# 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

June 18, 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 firm 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, firm size distribution, top income share

## 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 was stable at around 8% of national income during the 1960s and 70s, but this pattern was broken in the 1980s. Since then, the top 1% share has grown to 18% by 2010. 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.

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

*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.

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 $\lambda$, 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 $\lambda$ 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 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) argues 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. While our paper shares the view with theirs that a tax cut is an important factor, there is also a difference. In our model, unlike theirs, 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.[1]

Recently, several papers have built models to understand why income distribution follows a Pareto distribution. There are two types of approaches in the literature. The one is the approach that explains Pareto's law of incomes from assuming other distributions that 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 decreasing Pareto exponent of the income distribution, 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.

The other is the approach that explains Pareto's law of incomes from 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 household problem in which households optimally make saving and bequest decisions. These models are not dynamic general equilibrium models because they only consider the household side problem and do not consider the firm side. Nirei (2009) extend the framework to a Bewley-type model and derive the Pareto's law 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.

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 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, 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.

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

[1]The reason that a cut in top marginal income tax rate itself does not matter in our model is that top marginal income tax in our model plays the same role as dividend tax in the "new view" of dividend taxation (Sinn, 1991; McGrattan and Prescott, 2005).

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 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 $\nu$. 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. $\nu$ is the pension premium rate on 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 $s_{i,t}$ and risk-free market portfolio $b_{i,t}$. The remaining $L-N$ of households are workers and can only hold risk-free assets $b_{i,t}$ (therefore, for them, $s_{i,t}=0$). Each entrepreneur leaves the firm by a Poisson hazard rate $p_f$, and becomes a worker (referred to as a former entrepreneur). 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 by holding 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 $\tau^e$ is the tax rate on risky asset and $B_{i,t}$ is the Wiener process. We interpret $\tau^e$ 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

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

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.

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

subject to (2), where $V^i(a_{i,t}, t)$ denotes value functions with household characteristics $i$: if the household is an entrepreneur $i = e$, if worker $i = w$. Note that if the household is an entrepreneur, household characteristics in the next period, denoted by $i'$, can be both entrepreneur and worker, while if the household is a worker, denoted by $i'$, is also worker.

The household problem is a variant of Merton's (1969) dynamic portfolio problem (see Appendix A for derivations), 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.

## 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 $k_j$. 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 $\iota$ 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. We make these assumptions to simplify the analysis. These assumptions prevent the entrepreneurs from choosing the dividend and investment plans in order to suit their own stochastic discount factors rather than the market's.

### 2.2.1 Financial intermediary's problem

A financial intermediary maximizes the residual profit:

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

where $s^f_{j,t}$ is the shares of firm $j$ owned by the financial intermediary and $\tau^f$ is the tax on the dividend. We interpret $\tau^f$ in the numerical analysis as the combination of capital gains and corporate income taxes. Note that $\tau^f$ is different from $\tau^e$, the tax on risky assets received by entrepreneurs. The solution of the problem leads to

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

### 2.2.2 Firm's problem

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

$$y_{j,t} = z_{j,t} k^{\alpha}_{j,t} \ell^{1-\alpha}_{j,t}.$$

The productivity of the firm evolves as

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

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 (we will discuss how the deal is conducted in the next section).

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

$$r^f_t q_{j,t} dt = \mathsf{E}_t \left[ \max_{dk_{j,t}, \ell_{j,t}} (1 - \tau^f - \iota) d_{j,t} dt + dq_{j,t} \right], \tag{8}$$

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 $\delta$ is the depreciation rate.

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

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

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

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, as mentioned in the "new view" literature of dividend taxation (Sinn, 1991; McGrattan and Prescott, 2005), because the taxes in the model are imposed on dividends, 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. Goods that a mass $N$ of firms produce are aggregated according to

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

(Throughout the paper, we denote the aggregate variables by upper case letters.) We assume that the aggregate good $Y$ is produced competitively.

The market clearing condition for final goods is

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

where $A_{e,t}$ is the total assets of entrepreneurs and $F_t$ is aggregate financial asset. (Since $\iota$ is real transaction cost, this term should also be subtracted.) 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 that owned by the entrepreneur and $s_j^f$ is that 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 whose exponent is close to unity 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 $\mu_{k,t}$ and $\mu_{d,t}$ is 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, $z_{j,t}^{\frac{\rho}{1-\rho}}$. Therefore, the firm size distribution depends only on the heterogeneous component.

## 3.2 Restructuring

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 pay $\overline{q}_{t+dt}\overline{\ell}_{t+dt}\left( z_{\min}^{\frac{\rho}{1-\rho}} - z_{j,t+dt}^{\frac{\rho}{1-\rho}} \right)$ as the restructuring cost to increase the firm size to. We denote the total payment of the restructured firms in the economy as $Q_{\text{restructuring},t+dt}$. Here, we analyze how the firm values needed for restructuring are collected from other firms.

We assume that from at each instant, a fraction $m\left(\frac{\rho}{1-\rho}\right)dt$ of each firm's value $q_{j,t}$, whose value comes from underlying productivity and capital, is sold to firms whose employment is less than $\ell_{\min}$ 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 $z_{j,t}^{\frac{\rho}{1-\rho}}$.) After the sellout, firm's capital and productivity decrease by $m\left(\frac{\rho}{1-\rho}\right)k_{j,t}dt$, and $mz_{j,t}dt$ respectively. 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 firm values needed for restructuring has to equate 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)

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

## 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 sellout detrended by $e^{g_z t}$, where $g_z$ is a constant (the value of $g_z$ 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:

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

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

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

With this $\lambda$, 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 also 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, under these assumptions, Zipf's law of firms holds. For the above equation shows that $\lambda > \rho/(1-\rho)$ and that $\lambda$ becomes close to $\rho/(1-\rho)$ if $\ell_{\min}$ is sufficiently small compared with the average employment level $L/N$. Then, Zipf's law approximately holds for the firm size distribution, e.g., the distribution of $\ell_{j,t}$ because the firm size distribution, such as the distribution of $\ell_{j,t}$, cross-sectionally obeys to $\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}$, whose Pareto exponent is $\lambda/\left(\frac{\rho}{1-\rho}\right)$.

# 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 (workers by birth), 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 $\left\{\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t\right\}$. The evolution of these variables is computed at each $t$ as follows:

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

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

2. $\tilde{C}_t = v\tilde{A}_t$ and $\tilde{F}_t = \tilde{A}_t - \tilde{H}_t$. Then,

$$\frac{d\tilde{K}_t}{dt} = \tilde{Y}_t - \delta \tilde{K}_t - \tilde{C}_t - \iota \left(1 - \frac{\tilde{A}_{e,t} x_t}{\tilde{F}_t}\right) \tilde{D}_t - g\tilde{K}_t,$$

$$\tilde{D}_t = (1-(1-\alpha)\rho)\tilde{Y}_t - (\delta + g + \mu_{\overline{\mathsf{E}}})\tilde{K}_t - \frac{d\tilde{K}_t}{dt},$$

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):

$$\mu_{q,t} = \left\{ \left( \frac{1-\tau^e}{(1-\tau^f-\iota)} - 1 \right) \frac{1}{\int_t^\infty \exp\left\{ -\int_t^u (r_s^f - \mu_{d,s}) ds \right\} du} + r_t^f \right\},$$

$$\sigma_{q,t} = \left( \frac{\rho}{1-\rho} \right) \sigma_z$$

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

where

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

3. We can compute $d\tilde{A}_t/dt$ by summing the following equations:

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

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

$$\frac{d\tilde{A}_{f,t}}{dt} = (\mu_{aw,t} - g)\, \tilde{A}_{f,t} + p_f \tilde{A}_{e,t} - \nu \tilde{A}_{f,t}.$$

$d\tilde{H}_t/dt$ can be computed by

$$\frac{d\tilde{H}_t}{dt} = -(\tilde{w}_t + \tilde{tr}_t)L + (\nu + r_t^f - g)\tilde{H}_t, \tag{25}$$

where

$$\tilde{w}_t = (1-\alpha)\rho \tilde{Y}_t / L,$$

$$\tilde{tr}_t = \left\{ \frac{\tilde{A}_{e,t} x_{e,t}}{\tilde{F}_t} \tau^e + \left(1 - \frac{\tilde{A}_{e,t} x_{e,t}}{\tilde{F}_t}\right) \tau^f \right\} \tilde{D}_t / L.$$

# 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. The asset distribution of households is

## 5.1 Asset distribution of entrepreneurs

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

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

where $\mu_{ae,t}$ and $\sigma_{ae,t}$ 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 entrepreneurs who are alive at $t$, relative to their initial asset, $\ln(a_{e,t}/h_{t-t'}) = \ln\tilde{a}_{e,t} - (\ln\tilde{h}_{t-t'} - gt')$, follows a normal distribution with mean $(\mu_{ae} - \sigma^2_{ae}/2)t'$ and variance $\sigma^2_{ae}t'$.

By combining the above property and the constant probability of death assumption, the asset distribution of entrepreneurs is obtained. As Benhabib et al. (2012) and Toda (2012) show, the probability density function becomes a double-Pareto distribution (see Appendix C for the derivations)[2]

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

where

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

The result shows that $\psi_1$ is the Pareto exponent for entrepreneurs at the upper-tail.

[2]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 $f_l(\ln\tilde{a}_i)$ such that

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

## 5.2 Asset distribution of innate workers

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

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

where $\mu_{aw}$ is the drift part of the worker's asset process.

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

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

Under the parameter values in numerical analysis, the trend growth of worker's asset is lower than the trend growth of the economy, i.e., $\mu_{aw} \leq g$. Then, the detrended asset level of the innate workers becomes less than $\tilde{h}$.

## 5.3 Asset distribution of former entrepreneurs

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

Under the settings, we can analytically derive the steady state asset distribution of the former entrepreneurs. Because under the parameter values in numerical analysis $\mu_{aw} \leq g$, here, we report the distribution for that case (for the $\mu_{aw} > g$ case, see Appendix C):

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

This shows that the Pareto exponent for former entrepreneurs at the upper-tail is the same as that for entrepreneurs.

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

There are two remarks on the household asset and income distribution. First, the Pareto exponent at the upper-tail for all of the households is that of entrepreneurs, whose Pareto exponent is $\psi_1$. This is because the sum of variables each of which follows Pareto distribution also follows a Pareto distribution at the upper-tail and the distribution of smallest Pareto exponent dominates (see e.g., Gabaix, 2009).

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, $\psi_1$. This is because the household's consumption and income is proportional to the household's asset level.

## 6 Numerical Analysis

In this section, we suppose that at 1970 a cut in top marginal tax rates suddenly and permanently in an unexpected way, and numerically analyze how the tax cut affects top incomes.

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 1970, the tax rate on risky asset $\tau^e$ becomes relatively lower than the tax rate on risk-free asset $\tau^f$, 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 $\tau^e$ is equal to the ordinary income tax that is imposed on the CEO pay in the real world. We assume that the tax on risk-free assets $\tau^f$ is the sum of taxes that are imposed on dividends when investors buy the equity of the firm. We calculate the tax rate of risk-free assets, $\tau^f$ by $1-(1-\tau^{\text{cap}})(1-\tau^{\text{corp}})$, where $\tau^{\text{cap}}$ and $\tau^{\text{corp}}$ are the marginal tax rates for capital gains and corporate income.

The tax rates are calibrated using top statutory marginal federal tax rates reported in Saez et al. (2012) (see Figure 1 and Tabletab:taxes). We use the top marginal tax rates because we focus on inequality at the upper-tail.

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 $\nu$ that the average length of life after a household begins to work is 50 years.

$\rho$ is set to 0.7, which implies that of the firm's sales 30% is rent. The value of $\rho$ is lower than the standard one. There are two reasons for the value. 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. To take it into account, we set a lower $\rho$. Second, if $\rho$ is high, the total value of entrepreneur's risky assets exceeds the total value of financial assets in the economy. To avoid this, a low $\rho$ should be chosen. We assume for $p_f$ 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 avoid this, 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 moves in the same direction.[3]

For the calibration of the firm-level volatility, we consider the 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, $\iota$, is calibrated to match the Pareto exponent in the pre-1970 steady state with the data that is around 2.4.

Insert Table 2 here.

## 6.3 Computation of the transition dynamics

We compute the Pareto exponent of household's asset and income distribution and the top 1% income share before and after 1970. We assume that before 1970 the economy is in the pre-1970 steady state. In our experiment, taxes change suddenly and permanently and unexpectedly at 1970, and the economy moves toward the post-1970 steady state.

We model the transition dynamics after 1970 in the following way. First, the dynamics of aggregate variables are computed separately. To compute the dynamics of aggregate variables $\left\{\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t\right\}$ explained in Section 4.2, we need to pin down their initial values. We suppose that at 1970 when the tax change occurs, the aggregate capital stock is the same as that in the pre-1970 steady state. We also suppose that asset shares of entrepreneurs, innate workers, and former entrepreneurs, $A_{e,1970}/A_{1970}$, $A_{w,1970}/A_{1970}$, $A_{f,1970}/A_{1970}$, are the same as those in the pre-1970 steady state. The remaining initial variables, $\tilde{A}_{1970}$ and $\tilde{H}_{1970}$ are determined by the shooting algorithm by the following steps:

1. Set $\tilde{A}_{1970}$. Set also the upper and lower bound of $\tilde{A}_t$, $\tilde{A}_H$ and $\tilde{A}_L$.
    (a) Set $\tilde{H}_{1970}$ 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 $\tilde{H}_{1970}$ by backwardly solving (25) with the terminal condition

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

    where the variables with asterisks are those in the post-1970 steady state and $T = \arg\min_T \sqrt{(K_T - K^*)^2 + (C_T - C^*)^2}$.
    (c) Repeat until $|\tilde{H}^{\text{new}}_{1970} - \tilde{H}^{\text{old}}_{1970}| < \varepsilon$.

[3]A possible story behind the assumption is that these productivity shocks are caused by managerial decisions.

2. If $\tilde{A}_t$ upwardly diverges, $\tilde{A}_{1970}^{\text{new}} = (\tilde{A}_{1970}^{\text{old}} + \tilde{A}_L)/2$ and redo the procedure. Otherwise, set $\tilde{A}_{1970}^{\text{new}} = (\tilde{A}_H + \tilde{A}_{1970}^{\text{old}})/2$.

3. Repeat the procedure until $|\tilde{A}_{1970}^{\text{new}} - \tilde{A}_{1970}^{\text{old}}| < \varepsilon$.

Note that since $\tilde{C}_t = v\tilde{A}_t$, the above procedure is similar to the standard shooting algorithm used in 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-1970 steady state, the economy switches to the post-1970 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 and the Fokker-Planck equations, we compute the probability density functions of asset distribution for entrepreneurs and former entrepreneurs, $f_e(\ln \tilde{a}_{i,t}, t)$ and $f_f(\ln \tilde{a}_{i,t}, t)$:

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

Here, we compute the asset distribution at the upper-tail, using a numerical method of partial differential equations.[4] 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 $\tilde{a}_{i,t}$, $\tilde{a}_{\text{LB}}$, which is set to be higher than $\tilde{h}$ at the pre- and post-1970 steady state, $f_i(\ln \tilde{a}_{\text{LB}}, t)$ moves linearly during the 50 years from that of the pre-1970 steady state to that of the post-1970 steady state.

## 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-1970 and post-1970 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 1970, although the model's prediction is somewhat lower in the level than data. Perhaps, other factors like the differences in talents also account for the level of the top income share.

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 volatility at the top firms where the richest CEOs work is an important factor to understand the evolution of top incomes,

[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 1970 to 2030.

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 growing inequality in the model is that by the tax change it is 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-1970 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).[5]

We plot the "percent-percent" incentives constructed from Frydman and Saks (2010) and $x_{e,t}$ in the model in Figure 6.[6] 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.

## 7 Conclusion

We have proposed a model of asset and income inequalities, consistent with (1) firm-size distribution (2) and household asset and income inequalities at the upper-tail. Our model matches with the decline in the Pareto exponent of income distribution and the trend in top 1% share. On the other hand, there are also 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.

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

[6] 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).

# 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.

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 (2009) "Power Laws in Economics and Finance," *Annual Review of Economics*, Vol. 1, No. 1, pp. 255–294.

Gabaix, Xavier and Augustin Landier (2008) "Why Has CEO Pay Increased So Much?" *The 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.

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

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

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, 07.

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

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," *The 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*: Oxford: 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. By Ito's formula, $V^i(a_{i,t},t)$ is rewritten as follows:

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

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,

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

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

$$
\begin{aligned}
0 =& \max_{c_{i,t}, x_{i,t}} \ln c_{i,t} - (\beta+\nu)V_t^i + \frac{\partial V_t^i}{\partial t} + \mu_{a,t} a_{i,t} \frac{\partial V_t^i}{\partial a_{i,t}} + \frac{(\sigma_{a,t} a_{i,t})^2}{2} \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\
&+ p_f \left( V_t^{i'} - V_t^i \right) \\
=& \max_{c_{i,t}, x_{i,t}} \ln c_{i,t} - (\beta+\nu)V_t^i + \frac{\partial V_t^i}{\partial t} + \frac{\sigma_{q,t}^2}{2} x_{i,t}^2 a_{i,t}^2 \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\
&+ ((\nu+\mu_{q,t}) x_{i,t} a_{i,t} + (\nu + r_t^f)(1-x_{i,t}) a_{i,t} - c_{i,t}) \frac{\partial V_t^i}{\partial a_{i,t}}, \\
&+ p_f \left\{ V^{i'}(a_{i,t}, t) - V^i(a_{i,t}, t) \right\}. \qquad (26)
\end{aligned}
$$

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

$$
c_{i,t}^{-1} = \frac{\partial V_t^i}{\partial a_{i,t}}, \qquad (27)
$$

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

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

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

We obtain this solution by guess-and-verify. The first-order condition (27) becomes:

$$
(v_{i,t})^{-1} = B_t^i. \qquad (29)
$$

Condition (28) is rewritten as:

$$
x_{i,t} = \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}. \qquad (30)
$$

HJB (26) is expressed as:

$$
\begin{aligned}
0 =& \ln(v_{i,t} a_{i,t}) - (\beta+\nu) B_t^i \ln a_{i,t} + \frac{\partial B_t^i}{\partial t} \ln a_{i,t} - B_t^i a_{i,t}^{-2} (x_{i,t})^2 a_{i,t}^2 \frac{(\sigma_{q,t})^2}{2} \\
&+ B_t^i a_{i,t}^{-1} ((\nu+\mu_{q,t}) x_{i,t} a_{i,t} + (\nu + r_t^f)(1-x_{i,t}) a_{i,t} - v_{i,t} a_{i,t}) \\
&+ p_f \left\{ B_t^{i'} \ln a_{i,t} - B_t^i \ln a_{i,t} \right\}.
\end{aligned}
$$

By solving this equation, we obtain

$$
v_{i,t} = \nu + \beta. \qquad (31)
$$

# 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 $k_{j,t}$ and $z_{j,t}$. By applying Ito's formula to $q_{j,t}$, we obtain

$$
\begin{aligned}
dq(k_{j,t}, z_{j,t}, t) =& \left(\frac{\partial q_{j,t}}{\partial t}dt + \frac{\partial q_{j,t}}{\partial z_{j,t}}dz_{j,t} + \frac{\partial q_{j,t}}{\partial k_{j,t}}dk_{j,t}\right) + \frac{1}{2}\frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2}(dz_{j,t})^2 \\
=& \left(\frac{\partial q_{j,t}}{\partial t} + \mu_z\frac{\partial q_{j,t}}{\partial z_{j,t}} + \frac{1}{2}\sigma_z^2\frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2}\right)dt + \frac{\partial q_{j,t}}{\partial k_{j,t}}dk_{j,t} \\
&+ \sigma_z\frac{\partial q_{j,t}}{\partial z_{j,t}}dB_{j,t}.
\end{aligned}
$$

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

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

By the envelope theorem,

$$
\begin{aligned}
r_t^f\frac{\partial q_{j,t}}{\partial k_{j,t}}dt =&(1-\tau^f-\iota)\left(\frac{\partial p_{j,t}y_{j,t}}{\partial k_{j,t}}dt - \delta dt\right) \\
&+ \left((1-\delta_b)\frac{\partial V((1-\delta_b)k_{j,t},0,t)}{\partial k_{j,t}} - \frac{\partial V(k_{j,t},z_{j,t},t)}{\partial k_{j,t}}\right) \\
\Longrightarrow r_t^f =& \frac{\partial p_{j,t}y_{j,t}}{\partial k_{j,t}} - \delta.
\end{aligned}
$$

Therefore, we obtain the conditions

$$
\begin{aligned}
r_t^f &= \frac{\partial p_{j,t}y_{j,t}}{\partial k_{j,t}} - \delta, \\
w_t &= \frac{\partial p_{j,t}y_{j,t}}{\partial \ell_{j,t}}.
\end{aligned}
$$

## 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,

$$
\ell_{j,t} = \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho} z_{j,t}^{\rho}k_{j,t}^{\alpha\rho}\right)^{\frac{1}{1-(1-\alpha)\rho}}. \tag{32}
$$

On the other hand, from (9),

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

By substituting (32) into (33) and rearranging,

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

where $\phi \equiv \frac{\rho}{1-(1-\alpha)\rho}$. Substituting (34) into (32),

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

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

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

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

where $\overline{\mathsf{E}}$ is the operator of the cross-sectional average of the whole firms. Then, substituting (36) into (35),

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

Rewriting (34),

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

Substituting (37) into (38),

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

Next, we derive $Y$. Substituting (15) and (39) 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 $Y = \left( \int_0^N \left( \frac{1}{N} \right)^{1-\rho} y_{j,t}^{\rho} dj \right)^{\frac{1}{\rho}}$,

$$\left( \frac{Y_t}{N} \right)^{1-\rho} = \left( \frac{\alpha\rho}{\mathsf{MPK}_t} \right)^{\frac{\alpha(1-\rho)}{1-\alpha}} \left( \frac{L}{N} \right)^{1-\rho} \overline{\mathsf{E}} \left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}^{(1-\rho)\left[ \frac{1-\alpha\rho}{(1-\alpha)\rho} - 1 \right]}. \tag{40}$$

Substituting (40) into (39),

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

Substituting (15) and (41) into (40)

$$p_{j,t} y_{j,t} = Y_t^{1-\rho} y_{j,t}^{\rho} \tag{42}$$

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

Rewriting (15),

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

Rewriting (43),

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

Rewriting (41),

$$k_{j,t} = \overline{k}_t \overline{\ell}_t z_{j,t}^{\frac{\rho}{1-\rho}}, \text{ where } \overline{k}_t \equiv \left( \frac{\alpha\rho}{\mathsf{MPK}_t} \overline{\mathsf{E}} \left\{ z_{j,t}^{\frac{\rho}{1-\rho}} \right\}^{\frac{1-\rho}{\rho}} \right)^{\frac{1}{1-\alpha}}. \tag{44}$$

From (44),

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

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,

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

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 $z_{\min}$ and $g$ is the trend growth rate of the aggregate economy.

$d_{j,t}dt$ is computed using these results and 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:

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

## 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[7], we obtain

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

[7]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,

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

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

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

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,

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

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

$$\begin{aligned}
\frac{(1-\tau^e)d_{j,t}+dq_{j,t}}{q_{j,t}} =& \left\{\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}-1\right)\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du} + r_t^f\right\}dt \\
&+ \left(\frac{\rho}{1-\rho}\right)\sigma_z\left\{1-\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}\right)\frac{\overline{k}_t}{\overline{d}_t}\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}\right\}dB_{j,t}.
\end{aligned} \tag{45}$$

Note that if $(r_t^f - \mu_{d,t})$ is constant, $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du = 1/(r^f-\mu_d)$ 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 transition dynamics. We calculated the value by the following steps. If we know the value of the aggregate asset $A_t$ and aggregate human asset $H_t$, we can compute aggregate financial asset $F_t$ by $A_t - H_t$. On the other hand, integrating (19), we obtain

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

Substituting this result into (45), we can compute the return of risky asset in transition.

## B.4 Derivations on the restructuring

This appendix shows the derivations of the restructuring at Section 3.2. Let $\tilde{z}_{j,t} \equiv z_{j,t}/e^{g_z t}$ . Then, $Q_{\text{restructuring},t+dt}$ is written as follows:

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

where $\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\}$ 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),

$$\begin{aligned}
\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\} = & \int_{\ln \tilde{z}_{\min}}^{\infty} d(\ln \tilde{z}_{j,t}) \int_{-\infty}^{\ln \tilde{z}_{\min}} d(\ln \tilde{z}_{j,t+dt}) \\
& \left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\right) f_z(\ln \tilde{z}_{j,t}) f_z(\ln \tilde{z}_{j,t+dt}|\ln \tilde{z}_{j,t}) \\
= & \int_{\ln \tilde{z}_{\min}}^{\infty} d(\ln \tilde{z}_{j,t}) \int_{-\infty}^{\ln \tilde{z}_{\min}} d(\ln \tilde{z}_{j,t+dt}) \\
& \left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\right) C_0 e^{-\lambda \ln \tilde{z}_{j,t}} \frac{1}{\sqrt{2\pi\sigma_z^2 dt}} e^{-\frac{\left(\ln \tilde{z}_{j,t+dt} - (\ln \tilde{z}_{j,t} + \tilde{\mu}_z dt)\right)^2}{2\sigma_z^2 dt}},
\end{aligned}$$

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

$$\begin{aligned}
\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\frac{\rho}{1-\rho}}\right\} m \left(\frac{\rho}{1-\rho}\right) = & \lim_{dt \to 0+} \frac{\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+dt}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\}}{dt} \\
= & \lim_{t' \to 0+} \frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\middle|\tilde{z}_{j,t+t'} \leq \tilde{z}_{\min}\right\}}{dt'}.
\end{aligned} \tag{46}$$

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

$$
\begin{aligned}
&\frac{d\mathsf{E}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}} \middle| \tilde{z}_{j,t+t'} \leq \tilde{z}_{\min}\right\}}{dt'} \\
&= \int_{\ln \tilde{z}_{\min}}^{\infty} d(\ln \tilde{z}_{j,t}) \int_{-\infty}^{\ln \tilde{z}_{\min}} d(\ln \tilde{z}_{j,t+t'}) \\
&\quad \frac{d}{dt'}\left(\left(\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}}\right) C_0 e^{-\lambda \ln \tilde{z}_{j,t}} \frac{1}{\sqrt{2\pi\sigma_z^2 t'}} e^{-\frac{\left(\ln \tilde{z}_{j,t+t'} - (\ln \tilde{z}_{j,t} + \tilde{\mu}_z t')\right)^2}{2\sigma_z^2 t'}}\right) \\
&= \frac{C_0 e^{-\left(\lambda - \left(\frac{\rho}{1-\rho}\right)\right) \ln \tilde{z}_{\min}} \left(\frac{\rho}{1-\rho}\right)}{4\left(\lambda - \left(\frac{\rho}{1-\rho}\right)\right)} \\
&\quad \times \left(e^{\frac{1}{2}\lambda(2\tilde{\mu}_z + \lambda\sigma_z^2)t'}(2\tilde{\mu}_z + \lambda\sigma_z^2)\,\mathrm{Erfc}\left[\frac{(\tilde{\mu}_z + \lambda\sigma_z^2)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right. \\
&\quad \left. - e^{\frac{1}{2}\left(\frac{\rho}{1-\rho}\right)\left(2\tilde{\mu}_z + \left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right)t'} \left(2\tilde{\mu}_z + \left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right) \mathrm{Erfc}\left[\frac{\left(\tilde{\mu}_z + \left(\frac{\rho}{1-\rho}\right)\sigma_z^2\right)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right).
\end{aligned}
$$

By combining these results and taking the limit, we obtain

$$
\lim_{t' \to 0+} \frac{d\mathsf{E}\left\{\tilde{z}_{\min}^{\frac{\rho}{1-\rho}} - \tilde{z}_{j,t+t'}^{\frac{\rho}{1-\rho}} \middle| \tilde{z}_{j,t+t'} \leq \tilde{z}_{\min}\right\}}{dt'} = \frac{1}{4} C_0 e^{-\left(\lambda - \left(\frac{\rho}{1-\rho}\right)\right) \ln \tilde{z}_{\min}} \left(\frac{\rho}{1-\rho}\right) \sigma_z^2
$$

Substituting this result into (46), 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 5.

## C.1 Derivations on the asset distribution of entrepreneurs

$f_e(\ln \tilde{a}_i | t')$, the probability density function of entrepreneurs at age $t'$ whose detrended log wealth level is $\ln \tilde{a}_i$, is

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

Then, the probability density function of entrepreneur's asset distribution, $f_e(\ln \tilde{a}_i)$, is

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

where

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

is the probability density of entrepreneurs whose age is $t'$. By applying the formula

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

to the above equation, we obtain $f_e(\ln \tilde{a}_i)$ in Section 5.1.

## C.2 Derivations on the asset distribution of innate workers

The asset distribution of innate workers is calculated as follows:

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

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

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

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

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

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

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

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

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

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

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

| | pre-1970 | post-1970 |
|---|---|---|
| ordinary income tax, $\tau^{\text{ord}}$ | 0.75 | 0.40 |
| corporate income tax, $\tau^{\text{corp}}$ | 0.50 | 0.35 |
| capital gain tax, $\tau^{\text{cap}}$ | 0.25 | 0.25 |
| $\tau^e$ | 0.75 | 0.40 |
| $\tau^f$ | 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, $\tau^e$, is set to be equal to $\tau^{ord}$. The tax rate on risk-free assets, $\tau^f$, is calculated by $1-(1-\tau^{\text{cap}})(1-\tau^{\text{corp}})$,

| | | |
|---|---|---|
| $\beta$ | discount rate | 0.04 |
| $\nu$ | prob. of death | 1/50 |
| $\alpha$ | capital share | 1/3 |
| $\delta$ | depreciation rate | 0.1 |
| $g$ | steady state growth rate | 0.02 |
| $\rho$ | 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 |
| $\iota$ | 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.

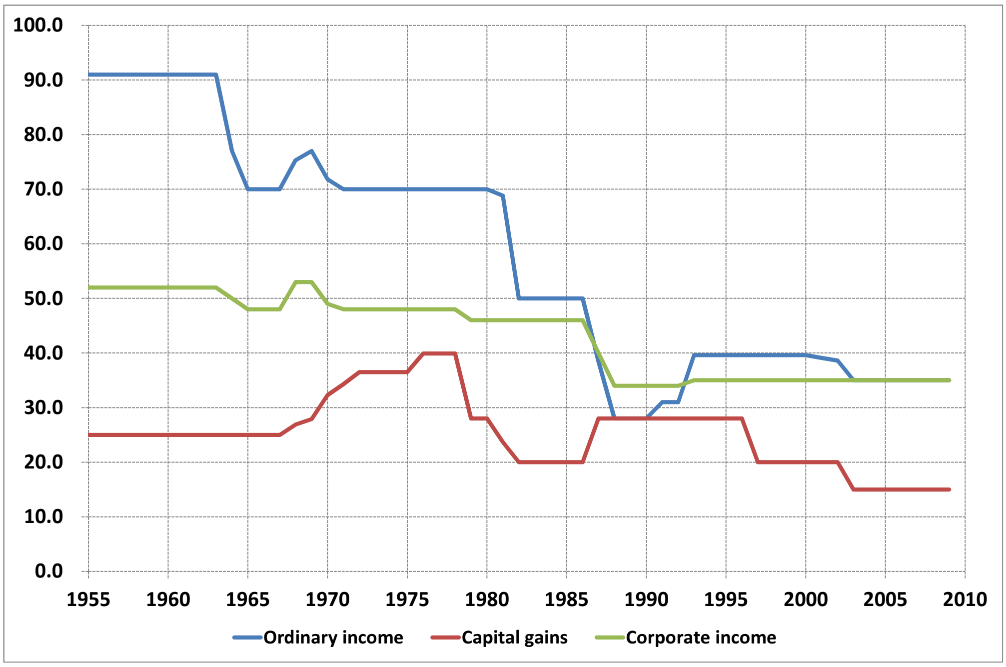


Figure 1: Federal tax rates
Note: The data are taken from Table A1 of Saez et al. (2012).

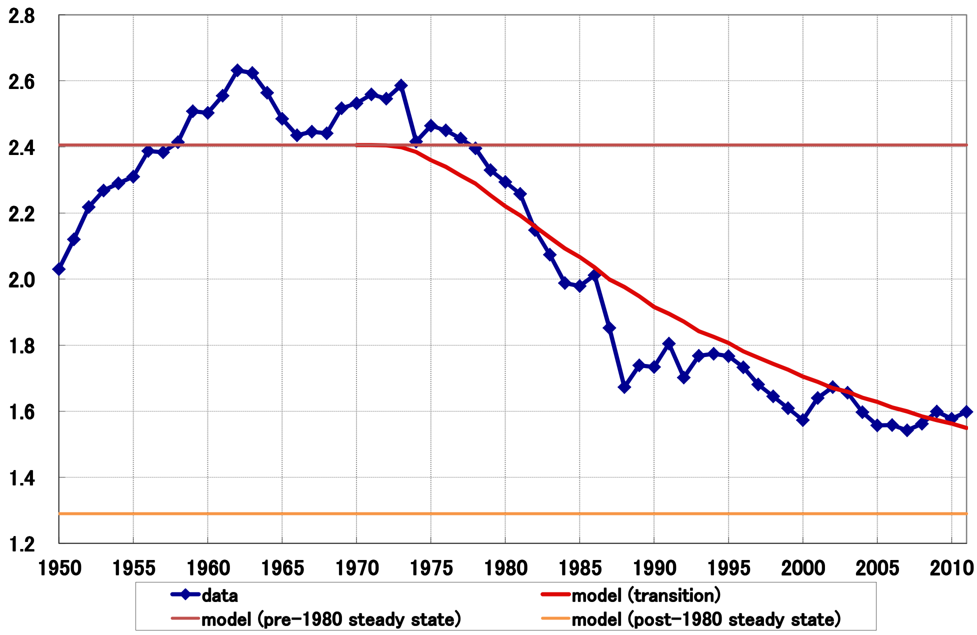


Figure 2: Pareto exponent: Case A
Note: Data are taken from Alvaredo et al. (2013).

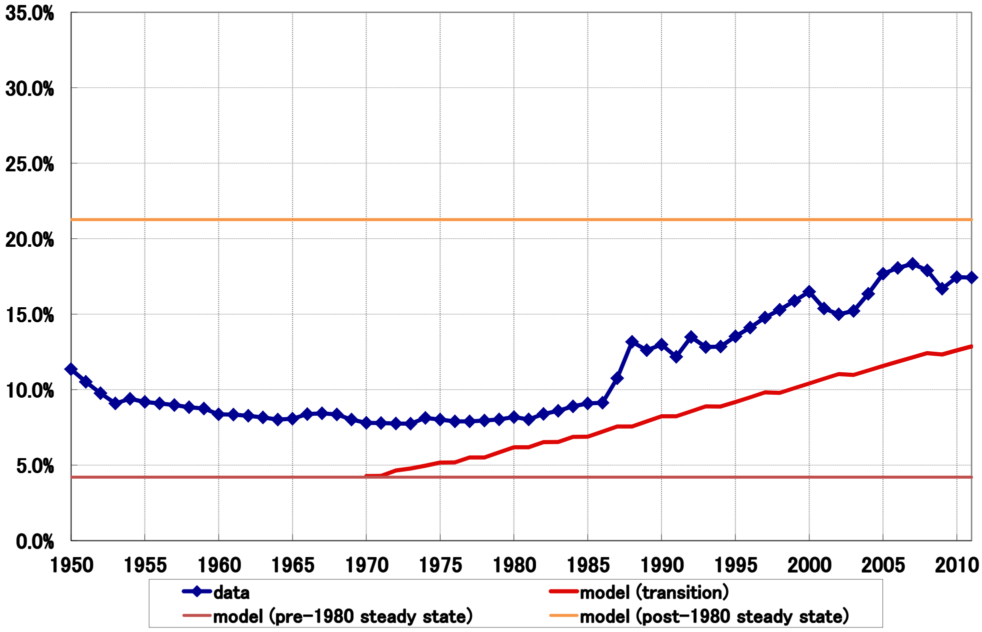


Figure 3: Top 1% share: Case A
Note: Data are taken from Alvaredo et al. (2013).

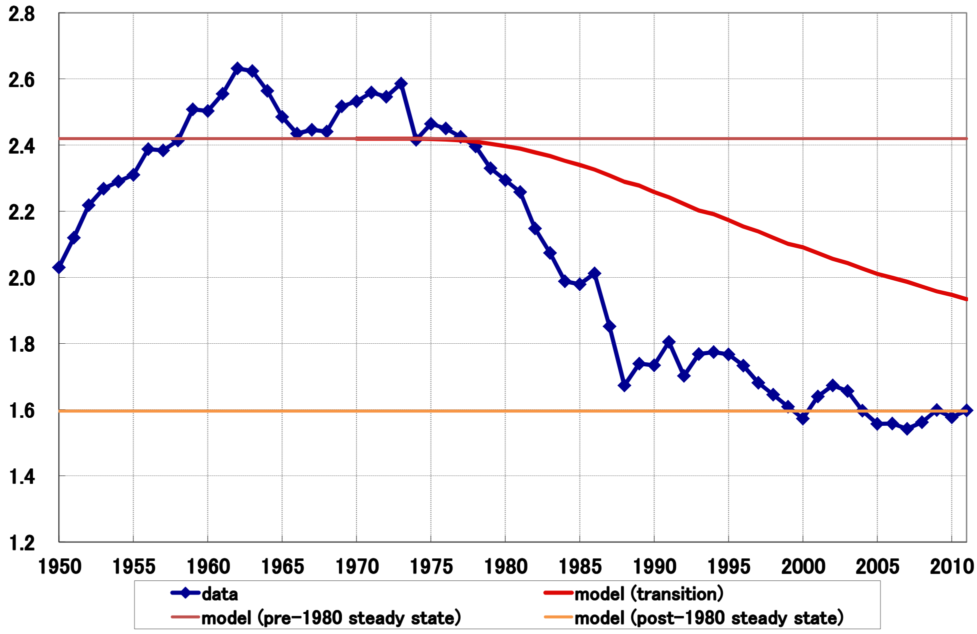


Figure 4: Pareto exponent: Case B
Note: Data are taken from Alvaredo et al. (2013).

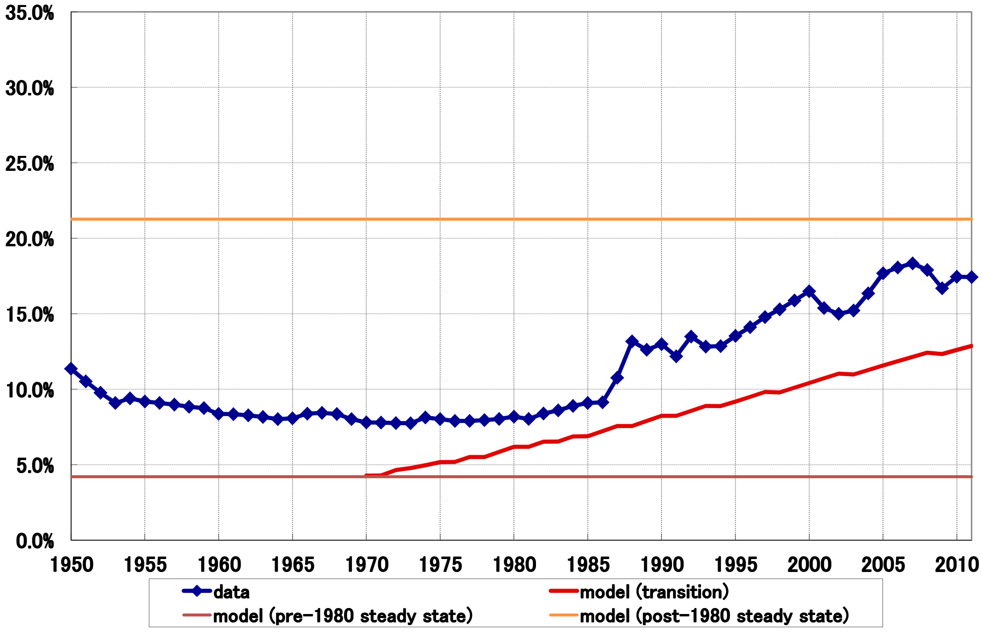


Figure 5: Top 1% share: Case B
Note: Data are taken from Alvaredo et al. (2013).

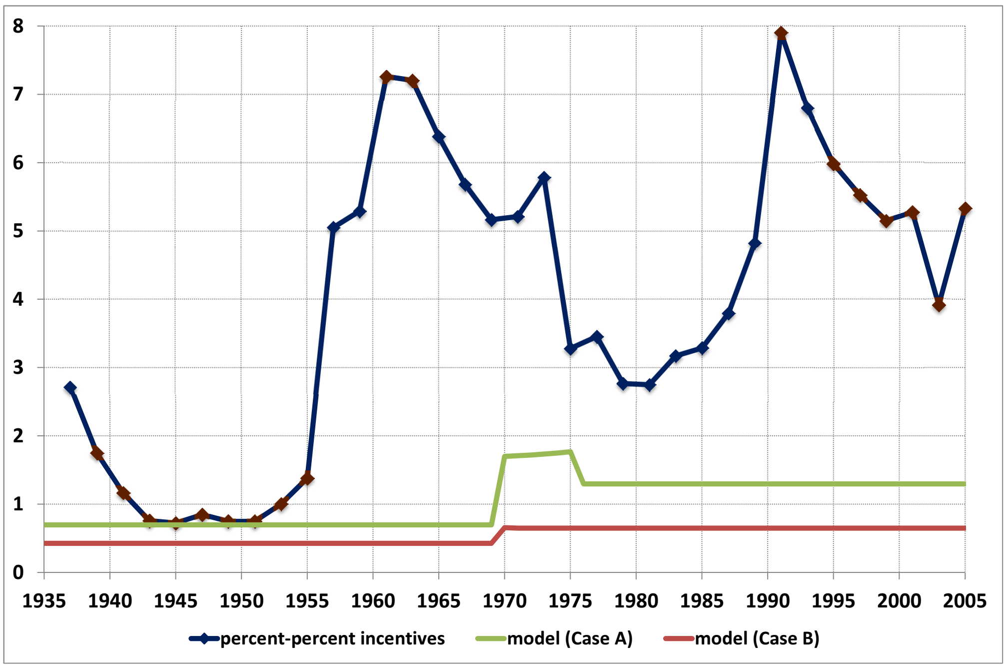


Figure 6: "Percent-percent" incentives
Notes: The "percent-percent" data are calculated by dividing "dollar change in wealth for a 1% increase in firm rate of return" by "total compensation," both of which are estimated in Frydman and Saks (2010). These data correspond to the median value of the fifty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