Pareto Distributions and the Evolution of Top Incomes in the U.S.

Aoki, Shuhei and Nirei, Makoto

3 July 2013

Pareto Distributions and the Evolution of Top Incomes in the U.S.

Shuhei Aoki*
Faculty of Economics, Hitotsubashi University

Makoto Nirei†
Institute of Innovation Research, Hitotsubashi University

July 3, 2013

Abstract

This paper presents a dynamic general equilibrium model with heterogeneous firms and entrepreneur's portfolio choice. We analytically show that this model generates the Pareto distribution of top income earners and Zipf's law of firms at the steady state. The differential equation for the probability density distribution of income is derived and numerically evaluated. In the model, CEOs respond to a tax cut by increasing their share of stocks of their own firms, thereby increasing the diffusion of their wealth. The calibrated model shows that the transition path matches with the decline of the Pareto exponent of the income distribution and the trend of top 1% income share in the U.S. in recent decades. We argue that the low marginal income tax at the top bracket of income could lead to the higher dispersion of income among the top income earners, which results in the higher concentration of income in the top income group.

JEL Codes: D31, L11, O40
Keywords: income distribution; wealth distribution; Pareto exponent; top income share; firm size distribution; Zipf's law

1 Introduction

There has been a secular trend towards the concentration of income on top earners in the U.S. economy for the last three decades. According to Alvaredo et al. (2013), the income share of top 1% earners declined from around 18% to 8% after the 1930s but the trend has reversed during the 1970s. Since then, the top 1% share has grown to 18% by 2010, now on a par with the prewar level. Piketty and Saez (2003) found that this trend is particular in the very top percentile, while the concentration on the lesser percentile group has been much milder.

*Address: 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. Phone: +81 (42) 580-8477. Fax: +81 (42) 580-8195. Email: shuhei.aoki@gmail.com.

†Address: 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Phone: +81 (42) 580-8417. Fax: +81 (42) 580-8410. E-mail: nirei@iir.hit-u.ac.jp.

Along with the increasing trend of the top income share, a widening dispersion of income within the top income group has been also observed over the same periods. It is known that the tail part of income follows a Pareto distribution very well. When income follows a Pareto distribution with exponent λ, the ratio of the number of people who earn more than x_1 to those who earn more than x_2, for any income levels x_1 and x_2, is $(x_1/x_2)^\lambda$. Thus, the Pareto exponent λ is a measure of equality among the riches. The estimated Pareto exponent shows a close connection with the top income share historically. It declined from 2.46 in 1975 to 1.6 in 2010 along with the secular increase in the top 1% share.

In this paper, we argue that the concentration of income in the last three decades was driven by the economic force that caused income dispersion of top earners. Among the driving forces of dispersion among the rich, we pay special attention to the decrease in the marginal income tax rate, the importance of stock-related income through the widespread use of employee stock options for CEOs or founding entrepreneur's share ownership, and the changing volatility of firm's risk environment. We present a model of portfolio choice by CEOs, who can invest in their own firms' risky stocks or in risk-free assets. The dispersion of CEO's income is determined by the extent of the risk taken in their after-tax returns of portfolio.

We develop a dynamic general equilibrium model with heterogeneous firms and the CEO's portfolio choice. In this type of model, the distribution of CEO's pay can be strongly affected by the distribution of firm size. It is known that the firm size distribution follows a Zipf's law, a special case of Pareto distribution with exponent $\lambda = 1$. As a discipline for our approach, we require our model to generate the Zipf's law of firms, while our main focus is the Pareto distribution of income.

The contribution of the paper is summarized as follows. First, this paper presents a parsimonious neoclassical growth model that generates Zipf's law of firms and Pareto's law of incomes. The model is simple enough to allow the analytical derivation of the stationary distributions of firms and income. Second, we obtain an analytical expression for the evolution of probability density distribution of income in the transition path. Using this expression, we can implement numerical computation of the transition dynamics of income distribution after an unanticipated and permanent cut in top marginal income tax rate. Third, we calibrate the model parameters and show that the transition path matches the decline of the Pareto exponent of the income distribution and the trend of increasing top income share in the last three decades. Hence, we argue that the calibrated analysis of our model predicts that the tax cut and CEOs' response to tax in their portfolio can explain the widening dispersion and more concentration of income. The numerical exercises show that the change in firm's volatility explains little dispersion of top income, since the portfolio choice responds to the change in order to mitigate the impact of the firm's volatility on the risk of CEO's portfolio. The calibrated model brings out testable implications on CEO portfolios and future development of inequality under the current tax rate level.

Much has been debated about the causes of the concentration of income in recent decades. Among them, Piketty and Saez (2003) argue that a cut in top marginal income tax rate is one of the plausible interpretations, compared with other interpretations such as skill-biased technical change. Our paper shares the

view with theirs that a tax cut is an important factor. Our model differs from theirs in that a cut in top marginal income tax rate itself does not matter, while a cut in top marginal income tax rate relative to other taxes, such as capital gains and corporate taxes, does matter. The reason that a cut in top marginal income tax rate itself does not matter in our model comes from the property that top marginal income tax in our model plays the same role as dividend tax in the "new view" of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005).

Recently, several papers have built models to understand why income distribution follows a Pareto distribution. There are two types of approaches in the literature. First approach explains Pareto's law of incomes by the assumption that other distributions follow certain types of distributions. Gabaix and Landier (2008) take this approach. They construct a model of the CEO pay, which assumes that the firm size distribution follows Zipf's law and that the CEO's talent follows a certain distribution. Under the settings, they show that the CEO pay distribution follows a Pareto distribution. An advantage of their model is that their model is consistent with the two stylized facts, i.e., Zipf's law of firms and Pareto's law of incomes. Jones and Kim (2012) extend the model to be consistent with the recent decline in the Pareto exponent of the income distribution in the U.S., which is assumed to be constant in Gabaix and Landier (2008). Compared with the papers taking this approach, our paper's contribution is to build a model that generates the Zipf's law and the Pareto's law both from the productivity shocks of firms without assuming certain types of distributions.

Second approach explains Pareto's law of incomes by idiosyncratic shocks. Using a household model with a consumption function, Nirei and Souma (2007) show that idiosyncratic shocks on the household asset returns generate Pareto's law of assets and incomes. Benhabib et al. (2011) show a similar result in the model of households who optimally make saving and bequest decisions. These models are not dynamic general equilibrium models in a sense that they only consider the household side problem and do not consider the firm side. Nirei (2009) extends the framework to a Bewley-type model and shows that idiosyncratic shocks on firms' productivities generate the Pareto's law of incomes in dynamic general equilibrium environment. Toda (2012) also builds a similar but more analytically tractable dynamic general equilibrium model and derive the Pareto's law. Our paper belongs to this approach. Compared with the previous studies, this paper features a model that can explain the Zipf's law of firms, and analyzes how the recent tax cut affects the evolution of top incomes.[1]

Perhaps, the closest paper to ours is Kim (2013), who following the latter approach, builds a model of human capital accumulation with idiosyncratic shocks that generates the Pareto's law of incomes. Using the model, she analyzes how a cut in top marginal income tax in recent decades affects the Pareto exponent of income distribution. Compared with her paper, our paper's contribution is to build a model that also explains Zipf's law of firms, both from the same shocks that generate the Pareto's law of incomes. In addition, because the mechanism through which a tax cut affects top incomes is different between hers and ours, the predictions of the models are also different. For example, in her model,

[1]Especially, this paper's model is consistent with the fact that the firm's productivity distribution also follows a Pareto distribution (Mizuno et al., 2012).

an income tax cut encourages human capital accumulation among top income earners, which would result in the labor productivity increase in the U.S. in recent decades compared with the previous periods and other countries such as France. In contrast, in our model, a tax cut does not directly affect capital accumulation.

Finally, our model is also closely related with the general equilibrium models of the firm size distribution that explain Zipf's law of firms (for a survey, see Luttmer, 2010). The basic mechanism employed in our paper to generate Zipf's law of firms draws on the literature. Compared with the literature, our firm side formulation is a rather simplified one, because our focus is to understand the evolution of top incomes.

The organization of the paper is follows. Section 2 sets up a dynamic general equilibrium model. Section 3 discusses the firm side properties of the model and derives Zipf's law of firms. Section 4 defines the equilibrium of the model and how to solve the model. After defining the equilibrium, Section 5 illustrates how in the steady state the household asset and income distribution follows a Pareto distribution. Section 6 analyzes how a tax cut affects top incomes in our model and contrasts the results with data. Finally, Section 7 concludes.

2 Model

It is well known that the distribution of certain types of stochastic processes follows a Pareto distribution. The purpose of the model presented here is to incorporate the types of stochastic processes into otherwise standard general equilibrium model with incomplete markets and to replicate Pareto distributions observed as stylized facts. Key assumptions that generate Zipf's law of firms are that firm's productivity is affected by multiplicative idiosyncratic shocks and that there is a lower bound for the firm size. Similarly, key assumptions that generate Pareto's law of household's assets and incomes are that household's asset is affected by multiplicative idiosyncratic shocks and that each household faces a constant probability of death (i.e., the perpetual youth assumption). In the next sections, we discuss how these properties generate these laws.

2.1 Household's problem

There is a continuum of households with a mass L. As in Blanchard (1985), each household is discontinued by a Poisson hazard rate ν. Households participate in a pension program. If a household dies, all of his non-human wealth is distributed to living households. Instead, if a household does not die, he obtains a part of the capital of the dead. The amount he gets is proportional to his wealth, that is, the pension premium rate ν times his wealth.

The households consist of entrepreneurs and workers. A mass N of households are entrepreneurs referred to as entrepreneurs and can hold the shares of his firm $s_{i,t}$ referred to as risky asset and risk-free market portfolio $b_{i,t}$. We assume that the risky asset is affected by uninsurable idiosyncratic shocks. Each entrepreneur leaves the firm by a Poisson hazard rate p_f, and becomes a worker. We refer to such households as former entrepreneurs.[2]

[2] We introduce the former entrepreneurs for a purely quantitative reason. The qualitative results of paper are intact even when $p_f = 0$. Quantitatively, if we do not introduce the former

The remaining $L-N$ of households are workers and can only hold risk-free assets $b_{i,t}$ (therefore, for them, $s_{i,t}=0$). There are two types of workers. The one is the former entrepreneurs and the other is workers by birth. We refer to the latter type of workers as innate workers.

These households maximize expected discounted utility by choosing sequences of consumption and asset portfolio. Let $q_{i,t}$ and $d_{i,t}$ be the price and dividend of the risky asset. The return of the risky asset is described by the following stochastic process:

$$((1-\tau^e)d_{i,t}dt + dq_{i,t})/q_{i,t} = \mu_{q,t}dt + \sigma_{q,t}dB_{i,t},$$

where τ^e is the tax rate on risky asset and $B_{i,t}$ is a Wiener process. We interpret τ^e as top marginal tax rate on ordinary income in the numerical analysis. Risk-free assets yield a net return r_t^f with certainty. The sum of the two asset holdings constitutes a financial wealth $s_{i,t}q_{i,t}+b_{i,t}$.

These households earn a constant labor income flow w_t and obtain government transfers tr_t. The human asset is defined by $h_t = \int_t^\infty (w_u + tr_u)e^{-\int_t^u(\nu+r_s^f)ds}du$. The labor income flow is expressed as an annuity payment

$$w_t + tr_t = (\nu + r_t^f)h_t - dh_t/dt. \tag{1}$$

Let $a_{i,t} = s_{i,t}q_{i,t} + b_{i,t} + h_t$ denote total wealth of a household. The accumulation of total wealth grows according to the following process:

$$\begin{aligned} da_{i,t} &= (\nu(s_{i,t}q_{i,t}+b_{i,t}) + \mu_{q,t}s_{i,t}q_{i,t} + r_t^f b_{i,t} + (\nu + r_t^f)h_t - c_{i,t})dt \\ &\quad + \sigma_{q,t}s_{i,t}q_{i,t} \\ &= \mu_{a,t}a_{i,t}dt + \sigma_{a,t}a_{i,t}dB_{i,t}. \end{aligned} \tag{2}$$

where $\mu_{a,t}a_{i,t} \equiv \nu a_{i,t} + \mu_{q,t}x_{i,t}a_{i,t} + r_t^f(1-x_{i,t})a_{i,t} - c_{i,t}$, $\sigma_{a,t}a_{i,t} \equiv \sigma_{q,t}x_{i,t}a_{i,t}$, and $x_{i,t}$ is the share of $a_{i,t}$ invested in the risky asset. Note that $dB_{i,t}$ is a multiplicative shock to the asset accumulation in that the shock is multiplied by the current asset level $a_{i,t}$.

Household's dynamic programming problem is specified as follows.

$$V^i(a_{i,t}, \mathbf{S}_t, t) = \max_{c_{i,t}, x_{i,t}} \ln c_{i,t}dt + e^{-(\beta+\nu)dt}\,\mathsf{E}_t[V^{i'}(a_{i,t+dt}, \mathbf{S}_{t+dt}, t+dt)] \tag{3}$$

subject to (2), where $\mathbf{S}_t$ is a set of aggregate variables that describes the aggregate dynamics of the model (for the definition, see Section 4.2), V^i denotes value functions of household characteristics i: if the household is an entrepreneur $i=e$, if he is an innate worker $i=w$, and if he is a former entrepreneur $i=f$. Note that if the household is an entrepreneur, the household characteristics in the next period, denoted by i', can be both entrepreneur and worker. If the household is an innate worker or a former entrepreneur, $i'=i$.

The household problem is a variant of Merton's dynamic portfolio problem (Merton, 1969, 1971, 1973, and Campbell and Viceira, 2002). It is well known that the solution of the problem under the log coincides with the myopic rules,

entrepreneurs and all of the entrepreneurs remain the positions, the mobility of household's asset or income level becomes too slow or the Pareto exponent of income distribution becomes too low, compared with the data.

whose solution is :

$$x_{i,t} = \begin{cases} \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}, & \text{if } i = e, \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

$$v_{i,t} = \nu + \beta, \tag{5}$$

where $v_{i,t}$ is the consumption-wealth ratio (see Appendix A for derivations).

In the model, we assume that entrepreneurs can hold either risky assets of his own firm or risk-free assets. We can relax the assumption and allow entrepreneurs to hold risky assets of other firms whose expected returns are as low as that of risk-free assets, r^f, due to transaction costs explained in the next section. Then, because the shocks on risky assets are assumed to be uncorrelated with one another, the portfolio share of another firm's risky asset $x'_{i,t}$, becomes $(r_t^f - r_t^f)/(\sigma'_{q,t})^2 = 0$, where $\sigma'_{q,t}$ is the volatility of this risky asset, (see Campbell and Viceira, 2002). It implies that the results are unchanged by relaxing the assumption.

2.2 Firms and the financial market

A continuum of firms with a mass N produces differentiated goods. As in McGrattan and Prescott (2005), each firm issues shares, and owns and self-finances capital $k_{j,t}$. As noted above, the entrepreneur of the firm can directly own the shares of his firm. Financial intermediaries also own the shares of the firm and by combining the shares, issue risk-free market portfolio to households, which diversifies the idiosyncratic shocks of the firms. The financial intermediaries incur ι per dividend $d_{j,t}$ as transaction costs. We assume that financial intermediaries possess the majority shares, or that when an entrepreneur possesses his firm's shares, they are preferred stocks without voting rights. Then, following the interest of financial intermediaries, firms maximize the expected profits, and the market value of a firm is priced at the net present value of the after-tax profits discounted by the risk-free rate r_t^f. We make these assumptions to simplify the analysis.

2.2.1 Financial intermediary's problem

We assume that returns and risks on risky assets are ex ante identical and that shocks on the risky assets are uncorrelated with each other. A financial intermediary maximizes the residual profit by diversifying the risks on risky assets and issuing risk-free assets:

$$\max_{s_{j,t}^f} \mathsf{E}_t \left[\left(\int_0^N \left\{ (1 - \tau^f - \iota) d_{j,t} dt + dq_{j,t} \right\} s_{j,t}^f dj \right) \right] - r_t^f dt \left(\int_0^N q_{j,t} s_{j,t}^f dj \right),$$

where $s_{j,t}^f$ is the shares of firm j owned by the financial intermediary and τ^f is the tax on the dividend. We interpret τ^f in the numerical analysis as the combination of capital gains and corporate income taxes. The solution of the problem leads to

$$r_t^f q_{j,t} dt = \mathsf{E}_t[(1 - \tau^f - \iota) d_{j,t} dt + dq_{j,t}]. \tag{6}$$

2.2.2 Firm's problem

There are heterogeneous firms in the economy. The production function of firm j is

$$y_{j,t} = z_{j,t} k_{j,t}^{\alpha} \ell_{j,t}^{1-\alpha}.$$

The productivity of the firm evolves as

$$dz_{j,t} = \mu_z z_{j,t} dt + \sigma_z z_{j,t} dB_{j,t},$$

where $B_{j,t}$ is a Wiener process, which is uncorrelated with shocks in other firms. Note that $dB_{j,t}$ is a multiplicative shock to the productivity growth because the shock is multiplied by its productivity level $z_{j,t}$.

In order to derive the property that the firm size distribution is a Pareto distribution, we impose the following assumptions on the minimum level of firm size. We assume that there is a minimum level of employment $\ell_{\min}$, i.e.,

$$\ell_{j,t} \geq \ell_{\min}. \tag{7}$$

A firm whose optimal employment size is less than $\ell_{\min}$ is restructured. More precisely, we define the productivity level $z_{\min}$ as the one at which, when the firm optimally chooses labor (following (8) below), $\ell_{j,t} = \ell_{\min}$. We assume that the firm whose productivity $z_{j,t}$ is less than $z_{\min}$ has to be restructured in the way that the firm buys productivities and accompanying capitals from other firms at the market price to increase the firm size. Correspondingly, we assume that each firm sells a constant fraction of the firm to the restructuring firms (we will discuss how the deals are conducted in the next section).

A firm chooses the investment level $dk_{j,t}$ and employment $\ell_{j,t}$ to maximize the profit:

$$r_t^f q(k_{j,t}, z_{j,t}, \mathbf{S}_t, t) dt = \mathsf{E}_t \left[\max_{dk_{j,t}, \ell_{j,t}} (1 - \tau^f - \iota) d_{j,t} dt + dq(k_{j,t}, z_{j,t}, \mathbf{S}_t, t) \right], \tag{8}$$

The dividend $d_{j,t}$ consists of

$$d_{j,t} dt = (p_{j,t} y_{j,t} - w_t \ell_{j,t} - \delta k_{j,t})\, dt - dk_{j,t},$$

where $p_{j,t}$ and $y_{j,t}$ are the price and quantity of the good produced by the firm, $k_{j,t}$ is the capital, w_t is the wage rate, and δ is the depreciation rate.

By solving the firm's problem, we obtain the conditions (see Appendix B for details):

$$\mathsf{MPK}_t \equiv r_t^f + \delta = \frac{\partial p_{j,t} y_{j,t}}{\partial k_{j,t}}, \tag{9}$$

$$w_t = \frac{\partial p_{j,t} y_{j,t}}{\partial \ell_{j,t}}. \tag{10}$$

There are two remarks about the firm's problem. First, in the model, the MPK becomes the same among firms because the stochastic discount factor of those who own diversified bonds is not correlated with the shock of firm j. Second, because the taxes in the model are imposed on dividends as in the "new view" literature of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005), they do not affect the marginal product of capital (MPK).

2.3 Aggregation and market conditions

We consider the market conditions for the aggregate economy. (Throughout the paper, we denote the aggregate variables by upper case letters.) Goods that a mass N of firms produces are aggregated according to

$$Y_t = \left(\int_0^N \left(\frac{1}{N} \right)^{1-\rho} y_{j,t}^{\rho} dj \right)^{\frac{1}{\rho}}. \tag{11}$$

We assume that the aggregate good Y is produced competitively.

The market clearing condition for final goods is

$$C_t + \frac{dK_t}{dt} - \delta K_t + \iota \left(1 - \frac{A_{e,t} x_{e,t}}{Q_t} \right) D_t = Y_t, \tag{12}$$

where $A_{e,t}$ is the total assets of entrepreneurs and Q_t is aggregate financial asset. The last term in the left hand side of the equation indicates that a part of the final goods is consumed as transaction costs. The labor market clearing condition is

$$\int_0^N \ell_j dj = L. \tag{13}$$

The market clearing condition for the shares of firms is

$$s_j + s_j^f = 1, \tag{14}$$

where s_j is the shares owned by the entrepreneur according to (4) and s_j^f is the shares owned by financial intermediaries. We assume that government transfers are adjusted so that tax revenues equal government transfers period by period.

3 Firm side Properties

Before defining and solving the model, we review the following firm side properties of the model. First, in this model, given r_t^f, the firm side variables such as $\ell_{j,t}$, $k_{j,t}$, and $d_{j,t}$ can be obtained as the closed-form expressions. These variables can be written as a product of the components common across firms and the heterogeneous component. Second, the distribution of firm's productivity is obtained independently of other variables and is a Pareto distribution that establishes Zipf's law of firms when the minimum employment level $\ell_{\min}$ is sufficiently small.

3.1 Firm side variables

Employing firm's FOCs (9) and (10) together with the aggregate condition (11) and the labor market condition (13), the firm's variables can be written as

follows (for the derivations, see Appendices B.2 and B.3):

$$\ell_{j,t} = \overline{\ell}_t \; z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{\ell}_t \equiv \left(\frac{L/N}{\overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}} \right), \tag{15}$$

$$p_{j,t} y_{j,t} = \overline{py}_t \overline{\ell}_t \; z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{py}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}}, \tag{16}$$

$$k_{j,t} = \overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{k}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}}, \tag{17}$$

$$d_{j,t} dt = \overline{d}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}} dt - \left(\frac{\rho}{1-\rho} \right) \sigma_z \overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}} dB_{j,t}, \tag{18}$$

$$\text{where } \overline{d}_t \equiv (1-(1-\alpha)\rho)\overline{py}_t - (\delta + \mu_{k,t})\,\overline{k}_t,$$

$$q_{j,t} = \overline{q}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \quad \text{where } \overline{q}_t \equiv (1-\tau^f-\iota)\overline{d}_t \int_t^\infty \exp\left\{ -\int_t^u (r_s^f - \mu_{d,s}) ds \right\} du. \tag{19}$$

Note that $\overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}$ is the average of $z_{j,t}^{\frac{\rho}{1-\rho}}$ over the firm size distribution (we will show later that the average exists), and $\mu_{k,t}$ and $\mu_{d,t}$ are the expected growth rate of $k_{j,t}$ and $d_{j,t}$.

In the above equations, each variables have the common components such as $\overline{\ell}_t$ and $\overline{py}_t$ and the heterogeneous component, $z_{j,t}^{\frac{\rho}{1-\rho}}$. Thus, the size distributions of the firm side variables depend only on the heterogeneous component.

3.2 Restructuring

Before deriving the firm size distribution, we analyze how restructuring firms buy the assets of other firms. In each small time interval some firms decrease their productivities from t and $t+dt$ to $z_{j,t+dt} < z_{\min}$. Then, these firms have to buy $\overline{q}_{t+dt}\overline{\ell}_{t+dt}\left(z_{\min}^{\frac{\rho}{1-\rho}} - z_{j,t+dt}^{\frac{\rho}{1-\rho}} \right)$ from other firms to increase the firm size. We denote the total of these payments in the restructured firms as $Q_{\text{restructuring},t+dt}$.

At each instant, each firm sells a constant fraction $m\left(\frac{\rho}{1-\rho}\right) dt$ of the firm's value to these restructuring firms at the market price $m\left(\frac{\rho}{1-\rho}\right) q_{j,t} dt$. (The adjustment term $\left(\frac{\rho}{1-\rho}\right)$ enters because firm side variables are proportional to $z_{j,t}^{\frac{\rho}{1-\rho}}$.) The total value of the sellouts is

$$m\left(\frac{\rho}{1-\rho}\right) dt \int_0^N q_{j,t} dj = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}\overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\} m\left(\frac{\rho}{1-\rho}\right) dt.$$

Since the demand of restructuring firms equates the supply:

$$Q_{\text{restructuring},t+dt} = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}\overline{\mathsf{E}}\left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\} m\left(\frac{\rho}{1-\rho}\right) dt. \tag{20}$$

Rearranging this equation and taking the limit as dt approaches zero from above,

we obtain (see Appendix B.4 for details)

$$m = \left(\lambda - \frac{\rho}{1-\rho}\right)\frac{\sigma_z^2}{4}, \tag{21}$$

where λ is the Pareto exponent of the firm size distribution pinned down in the next section.

3.3 Firm size distribution

We detrend the firm's productivity to derive the invariant productivity distribution. Let $\tilde{z}_{j,t}$ be the firm's productivity level after selling a part of the firm's assets to restructuring firms, detrended by $e^{g_z t}$ (g_z is a constant whose value is determined below). The firm's detrended productivity growth after sellout is

$$\begin{aligned} d\tilde{z}_{j,t} &= (\mu_z - g_z - m)\,\tilde{z}_{j,t}dt + \sigma_z \tilde{z}_{j,t} dB_{j,t}, \\ \text{or, } d\ln\tilde{z}_{j,t} &= \left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)dt + \sigma_z dB_{j,t}. \end{aligned} \tag{22}$$

The Fokker-Planck equation for the probability density $f_z(\ln\tilde{z}_{j,t}, t)$ for firm's productivity is

$$\frac{\partial f_z(\ln\tilde{z}_{j,t},t)}{\partial t} = -\left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)\frac{\partial f_z(\ln\tilde{z}_{j,t},t)}{\partial \ln\tilde{z}_{j,t}} + \frac{\sigma_z^2}{2}\frac{\partial^2 f_z(\ln\tilde{z}_{j,t},t)}{\partial(\ln\tilde{z}_{j,t})^2}.$$

In this paper, we assume the invariant distribution for firms, i.e., $\partial f_z(\ln\tilde{z}_{j,t},t)/\partial t = 0$. When the invariant distribution exists, the Fokker-Planck equation has a solution in exponential form,

$$f_z(\ln\tilde{z}_{j,t}) = C_0 \exp(-\lambda \ln\tilde{z}_{j,t}), \tag{23}$$

where the coefficients satisfy:

$$C_0 = \lambda \tilde{z}_{\min}^{\lambda}, \quad \lambda = -2\left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)/\sigma_z^2. \tag{24}$$

(23) shows that the distribution of $\ln\tilde{z}_{j,t}$ follows an exponential distribution. By change of variables, it can also be shown that the distribution of $\tilde{z}_{j,t}$ follows a Pareto distribution whose Pareto exponent is λ.

In this model, the exogenous parameter $\ell_{\min}$ pins down λ and g_z. From the restriction on $\ell_{\min}$ and (15), we obtain the Pareto exponent for $\tilde{z}_{j,t}$ as,

$$\lambda = \frac{1}{1 - \frac{\ell_{\min}}{L/N}}\left(\frac{\rho}{1-\rho}\right). \tag{25}$$

With this λ, we obtain the rescaling parameter g_z that assures the existence of the invariant distribution of $\tilde{z}_{j,t}$.

There are four remarks on the firm size distribution. First, we obtain a constant rescaled mean $\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\}$ for a constant $\tilde{z}_{\min}$ as follows:

$$\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\} = \int_{\tilde{z}_{\min}}^{\infty} \tilde{z}^{\frac{\rho}{1-\rho}} f_z(\ln\tilde{z})\frac{\partial \ln\tilde{z}}{\partial\tilde{z}}d\tilde{z} = \frac{C_0 \tilde{z}_{\min}^{-\left(\lambda - \frac{\rho}{1-\rho}\right)}}{\lambda - \frac{\rho}{1-\rho}}.$$

it is shown that when $\tilde{z}_{\min}$ is a constant $\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\}$ is also a constant.

Second, the growth rate of the aggregate output is $g \equiv g_z/(1-\alpha)$. We can confirm this property by detrending and aggregating (16).

Third, the expected growth rate of $\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}$ is negative. It means that the expected growth rate of the detrended firm side variables is negative, while the mean of the detrended firm size distribution, which is proportional to $\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\}$, is constant. This is a key property that generates a Pareto distribution with a finite distributional mean.

Fourth, Zipf's law approximately holds for the firm size distribution, e.g., the distribution of $\ell_{j,t}$. This is because the firm size distribution cross-sectionally obeys to $\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}$, whose Pareto exponent is $\lambda/\left(\frac{\rho}{1-\rho}\right)$. (25) shows that $\lambda/\left(\rho/(1-\rho)\right) > 1$ and that if $\ell_{\min}$ is sufficiently small compared with the average employment level L/N, $\lambda/\left(\rho/(1-\rho)\right)$ becomes close to 1.

4 Equilibrium and Solution of the Model

In this model, because the household policy functions are independent of the household's wealth level, the dynamics of aggregate variables are obtained independent of the heterogeneity within entrepreneurs and workers.

4.1 Definition of a competitive equilibrium

A competitive equilibrium of the model given initial aggregate capital K_0, initial asset shares of entrepreneurs, innate workers, and former entrepreneurs, $A_{e,0}/A_0$, $A_{w,0}/A_0$, $A_{f,0}/A_0$, and the stationary detrended firm size distribution, is a set of variables, $\left\{A_{e,t}, A_{w,t}, A_{f,t}, H_t, K_t, Y_t, C_t, d_{j,t}, r_t^f, w_t, tr_t, v_t, x_t\right\}$, that satisfies the following conditions:

- household's decisions on the portfolio choice (4) and (5) and the law of motion for human and total assets (1) and (2),
- firm's decisions (15)–(19),
- and the market clearing conditions (12) and (14).

4.2 Solution of the model

Let variables with tilde such as $\tilde{K}_t$ be the variables detrended by e^{gt}. The aggregate dynamics of the detrended variables can be reduced to the differential equations of $\mathbf{S}_t = \left\{\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t\right\}$. The evolution of these variables is computed at each t as follows:

1. Given $\tilde{K}_t$,

$$\mathsf{MPK}_t = \alpha\rho\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}} \Big/ \left(\frac{\tilde{K}_t}{L}\right)^{1-\alpha}.$$

2. $\tilde{C}_t = v\tilde{A}_t$ and $\tilde{Q}_t = \tilde{A}_t - \tilde{H}_t$. Given MPK_t, $\tilde{Y}_t = \overline{py}_t/e^{gt}$ is pinned down. Then,

$$\begin{aligned}\frac{d\tilde{K}_t}{dt} =& \tilde{Y}_t - \delta\tilde{K}_t - \tilde{C}_t - \iota\left(1 - \frac{\tilde{A}_{e,t}x_t}{\tilde{Q}_t}\right)\tilde{D}_t - g\tilde{K}_t,\\ \tilde{D}_t =& (1-(1-\alpha)\rho)\tilde{Y}_t - (\delta + g + \mu_{\overline{\mathsf{E}}})\tilde{K}_t - \frac{d\tilde{K}_t}{dt},\end{aligned}$$

and $x_{e,t}$ are jointly determined.

Note that, here, the expected return and volatility of a risky asset are jointly determined as follows (see Appendix B.3 for details of the derivations):

$$\begin{aligned}\mu_{q,t} =& \left\{\left(\frac{1-\tau^e}{(1-\tau^f-\iota)} - 1\right)\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du} + r_t^f\right\},\\ \sigma_{q,t} =& \left(\frac{\rho}{1-\rho}\right)\sigma_z\\ &\times\left\{1 - \left(\frac{1-\tau^e}{(1-\tau^f-\iota)}\right)\frac{\tilde{K}_t}{\tilde{D}_t}\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du}\right\},\end{aligned}$$

where $\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du$ is computed by the following equation:

$$\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du = \frac{\tilde{Q}_t}{(1-\tau^f-\iota)\tilde{D}_t}.$$

3. The assets for the three types of households evolve as:

$$\begin{aligned}\frac{d\tilde{A}_{e,t}}{dt} &= (\mu_{ae,t} - g)\,\tilde{A}_{e,t} + (\nu + p_f)N\tilde{H}_t/L - (\nu + p_f)\tilde{A}_{e,t},\\ \frac{d\tilde{A}_{w,t}}{dt} &= (\mu_{aw,t} - g)\,\tilde{A}_{w,t} + (\nu L - (\nu + p_f)N)\tilde{H}_t/L - \nu\tilde{A}_{w,t},\\ \frac{d\tilde{A}_{f,t}}{dt} &= (\mu_{aw,t} - g)\,\tilde{A}_{f,t} + p_f\tilde{A}_{e,t} - \nu\tilde{A}_{f,t},\end{aligned}$$

where $\mu_{ae,t}$ and $\mu_{aw,t}$ are calculated according to (4) and (5). The human asset evolves as

$$\frac{d\tilde{H}_t}{dt} = -(\tilde{w}_t + \tilde{tr}_t)L + (\nu + r_t^f - g)\tilde{H}_t, \tag{26}$$

where

$$\begin{aligned}\tilde{w}_t =& (1-\alpha)\rho\tilde{Y}_t/L,\\ \tilde{tr}_t =& \left\{\frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\tau^e + \left(1 - \frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\right)\tau^f\right\}\tilde{D}_t/L.\end{aligned}$$

5 Household's Asset Distributions in the Steady State

In this model, the steady state household asset distribution can be derived analytically. We show below that the distributions of entrepreneurs, innate workers, and former entrepreneurs are all Pareto distributions. We also discuss that the asset, income, and consumption distributions of the total households follow a Pareto distribution at the upper-tail whose Pareto exponent coincides with that of the asset distribution of entrepreneurs.

5.1 Asset distribution of entrepreneurs

Individual entrepreneur's asset, $\tilde{a}_{e,t}$, if he does not die, evolves as

$$d\ln\tilde{a}_{e,t} = \left(\mu_{ae} - g - \frac{\sigma_{ae}^2}{2}\right)dt + \sigma_{ae}dB_{i,t},$$

where μ_{ae} and σ_{ae} are the drift and diffusion parts of the entrepreneur's asset process. Since they are constants in the steady state, we omit time subscript.

The initial asset of entrepreneurs with age t' at period t is $h_{t-t'}$. The relative asset of the entrepreneurs who are alive at t, relative to their initial asset is in a logarithmic expression, $\ln(a_{e,t}/h_{t-t'}) = \ln\tilde{a}_{e,t} - (\ln\tilde{h}_{t-t'} - gt')$, which follows a normal distribution with mean $(\mu_{ae} - \sigma_{ae}^2/2)t'$ and variance $\sigma_{ae}^2 t'$.

By combining the above property with the constant probability of death assumption, the asset distribution of entrepreneurs is obtained. The probability density function of log assets becomes a double-exponential distribution (see Appendix C for the derivations in this section)[3]

$$f_e(\ln\tilde{a}_i) = \begin{cases} f_{e1}(\ln\tilde{a}_i) \equiv \frac{(\nu+p_f)N}{L}\frac{1}{\theta} & \exp\left[-\psi_1(\ln\tilde{a}_i - \ln\tilde{h})\right] \\ & \text{if } \tilde{a}_i \geq \tilde{h}, \\ f_{e2}(\ln\tilde{a}_i) \equiv \frac{(\nu+p_f)N}{L}\frac{1}{\theta} & \exp\left[\psi_2(\ln\tilde{a}_i - \ln\tilde{h})\right] \\ & \text{otherwise}, \end{cases}$$

where

$$\begin{aligned} \psi_1 \equiv& \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2}\left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} - 1\right), \\ \psi_2 \equiv& \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2}\left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} + 1\right), \\ \theta \equiv& \sqrt{2(\nu+p_f)\sigma_{ae}^2 + (\mu_{ae} - g - \sigma_{ae}^2/2)^2}. \end{aligned}$$

This result shows that the asset distribution of entrepreneurs follows a double-Pareto distribution (Benhabib et al., 2012 and Toda, 2012), whose Pareto exponent at the upper-tail is ψ_1.

[3] We normalize the probability density functions of entrepreneurs, innate workers, and former entrepreneurs, $f_e(\ln\tilde{a}_i)$, $f_w(\ln\tilde{a}_i)$, and $f_f(\ln\tilde{a}_i)$ such that

$$\int_{-\infty}^{\infty}\left\{f_e(\ln\tilde{a}_i) + f_w(\ln\tilde{a}_i) + f_f(\ln\tilde{a}_i)\right\}d(\ln\tilde{a}_i) = 1.$$

5.2 Asset distribution of innate workers

Individual worker's asset, $\tilde{a}_{w,t}$, if he does not die, evolves as

$$d \ln \tilde{a}_{w,t} = (\mu_{aw} - g)\, dt,$$

where μ_{aw} is the drift part of the worker's asset process.

Under the asset process, the asset distribution of innate workers becomes

$$f_w(\ln \tilde{a}_i) = \begin{cases} \frac{\nu L - (\nu + p_f)N}{L} \frac{1}{|\mu_{aw} - g|} \exp\left(-\frac{\nu}{\mu_{aw} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) & \text{if } \frac{\ln \tilde{a}_i - \ln \tilde{h}}{\mu_{aw} - g} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The result shows that the log assets of innate workers follow an exponential distribution and that the assets of innate workers follow a Pareto distribution. With the parameter values in numerical analysis, the trend growth of worker's asset is close to the trend growth of the economy, i.e., $\mu_{aw} \approx g$. Then, the detrended assets of the innate workers are concentrated on the level around $\tilde{h}$.

5.3 Asset distribution of former entrepreneurs

The asset distribution of former entrepreneurs depends on the asset distribution of entrepreneurs, the Poisson rate p_f with which each entrepreneur leaves the firm, and the asset process after he becomes a worker.

The steady state asset distribution of the former entrepreneurs under the settings can be analytically derived. Here, for brevity, we only report the case where $\mu_{aw} \geq g$ (for the $\mu_{aw} < g$ case, see Appendix C):

$$f_f(\ln \tilde{a}_i) = \begin{cases} \frac{p_f}{\nu - \psi_1(\mu_{aw} - g)} f_{e1}(\ln \tilde{a}_i) - \left(\frac{1}{\nu - \psi_1(\mu_{aw} - g)} - \frac{1}{\nu + \psi_2(\mu_{aw} - g)}\right) p_f f_{e1}(\ln \tilde{h}) \\ \quad \times \exp\left(-\frac{\nu}{\mu_{aw} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) \quad \text{if } \ln \tilde{a}_i \geq \ln \tilde{h}, \\ \frac{p_f}{\nu + \psi_2(\mu_{aw} - g)} f_{e2}(\ln \tilde{a}_i) \quad \text{otherwise.} \end{cases}$$

The probability density function for $\tilde{a}_i \geq \tilde{h}$ consists of two exponential terms. The second term, representing the distribution of innate workers ' distribution, declines faster as an asset level increases than the first term, the distribution of entrepreneurs. Therefore, the Pareto exponent of the former entrepreneur's asset distribution becomes the same as that for entrepreneurs in the tail (the same result applies to the case where $\mu_{aw} < g$).

5.4 Pareto exponents of asset and income distributions for all of the households

We make two remarks on the household asset and income distributions. First, the Pareto exponent at the upper-tail for all of the households is that of entrepreneurs, ψ_1. This is because, as noted above, the distribution of the smallest Pareto exponent dominates at the upper-tail.

Second, in this model, the consumption and income distributions at the upper-tail are also Pareto distributions with the same Pareto exponent as that of assets, ψ_1. This is because the consumption and income of a household are proportional to the household's asset level.

6 Numerical Analysis

In this section, we numerically analyze how the cut in top marginal tax rate accounts for the evolution of top incomes in recent decades, by assuming that at 1975 the tax cut occurs in an unexpected and permanent way.

There are three reasons to choose 1975 as the year of the structural change. First, several empirical studies suggest inequality has begun to grow since the 1970s (see e.g., Katz and Murphy, 1992 and Piketty and Saez, 2003). Second, political scientists such as Hacker and Pierson (2010) argue that the U.S. politics transformed during the 1970s in favor of industries, which might have affected entrepreneurs' future expectations on tax rates. Third, the top marginal earned income tax declined from 77% to 50% during the 1970s alone (see Figure 1), which was followed by subsequent cuts in other taxes in the 1980s. These evidences suggest that a structural change occurred during the 1970s (we choose 1975 as the median).

In our model, a tax cut affects top incomes by changing entrepreneur's incentive to invest in the risky assets. In the tax parameters calibrated below, after 1975, the tax rate on risky asset τ^e becomes relatively lower than the tax rate on risk-free asset τ^f, which induces entrepreneurs to increase the share of risky assets in their asset portfolios. This is why the Pareto exponent declines and the top income share increases in our model.

6.1 Tax rates

We assume that the tax on risky assets τ^e is equal to the ordinary income tax that is imposed on the CEO pay. We assume that the tax on risk-free assets τ^f is the sum of taxes that are imposed on dividends when investors buy the equities of firms. We calculate the tax rate of risk-free assets, τ^f by $1-(1-\tau^{\text{cap}})(1-\tau^{\text{corp}})$, where τ^{cap} and τ^{corp} are the marginal tax rates for capital gains and corporate income. These tax rates are calibrated using top statutory marginal federal tax rates reported in Saez et al. (2012) (see Figure 1 and Table 1).

Insert Figure 1 here.
Insert Table 1 here.

6.2 Calibration

The parameters are chosen to roughly match the annual data. The first five parameters at Table 2 are standard values. For example, we assume for ν that the average length of life after a household begins to work is 50 years.

ρ is set to 0.7, which implies that 30% of firm's sales is rent. The value of ρ is lower than the standard one. There are two reasons for this. First, model's treatment of entrepreneur's income is different from the data: in our model, entrepreneur's income mostly comes from firm's dividend, while in the data, the CEO pay is in most situations categorized in the labor income. A lower ρ is chosen to take it into account. Second, if ρ is high, under the situation that entrepreneurs choose $s_{i,t}$ according to (4), the total value of entrepreneur's risky

assets exceeds the total value of financial assets in the economy. To avoid this, a low ρ should be chosen.

For p_f, we assume that the CEO's average term of office is 20 years. $\ell_{\min}$ is set to unity, which implies that the minimum employment level is one person. We assume that $L = 1.0$ and $N = 0.05$, which implies that the average employment per a firm is 20 persons, which is consistent with the data reported in Davis et al. (2007). Under the settings, the Pareto exponent of the firm size distribution in the model is $1/(1 - 0.05) \approx 1.0526$, which is roughly consistent with Zipf's law. Note that under these parameters, for small-sized firms, the value of an entrepreneur's risky asset calculated by (4) exceeds the value of his firm. To resolve this problem, we assume that such an entrepreneur jointly runs business with other entrepreneurs so that the asset value of the entrepreneurs' risky assets does not exceeds the value of the joint firms. We assume that the productivity shocks of the joint firms move in the same direction. A possible story behind the assumption is that these productivity shocks are caused by managerial decisions.

For the calibration of the firm-level volatility, we consider two cases. In Case A, we use the average firm-level volatility of publicly traded firms. In Case B, we use the average firm-level volatility of both publicly traded and privately held firms. These values are taken from Davis et al. (2007). In each case, the transaction costs of financial intermediaries, ι, is calibrated to match the Pareto exponent in the pre-1975 steady state with 2.4. that is close to the data around 1975.

Insert Table 2 here.

6.3 Computation of the transition dynamics

We compute the Pareto exponent of household's income (or asset) distribution and the top 1% income share before and after 1975. We assume that before 1975 the economy is in the pre-1975 steady state. In our experiment, taxes change unexpectedly and permanently at 1975, and the economy moves toward the post-1975 steady state.

We model the transition dynamics after 1975 in the following way. First, the dynamics of aggregate variables are computed separately. To compute the dynamics of a set of the aggregate variables $\mathbf{S}_t = \left\{ \tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t \right\}$ explained in Section 4.2, we need to pin down their initial values. We suppose that at 1975 when the tax change occurs, the aggregate capital stock is the same as that in the pre-1975 steady state. We also suppose that asset shares of entrepreneurs, innate workers, and former entrepreneurs, $A_{e,1975}/A_{1975}$, $A_{w,1975}/A_{1975}$, $A_{f,1975}/A_{1975}$, are the same as those in the pre-1975 steady state. The remaining initial variables, $\tilde{A}_{1975}$ and $\tilde{H}_{1975}$ are determined by the shooting algorithm by the following steps:

1. Set $\tilde{A}_{1975}$. Set also the upper and lower bound of $\tilde{A}_t$, $\tilde{A}_H$ and $\tilde{A}_L$.
   (a) Set $\tilde{H}_{1975}$ and compute the dynamics of aggregate variables as explained in Section 4.2. Stop the computation if $\tilde{A}_t$ hits the upper or lower bound, $\tilde{A}_H$ or $\tilde{A}_L$.

(b) Update $\tilde{H}_{1975}$ by backwardly solving (26) with the terminal condition

$$\tilde{H}_T = \frac{(1-\alpha)\rho\tilde{p}y^* + \tilde{tr}^*}{\nu + r^{f*} - g},$$

where the variables with asterisks are those in the post-1975 steady state and $T = \arg\min_T \sqrt{(K_T - K^*)^2 + (C_T - C^*)^2}$.

(c) Repeat (a) and (b) until $|\tilde{H}^{\text{new}}_{1975} - \tilde{H}^{\text{old}}_{1975}| < \varepsilon$.

2. Repeat the procedure and find the initial value $\tilde{A}_{1975}$ whose sequence of $\left\{\tilde{K}_t, \tilde{C}_t\right\}$ comes the closest to those at the post-1975 steady state, i.e., whose $\min_T \sqrt{(K_T - K^*)^2 + (C_T - C^*)^2}$ is the smallest.

Note that since $\tilde{C}_t = v\tilde{A}_t$, the above procedure is similar to the standard shooting algorithm used in standard growth models. In computation of the variables used below, we assume that after time T^* when the dynamics of K_t and C_t are in the closest distance to the post-1975 steady state, the economy switches to the post-1975 steady state.

Next, from the aggregate variables calculated above, we compute the variables related to entrepreneur's and worker's asset processes, $\mu_{ae,t}$, $\sigma_{ae,t}$, and $\mu_{aw,t}$. Using these variables we compute the asset (and thus income) distribution at the upper-tail. The transition dynamics of the distribution can be computed by numerically solving the Fokker-Planck equations for the asset distributions of entrepreneurs and workers, $f_e(\ln\tilde{a}_{i,t}, t)$ and $f_l(\ln\tilde{a}_{i,t}, t) \equiv f_w(\ln\tilde{a}_{i,t}, t) + f_f(\ln\tilde{a}_{i,t}, t)$:[4]

$$\begin{aligned}
\frac{\partial f_e(\ln\tilde{a}_{i,t}, t)}{\partial t} &= -\left(\mu_{ae,t} - \frac{\sigma^2_{ae,t}}{2} - g\right)\frac{\partial f_e(\ln\tilde{a}_{i,t}, t)}{\partial \ln\tilde{a}_{i,t}} \\
&\quad + \frac{\sigma^2_{ae,t}}{2}\frac{\partial^2 f_e(\ln\tilde{a}_{i,t}, t)}{\partial(\ln\tilde{a}_{i,t})^2} - (\nu + p_f)f_e(\ln\tilde{a}, t), \\
\frac{\partial f_l(\ln\tilde{a}_{i,t}, t)}{\partial t} &= -(\mu_{aw,t} - g)\frac{\partial f_l(\ln\tilde{a}_{i,t}, t)}{\partial \ln\tilde{a}_{i,t}} - (\nu - p_f)f_e(\ln\tilde{a}, t).
\end{aligned}$$

We impose boundary conditions that $\lim_{\tilde{a}_{i,t}\to\infty} f_i(\ln\tilde{a}_{i,t}, t) = 0$ and that at the lower bound of $\tilde{a}_{i,t}$, $\tilde{a}_{\text{LB}}$, $f_i(\ln\tilde{a}_{\text{LB}}, t)$ moves linearly during the 50 years from that of the pre-1975 steady state to that of the post-1975 steady state.[5]

6.4 Pareto exponent and the top 1% income share

Figures 2 and 3 plot the model predictions of the Pareto exponent and the top 1% share of the income distribution for Case A together with data. Data are taken from Alvaredo et al. (2013). For the model prediction, we plot the two steady states for the pre-1975 and post-1975 periods, and the transition path between them.

We find that the model traces data for the Pareto exponent well. The model also captures the trend in the top 1% share after 1975, although the model's

[4] We use the partial differential equations solver in Matlab. We set the 2000 mesh points to $\ln\tilde{a}_{i,t}$ between $\ln\tilde{a}_{\text{LB}}$ to 100 and 500 mesh points to time t between 1975 to 2030.

[5] $\tilde{a}_{\text{LB}}$ is set to be higher than $\tilde{h}$ at the pre- and post-1975 steady state.

prediction is somewhat lower in the level than data. Perhaps, other factors such as the differences in talents account for the gap between them.

The corresponding results for Case B are graphed in Figures 4 and 5. The model's transitions of the Pareto exponent and the top 1% share become slower than those in Case A. The reason is that in Case B, firm's volatility becomes higher. This makes $x_{e,t}$ lower by (4), which results in lower volatility of entrepreneur's asset. This perhaps implies that the lower firm-level volatility at the top firms where the richest CEOs are employed is an important factor to understand the evolution of top incomes.

Insert Figure 2 here.
Insert Figure 3 here.
Insert Figure 4 here.
Insert Figure 5 here.

6.5 Incentive pay for CEOs

The reason for the evolution of top incomes after the tax change in the model is that it becomes more profitable for CEOs to hold risky assets. It means that entrepreneur's portfolio share of risky assets, $x_{e,t}$, increases in the post-1975 periods, possibly through utilizing employee stock options. Here, we compare $x_{e,t}$ in the model with the empirical counterpart of $x_{e,t}$.

An empirical counterpart of $x_{e,t}$ for corporate CEOs is called as "percent-percent" incentives, which is defined by

$$\frac{x\% \text{ increase in pay}}{1\% \text{ increase in firm rate of return}}.$$

The concept of "percent-percent" measure is used by Murphy (1985), Gibbons and Murphy (1992), Rosen (1992), and Edmans et al. (2009).[6]

We plot the "percent-percent" incentives constructed from Frydman and Saks (2010) and $x_{e,t}$ in the model in Figure 6.[7] We confirm that the data and model are in the same order. Of course, our model is not intended to explain the fluctuations in the "percent-percent" incentive itself, and the model cannot explain why the "percent-percent" incentives increase around the late 1950s. Further research is needed to understand the empirical facts.

Insert Figure 6 here.

[6] Edmans et al. (2009) also argue that the "percent-percent" incentives are cross-sectionally independent of the firm size. This property is satisfied in our model.

[7] The "percent-percent" data are calculated by dividing "dollar change in wealth for a 1% increase in firm rate of return" by "total compensation," both of which are taken from Figures 5 and 6 of Frydman and Saks (2010).

7 Conclusion

We have proposed a model of asset and income inequalities that explains Zipf's law of firms and Pareto's law of incomes both from the same productivity shocks of firms. Empirical studies show that the Pareto exponent of income varies over time while the Zipf's law of firm size is quite stable. This paper consistently explains these distributions with an analytically tractable model. We derive closed-form expressions for the stationary distributions of firm size and individual income. The transition dynamics of those distributions are also explicitly derived, which is then used for numerical analysis.

Our model features an entrepreneur who can invest in their own firms as well as risk-free assets. The entrepreneur incurs a substantial transaction cost if he diversifies the risk of his portfolio returns. When a tax on risky returns is reduced, the entrepreneur increases the share of his own firms. This in turn increases the variance of his portfolio returns, which results in a wider dispersion of wealth among entrepreneurs.

By calibrating the model, we have analyzed to what extent the changes in tax rates account for the recent evolution of top incomes in the U.S. We find that the model matches with the decline in the Pareto exponent of income distribution and the trend in top 1% share. There remain some discrepancies between the model and data. For example, model's prediction of top 1% share is somewhat lower than the data. Further research is needed for understanding the causes of discrepancies.

References

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez (2013) "The World Top Incomes Database." Data are extracted at 28/05/2013.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011) "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, Vol. 79, No. 1, pp. 123–157.

——— (2012) "The Distribution of Wealth in the Blanchard-Yaari Model," mimeo.

Blanchard, Olivier J (1985) "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, Vol. 93, No. 2, pp. 223–47.

Campbell, John Y. and Luis M. Viceira (2002) *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*: Oxford University Press.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda (2007) "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," in *NBER Macroeconomics Annual 2006, Volume 21*: National Bureau of Economic Research, Inc, pp. 107–180.

Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009) "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies*, Vol. 22, No. 12, pp. 4881–4917.

Frydman, Carola and Raven E. Saks (2010) "Executive Compensation: A New View from a Long-Term Perspective, 1936–2005," *Review of Financial Studies*, Vol. 23, No. 5, pp. 2099–2138.

Gabaix, Xavier and Augustin Landier (2008) "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, Vol. 123, No. 1, pp. 49–100.

Gibbons, Robert and Kevin J Murphy (1992) "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, Vol. 100, No. 3, pp. 468–505.

Hacker, Jacob S. and Paul Pierson (2010) *Winner-Take-All Politics: How Washington Made the Rich Richer–and Turned Its Back on the Middle Class*: Simon and Schuster.

Jones, Charles I. and Jihee Kim (2012) "Exploring the Dynamics of Top Income Inequality," mimeo.

Katz, Lawrence F and Kevin M Murphy (1992) "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, Vol. 107, No. 1, pp. 35–78.

Kim, Jihee (2013) "The Effect of the Top Marginal Tax Rate on Top Income Inequality," mimeo.

Luttmer, Erzo G.J. (2010) "Models of Growth and Firm Heterogeneity," *Annual Review of Economics*, Vol. 2, No. 1, pp. 547–576.

McGrattan, Ellen R. and Edward C. Prescott (2005) "Taxes, Regulations, and the Value of U.S. and U.K. Corporations," *Review of Economic Studies*, Vol. 72, No. 3, pp. 767–796.

Merton, Robert C. (1969) "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics*, Vol. 51, No. 3, pp. 247–57.

——— (1971) "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, Vol. 3, No. 4, pp. 373–413.

——— (1973) "An Intertemporal Capital Asset Pricing Model," *Econometrica*, Vol. 41, No. 5, pp. 867–87.

Mizuno, Takayuki, Atushi Ishikawa, Shouji Fujimoto, and Tsutomu Watanabe (2012) "Power Laws in Firm Productivity," *Progress of Theoretical Physics Supplement*, Vol. 194, pp. 122-134.

Murphy, Kevin J. (1985) "Corporate Performance and Managerial Remuneration: An Empirical Analysis," *Journal of Accounting and Economics*, Vol. 7, No. 1–3, pp. 11–42.

Nirei, Makoto (2009) "Pareto Distributions in Economic Growth Models," IIR Working Paper 09-05, Institute of Innovation Research, Hitotsubashi University.

Nirei, Makoto and Wataru Souma (2007) "A Two Factor Model Of Income Distribution Dynamics," *Review of Income and Wealth*, Vol. 53, No. 3, pp. 440–459.

Piketty, Thomas and Emmanuel Saez (2003) "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 1–39.

Rosen, Sherwin (1992) "Contracts and the Markets for Executives," in Werin, Lars and Hans Wijkander eds. *Main Currents in Contract Economics*: Blackwell Press.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz (2012) "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, Vol. 50, No. 1, pp. 3–50.

Sinn, Hans-Werner (1991) "Taxation and the Cost of Capital: The 'Old' View, the 'New' View, and Another View," in *Tax Policy and the Economy, Volume 5*: National Bureau of Economic Research, Inc, pp. 25–54.

Toda, Alexis Akira (2012) "Asset Pricing and Wealth Distribution with Heterogeneous Investment Returns," mimeo.

A Derivations for the household's problem

This appendix shows the derivations of the household problem in Section 2.1. As shown in Section 4.2, the aggregate dynamics of the model is described by $\mathbf{S}_t$, whose evolution can be written as

$$d\mathbf{S}_t = \mu_{\mathbf{S}}(\mathbf{S}_t)dt.$$

By Ito's formula, $V^i(a_{i,t}, \mathbf{S}_t, t)$ is rewritten as follows:

$$\begin{aligned} dV^i(a_{i,t}, \mathbf{S}_t, t) =& \frac{\partial V_t^i}{\partial t}dt + \frac{\partial V_t^i}{\partial a_{i,t}}da_{i,t} + \frac{1}{2}\frac{\partial^2 V_t^i}{\partial a_{i,t}^2}(da_{i,t})^2 + \frac{\partial V_t^i}{\partial \mathbf{S}_t}\cdot d\mathbf{S}_t \\ &+ \left(V^{i'}(a_{i,t}, \mathbf{S}_t, t) - V^i(a_{i,t}, \mathbf{S}_t, t)\right) dJ_{i,t}, \end{aligned}$$

where $J_{i,t}$ is the Poisson jump process describing the probability of leaving his firm:

$$dJ_{i,t} = \begin{cases} 0 & \text{with probability } 1 - p_f dt \\ 1 & \text{with probability } p_f dt. \end{cases}$$

Thus,

$$\frac{\mathsf{E}_t[dV_t^i]}{dt} = \frac{\partial V_t^i}{\partial t} + \mu_{a,t}a_{i,t}\frac{\partial V_t^i}{\partial a_{i,t}} + \frac{(\sigma_{a,t}a_{i,t})^2}{2}\frac{\partial^2 V_t^i}{\partial a_{i,t}^2} + \mu'_{\mathbf{S}}(\mathbf{S}_t)\cdot\frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f\left(V_t^{i'} - V_t^i\right)$$

Substituting in (3), we obtain a Hamilton-Jacobi-Bellman equation:

$$
\begin{aligned}
0 = & \max_{c_{i,t},x_{i,t}} \ln c_{i,t} - (\beta+\nu)V_t^i + \frac{\partial V_t^i}{\partial t} + \mu_{a,t}a_{i,t}\frac{\partial V_t^i}{\partial a_{i,t}} + \frac{(\sigma_{a,t}a_{i,t})^2}{2}\frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\
& + \mu'_{\mathbf{S}}(\mathbf{S}_t)\cdot\frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f\left(V_t^{i'} - V_t^i\right) \\
= & \max_{c_{i,t},x_{i,t}} \ln c_{i,t} - (\beta+\nu)V_t^i + \frac{\partial V_t^i}{\partial t} + \frac{\sigma_{q,t}^2}{2}x_{i,t}^2 a_{i,t}^2\frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\
& + ((\nu+\mu_{q,t})x_{i,t}a_{i,t} + (\nu+r_t^f)(1-x_{i,t})a_{i,t} - c_{i,t})\frac{\partial V_t^i}{\partial a_{i,t}}, \\
& + \mu'_{\mathbf{S}}(\mathbf{S}_t)\cdot\frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f\left\{V_t^{i'} - V_t^i\right\}. \qquad (27)
\end{aligned}
$$

First-order conditions with respect to $c_{i,t}$ and $x_{i,t}$ are summarized by:

$$
\begin{aligned}
c_{i,t}^{-1} &= \frac{\partial V_t^i}{\partial a_{i,t}}, \qquad (28) \\
x_{i,t} &= -\frac{\partial V_t^i/\partial a_{i,t}}{\left(\partial^2 V_t^i/\partial a_{i,t}^2\right)a_{i,t}}\frac{\mu_{q,t}-r_t^f}{\sigma_{q,t}^2}. \qquad (29)
\end{aligned}
$$

Following Merton (1969) and Merton (1971), this problem is solved by the following value function and linear policy functions:

$$
\begin{aligned}
V_t^i &= B_t^i \ln a_{i,t} + H(\mathbf{S}_t, t), \\
c_{i,t} &= v_{i,t}a_{i,t}, \\
q_{i,t}s_{i,t} &= x_{i,t}a_{i,t}, \\
b_{i,t} &= (1-x_{i,t})a_{i,t} - h_t.
\end{aligned}
$$

We obtain this solution by guess-and-verify. The first-order condition (28) becomes:

$$(v_{i,t})^{-1} = B_t^i. \qquad (30)$$

Condition (29) is rewritten as:

$$x_{i,t} = \frac{\mu_{q,t}-r_t^f}{\sigma_{q,t}^2}. \qquad (31)$$

Substituting these results into (27), we find that

$$v_{i,t} = \nu + \beta. \qquad (32)$$

B Derivations for the firm's problem

B.1 Derivations of the FOCs of firm's problem

This appendix shows the derivations of firm's problem at Section 2.2.2. $q_{j,t}$ is a function of $k_{j,t}$, $z_{j,t}$, and the aggregate dynamics $\mathbf{S}_t$ (see Appendix A). By

applying Ito's formula to $q_{j,t}$, we obtain

$$\begin{aligned}
dq(k_{j,t}, z_{j,t}, \mathbf{S}_t, t) &= \left(\frac{\partial q_{j,t}}{\partial t} dt + \frac{\partial q_{j,t}}{\partial z_{j,t}} dz_{j,t} + \frac{\partial q_{j,t}}{\partial k_{j,t}} dk_{j,t} + \frac{\partial q_{j,t}}{\partial \mathbf{S}_t} \cdot d\mathbf{S}_t \right) + \frac{1}{2} \frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2} (dz_{j,t})^2 \\
&= \left(\frac{\partial q_{j,t}}{\partial t} + \mu_z \frac{\partial q_{j,t}}{\partial z_{j,t}} + \frac{1}{2} \sigma_z^2 \frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2} \right) dt + \frac{\partial q_{j,t}}{\partial k_{j,t}} dk_{j,t} + \mu'_{\mathbf{S}}(\mathbf{S}_t) \cdot \frac{\partial q_{j,t}}{\partial \mathbf{S}_t} \\
&\quad + \sigma_z \frac{\partial q_{j,t}}{\partial z_{j,t}} dB_{j,t}.
\end{aligned}$$

The FOCs for $\ell_{j,t}$ and $dk_{j,t}$ are

$$\begin{aligned}
(1 - \tau^f - \iota) &= \frac{\partial q_{j,t}}{\partial k_{j,t}}, \\
w_t &= \frac{\partial p_{j,t} y_{j,t}}{\partial \ell_{j,t}}.
\end{aligned}$$

By the envelope theorem,

$$r_t^f \frac{\partial q_{j,t}}{\partial k_{j,t}} dt = (1 - \tau^f - \iota) \left(\frac{\partial p_{j,t} y_{j,t}}{\partial k_{j,t}} dt - \delta dt \right).$$

By rearranging the equation, we obtain

$$r_t^f = \frac{\partial p_{j,t} y_{j,t}}{\partial k_{j,t}} - \delta.$$

B.2 Derivations on the firm side variables

This appendix shows the derivations of the firm side variables at Section 3.1. From (10)

$$w_t = (1 - \alpha)\rho \left(\frac{Y_t}{N} \right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho} \ell_{j,t}^{(1-\alpha)\rho - 1}.$$

Rewriting this,

$$\ell_{j,t} = \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N} \right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho} \right)^{\frac{1}{1-(1-\alpha)\rho}}. \tag{33}$$

On the other hand, from (9),

$$\mathsf{MPK}_t = \alpha\rho \left(\frac{Y_t}{N} \right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho - 1} \ell_{j,t}^{(1-\alpha)\rho}. \tag{34}$$

By substituting (33) into (34) and rearranging,

$$k_{j,t}^{\frac{\alpha\rho}{1-(1-\alpha)\rho}} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N} \right)^{1-\rho} \right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N} \right)^{1-\rho} \right)^{\frac{\alpha\rho(1-\alpha)\rho}{(1-\rho)(1-(1-\alpha)\rho)}} z_{j,t}^{\phi \frac{\alpha\rho}{1-\rho}}, \tag{35}$$

where $\phi \equiv \frac{\rho}{1-(1-\alpha)\rho}$. Substituting (35) into (33),

$$\ell_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{1-\alpha\rho}{1-\rho}} z_{j,t}^{\frac{\rho}{1-\rho}} \tag{36}$$

By substituting this equation into the labor market condition (13) and rearranging,

$$\left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{1-\alpha\rho}{1-\rho}} = \frac{L}{N}\frac{1}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}, \tag{37}$$

$$\text{or, } \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{(1-\alpha)\rho}{1-\rho}} = \left\{\left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{-\alpha\rho}{1-\rho}} \frac{L}{N}\frac{1}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}\right\}^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}. \tag{38}$$

where $\overline{\mathsf{E}}$ is the operator of the cross-sectional average of all firms. Then, substituting (37) into (36),

$$\ell_{j,t} = \frac{L}{N}\left(\frac{z_{j,t}^{\frac{\rho}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}\right). \tag{39}$$

Rewriting (35),

$$k_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{1-(1-\alpha)\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{(1-\alpha)\rho}{1-\rho}} z_{j,t}^{\frac{\rho}{1-\rho}}. \tag{40}$$

Substituting (38) into (40),

$$k_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{1}{1-\alpha\rho}} \left(\frac{L}{N}\right)^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} \left(\frac{z_{j,t}^{\frac{\rho}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}}\right). \tag{41}$$

Next, we derive Y. Substituting (39) and (41) into $y_{j,t} = z_{j,t}k_{j,t}^{\alpha}\ell_{j,t}^{1-\alpha}$, and rearranging,

$$y_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{\alpha}{1-\alpha\rho}} \left(\frac{L}{N}\right)^{\frac{1-\alpha}{1-\alpha\rho}} \left(\frac{z_{j,t}^{\frac{1}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{(1-\alpha)}{1-\alpha\rho}}}\right).$$

Substituting this equation into $Y = \left(\int_0^N \left(\frac{1}{N}\right)^{1-\rho} y_{j,t}^{\rho} dj\right)^{\frac{1}{\rho}}$,

$$\left(\frac{Y_t}{N}\right)^{1-\rho} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha(1-\rho)}{1-\alpha}} \left(\frac{L}{N}\right)^{1-\rho} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{(1-\rho)\left[\frac{1-\alpha\rho}{(1-\alpha)\rho}-1\right]}. \tag{42}$$

Substituting (42) into (41),

$$k_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}} \left(\frac{L}{N}\right)\left(\frac{z_{j,t}^{\frac{\rho}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}\right)$$

$$= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}} \ell_{j,t}. \tag{43}$$

Substituting (39) and (43) into (42)

$$p_{j,t}y_{j,t} = Y_t^{1-\rho} y_{j,t}^{\rho} \tag{44}$$

$$= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}} \left(\frac{L}{N}\right)\left(\frac{z_{j,t}^{\frac{\rho}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}\right)$$

$$= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}} \ell_{j,t}. \tag{45}$$

Rewriting (39),

$$\ell_{j,t} = \overline{\ell}_t \; z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{\ell}_t \equiv \left(\frac{L/N}{\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}}\right).$$

Rewriting (45),

$$p_{j,t}y_{j,t} = \overline{py}_t \overline{\ell}_t \; z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{py}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}\frac{1}{1-\alpha}}.$$

Rewriting (43),

$$k_{j,t} = \overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{k}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}}\right)^{\frac{1}{1-\alpha}}. \tag{46}$$

From (46),

$$dk_{j,t} = d(\overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}})$$
$$= \frac{d\overline{k}_t \overline{\ell}_t}{dt} z_{j,t}^{\frac{\rho}{1-\rho}} dt + \overline{k}_t \overline{\ell}_t d\left(z_{j,t}^{\frac{\rho}{1-\rho}}\right).$$

Note that

$$d\left(z_{j,t}^{\frac{\rho}{1-\rho}}\right) = \left\{\left(\frac{\rho}{1-\rho}\right)\mu_z + \left(\frac{\rho}{1-\rho}\right)\left(\frac{\rho}{1-\rho} - 1\right)\frac{\sigma_z^2}{2}\right\} z_{j,t}^{\frac{\rho}{1-\rho}} dt + \left(\frac{\rho}{1-\rho}\right)\sigma_z z_{j,t}^{\frac{\rho}{1-\rho}} dB_{j,t}.$$

Then,

$$dk_{j,t} = d(\overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}})$$
$$= \frac{d\overline{k}_t \overline{\ell}_t}{dt} z_{j,t}^{\frac{\rho}{1-\rho}} dt + \overline{k}_t \overline{\ell}_t d\left(z_{j,t}^{\frac{\rho}{1-\rho}}\right)$$
$$= k_{j,t}\left\{\mu_{k,t}dt + \left(\frac{\rho}{1-\rho}\right)\sigma_z dB_{j,t}\right\}.$$

where

$$\mu_{k,t} \equiv g - \frac{1}{1-\alpha}\frac{dr_t^f/dt}{\mathsf{MPK}_t} + \left(\frac{\rho}{1-\rho}\right)\left\{(\mu_z - g_z) + \left(\frac{\rho}{1-\rho} - 1\right)\frac{\sigma_z^2}{2}\right\}.$$

g_z is the growth rate of $\overline{\mathsf{E}}\left\{z_{j,t}^{\frac{\rho}{1-\rho}}\right\}$ and $g = g_z/(1-\alpha)$.

$d_{j,t}dt$ is computed by substituting these results into the following relationship:

$$\begin{aligned} d_{j,t}dt =& (p_{j,t}y_{j,t} - w_t\ell_{j,t} - \delta k_{j,t})dt - dk_{j,t} \\ =& (1-(1-\alpha)\rho)p_{j,t}y_{j,t}dt - \delta k_{j,t}dt - dk_{j,t}. \end{aligned}$$

Then, $d_{j,t}dt$ is rewritten as follows:

$$\begin{aligned} d_{j,t}dt =& \overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}dt - \left\{\left(\frac{\rho}{1-\rho}\right)\sigma_z dB_{j,t}\right\}\overline{k}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \\ \text{where } \overline{d}_t \equiv& (1-(1-\alpha)\rho)\overline{py}_t - (\delta + \mu_{k,t})\,\overline{k}_t. \end{aligned}$$

B.3 Returns on risky assets

This appendix explains the derivation of the returns on risky assets at Sections 3.1 and 4.2. Multiplying (6) by $e^{-\int_t^u r_s^f ds}$ and integrating[8], we obtain

$$q_{j,t} = \mathsf{E}_t\left[\int_t^\infty (1-\tau^f-\iota)d_{j,u}e^{-\int_t^u r_s^f ds}du\right].$$

By further rearranging the above equation,

$$q_{j,t} = \int_t^\infty (1-\tau^f-\iota)e^{-\int_t^u r_s^f ds}\,\mathsf{E}_t\left[d_{j,u}\right]du.$$

Because

$$\begin{aligned} \mathsf{E}_t[d_{j,u}] =& \overline{d}_u\overline{\ell}_u\,\mathsf{E}_t[z_{j,u}^{\frac{\rho}{1-\rho}}] \\ =& \overline{d}_t\overline{\ell}_t\frac{\overline{d}_u\overline{\ell}_u}{\overline{d}_t\overline{\ell}_t} \times \exp\left\{\int_t^u\left(\left(\frac{\rho}{1-\rho}\right)\mu_z + \left(\frac{\rho}{1-\rho}\right)\left(\frac{\rho}{1-\rho}-1\right)\frac{\sigma_z^2}{2}\right)ds\right\}\cdot z_{j,t}^{\frac{\rho}{1-\rho}} \\ =& \overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}\exp\left\{\int_t^u\left(\frac{d\ln(\overline{d}_s\overline{\ell}_s)}{ds} + \left(\frac{\rho}{1-\rho}\right)\mu_z + \left(\frac{\rho}{1-\rho}\right)\left(\frac{\rho}{1-\rho}-1\right)\frac{\sigma_z^2}{2}\right)ds\right\} \\ \equiv& \overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}\exp\left\{\int_t^u \mu_{d,s}ds\right\}. \end{aligned}$$

[8]The Ito process version of integration by parts

$$\int_t^T X_{j,s}dY_{j,s} = X_{j,T}Y_{j,T} - X_{j,t}Y_{j,t} - \int_t^T Y_{j,s}dX_{j,s} - \int_t^T dX_{j,s}dY_{j,s}.$$

is used here. Define $\Delta_{t,u} \equiv e^{-\int_t^u r_s^f ds}$. Then,

$$\int_t^\infty \Delta_{t,u}dq_{j,u} = q_{j,u}\Delta_{t,u}\,|_t^\infty - \int_t^\infty q_{j,u}(-r_u^f)\Delta_{t,u}du$$

Therefore,

$$q_{j,t} = \overline{q}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \quad \text{where } \overline{q}_t \equiv (1-\tau^f-\iota)\overline{d}_t \int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\} du.$$

Then,

$$\begin{aligned} dq_{j,t} =& q_{j,t}\frac{d\ln(\overline{d}_t\overline{\ell}_t)}{dt}dt + q_{j,t}\frac{d(z_{j,t}^{\frac{\rho}{1-\rho}})}{z_{j,t}^{\frac{\rho}{1-\rho}}} + q_{j,t}\frac{-1+(r_t^f-\mu_{d,t})\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}dt \\ =& \left\{-(1-\tau^f-\iota)\overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}} + r_t^f q_{j,t}\right\}dt + q_{j,t}\left(\frac{\rho}{1-\rho}\right)\sigma_z dB_{j,t}. \end{aligned}$$

Using $d_{j,t} = \overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}$, the return of a risky asset is

$$\begin{aligned} \frac{(1-\tau^e)d_{j,t}+dq_{j,t}}{q_{j,t}} =& \left\{\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}-1\right)\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du} + r_t^f\right\}dt \\ &+ \left(\frac{\rho}{1-\rho}\right)\sigma_z\left\{1-\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}\right)\frac{\overline{k}_t}{\overline{d}_t}\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}\right\}dB_{j,t}. \end{aligned} \tag{47}$$

Note that if $(r_t^f - \mu_{d,t})$ is constant as in the steady state, $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du = 1/(r^f-\mu_d)$ and

$$q_{j,t} = \frac{(1-\tau^f-\iota)\overline{d}_t\overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}}{r^f-\mu_d}.$$

We need to know the value of $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du$ to compute the return on risky assets. We calculated the value as follows. Integrating (19), we obtain

$$\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du = \frac{Q_t}{(1-\tau^f-\iota)\overline{d}_t L}.$$

If we know the value of $Q_t = A_t - H_t$ and $(1-\tau^f-\iota)\overline{d}_t L$, we can calculate the value of $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du$.

B.4 Derivations on the restructuring

This appendix shows the derivations of the restructuring at Section 3.2. Let $\tilde{z}_{j,t}$ be the firm's productivity level after selling a part of the firm's assets to restructuring firms, detrended by $e^{g_z t}$. Then, $Q_{\text{restructuring},t+dt}$ is written as follows:

$$Q_{\text{restructuring},t+dt} = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}e^{g_z t}\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt} \le \tilde{z}_{\min}\right\}.$$

where $\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt} \le \tilde{z}_{\min}\right\}$ is the expectation of $\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}$ conditional on that $\tilde{z}_{j,t+dt}$ is lower than $\tilde{z}_{\min}$. Since the evolution of $\tilde{z}_{j,t}$ follows

(22) and the distribution follows (23),

$$\begin{aligned}\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt}\leq\tilde{z}_{\min}\right\}&=\int_{\ln\tilde{z}_{\min}}^{\infty}d(\ln\tilde{z}_{j,t})\int_{-\infty}^{\ln\tilde{z}_{\min}}d(\ln\tilde{z}_{j,t+dt})\\&\quad\left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\right)f_z(\ln\tilde{z}_{j,t})f_z(\ln\tilde{z}_{j,t+dt}|\ln\tilde{z}_{j,t})\\&=\int_{\ln\tilde{z}_{\min}}^{\infty}d(\ln\tilde{z}_{j,t})\int_{-\infty}^{\ln\tilde{z}_{\min}}d(\ln\tilde{z}_{j,t+dt})\\&\quad\left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\right)C_0e^{-\lambda\ln\tilde{z}_{j,t}}\\&\quad\times\frac{1}{\sqrt{2\pi\sigma_z^2dt}}e^{-\frac{\left(\ln\tilde{z}_{j,t+dt}-(\ln\tilde{z}_{j,t}+\tilde{\mu}_zdt)\right)^2}{2\sigma_z^2dt}},\end{aligned}$$

where $\tilde{\mu}_z\equiv\mu_z-g_z-\sigma_z^2/2-m$, $f_z(\ln\tilde{z}_{j,t+dt}|\ln\tilde{z}_{j,t})$ is the distribution of $\ln\tilde{z}_{j,t+dt}$ conditional on $\ln\tilde{z}_{j,t}$, which follows a normal distribution, and $f_z(\ln\tilde{z}_{j,t})$ is the steady state firm size distribution.

Under the setup, taking the limit as dt approaches zero from above, (20) becomes

$$\begin{aligned}\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\}m\left(\frac{\rho}{1-\rho}\right)&=\lim_{dt\to0+}\frac{\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt}\leq\tilde{z}_{\min}\right\}}{dt}\\&=\lim_{t'\to0+}\frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}}{dt'}.\end{aligned}\tag{48}$$

$d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}/dt'$ can be further calculated:

$$\begin{aligned}&\frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}}{dt'}\\&=\int_{\ln\tilde{z}_{\min}}^{\infty}d(\ln\tilde{z}_{j,t})\int_{-\infty}^{\ln\tilde{z}_{\min}}d(\ln\tilde{z}_{j,t+t'})\\&\quad\frac{d}{dt'}\left(\left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\right)C_0e^{-\lambda\ln\tilde{z}_{j,t}}\frac{1}{\sqrt{2\pi\sigma_z^2t'}}e^{-\frac{\left(\ln\tilde{z}_{j,t+t'}-(\ln\tilde{z}_{j,t}+\tilde{\mu}_zt')\right)^2}{2\sigma_z^2t'}}\right)\\&=\frac{C_0e^{-\left(\lambda-\left(\frac{\rho}{1-\rho}\right)\right)\ln\tilde{z}_{\min}}\left(\frac{\rho}{1-\rho}\right)}{4\left(\lambda-\left(\frac{\rho}{1-\rho}\right)\right)}\\&\quad\times\left(e^{\frac{1}{2}\lambda(2\tilde{\mu}_z+\lambda\sigma_z^2)t'}(2\tilde{\mu}_z+\lambda\sigma_z^2)\operatorname{Erfc}\left[\frac{(\tilde{\mu}_z+\lambda\sigma_z^2)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right.\\&\quad\left.-e^{\frac{1}{2}\left(\frac{\rho}{1-\rho}\right)\left(2\tilde{\mu}_z+\left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right)t'}\left(2\tilde{\mu}_z+\left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right)\operatorname{Erfc}\left[\frac{\left(\tilde{\mu}_z+\left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right).\end{aligned}$$

By combining these results and taking the limit, we obtain

$$\lim_{t'\to0+}\frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}}-\tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}}{dt'}=\frac{1}{4}C_0e^{-\left(\lambda-\left(\frac{\rho}{1-\rho}\right)\right)\ln\tilde{z}_{\min}}\left(\frac{\rho}{1-\rho}\right)\sigma_z^2$$

Substituting this result into (48), we finally obtain

$$m = \left(\lambda - \frac{\rho}{1-\rho}\right)\frac{\sigma_z^2}{4}.$$

C Derivations on Household's Asset Distributions in the Steady State

This appendix shows the derivations of the household asset distributions at Section 5.

C.1 Derivations on the asset distribution of entrepreneurs

Discussion in Section 5.1 indicates that the probability density function of entrepreneurs at age t' whose detrended log wealth level is $\ln \tilde{a}_i$ is

$$f_e(\ln \tilde{a}_i|t') = \frac{1}{\sqrt{2\pi\sigma_{ae}^2 t'}} \exp\left(-\frac{(\ln \tilde{a}_i - (\ln \tilde{h} + (\mu_{ae} - g - \sigma_{ae}^2/2)t'))^2}{2\sigma_{ae}^2 t'}\right).$$

The probability density of entrepreneurs whose age is t' is

$$f_e(t') = \frac{(\nu + p_f)N}{L} \exp\left(-(\nu + p_f)t'\right).$$

By combining them, we can calculate the probability density function of entrepreneur's asset distribution, $f_e(\ln \tilde{a}_i)$, by

$$f_e(\ln \tilde{a}_i) = \int_0^\infty dt' \; f_e(t') f_e(\ln \tilde{a}_i|t'),$$

To derive $f_e(\ln \tilde{a}_i)$ in Section 5.1, we apply to the above equation the following formula:

$$\int_0^\infty \exp(-at - b^2/t)/\sqrt{t}dt = \sqrt{\pi/a}\exp(-2b\sqrt{a}), \quad \text{for } a > 0, b > 0.$$

C.2 Derivations on the asset distribution of innate workers

The asset distribution of innate workers is calculated as follows:

$$\begin{aligned}
f_w(\ln \tilde{a}_i) &= \int_0^\infty dt' \; f_w(t') f_w(\ln \tilde{a}_i|t') \\
&= \int_0^\infty dt' \; \frac{\nu L - (\nu + p_f)N}{L} \exp(-\nu t') \cdot \mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{aw} - g)t') \\
&= \int_{\ln \tilde{h}}^{\ln \tilde{h} + (\mu_{aw} - g)\infty} \frac{dt'}{d\ln \tilde{a}_i} d(\ln \tilde{a}_i) \; \cdot \frac{\nu L - (\nu + p_f)N}{L} \exp\left(-\frac{\nu}{\mu_{aw} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) \\
&\quad \times \mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{aw} - g)t') \\
&= \begin{cases} \frac{\nu L - (\nu + p_f)N}{L} \frac{1}{|\mu_{aw} - g|} \exp\left(-\frac{\nu}{\mu_{aw} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) & \text{if } \frac{\ln \tilde{a}_i - \ln \tilde{h}}{\mu_{aw} - g} \geq 0, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Note that $\mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{aw} - g)t')$ is a unit function that takes 1 if $\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{aw} - g)t')$ and 0 otherwise.

C.3 Derivations on the asset distribution of former entrepreneurs

The asset distribution of former entrepreneurs is derived as follows. Let $t'_m \equiv (\ln \tilde{a}_i - \ln \tilde{h})/(\mu_{aw} - g)$. First, we consider the case where $\mu_{aw} \geq g$. If $\ln \tilde{a}_i \geq \ln \tilde{h}$, then

$$\begin{aligned}
f_f(\ln \tilde{a}_i) = &\int_0^{t'_m} dt' \; p_f f_{e1}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
&+ \int_{t'_m}^{\infty} dt' \; p_f f_{e2}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
= &\left[\frac{-p_f}{\nu - \psi_1(\mu_{aw} - g)} f_{e1}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \right]_0^{t'_m} \\
&+ \left[\frac{-p_f}{\nu + \psi_2(\mu_{aw} - g)} f_{e2}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \right]_{t'_m}^{\infty} \\
= &\frac{p_f}{\nu - \psi_1(\mu_{aw} - g)} \left\{ -f_{e1}(\ln \tilde{a}_i - (\mu_{aw} - g)t'_m) \times \exp(-\nu t'_m) + f_{e1}(\ln \tilde{a}_i) \right\} \\
&+ \frac{p_f}{\nu + \psi_2(\mu_{aw} - g)} \left\{ -0 + f_{e2}(\ln \tilde{a}_i - (\mu_{aw} - g)t'_m) \times \exp(-\nu t'_m) \right\}.
\end{aligned}$$

By substituting into the above equation the following relations: $\ln \tilde{a}_i - (\mu_{aw} - g)t'_m = \ln \tilde{h}$, $f_{e1}(\ln \tilde{h}) = f_{e2}(\ln \tilde{h})$, and $t'_m = (\ln \tilde{a}_i - \ln \tilde{h})/(\mu_{aw} - g)$, we obtain

$$\begin{aligned}
f_f(\ln \tilde{a}_i) = &\frac{p_f}{\nu - \psi_1(\mu_{aw} - g)} f_{e1}(\ln \tilde{a}_i) \\
&- \left(\frac{1}{\nu - \psi_1(\mu_{aw} - g)} - \frac{1}{\nu + \psi_2(\mu_{aw} - g)} \right) p_f f_{e1}(\ln \tilde{h}) \\
&\times \exp\left(-\frac{\nu}{\mu_{aw} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right).
\end{aligned}$$

If $\ln \tilde{a}_i < \ln \tilde{h}$,

$$\begin{aligned}
f_f(\ln \tilde{a}_i) = &\int_0^{\infty} dt' p_f f_{e2}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
= &\frac{p_f}{\nu + \psi_2(\mu_{aw} - g)} f_{e2}(\ln \tilde{a}_i).
\end{aligned}$$

Next, we consider the case where $\mu_{aw} < g$. If $\ln \tilde{a}_i \geq \ln \tilde{h}$, then

$$\begin{aligned}
f_f(\ln \tilde{a}_i) = &\int_0^{\infty} dt' p_f f_{e1}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
= &\frac{p_f}{\nu - \psi_1(\mu_{aw} - g)} f_{e1}(\ln \tilde{a}_i).
\end{aligned}$$

If $\ln \tilde{a}_i < \ln \tilde{h}$,

$$
\begin{aligned}
f_f(\ln \tilde{a}_i) = & \int_0^{t'_m} dt' p_f f_{e2}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
& + \int_{t'_m}^{\infty} dt p_f f_{e1}(\ln \tilde{a}_i - (\mu_{aw} - g)t') \times \exp(-\nu t') \\
& = \frac{p_f}{\nu + \psi_2(\mu_{aw} - g)} f_{e2}(\ln \tilde{a}_i) \\
& - \left(\frac{1}{\nu + \psi_2(\mu_{aw} - g)} - \frac{1}{\nu - \psi_1(\mu_{aw} - g)} \right) p_f f_{e1}(\ln \tilde{h}) \\
& \times \exp\left(-\frac{\nu}{\mu_{aw} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right).
\end{aligned}
$$

	pre-1975	post-1975
ordinary income tax, τ^{ord}	0.75	0.40
corporate income tax, τ^{corp}	0.50	0.35
capital gain tax, τ^{cap}	0.25	0.25
τ^e	0.75	0.40
τ^f	0.63	0.51

Table 1: Tax rates

Notes: The figures of those in the upper half of the Table are calibrated from the top statutory marginal federal tax rates in Figure 1, which is taken from Saez et al. (2012). The tax rate on risky assets, τ^e, is set to be equal to τ^{ord}. The tax rate on risk-free assets, τ^f, is calculated by $1-(1-\tau^{\text{cap}})(1-\tau^{\text{corp}})$,

β	discount rate	0.04
ν	prob. of death	1/50
α	capital share	1/3
δ	depreciation rate	0.1
g	steady state growth rate	0.02
ρ	elasticity of substitution	0.7
p_f	prob. of entrepreneur's quit	1/20
$\ell_{\min}$	min. level of employment	1
L	fraction of population	1.0
N	fraction of employees	0.05

		Case A	Case B
$\left(\frac{\rho}{1-\rho}\right)\sigma_z$	firm-level vol. of employment	0.25	0.45
ι	transaction costs of fin. intermed.	0.215	0.243

Table 2: Calibrated parameters

Notes: The figures on the firm-level volatility of employment are taken from Figure 2.6 of Davis et al. (2007). Case A corresponds to the case where the firm-level volatility is equal to that of publicly traded firms in the data and Case B corresponds to the case where the firm-level volatility is equal to that of both publicly traded and privately held firms in the data.

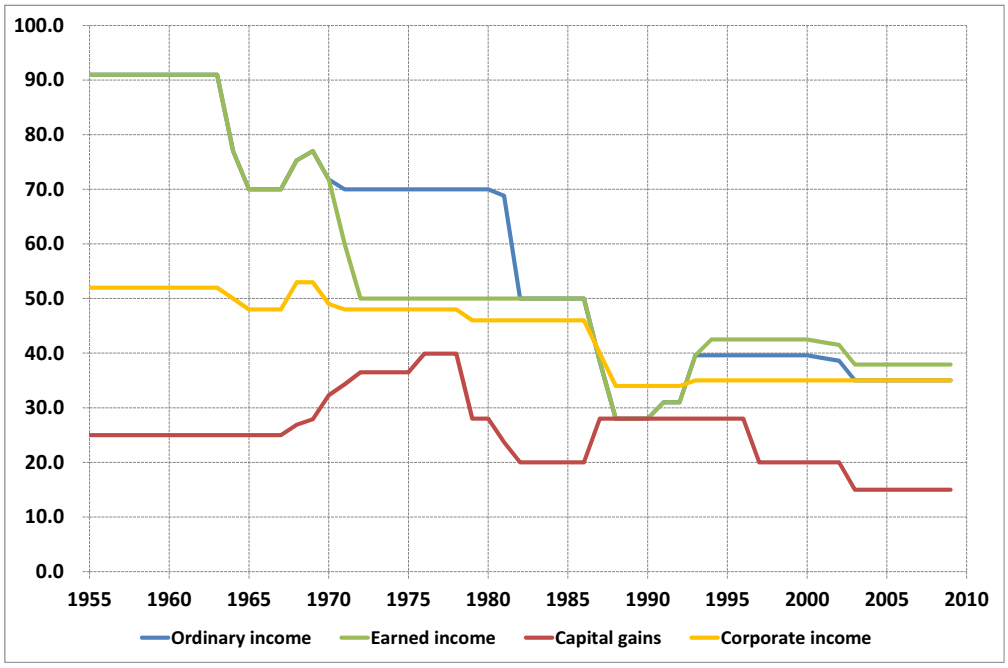

Figure 1: Federal tax rates
Note: The data are taken from Table A1 of Saez et al. (2012).

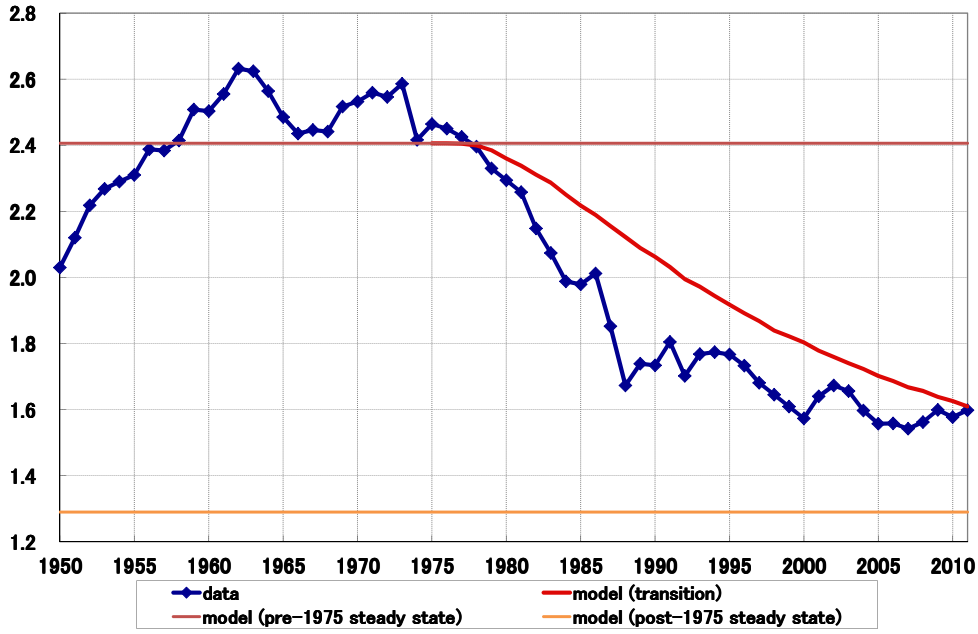

Figure 2: Pareto exponent: Case A
Note: Data are taken from Alvaredo et al. (2013).

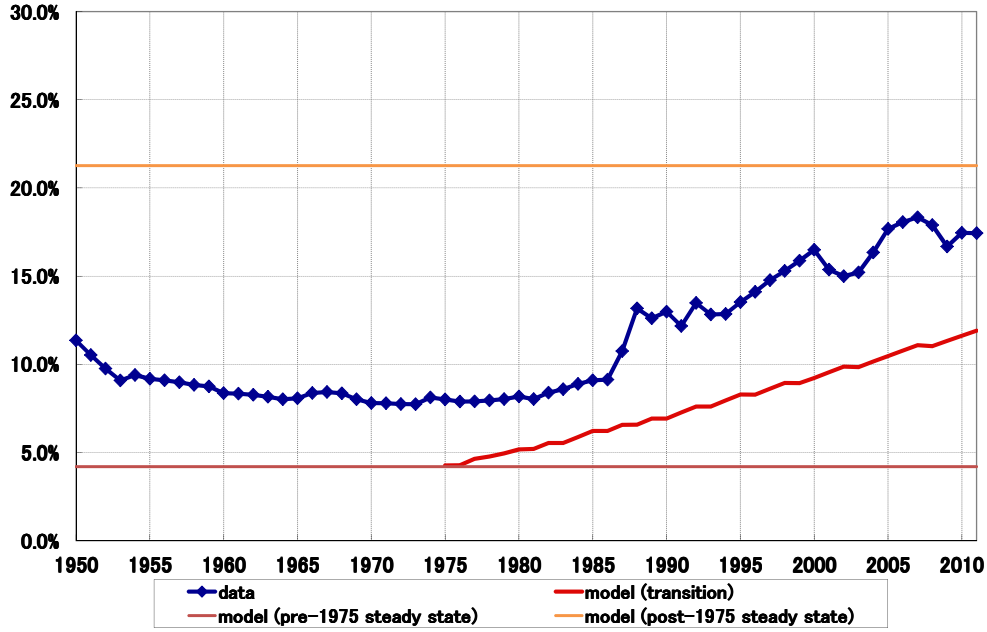

Figure 3: Top 1% share: Case A
Note: Data are taken from Alvaredo et al. (2013).

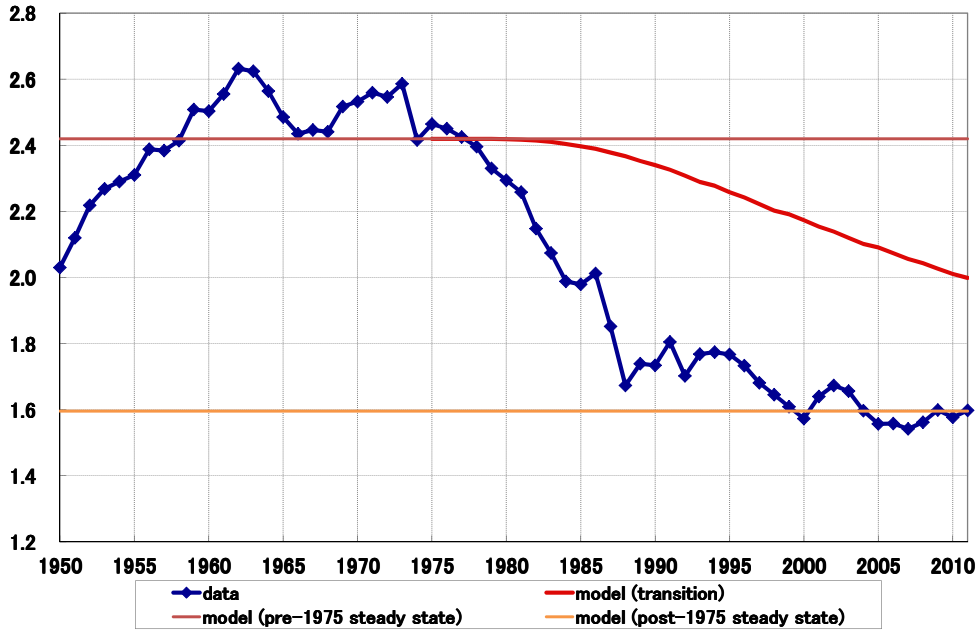

Figure 4: Pareto exponent: Case B
Note: Data are taken from Alvaredo et al. (2013).

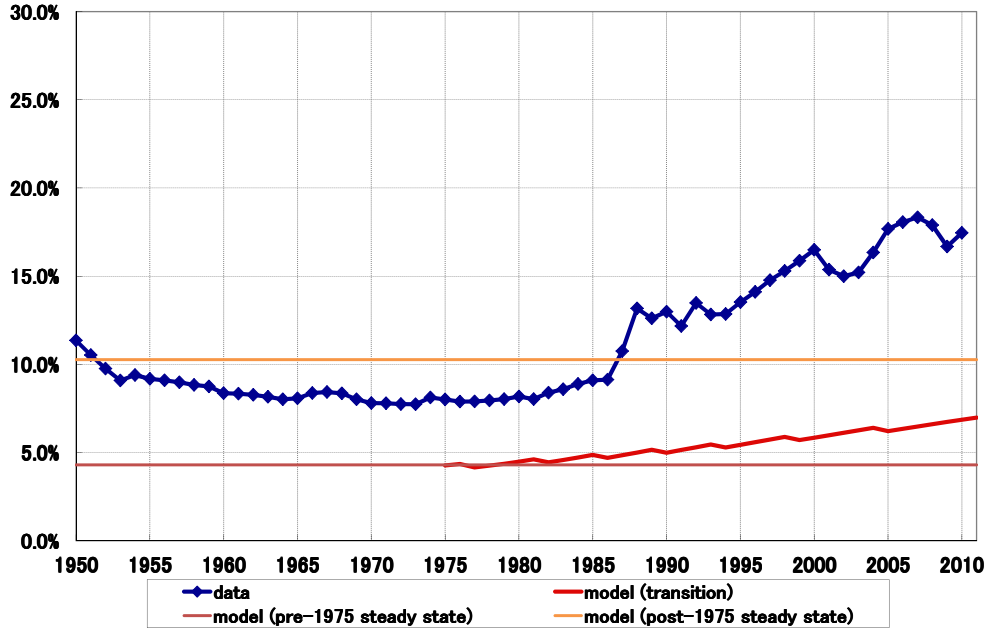

Figure 5: Top 1% share: Case B
Note: Data are taken from Alvaredo et al. (2013).

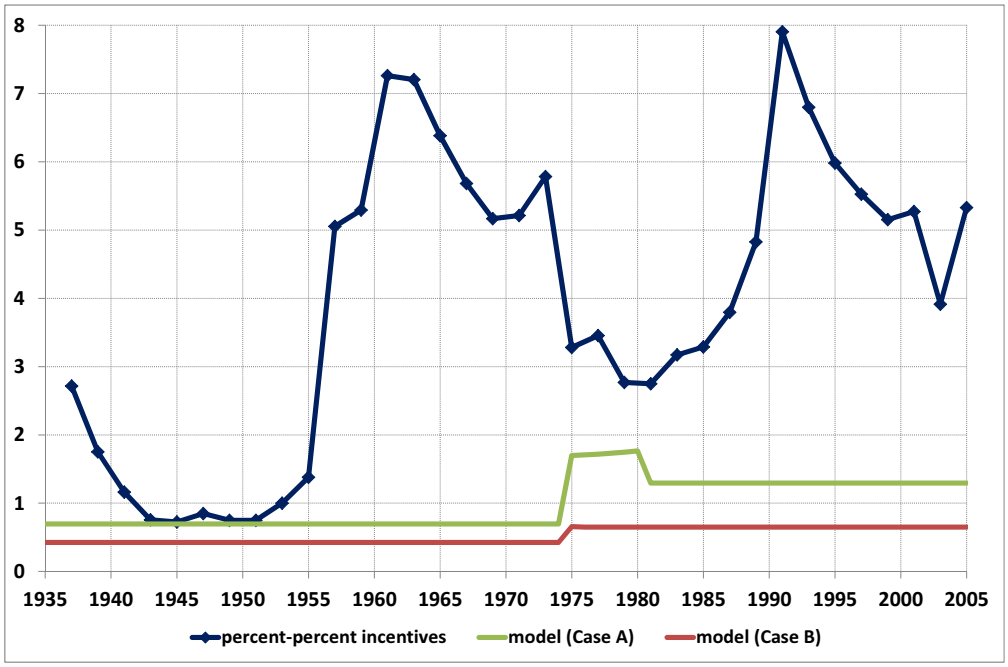

Figure 6: "Percent-percent" incentives
Notes: The "percent-percent" data are calculated by dividing "dollar change in wealth for a 1% increase in firm rate of return" by "total compensation," both of which are estimated in Frydman and Saks (2010). These data correspond to the median value of the fifty largest firms.