_0) \\
= & \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, g, \epsilon_{t-1}, \eta_t) \cdot \pi_z(g, g, z_{t-2} \ldots, z_0) \\
& + \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, b, \epsilon_{t-1}, \eta_t) \cdot \pi_z(g, b, z_{t-2} \ldots, z_0).
\end{aligned} \tag{8}
$$

The second part can be expressed in a similar way.

Let $P_{t-1}(g, \epsilon)$ be the probability that, for given $\{\{i_s\}_0^{t-1}, \{\eta_s\}_0^{t-1}\}$, the realization of $\{z_s\}_0^{t-1}$ induces (i) $z_{t-1} = g$ and (ii) $\epsilon_{t-1} = \epsilon$.

Further, let $\pi_z(z_t|z_{t-1})$ be the conditional probability. Then, the first part of (8) can be rewritten as:

$$
\begin{aligned}
& \sum_{z_{t-2}} \cdots \sum_{z_0} \bar{g}_t(i_t, g, g, \epsilon_{t-1}, \eta_t) \cdot \pi_z(g, g, z_{t-2} \ldots, z_0) \\
= & \ \bar{g}_t(i_t, g, g, 1, \eta_t) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 1) + \bar{g}_t(i_t, g, g, 0, \eta_t) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 0).
\end{aligned} \tag{9}
$$

Here, $\bar{g}_t(i_t, g, g, 1, \eta_t)$ is either 0 or 1, and can be calculated by (7). $\pi_z(g|g)$ is given by the Markov transition matrix. Thus, given $P_{t-1}(z_{t-1}, \epsilon_{t-1})$, $\hat{y}_t$ can be calculated only from the information of i_t and η_t, using (9) for all possible combinations of z_{t-1} and z_t.

One could calculate $P_{t-1}(z_{t-1}, \epsilon_{t-1})$ recursively. First, notice that again $P_{t-1}(z_{t-1}, 0) + P_{t-1}(z_{t-1}, 1) = \pi_z(z_{t-1})$, where $\pi_z(z_{t-1})$ can be mechanically calculated from the Markov transition matrix and z_0. Thus, we only need to keep track of $P_{t-1}(z_{t-1}, 1)$. To obtain $P_t(g, 1)$, we have to calculate the probability of (i) $z_t = g$ and (ii) $\epsilon_t = 1$. This can be done by just picking up the $z_t = g$ part in the $\hat{y}_t$ calculation (sum of (9) for all possible combinations of z_{t-1} and z_t), since $\bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}, \eta_t) = 1$ if $\epsilon_t = 1$ and $\bar{g}_t(i_t, z_t, z_{t-1}, \epsilon_{t-1}, \eta_t) = 0$ if $\epsilon_t = 0$. That is,

$$
\begin{aligned}
P_t(g, 1) = & \quad \bar{g}_t(i_t, g, g, 1, \eta_t) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 1) \\
& + \bar{g}_t(i_t, g, g, 0, \eta_t) \cdot \pi_z(g|g) \cdot P_{t-1}(g, 0) \\
& + \bar{g}_t(i_t, g, b, 1, \eta_t) \cdot \pi_z(g|b) \cdot P_{t-1}(b, 1) \\
& + \bar{g}_t(i_t, g, b, 0, \eta_t) \cdot \pi_z(g|b) \cdot P_{t-1}(b, 0).
\end{aligned} \tag{10}
$$

In the same way,

$$
\begin{aligned}
P_t(b,1) \quad = & \qquad \bar{g}_t(i_t, b, g, 1, \eta_t) \cdot \pi_z(b|g) \cdot P_{t-1}(g,1) \\
& + \quad \bar{g}_t(i_t, b, g, 0, \eta_t) \cdot \pi_z(b|g) \cdot P_{t-1}(g,0) \\
& + \quad \bar{g}_t(i_t, b, b, 1, \eta_t) \cdot \pi_z(b|b) \cdot P_{t-1}(b,1) \\
& + \quad \bar{g}_t(i_t, b, b, 0, \eta_t) \cdot \pi_z(b|b) \cdot P_{t-1}(b,0).
\end{aligned}
\tag{11}
$$

Note that

$$\hat{y}_t = P_t(g,1) + P_t(b,1)$$

by construction.

B.4 Algorithm

Denote the last period with aggregate fluctuations as $t = 0$.

1. For each agent, given z_0, ϵ_0, and η_0, simulate i_t and η_t, and thus obtain the sequence of P_t for $t = 1, 2, ...$ Sum up and obtain the aggregate labor supply $\bar{n}_t$ at $t = 1, 2, ...$ Check when $\bar{n}_t$ and the distribution of the labor supply settles down. Call that period N_1. (We used $N_1 = 40$ to 100, depending on the starting point.) From the process of P_t, we can obtain the time series of $\hat{y}$. Instead of using P_t in the individual decision problem, we approximate this process of $\hat{y}$ by a finite-state Markov process and use this as the individual shocks.

2. Pick $N_2 > N_1$. (We used $N_2 = 1000$.) Use the average of the law of motions in $t < 0$ to guess $\bar{k}_t$, $t = 1, 2, ..., N_2$. Call them $\bar{k}_t^0$.

3. Given the law of motion for $\bar{k}$ and the stationary value of $\bar{n}_t$, perform the value-function iteration (part of Krusell-Smith's [1998] method) to obtain the value function for the periods $t = N_1 + 1, ..., N_2$. Note that the decision problem is

$$V(k, \hat{y}, \eta, \beta; \bar{k}) = \max_{c,k'} \left\{ \log c + \beta E \left[V(k', \hat{y}', \eta', \beta'; \bar{k}') \, | \hat{y}, \eta, \beta \right] \right\}$$

subject to

$$c + k' = rk + w\phi(\eta)\hat{y} + (1 - \delta)k,$$

$$\bar{k}' = H(\bar{k}),$$

$$k' \geq \underline{k},$$

$$r \text{ and } w \text{ calculated from } \bar{n} \text{ and } \bar{k}.$$

4. From $t = N_1$, work backwards to obtain value functions and decision rules for $t = 1, ..., N_1$.

$$V_{N_1}(k, \hat{y}, \eta, \beta) = \max_{c,k'} \left\{ \log c + \beta E \left[V(k', \hat{y}', \eta', \beta'; \bar{k}_{N1+1}) \,|\hat{y}, \eta, \beta \right] \right\}$$

subject to

$$c + k' = rk + w\phi(\eta)\hat{y} + (1 - \delta)k,$$

$$k' \geq \underline{k},$$

$$r \text{ and } w \text{ calculated from } \bar{n}_{N_1} \text{ and } \bar{k}^0_{N_1}.$$

5. Simulate the economy from $t = 0$, using the initial distribution at $t = 0$ and the decision rules obtained above.

6. Compare the simulated path of $\bar{k}$ and $\bar{k}^0$. If they are not close enough, update the $\bar{k}$ sequence by the weighted average. Also obtain the new prediction rule for $\bar{k}$ by performing OLS for the new $\bar{k}_t$, $t = N_1, ..., N_2$.

C How to Find the Welfare Cost

The welfare cost from business cycles is defined (following Lucas) as λ which satisfies

$$E_0 \left[\sum_{t=0}^{\infty} (\prod_{j=0}^{t} \beta_j) U((1 + \lambda) c_t^o) \right] = E_0 \left[\sum_{t=0}^{\infty} (\prod_{j=0}^{t} \beta_j) U(c_t^s) \right],$$

where $\{c_t^o\}_{t=0}^{\infty}$ is the consumption stream in the original economy (with business cycles), and $\{c_t^s\}_{t=0}^{\infty}$ is the consumption stream in the "smoothed" economy (without business cycles).

For logarithmic utility, some calculation yields

$$\lambda = \exp[(V_s - V_o)/B] - 1,$$

where $V_o \equiv E_0 \left[\sum_{t=0}^{\infty}(\prod_{j=0}^{t} \beta_j)U(c_t^o)\right]$ is the expected (average) discounted utility under the original economy, $V_s \equiv E_0 \left[\sum_{t=0}^{\infty}(\prod_{j=0}^{t} \beta_j)U(c_t^s)\right]$ is the expected (average) discounted utility under the smoothed economy, and $B \equiv E_0 \left[\sum_{t=0}^{\infty}(\prod_{j=0}^{t} \beta_j)\right]$.

D More Tables

	asset level				
	constrained	bottom 1%	50%	99%	99.9%
$\eta = u$, $\epsilon = 0$	0.172%	0.087%	-0.010%	0.003%	0.146%
$\eta = u$, $\epsilon = 1$	0.139%	0.112%	0.039%	0.019%	0.156%
$\eta = s$, $\epsilon = 0$	0.064%	0.004%	-0.075%	0.031%	0.109%
$\eta = s$, $\epsilon = 1$	0.038%	0.026%	-0.010%	0.057%	0.126%

Table 3: The value of λ for $\beta = 0.995$, $\bar{k} = 144$, $z = g$

	asset level				
	constrained	bottom 1%	50%	99%	99.9%
$\eta = u$, $\epsilon = 0$	0.752%	0.402%	0.166%	0.140%	0.288%
$\eta = u$, $\epsilon = 1$	0.349%	0.197%	0.096%	0.123%	0.278%
$\eta = s$, $\epsilon = 0$	0.586%	0.313%	0.097%	0.182%	0.253%
$\eta = s$, $\epsilon = 1$	0.169%	0.055%	0.011%	0.153%	0.239%

Table 4: The value of λ for $\beta = 0.995$, $\bar{k} = 146$, $z = b$

	asset level				
	constrained	bottom 1%	50%	99%	99.9%
$\eta = u$, $\epsilon = 0$	0.660%	0.413%	0.173%	0.082%	0.234%
$\eta = u$, $\epsilon = 1$	0.262%	0.207%	0.102%	0.066%	0.225%
$\eta = s$, $\epsilon = 0$	0.500%	0.323%	0.106%	0.152%	0.192%
$\eta = s$, $\epsilon = 1$	0.087%	0.065%	0.020%	0.124%	0.178%

Table 5: The value of λ for $\beta = 0.995$, $\bar{k} = 144$, $z = b$

References

[1] Aiyagari, S. Rao (1994). "Uninsured Idiosyncratic Risk and Aggrigate Saving," *Quarterly Journal of Economics* 109, 659–684.

[2] Atkeson, Andrew and Christopher Phelan (1994). "Reconsidering the Costs of Business Cycles with Incomplete Markets," in Stanley Fischer and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual 1994*, National Bureau of Economic Research.

[3] Barlevy, Gadi (2002). "The Cost of Business Cycles under Endogenous Growth," mimeo. Northwestern University.

[4] Belzil, Christian and Jörgen Hansen (2003). "Structural Estimates of the Intergenerational Education Correlation," *Journal of Applied Econometrics* 18, 679–696.

[5] Burdía Rodríguez, Santiago; Javier Díaz-Giménez; Vincenzo Quadrini; and José-Victor Ríos-Rull (2002). "Updated Facts on the U.S. Distributions of Earnings, Income, and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review* 26, 2–35.

[6] Card, David (1995). "Earnings, Schooling, and Ability Revisited." In Solomon Polachek, ed., *Research in Labor Economics* Vol. 14. JAI Press: Greenwich Connecticut.

[7] Hansen, Gary D. and Ayşe İmrohoroğlu (1992). "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard," *Journal of Political Economy* 100, 118–142.

[8] İmrohoroğlu, Ayşe (1989). "The Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy* 97, 1364-1383.

[9] Keane, Michael and Eswar Prasad (1993). "Skill Levels and the Cyclical Variability of Employment, Hours, and Wages," *IMF Staff Papers* 40, 711–743.

[10] Krusell, Per and Anthony A. Smith, Jr. (1998). "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy* 106, 867-896.

[11] Krusell, Per and Anthony A. Smith, Jr. (1999). "On the Welfare Effects of Eliminating Business Cycles," *Review of Economic Dynamics* 2, 245-272.

[12] Krusell, Per and Anthony A. Smith, Jr. (2002). "Revisiting the Welfare Effects of Eliminating Business Cycles," mimeo. University of Rochester and Carnegie Mellon University.

[13] Lucas, Robert E. Jr. (1987). *Models of Business Cycles*, Basil Blackwell: New York.

[14] Lucas, Robert E. Jr. (2003). "Macroeconomic Priorities," *American Economic Review* 93, 1–14.

[15] Mayer, Adalbert (2002). "Education, Self-Selection and Intergenerational Transmission of Abilities," mimeo. University of Rochester.

[16] Mincer, Jacob (1991). "Education and Unemployment," NBER Working Paper, #3838.

[17] Murphy, Kevin M., and Finis Welch (1992). "The Structure of Wages," *Quarterly Journal of Economics* 107, 285-326.

[18] Obstfeld, Maurice (1994). "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability," *European Economic Review* 38, 1471-86.

[19] Raisian, John (1983). "Contracts, Job Experience, and Cyclical Labor Market Adjustments," *Journal of Labor Economics* 1, 152–170.

[20] Statistics Canada (1998). *Education Quarterly Review* Vol.5, no.2.

[21] Topel, Robert (1993). "What Have We Learned from Empirical Studies of Unemployment and Turnover?" *American Economic Review* 83, 110-115.

[22] Wolff, Edward N. (1995). *Top Heavy,* The New Press: New York.