A Politico-Growth Theory of Local Fiscal Policy and Regional Redistribution

Lieyuan Yang *

Department of Economics, University of Minnesota

October 28, 2025

Abstract

This paper proposes a politico-economic theory of regional dynamics in an environment comprising a set of small open economies, where local governments allocate fiscal spending between public goods and investment without commitment and the central government persistently redistributes fiscal resources across regions. I analytically characterize local fiscal policy, aggregate dynamics, and the steady state under a progressive redistribution scheme. The framework is used to compute the time path of optimal progressivity subject to political implementability constraints.

Keywords: Fiscal Policy, Progressivity, Regional Redistribution, Markov Equilibrium, Political Economy, Economic Growth, Local Government, Government Investment, Public Goods, Intergenerational Conflict

JEL codes: D72, E60, E62, H40, H70, O11, O23, R10

1 Introduction

Over the past three decades, China has significantly expanded transfers and grants to less-developed regions through fiscal redistribution, the Great Western Development

*University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455. Email: yang7481@umn.edu. I am deeply grateful to Kjetil Storesletten for his invaluable guidance and support.

Program, and targeted poverty alleviation. Similarly, the European Union provides substantial cohesion and recovery funds to support disadvantaged member states and regions (Becker, Egger and Von Ehrlich, 2010, 2013). In the United States, the federal government channels large intergovernmental grants for healthcare, transportation, education, and social welfare to individual states (Clemens and Veuger, 2023). Other federal systems—such as Switzerland, Germany, and Canada—also operate fiscal equalization schemes that transfer resources across jurisdictions (Buettner, 2006). While such redistribution practices are widespread and long-standing, the theoretical understanding of regional economic dynamics under redistribution and the implementability of optimal policy remains limited.

To address these issues, this paper constructs a general equilibrium politico-growth model of local fiscal policy and regional redistribution. I extend the neoclassical growth model along two dimensions. First, the economy consists of a set of small-open-economy regions. Each region is governed by a one-period-tenured local government that purchases public goods and invests in local infrastructure to maximize the welfare of current local households. Growth is driven by infrastructure investment, which reflects the intergenerational conflict between young and old generations and the dynamics of political turnover. The elderly prefer greater spending on public goods, while the young favor investment that expands future fiscal capacity. The interaction between current and future governments creates a strategic dynamic in local fiscal choices. Second, the central government redistributes fiscal resources across regions through a progressive tax scheme. Richer local governments face higher marginal tax rates. The central government, with full commitment, chooses redistribution policy to maximize the welfare of all generations, balancing regional equity against long-run economic growth.

The demographic and governmental structure of the model follows Song, Storesletten and Zilibotti (2012). The overlapping generation (OLG) households live for two periods and derive utility from public goods when old. I characterize local fiscal policy in a Markov equilibrium, where each one-period-tenured local government allocates spending between public goods and infrastructure investment, taking the policy of its successor as given. Public goods directly benefit the elderly, while investment enhances future public goods provision, benefiting today's young when they age. Meanwhile, the central government redistributes fiscal resources through the log-linear progressive tax function following Heathcote, Storesletten and Violante (2017) (HSV hereafter). Current progressivity transfers resources from high- to low-productivity regions, reducing regional inequality. Future progressivity lowers the return to local government investment, inducing a shift in local fiscal spending away from investment toward public goods, thereby dampening economic growth. The model is analytically tractable and yields closed-form expressions for local Markov policy, the transition path under time-varying redistribution policy, and the steady state.

The structure of the model makes the central government's Ramsey problem over time-varying progressivity straightforward to compute. This stands in sharp contrast to canonical heterogeneous agent Ramsey settings such as Aiyagari (1995), which require tracking the full cross-sectional distribution of state variables and handling non-recursive forward-looking implementability constraints (Marcet and Marimon, 2019). In such environments, solving for time-varying policy often entails optimizing over hundreds or thousands of variables (Acikgoz et al., 2023). In my setting, the local government's problem admits a closed-form solution, allowing the planner's implementability constraint to be expressed as tractable laws of motion for state variables. Moreover, with log-normally distributed idiosyncratic shocks, these state variables aggregate cleanly, and the heterogeneous-region economy therefore collapses to a pseudo-representative formulation.[1] As a result, the optimal policy path is characterized by a low-dimensional system of difference equations that can be efficiently solved using a standard shooting algorithm.

The theory highlights key forces shaping optimal redistribution policy. On the one hand, the central government has a redistributive motive to increase progressivity and reduce regional inequality. On the other hand, intergenerational conflict constrains the central government's ability to do so. In the model, local governments are one-period-tenured and choose fiscal policy to maximize the welfare of current generations. Consequently, under political pressure from the elderly, local governments tend to favor higher public goods provision at the expense of investment. While greater progressivity

[1]This approach is analogous to Acharya and Dogra (2020); Acharya, Challe and Dogra (2023), who show that with CARA utility and normally distributed income shocks, the household Euler equation aggregates linearly, yielding a tractable pseudo-representative agent formulation in a heterogeneous agent New Keynesian model.

is desirable to reduce regional inequality, the central government has to implement a lower one to preserve long-run growth and maintain the welfare of future generations by encouraging local government investment. The model therefore delivers normative implications for redistribution design, particularly in the context of demographic change. I calibrate the model and compute time-varying optimal redistribution policy under different model specifications.

This paper builds on and contributes to several strands of literature. First, it closely relates to the long-standing theoretical tradition that incorporates government spending or public capital into the growth model pioneered by Barro (1990). Key contributions include Barro and Sala-i Martin (1992); Baxter and King (1993); Turnovsky (2000); Leeper, Walker and Yang (2010); Arrow and Kruz (2013); Brandt, Tombe and Zhu (2013); Bouakez, Guillard and Roulleau-Pasdeloup (2017); Boehm (2020); Ramey (2020); Henderson et al. (2022). Most recently, Song and Xiong (2023) emphasizes the role of local government investment in China's economic growth. This paper contributes by analyzing the political determinants of government spending and their implications for long-run economic growth.

Second, this paper is methodologically similar to the literature that characterizes public policy as a Markov perfect equilibrium. Key contributions include Krusell and Rios-Rull (1999), Klein and Ríos-Rull (2003) and Klein, Krusell and Rios-Rull (2008) on taxation; Song, Storesletten and Zilibotti (2012) and Müller, Storesletten and Zilibotti (2016) on public expenditure and debt; Debortoli, Nunes and Yared (2017) on government debt maturity; Benigno et al. (2016) and Bianchi and Mendoza (2018) on macroprudential policy; Chari, Dovis and Kehoe (2020) on financial repression; Moser and Yared (2022) on pandemic lockdown; and Bhattarai, Eggertsson and Gafarov (2023) on quantitative easing. My work contributes by providing an analytical framework that characterizes optimal local government investment in a Markov perfect equilibrium and examines its implications for economic growth and the optimal design of regional redistribution.

Moreover, this paper relates to the literature on Ramsey-style optimal tax design (Ramsey, 1927). I model regional redistribution as a progressive tax on local fiscal income using the HSV-style log-linear tax function (Heathcote, Storesletten and Violante, 2017). While most existing studies apply the HSV framework to income taxation (e.g., Fajgelbaum et al., 2019; Sachs, Tsyvinski and Werquin, 2020; Heathcote and Tsujiyama, 2021; Berger, Herkenhoff and Mongey, 2025), my work differs by using it to study redistribution across regions. Solving the Ramsey problem with time-varying policy during transitional dynamics in heterogeneous-agent models is computationally challenging. Recent advances include the recursive formulation in Marcet and Marimon (2019), the tractable first-order characterization in Acikgoz et al. (2023), parametric optimal policy paths in Dyrda and Pedroni (2023), and truncation methods in Le Grand, Martin-Baillon and Ragot (forthcoming) and Le Grand and Ragot (2025). I contribute by constructing an environment that reduces the heterogeneous-agent Ramsey problem to a pseudo-representative setting, in which the non-recursive implementability constraint is replaced by tractable laws of motion for aggregate state variables. This structure enables efficient computation of the optimal policy path.

The remainder of the paper is organized as follows. Section 2 introduces the economic environment. Section 3 characterizes the local fiscal Markov policy, aggregate dynamics, and the steady state under regional redistribution. Section 4 presents the central government's planning problem and analyzes the design of optimal redistribution policy. Section 5 calibrates the model and examines its quantitative implications. Section 6 discusses empirical implications. Section 7 concludes.

2 Economic Environment

This section describes the economic environment. The economy consists of a unit measure of regions, indexed by $i \in [0, 1]$. Each region is a small open economy populated by two-period-lived OLG households who inelastically supply unit labor and operate the local firm in the first period and live off savings in the second period. Households derive utility from local public goods g when they are old. Each region has a representative firm and a one-period-tenured local government that chooses public goods g and next-period public capital k_G. A permanent central government governs the entire economy. Regions differ in local productivity A and accumulated public capital k_G. Capital fully depreciates after one period.

2.1 Regional Redistribution policy

The central government redistributes fiscal resources using the HSV tax function:

$$T(x) = x - \lambda x^{1-\phi} \tag{1}$$

where $T(x)$ denotes the net tax paid to the central government, x is the pre-redistribution fiscal income, and $\lambda(x)^{1-\phi}$ corresponds to the post-redistribution fiscal income. The parameter $1-\phi$ determines the pass-through from pre- to post-redistribution fiscal income. A higher ϕ indicates a higher degree of progressivity. The HSV function nests two extreme case: the linear tax scheme when $\phi = 0$ and the uniform post-redistribution income when $\phi = 1$.

In Yang (2025), I show that the HSV function provides a good description of China's regional redistribution system and the estimated progressivity has increased dramatically over time (see Figure D1).

2.2 Timeline

The timeline is described as follows. The central government first announces and commits to a path of central fiscal policy $\Phi = \{\phi_t, \lambda_t, \tau_t\}_{t=0}^{\infty}$. In each period, given the tax rate τ_t, local firms produce a homogeneous good yit and pay taxes $\tau_t y_{it}$. Young agents supply one unit of labor, operate firms, consume, and save, while old agents live off their savings. Local governments collect taxes from firms and pay taxes to (or receive subsidies from) the central government according to the redistribution policy (ϕ_t, λ_t). They then choose the level of local public goods g_{it} and next-period public capital k_{Git+1}, subject to the post-redistribution budget constraint.

2.3 Production and Firm

The production function is Cobb-Douglas with constant returns to scale:

$$y_{it} = A_{it} k_{Git}^{\alpha} k_{it}^{\theta} n_{it}^{1-\alpha-\theta} \tag{2}$$

where α and θ are output elasticities of public and private capital, respectively. Total factor productivity (TFP), A_{it}, consists of the permanent component A_i and a transitory component ε_{it}:

$$\log A_{it} = \log A_i + \log \varepsilon_{it} \tag{3}$$

This assumption is made for analytical convenience. Appendix B extends the analysis to a richer process with both persistent and transitory TFP shocks.

Firms rent capital at an exogenous constant interest rate R and hire labor at the local wage rate w_{it} to maximize after-tax profits:

$$\max_{k_{it},n_{it}} (1-\tau_t)y_{it} - Rk_{it} - w_{it}n_{it} \tag{4}$$

where y_{it} is given by the production function (2). Imposing the labor market clearing condition $n_{it} = 1$ and solving for the firm's optimization problem yields output, the wage rate, and profits:

$$\begin{aligned}
y_{it} &= \left(\frac{\theta(1-\tau_t)}{R}\right)^{\frac{\theta}{1-\theta}} A_{it}^{\frac{1}{1-\theta}} k_{Git}^{\frac{\alpha}{1-\theta}} \\
w_{it} &= (1-\alpha-\theta)\left(\frac{\theta}{R}\right)^{\frac{\theta}{1-\theta}} (1-\tau_t)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}} \\
\pi_{it} &= \alpha\left(\frac{\theta}{R}\right)^{\frac{\theta}{1-\theta}} (1-\tau_t)^{\frac{1}{1-\theta}} A_{it}^{\frac{1}{1-\theta}} k_{Git}^{\frac{\alpha}{1-\theta}}
\end{aligned} \tag{5}$$

2.4 Household

The young agent inelastically supplies unit labor and operates the local firm. Households derive utility from household consumption c in both periods and from local public goods g when old. This assumption follows Song, Storesletten and Zilibotti (2012); Müller, Storesletten and Zilibotti (2016). The utility function is assumed to be logarithmic. The problem of the young agent is given by:

$$\begin{aligned}
\max_{c_{it}^Y, c_{it+1}^O} U_{it} &= \log c_{it}^Y + \beta(\log c_{it+1}^O + \chi \log g_{it+1}) \\
\text{s.t.} \quad & c_{it}^Y + s_{it+1} \leq w_{it} + \pi_{it} \\
& c_{it+1}^O \leq (1+R)s_{it+1}
\end{aligned} \tag{6}$$

where g_{it+1} is exogenously chosen by the local government, β is the discount factor, and χ governs the relative preference on public goods. Solving this problem yields optimal consumption allocations:

$$c_{it}^Y = \frac{1}{1+\beta}(1-\theta)(1-\tau_t)y_{it}, \quad c_{it+1}^O = \frac{\beta(1+R)}{1+\beta}(1-\theta)(1-\tau_t)y_{it} \tag{7}$$

and the indirect utility of young and old agents of as functions of local output y, local public goods g, and the policy parameter τ (omitting the redundant terms):

$$\begin{aligned} U_{it}^O(y_{i,t-1}, g_{it}) &= \log y_{it-1} + \chi \log g_{it} + \log(1-\tau_{t-1}) \\ U_{it}^Y(y_{it}, g_{it+1}) &= (1+\beta)\log y_{it} + (1+\beta)\log(1-\tau_t) + \beta\chi \log g_{it+1} \end{aligned} \tag{8}$$

2.5 Local Government

Local governments have a one-period tenure and choose current public goods and next-period public capital to maximize a weighted sum of young and old agents' utilities. The weight captures the relative political influence of each generation, reflecting factors such as group size, voting turnout, and political responsiveness in a probabilistic voting model like Persson and Tabellini (2002) and Song, Storesletten and Zilibotti (2012). Let ω denote the weight on the old agent. Throughout, I refer ω as the local government Pareto weight on the old. The local government problem is given by:

$$\begin{aligned} \max_{g_{it}, k_{Git+1}} \quad & \omega U_{it}^O(y_{i,t-1}, g_{it}) + (1-\omega) U_{it}^Y(y_{it}, g_{it+1}) \\ \text{s.t.} \quad & g_{it} + k_{Git+1} = \lambda_t(\tau_t y_{it})^{1-\phi_t} \end{aligned} \tag{9}$$

where U_{it}^O and U_{it}^Y are indirect utilities given by Equation (8). Local government debt is not allowed. The assumption that local governments care only about current generations is made for tractability and can be micro-founded using a probabilistic voting model. Appendix C extends the model to allow local governments to care about future generations.

2.6 Central Government

The central government is committed to choosing the path of policy $\Phi = \{\phi_t, \lambda_t, \tau_t\}_{t=0}^{\infty}$. In each period, the central budget must satisfy a balanced budget constraint:

$$\int \tau_t y_{it} di = \int \lambda_t (\tau_t y_{it})^{1-\phi_t} di \quad \forall t \tag{10}$$

where central government debt is also ruled out. If regional output is log-normally distributed, i.e. $\log y_{it} \sim \mathcal{N}(\mu_{y_t}, \sigma_{y_t}^2)$, the equilibrium value of λ_t can be given by:

$$\log \lambda_t = \phi_t \mu_{y_t} + \frac{1-(1-\phi_t)^2}{2} \sigma_{y_t}^2 + \phi \log \tau_t \tag{11}$$

where μ_{y_t} and $\sigma_{y_t}^2$ denote the mean and variance of log output, respectively.

3 Equilibrium

This section analytically characterizes local fiscal policy, aggregate dynamics, and the steady state in equilibrium. Full details and proofs are provided in Appendix A. The equilibrium consists of local fiscal policy function and the induced laws of motion for regional and aggregate public capital. In each region, local fiscal policy is determined by the dynamic game between successive local governments. I restrict attention to Markov-perfect equilibria where local governments condition their strategy solely on current states $(A_i, \varepsilon_{it}, k_{Git})$, which include the permanent and transitory components of productivity and the stock of public capital. Since each region is small, its fiscal policy does not affect the central government policy. Therefore, local governments treat the path of central policy Φ as exogenously given. I start by formally defining the equilibrium.

Definition 1 (Markov-perfect equilibrium)**.** *Given the initial condition $\{A_{i0}, k_{Gi0}\}_i$, the policy sequence $\Phi = \{\phi_t, \lambda_t, \tau_t\}_{t=0}^{\infty}$, and the state variable for a region $s_{it} = (A_i, \varepsilon_{it}, k_{Git})$, a Markov-perfect equilibrium (MPE) consists of allocations $\{y_{it}, k_{it}, k_{Git}, c_{it}^Y, c_{it+1}^O, g_{it}\}_{i,t}$, wage rates $\{w_{it}\}_{i,t}$, the local fiscal policy rules $\mathcal{K} : \mathbb{R}_+^3 \to \mathbb{R}_+$, $k_{Git+1} = \mathcal{K}(s_{it}; \Phi)$, and $\mathcal{G} : \mathbb{R}_+^3 \to \mathbb{R}_+$, $g_{it} = \mathcal{G}(s_{it}; \Phi)$, and the distribution over states $\Omega(s_{it})$, such that:*

1. *Given the regional state s_{it} and output tax rate τ_t, (y_{it}, k_{it}) solves the firm's problem* (4) *and (c_{it}^Y, c_{it+1}^O) solves the household problem* (6).

2. *Given the central fiscal policy Φ, firm production plan, and household optimal allocations, the policy rules $\mathcal{K}(s_{it};\Phi)$ and $\mathcal{G}(s_{it};\Phi)$ solves the local government's problem* (9), *where $g_{it+1} = \mathcal{G}((A_i, \varepsilon_{it+1}, \mathcal{K}(s_{it};\Phi));\Phi)$.*

3. *The wage rate w_{it} clears the regional labor market: $n_{it} = 1$.*

4. *The central government budget constraint* (10) *holds in each period.*

5. *The distribution $\Omega(s_{it+1})$ is induced by the TFP process* (3) *and the policy rule $k_{Git+1} = \mathcal{K}(s_{it};\Phi)$.*

Definition 2 (Stationary Markov-perfect equilibrium)**.** *A stationary Markov-perfect equilibrium (SMPE) is a MPE in which the distribution $\Omega(s)$ is time-invariant under the constant policy $\bar{\Phi} = (\phi, \lambda, \tau_y)$*

To make the aggregate dynamics tractable, I impose two distributional assumptions. First, both permanent and transitory components of TFP are log-normally distributed:

$$\log A_i \sim \mathcal{N}(-\frac{\sigma_a^2}{2}, \frac{\sigma_a^2}{2}), \quad \log \varepsilon_{it} \sim \mathcal{N}(-\frac{\sigma_\varepsilon^2}{2}, \frac{\sigma_\varepsilon^2}{2}), \quad \mathrm{cov}(\log A_i, \log \varepsilon_{it}) = 0$$

Second, the joint distribution of the initial state (A_{i0}, k_{Gi0}) is also assumed to be log-normal.

3.1 Local Fiscal Policy

The local government allocates resources between current public goods and future public capital to maximize the weighted welfare of old and young agents. On the one hand, increasing public goods directly raises the utility of the current old agents. On the other hand, investment in future public capital raises output and fiscal income for the next-period government, which in turn can fund more public goods for the current young agents once they become old. Taking the current state s_{it} and the policy rule of the succeeding government as given, the local government's optimality condition is characterized by the generalized Euler equation (GEE):

$$\frac{\omega}{g_{it}} = \mathbb{E}\left[\frac{(1-\omega)\beta}{g_{it+1}}\frac{\partial\mathcal{G}(A_i,\varepsilon_{it+1},k_{Git+1};\Phi)}{\partial k_{Git+1}}\middle| s_{it}\right] \tag{12}$$

This GEE (12) equates the marginal cost and the expected marginal benefit of government investment. The left-hand side represents the marginal cost: allocating more resources to investment crowds out current public goods provision, reducing utility for the old agent. The right-hand side captures the expected marginal benefit: higher public investment increases future output and fiscal income, allowing the succeeding local government to provide more public goods, thereby improving utility for young agents when they grow old.

Let $\kappa_{t+1} = k_{Git+1}/[\lambda_t(\tau_t y_{it})^{1-\phi_t}]$ denote the investment share in post-redistribution fiscal income. I guess and verify that κ_{t+1} is determined purely by policy and preference parameters and is independent of regional productivity levels or the successor government's choice κ_{t+2}. The intuition is as follows: higher productivity A_{it} or a lower successor investment share κ_{t+2} raises the marginal return to investment (the derivative term on the right-hand side of GEE (12)), but also reduces the marginal utility of next-period public goods (due to larger g_{it+1}). With logarithmic utility, these two effects exactly cancel out, leaving the optimal investment share invariant to productivity levels and successor behavior. Proposition 1 characterizes the optimal local government investment.

Proposition 1 (Local fiscal policy)**.** *The local government investment in the equilibrium is given by:*

$$\log k_{Git+1} = (1-\phi_t)\log y_{it} + \log\kappa(\phi_{t+1}) + \log\tau_t + \left[\phi\mu_{y_t} + \frac{\sigma_{y_t}^2}{2} - \frac{(1-\phi_t)^2\sigma_{y_t}^2}{2}\right] \tag{13}$$

where $\kappa(\phi_{t+1})$ is the investment share of post-redistribution fiscal income, given by:

$$\kappa(\phi_{t+1}) = \frac{\alpha(1-\phi_{t+1})(1-\omega)\beta}{(1-\theta)\omega + \alpha(1-\phi_{t+1})(1-\omega)\beta} \tag{14}$$

It is straightforward to verify that, when public goods are computed using the local government budget constraint, the fiscal policy rule given by Equation (13) satisfies the GEE (12). Investment k_{Git+1} depends on the current state $(A_i, \varepsilon_{it}, k_{Git})$ through the local output term y_{it} and on the policy parameters $(\phi_t, \phi_{t+1}, \tau_t)$. This structure satisfies the definition of a Markov-perfect equilibrium, as policy depends only on current-period states and parameters. The expression for λ_t is substituted using its equilibrium value in Equation (11).

Local government investment consists of four components. The first term captures the responsiveness of investment to productivity and capital stock. High-productivity regions invest more, but current progressivity ϕ_t lowers this elasticity. A more progressive current redistribution scheme reallocates fiscal resources from high- to low-productivity regions, reducing the pass-through from permanent productivity, transitory shocks, and public capital stock to next-period investment.

The second term reflects the investment share in post-redistribution fiscal income. This term captures the distortion induced by future progressivity: a higher future progressivity ϕ_{t+1} lowers the return to public investment, motivating local governments to shift spending away from investment toward public goods. Notably, the investment share $\kappa(\phi_{t+1})$ is increasing in the discount factor β and decreasing in the local old-generation weight ω. A higher β implies that young agents place more value on future public goods, incentivizing greater investment. Conversely, a higher ω strengthens the political influence of old agents, inducing a reallocation toward public goods.

The third term captures the effect of the output tax rate. A higher tax rate increases the overall fiscal capacity of the local government, thereby expanding investment. This effect does not interact with progressivity ϕ, due to the fact that the HSV tax function is homogeneous of degree one when λ has an elasticity of ϕ with respect to the scale of y as shown in Equation (11). The fourth bracketed term arises from the effect of cross-regional heterogeneity. As progressivity increases, redistribution becomes more aggressive in an unequal economy, raising the level of investment for all regions through the equilibrium adjustment of λ_t. This reflects the general equilibrium effect that greater inequality magnifies the redistributive force of a given level of progressivity.

3.2 Aggregate Dynamics

Since the initial distribution of regional states is assumed to be log-normal and the law of motion indicated by Equation (13) is log-linear, the cross-sectional distribution of states remains log-normal in every period. Although the primitive state variables are productivity (A_i, ε_{it}) and public capital k_{Git}, it is more convenient to track the distribution of log output and the covariance between productivity and public capital as sufficient statistics. Let $\sigma_{Ak_G,t} = \text{cov}(\log A_{it}, \log k_{Git})$ denote this covariance and $Y_t = \int y_{it} di$ denote the aggregate output. Because the distribution of productivity is time-invariant, the set of aggregate moments $(\log Y_t, \sigma^2_{y_t}, \sigma_{Ak_G,t})$ is sufficient to characterize the macroeconomic state. All other aggregates of interest—such as aggregate capital and capital dispersion—are affine functions of these moments. Proposition 2 characterizes the evolution of this state vector as a log-linear dynamic system governing the distribution of output and the allocation of public capital across regions.

Proposition 2 (Aggregate dynamics)**.** *The aggregate dynamics are characterized by the evolutions of aggregate output Y_t, the variance of log output $\sigma^2_{y_t}$, and the covariance between productivity and public capital $\sigma_{Ak_G,t}$:*

$$\begin{bmatrix} \log Y_{t+1} \\ \sigma^2_{y_{t+1}} \\ \sigma_{Ak_G,t+1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2(\phi_t) & A_3(\phi_t) \\ 0 & B_2(\phi_t) & B_3(\phi_t) \\ 0 & 0 & C_3(\phi_t) \end{bmatrix} \begin{bmatrix} \log Y_t \\ \sigma^2_{y_t} \\ \sigma_{Ak_G,t} \end{bmatrix} + \begin{bmatrix} A_0(\phi_t, \phi_{t+1}, \tau_t, \tau_{t+1}) \\ B_0(\phi_t) \\ C_0(\phi_t) \end{bmatrix} \tag{15}$$

where the coefficient mappings $\{A_j(\cdot), B_j(\cdot), C_j(\cdot)\}_{j\in\{0,1,2,3\}}$ are deterministic functions of the policy sequence and the structural parameters, defined in Appendix A.2.

Redistribution policy influences the distribution of output through three main channels. First, current progressivity ϕ_t redistributes fiscal resources from high- to low-productivity regions. This dampens the correlation between productivity and public capital stock by (1) attenuating the transmission from current correlation to future correlation, captured by the multiplier $C_3(\phi_t)$, and (2) weakening the pass-through from permanent productivity component to investment, captured by $C_0(\phi_t)$. This channel has two further implications. First, a lower correlation between public capital and productivity compresses output dispersion, reflected in the coefficients $B_3(\phi_t)$ and $B_0(\phi_t)$.

Second, redirecting capital from high- to low-productivity regions induces misallocation, reducing aggregate efficiency and output, as captured by $A_3(\phi_t)$ and the ϕ_t-term in $A_0(\cdot)$. All of these terms are falling in current progressivity ϕ_t.

Second, progressivity reduces the pass-through from pre- to post-redistribution fiscal income, narrowing the investment gap between rich and poor regions. This compression directly attenuates the transmission from current to future output dispersion. This strength of this channel is captured by $B_2(\phi_t)$, which is falling in ϕ_t. Furthermore, because output is log-normally distributed, holding the mean of output, an increase in output dispersion requires that output losses in low-productivity regions outweigh the gains in high-productivity regions. As a result, higher $\sigma^2_{y_t}$ reduces aggregate output growth. This is captured by $A_2(\phi_t) < 0$. However, since progressivity reduces the pass-through from output to investment, it compresses the dispersion in investment and thereby mitigates the adverse growth effect of output inequality. That is, $\partial A_2(\phi_t)/\partial \phi_t > 0$, so higher progressivity dampens the negative impact of output dispersion on output growth.

Third, future progressivity reduces the return to investment, inducing local governments to shift fiscal spending away from investment toward public goods. This reallocation lowers aggregate output in the next period. With logarithmic utility, this effect is uniform across regions and does not depend on local productivity or capital levels. It is captured by the ϕ_{t+1}-term in the constant component of the output law of motion, $A_0(\cdot)$, with $\partial A_0(\cdot)/\partial \phi_{t+1} < 0$. This reflects the distortionary effect of future redistribution policy on fiscal incentives: higher expected progressivity reduces the marginal benefit of investing today, leading to lower capital accumulation and output tomorrow.

I define TFP as the Solow residual, which is calculated by $\log \mathrm{TFP}_t = \log Y_t - \alpha \log K_{Gt} - \theta \log K_t$, where K_{Gt} and K_t are aggregate public and private capital, respectively. Productivity A_{it} satisfies the mean-preserving spread in every period and does not include any drift or deterministic trend component. Therefore, aggregate TFP in my model captures how effectively capital is allocated across regions. Proposition 3 characterizes the evolution of aggregate TFP relative to a *non-redistribution* economy ($\phi = 0$).

Proposition 3 (Aggregate TFP growth)**.** *Let* TFP* *denote aggregate TFP in the non-redistribution economy* ($\phi = 0$)*. The growth in aggregate TFP relative to the non-redistribution economy is given by*

$$\Delta \log \text{TFP}_{t+1} - \Delta \log \text{TFP}^*_{t+1} = \underbrace{\frac{\alpha(1-\theta-\alpha)[1-(1-\phi_t)^2]}{(1-\theta)} \frac{\sigma^2_{yt}}{2}}_{\text{Insurance effect}} \underbrace{- \frac{\alpha\phi_t}{(1-\theta)^2}(\sigma^2_a + \alpha\sigma_{Ak_G,t})}_{\text{Distortionary effect}}$$

First, as discussed above, greater output dispersion leads to an inefficient capital allocation: capital is disproportionately directed toward low-productivity regions, which diminishes aggregate efficiency. Progressivity mitigates this TFP loss by compressing the pass-through from current output to future investment, thereby reducing the transmission of output dispersion. This channel is captured by the first term, labeled as the "*insurance effect*". Second, an efficient allocation requires that capital is allocated proportional to productivity. Redistribution towards low-productivity regions weakens this allocation by attenuating the propagation of the capital-productivity covariance and by reducing the responsiveness of investment to the permanent productivity component. This distortion lowers allocative efficiency and is captured by the second term, labeled as the "*distortionary effect*".

3.3 Steady State

Proposition 4 characterizes steady-state aggregate output and TFP as functions of time-invariant policy parameters, structural parameters, and distribution parameters of productivity.

Proposition 4 (Steady state)**.** *In the steady state, the aggregate output is given by:*

$$\log Y = \frac{1}{1-\theta-\alpha}\left[\alpha \log \tau + \theta \log\left(\frac{\theta(1-\tau)}{R}\right) + \alpha \log \kappa(\phi) + \log \text{TFP}\right] \tag{16}$$

where aggregate TFP is given by:

$$\log \mathrm{TFP} = \underbrace{\frac{\theta+\alpha}{1-\theta-\alpha}\frac{\sigma_a^2}{2} + \frac{\theta-\alpha}{1-\theta+\alpha}\frac{\sigma_\varepsilon^2}{2}}_{\text{TFP in non-redistribution economy}} \\ \underbrace{-\frac{\alpha(1-\theta)\phi^2}{(1-\theta-\alpha)(1-\theta-\alpha(1-\phi))^2}\frac{\sigma_a^2}{2}}_{\text{Distortionary effect}} + \underbrace{\frac{\alpha(1-\theta)[1-(1-\phi)^2]}{(1-\theta+\alpha)[(1-\theta)^2-\alpha^2(1-\phi)^2]}\frac{\sigma_\varepsilon^2}{2}}_{\text{Insurance effect}} \tag{17}$$

In Equation (16), the steady-state aggregate output consists of four components. The first term reflects the positive effect of a higher output tax rate, which raises fiscal capacity and enables more public investment. The second term captures the distortionary effect of output taxation on firms' production decisions, which lowers output. The third term reflects the spending share of public investment, $\kappa(\phi)$, which scales the level of public capital and aggregate output and is falling in progressivity. The fourth term is aggregate TFP, which summarizes allocative efficiency across regions.

Equation (17) further decomposes steady-state TFP into the non-redistribution component and two policy-induced components. The first line represents the steady-state aggregate TFP in the non-redistribution economy. With the mean-preserving spread, a greater permanent productivity variance raises TFP by allocating capital towards more productive regions. The effect of the variance of the transitory shock σ_ε^2 is ambiguous: since firms adjust statically each period, dispersion in the transitory shock can raise efficiency on the firm side, while the local-government mapping from the transitory shock to the next-period public capital can generate misallocation. The second line captures the effect of redistribution. The distortionary effect (first term) arises because progressivity weakens the link between permanent productivity component and investment, lowering TFP. In contrast, the insurance effect (second term) reflects how progressivity reduces the responsiveness of local government investment to transitory shocks, improving TFP.

This section analytically characterizes the optimal local fiscal policy, the induced aggregate dynamics, and steady-state aggregates. In brief, a higher future progressivity of regional redistribution lowers the fiscal return to public investment, shifting local fiscal spending from investment to public goods and thereby lowering aggregate output. With logarithmic utility, this effect is uniform across regions. Current progressivity reallocates resources from high- to low-productivity regions and compresses regional inequality. Its net impact on aggregate efficiency depends on the interaction between redistribution and the permanent versus transitory components of TFP. Building on this framework, the next section formulates the central government's planning problem and characterizes the optimal policy.

4 Planning Problem and Optimal Policy

What is the optimal path of redistribution? Due to the concavity of utility function, the central government has an incentive to redistribute toward low-productivity regions. However, higher progressivity discourages local public investment, lowers future output, and induces capital misallocation. A Ramsey planner chooses the path of redistribution to balance equity across regions with aggregate efficiency and growth over time.

Let $U_t^m = \int U_{it}^m di$ denote the average utility of generation $m \in \{Y, O\}$ born in period t, where U_{it}^m is the indirect household utility given by Equation (8). The central government announces and commits to a policy path $\{\phi_t, \tau_t\}_{t=0}^{\infty}$. Following Farhi and Werning (2007), the planner assigns geometrically declining Pareto weights ψ^{t+1} to generations born in period t. The social welfare of implementing the policy sequence $\{\phi_t, \tau_t\}_{t=0}^{\infty}$ is given by:

$$\begin{aligned}\mathcal{W}(\{\phi_t\}_{t=0}^{\infty}, \{\tau_t\}_{t=0}^{\infty}) &= \beta U_0^O + \sum_{t=0}^{\infty} \psi^{t+1} U_t^Y \\ &= \sum_{t=0}^{\infty} \psi^{t+1} \left[\Omega_1 \log Y_t - \Omega_2(\phi_t) \frac{\sigma_{y_t}^2}{2} + \Omega_3(\phi_{t+1}, \tau_t) \right]\end{aligned} \tag{18}$$

where the coefficient mappings $\{\Omega_j\}_{j \in \{1,2,3\}}$ are deterministic functions of policy and structural parameters defined in Appendix A.5.

The social welfare given by Equation (18) can be decomposed into three components. The first term reflects the welfare from the aggregate output level. The coefficient Ω_1 depends only on preference parameters. The second term represents the disutility from output dispersion. Since $\partial \Omega_2(\phi_t)/\partial \phi_t < 0$, higher progressivity reduces output dispersion and improves social welfare. The third term is the combination of policy and structural parameters. It reflects the welfare consequence of distortionary output taxation and the local government's choice of investment and public goods.

The central government's Ramsey problem is to maximize the social welfare function (18) subject to the laws of motion for aggregate state variables (15), which serve as the implementability constraint. The planner chooses the entire policy path $\{\phi_t, \tau_t\}_{t=0}^{\infty}$, taking into account its dynamic effects on the endogenous distribution of output and capital. The first-order conditions characterizing the optimal policy are derived in Appendix A.5. This section discusses the main analytical results.

Output tax rate. The optimal output tax rate is time-invariant and is given by:

$$\tau_t^* = \tau^* = \frac{(1-\theta)\beta\chi + \psi^2\alpha(1+\beta)}{\beta\chi + \psi(1+\beta)}$$

The optimal tax rate is increasing in the discount factor β, relative preference over public goods χ, and the output elasticity of public capital α. A higher β implies that future public goods are valued more by young agents, a larger χ raises the marginal utility of public goods, and a higher α increases the productivity of public capital. All cases raises the return to increasing taxes. In contrast, the optimal tax rate falls with the output elasticity of private capital θ, which captures the distortionary effect of taxation on firm behavior. The effect of the planner's Pareto weight on future generations ψ is ambiguous: a higher ψ raises the value of future output, favoring higher taxes and investment, but also raises the relative weight on young agents who prefers cutting public goods and boosting output.

From Equation (A5), the planner's trade-off in raising the output tax rate includes a reduction output due to distortions in firm production, a lower wage rate, higher public goods provision, and greater public investment. With logarithmic utility and a linear tax scheme, the marginal effects of tax rate changes on these outcomes are independent of any states. From Equation (A7), the Lagrangian multiplier on the law of motion for $\log Y_t$ is constant over time, implying that the shadow price of future log output is time-invariant. As a result, the planner's trade-off is identical across periods, which yields a time-invariant optimal tax rate.

Optimal Progressivity. The closed-form solution for optimal progressivity is unavailable; I solve it numerically below. Here, I outline the economic forces shaping optimal progressivity. Three main considerations arise. First, the concavity of utility generates a redistributive motive for the utilitarian social planner. Inequality is costly, so the central government values higher progressivity as a means to transfer resources toward low-productivity regions. Second, redistribution entails efficiency costs. Allocating fiscal resources away from high-productivity regions reduces the correlation between capital and productivity, distorts investment incentives, and ultimately dampens aggregate output and growth. This efficiency-equity trade-off tempers the planner's redistributive motive. Third, there is a conflict between central and local governments, rooted in intergenerational and institutional frictions. Local governments represent current generations and have short time horizons. They have a tendency to increase current public goods and underweights the long-run returns to public capital. In contrast, the central government values future generations and seeks to preserve long-run economic growth. Hence, while higher progressivity is desirable to reduce regional inequality, the planner's ability to implement such policy is constrained by the preferences of selfish short-lived households and the decentralized behavior of local governments

To isolate the conflict-of-interest channel from heterogeneity, consider a representative-region benchmark with $\sigma_a^2 = \sigma_\varepsilon^2 = \sigma_y^2 = 0$. Under this restriction, the optimality condition (A6) yields the optimal progressivity in the representative-region model:

$$1 - \phi^{R*} = \frac{\omega}{1-\omega} \frac{\psi(1+\beta) + \chi\beta}{\beta^2} \frac{(1-\theta)\psi}{\chi(1-\theta-\alpha\psi)} \tag{19}$$

Optimal progressivity in the representative-region model is decreasing in ω and ψ and is increasing in β. First, ω features the intergenerational conflict. A higher ω places more weight on the old, inducing local governments to shift resources from investment to public goods. To offset this short-termism and preserve long-run growth, the central government must reduce progressivity and strengthen investment incentives. Similarly, a higher ψ raises the central planner's weight on future generations. This increases the planner's concern for long-run output, reinforcing the need to reduce progressivity and stimulate investment. By contrast, a higher β increases the young generation's valuation of future public goods. This raises local governments' own incentives to invest, thereby relaxing the central government's constraint and permitting higher progressivity.

When heterogeneity is introduced, the welfare gains from reducing inequality can partially offset the welfare losses from reduced local investment. In this case, a closed-form solution for optimal progressivity is no longer available, and I compute the optimal policy path numerically.

5 Quantitative Analysis

5.1 Algorithm

Solving heterogeneous-agent Ramsey problems is typically computationally demanding, due to the need to track the full cross-sectional distribution of state variables and to handle non-recursive forward-looking implementability constraints. In particular, solving time-varying policy during transitional dynamics require optimizing over hunderds or thousands of variables.

In contrast, my framework remains tractable for two key reasons. First, the model admits a closed-form solution to the local government's problem, so the implementability constraint in the planner's problem can be expressed as the tractable law of motion for state variables, avoiding forward-looking constraints. Second, under the log-normal specification of regional heterogeneity, the economy aggregates cleanly. The planner's problem reduces to a pseudo-representative formulation that tracks only two additional sufficient statistics, the variance of output, $\sigma^2_{y_t}$, and the covariance between productivity and public capital, $\sigma_{Ak_G,t}$. These two statistics, along with aggregate output, fully characterize the evolution of the economy. Under this structure, the planner's optimality conditions form a system of difference equations, which can be solved using a standard shooting algorithm:

1. Use the steady-state condition (A4) and the optimality conditions (A6), (A8), and (A9) to solve for the steady-state optimal policy and states.

2. Assume convergence to the steady state by a sufficiently large terminal period T and guess a path for the optimal policy.

3. Update the state forward.

4. Update the optimal policy and the associated Lagrangian multipliers backward using the planner's optimality conditions.

5. Iterate on steps 3 and 4 until convergence.

As a robustness check, I also implement a neural network projection method following Fernández-Villaverde (2025) to approximate the policy function $\phi_t = \mathcal{P}(\sigma^2_{y_t}, \sigma_{Ak_G,t})$. Since the optimal policy does not depend on the level of aggregate output, the projection only requires the two sufficient statistics. The three unused optimality conditions (A6), (A8), and (A9), form a closed system for the policy ϕ_t and the two associated Lagrangian multipliers. Figure D2 plots the policy surface $\phi = \mathcal{P}(\sigma^2_y, \sigma_{Ak_G})$ and Figure D3 presents the path for optimal policy obtained from neural network approximation. The shooting algorithm and the neural network projection produce numerically identical results. In what follows, I report results based on the shooting algorithm which is computationally more efficient.

5.2 Calibration

The model is calibrated to match empirical moments during 2013-2017 in China, which I treat as the steady state. The length of a period is 30 years. I set $\alpha = 0.185$ based on my IV estimate in Yang (2025) and $\theta = 0.315$ to target a 50% labor income share. The discount factor and interest rate are set as $\beta = (0.973)^{30}$ and $R = (1.04)^{30}$ following Song, Storesletten and Zilibotti (2012). The local government Pareto weight on the old, $\omega = 10\%$, is chosen to match the government investment share and coincidentally matches the 10% elderly population share. Productivity variance parameters $\sigma^2_a = 0.1$ and $\sigma^2_\varepsilon = 0.002$ are set to match the cross-sectional moments of TFP. The relative preference for public goods $\chi = 0.980$ and the central government Pareto weight $\psi = (0.984)^{30}$ is set to rationalize the steady-state policy $\tau = 0.3$ and $\phi = 0.485$. Table 1 summarizes the model parameterization.

5.3 Quantitative Results of Optimal Policy

I assume that the economy starts at the steady state under $\phi = 0$, while the output tax rate τ is always set at the optimal level. This scenario is labeled as the baseline model.

Table 1: Model Parameterization

Parameter	Value	Source
α	0.185	IV estimate from Yang (2025)
θ	0.315	Labor income share 50%
β	$(0.973)^{30}$	Song, Storesletten and Zilibotti (2012)
R	$(1.04)^{30}$	Song, Storesletten and Zilibotti (2012)
ω	10%	Spending share of government investment 35%
χ	0.980	Government income to GDP ratio 30%
ψ	$(0.984)^{30}$	Estimated progressivity 0.485
σ_a^2	0.1	$\text{var}(\log A_{it})$
σ_ε^2	0.002	$\text{var}(\Delta \log A_{it})$

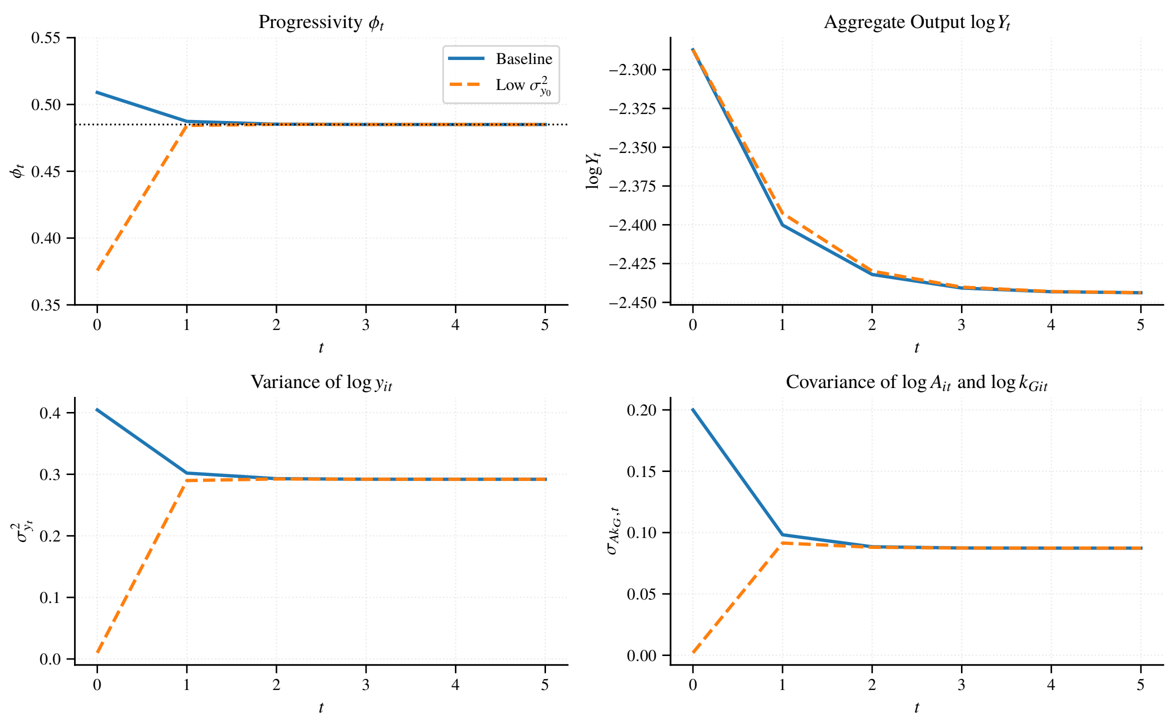
As output dispersion is quite higher in the steady state under $\phi = 0$, I consider an alternative scenario where the economy starts with very low output dispersion. Figure 1 compares the optimal progressivity paths for these scenarios.

First, the optimal policy in the steady state is invariant to the initial state. From Equations (16) and (A4), the steady state is fully pinned down by policy and structural parameters and does not depend on initial conditions. Substituting the steady-state expressions into the optimality conditions shows that the policy parameters and associated Lagrangian multipliers are independent of initial states. Therefore, the steady-state optimal policy is invariant to the initial state, consistent with the result in Acikgoz et al. (2023).

Second, the initial state does matter for the transition path of optimal policy. Raising progressivity involves a trade-off between reducing inequality and lowering future output. In the baseline case, the initial output dispersion is high, so progressivity generates substantial welfare gains by compressing regional inequality. As a result, the central government sets a relatively high level of progressivity in the initial period. In contrast, when the economy begins with low dispersion, the central government opts for lower initial progressivity to preserve investment incentives. Since capital is fully depreciated after one period and the TFP distribution is stationary, the economy converges to the steady state typically by period 2.

What factors determine the implementation of optimal policy? Figure 2 compares optimal policy paths under different parameter values. First, consistent with Equation (19), a lower ψ increases optimal progressivity, while lower β and ω reduce it. Raising

Figure 1: Path of Optimal Progressivity and Aggregate Dynamics under Optimal Policy



Note: This figure plots the transitional dynamics of the economy under the optimal policy. The top left panel shows optimal progressivity ϕ_t; the top right panel shows aggregate output $\log Y_t$; the bottom left panel shows the variance of log output $\sigma^2_{y_t}$; and the bottom right panel shows the covariance between productivity and public capital $\sigma_{Ak_{G,t}}$. Solid lines represent the baseline model where the economy starts from the steady state under $\phi = 0$, and dashed lines represent the case where the economy starts with low initial dispersion.

progressivity trades off lower inequality against lower future output. When ψ is lower, the central government puts less weights on future generations and thus favors higher progressivity to reduce inequality. In contrast, when β is lower, young agents care less about future public goods, weakening local governments' incentive to invest. Similarly, when ω is higher, local governments place higher weight on the old agents and shift spending toward public goods. In both cases, local investment falls, and to preserve future welfare, the central government responds by setting lower progressivity to restore investment incentives.

Local Pareto weights on the old agents, ω, can be interpreted as reflecting the elderly population share. During a demographic transition with population aging, how should the central government adjust redistribution policy? I conduct a policy experiment where the economy starts at the steady state under the previously computed optimal policy, and ω gradually increases over time. Figure 3 plots the path of rising ω (right panel) and the corresponding optimal policy(left panel). As the elderly polution share increases from 10% to 15%, the central government responds by steadily reducing progressivity. As aging raises the local government's weight on the old, the incentive to invest declines. To counteract this distortion and preserve future output, the central government lowers progressivity to strengthen investment incentives.

Other factors play important roles as well. In Figure 2, a lower α dampens the distortionary effects of reduced public investment and capital misallocation, thereby permitting a higher degree of progressivity. Greater dispersion in permanent TFP component increases regional inequality and thus raises the redistributive motive, leading to higher optimal progressivity. Conversely, a lower χ reduces the welfare gain from compressing inequality in local public goods, which lowers the incentive for redistribution. Figure D4 shows the path of optimal policy when the economy starts with lower output dispersion. In this case, initial progressivity is lower, while the steady-state policy remains unchanged.

Figure 2: Optimal Progressivity Paths Under Different Parameter Values

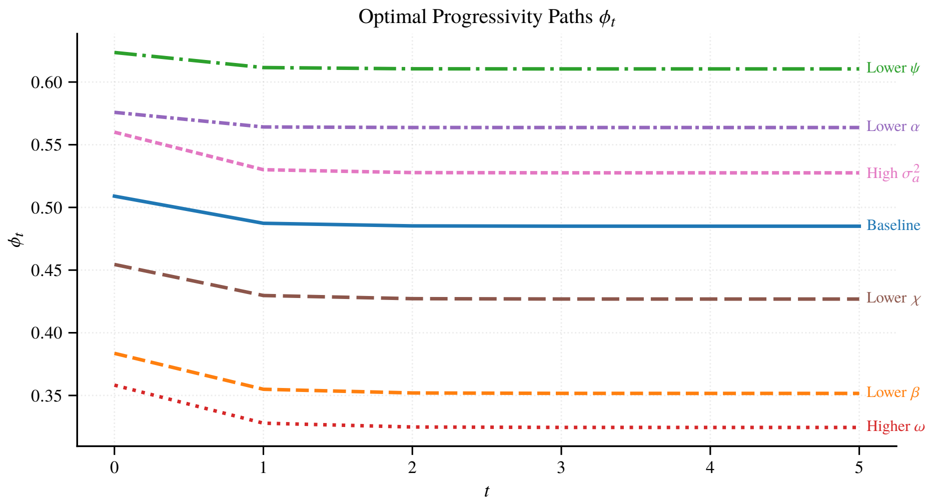
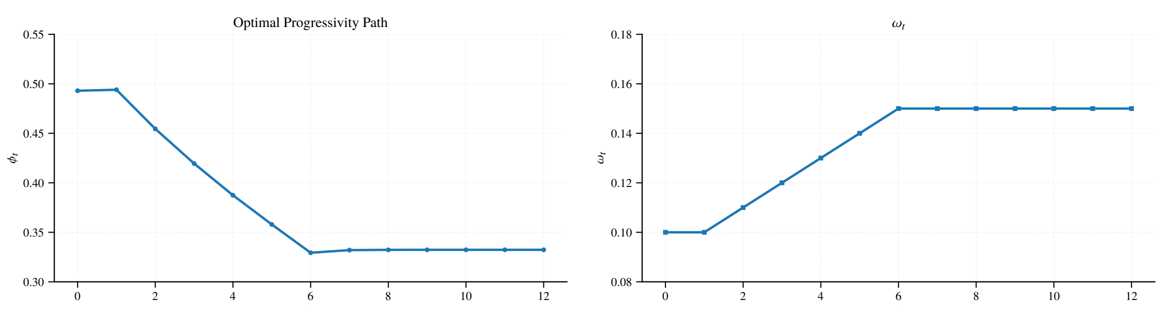


Figure 3: Optimal Progressivity During the Demographic Change



Note: This figure plot the policy experiment: the elderly population share increases over time (right panel: increasing ω_t) and the optimal progressivity decreases in response (left panel).

6 Empirical Implications

Although the model is primarily theorectical, it generates several testable implications that resonate with observed patterns in the Chinese economy under rising progressivity of regional redistribuion.

Declining propensity to invest. As rising progressivity lowers the return to government investment, local governments reallocate fiscal resources from investment toward public goods. This mechanism is consistent with the empirical findings of Song and Xiong (2023), who document that the ratio of infrastructure investment to local government revenue declined from 55.1% during 1999–2002 to 33.8% in 2013–2017.

Demographic change. Aging demographics can further reduce the propensity to invest. Equation (14) shows that local investment is decreasing in the Pareto weight on old agents, ω, which can be interpreted as the elderly population share. Notably, the calibrated value of ω aligns with China's observed elderly share. As the population ages, local governments face growing pressure to expand expenditure on pensions and healthcare, which reallocates fiscal resources away from investment. This mechanism also contributes to declining infrastructure spending.

While the model abstracts from cross-regional heterogeneity in ω, the relationship between fiscal investment and demographics can still be examined empirically. Using provincial-level data from 2016–2017, I estimate a negative correlation between the elderly population share and government investment. Specifically, regressing the share of infrastructure investment in total local government expenditure on the elderly population share yields a coefficient of –1.36 (standard error = 0.62, $R^2 = 0.15$).

This mechanism also sheds light on why China's fiscal progressivity plateaued after 2012. As the elderly population increased, local governments prioritized public goods and service provision. Further increases in progressivity would have amplified the investment disincentive, compromising long-run growth.

Regional convergence and TFP slowdown. The model predicts that progressivity-induced redistribution from high- to low-productivity regions reduces output dispersion and promotes regional convergence. This prediction aligns with the findings of Yang (2025), who document a decline in the variance of log output from 0.26 in 2004 to 0.17 in 2013. A standard Barro-style convergence regression estimates an annual convergence rate of 2.4% during this period, implying that it takes less than 30 years for any two provinces to cut GDP differences by half.

Moreover, as the estimated variance of transitory shock is relatively small, the distortionary effects dominate its insurance benefits, resulting in capital misallocation. Yang (2025) further show that aggregate TFP growth declined from over 5% in 2002 to just 1% in 2013, suggesting substantial efficiency losses.

7 Conclusion

This paper develops a tractable politico-growth model of regional redistribution and regional dynamics. The model analytically characterizes local fiscal policy, aggregate dynamics, and the steady state under a progressive redistribution scheme. Current progressivity reallocates resources toward low-productivity regions, reducing regional inequality at the cost of potential misallocation. Future progressivity reduces the returns to local government investment, shifting fiscal spending away from investment toward public goods and depressing long-run growth. While the model is primarily theoretical, it generates predictions consistent with salient features of the Chinese economy following increased redistribution.

A key advantage of this framework is its tractability, which enables the characterization of time-varying optimal policy. While raising progressivity to reduce inequality is desirable, political pressure from older generations to expand public goods constrains the central government's ability to do so. This has important policy implications in the context of demographic transition and population aging. The model can further be extended to incorporate additional economic forces such as migration, agglomeration, and financial frictions in regional development.

That said, the tractability comes at a cost. The assumptions of logarithmic utility and full capital depreciation are critical for closed-form solutions and aggregation. Log utility implies a constant shadow value of capital, abstracting from the role of capital scarcity in shaping optimal progressivity over the development process. Full depreciation induces rapid convergence, limiting the realism of transitional dynamics. Future research could relax these assumptions to capture dynamics of optimal redistribution in richer environments.

References

Acharya, Sushant, and Keshav Dogra. 2020. "Understanding HANK: Insights from a PRANK." *Econometrica*, 88(3): 1113–1158.

Acharya, Sushant, Edouard Challe, and Keshav Dogra. 2023. "Optimal monetary policy according to HANK." *American Economic Review*, 113(7): 1741–1782.

Acikgoz, Omer, Marcus Hagedorn, Hans Aasnes Holter, and Yikai Wang. 2023. "The optimum quantity of capital and debt." *Available at SSRN 3185088*.

Aiyagari, S Rao. 1995. "Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting." *Journal of political Economy*, 103(6): 1158–1175.

Arrow, Kenneth J, and Mordecai Kruz. 2013. *Public investment, the rate of return, and optimal fiscal policy.* RFF Press.

Barro, Robert J. 1990. "Government spending in a simple model of endogeneous growth." *Journal of political economy*, 98(5, Part 2): S103–S125.

Barro, Robert J, and Xavier Sala-i Martin. 1992. "Public finance in models of economic growth." *The Review of Economic Studies*, 59(4): 645–661.

Baxter, Marianne, and Robert G King. 1993. "Fiscal policy in general equilibrium." *The American Economic Review*, 315–334.

Becker, Sascha O, Peter H Egger, and Maximilian Von Ehrlich. 2010. "Going NUTS: The effect of EU Structural Funds on regional performance." *Journal of Public Economics*, 94(9-10): 578–590.

Becker, Sascha O, Peter H Egger, and Maximilian Von Ehrlich. 2013. "Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects." *American Economic Journal: Economic Policy*, 5(4): 29–77.

Benigno, Gianluca, Huigang Chen, Christopher Otrok, Alessandro Rebucci, and Eric R Young. 2016. "Optimal capital controls and real exchange rate policies: A pecuniary externality perspective." *Journal of monetary economics*, 84: 147–165.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2025. "Minimum Wages, Efficiency, and Welfare." *Econometrica*, 93(1): 265–301.

Bhattarai, Saroj, Gauti B Eggertsson, and Bulat Gafarov. 2023. "Time consistency and duration of government debt: A model of quantitative easing." *The Review of Economic Studies*, 90(4): 1759–1799.

Bianchi, Javier, and Enrique G Mendoza. 2018. "Optimal time-consistent macroprudential policy." *Journal of Political Economy*, 126(2): 588–634.

Boehm, Christoph E. 2020. "Government consumption and investment: Does the composition of purchases affect the multiplier?" *Journal of Monetary Economics*, 115: 80–93.

Bouakez, Hafedh, Michel Guillard, and Jordan Roulleau-Pasdeloup. 2017. "Public investment, time to build, and the zero lower bound." *Review of Economic Dynamics*, 23: 60–79.

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor market distortions across time, space and sectors in China." *Review of economic dynamics*, 16(1): 39–58.

Buettner, Thiess. 2006. "The incentive effect of fiscal equalization transfers on tax policy." *Journal of Public Economics*, 90(3): 477–497.

Chari, Varadarajan Venkata, Alessandro Dovis, and Patrick J Kehoe. 2020. "On the optimality of financial repression." *Journal of Political Economy*, 128(2): 710–739.

Clemens, Jeffrey, and Stan Veuger. 2023. "Intergovernmental grants and policy competition: concepts, institutions, and evidence." National Bureau of Economic Research.

Debortoli, Davide, Ricardo Nunes, and Pierre Yared. 2017. "Optimal time-consistent government debt maturity." *The Quarterly Journal of Economics*, 132(1): 55–102.

Dyrda, Sebastian, and Marcelo Pedroni. 2023. "Optimal fiscal policy in a model with uninsurable idiosyncratic income risk." *The Review of Economic Studies*, 90(2): 744–780.

Fajgelbaum, Pablo D, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. 2019. "State taxes and spatial misallocation." *The Review of Economic Studies*, 86(1): 333–376.

Farhi, Emmanuel, and Ivan Werning. 2007. "Inequality and social discounting." *Journal of political economy*, 115(3): 365–402.

Fernández-Villaverde, Jesús. 2025. "Deep Learning for Solving Economic Models." National Bureau of Economic Research.

Heathcote, Jonathan, and Hitoshi Tsujiyama. 2021. "Optimal income taxation: Mirrlees meets Ramsey." *Journal of Political Economy*, 129(11): 3141–3184.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. "Optimal tax progressivity: An analytical framework." *The Quarterly Journal of Economics*, 132(4): 1693–1754.

Henderson, J Vernon, Dongling Su, Qinghua Zhang, and Siqi Zheng. 2022. "Political manipulation of urban land markets: Evidence from China." *Journal of Public Economics*, 214: 104730.

Klein, Paul, and José-Víctor Ríos-Rull. 2003. "Time-consistent optimal fiscal policy." *International Economic Review*, 44(4): 1217–1245.

Klein, Paul, Per Krusell, and Jose-Victor Rios-Rull. 2008. "Time-consistent public policy." *The Review of Economic Studies*, 75(3): 789–808.

Krusell, Per, and Jose-Victor Rios-Rull. 1999. "On the size of US government: political economy in the neoclassical growth model." *American Economic Review*, 89(5): 1156–1181.

Leeper, Eric M, Todd B Walker, and Shu-Chun S Yang. 2010. "Government investment and fiscal stimulus." *Journal of monetary Economics*, 57(8): 1000–1012.

Le Grand, François, Alaïs Martin-Baillon, and Xavier Ragot. forthcoming. "Should monetary policy care about redistribution? Optimal monetary and fiscal policy with heterogeneous agents." *The Review of Economic Studies.*

Le Grand, François, and Xavier Ragot. 2025. "Optimal Fiscal Policy with Heterogeneous Agents and Capital: Should We Increase or Decrease Public Debt and Capital Taxes?" *Journal of Political Economy*, 133(7): 2320–2369.

Marcet, Albert, and Ramon Marimon. 2019. "Recursive contracts." *Econometrica*, 87(5): 1589–1631.

Moser, Christian, and Pierre Yared. 2022. "Pandemic lockdown: The role of government commitment." *Review of Economic Dynamics*, 46: 27–50.

Müller, Andreas, Kjetil Storesletten, and Fabrizio Zilibotti. 2016. "The political color of fiscal responsibility." *Journal of the European Economic Association*, 14(1): 252–302.

Persson, Torsten, and Guido Tabellini. 2002. *Political economics: explaining economic policy.* MIT press.

Ramey, Valerie A. 2020. *The macroeconomic consequences of infrastructure investment.* Vol. 10, National Bureau of Economic Research Cambridge, MA.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The economic journal*, 37(145): 47–61.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin. 2020. "Nonlinear tax incidence and optimal taxation in general equilibrium." *Econometrica*, 88(2): 469–493.

Song, Zheng, and Wei Xiong. 2023. "The Mandarin model of growth."

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2012. "Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt." *Econometrica*, 80(6): 2785–2803.

Turnovsky, Stephen J. 2000. "Fiscal policy, elastic labor supply, and endogenous growth." *Journal of Monetary Economics*, 45(1): 185–210.

Yang, Lieyuan. 2025. "Growth with Regional Redistribution." working paper.

A Proof

This section provides all details and proofs in the main text.

A.1 Proof of Proposition 1

Before I prove Proposition 1, it is necessary to solve for firm and household problems. For the firm problem, plugging the production function (2) into the firm's problem (4) and taking first-order conditions with respect to k_i and n_i yields

$$\{k_i\}: \quad \theta(1-\tau_y)A_i k_{Gi}^{\alpha} k_i^{\theta-1} n_i^{1-\alpha-\theta} = (1+\tau_{ki})R$$
$$\{n_i\}: \quad (1-\alpha-\theta)(1-\tau_y)A_i k_{Gi}^{\alpha} k_i^{\theta} n_i^{-\alpha-\theta} = w_i$$

This can pin down the firm's labor-capital ratio with the interest rate and wage rate. With $n_i = 1$ we have

$$\frac{\theta}{1-\alpha-\theta}\frac{n_i}{k_i} = \frac{(1+\tau_{ki})R}{w_i} \quad \Rightarrow \quad k_i = \frac{\theta}{1-\alpha-\theta}\frac{w_i}{(1+\tau_{ki})R} \tag{A1}$$

Plugging (A1) back into firm's FOCs can solve for the local wage rate w_i

$$w_i = (1-\alpha-\theta)\left(\frac{\theta}{(1+\tau_{ki})R}\right)^{\frac{\theta}{1-\theta}}(1-\tau_y)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}}$$

This can further pin down the firm's capital, output, and profit

$$k_i = \left(\frac{\theta(1-\tau_y)}{R}\right)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}}$$
$$y_i = \left(\frac{\theta(1-\tau_y)}{R}\right)^{\frac{\theta}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}}$$
$$\pi_i = \alpha\left(\frac{\theta}{R}\right)^{\frac{\theta}{1-\theta}}(1-\tau_y)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}}$$

Also, it is easy to check that $w_{it} + \pi_{it} = (1-\theta)(1-\tau_t)y_{it}$

For the household problem, using c_{it+1}^{O} to express s_{it+1} and replacing s_{it+1} in the first-period budget constraint yields:

$$c_{it}^{Y}+\frac{c_{it+1}^{O}}{1+R}=(1-\theta)(1-\tau_t)y_{it}$$

First-order conditions with respect to c_{it}^{Y} and c_{it+1}^{O} yield:

$$\frac{1}{c_{it}^{Y}}=\frac{\beta(1+R)}{c_{it+1}^{O}}$$

Plugging this into the budget constraint yields the allocation in Equation (7).

For the central government budget constraint, when output is assumed to be log-normally distributed with the distribution $\log y_{it}\sim\mathcal{N}(\mu_{y_t},\sigma_{y_t}^2)$, Euqation (10) becomes:

$$\tau_t\exp\left(\mu_{y_t}+\frac{\sigma_{y_t}^2}{2}\right)=\lambda_t\tau_t^{1-\phi_t}\exp\left[(1-\phi_t)\mu_{y_t}+\frac{(1-\phi_t)^2\sigma_{y_t}^2}{2}\right]$$

Solving for λ_t and taking logs yield Equation (11). This assumption of log-normal distribution of output is reasonable if the distribution of the initial state is log-normal.

Now, let prime denote the variable in the next period. Guess the optimal choice of local government as $k'_{Gi}=\kappa'\lambda(\tau y_i)^{1-\phi}$, $g_i=(1-\kappa')\lambda(\tau y_i)^{1-\phi}$, and $g'_i=(1-\kappa'')\lambda'(\tau' y'_i)^{1-\phi'}$. Then,

$$\frac{\partial\mathcal{G}(A'_i,k'_{Gi})}{\partial k'_{Gi}}=\frac{\alpha(1-\phi')}{1-\theta}\frac{(1-\kappa'')\lambda'(\tau' y'_i)^{1-\phi'}}{k'_{Gi}}$$

Therefore, the GEE (12) can be rewritten as

$$\begin{aligned}\frac{\omega}{(1-\kappa')\lambda(\tau y_i)^{1-\phi}}&=\mathbb{E}\left[\frac{(1-\omega)\beta}{(1-\kappa'')\lambda'(\tau' y'_i)^{1-\phi'}}\frac{\alpha(1-\phi')}{1-\theta}\frac{(1-\kappa'')\lambda'(\tau' y'_i)^{1-\phi'}}{k'_{Gi}}\middle| A\right]\\&=\frac{(1-\omega)\beta}{\kappa'\lambda(\tau y_i)^{1-\phi}}\frac{\alpha(1-\phi')}{1-\theta}\end{aligned}$$

which gives the solution for κ'.

A.2 Proof of Proposition 2

Because (1) the distribution of initial state is log-normal, (2) output is log-linear in current public capital, and (3) the law of motion of public capital is log-linear given by Equation (13), so the distribution of public capital and output in every period is log-normal. If $\log y_{it} \sim \mathcal{N}(\mu_{y_t}, \sigma^2_{y_t})$, then

$$\log Y_t = \mu_{y_t} + \frac{\sigma^2_{y_t}}{2} \tag{A2}$$

Let $\mu_{k_{Gt}} = \mathbb{E}[\log k_{Git}]$, $\sigma^2_{k_{Gt}} = \mathrm{var}(\log k_{Git})$, $\sigma^2_A = \sigma^2_a + \sigma^2_\varepsilon$ and $\mu_A = -\sigma^2_A/2$. From Equation (13):

$$\begin{aligned}
\mu_{k_{Gt+1}} &= \mu_{y_t} + \frac{\sigma^2_{y_t}}{2} - \frac{(1-\phi_t)^2\sigma^2_{y_t}}{2} + \log\kappa(\phi_{t+1}) + \log\tau_t \\
&= \frac{1}{1-\theta}\mu_A + \frac{\alpha}{1-\theta}\mu_{k_{Gt}} + \frac{\theta}{1-\theta}\log\left(\frac{\theta(1-\tau_t)}{R}\right) \\
&\quad + \frac{1-(1-\phi_t)^2}{2}\left(\frac{1}{(1-\theta)^2}\sigma^2_A + \frac{\alpha^2}{(1-\theta)^2}\sigma^2_{k_{Gt}} + \frac{2\alpha}{(1-\theta)^2}\sigma_{Ak_G,t}\right) + \log\kappa(\phi_{t+1}) + \log\tau_t \\
\sigma^2_{k_{Gt+1}} &= (1-\phi_t)^2\sigma^2_{y_t} = (1-\phi_t)^2\left(\frac{1}{(1-\theta)^2}\sigma^2_A + \frac{\alpha^2}{(1-\theta)^2}\sigma^2_{k_{Gt}} + \frac{2\alpha}{(1-\theta)^2}\sigma_{Ak_G,t}\right) \\
\sigma_{Ak_G,t+1} &= \underbrace{\frac{1-\phi_t}{1-\theta}\sigma^2_a}_{C_0(\phi_t)} + \underbrace{\frac{\alpha(1-\phi_t)}{1-\theta}}_{C_3(\phi_t)}\sigma_{Ak_G,t}
\end{aligned}$$

Now, we track the evolution of the economy through the distribution of public capital k_G. From Equation (5), the distribution and the evolution of output can be characterized as:

$$\begin{aligned}
\mu_{y_t} &= \frac{1}{1-\theta}\mu_A + \frac{\alpha}{1-\theta}\mu_{k_{Gt}} + \frac{\theta}{1-\theta}\log\left(\frac{\theta(1-\tau_t)}{R}\right) \\
\sigma^2_{y_t} &= \frac{1}{(1-\theta)^2}\sigma^2_A + \frac{\alpha^2}{(1-\theta)^2}\sigma^2_{k_{Gt}} + \frac{2\alpha}{(1-\theta)^2}\sigma_{Ak_G,t} \\
\mu_{y_{t+1}} &= \frac{1}{1-\theta}\mu_A + \frac{\alpha}{1-\theta}\left(\mu_{y_t} + \frac{\sigma^2_{y_t}}{2} - \frac{(1-\phi_t)^2\sigma^2_{y_t}}{2} + \log\kappa(\phi_{t+1}) + \log\tau_t\right) + \frac{\theta}{1-\theta}\log\left(\frac{\theta(1-\tau_{t+1})}{R}\right) \\
\sigma^2_{y_{t+1}} &= \underbrace{\frac{\alpha^2(1-\phi_t)^2}{(1-\theta)^2}}_{B_2(\phi_t)}\sigma^2_{y_t} + \underbrace{\frac{2\alpha^2(1-\phi_t)}{(1-\theta)^3}}_{B_3(\phi_t)}\sigma_{Ak_G,t} + \underbrace{\frac{\sigma^2_A}{(1-\theta)^2} + \frac{2\alpha(1-\phi_t)}{(1-\theta)^3}\sigma^2_a}_{B_0(\phi_t)}
\end{aligned} \tag{A3}$$

Using Equation (A2) to replace expressions of Equation (A3) yields:

$$\log Y_{t+1} = \underbrace{\frac{\alpha}{1-\theta}}_{A_1} \log Y_t \underbrace{-\frac{\alpha(1-\theta-\alpha)}{(1-\theta)^2}\frac{(1-\phi_t)^2}{2}}_{A_2(\phi_t)} \sigma_{y_t}^2 + \underbrace{\frac{\alpha^2(1-\phi_t)}{(1-\theta)^3}}_{A_3(\phi_t)} \sigma_{Ak_G,t}$$
$$+ \underbrace{\frac{\alpha}{1-\theta}(\log \kappa(\phi_{t+1}) + \log \tau_t) + \frac{\theta}{1-\theta}\log\left(\frac{\theta(1-\tau_{t+1})}{R}\right) + \frac{\theta}{(1-\theta)^2}\frac{\sigma_A^2}{2} + \frac{\alpha(1-\phi_t)\sigma_a^2}{(1-\theta)^3}}_{A_0(\phi_{t+1},\phi_t,\tau_t,\tau_{t+1})}$$

This completes the proof of Proposition 2.

A.3 Proof of Proposition 3

Taking the integral of Equation (5) over all regions yields:

$$\log K_{Gt+1} = \log Y_t + \log \kappa(\phi_{t+1}) + \log \tau_t$$

Using the definition of aggregate TFP yields

$$\log \mathrm{TFP}_t = \frac{\theta}{1-\theta}\frac{\sigma_A^2}{2} - \frac{\alpha(1-\alpha-\theta)}{1-\theta}\frac{\sigma_{k_{Gt}}^2}{2} + \frac{\alpha}{1-\theta}\sigma_{Ak_G,t}$$

Using $\sigma_{k_{Gt+1}}^2 = (1-\phi_t)^2\sigma_{y_t}^2$ and the law of motion of σ_{Ak_G}, we have:

$$\begin{aligned}
\sigma_{k_{Gt+1}}^2 - \sigma_{k_{Gt}}^2 &= (1-\phi_t)^2\left(\frac{1}{(1-\theta)^2}\sigma_A^2 + \frac{\alpha^2}{(1-\theta)^2}\sigma_{k_{Gt}}^2 + \frac{2\alpha}{(1-\theta)^2}\sigma_{Ak_G,t}\right) - \sigma_{k_{Gt}}^2 \\
&= \frac{(1-\phi)^2}{(1-\theta)^2}\sigma_A^2 + \frac{2\alpha(1-\phi_t)^2}{(1-\theta)^2}\sigma_{Ak_G,t} + \frac{\alpha^2(1-\phi_t)^2-(1-\theta)^2}{(1-\theta)^2}\sigma_{k_{Gt}}^2 \\
\sigma_{Ak_G,t+1} - \sigma_{Ak_G,t} &= \frac{1-\phi}{1-\theta}\sigma_a^2 + \frac{\alpha(1-\phi)}{1-\theta}\sigma_{Ak_G,t} - \sigma_{Ak_G,t} \\
&= \frac{1-\phi_t}{1-\theta}\sigma_a^2 - \frac{1-\theta-\alpha(1-\phi_t)}{1-\theta}\sigma_{Ak_G,t}
\end{aligned}$$

Therefore, the growth in log aggregate TFP can be expressed as:

$$\begin{aligned}
\log \frac{\mathrm{TFP}_{t+1}}{\mathrm{TFP}_t} &= -\frac{\alpha(1-\alpha-\theta)}{2(1-\theta)}\left[\frac{(1-\phi_t)^2}{(1-\theta)^2}\sigma_A^2 + \frac{2\alpha(1-\phi_t)^2}{(1-\theta)^2}\sigma_{Ak_G,t} + \frac{\alpha^2(1-\phi_t)^2-(1-\theta)^2}{(1-\theta)^2}\sigma_{k_{Gt}}^2\right] \\
&+ \frac{\alpha}{1-\theta}\left[\frac{1-\phi}{1-\theta}\sigma_a^2 - \frac{1-\theta-\alpha(1-\phi_t)}{1-\theta}\sigma_{Ak_G,t}\right] \\
&= \frac{\alpha(1-\theta+\alpha)}{(1-\theta)^3}\frac{\sigma_a^2}{2} - \frac{\alpha(1-\alpha-\theta)}{(1-\theta)^3}\frac{\sigma_\varepsilon^2}{2} \\
&- \frac{\alpha(1-\theta-\alpha)(1-\theta+\alpha)}{(1-\theta)^3}\sigma_{Ak_G,t} + \frac{\alpha(1-\alpha-\theta)^2(1-\theta+\alpha)}{(1-\theta)^3}\frac{\sigma_{k_{Gt}}^2}{2} \\
&+ \frac{\alpha(1-\theta-\alpha)[1-(1-\phi_t)^2]}{(1-\theta)^3}\frac{\sigma_\varepsilon^2}{2} - \frac{\alpha[2\alpha\phi_t+(1-\theta-\alpha)\phi_t^2]}{(1-\theta)^3}\frac{\sigma_a^2}{2} \\
&+ \frac{\alpha^3(1-\alpha-\theta)[1-(1-\phi_t)^2]}{(1-\theta)^3}\frac{\sigma_{k_{Gt}}^2}{2} \\
&+ \frac{\alpha^2}{(1-\theta)^3}\left[(1-\theta-\alpha)(1-(1-\phi_t)^2)-(1-\theta)\phi\right]\sigma_{Ak_G,t} \\
&= \frac{\alpha(1-\theta+\alpha)}{(1-\theta)^3}\frac{\sigma_a^2}{2} - \frac{\alpha(1-\alpha-\theta)}{(1-\theta)^3}\frac{\sigma_\varepsilon^2}{2} \\
&- \frac{\alpha(1-\theta-\alpha)(1-\theta+\alpha)}{(1-\theta)^3}\sigma_{Ak_G,t} + \frac{\alpha(1-\alpha-\theta)^2(1-\theta+\alpha)}{(1-\theta)^3}\frac{\sigma_{k_{Gt}}^2}{2} \\
&+ \frac{\alpha(1-\theta-\alpha)[1-(1-\phi_t)^2]}{(1-\theta)}\frac{\sigma_y^2}{2} - \frac{\alpha\phi_t}{(1-\theta)^2}(\sigma_a^2+\alpha\sigma_{Ak_G,t})
\end{aligned}$$

The second equality follows:

$$\sigma_{k_{Gt}}^2 = \frac{(1-\theta)^2}{\alpha^2}\sigma_{y_t}^2 - \frac{1}{\alpha^2}\sigma_A^2 - \frac{2}{\alpha}\sigma_{Ak_G,t}$$

The last equality follows:

$$\begin{aligned}
-\frac{\alpha[2\alpha\phi+(1-\theta-\alpha)\phi_t^2]}{(1-\theta)^3}\frac{\sigma_a^2}{2} &= \frac{\alpha[-2\alpha\phi_t-(1-\theta-\alpha)\phi_t^2]}{(1-\theta)^3}\frac{\sigma_a^2}{2} \\
&= \frac{\alpha[2(1-\theta-\alpha)\phi_t-(1-\theta-\alpha)\phi_t^2-2(1-\theta)\phi_t]}{(1-\theta)^3}\frac{\sigma_a^2}{2} \\
&= \frac{\alpha[(1-\theta-\alpha)(1-(1-\phi_t)^2)-2(1-\theta)\phi_t]}{(1-\theta)^3}\frac{\sigma_a^2}{2}
\end{aligned}$$

and

$$\sigma_{y_t}^2 = \frac{\sigma_a^2+\sigma_\varepsilon^2}{(1-\theta)^2} + \frac{\alpha^2}{(1-\theta)^2}\sigma_{k_{Gt}}^2 + \frac{2\alpha}{(1-\theta)^2}\sigma_{Ak_G,t}$$

This completes the proof of Proposition 3.

A.4 Proof of Proposition 4

For any $\phi \in (0,1)$ and $\tau \in (0,1)$, the steady state exists if $\alpha/(1-\theta) < 1$. When $\phi_t \to \phi$ and $\tau_t \to \tau$ as $t \to \infty$, imposing the steady-state condition yields:

$$
\begin{aligned}
\mu_y &= \frac{1}{1-\theta-\alpha}\left[\mu_A + \theta \log\left(\frac{\theta(1-\tau)}{R}\right) + \alpha\left(\log\kappa(\phi) + \log\tau + \frac{1-(1-\phi)^2}{2}\sigma_y^2\right)\right] \\
\sigma_y^2 &= \frac{\sigma_\varepsilon^2}{(1-\theta)^2 - \alpha^2(1-\phi)^2} + \frac{\sigma_a^2}{(1-\theta-\alpha(1-\phi))^2} \\
\mu_{k_G} &= \frac{1}{1-\theta-\alpha}\left[\mu_A + \theta \log\left(\frac{\theta(1-\tau)}{R}\right) + (1-\theta)\left(\log\kappa(\phi) + \log\tau + \frac{1-(1-\phi)^2}{2}\sigma_y^2\right)\right] \\
\sigma_{k_G}^2 &= \frac{(1-\phi)^2\sigma_\varepsilon^2}{(1-\theta)^2 - \alpha^2(1-\phi)^2} + \frac{(1-\phi)^2\sigma_a^2}{(1-\theta-\alpha(1-\phi))^2} \\
\sigma_{Ak_G} &= \frac{1-\phi}{1-\theta-\alpha(1-\phi)}\sigma_a^2
\end{aligned}
\tag{A4}
$$

Again using the property of log-normal distribution yields aggregate quantities in the steady state:

$$
\begin{aligned}
\log Y &= \frac{1}{1-\theta-\alpha}\left[\mu_A + \theta\log\left(\frac{\theta(1-\tau)}{R}\right) + \alpha(\log\kappa + \log\tau) + \frac{1-\theta-\alpha(1-\phi)^2}{2}\sigma_y^2\right] \\
\log K_G &= \frac{1}{1-\theta-\alpha}\left[\mu_A + \theta\log\left(\frac{\theta(1-\tau)}{R}\right) + (1-\theta)(\log\kappa + \log\tau) + \frac{1-\theta-\alpha(1-\phi)^2}{2}\sigma_y^2\right] \\
\log K &= \frac{1}{1-\theta-\alpha}\left[\mu_A + (1-\alpha)\log\left(\frac{\theta(1-\tau)}{R}\right) + \alpha(\log\kappa + \log\tau) + \frac{1-\theta-\alpha(1-\phi)^2}{2}\sigma_y^2\right]
\end{aligned}
$$

$$
\begin{aligned}
\log \mathrm{TFP} &= \mu_A + (1-\theta-\alpha(1-\phi)^2)\frac{\sigma_y^2}{2} \\
&= \underbrace{\frac{\theta+\alpha}{1-\theta-\alpha}\frac{\sigma_a^2}{2} + \frac{\theta-\alpha}{1-\theta+\alpha}\frac{\sigma_\varepsilon^2}{2}}_{\text{TFP in non-redistribution economy}} \\
&\quad \underbrace{-\frac{\alpha(1-\theta)\phi^2}{(1-\theta-\alpha)(1-\theta-\alpha(1-\phi))^2}\frac{\sigma_a^2}{2}}_{\text{Distortionary effects}} + \underbrace{\frac{\alpha(1-\theta)[1-(1-\phi)^2]}{(1-\theta+\alpha)[(1-\theta)^2-\alpha^2(1-\phi)^2]}\frac{\sigma_\varepsilon^2}{2}}_{\text{Insurance effects}}
\end{aligned}
$$

A.5 Detials on Planning Problem

In period 0, the central government announces and is committed to the policy $(\{\phi_t\}_{t=0}^{\infty}, \{\tau_t\}_{t=0}^{\infty})$. Define the social welfare as:

$$
\begin{aligned}
&\mathcal{W}(\{\phi_t\}_{t=1}^{\infty}, \{\tau_t\}_{t=1}^{\infty}) \\
&= \beta \int U_t^O(y_{i,-1}, g_{i0})di + \sum_{t=0}^{\infty} \psi^{t+1} \int U_t^Y(y_{it}, g_{it+1})di \\
&= \sum_{t=0}^{\infty} \psi^{t+1} \int (1+\beta)\log y_{it} + \chi\frac{\beta}{\psi}\log g_{it} + (1+\beta)\log(1-\tau_t)di \\
&= \sum_{t=0}^{\infty} \psi^{t+1} \left[(1+\beta)\left(\log Y_t - \frac{\sigma_{y_t}^2}{2} + \log(1-\tau_t)\right) + \chi\frac{\beta}{\psi}\left(\log G_t - \frac{(1-\phi_t)^2\sigma_{y_t}^2}{2}\right)\right] \\
&= \underbrace{(1+\beta+\chi\beta/\psi)}_{\Omega_1} \sum_{t=0}^{\infty} \psi^{t+1} \log Y_t \\
&- \sum_{t=0}^{\infty} \psi^{t+1} \underbrace{(1+\beta+(1-\phi_t)^2\chi\beta/\psi)}_{\Omega_2(\phi_t)} \frac{\sigma_{y_t}^2}{2} \\
&+ \sum_{t=0}^{\infty} \psi^{t+1} \underbrace{[(1+\beta)\log(1-\tau_t) + \chi\beta/\psi(\log(1-\kappa(\phi_{t+1})) + \log\tau_t)]}_{\Omega_3(\phi_{t+1},\tau_t)}
\end{aligned}
$$

The state in period -1 and 0 is taken as given. The fourth equality follows from the fact that $G_t = (1-\kappa(\phi_{t+1}))\tau_t Y_t$ and $G_t = \int g_{it}di$.

Define the Lagrangian multipliers as $\psi^{t+1}M_t^Y$, $\psi^{t+1}M_t^{\sigma_y}$, $\psi^{t+1}M_t^{\sigma_{AkG}}$. Then, the optimality conditions are:

$$
\{\tau_t\}: \quad -\psi\frac{1+\beta}{1-\tau_t} + \frac{\beta\chi}{\tau_t} + \psi M_t^Y \frac{\alpha}{1-\theta}\frac{1}{\tau_t} - M_{t-1}^Y \frac{\theta}{1-\theta}\frac{1}{1-\tau_t} = 0 \tag{A5}
$$

$$
\begin{aligned}
\{\phi_t\}: \quad & \beta\chi(1-\phi_t)\sigma_{y_t}^2 \\
& + \psi M_t^Y \left(\frac{\alpha(1-\theta-\alpha)(1-\phi_t)}{(1-\theta)^2}\sigma_{y_t}^2 - \frac{\alpha^2}{(1-\theta)^3}\sigma_{Ak_G,t} - \frac{\alpha}{(1-\theta)^3}\sigma_a^2\right) \\
& - \psi M_t^{\sigma_y} \left(\frac{2\alpha^2(1-\phi_t)}{(1-\theta)^2}\sigma_{y_t}^2 + \frac{2\alpha^2}{(1-\theta)^3}\sigma_{Ak_G,t} + \frac{2\alpha}{(1-\theta)^3}\sigma_a^2\right) \\
& - \chi\beta/\psi\frac{\kappa'(\phi_t)}{1-\kappa(\phi_t)} - \psi M_t^{\sigma_{Ak_G}} \frac{1}{1-\theta}(\sigma_a^2 + \alpha\sigma_{Ak_G,t}) + M_{t-1}^Y \frac{\alpha}{1-\theta}\frac{\kappa'(\phi_t)}{\kappa(\phi_t)} = 0
\end{aligned} \tag{A6}
$$

$$
\{\log Y_t\}: \quad \psi(1+\beta+\chi\beta/\psi) + \frac{\alpha}{1-\theta}\psi M_t^Y - M_{t-1}^Y = 0 \tag{A7}
$$

$$
\begin{aligned}
\{\sigma^2_{yt}\}: \quad & -\psi(1+\beta+\chi\beta/\psi(1-\phi_t)^2)\frac{1}{2} - \psi M_t^Y \frac{\alpha(1-\theta-\alpha)}{(1-\theta)^2}\frac{(1-\phi_t)^2}{2} \\
& + \psi M_t^{\sigma_y}\frac{\alpha^2(1-\phi_t)^2}{(1-\theta)^2} - M_{t-1}^{\sigma_y} = 0
\end{aligned} \tag{A8}
$$

$$
\{\sigma_{A_t k_{Gt}}\}: \quad \psi M_t^Y \frac{\alpha^2(1-\phi_t)}{(1-\theta)^3} + \psi M_t^{\sigma_{Ak_G}}\frac{\alpha(1-\phi_t)}{1-\theta} - M_{t-1}^{\sigma_{Ak_G}} + \psi\frac{2\alpha^2(1-\phi_t)}{(1-\theta)^3}M_t^{\sigma_y} = 0 \tag{A9}
$$

From Equation (A7), $M_t^Y = M_{t-1}^Y$: the shadow value of future log output is constant. Otherwise M_t^Y will explode as $t \to \infty$. This pins down M_t^Y, and, with Equation (A5), pins down τ_t^*:

$$
M_t^Y = M^{Y*} = \frac{(1-\theta)\psi(1+\beta+\chi\beta/\psi)}{1-\theta-\alpha\psi} \tag{A10}
$$

$$
\tau_t^* = \tau^* = \frac{(1-\theta)\beta\chi + \psi^2\alpha(1+\beta)}{\beta\chi+\psi(1+\beta)}
$$

Two useful expressions:

$$
\begin{aligned}
\frac{\kappa'(\phi)}{\kappa(\phi)} &= -\frac{1}{1-\phi} + \frac{\alpha(1-\omega)\beta}{(1-\theta)\omega + \alpha(1-\phi)(1-\omega)\beta} \\
&= -\frac{(1-\theta)\omega}{(1-\phi)[(1-\theta)\omega + \alpha(1-\phi)(1-\omega)\beta]} \\
\frac{\kappa'(\phi)}{1-\kappa(\phi)} &= -\frac{\alpha(1-\omega)\beta}{(1-\theta)\omega + \alpha(1-\phi)(1-\omega)\beta}
\end{aligned} \tag{A11}
$$

Then, in a representative region model, imposing $\sigma_a^2 = \sigma_\varepsilon^2 = \sigma_y^2 = 0$ yields the optimality condition for ϕ:

$$
-\chi\beta/\psi\frac{\kappa'(\phi_t)}{1-\kappa(\phi_t)} + M_{t-1}^Y\frac{\alpha}{1-\theta}\frac{\kappa'(\phi_t)}{\kappa(\phi_t)} = 0
$$

Plugging the Equations (A10) and (A11) into the optimality condition yields

$$
1-\phi^{R*} = \frac{\omega}{1-\omega}\frac{\psi(1+\beta)+\chi\beta}{\beta^2}\frac{(1-\theta)\psi}{\chi(1-\theta-\alpha\psi)}
$$

B TFP Process with Persistent Shock

This section provides an extension where the TFP process includes a persistent shock ζ_{it} and a transitory shock ε_{it}

$$\log A_{it} = \log \zeta_{it} + \log \varepsilon_{it}$$

where the persistent shock ζ_{it} follows an AR(1) process:

$$\log \zeta_{it} = \rho \log \zeta_{it-1} + \log \xi_{it}$$

with the distribution $\log \xi_{it} \sim \mathcal{N}(-\sigma_{\xi}^2/2, \sigma_{\xi}^2)$ and $\log \varepsilon_{it} \sim \mathcal{N}(-\sigma_{\varepsilon}^2/2, \sigma_{\varepsilon}^2)$.

In this environment, we should track the distribution of ζ and k_G:

$$\mu_{\zeta_{t+1}} = \rho \mu_{\zeta} - \frac{\sigma_{\xi}^2}{2}$$
$$\sigma_{\zeta_{t+1}}^2 = \rho^2 \sigma_{\zeta_t}^2 + \sigma_{\xi}^2$$

Then, the distribution of productivity is given by:

$$\mu_{A_{t+1}} = \mu_{\zeta_{t+1}} - \frac{\sigma_{\varepsilon}^2}{2} = \rho \mu_{\zeta} - \frac{\sigma_{\varepsilon}^2 + \sigma_{\xi}^2}{2}$$
$$\sigma_{A_{t+1}}^2 = \sigma_{\zeta_{t+1}}^2 + \sigma_{\varepsilon}^2 = \rho^2 \sigma_{\zeta_t}^2 + \sigma_{\xi}^2 + \sigma_{\varepsilon}^2$$

The mean and the variance of k_G is structurally unchanged, but the covariance between A_{it+1} and k_{Git+1} is given as:

$$\sigma_{Ak_G,t+1} = \frac{\rho(1-\phi)}{1-\theta} \sigma_{\zeta_t}^2 + \frac{\rho(1-\phi)\alpha}{1-\theta} \sigma_{Ak_G,t}$$

In the steady state, the distribution aof the persistent component and the productivity is:

$$\mu_\zeta = -\frac{\sigma_\xi^2}{2(1-\rho)}$$
$$\sigma_\zeta^2 = \frac{\sigma_\xi^2}{1-\rho^2}$$
$$\mu_A = -\frac{\sigma_\xi^2}{2(1-\rho)} - \frac{\sigma_\varepsilon^2}{2}$$
$$\sigma_A^2 = \frac{\sigma_\xi^2}{1-\rho^2} + \sigma_\varepsilon^2$$

Other distribution parameters are given by:

$$\sigma_{Ak_G} = \frac{\rho(1-\phi)}{1-\theta-\rho\alpha(1-\phi)}\sigma_\zeta^2 = \frac{\rho(1-\phi)}{1-\theta-\rho\alpha(1-\phi)}\frac{\sigma_\xi^2}{1-\rho^2}$$
$$\sigma_y^2 = \frac{1}{(1-\theta)^2-\alpha^2(1-\phi)^2}\left[\frac{1-\theta+\rho\alpha(1-\phi)}{1-\theta-\rho\alpha(1-\phi)}\frac{\sigma_\xi^2}{1-\rho^2} + \sigma_\varepsilon^2\right]$$

The model still remain tractable.

C Local Governments Care Future Generations

If local governments care about future generations, the local government problem is rewritte as:

$$\max_{g_i, k_{Gi}} \quad \beta U_i^O(y_{i-1}, g_{i0}) + \sum_{t=0}^{\infty} \psi^{t+1} U_i^Y(y_{it}, g_{it+1})$$
$$\text{s.t.} \quad g_{it} + k_{Git+1} = \lambda_t(\tau_t y_{it})^{1-\phi_t}$$

Define the policy function as $k'_G = h(k_G)$. The Bellman Equation:

$$v(k_G) = \frac{(1+\beta)\alpha}{1-\theta}\log k_G + \frac{\beta\chi}{\psi}\log g + \psi v(h(k_G))$$
$$\text{s.t.} \quad g + h(k_G) = \lambda(\tau y)^{1-\phi}$$

The first-order condition with respect to k'_G yields:

$$\frac{\frac{\beta\chi}{\psi}}{\lambda(\tau y)^{1-\phi}} = \psi v'(h(k_G))$$

The first-oder condition with respect to k_G (envelope condition) yields:

$$
\begin{aligned}
v'(k_G) &= \frac{(1+\beta)\alpha}{1-\theta}\frac{1}{k_G} + \frac{\frac{\beta\chi}{\psi}[(1-\phi)\frac{\alpha}{1-\theta}\lambda(\tau y)^{1-\phi}\frac{1}{k_G} - h'(k_G)]}{\lambda(\tau y)^{1-\phi} - h(k_G)} + \psi v'(h(k_G))h'(k_G) \\
&= \frac{(1+\beta)\alpha}{1-\theta}\frac{1}{k_G} + \frac{\frac{\beta\chi}{\psi}}{\lambda(\tau y)^{1-\phi} - h(k_G)}\frac{\alpha(1-\phi)}{1-\theta}\frac{\lambda(\tau y)^{1-\phi}}{k_G}
\end{aligned}
$$

Plugging the envelope condition into the first-order condition yields the Generalized Euler Equation (GEE):

$$
\frac{\frac{\beta\chi}{\psi}}{\lambda(\tau y)^{1-\phi} - h(k_G)} = \frac{\psi(1+\beta)\alpha}{1-\theta}\frac{1}{h(k_G)} + \frac{(\beta+\psi)\chi}{\lambda'(\tau' y')^{1-\phi'} - h(h(k_G))}\frac{\alpha(1-\phi')}{1-\theta}\frac{\lambda'(\tau' y')^{1-\phi'}}{h(k_G)}
$$

Guess: $k_G' = \tilde{\kappa}'\lambda(\tau y)^{1-\phi}$, which gives:

$$
\begin{aligned}
\frac{\frac{\beta\chi}{\psi}}{(1-\tilde{\kappa}')\lambda(\tau y)^{1-\phi}} &= \frac{\alpha\psi(1+\beta)}{1-\theta}\frac{1}{\tilde{\kappa}'\lambda(\tau y)^{1-\phi}} + \frac{\lambda'(\tau' y')^{1-\phi'}}{(1-\tilde{\kappa}'')\lambda'(\tau' y')^{1-\phi'}}\frac{(\beta+\psi)\chi\alpha(1-\phi')}{1-\theta}\frac{1}{\tilde{\kappa}'\lambda(\tau y)^{1-\phi}} \\
&= \frac{\alpha\psi(1+\beta)}{1-\theta}\frac{1}{\tilde{\kappa}'\lambda(\tau y)^{1-\phi}} + \frac{1}{(1-\tilde{\kappa}'')}\frac{(\beta+\psi)\chi\alpha(1-\phi')}{1-\theta}\frac{1}{\tilde{\kappa}'\lambda(\tau y)^{1-\phi}}
\end{aligned}
$$

The GEE is simplified to:

$$
\frac{\frac{\beta\chi}{\psi}}{1-\tilde{\kappa}'} = \frac{\alpha}{1-\theta}\frac{1}{\tilde{\kappa}'}\left(\psi(1+\beta) + \frac{(\beta+\psi)\chi(1-\phi')}{1-\tilde{\kappa}''}\right)
$$

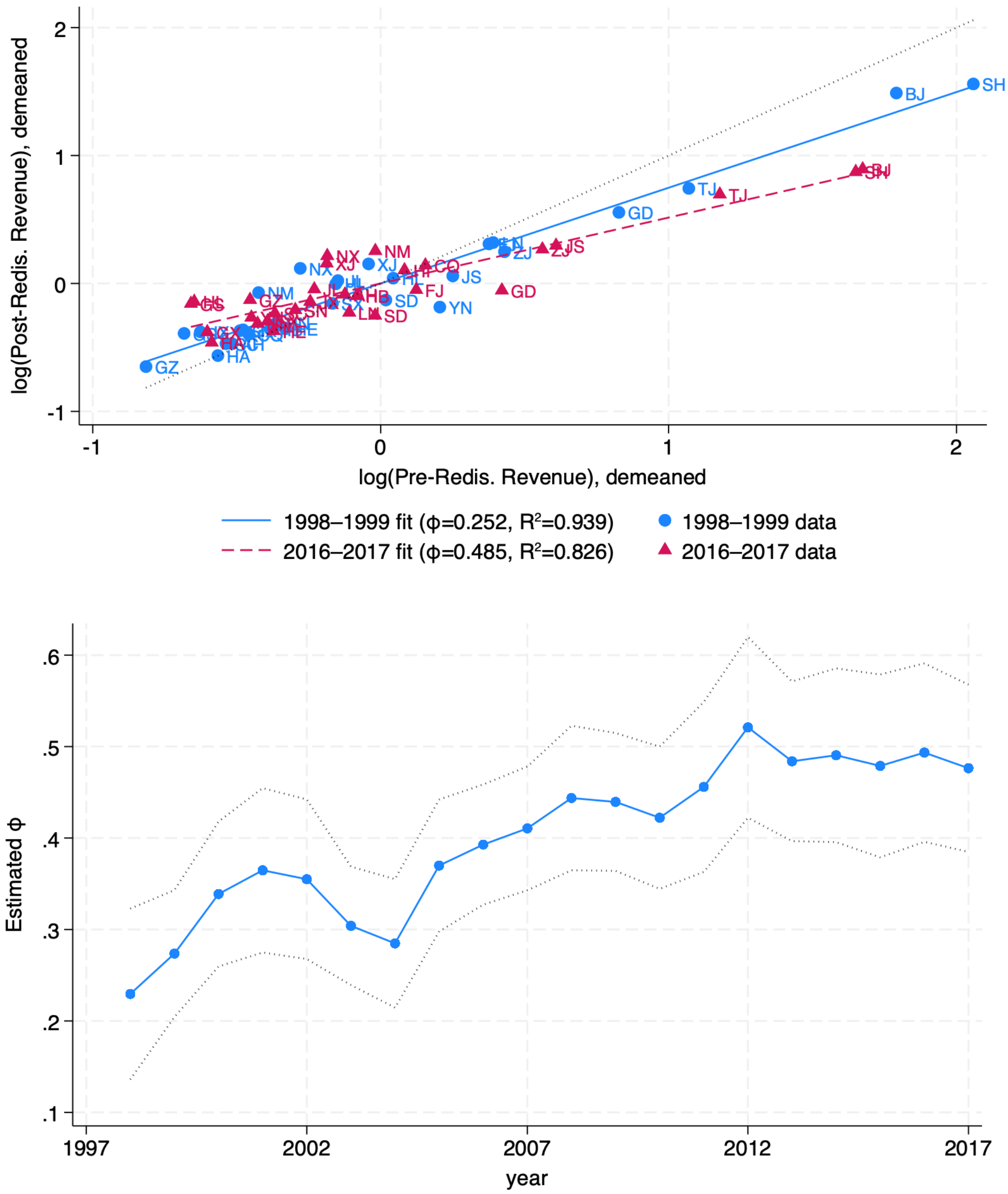
Therefore, $\tilde{\kappa}'$ is recursively given by:

$$
\begin{aligned}
\frac{\tilde{\kappa}'}{1-\tilde{\kappa}'} &= \frac{\alpha\psi}{(1-\theta)\beta\chi}\left(\psi(1+\beta) + \frac{(\beta+\psi)\chi(1-\phi')}{1-\tilde{\kappa}''}\right) \\
&= \frac{\alpha\psi^2(1+\beta)}{(1-\theta)\beta\chi} + \frac{\alpha\psi(1-\phi')}{(1-\tilde{\kappa}'')(1-\theta)}
\end{aligned}
$$

The model still maintain tractability.

D Additional Results

Figure D1: Representation of the Regional Redistribution System through the HSV Tax Function and Rising Progressivity



Note: This figure are from Yang (2025). The top panel plots the log of post-redistribution revenue per capita against the log of pre-redistribution revenue per capita across different periods. To ensure annual comparability, both pre and post-redistribution are demeaned. The bottom panel presents the OLS estimation of ϕ using Equation (1) for each year. The dotted line represents the 95% confidence interval. Tibet and Qinghai are excluded.

Figure D2: Approximation of ϕ_t as a function $\mathcal{P}(\sigma^2_{y_t}, \sigma_{Ak_G,t})$, Neural Network Approximation

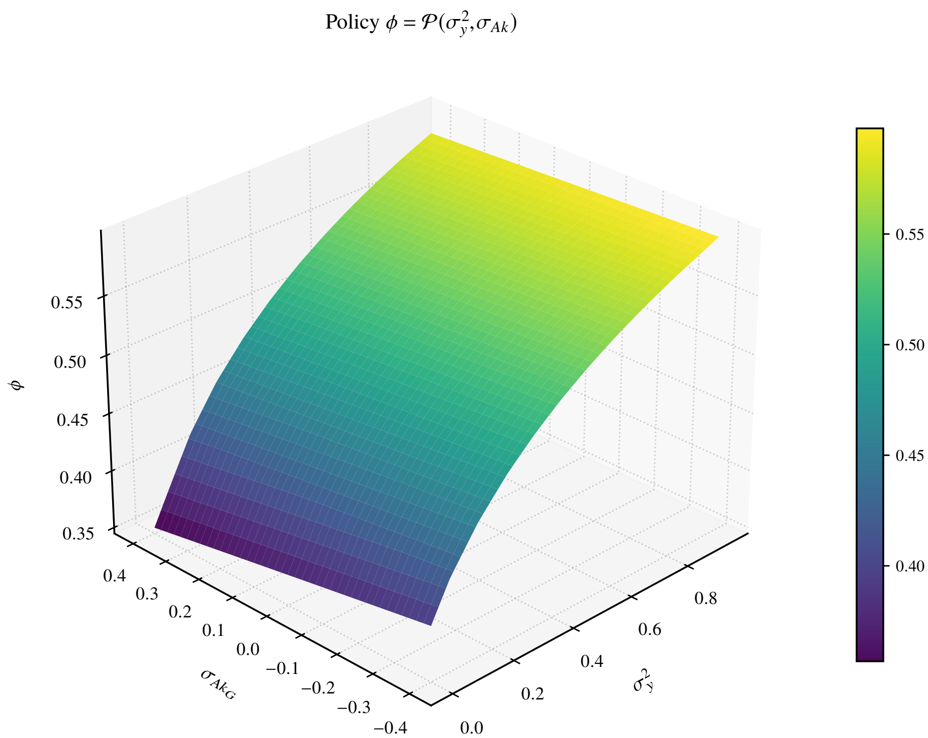


Figure D3: The Path of Optimal Policy and Aggregate Dynamics, Neural Network Approximation

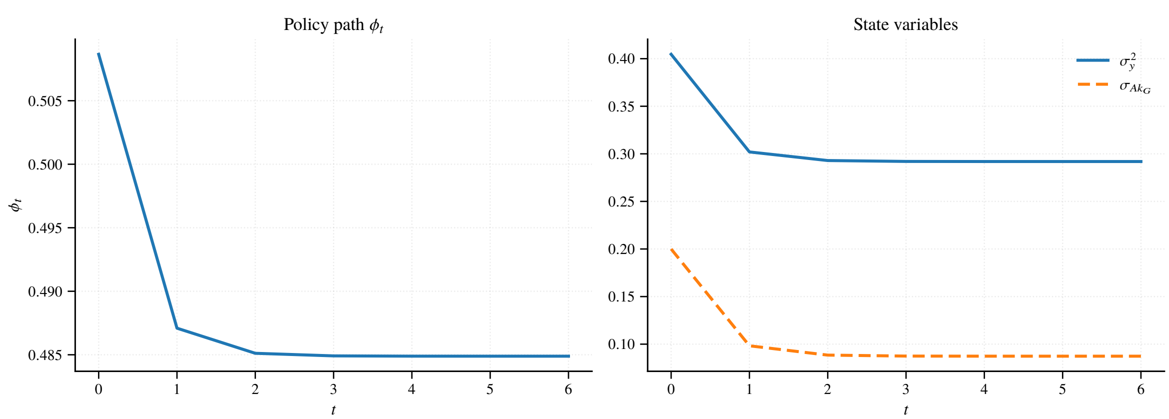


Figure D4: Path of Optimal Progressivity Under Different Parameter Values, Starting with Lower Output dispersion