Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S.

Aoki, Shuhei and Nirei, Makoto

Department of Economics, Shinshu University, Policy Research Institute, Ministry of Finance, Japan

15 September 2016

Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S.*

Shuhei Aoki†
Department of Economics, Shinshu University

Makoto Nirei‡
Policy Research Institute, Ministry of Finance, Japan

September 15, 2016

Abstract

We construct a tractable neoclassical growth model that generates Pareto's law of income distribution and Zipf's law of the firm size distribution from idiosyncratic, firm-level productivity shocks. Executives and entrepreneurs invest in risk-free assets as well as their own firms' risky stocks, through which their wealth and income depend on firm-level shocks. By using the model, we evaluate how changes in tax rates can account for the evolution of top incomes in the U.S. The model matches the decline in the Pareto exponent of the income distribution and the trend of the top 1% income share in recent decades.

JEL Codes: D31, L11, O40

*First draft: June 18, 2013. A previous version of this paper was entitled "Pareto Distributions and the Evolution of Top Incomes in the U.S." This research began when the authors were affiliated at the Institute of Innovation Research, Hitotsubashi University. We thank Michio Suzuki, Tomoaki Yamada, Kozo Ueda, and Shin-ichi Fukuda for their helpful comments and discussions. This work was supported by JSPS KAKENHI Grant Numbers JP23243050, JP24730238, JP16K17114.

†Address: 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan. Phone: +81 (263) 37-2330. Fax: +81 (263) 37-2344. Email: shuhei.aoki@gmail.com.

‡3-1-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8940. Phone: +81 (3) 3581-4111. Fax: +81 (3) 5251-2113. Email: mnirei@gmail.com.

Keywords: income distribution; wealth distribution; Pareto exponent; top income share; firm size distribution; Zipf's law

1 Introduction

For the past three decades, there has been a secular trend of the concentration of income among the top earners in the U.S. economy. According to Alvaredo et al. (2013), the top 1% income share, the share of the total income accruing to the richest 1% of the population, grew from around 8% in the 1970s to 18% by 2010, on par with the high level of concentration in the 1930s.

Along with the increasing trend in the top income share, a widening dispersion of income within the top income group has also been observed over the same periods. The right tail of the income distribution is well fitted by a Pareto distribution, as known as Pareto's law of incomes. When income follows a Pareto distribution with a slope parameter λ, the ratio of the number of people who earn more than x_1 to those who earn more than x_2 is equal to $(x_1/x_2)^{-\lambda}$ for any income levels x_1 and x_2. Thus, the parameter λ, which is called the Pareto exponent, measures the degree of equality among the rich. Notably, the estimated Pareto exponent historically shows a close connection with the top income share (see, e.g., Atkinson et al., 2011). The Pareto exponent declined from 2.5 in 1970 to 1.6 in 2010, implying that a widening dispersion of income within the top income group occurred along with a secular increase in the top 1% income share.

The purpose of this study is to develop a tractable dynamic general equilibrium model that explains Pareto's law and to analyze the causes of income concentration and dispersion. We pay special attention to the top marginal income tax rate as a driving force of income dispersion among the rich, in line with Piketty and Saez (2003). Piketty et al. (2011) report that among OECD countries, the countries that have experienced a sharp rise in their top 1% income share are also the ones that have reduced the top marginal income tax rate drastically. This study examines how a decrease in the top marginal income tax rate contributes to income concentration and dispersion in a heterogeneous-agent dynamic general equilibrium model.

While our main focus is on the income distribution, we require the model to be consistent with firm-side stylized facts because a substantial part of top income in recent decades has been derived from business income such as corporate executive compensation and entrepreneurial income (Piketty and Saez, 2003; Atkinson et al., 2011; Bakija et al., 2012). Although executives and entrepreneurs are different in many respects, they are similar in that their earnings strongly depend on firms' performance (Bitler et al., 2005; Moskowitz and Vissing-Jorgensen, 2002; Edmans et al., 2009; Clementi and Cooley, 2009; Frydman and Saks, 2010). This is clear for an entrepreneur, and it increasingly holds true for an executive because of the widespread use of stock options as a form of executive compensation (see Frydman and Jenter, 2010 for a survey). Since a firm's performance is determined by its productivity in standard neoclassical models, a model of the evolution of top income in this framework should be consistent with the stylized facts of firm productivity. One of these facts is Zipf's law of firms, which states that the firm size distribution follows a special case of a Pareto distribution with exponent $\lambda = 1$. Zipf's law is closely related to Gibrat's law, which observes that a firm's growth rate is independent of its size (Gabaix, 2009; Luttmer, 2010).[1] For example, Luttmer (2007) generates Zipf's law from firms' idiosyncratic productivity shocks in standard models. We construct our model in line with this literature.[2]

The contribution of this study is summarized as follows. First, we present a parsimonious neoclassical growth model that generates Zipf's and Gibrat's laws of firms and Pareto's law of incomes from idiosyncratic, firm-level productivity shocks. In the model, the dispersion of firm size and value solely results from the firm-level productivity shocks. Executives and entrepreneurs (collectively

[1]Some deviations from Gibrat's law are reported for young and small firms, as pointed out by Gabaix (2009) and Luttmer (2010). However, since our focus is on the right tail of income that is mainly affected by large productive firms, we set aside this issue in our analysis.

[2]Our model is consistent with another observation that the firm productivity distribution also follows a Pareto distribution (Mizuno et al., 2012).

called entrepreneurs in our model) can invest in their own firms' risky stocks or in risk-free assets. The dispersion of entrepreneurs' income is determined by the risk taken with their after-tax portfolio returns. To develop the model, we introduce transferable product lines and financial intermediaries that are new to the literature. The model is simple enough to allow for the analytical derivation of the stationary distributions of firms and income. Furthermore, the household income process is determined by partial differential equations (PDEs), enabling the straightforward numerical computation of an equilibrium transition path.

Second, by using the model, we evaluate how an unanticipated and permanent cut in the top marginal income tax rate affects the evolution of top incomes. A tax cut that favors risky stocks relative to risk-free assets would induce entrepreneurs to hold more risky stocks, leading to a more diffusive income process and a more dispersed distribution of entrepreneurs' income and wealth. Similarly, the tax cut would induce managers and firms to redesign their contracts toward an increased share of executive stock options to capture the benefit of the tax cut. To model this effect, we regard top marginal income tax in the real world as a tax on the risky stocks of entrepreneurs' and executives' own firms in the model, whereas taxes on equities in the real world are a tax on risk-free assets that are converted from a large variety of risky securities by financial intermediaries in the model. In the transition dynamics, a one-time tax cut leads to a slow-moving evolution of the distribution. The evolution of the distribution is analytically derived as PDEs. By using the PDEs with calibrated parameters, we numerically compute the transition dynamics of the income distribution assuming that the tax cut occurred in 1970. We show that our model matches the decline in the Pareto exponent of the income distribution and the trend of increasing top income shares observed in the U.S. in recent decades.

Third, we explore the general equilibrium implications of our model. Our model implies that a tax cut has no quantitative effects on the per-capita output and the capital–output ratio of the aggregate economy. In our model, a cut in the tax imposed on a financial asset does not quantitatively affect the return of the asset, because the asset price endogenously changes to offset the effect of tax change. The stable asset return leads to stable per-capita output and capital–output ratio. This irrelevance of tax relates to the well-known property of the "new" explanation of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005)—namely, that a change in dividend tax alone does not affect investment decisions. This property of the model is consistent with the stable growth rate of per-capita GDP and of the capital–output ratio observed in the postwar U.S. The irrelevance of tax to asset returns also produces implications with inequality. While we share views on the importance of a tax cut with Piketty and Saez (2003) and Piketty et al. (2011), in our model, a cut in the top marginal income tax rate will not in itself affect the income distribution, if there are no alternative assets. Instead, an income tax cut relative to other taxes—such as capital gains or corporate taxes—will affect the distribution through changes in entrepreneurs' portfolio choices. We show that the model's predictions on portfolio choice are consistent with observed measures of executives' incentive pay.

Our study builds on several others that investigate why the income distribution follows a Pareto distribution. Gabaix and Landier (2008) construct a model of executive pay. By assuming that the firm size distribution follows Zipf's law and the CEO's talent follows a certain distribution, they show that the CEO's pay follows a Pareto distribution. By using the model, they interpret that rising CEO pay in the U.S. in recent decades has resulted from rising firm values. Their model has the advantage of being consistent with both Zipf's law of firms and Pareto's law of incomes, similar to ours. However, their model predicts a constant Pareto exponent. Jones and Kim (2014) extend the model to be consistent with the recent decline in the Pareto exponent of the income distribution in the U.S. Compared to these studies, the contribution of ours is to build a model that generates both Zipf's and Pareto's laws from the productivity shocks of firms, without assuming particular underlying distributions.

Another thread of the literature, dating back to Champernowne (1953) and Wold and Whittle (1957), explains Pareto's law of incomes by idiosyncratic shocks on household wealth.[3] Most of these studies use partial equilibrium or endowment models that abstract from production.[4] As Jones (2015) notes, however, analyses that abstract from general equilibrium forces tend to generate unsatisfactory comparative statics. Recently, Nirei and Aoki (2016) and Benhabib et al. (2015) extend the framework to standard Bewley models, that is, dynamic general equilibrium models of heterogeneous households with production, and show that idiosyncratic shocks on firms' productivities generate Pareto's law of incomes in the environment. Our study follows this approach. In our model, behaviors of both households and firms are essential in determining Pareto's law. Compared with previous studies, we feature a model that explains not only Pareto's law of incomes but also Zipf's law of firms, both generated from the productivity shocks of firms. Previous studies can explain only one of these laws, because the entrepreneur of a firm possesses all of the firm's capital and thus the entrepreneur's wealth becomes proportional to the firm's size. We resolve this problem by incorporating the entrepreneur's portfolio choice into the model, in which an entrepreneur owns only a part, not all, of the firm's residual claim. This feature of the model characterizes our explanation as to how the recent tax cut has affected the evolution of top incomes.

The closest studies to ours are perhaps Kim (2013) and Jones and Kim (2014). Kim (2013) builds a human capital accumulation model with idiosyncratic shocks that generate Pareto's law of incomes and analyzes the impact of the cut in top marginal income tax in recent decades on the Pareto exponent of the income distribution. Jones and Kim (2014) extend the human capital model to an endogenous growth setting, incorporating creative destruction. In contrast to their studies, we build a model that also explains Zipf's law of firms.[5] The

[3]This approach requires some additional features to prevent the income and wealth distributions from diverging in order to obtain Pareto's law. The overlapping generations setting used by Dutta and Michel (1998) and Benhabib et al. (2011 and 2016), and the lower bound on savings used in Nirei and Souma (2007), Nirei and Aoki (2016) and Benhabib et al. (2015) are examples of the features that prevent the distribution from diverging.

[4]Exceptions include Dutta and Michel (1998) and Toda (2014) who construct general equilibrium models with production. The properties of these models are similar to those of endowment models, as they are AK (and AL) type models in which the asset returns and wage income are independent of allocation in production.

[5]Kim (2013) does not consider the firm-side problem. In Jones and Kim (2014),

model predictions also differ. For example, an income tax cut in their model encourages human capital accumulation among top income earners, resulting in a higher per-capita output in the U.S. in recent decades than in previous periods or in other countries such as France. By contrast, in our model, an income tax cut does not directly affect capital accumulation.

Finally, our model is closely related to the general equilibrium models of firm size distribution that explain Zipf's law of firms (for a survey, see Luttmer, 2010). Following the literature, we generate Zipf's law of firms through Gibrat's law and a minimum limit of firms. As an extension of this literature, we devise a model of firms with multiple product lines and entrepreneur-specific shocks that yields a reflected random growth in firm size.

The rest of the paper is organized as follows. Section 2 presents the dynamic general equilibrium model. Section 3 discusses the firm-side properties of the model and derives Zipf's law of firms. Section 4 describes the aggregate dynamics of the model and defines the equilibrium. Section 5 illustrates how the household wealth and income distributions follow a Pareto distribution in the steady state. Section 6 analyzes how a tax cut affects top incomes in our model and compares the results with the data. Finally, in Section 7, we present our concluding remarks.

2 Model

We build a Bewley economy with a continuum of households (workers and entrepreneurs), a continuum of firms, and financial intermediaries. A simple sketch of the model is as follows. Each firm has a continuum of products. Firms can trade the products, by which they maintain the minimum number of employees that is exogenously imposed. Each firm bears an idiosyncratic productivity shock that is specific to entrepreneurs who manage the firm. Thus, the idiosyncratic shock hits the production of all the products of a firm. This property results in Gibrat's law, which generates Zipf's law of firm size

the entrepreneurs' income distribution becomes proportional to the firm size distribution because each entrepreneur acquires all of the firms' rent.

by combining with the minimum firm size requirement. Competitive financial intermediaries convert a proportion of risky stocks into risk-free assets. Entrepreneurs are compensated by stocks, and they choose how to divide their portfolio between risky stocks and risk-free assets. The value of the stocks depends on the firms' idiosyncratic productivity shocks. Thus, the productivity shocks generate Pareto distributions of entrepreneurs' wealth and income. In this environment, income tax affects the Pareto distribution through the portfolio choice of the entrepreneurs. In what follows, we present a formal dynamic general equilibrium model and derive the PDEs that describe the transition of the wealth distribution.

2.1 Households

There is a continuum of households with measure 1. As in Blanchard (1985), each household is discontinued by a Poisson hazard rate ν and is replaced by a newborn household that has no bequest.[6] Households participate in a pension program. If a household dies, all of its non-human capital is distributed to living households. A living household receives the pension premium rate ν times its financial assets.

Households consist of entrepreneurs and workers. Measure E of households are entrepreneurs and the remaining $1 - E$ are workers. An entrepreneur as well as a worker provides one unit of labor and earns wage income w_t. Households also receive a government transfer tr_t. Among these households, only entrepreneurs manage firms. An entrepreneur has the benefit of holding the stocks of his firm relatively cheaper, as is explained subsequently. Whether a household is born as an entrepreneur or a worker is exogenously determined. An entrepreneur stochastically switches to a worker at constant hazard rate p_f.[7] Hence, there are two types of workers: workers who were born as workers, whom we call innate workers, and workers who were born as entrepreneurs,

[6] Our model assumes away the bequest motive of households. A justification for the assumption is that as Kaplan and Rauh (2013) argue, "[t]hose in the Forbes 400 are less likely to have inherited their wealth or to have grown up wealthy."

[7] It is possible to extend the model by incorporating the transition of a worker to become an entrepreneur.

whom we call former entrepreneurs.[8]

Household i chooses sequences of consumption $c_{i,s}$ and an asset portfolio to maximize the expected discounted log utility

$$\mathsf{E}_t \int_t^\infty e^{-(\beta+\nu)s} \ln c_{i,s} \; ds,$$

where β is the discount rate. A worker holds his wealth in (i) a risk-free bond $b_{i,t}$ and (ii) human capital h_t that consists of current and future wage incomes w_t and government transfers tr_t. The risk-free bond yields return r_t^f (and pension premium ν) with certainty. The human capital is defined by $h_t \equiv \int_t^\infty (w_u + tr_u) e^{-\int_t^u (\nu + r_s^f) ds} du$, whose return is

$$(\nu + r_t^f) h_t = (w_t + tr_t) + dh_t/dt.$$

An entrepreneur can hold (iii) risky stocks of his firm $s_{i,t}$ as an asset in addition to (i) and (ii). Let $q_{i,t}$ and $d_{i,t}$ be the price and dividend of the risky stocks, respectively. Then, the return on the risky stock 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 the risky stock, $B_{i,t}$ is the Wiener process, and $\mu_{q,t}$ and $\sigma_{q,t}$ are endogenous parameters determined in equilibrium. Note that the risky stocks obtained by entrepreneurs in the model capture the incentive scheme for executive compensation in the real world. Therefore, we calibrate the tax on risky stocks τ^e by the top marginal income tax rate in our numerical analysis. In Section 6.7, we compare our formulation of executive pay with that in previous studies and compare our model's prediction with the data.

Let $a_{i,t} \equiv s_{i,t} q_{i,t} + b_{i,t} + h_t$ denote the wealth of a household (we refer to the

[8] In the model, either the death rate ν or the rate of exiting entrepreneur p_f must be strictly positive in order to generate Pareto's law of incomes. We introduce both types of hazard events for a quantitative reason. Without either of these two types, the mobility of a household's wealth or income level becomes very slow, or the Pareto exponent of the income distribution becomes very low, compared with the data (see Gabaix et al., 2015 and Jones and Kim, 2014).

sum of assets as wealth) and $x_{i,t} \equiv s_{i,t}q_{i,t}/a_{i,t}$ the share of risky stocks. The household's wealth accumulates according to the following process:

$$\begin{aligned} da_{i,t} &= (\nu(s_{i,t}q_{i,t} + b_{i,t}) + \mu_{q,t}s_{i,t}q_{i,t} + r^f_t b_{i,t} + (\nu + r^f_t)h_t - c_{i,t})dt + \sigma_{q,t}s_{i,t}q_{i,t}dB_{i,t} \\ &= \mu_{a,t}a_{i,t}dt + \sigma_{a,t}a_{i,t}dB_{i,t}, \end{aligned} \tag{1}$$

where

$$\begin{aligned} \mu_{a,t} &\equiv \nu + \mu_{q,t}x_{i,t} + r^f_t(1 - x_{i,t}) - c_{i,t}/a_{i,t} \\ \sigma_{a,t} &\equiv \sigma_{q,t}x_{i,t}. \end{aligned}$$

Note that $dB_{i,t}$ forms a multiplicative shock to the current wealth level $a_{i,t}$.

Let V^i denote the value function of household characteristics i, where $i = e$ if the household is an entrepreneur, $i = \ell$ if he is a worker, $i = w$ if he is an innate worker, and $i = f$ if he is a former entrepreneur. An innate worker w and a former entrepreneur f do not change their household characteristics (i.e., $i' = i$, where i' denotes the characteristics in the next period), while an entrepreneur e may change to f in the next period. Let $\mathbf{S}_t$ denote a set of aggregate state variables defined in Section 4. By using these notations, the household's dynamic programming problem is specified as follows:

$$V^i(a_{i,t}, \mathbf{S}_t) = \max_{c_{i,t}, x_{i,t}} \ln c_{i,t}dt + e^{-(\beta+\nu)dt}\,\mathsf{E}_t[V^{i'}(a_{i,t+dt}, \mathbf{S}_{t+dt})] \tag{2}$$

subject to (1).

The household problem is a variant of Merton's dynamic portfolio problem (Campbell and Viceira, 2002 for a reference). It is well known that the solution to the problem under the log utility follows the myopic rules,

$$x_{i,t} = \begin{cases} \frac{\mu_{q,t} - r^f_t}{\sigma^2_{q,t}}, & \text{if } i = e, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

$$v_{i,t} = \beta + \nu, \tag{4}$$

where $v_{i,t}$ is the consumption–wealth ratio (see Appendix A for the derivations), and satisfies the transversality condition

$$\lim_{T\to\infty} e^{-(\beta+\nu)T}\,\mathsf{E}_0\left[V^i(a_{i,T},\mathbf{S}_T)\right]=0. \tag{5}$$

Note that the household decision rules (3)–(5) do not depend on the probability of an entrepreneur switching to a worker p_f. This property, which results from the log utility assumption, is convenient when we numerically solve transition dynamics.

2.2 Firms and the financial market

A continuum of firms with measure E produces differentiated goods. Each firm is managed by an entrepreneur. As in McGrattan and Prescott (2005), each firm issues shares, and owns and self-finances capital. We assume that the ownership of a share of a firm incurs transaction costs, except for the case where an entrepreneur directly owns stocks of his firm in the form of non-voting shares. Financial intermediaries also own the firm's shares by bearing the transaction costs. The financial intermediaries combine the shares of all firms and issue risk-free bonds to households. Thus, financial intermediaries provide the means for households to diversify the firms' idiosyncratic shocks. At the competitive level of the risk-free rate, workers prefer to hold risk-free bonds rather than to own shares by paying transaction costs. The transaction cost is denoted by ι per dividend $d_{e,t}$ of a firm managed by entrepreneur e. Since financial intermediaries own all of the voting shares, firms maximize expected profits following the interest of financial intermediaries. Then, the market value of a firm achieves 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

In this model, returns and risks on risky stocks are ex ante identical across firms, and shocks on the risky stocks are uncorrelated with each other. Then, a financial intermediary maximizes residual profit by diversifying the risks on risky stocks and issuing risk-free assets as follows:

$$\max_{s^f_{e,t}} \; \mathsf{E}_t \left[\int_0^E \left\{ (1-\tau^f)(1-\iota) d_{e,t} dt + dq_{e,t} \right\} s^f_{e,t} de \right] - r^f_t dt \left(\int_0^E q_{e,t} s^f_{e,t} de \right),$$

where $s^f_{e,t}$ is the shares of firm e owned by the financial intermediary and τ^f is the dividend tax, which is different from the tax rate on risky stocks τ^e. We interpret τ^f in the numerical analysis as a combination of capital gains and corporate income taxes. In Section 6, we account for the evolution of top incomes by the change in the difference between these tax rates. The solution of the problem leads to

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

2.2.2 Firm's problem

Firm e owns a continuum measure $\bar{n}(e)$ portfolio of product lines, and each product line produces a differentiated good. The total measure of product lines in the economy is constant and normalized to 1. Firms can buy and sell the product lines through merger and acquisition (M&A), as we explain more precisely later.

The product $n \in [0, \bar{n}(e)]$ of firm $e \in [0, E]$ is produced with a Cobb–Douglas production technology:

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

where $y_{n,e,t}$ is output, $z_{n,e,t}$ is productivity, $k_{n,e,t}$ is the capital input, and $\ell_{n,e,t}$ is the labor input. The productivity of the product line evolves as

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

where $B_{e,t}$ follows the Wiener process. Note that $dB_{e,t}$ does not depend on n. That is, we assume that productivity shocks are perfectly correlated between the product lines within firm e, but uncorrelated with shocks in other firms. A possible interpretation of the correlation of shocks is that the shocks are caused by managerial decisions. Note that the productivity levels, rather than shocks, can be different between product lines even within a firm. This can occur when the initial productivity levels vary across products or when a firm buys product lines from other firms. $dB_{e,t}$ is a multiplicative shock to productivity, because the shock is multiplied by its productivity level $z_{n,e,t}$. Under the formulation, when the firm's size is proportional to its productivity, as shown below, Gibrat's law of firms holds; that is, the growth rate of the firm is independent of the firm's size.

The above setting is reminiscent of those in Klette and Kortum (2004) and Luttmer (2011), who construct models of the firm heterogeneity. We construct such a model to derive Zipf's law in a tractable way. There are a few remarks about our model. First, the product lines in our model are continuous, while in Klette and Kortum (2004) and Luttmer (2011), they are discrete. This is for tractability and ease of calculation. Second, each product line incurs productivity shocks that are common within a firm. The setting is different from Klette and Kortum (2004) and Luttmer (2011), in which shocks affect the number of product lines firms own and product lines do not incur productivity shocks.

A firm chooses the investment level $dk_{n,e,t}$ and employment $\ell_{n,e,t}$ of a product line to maximize the value of the product line $q_{n,e,t} = q(k_{n,e,t}, z_{n,e,t}, \mathbf{S}_t)$:

$$r_t^f q(k_{n,e,t}, z_{n,e,t}, \mathbf{S}_t)dt = \mathsf{E}_t \left[\max_{dk_{n,e,t}, \ell_{n,e,t}} (1-\tau^f)(1-\iota) d_{n,e,t} dt + dq(k_{n,e,t}, z_{n,e,t}, \mathbf{S}_t) \right]. \tag{7}$$

Here, the dividend $d_{n,e,t}$ of the product line consists of

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

where $p_{n,e,t}$ and $y_{n,e,t}$ are, respectively, the price and quantity of the good, $k_{n,e,t}$ is the capital invested in the product line, w_t is the wage rate, and δ is the depreciation rate. The value and dividends of a firm are equal to the sums of

$q_{n,e,t}$ and $d_{n,e,t}$ over the firm's product portfolio:

$$q_{e,t} = \int_0^{\bar{n}(e)} q_{n,e,t} \, dn, \quad \text{and } d_{e,t} = \int_0^{\bar{n}(e)} d_{n,e,t} \, dn.$$

By solving the maximization problem, we obtain the following conditions (see Appendix B for details):

$$\mathsf{MPK}_t \equiv r_t^f + \delta = \frac{\partial p_{n,e,t} y_{n,e,t}}{\partial k_{n,e,t}}, \tag{8}$$

$$w_t = \frac{\partial p_{n,e,t} y_{n,e,t}}{\partial \ell_{n,e,t}}. \tag{9}$$

Note that in the model, the marginal product of capital (MPK) becomes the same across product lines and between firms, because the stochastic discount factor of financial intermediaries is not correlated with the shock of firm e. In addition, note that taxes do not distort MPK because the taxes in the model are imposed on dividends. As argued in the "new view" literature of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005), these do not distort MPK.

A key factor to obtain Zipf's law of firm size is to impose a minimum level of firm size (Gabaix, 2009; Luttmer, 2010). Following Rossi-Hansberg and Wright (2007), we assume a minimum level of employment $\ell_{\min}$ for each firm, that is,

$$\int_0^{\bar{n}(e)} \ell_{n,e,t} \, dn \geq \ell_{\min}.$$

We assume that a firm maintains the minimum level of employment by purchasing product lines from other firms when the firms' employment level becomes smaller than $\ell_{\min}$. At a price equal to the value of a product line $q_{n,e,t}$, all firms are indifferent between buying and selling the product line. An acquiring firm pays the price by newly issuing stocks. For simplicity, we assume that all firms sell a proportion of their product lines, so that the value of the product lines sold during time $[t, t + dt]$ is equal to $mq_{e,t}dt$. The rate m is determined endogenously, so that the transactions of product lines clear in aggregate.

Provided that the initial distribution of $z_{n,e,0}$ within firm e is non-degenerate, it is always possible to find a reallocation of product lines that satisfies the needs of acquiring firms.

Our model of M&A provides a convenient mechanism by which the minimum size is maintained and, at the same time, a firm's value $q_{e,t}$ is linearly related to productivity, as we show in Section 3. From this linearity, we confirm that an acquiring firm can recover the minimum employment level $\ell_{\min}$ by purchasing product lines with values totaling $q_{\min} - q_{e,t}$.

2.3 Aggregation and market conditions

We now consider the market conditions for the aggregate economy. We use upper-case letters to denote the aggregate variables throughout the paper. Goods produced in the product lines are aggregated according to

$$Y_t = \left(\int_0^E \int_0^{\bar{n}(e)} y_{n,e,t}^{\frac{\phi-1}{\phi}} \, dnde \right)^{\frac{\phi}{\phi-1}}, \quad \phi > 1. \tag{10}$$

We assume that the aggregate good Y is produced competitively and normalize the price of the aggregate good to 1. The other aggregate variables are simply summed up over households or the product lines of firms. For example, let C_t and K_t be the aggregate consumption and capital. Then, $C_t \equiv \int_0^1 c_{i,t} \, di$ and $K_t \equiv \int_0^E \int_0^{\bar{n}(e)} k_{n,e,t} \, dnde$.

The market-clearing condition for final goods is

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

where $A_{e,t}$ is the wealth (the sum of financial assets and human capital) owned by entrepreneurs, Q_t is the aggregate financial asset, and D_t is the aggregate dividends. $(1 - A_{e,t} x_{e,t}/Q_t)$ is the share of stocks owned by financial intermediaries in the aggregate financial asset. Thus, the last term on the left-hand side of the equation indicates the proportion of the final goods used for transaction costs when financial intermediaries convert the stocks into risk-free bonds.

The total measures of existing product lines and labor supply are normalized to 1. Thus, the market-clearing condition for product lines is

$$\int_0^E \int_0^{\bar{n}(e)} dnde = \int_0^E \bar{n}(e)de = 1.$$

Correspondingly, the labor market-clearing condition is

$$\int_0^E \int_0^{\bar{n}(e)} \ell_{n,e,t} \; dnde = 1. \tag{11}$$

The market-clearing condition for the shares of firm e is

$$s_{e,t} + s_{e,t}^f = 1,$$

where $s_{e,t}$ is the shares owned by firm e's entrepreneur according to (3) and $s_{e,t}^f$ is the shares owned by financial intermediaries. We assume that all tax revenues are rebated to households as lump-sum government transfers in each period. Finally, the market-clearing condition for the risk-free bonds is

$$\int_0^1 b_{i,t}di = \int_0^E q_{e,t}s_{e,t}^f de.$$

3 Firm-Side Properties

Before we define the equilibrium and solve the model, we review some of the firm-side properties. Closed-form expressions for the product line variables $(\ell_{n,e,t}, k_{n,e,t}, d_{n,e,t})$ are obtained, given r_t^f. The heterogeneity of the product line variables stems solely from productivity. We then show that the stationary distribution of firm productivity depends only on the minimum employment level $\ell_{\min}$ and the entrepreneur measure E, and that Zipf's law of firm size is obtained when $\ell_{\min}$ is sufficiently small.

3.1 Firm-side variables

We express the product line variables as functions of relative productivity, which we denote by $\tilde{z}_{n,e,t} \equiv z_{n,e,t}^{\phi-1}/\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}$. Note that $\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}$ is the average of $z_{n,e,t}^{\phi-1}$ over all product lines in the economy. We obtain the following relations by using the firm's first-order conditions (FOCs) (8) and (9), together with the aggregation condition (10) and the labor market condition (11) (see Appendix B for the derivations):

$$\ell_{n,e,t} = \frac{p_{n,e,t}y_{n,e,t}}{\overline{py}_t} = \frac{k_{n,e,t}}{\overline{k}_t} = \frac{q_{n,e,t}}{\overline{q}_t} = \tilde{z}_{n,e,t}, \tag{12}$$

$$dd_{n,e,t} = \overline{d}_t \tilde{z}_{n,e,t}dt - (\phi-1)\sigma_z \overline{k}_t \tilde{z}_{n,e,t} dB_{e,t}, \tag{13}$$

where

$$\overline{py}_t \equiv \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}}, \tag{14}$$

$$\overline{k}_t \equiv \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}\right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}}, \tag{15}$$

$$\overline{q}_t \equiv \overline{d}_t \int_t^{\infty} (1-\tau^f)(1-\iota)\exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\} du, \tag{16}$$

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

where $\mu_{k,t}$ and $\mu_{d,t}$ are the expected growth rates of $k_{n,e,t}$ and $d_{n,e,t}$, respectively. Note that the dispersion of the product line variables is solely determined by relative productivity $\tilde{z}$. This property significantly simplifies the computation of the transition paths.

3.2 Zipf's law of firm size

In this study, we measure the size of a firm by its employment. By using (12), the employment growth of a firm is derived as

$$d\ln \ell_{e,t} = -\left(\frac{(\phi-1)^2\sigma_z^2}{2} + m\right)dt + (\phi-1)\sigma_z dB_{e,t}. \tag{17}$$

In the above equation, m is the measure of the product lines sold to other firms. Given this differential equation for firm size (employment), the Fokker–Planck equation (also called the Kolmogorov forward equation) for the probability density of the firm size distribution $f_s(\ln \ell, t)$ is obtained as

$$\frac{\partial f_s(\ln \ell, t)}{\partial t} = -\left(\frac{(\phi - 1)^2 \sigma_z^2}{2} + m\right) \frac{\partial f_s(\ln \ell, t)}{\partial \ln \ell} + \frac{(\phi - 1)^2 \sigma_z^2}{2} \frac{\partial^2 f_s(\ln \ell, t)}{\partial (\ln \ell)^2}.$$

In this section, we solve an invariant distribution for firms. When $\partial f_s(\ln \ell, t)/\partial t = 0$, the Fokker–Planck equation with the constraint $\ell_{e,t} \geq \ell_{\min}$ has a solution in an exponential form,

$$f_s(\ln \ell) = F_0 \exp(-\lambda \ln \ell), \tag{18}$$

where the coefficients satisfy

$$F_0 = \lambda \ell_{\min}^{\lambda}, \quad \lambda = 1 + \frac{m}{(\phi - 1)^2 \sigma_z^2 / 2}. \tag{19}$$

Equation (18) shows that the distribution of $\ln \ell_{e,t}$ follows an exponential distribution. Through a change of variables, it is shown that the distribution of $\ell_{e,t}$ follows a Pareto distribution whose Pareto exponent is λ. When $\ell_{e,t}$ follows a Pareto distribution, we obtain

$$1 = \int_0^E \ell_{e,t} de = E \times \int_{\ell_{\min}}^{\infty} \ell_{e,t} f_s(\ln \ell_{e,t}) \frac{d \ln \ell_{e,t}}{d \ell_{e,t}} \; d\ell_{e,t} = \frac{\lambda \ell_{\min}}{\lambda - 1}.$$

By rearranging this equation, we obtain

$$\lambda = \frac{1}{1 - \frac{\ell_{\min}}{1/E}}. \tag{20}$$

This equation shows that λ approaches 1 if $\ell_{\min}$ is sufficiently small compared with average employment per firm $1/E$. Hence, we obtain Zipf's law for firms' employment $l_{e,t}$ as well as for firms' sales $p_{e,t} y_{e,t}$ or capital input $k_{e,t}$, when the minimum size of a firm is sufficiently small.

4 Aggregate Dynamics and Equilibrium of the Model

In this model, we obtain the dynamics of the aggregate variables independently of the heterogeneities within entrepreneurs, innate workers, and former entrepreneurs. This separation between aggregates and cross-sectional heterogeneity stems from the model property that the household's policies are independent of its wealth level and that the firm's policies are linear in relative productivity. We first summarize these properties and then define the equilibrium of the model.

4.1 Aggregate dynamics of the model

The growth rate of the aggregate output on the balanced growth path is

$$g \equiv \left\{ \left(\mu_z - \frac{\sigma_z^2}{2} \right) + (\phi - 1)\frac{\sigma_z^2}{2} \right\} / (1 - \alpha),$$

which is confirmed by aggregating (14). We detrend the aggregate variables by growth rate g and denote them by tilde, for example $\tilde{K}_t \equiv K_t / \left(e^{gt} \cdot \overline{\mathsf{E}} \left\{ z_{n,e,0}^{\phi-1} \right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}} \right)$.

Let $\tilde{A}_{e,t}$, $\tilde{A}_{w,t}$, and $\tilde{A}_{f,t}$ denote the detrended aggregate wealth (the sum of financial assets and human capital) of entrepreneurs, innate workers, and former entrepreneurs, respectively, while $\tilde{H}_t$ is detrended aggregate human capital. The sum of $\tilde{A}_{e,t}$, $\tilde{A}_{w,t}$, and $\tilde{A}_{f,t}$ is equal to the aggregate wealth of all households $\tilde{A}_t$. We denote the set of the detrended aggregate variables by $\tilde{\mathbf{S}}_t \equiv (\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t)$, whereas $\mathbf{S}_t \equiv e^{gt}\tilde{\mathbf{S}}_t$ denotes the original aggregate variables.

We show below that the aggregate dynamics of the detrended variables are reduced to ordinary differential equations:

$$\frac{d\tilde{\mathbf{S}}_t}{dt} = \mu_{\tilde{\mathbf{S}}}(\tilde{\mathbf{S}}_t) \equiv \left(\frac{d\tilde{A}_{e,t}}{dt}, \frac{d\tilde{A}_{w,t}}{dt}, \frac{d\tilde{A}_{f,t}}{dt}, \frac{d\tilde{H}_t}{dt}, \frac{d\tilde{K}_t}{dt} \right), \tag{21}$$

and price variables $(r_t^f, \mu_{q,t}, \sigma_{q,t})$ are functions of $\tilde{\mathbf{S}}_t$. Given $\tilde{\mathbf{S}}_t$, the aggregate dynamics (21) and price functions are obtained through the following steps:

1. Given $\tilde{K}_t$, from (15), r_t^f and MPK_t are obtained by

$$r_t^f + \delta = \mathsf{MPK}_t = \frac{\alpha(\phi-1)/\phi}{\tilde{K}_t^{1-\alpha}}.$$

2. Given $\tilde{A}_t$ and $\tilde{H}_t$, we obtain $\tilde{C}_t = (\beta+\nu)\tilde{A}_t$ from (4) and $\tilde{Q}_t = \tilde{A}_t - \tilde{H}_t$. Given MPK_t, $\tilde{Y}_t = \overline{py}_t/e^{gt}$ is pinned down. Given the variables obtained above and (3), $d\tilde{K}_t/dt$ is, jointly with $\tilde{D}_t$ and $x_{e,t}$, computed by the following equations,

$$\begin{aligned}\frac{d\tilde{K}_t}{dt} =& \tilde{Y}_t - \delta\tilde{K}_t - \tilde{C}_t - \iota\left(1 - \frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\right)\tilde{D}_t - g\tilde{K}_t,\\ \tilde{D}_t =& (1-(1-\alpha)(\phi-1)/\phi)\tilde{Y}_t - (\delta+g)\tilde{K}_t - \frac{d\tilde{K}_t}{dt},\end{aligned}$$

and (3). Note that $\mu_{q,t}$ and $\sigma_{q,t}$ in (3) are the functions of $\tilde{K}_t$, $\tilde{Q}_t$, and $\tilde{D}_t$ (see Appendix B.2).

3. Given the variables obtained above, $(d\tilde{A}_{e,t}/dt, d\tilde{A}_{w,t}/dt, d\tilde{A}_{f,t}/dt)$ are computed as follows:

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

where $\mu_{ae,t}$ and $\mu_{a\ell,t}$ are the $\mu_{a,t}$'s of an entrepreneur and a worker, respectively, and are computed by (1) and (4). Finally, given the variables obtained above, $d\tilde{H}_t/dt$ is computed by

$$\frac{d\tilde{H}_t}{dt} = -(\tilde{w}_t + \tilde{tr}_t) + (\nu + r_t^f - g)\tilde{H}_t, \tag{22}$$

where $\tilde{w}_t = ((1-\alpha)(\phi-1)/\phi)\tilde{Y}_t$ and $\tilde{tr}_t = \left\{\frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\tau^e + \left(1 - \frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\right)\tau^f\right\}\tilde{D}_t$.

4.2 Definition of a competitive equilibrium

By using the property of the aggregate dynamics, we now define the equilibrium of the model. To simplify the analysis, we specify the initial conditions in the following manner. First, the initial capital of a product line is proportional to the product line's productivity, that is, $\tilde{k}_{n,e,0} \propto \tilde{z}_{n,e,0}$. Then, the initial value of the product line is also proportional to productivity, that is, $\tilde{q}_{n,e,0} = \tilde{z}_{n,e,0}\tilde{Q}_0$, where $\tilde{Q}_0 = \tilde{A}_0 - \tilde{H}_0$. Second, the initial firm size distribution follows (18) and (19). Third, we assume that all stocks are initially owned by households, and except for those held by entrepreneurs, these stocks are sold to financial intermediaries in period 0.[9] Let $s^i_{e,0}$ be the initial shares of firm e held by household i (then, $\int_0^1 s^i_{e,0} di = 1$).

A competitive equilibrium of the model, given the law of motion of the product line's productivities $\{z_{n,e,t}\}_t$, the initial capital of product lines in firms $\tilde{k}_{n,e,0} \propto \tilde{z}_{n,e,0}$, the initial shares of firms held by households $s^i_{e,0}$, taxes τ^e and τ^f, and the measure of entrepreneurs E, is a set of household variables $\{x_{i,t}, v_{i,t}, \tilde{a}_{i,t}\}_{i,t}$, price variables $\tilde{q}_{e,0}$ and $\{w_t, r^f_t, \mu_{q,t}, \sigma_{q,t}\}_t$, and aggregate variables $\{\tilde{\mathbf{S}}_t\}_t$, such that

- the household variables $\{x_{i,t}, v_{i,t}, \tilde{a}_{i,t}\}_t$, where $\tilde{a}_{i,0} = \int_0^E \tilde{q}_{e,0} s^i_{e,0} de + \tilde{H}_0$, are chosen according to the household's decisions on (3) and (4), and the law of motion for wealth (1), and satisfy the transversality condition (5),
- the price variables $\tilde{q}_{e,0}$ and $\{w_t, r^f_t, \mu_{q,t}, \sigma_{q,t}\}_t$ are determined so that markets for labor, final goods, product lines, shares, and risk-free bonds clear, given $\tilde{\mathbf{S}}_t$ and initial price condition $\tilde{q}_{n,e,0} = \tilde{z}_{n,e,0}\tilde{Q}_0$,
- and the aggregate variables $\{\tilde{\mathbf{S}}_t\}_t$ evolve according to (21).

[9] We assume that the sellout to financial intermediaries is mandatory. We can relax the assumption and allow households by paying transaction costs ι to hold risky stocks of the firms not managed by them.

5 Stationary Distribution of Households' Wealth

In this model, stationary wealth distributions are derived analytically for each type of household. We show below that the wealth distributions of entrepreneurs, innate workers, and former entrepreneurs are all Pareto distributions. We also discuss that the wealth, income, and consumption distributions of all households follow a Pareto distribution at the upper tail, whose Pareto exponent coincides with that of the wealth distribution of entrepreneurs.

5.1 Wealth distribution of entrepreneurs

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

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

We omit the time subscript for variables that are constant in the steady state.

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

We obtain the wealth distribution of entrepreneurs by combining the above property with the assumption of the constant probability of death. The probability density function of the log wealth of entrepreneurs, $f_e(\ln \tilde{a})$, becomes a double-exponential distribution (see Appendix D for the derivations in this section).[10]

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

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

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

where

$$\psi_1 \equiv \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2}\left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} - 1\right), \psi_2 \equiv \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2}\left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} + 1\right),$$

$$\theta \equiv \sqrt{2(\nu + p_f)\sigma_{ae}^2 + (\mu_{ae} - g - \sigma_{ae}^2/2)^2}.$$

This result shows that the wealth distribution of entrepreneurs follows a double-Pareto distribution (Reed, 2001; Benhabib et al., 2016; Toda, 2014), whose Pareto exponent at the upper tail is ψ_1.

5.2 Wealth distribution of innate workers

A worker's wealth $\tilde{a}_{\ell,t}$, if he does not die, evolves as

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

Under the wealth process, the probability density function of innate workers, $f_w(\ln \tilde{a})$, becomes

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

The result shows that the log wealth of innate workers follows an exponential distribution, which implies that their wealth levels follow a Pareto distribution. With the parameter values in the numerical analysis, the trend growth of workers' wealth is close to the trend growth of the economy, that is, $\mu_{a\ell} \approx g$. Then, the detrended wealth of innate workers is concentrated at the level around $\tilde{h}$.

5.3 Wealth distribution of former entrepreneurs

The wealth distribution of former entrepreneurs is determined by the entrepreneurs' wealth distribution, the Poisson rate p_f with which each entrepreneur leaves the firm, and the wealth process after the entrepreneur becomes a worker. We can analytically derive the stationary wealth distribution of former entrepreneurs.

Here, for brevity, we only report the probability density function of former entrepreneurs, $f_f(\ln \tilde{a})$, for the case $\mu_{a\ell} \geq g$:

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

The probability density function for the region $\tilde{a} \geq \tilde{h}$ consists of two exponential terms. As the wealth level increases, the second exponential term, which represents the innate workers' distribution, declines faster than the first term, the entrepreneurs' distribution. Therefore, the Pareto exponent of the former entrepreneurs' wealth distribution becomes the same as that of entrepreneurs in the tail of the distribution (the same result applies to the case $\mu_{a\ell} < g$).

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

The distributions of entrepreneurs, innate workers, and former entrepreneurs determine the overall wealth distribution of households. We make two remarks on the overall distribution. First, the Pareto exponent at the upper tail of the households' wealth distribution is the same as that of the entrepreneurs' wealth distribution ψ_1. This is because the distribution of the smallest Pareto exponent dominates at the upper tail as noted above (see, e.g., Gabaix, 2009).

Second, the income and consumption distributions at the upper tail also follow the Pareto distribution with the same Pareto exponent as that of wealth ψ_1. This is because, in our model, the income and consumption of a household are always proportional to the household's wealth level.

6 Numerical Analysis

In this section, we numerically analyze how a reduction in the top marginal tax rate accounts for the evolution of top incomes in recent decades. In the baseline experiment, we assume that an unexpected and permanent tax cut occurs in 1970. As a robustness check, we also conduct numerical exercises feeding the exact time path for these taxes into our model.

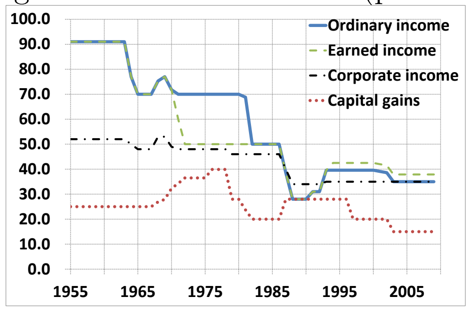
We choose 1970 as the year of the structural change, based on several empirical studies suggesting that inequality began to grow after the 1970s (see, e.g., Katz and Murphy, 1992; Piketty and Saez, 2003). Some political scientists also point out that U.S. politics began to favor industries after the 1970s (Hacker and Pierson, 2010). Indeed, top marginal earned income tax declined from 77% to 50% around 1970 alone (see Figure 1). This would make entrepreneurs anticipate a subsequent cut in top earned income tax, the most important variable in our analysis to account for the evolution of top incomes. These factors suggest that a structural change has occurred since the 1970s.

In our model, a tax cut affects top incomes by changing entrepreneurs' incentives to invest in risky stocks. In the tax parameter set we calibrate below, the tax rate on risky stock τ^e becomes lower after 1970 relative to the tax rate on the risk-free asset τ^f. This shift in tax structure induces entrepreneurs to increase the share of risky stocks in their asset portfolios, which leads to a decline in the Pareto exponent and an increase in top income share in our model.

6.1 Tax rates

In our model, entrepreneurs' holdings of own risky stocks correspond to the incentive pay for executives, such as employee stock options. Thus, we set the tax on risky stocks τ^e in our model to be equal to the top marginal earned income tax imposed on top executive pay. Meanwhile, the tax on risk-free assets τ^f captures the taxes that households bear when they hold equities through financial intermediaries. Thus, we set the tax on risk-free assets according to $1 - \tau^f = (1 - \tau^{\text{cap}})(1 - \tau^{\text{corp}})$, where τ^{cap} and τ^{corp} are the marginal tax rates for capital gains and corporate income, respectively. These tax rates are calibrated by using the top statutory marginal federal tax rates reported in Saez et al. (2012) (see Figure 1 and Table 1).

Figure 1: Federal tax rates (percent)



Note: The data are taken from Table A1 of Saez et al. (2012).

Table 1: Tax rates (percent)

	Pre-1970	Post-1970
Earned income tax, τ^{ord}	71.8	37.9
Capital gain tax, τ^{cap}	32.3	15.0
Corporate income tax, τ^{corp}	49.0	35.0
τ^{e}	71.8	37.9
τ^{f}	65.5	44.8

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

6.2 Calibration

The parameters are calibrated to the annual frequency data as in Table 2. The first five parameter values are standard. We assume for ν that the average length of life after a household begins working is 50 years. ϕ is set to 3.33, implying that 30% of a firm's sales is rent. The value of ϕ is lower than the standard value, because our model's treatment of entrepreneurial income is different from the data—in our model, an entrepreneur's income derives mainly from the firm's dividend, whereas in the data, executive compensation is categorized as labor income in most situations. A lower ϕ is chosen to take this into account. In addition, if ϕ is too high, the total value of an entrepreneur's risky stocks may exceed the total value of financial assets in the economy, provided that entrepreneurs choose $s_{i,t}$ according to (3). A low ϕ should be chosen to avoid this.

For p_f, we assume that the average term of office of an entrepreneur or an executive is 20 years. $\ell_{\min}$ is set to unity, that is, the minimum employment level is one person. The fraction of entrepreneurs in all households is set as $E = 0.05$, implying that the average employment of a firm is 20 persons. This is consistent with the data reported in Davis et al. (2007). Under $E = 0.05$, the Pareto exponent of the firm size distribution in the model is $1/(1-0.05) \approx 1.0526$, which is consistent with Zipf's law.[11] Note that the Pareto exponent of firm size does not depend on the tax rate. In our model, a tax cut affects only the income distribution but not the firm size distribution, which we find consistent with the data.

To calibrate firm-level volatility, we consider two cases. In Case A, we match with the average firm-level volatility of publicly traded firms, and in Case B, we match with that of both publicly traded and privately held firms. We match the estimates of firm-level employment volatility in Davis et al. (2007) with the model counterpart $(\phi - 1)\sigma_z$. The calibrated values are shown in Table 2. In Cases A and B, the transaction cost of financial intermediaries ι is calibrated to match the Pareto exponent in the pre-1970 steady state with the 1970 observation 2.53.

To cross-check the calibration of firm-level volatility using employment data, we compare the calibrated values with the firm's asset value volatilities. In our model, asset value volatility coincides with employment volatility $(\phi - 1)\sigma_z$. Moskowitz and Vissing-Jorgensen (2002, Table 6) report the standard deviation of the market equity returns of all public firms between 1953 to 1999 to be 17.0% and that of the smallest decile of public firms to be 41.1%. Several studies report the magnitude of the idiosyncratic volatility of stocks or risky assets owned by households. Flavin and Yamashita (2002), using the Panel Study of Income Dynamics from 1968 to 1992, find that the standard deviation

[11]Note that under these parameters, for small firms, the value of an entrepreneur's risky stock calculated by (3) exceeds the value of his firm. To resolve this problem, we assume that such an entrepreneur jointly runs a business with other entrepreneurs, such that the asset value of the entrepreneurs' risky stocks does not exceed the value of the joint firms. We assume that the productivity shocks of the joint firms move in the same direction. A possible reason for this assumption is that productivity shocks are caused by managerial decisions.

Table 2: Calibrated parameters

β	Discount rate	0.03
ν	Probability of death	1/50
α	Capital share	0.36
δ	Depreciation rate	0.06
g	Steady-state growth rate	0.02
ϕ	Elasticity of substitution	3.33
p_f	Probability of an entrepreneur quitting	1/20
$\ell_{\min}$	Minimum level of employment	1
E	Share of entrepreneurs in households	0.05

		Case A	Case B
$(\phi-1)\sigma_z$	Firm-level volatility of employment	25%	45%
ι	Transaction costs of financial intermediaries	0.502	0.601

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

of stocks owned by U.S. households is 24.2%. Calvet et al. (2007) report that the idiosyncratic volatility of assets in the portfolio of Swedish households around 2000 is 21.1%. Fagereng et al. (2016) find that the standard deviation of risky assets of Norwegian households in 2013 is 23.4%. In sum, the estimates on asset value volatilities fall near the range between Cases A and B.

6.3 Computation of transition dynamics

We compute the Pareto exponent of the household's 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 unexpectedly and permanently in 1970, and the economy moves toward the post-1970 steady state.

An advantage of our model is that the dynamics of the aggregate variables can be computed separately from the dynamics of the cross-sectional distributions. The transition dynamics of a set of the aggregate variables, $\tilde{\mathbf{S}}_t \equiv \mathbf{S}_t/e^{gt} =$

$(\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t)$ defined in Section 4.1, are computed by a shooting algorithm that pins down their initial values. When taxes change unexpectedly in 1970, prices also change suddenly, which affects the wealth distribution. We assume that while perfect risk-sharing for the unexpected change in asset values is achieved for the risk-free bonds and human capital, it does not work for the risky assets. Then, the wealth shares of entrepreneurs, innate workers, and former entrepreneurs, namely $A_{e,1970}/A_{1970}$, $A_{w,1970}/A_{1970}$, and $A_{f,1970}/A_{1970}$, change accordingly. The remaining initial variables, $\tilde{A}_{1970}$ and $\tilde{H}_{1970}$, are determined by using the shooting algorithm (for details, see Appendix C.1).

Next, given the transition of the aggregate variables calculated above, we compute the variables that determine the entrepreneurs' and workers' wealth processes, $\mu_{ae,t}$, $\sigma_{ae,t}$, and $\mu_{a\ell,t}$. By using these variables, the transition dynamics of the distribution can be computed by numerically solving the Fokker–Planck equations for the wealth distributions of entrepreneurs and workers, $f_e(\ln \tilde{a}, t)$ and $f_\ell(\ln \tilde{a}, t) \equiv f_w(\ln \tilde{a}, t) + f_f(\ln \tilde{a}, t)$, respectively, as follows:[12]

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

We impose the boundary conditions that $\lim_{\tilde{a}\to\infty} f_i(\ln \tilde{a}, t) = 0$ and that the probability density function of the wealth distribution at the lower bound $\tilde{a}_{\mathrm{LB}}$, $f_i(\ln \tilde{a}_{\mathrm{LB}}, t)$, moves linearly for 50 years between the pre-1970 and the post-1970 steady-state values.[13] Finally, we define a household's income as $\nu a_{i,t} + \mu_{q,t} x_{i,t} a_{i,t} + r_t^f (1 - x_{i,t}) a_{i,t}$. The income distribution can be computed after the aggregate dynamics and wealth distributions are obtained.

[12] We use the PDE solver in MATLAB. We set 44000 mesh points to $\ln \tilde{a}$ between the lower bound $\ln \tilde{a}_{\mathrm{LB}}$ (see footnote 13 for details) and 110 and 500 mesh points to time t between 1970 and 2020.

[13] $\tilde{a}_{\mathrm{LB}}$ is set to be higher than $\tilde{h}$ at the pre- and post-1970 steady states.

6.4 Aggregate transition after the tax cut

Before analyzing the evolution of the income and wealth distributions, we first look at the aggregate dynamics of the transition economy. An important implication of the model is that a tax cut does not significantly affect capital accumulation or the capital–output ratio of the economy. This result comes from the property that investment in capital is financed by retained earnings. Then, the tax change does not affect the return on stocks $((1-\tau^e)d_{i,t}dt + dq_{i,t})/q_{i,t}$, because $q_{i,t}$ in the denominator of the equation changes to offset the effect of tax change $(1-\tau^e)$ in the numerator.

Figures 2(a) and 2(b) plot the computed transitions of detrended per-capita output and the capital–output ratio of the model economy under Cases A and B. In the plot, per-capita output is normalized to 1, before the tax cut and the transition for 50 years after the tax cut is shown. Although the variables increase after the tax cut, we note that the magnitudes are quantitatively negligible: for example, detrended per-capita output only increases by 1 percentage point 20 years after the tax cut. Thus, the computed transition confirms our prediction that a tax cut has almost no quantitative impacts on per-capita output or the capital–output ratio.

Figures 2(c) and 2(d) plot the transition of prices. We observe that the risk-free rate and detrended wage rate are almost unchanged after the tax cut. This is another consequence of the fact that the tax cut has negligible effects on capital accumulation in our model.

The prediction of the model that the tax change has negligible effects on capital accumulation is in sharp contrast to previous models of the income distribution such as Nirei and Aoki (2016), Toda (2014), and Kim (2013). We note that the prediction is consistent with the facts in the U.S. that the capital–output ratio has not changed significantly over the post-World War II years; nor has the level of per-capita GDP increased above the trend line recently. In contrast to the capital–output ratio, the value of detrended financial asset $\tilde{Q}_t$ jumps after the tax cut, which is caused by the increase in after-tax dividends. This mechanism is the same as the model in McGrattan and Prescott (2005) and consistent with their interpretation on the rise in the equity value of U.S. firms since the 1970s.

Figures 2(e) and 2(f) show the wealth accumulation rates of entrepreneurs and workers before and after the tax cut. Note that the gap in wealth accumulation rates between entrepreneurs and workers significantly and permanently widens after the tax cut. This widened gap is caused by the increased difference in risky and risk-free returns. The key mechanism of rising inequality in our model is this increased difference between the risky and risk-free rates after the tax cut.

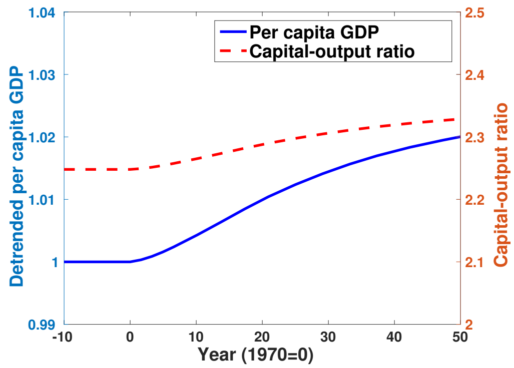
6.5 Pareto exponent and the top 1% income share

Figures 3(a) and 3(c) show the Pareto exponent of the income distribution in the calibrated model for Cases A and B, along with the historical U.S. Pareto exponent. While we referred to the "Pareto exponent" in the model analysis as an asymptotic exponent in the right tail distribution, we need to fix the tail range when we estimate the exponent with finite data. We calculate the exponent from the slope of the complementary cumulative distribution of household wealth $\Pr(\tilde{a}_{i,t} > \tilde{a})$ between the top 0.1% and top 1% in the calibrated model as well as in the U.S. data. We hereafter refer to this as an "empirical Pareto exponent." For the model prediction, we plot the stationary empirical Pareto exponents for the pre- and post-1970 periods and the transition path of the empirical Pareto exponent between them.

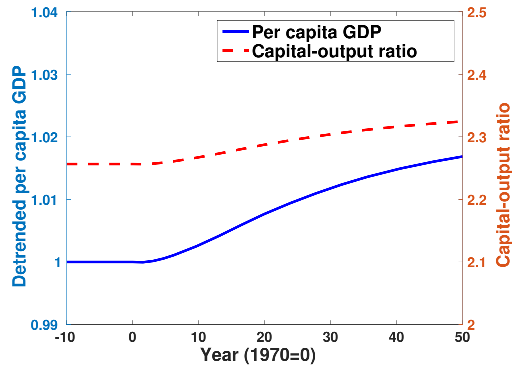
We find that in both Cases A and B, the model traces data for the empirical Pareto exponent well. Although ι is set to match the level of the empirical Pareto exponent at the initial steady state, it is nontrivial that the model matches both the level and the changes in the empirical Pareto exponent afterward. For example, suppose that we need to set a low (high) ι to match the empirical Pareto exponent at the initial steady state. Then, the changes in the empirical Pareto exponent during the transition become slower (faster) than the data because the volatility of each entrepreneur's wealth decreases (increases).

Figures 3(b) and 3(d) plot the top 1% income shares for Cases A and B.

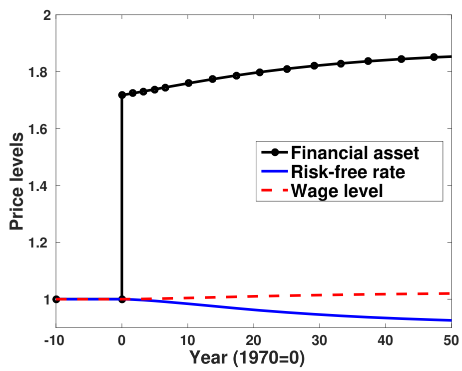
Figure 2: Response of the aggregate variables after the tax cut



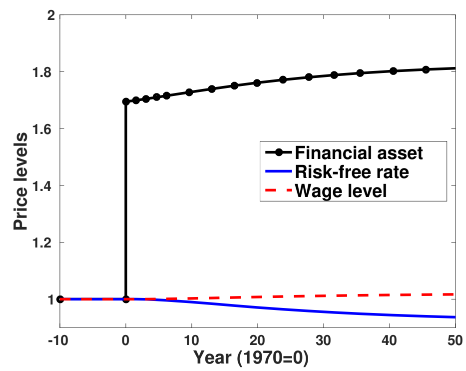
(a) Output and capital: Case A



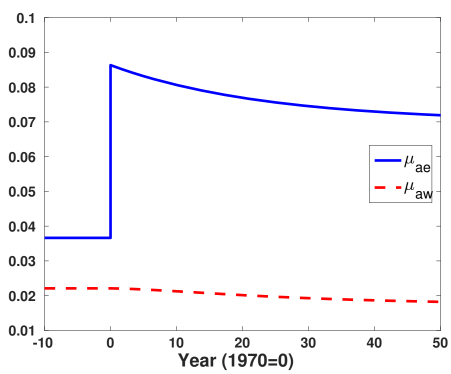
(b) Output and capital: Case B



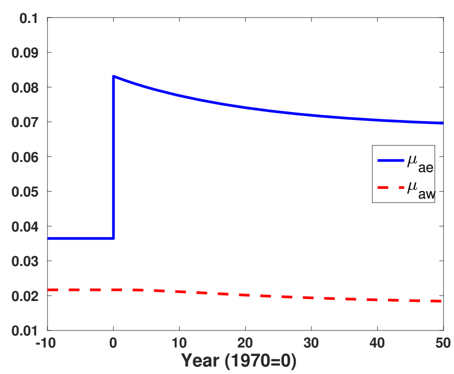
(c) Price levels: Case A



(d) Price levels: Case B



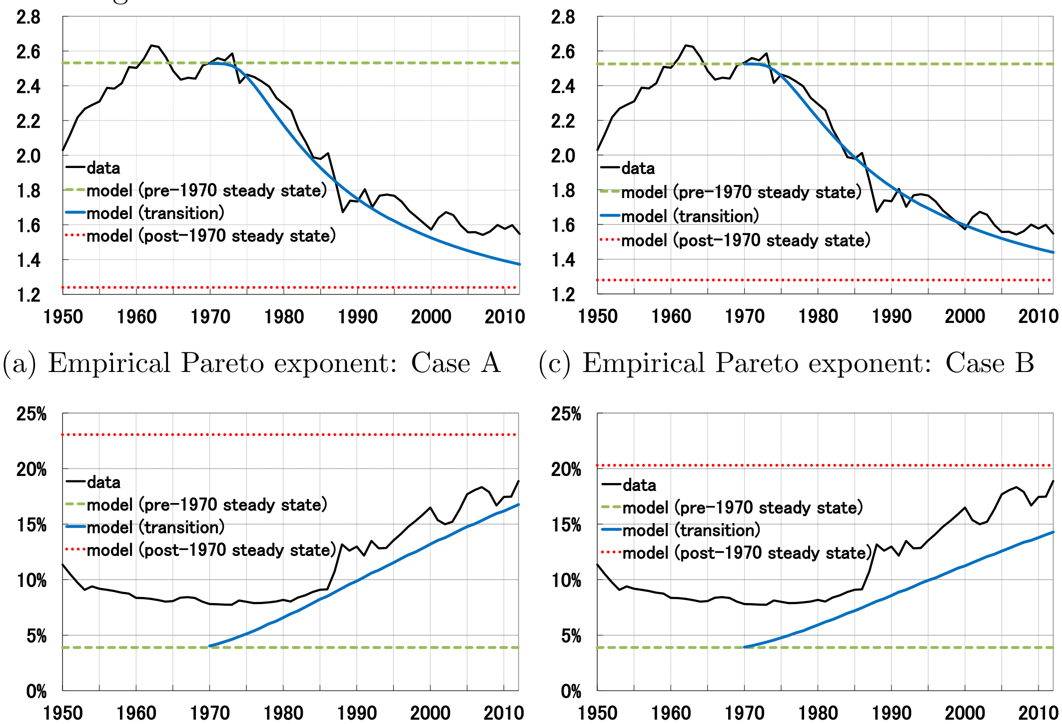
(e) Wealth growth rates: Case A



(f) Wealth growth rates: Case B

Notes: Figures (a) and (b) plot detrended per-capita output (the first axis) and the capital–output ratio (the second axis). Figures (c) and (d) plot the detrended aggregate financial wealth value, risk-free rate, and detrended wage rate. Figures (e) and (f) plot the wealth growth rates of an entrepreneur and a worker, μ_{ae} and μ_{aw}. The horizontal axis shows the years after the tax cut. Detrended per-capita output and prices before the tax cut are normalized to 1.

Figure 3: The evolution of the income and wealth distributions



(a) Empirical Pareto exponent: Case A (c) Empirical Pareto exponent: Case B

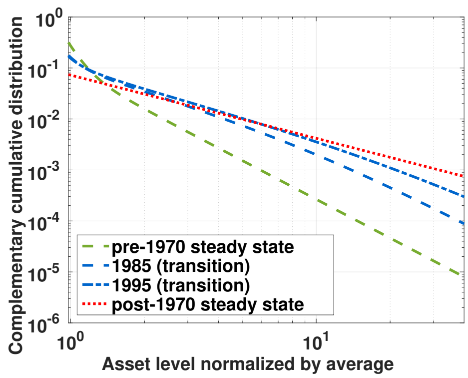
(b) Top 1% income share: Case A (d) Top 1% income share: Case B

Note: Data are taken from Alvaredo et al. (2013). The empirical Pareto exponent is calculated in the range between the top 0.1% and top 1%.

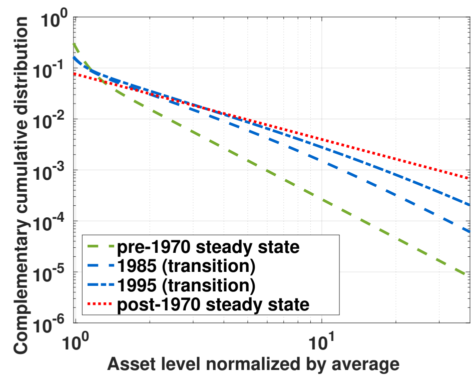
Under these specifications, the model captures the trend in the top 1% share of income after 1970, although the model's prediction of the pre-1970 steady state is lower in level than the data reveal. Other factors, such as rewards for executives' talents as argued by Gabaix and Landier (2008) and bargaining and rent extraction by executives as emphasized by Piketty et al. (2011), may account for this gap. Note that the top 1% income share in Case B increases somewhat more slowly than that in Case A. This is because the firm's volatility becomes higher in Case B. This makes $x_{e,t}$ lower by (3), which results in the lower volatility of the entrepreneur's wealth.

In Figure 4, we plot the complementary cumulative distributions of the household's detrended wealth $\Pr(\tilde{a}_{i,t} > \tilde{a})$ at the pre- and post-1970 steady states and the transition paths. We find that the wealth distribution converges

Figure 4: Household's wealth distributions



(a) Case A



(b) Case B

Notes: The figures plot the complementary cumulative distributions of a household's wealth $\Pr(\tilde{a}_{i,t} > \tilde{a})$ normalized by the average wealth each year. For example, "1985 (transition)" indicates the wealth distribution in 1985 normalized by the 1985 average wealth. The figure on the left presents the distributions for Case A, whereas the figure on the right presents them for Case B.

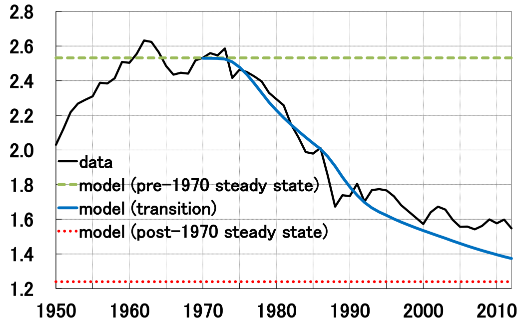
to the new distribution from the low wealth region first, whereas the convergence is slow in the high wealth region. We also find that the convergence is somewhat faster in Case A than in Case B, similar to the computed transition of the empirical Pareto exponent.

Gabaix et al. (2015) show that standard models with random wealth growth cannot generate the rise in inequality as fast as that observed in the data, unless the wealth growth process includes a high growth type. Our model is consistent with their view. In our model, the heterogeneity in the mean wealth growth rate between entrepreneurs and workers and the probability of an entrepreneur becoming a worker p_f generate the rapid decline in the empirical Pareto exponent.

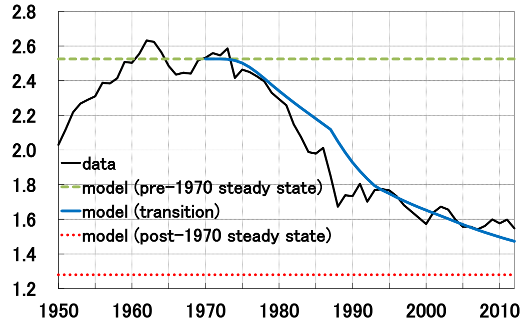
6.6 Gradual change in tax rates

In the benchmark cases, we assume a sudden tax change in 1970. This might seem a too convenient assumption, because the actual tax changes were more gradual. To check the robustness of our analysis above, we compute the

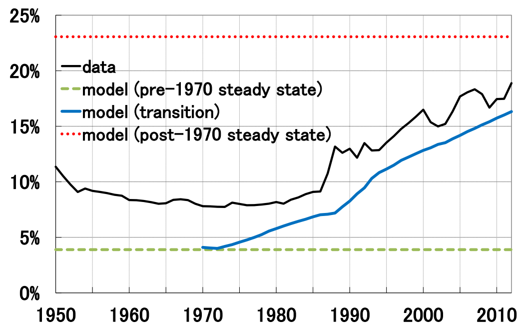
Figure 5: The evolution of the income distribution: changing taxes



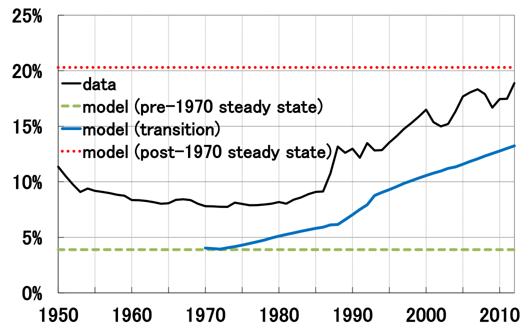
(a) Empirical Pareto exponent: Case A



(c) Empirical Pareto exponent: Case B



(b) Top 1% income share: Case A



(d) Top 1% income share: Case B

Note: Data are taken from Alvaredo et al. (2013).

transition path under the exact time series of the tax rates. Owing to the log utility, in which a household's decision rule is myopic, the model economy can be computed in the same way as before. The results are shown in Figure 5 and are similar to the benchmark cases.

6.7 Incentive pay for executives

In reality, executives obtain incentive pay such as stock options, whose value moves in line with the firm's performance. In our model, this is represented by entrepreneurs holding risky stocks of their firms. Here, we discuss whether our formulation is realistic.

Our formulation of executive pay is similar to those of Edmans et al. (2009) and Edmans et al. (2012). These studies theoretically derive that under the optimal incentive scheme in a moral hazard problem, as in our model, a proportion of the executive's wealth, denoted by $x_{e,t}$ in our model, is invested in his firm's stocks. Moreover, Edmans et al. (2009) find evidence that an empirical counterpart of $x_{e,t}$, (23), is cross-sectionally independent of firm size. This property is satisfied in theirs and in our models.

We can also check whether the value of $x_{e,t}$ in our model is quantitatively consistent with the empirical estimate. An empirical counterpart of $x_{e,t}$ is computed by

$$\frac{x\% \text{ increase in the executive's wealth}}{1\% \text{ increase in firm rate of return}}, \tag{23}$$

because from (1),

$$x_{e,t} = \frac{d(a_{e,t})/a_{e,t}}{\mu_{q,t}dt + \sigma_{q,t}dB_{e,t}}.$$

Clementi and Cooley (2009) estimate (23) from CEO compensation data in the U.S. for 1993–2008 provided by the EXECUCOMP database. The empirical value of (23) ranges from 1.14 to 1.24 (see Table 3). In our calibration, $x_{e,t}$ in the post-1970 steady state is 1.53 for Case A and 0.99 for Case B. Therefore, the empirical value of $x_{e,t}$ is between those of Case A and Case B.

Related to $x_{e,t}$, Edmans et al. (2009) define and provide empirical estimates of a wealth–performance sensitivity measure, which they refer to as B^I:

$$B^I \equiv \frac{x\% \text{ increase in the CEO's wealth}}{1\% \text{ increase in firm rate of return}} \times \frac{\text{the CEO's wealth}}{\text{the CEO's pay}}. \tag{24}$$

B^I is a slight modification of (23). Edmans et al. (2009) report that the empirical B^I measured in 1999 is 9.04 (Table 3). In our model, B^I in the post-1970 steady state is 12.71 for Case A and 8.45 for Case B.[14] The empirical values of B^I can also be calculated from the long-run data on CEO pay by

[14]The model's counterpart of B^I in (24) is calculated from

$$\frac{d(a_{e,t})/a_{e,t}}{\mu_{q,t}dt + \sigma_{q,t}dB_{e,t}} \frac{a_{e,t}}{\nu a_{i,t} + \mu_{q,t}x_{i,t}a_{i,t} + r_t^f(1 - x_{i,t})a_{i,t}} = \frac{x_{e,t}}{\mu_{a,t} + \beta + \nu}.$$

Table 3: Incentive elasticities

	Case A	Case B		Data
$x_{e,t}$ (Pre-1970)	0.71	0.45		
$x_{e,t}$ (Post-1970)	1.53	0.99	$x_{e,t}$ (1993–2008)	1.14–1.24
B^I (Pre-1970)	8.16	5.16	B^I (1961–2005)	5.04
B^I (Post-1970)	12.71	8.45	B^I (1999)	9.04

Notes: For the definitions, see (23)–(24). The values for $x_{e,t}$ (1993–2008) are the estimates by Clementi and Cooley (2009). B^I (1961–2005) is computed from Figures 5 and 6 in Frydman and Saks (2010). For B^I (1999), the estimate is provided by Edmans et al. (2009).

Frydman and Saks (2010).[15] Because the long-run values of B^I calculated from the data of Frydman and Saks (2010) are stable for 1961–2005, we only show the mean in Table 3. The empirical value, 5.06, is close to our model's pre-1970 steady state values of B^I, 8.16 for Case A and 5.16 for Case B.

There are also differences between our model and the models of Edmans et al. (2009) and Edmans et al. (2012). In their models, a single structural parameter, the disutility of effort, affects the proportion of an entrepreneur's wealth invested in his firm's stocks. By contrast, several factors affect this proportion in our model; for example, an increase in the volatility of firm value decreases the proportion of an entrepreneur's total wealth invested in risky stocks $x_{e,t}$ (see (3)). This prediction is consistent with the evidence surveyed in Frydman and Jenter (2010, Section 2.3).

6.8 Welfare analysis

To investigate how the tax change affected the welfare of households, we calculate the utility level of an entrepreneur and an innate worker (that is, a worker from the beginning of his life) in the pre- and post-1970 steady states. Table 4 shows the detrended initial utility level, defined by $V^i(\tilde{h}, \mathbf{S})$, under Cases A and B (for the details of the derivations, see Appendix E).

In both cases, the utility level of an entrepreneur becomes higher in the

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

Table 4: Welfare analysis

Case A	$V^e(\tilde{h}, \mathbf{S})$	$V^w(\tilde{h}, \mathbf{S})$
Pre-1970	8.66	7.21
Post-1970	12.01	6.74

Case B	$V^e(\tilde{h}, \mathbf{S})$	$V^w(\tilde{h}, \mathbf{S})$
Pre-1970	7.73	6.25
Post-1970	10.96	5.96

Notes: The table calculates the detrended initial utility level of an entrepreneur and an innate worker at the pre- and post-1970 steady states. The detrended initial utility level is defined by $V^i(\tilde{h}, \mathbf{S})$. The table on the left presents these calculations for Case A, whereas the table on the right presents them for Case B.

post-1970 steady state, whereas that of an innate worker becomes lower. These results are consistent with the view that the rich have benefited from the tax change at the expense of the poor.

7 Conclusion

We proposed a model of wealth and income inequalities that explains both Zipf's law of firms and Pareto's law of incomes from the idiosyncratic productivity shocks of firms. Empirical studies show that the Pareto exponent of income varies over time, whereas Zipf's law of firm size is 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 and are then used for the numerical analysis.

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

By calibrating the model, we analyzed the extent to which changes in tax rates account for the recent evolution of top incomes in the U.S. We find that the model matches the decline in the Pareto exponent of the income distribution and the trend in the top 1% income share.

There remain some discrepancies between the model and data. First, the model's prediction of the top 1% income share is somewhat lower than that seen in the data. Second, we did not attempt to account for top wealth shares. Note that in our model, the tail exponent of the wealth distribution is identical to that of the income distribution. This may seem counterfactual at first. However, it is important to note that the total wealth of a household in our model includes both financial and human assets. A quantitative analysis of the wealth distribution needs to be left for future research, which would appropriately take into account human wealth in the estimation.

References

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez (2013) "The World Top Incomes Database," http://topincomes.g-mond.parisschoolofeconomics.eu/. Data are extracted at 25/11/2013.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez (2011) "Top Incomes in the Long Run of History," *Journal of Economic Literature*, Vol. 49, No. 1, pp. 3–71.

Bakija, Jon, Adam Cole, and Bradley T. Heim (2012) "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data," mimeo.

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.

——— (2015) "The Wealth Distribution in Bewley Economies with Capital Income Risk," *Journal of Economic Theory*, Vol. 159, Part A, pp. 489–515.

——— (2016) "The Distribution of Wealth in the Blanchard-Yaari Model," *Macroeconomic Dynamics*, Vol. 20, No. 2, pp. 466–481.

Bitler, Marianne P., Tobias J. Moskowitz, and Annette Vissing-Jorgensen (2005) “Testing Agency Theory with Entrepreneur Effort and Wealth,” *Journal of Finance*, Vol. 60, No. 2, pp. 539–576.

Blanchard, Olivier J. (1985) “Debt, Deficits, and Finite Horizons,” *Journal of Political Economy*, Vol. 93, No. 2, pp. 223–247.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini (2007) “Down or Out: Assessing the Welfare Costs of Household Investment Mistakes,” *Journal of Political Economy*, Vol. 115, No. 5, pp. 707–747.

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

Champernowne, D.G. (1953) “A Model of Income Distribution,” *Economic Journal*, Vol. 63, No. 250, pp. 318–351.

Clementi, Gian Luca and Thomas F. Cooley (2009) “Executive Compensation: Facts,” NBER Working Papers 15426, National Bureau of Economic Research, Inc.

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*, Vol. 21, National Bureau of Economic Research, Inc., pp. 107–180.

Dutta, Jayasri and Philippe Michel (1998) “The Distribution of Wealth with Imperfect Altruism,” *Journal of Economic Theory*, Vol. 82, No. 2, pp. 379–404.

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.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov (2012) “Dynamic CEO Compensation,” *Journal of Finance*, Vol. 67, No. 5, pp. 1603–1647.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2016) "Heterogeneity and Persistence in Returns to Wealth," mimeo.

Flavin, Marjorie and Takashi Yamashita (2002) "Owner-Occupied Housing and the Composition of the Household Portfolio," *American Economic Review*, Vol. 92, No. 1, pp. 345–362.

Frydman, Carola and Dirk Jenter (2010) "CEO Compensation," *Annual Review of Financial Economics*, Vol. 2, No. 1, pp. 75–102.

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?" *Quarterly Journal of Economics*, Vol. 123, No. 1, pp. 49–100.

Gabaix, Xavier, Benjamin Moll, Jean-Michel Lasry, and Pierre-Louis Lions (2015) "The Dynamics of Inequality," mimeo.

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. (2015) "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality," *Journal of Economic Perspectives*, Vol. 29, No. 1, pp. 29–46.

Jones, Charles I. and Jihee Kim (2014) "A Schumpeterian Model of Top Income Inequality," mimeo.

Kaplan, Steven N. and Joshua Rauh (2013) "It's the Market: The Broad-Based Rise in the Return to Top Talent," *Journal of Economic Perspectives*, Vol. 27, No. 3, pp. 35–56.

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.

Klette, Tor Jakob and Samuel Kortum (2004) "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, Vol. 112, No. 5, pp. 986–1018.

Luttmer, Erzo G.J. (2007) "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, Vol. 122, No. 3, pp. 1103–1144.

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

——— (2011) "On the Mechanics of Firm Growth," *Review of Economic Studies*, Vol. 78, No. 3, pp. 1042–1068.

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.

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.

Moskowitz, Tobias J. and Annette Vissing-Jorgensen (2002) "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *American Economic Review*, Vol. 92, No. 4, pp. 745–778.

Nirei, Makoto and Shuhei Aoki (2016) "Pareto Distribution of Income in Neoclassical Growth Models," *Review of Economic Dynamics*, Vol. 20, pp. 25–42.

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.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva (2011) "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," NBER Working Papers 17616, National Bureau of Economic Research, Inc.

Reed, William J. (2001) "The Pareto, Zipf and Other Power Laws," *Economics Letters*, Vol. 74, No. 1, pp. 15–19.

Rossi-Hansberg, Esteban and Mark L.J. Wright (2007) "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, Vol. 97, No. 5, pp. 1639–1666.

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*, Vol. 5, National Bureau of Economic Research, Inc., pp. 25–54.

Toda, Alexis Akira (2014) "Incomplete Market Dynamics and Cross-sectional Distributions," *Journal of Economic Theory*, Vol. 154, pp. 310–348.

Wold, H.O.A. and P. Whittle (1957) "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, Vol. 25, No. 4, pp. 591–595.

A Derivations of the household problem

This appendix shows the derivations of the household problem in Section 2.1. As shown in Section 4.1, the aggregate dynamics of the model are described by $\mathbf{S}_t$, whose evolution can be written as

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

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

$$\begin{aligned} dV^i(a_{i,t}, \mathbf{S}_t) =& \frac{\partial V_t^i}{\partial a_{i,t}} da_{i,t} + \frac{1}{2}\frac{\partial^2 V_t^i}{\partial a_{i,t}^2}(da_{i,t})^2 + \frac{\partial V_t^i}{\partial \mathbf{S}_t} \cdot d\mathbf{S}_t \\ &+ \left(V^\ell(a_{i,t}, \mathbf{S}_t) - V^i(a_{i,t}, \mathbf{S}_t)\right) dJ_{i,t}, \end{aligned}$$

where $J_{i,t}$ is the Poisson jump process that describes the probability of an entrepreneur leaving his firm and becoming a worker.

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

Thus,

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

where $\mu'_{\mathbf{S}}(\mathbf{S}_t)$ is the transposed vector of $\mu_{\mathbf{S}}(\mathbf{S}_t)$. By substituting into (2), we obtain a Hamilton–Jacobi–Bellman equation as follows:

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

$$
\begin{aligned}
&= \max_{c_{i,t}, x_{i,t}} \ln c_{i,t} - (\beta+\nu)V_t^i + \frac{\sigma_{q,t}^2}{2} x_{i,t}^2 a_{i,t}^2 \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\
&\quad + ((\nu+\mu_{q,t})x_{i,t}a_{i,t} + (\nu + r_t^f)(1-x_{i,t})a_{i,t} - c_{i,t}) \frac{\partial V_t^i}{\partial a_{i,t}} \\
&\quad + \mu_{\mathbf{S}}'(\mathbf{S}_t) \cdot \frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f \left(V_t^\ell - V_t^i \right). \qquad (25)
\end{aligned}
$$

The FOCs with respect to $c_{i,t}$ and $x_{i,t}$ are summarized as follows:

$$
c_{i,t}^{-1} = \frac{\partial V_t^i}{\partial a_{i,t}}, \qquad (26)
$$

$$
x_{i,t} = \begin{cases} -\frac{\partial V_t^i / \partial a_{i,t}}{\left(\partial^2 V_t^i / \partial a_{i,t}^2\right) a_{i,t}} \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}, & \text{if } i = e, \\ 0, & \text{otherwise.} \end{cases} \qquad (27)
$$

Furthermore, (25) has to satisfy the transversality condition (5).

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

$$
\begin{aligned}
V_t^i &= B_t^i \ln a_{i,t} + H^i(\mathbf{S}_t), \qquad (28) \\
c_{i,t} &= v_{i,t} a_{i,t}, \\
q_{i,t} s_{i,t} &= x_{i,t} a_{i,t}, \\
b_{i,t} &= (1 - x_{i,t}) a_{i,t} - h_t.
\end{aligned}
$$

We obtain this solution by guess–and–verify. The FOC (26) becomes

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

Condition (27) is rewritten as

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

Substituting these results into (25), we find that

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

B Derivation of the firm problem

B.1 FOCs of the firm problem

This appendix shows the derivations of the firm problem described in Section 2.2.2. The value of a product line $q_{n,e,t} = q(k_{n,e,t}, z_{n,e,t}, \mathbf{S}_t)$ is a function of $k_{n,e,t}$, $z_{n,e,t}$, and the aggregate dynamics $\mathbf{S}_t$ (see Appendix A). By applying Ito's formula to $q_{n,e,t}$, we obtain

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

From the above equation, the FOCs of (7) for $\ell_{n,e,t}$ and $dk_{n,e,t}$ are

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

By the envelope theorem,

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

By rearranging the equation, we obtain

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

B.2 Firm-side variables

This appendix briefly explains the derivations of the firm-side variables described in Section 3.1 and used in Section 4.1. Our goal here is to rewrite the firm-side variables as the functions of MPK_t and exogenous variables. The basic strategy is as follows:

1. From FOCs (8) and (9), we rewrite $k_{n,e,t}$ and $\ell_{n,e,t}$ as the functions of MPK_t, w_t, Y_t, and exogenous variables.

   From (9),

$$w_t = (1-\alpha)(\phi-1)/\phi \; Y_t^{1-(\phi-1)/\phi} z_{n,e,t}^{(\phi-1)/\phi} k_{n,e,t}^{\alpha(\phi-1)/\phi} \ell_{n,e,t}^{(1-\alpha)(\phi-1)/\phi-1}.$$

   Rewriting this,

$$\ell_{n,e,t} = \left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t} Y_t^{1-(\phi-1)/\phi} z_{n,e,t}^{(\phi-1)/\phi} k_{n,e,t}^{\alpha(\phi-1)/\phi} \right)^{\frac{1}{1-(1-\alpha)(\phi-1)/\phi}}. \tag{29}$$

   On the other hand, from (8),

$$\mathsf{MPK}_t = \alpha(\phi-1)/\phi \; Y_t^{1-(\phi-1)/\phi} z_{n,e,t}^{(\phi-1)/\phi} k_{n,e,t}^{\alpha(\phi-1)/\phi-1} \ell_{n,e,t}^{(1-\alpha)(\phi-1)/\phi}. \tag{30}$$

   By substituting (29) into (30) and rearranging,

$$\begin{aligned} k_{n,e,t}^{\frac{\alpha(\phi-1)/\phi}{1-(1-\alpha)(\phi-1)/\phi}} = & \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} Y_t^{1-(\phi-1)/\phi} \right)^{\frac{\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} \\ & \times \left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t} Y_t^{1-(\phi-1)/\phi} \right)^{\frac{\alpha(\phi-1)/\phi(1-\alpha)(\phi-1)/\phi}{(1-(\phi-1)/\phi)(1-(1-\alpha)(\phi-1)/\phi)}} \\ & z_{n,e,t}^{\eta \frac{\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}}, \end{aligned} \tag{31}$$

   where $\eta \equiv \frac{(\phi-1)/\phi}{1-(1-\alpha)(\phi-1)/\phi}$. Substituting (31) into (29),

$$\begin{aligned} \ell_{n,e,t} = & \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} Y_t^{1-(\phi-1)/\phi} \right)^{\frac{\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} \\ & \times \left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t} Y_t^{1-(\phi-1)/\phi} \right)^{\frac{1-\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} z_{n,e,t}^{\phi-1} \end{aligned} \tag{32}$$

2. By using the labor market condition (11), we remove w_t from these equations.

By substituting (32) into the labor market condition (11) and rearranging,

$$\left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} \times \left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{1-\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} = \frac{1}{\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}}, \tag{33}$$

or,

$$\left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{(1-\alpha)(\phi-1)/\phi}{1-(\phi-1)/\phi}} = \left\{\left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{-\alpha(\phi-1)/\phi}{1-(\phi-1)/\phi}} \frac{1}{\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}}\right\}^{\frac{(1-\alpha)(\phi-1)/\phi}{1-\alpha(\phi-1)/\phi}}. \tag{34}$$

Here, $\overline{\mathsf{E}}$ is the operator of the cross-sectional average of all firms. Then, substituting (33) into (32),

$$\ell_{n,e,t} = \left(\frac{z_{n,e,t}^{\phi-1}}{\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}}\right). \tag{35}$$

Rewriting (31),

$$k_{n,e,t} = \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{1-(1-\alpha)(\phi-1)/\phi}{1-(\phi-1)/\phi}} \times \left(\frac{(1-\alpha)(\phi-1)/\phi}{w_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{(1-\alpha)(\phi-1)/\phi}{1-(\phi-1)/\phi}} z_{n,e,t}^{\phi-1}. \tag{36}$$

Substituting (34) into (36),

$$k_{n,e,t} = \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t}Y_t^{1-(\phi-1)/\phi}\right)^{\frac{1}{1-\alpha(\phi-1)/\phi}} \times \left(\frac{z_{n,e,t}^{\phi-1}}{\overline{\mathsf{E}}\left\{z_{n,e,t}^{\phi-1}\right\}^{\frac{(1-\alpha)(\phi-1)/\phi}{1-\alpha(\phi-1)/\phi}}}\right). \tag{37}$$

3. By using the results, the production function, and the aggregate good

function (10), we remove Y_t from the equations of the firm-side variables. Substituting (35) and (37) into $y_{n,e,t} = z_{n,e,t} k_{n,e,t}^{\alpha} \ell_{n,e,t}^{1-\alpha}$ and rearranging,

$$y_{n,e,t} = \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} Y_t^{1-(\phi-1)/\phi} \right)^{\frac{\alpha}{1-\alpha(\phi-1)/\phi}} \times \left(\frac{z_{n,e,t}^{\frac{1}{1-(\phi-1)/\phi}}}{\overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{(1-\alpha)}{1-\alpha(\phi-1)/\phi}}} \right).$$

Substituting this equation into $Y_t = \left(\int_0^E \int_0^{\bar{n}(e)} y_{n,e,t}^{(\phi-1)/\phi} \; dnde \right)^{\frac{1}{(\phi-1)/\phi}}$,

$$Y_t^{1-(\phi-1)/\phi} = \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \right)^{\frac{\alpha(1-(\phi-1)/\phi)}{1-\alpha}} \times \overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{(1-(\phi-1)/\phi)\left[\frac{1-\alpha(\phi-1)/\phi}{(1-\alpha)(\phi-1)/\phi} - 1 \right]}. \tag{38}$$

Substituting (38) into (37),

$$\begin{aligned} k_{n,e,t} &= \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}} \left(\frac{z_{n,e,t}^{\phi-1}}{\overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}} \right) \\ &= \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}} \ell_{n,e,t}. \end{aligned} \tag{39}$$

Substituting (35) and (39) into (38),

$$p_{n,e,t} y_{n,e,t} = Y_t^{1-(\phi-1)/\phi} y_{n,e,t}^{(\phi-1)/\phi} \tag{40}$$

$$\begin{aligned} &= \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}} \left(\frac{z_{n,e,t}^{\phi-1}}{\overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}} \right) \\ &= \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{1}{\phi-1}\frac{1}{1-\alpha}} \ell_{n,e,t}. \end{aligned} \tag{41}$$

Rewriting (35),

$$\ell_{n,e,t} = \overline{\ell}_t \; z_{n,e,t}^{\phi-1}, \text{ where } \overline{\ell}_t \equiv \left(\frac{1}{\overline{\mathsf{E}} \left\{ z_{n,e,t}^{\phi-1} \right\}} \right).$$

Rewriting (41),

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

Rewriting (39),

$$k_{n,e,t} = \overline{k}_t \overline{\ell}_t z_{n,e,t}^{\phi-1}, \text{ where } \overline{k}_t \equiv \left(\frac{\alpha(\phi-1)/\phi}{\mathsf{MPK}_t} \overline{\mathsf{E}} \left\{ z_{n,e,t}^{\phi-1} \right\}^{\frac{1}{\phi-1}} \right)^{\frac{1}{1-\alpha}}. \tag{42}$$

We obtain $\ell_{n,e,t}$, $p_{n,e,t} y_{n,e,t}$, and $k_{n,e,t}$ (12)–(15). To compute $dd_{n,e,t}$, we first need to compute $dk_{n,e,t}$.

4. We compute $dk_{n,e,t}$ as follows. From (42),

$$\begin{aligned} dk_{n,e,t} &= d(\overline{k}_t \overline{\ell}_t z_{n,e,t}^{\phi-1}) \\ &= \frac{d\overline{k}_t \overline{\ell}_t}{dt} z_{n,e,t}^{\phi-1} dt + \overline{k}_t \overline{\ell}_t \; dz_{n,e,t}^{\phi-1}. \end{aligned}$$

Note that

$$dz_{n,e,t}^{\phi-1} = \left((\phi-1) \left(\mu_z - \frac{\sigma_z^2}{2} \right) + \frac{(\phi-1)^2 \sigma_z^2}{2} \right) z_{n,e,t}^{\phi-1} dt + (\phi-1) \sigma_z z_{n,e,t}^{\phi-1} dB_{e,t}.$$

Then,

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

Here,

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

5. We obtain $dd_{n,e,t}$ (13) by substituting these results into the following relationship:

$$\begin{aligned} dd_{n,e,t} =& (p_{n,e,t}y_{n,e,t} - w_t\ell_{n,e,t} - \delta k_{n,e,t})dt - dk_{n,e,t} \\ =& (1-(1-\alpha)(\phi-1)/\phi)p_{n,e,t}y_{n,e,t}dt - \delta k_{n,e,t}dt - dk_{n,e,t}. \end{aligned}$$

Then, $dd_{n,e,t}$ is rewritten as follows:

$$dd_{n,e,t} = \overline{d}_t\overline{\ell}_t z_{n,e,t}^{\phi-1}dt - \overline{k}_t\overline{\ell}_t z_{n,e,t}^{\phi-1}(\phi-1)\sigma_z dB_{e,t},$$
$$\text{where } \overline{d}_t \equiv (1-(1-\alpha)(\phi-1)/\phi)\overline{py}_t - (\delta + \mu_{k,t})\,\overline{k}_t.$$

We obtain $q_{n,e,t}$ (16) through the following steps. Here, we allow taxes to change for the numerical analysis and add time subscript t to take into account tax changes.

6. By multiplying (6) by $e^{-\int_t^u r_s^f ds}$ and integrating,[16] we obtain

$$\begin{aligned} q_{n,e,t} =& \int_t^\infty \mathsf{E}_t\left[(1-\tau_u^f)(1-\iota)d_{n,e,u}e^{-\int_t^u r_s^f ds}\right]du \\ =& \int_t^\infty (1-\tau_u^f)(1-\iota)e^{-\int_t^u r_s^f ds}\,\mathsf{E}_t\left[d_{n,e,u}\right]du. \end{aligned}$$

[16]The Ito process version of integration by parts

$$\int_t^T X_s dY_s = X_tY_t - X_tY_t - \int_t^T Y_s dX_s - \int_t^T dX_s dY_s$$

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

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

7. $\mathsf{E}_t[d_{n,e,u}]$ in the above equation is further computed as follows:

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

By using this equation, we obtain (16):

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

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

$$q_{e,t} = \frac{(1-\tau^f)(1-\iota)d_{e,t}}{r^f - g}.$$

From the above results, we show the following properties that are used in Section 4.1.

1. The aggregate detrended dividend $\tilde{D}_t$ is obtained by aggregating $d_{n,e,t}dt$ (13) and detrending by e^{gt},

   $$\tilde{D}_t = (1-(1-\alpha)(\phi-1)/\phi)\tilde{Y}_t - (\delta+g)\tilde{K}_t - \frac{d\tilde{K}_t}{dt}.$$

   Here, we use the property

   $$\frac{1}{1-\alpha} \frac{dr_t^f/dt}{\mathsf{MPK}_t} = \frac{d\tilde{K}_t}{dt} / \tilde{K}_t.$$

2. Using the above relations, the return on a risky stock of firm e, $\left\{(1-\tau_t^e)d_{e,t}dt + dq_{e,t}\right\}/q_{e,t}$, is rewritten as the function of aggregate variables and exogenous shocks.

First note that for each product line,

$$
\begin{aligned}
dq_{n,e,t} =& q_{n,e,t}\frac{d\ln(\overline{d}_t\overline{\ell}_t)}{dt}dt + q_{n,e,t}\frac{dz^{\phi-1}_{n,e,t}}{z^{\phi-1}_{n,e,t}} \\
&+ q_{n,e,t}\left\{-\frac{(1-\tau^f_t)(1-\iota)}{\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du}\right. \\
&+ \left.\frac{(r^f_t-\mu_{d,t})\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u r^f_s-\mu_{d,s})ds\right\}du}{\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du}\right\}dt \\
=& \left\{-(1-\tau^f_t)(1-\iota)d_{n,e,t} + r^f_t q_{n,e,t}\right\}dt + q_{n,e,t}(\phi-1)\sigma_z dB_{e,t}.
\end{aligned}
$$

Integrating $d_{n,e,t}$ (13), $q_{n,e,t}$ (16), and the above relation, over the product lines of firm e, and substituting the results into the return on a risky stock $\{(1-\tau^e_t)d_{e,t}dt + dq_{e,t}\}/q_{e,t} = \mu_{q,t}dt + \sigma_{q,t}dB_{e,t}$, we obtain

$$
\begin{aligned}
\mu_{q,t} =& \left\{r^f_t + \frac{(1-\tau^e_t)-(1-\tau^f_t)(1-\iota)}{\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du}\right\}, \\
\sigma_{q,t} =& (\phi-1)\sigma_z \\
&\times \left\{1-\frac{\tilde{K}_t}{\tilde{D}_t}\frac{(1-\tau^e_t)}{\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du}\right\}.
\end{aligned}
$$

In order to compute the return on risky stocks from aggregate variables and exogenoush shocks, we need to know the value of $\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du$. We calculate the value as follows. Integrating (16), we obtain

$$
\int_t^\infty (1-\tau^f_u)(1-\iota)\exp\left\{-\int_t^u (r^f_s-\mu_{d,s})ds\right\}du = \frac{Q_t}{D_t}.
$$

C Details of the aggregate dynamics

C.1 Shooting algorithm

The initial values of aggregate total and human capital, $\tilde{A}_{1970}$ and $\tilde{H}_{1970}$, are determined by using the shooting algorithm through the following steps:

1. Set $\widetilde{HA}_{1970} \equiv \tilde{H}_{1970}/\tilde{A}_{1970}$. Further, set the upper and lower bounds of $\widetilde{HA}_{1970}$, $\widetilde{HA}_H$, and $\widetilde{HA}_L$.

   (a) Set $\tilde{A}_{1970}$. In addition, set the upper and lower bounds of $\tilde{A}_t$, $\tilde{A}_H$ and $\tilde{A}_L$.

   (b) Compute the dynamics of the aggregate variables as explained in Section 4.1. If the chosen path is above the saddle path, then adjust $\tilde{A}_{1970}$ down. If the chosen path is below the saddle path, then adjust $\tilde{A}_{1970}$ up.

   (c) By repeating the procedure, we obtain an appropriate $\tilde{A}_{1970}$.

2. Find year T where the distance of $(\tilde{K}_T, \tilde{C}_T)$ is closest to the post-1970 steady state-values, $(\tilde{K}^*, \tilde{C}^*)$.

3. Compute $\widetilde{HA}_T$. If the $\widetilde{HA}_T$ is above the post-1970 steady-state value, then adjust $\widetilde{HA}_{1970}$ down. Otherwise, adjust $\widetilde{HA}_{1970}$ up.

4. By repeating the procedure, we obtain an appropriate $\widetilde{HA}_{1970}$.

Note that since $\tilde{C}_t = v_{i,t}\tilde{A}_t$, the above procedure is similar to the shooting algorithm used in standard growth models. To compute the variables used below, we assume that after time T^*, when the dynamics of K_t and C_t are the closest to the post-1970 steady state, the economy switches to that steady state.

D Derivations of the household wealth distributions in the steady state

This appendix shows the derivations of the household wealth distributions described in Section 5.

D.1 Wealth distribution of entrepreneurs

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

$$f_e(t') = ((\nu + p_f)E) \exp\left(-(\nu + p_f)t'\right).$$

By combining them, we can calculate the probability density function of the entrepreneurs' wealth 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 the following formula to the above equation:

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

D.2 Wealth distribution of innate workers

We calculate the wealth distribution of innate workers as follows:

$$\begin{aligned} f_w(\ln \tilde{a}_i) =& f_w(t') f_w(\ln \tilde{a}_i | t') \left| \frac{dt'}{d \ln \tilde{a}_i} \right| \\ &= (\nu - (\nu + p_f)E) \exp(-\nu t') \cdot \mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{a\ell} - g)t') \cdot \frac{1}{|\mu_{a\ell} - g|} \end{aligned}$$

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

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

D.3 Wealth distribution of former entrepreneurs

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

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

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

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

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

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

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

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

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

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

E Details on the welfare analysis

In this appendix, we calculate the ex ante utilities of an entrepreneur and a worker in the steady state, which were used in Section 6.8. The value function is written as follows:

$$V^i(a_i, \mathbf{S}) = B_t^i \ln a_i + H^i(\mathbf{S}). \tag{43}$$

We then derive the utility (value function) of a worker $V^w(a_i, \mathbf{S})$. By substituting (3) and (4) into (25) and rearranging, we obtain $H^w(\mathbf{S})$ in (43) in the steady state as follows:

$$H^w(\mathbf{S}) = \frac{1}{\beta+\nu}\left[\ln(\beta+\nu) + \frac{r^f - \beta}{\beta+\nu}\right].$$

By using this equation, the value function of a worker in the steady state, whose total wealth is a_i, can be calculated by

$$V^w(a_i, \mathbf{S}) = \frac{\ln a_i}{\beta+\nu} + H^w(\mathbf{S}).$$

Next, from the above results, we derive the utility (value function) of an entrepreneur. From (25), we obtain $H^e(\mathbf{S})$ in (43) in the steady state as follows:

$$H^e(\mathbf{S}) = \frac{1}{\beta+\nu+p_f}\left[p^f H^w(\mathbf{S}) + \ln(\beta+\nu) + \frac{r^f - \beta + (\mu^q - r^f)x_e/2}{\beta+\nu}\right].$$

The value function of an entrepreneur in the steady state, whose total wealth is a_i, can be calculated by

$$V^e(a_i, \mathbf{S}) = \frac{\ln a_i}{\beta+\nu} + H^e(\mathbf{S}).$$

Section 6.8 calculates the detrended utility level defined by

$$V^i(\tilde{h}, \mathbf{S}) = \frac{\ln \tilde{h}}{\beta+\nu} + H^i(\mathbf{S}).$$