Regional Redistribution

Lieyuan Yang *

Department of Economics, University of Minnesota

September 13, 2025

Abstract

This paper studies redistribution and regional economic growth in China. The central government collects roughly half of fiscal revenue to subsidize undeveloped regions. I model regional redistribution as a progressive tax on local fiscal revenue and document a substantial rise in progressivity from around 0.2 in 1998 to 0.5 in 2017. Exploiting the 2002 tax reform as a natural experiment, I conduct an event study and show that the central-local tax-sharing rule shifts local fiscal spending away from infrastructure investment and slows regional growth. I develop an analytically tractable general equilibrium model of local fiscal policy under the progressive redistribution scheme. Quantitatively, rising progressivity largely compresses regional disparities and stimulates economic convergence at a modest efficiency cost. The current redistribution system is nearly utilitarian optimal.

Key words: Redistribution, Tax Progressivity, Ramsey Taxation, Fiscal Policy, Regional Economic Growth, Capital Misallocation, Efficiency, Social Welfare, Chinese Economy.

JEL codes: E60, H20, H50, H70, O11, O23, O40, R13, R50

*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. I also thank Anmol Bhandari, Kyle Herkenhoff, Loukas Karabarbounis, Hannes Malmberg, Simon Mongey, and Todd Schoellman for helpful comments.

1 Introduction

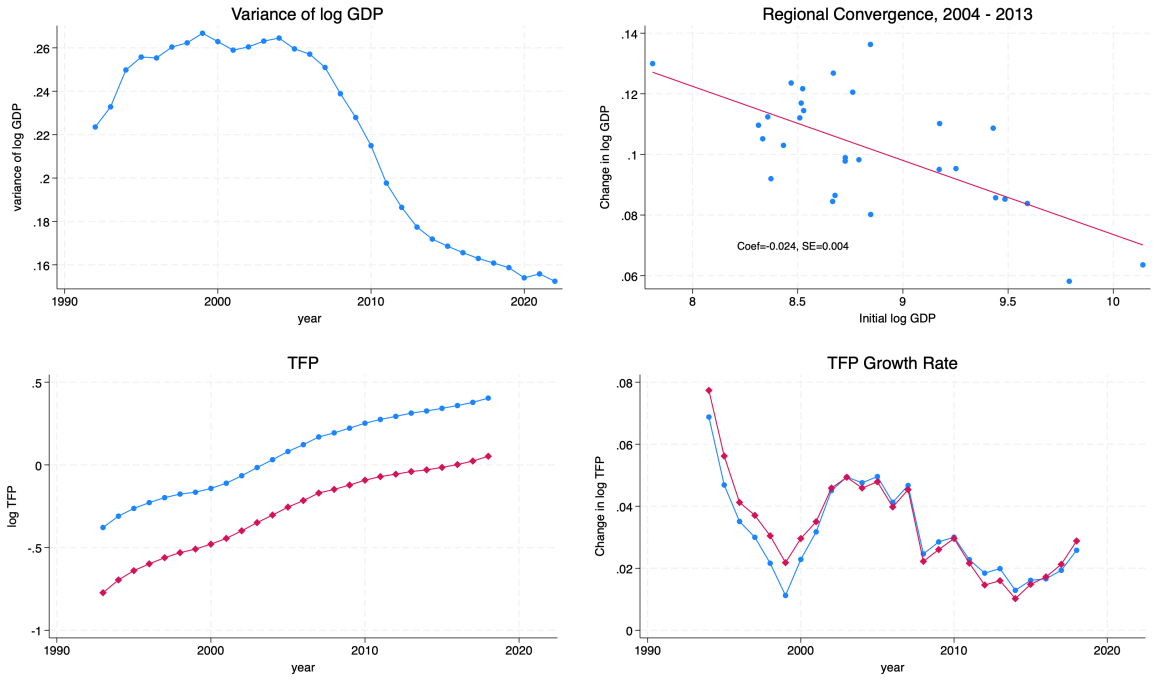
China has implemented one of the largest inter-governmental and inter-regional fiscal redistribution systems in the world. Following the 1994 tax-sharing reform, the central government began collecting 75% of value-added taxes. In 2002, it started to reallocate 50% of personal and corporate income taxes from local to central government, with a commitment to transfer most of these revenues to less-developed regions. By 2010, central taxes amounted to roughly 10% of GDP, 80% of which was redistributed to local governments. More importantly, per-capita transfers to poor and middle-income provinces were nearly twice those to rich provinces.

The large-scale redistribution has important implications for regional economic dynamics. Transfers to lagging regions can accelerate their growth and promote regional economic convergence. At the same time, channeling resources to low-productivity regions may exacerbate capital misallocation and dampen aggregate efficiency. Consistent with this tension, Figure 1 shows two striking patterns following the 2002 reform: a rapid decline in the dispersion of provincial GDP per capita and a sharp slowdown in aggregate total factor productivity (TFP) growth.

Motivated by these observations, this paper has two main goals. First, I bring rich empirical evidence to illustrate the question of how China's regional redistribution system evolves and how it influences local fiscal policy and regional economic growth. Specifically, I model the redistribution as a progressive tax-transfer scheme across local governments and document a sharp increase in progressivity following the 2002 reform. The increasingly progressive scheme reallocates resources towards less-developed regions and shifts local fiscal spending away from infrastructure investment toward government consumption. Second, I develop an analytically tractable general equilibrium model of local fiscal policy subject to the progressive redistribution scheme designed by the central government. The model aims to quantify its effects on output dispersion and aggregate productivity and to determine the optimal degree of progressivity.

My work contributes to the broad literature on capital misallocation and economic growth. Pioneered by Chari, Kehoe and McGrattan (2007), Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), a growing body of literature identifies the sources of distortions and examines their effects on economic growth, especially in China. Most

Figure 1: Regional Economic Convergence and Aggregate TFP Growth Slowdown



Note: Panels display the declining variance of log provincial GDP per capita (top-left), the β-convergence regression (top-right), measured TFP (bottom-left), and the TFP growth rate (bottom-right). In the bottom panels, circles plot $\log(\text{GDP}) - 0.185\log(\text{Infrastructure Capital}) - 0.315\log(\text{Non-infrastructure Capital})$; diamonds plot $\log(\text{GDP}) - 0.5\log(\text{Capital})$.

seminal contributions, such as Song, Storesletten and Zilibotti (2011) and David and Venkateswaran (2019), mainly focus on cross-sectoral or firm-level misallocation. My work contributes by identifying regional redistribution as a main source of distortions driving cross-regional capital allocation and by quantifying its effects on regional convergence and aggregate productivity.

This paper also contributes to the Ramsey tradition of optimal tax design (Ramsey, 1927). Improving social welfare while mitigating distortions is the central concern of redistribution policy. Following Heathcote, Storesletten and Violante (2017) (HSV henceforth), I model regional redistribution as a progressive tax on local fiscal revenue. While the previous literature has a thorough understanding of optimal income taxation, I extend the analysis to taxes and transfers across subnational governments and study the determinants of optimal regional redistribution policy.

To track the evolution of the regional redistribution system, I begin my analysis by estimating the degree of progressivity using the widely applied HSV tax function: $T(x) = x - \lambda x^{1-\phi}$, where $T(x)$ is the net central tax, x is pre-redistribution local fiscal revenue, ϕ governs the degree of progressivity, and λ determines the average tax level. A higher ϕ implies a more progressive redistribution scheme, with the marginal tax rate increasing in pre-redistribution fiscal revenue. I show that the HSV function closely fits China's regional redistribution pattern. Estimated progressivity rises substantially from around 0.2 in 1998 to 0.5 in 2017, raising the average marginal tax rate from 28% to 40%.

Alongside the rise in progressivity, several stylized facts emerge. Redistribution compresses the gap in fiscal resources between rich and poor provinces, but also distorts local fiscal policy. Local governments allocate spending between government consumption and infrastructure investment, with the latter being crucial for long-run growth and future tax revenue. As the central government claims a larger portion of taxes, the return to investment falls. In response, local governments shift spending from infrastructure toward government consumption, particularly in rich provinces facing higher marginal tax rates. As a result, infrastructure capital grows faster in less-developed provinces, resulting in regional convergence and capital misallocation.

To build the causal links among regional redistribution, local fiscal policy, and re-

gional economic growth, I exploit the 2002 tax reform as a natural experiment and conduct an event study. The reform reduced fiscal returns to infrastructure investment by allocating 50% of personal and corporate income taxes to the central government. Provinces more reliant on income taxes were more adversely affected and responded by cutting more public investment. I use the pre-reform share of income taxes in total tax revenue to measure local governments' exposure to the reform and find that high-share provinces experience slower post-reform growth in infrastructure capital and output. This pattern is consistent with the interpretation that the central-local tax-sharing rule reduces local governments' incentives to invest and slows regional growth.

Based on the empirical findings, I develop an analytically tractable general equilibrium model of local fiscal policy under the progressive redistribution scheme. The model embeds the HSV tax function into the mandarin model of Song and Xiong (2023), in which local governments actively invest in local infrastructure to boost economic growth. I extend their framework to analyze how local governments adjust fiscal policy in response to the redistribution rule. I also connect my framework to Hsieh and Klenow (2009) to assess how regional redistribution influences regional convergence and aggregate efficiency. In the baseline environment, regions are heterogeneous in productivity. Households derive utility from private consumption and public goods. Local governments, under the progressive redistribution scheme, purchase public goods and rent public capital to maximize local household utility. Public goods directly enter household utility, whereas public capital raises output and the equilibrium wage. Given the local governments' best response, the central government chooses the optimal degree of progressivity to maximize household welfare across regions.

To determine the implications of rising progressivity for local fiscal policy and regional economic growth, it is pivotal to estimate the output elasticity of infrastructure capital. Most prior work infers this elasticity from the infrastructure capital share, which implicitly assumes that capital allocation is socially optimal. However, in my framework, local fiscal policy is distorted by fiscal capacity and redistribution policy. To causally identify the output elasticity, I use the pre-2002 share of income taxes as an instrumental variable (IV). The logic is that the division of tax revenue between central and local governments influences regional economic performance exclusively through

local fiscal policy. It does not affect households' or firms' decisions directly as long as the statutory tax rates are unchanged. The IV estimate of the output elasticity is 0.185, exceeding the value inferred from capital shares and the OLS estimate, indicating substantial distortions in local fiscal policy and underinvestment in infrastructure.

To quantify the effects of progressivity on output dispersion and aggregate productivity, I treat the period from 2013 to 2017 as a new steady state when progressivity becomes stable. The model matches most cross-sectional moments in this steady state, as well as the post-reform transitional dynamics. In the new steady state, relative to the non-redistribution case, progressivity reduces the variance of log output from 0.26 to 0.19, while raising the efficiency loss from 2.7% to 3.6%. During the transition, eliminating the rise in progressivity would lower the annual convergence rate from 2.7% to 2.1%, diminish by about one third of the reduction in output dispersion, and reduce the decline in TFP relative to its efficient level from 1.7 to 1.3 percentage points.

The quantitative analysis highlights two competing forces in designing redistribution policy: reducing regional inequality and limiting the efficiency loss. The quantitative results indicate that the rising progressivity largely reduces regional disparities at a modest efficiency cost. The baseline model further implies that the utilitarian social planner will set the optimal progressivity at 0.575. The current progressivity, 0.485, is almost optimal: moving from the current system to the optimal one achieves negligible welfare gains.

This paper builds on and contributes to several strands of literature. First, my work is closely related to literature that employs the HSV form progressive tax function. The most relevant work is Heathcote, Storesletten and Violante (2017) that develops an analytically tractable framework to study economic forces shaping optimal progressivity. The most recent studies using the HSV tax function include Fajgelbaum et al. (2019); Sachs, Tsyvinski and Werquin (2020); Guerreiro, Rebelo and Teles (2022); Boar and Midrigan (2022); Ferriere et al. (2023); Borella et al. (2023); Auclert, Rognlie and Straub (2024); Berger, Herkenhoff and Mongey (2025); Fleck et al. (2025). Most of the work, however, only studies the progressive income tax. I contribute to the literature by using the HSV form tax function to analyze taxes on local governments and redistribution across regions. To my best knowledge, this paper is the first to study the progressivity

of inter-government transfers and inter-regional redistribution.

My work relates to the broader literature on capital misallocation and aggregate efficiency. Following Chari, Kehoe and McGrattan (2007), Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), a growing body of literature identifies the sources of distortions and examines their effects on aggregate efficiency, particularly in the context of China. For example, Brandt and Zhu (2010) and Song, Storesletten and Zilibotti (2011) underscore capital misallocation between state and non-state sectors. David and Venkateswaran (2019) identifies firm-level productivity-correlated distortions as the main source of misallocation. Adamopoulos et al. (2022) and Adamopoulos et al. (2024) analyze land market frictions for aggregate agricultural productivity. Bai, Jin and Lu (2024) studies misallocation under trade liberalization. Villamil, Wang and Xue (2021), Song and Xiong (2023) and Cheremukhin et al. (2024) discuss distortions introduced by political institutions.

There are two articles that particularly focus on the cross-regional factor allocation and aggregate efficiency in China. Brandt, Tombe and Zhu (2013) quantifies how market factor distortions shape cross-regional capital allocation. They argue that the effects of cross-regional misallocation on the TFP loss are modest compared to cross-sectoral misallocation. However, their analysis is restricted to the period from 1985 to 2007, before the substantial rise in progressivity reshaped the cross-regional distribution of fiscal resources. Tombe and Zhu (2019) documents cross-regional dispersion in output per worker and emphasizes the effects of goods- and labor-market frictions on aggregate productivity. My work contributes by identifying regional redistribution as a main source of distortions driving cross-regional capital allocation and by quantifying its effects on aggregate productivity.

Moreover, my work is also in line with literature on economic convergence. Prior work has documented convergence patterns across countries and regions (Barro, 1991; Mankiw, Romer and Weil, 1992; Barro and Sala-i Martin, 1992; Gennaioli et al., 2014; Jones, 2016) and examined factors contributing to convergence, such as human capital accumulation, the development of capital markets, technology diffusion, and structural transformation (Barro, 1996; Hall and Jones, 1999; Caselli and Coleman II, 2001; Hanushek and Woessmann, 2008; Rodrik, 2013; Kremer, Willis and You, 2022). Of

particular relevance for China, Hao et al. (2020) finds that reductions in the migration cost are central to regional income convergence in China between 2000 and 2015. Brandt, Kambourov and Storesletten (2025) identifies the change in entry barriers as a main source driving regional convergence in non-SOE manufacturing sectors between 1995 and 2008. While China's miraculous growth has received significant attention, its rapid regional convergence since the mid-2000s has been largely overlooked. My work contributes by documenting the regional convergence pattern and offering novel evidence on the role of regional redistribution in driving economic convergence.

The rest of the paper is organized as follows. Section 2 introduces the institutional background of China's tax-sharing system and presents some stylized facts. Section 3 conducts an event study to identify the causal effects of the tax-sharing rule on local fiscal incentives and regional economic growth. Based on the empirical evidence, Section 4 introduces the economic environment, Section 5 characterizes regional allocations and aggregate quantities in the equilibrium, and Section 6 formulates the closed-form social welfare function and explores the determinants of optimal progressivity. Section 7 discusses the model parameterization, with special attention to estimating the output elasticity. Section 8 discusses the model evaluation and the quantitative implications. Section 9 calculates the optimal degree of progressivity in various model scenarios. Section 10 concludes.

2 Institutional Background and Stylized Facts

This section first outlines institutional features of a series of tax reforms and the tax-sharing system. To track the evolution of the redistribution system, I model regional redistribution as a progressive tax on local fiscal revenue and estimate its progressivity. I then present time-series evidence on how the reform affected local governments' fiscal policy and regional economic growth, which motivates the theoretical framework.

2.1 Tax-Sharing Reform

The 1994 tax-sharing reform is one of the most significant policy changes in contemporary China. It has profoundly reshaped fiscal structures, supported the transition from

a planned economy to a market economy, and influenced land and housing markets, local government debts, shadow banking, and central-local relationships (Liu and Jia, 2008).

The initial objective of the tax reform was to mitigate the central government's fiscal deficit. Prior to 1994, the sharing of tax revenue between the central and local governments followed complex and region-specific rules. Some provinces paid a fixed share of fiscal revenue to the central government, such as Tianjin (46.5%), Shanxi (87.55%), and Anhui (77.5%). Others made lump-sum payments, like Shanghai (10.5 billion RMB), Shandong (0.3 billion), and Heilongjiang (0.29 billion). Some provinces paid annually increased lump-sum transfers. Guangdong paid 1.4 billion in 1988 with a 9% annual increase, and Hunan paid 0.8 billion in 1988 with a 7% annual increase. Other provinces paid an increasing proportional tax. For example, Beijing paid half of any increase in fiscal revenue above 4% in fiscal revenue to the central government (Zhou, 2012).

The pre-1994 system had two major drawbacks. First, it imposed uneven burdens across provinces. For instance, Shanghai faced disproportionately heavy obligations, limiting its ability to formulate sustainable fiscal plans. Second, central fiscal revenue grew more slowly than GDP, constraining its ability to manage the fiscal deficit. Figure 2 plots central and local fiscal revenue in the share of GDP. Before the reform, local fiscal revenue exceeded 10% of GDP, whereas central taxes were around 5%.

To address financial shortages, the central government implemented the 1994 tax-sharing reform. The reform addressed two key issues: (1) the classification of major tax categories in line with market-oriented reform, and (2) the share of fiscal revenue between the central and local governments. On the tax structure, the 1994 reform redefined major tax categories. Turnover taxes were restructured around the value-added tax (VAT), supplemented by excise taxes and business taxes. The VAT is levied broadly on production and circulation stages. The excise tax applies only to selected goods (e.g., tobacco, alcohol, cosmetics, automobiles, and jewelry) and aims to curb luxury consumption. The business tax, by contrast, targets specific services, intangible asset transfers, and real estate transactions. The reform also unified the corporate income tax system. Previously, state-owned enterprises (SOEs) and private firms were

subject to separate tax regimes. After the reform, all firms were taxed under a unified statutory corporate income tax of 33%. A nationwide personal income tax was also introduced for the first time.

On the tax-sharing rule, central taxes included excise taxes, tariffs, and corporate income taxes from central SOEs, while local taxes included business tax, income taxes from local enterprises and individuals, and several minor tax categories. Value-added tax (VAT) revenue was shared, with 75% allocated to the central government and 25% to local governments. The excise tax was entirely assigned to the central government, given its high profitability and the desire to avoid incentivizing local governments to promote industries such as tobacco and alcohol (Liu and Jia, 2008; Zhou, 2012). Table 1 summarizes the key points of the tax reforms.

Figure 3 shows the evolution of major tax categories in both GDP and total tax shares. The value-added tax emerged as the dominant tax category, accounting for around 5% of GDP and 40% of total tax revenue during the reform period. After the 1994 reform, central and local governments shared nearly the same proportion of tax revenue. Despite significant fiscal centralization, local governments remained responsible for expenditures on local infrastructure, education, health, etc. To support these obligations, the central government increased transfers to local governments, as shown in Figure 2.

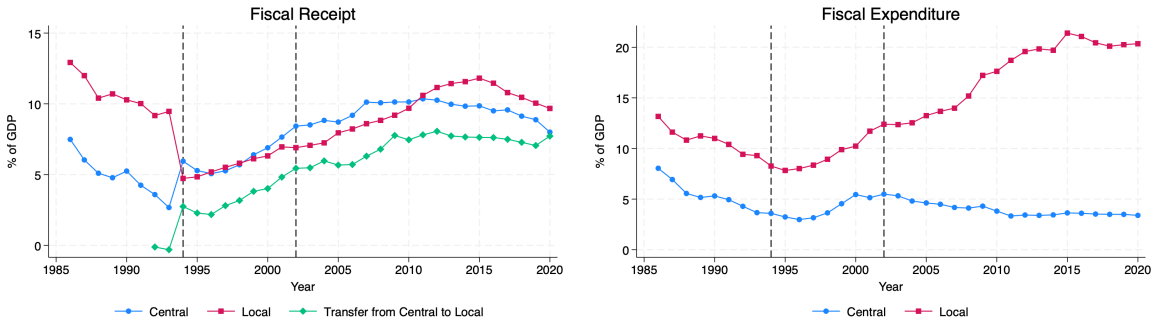
While the 1994 reform focused on strengthening central finances, the 2002 reform introduced a redistributive objective. In addition to receiving 75% of value-added taxes, the central government began collecting 50% of the corporate and personal income taxes in 2002, with the share rising to 60% in 2003. Crucially, all income taxes collected by the central government were designated for redistribution toward less developed regions. Figure 2 shows that transfers to local governments increased from roughly 2% of GDP in 1995 to over 5% in 2002. This significant rise implies that enhanced central revenues supported not only central expenditures but also greater redistribution. For instance, in 2000, Shanghai (China's wealthiest province) received 753 RMB per capita in transfers, while Gansu (one of the poorest) received 500 RMB. By 2010, Shanghai's transfers rose to 1,678 RMB per capita, while Gansu's reached 4,028 RMB. Notably, despite the growing central revenue share, local governments' expenditures continued

to rise, enabled by these expanding transfers.

The inclusion of income taxes in the tax-sharing system significantly strengthened the central government's fiscal capacity. As shown in Figure 3, income taxes are quantitatively important: corporate income taxes account for roughly 3% of GDP and 20% of total tax revenue, while personal income taxes contribute over 1% of GDP and slightly less than 10% of total revenue. Like the VAT following the 1994 reform, income taxes became a large and stable component of shared taxes. Their inclusion further enhanced the central government's ability to subsidize undeveloped regions, which substantially reshaped local fiscal conditions and central-local relations. The 2002 reform thus institutionalized a large-scale inter-governmental redistribution system, redirecting fiscal resources from rich to poor regions.

As noted in Table 1, an additional adjustment occurred in 2015, when the business tax was replaced by the VAT. Following the change, the VAT remained a shared tax, with 50% allocated to the central government. This change extended the logic of fiscal centralization initiated in 1994 and reinforced by the 2002 reform.

Figure 2: Fiscal Receipts and Expenditure for the Central and Local Governments and Net Transfer from the Central to Local Governments

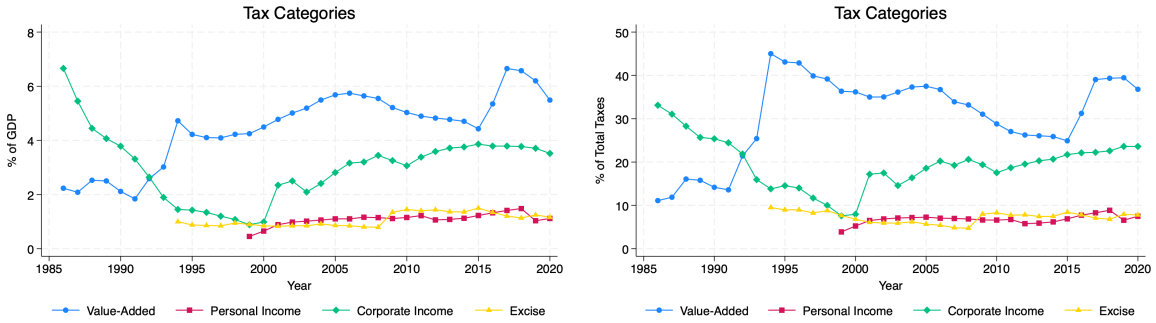


Note: The dashed lines marked the 1994 and the 2002 tax reforms. Data source: the NBS website.

2.2 Estimating Progressivity

To track the evolution of the redistribution system, I estimate progressivity using the widely applied HSV tax function. To align with the theoretical framework, I assume that the per capita net central tax is a function of per capita all fiscal revenues collected

Figure 3: Tax Categories Belong to or Shared by the Central Government



Note: Before 2000, corporate income taxes include only those from SOEs and collective enterprises; starting in 2001, they cover all firms. From 1985 to 1993, corporate income taxes also include SOE profits. Data source: the NBS website.

Table 1: Key Aspects of Each Tax Reform since 1994

Time	Key Points
1994	Central tax: excise tax, tariff, and taxes from central SOE. Local tax: business taxes, personal income taxes, and taxes from local firms. Shared tax: VAT (75% for central and 25% for local)
2002	The central government collects 50% of personal and corporate income taxes.
2003	The central government collects 60% of personal and corporate income taxes.
2015	Business tax was replaced by VAT. The central government collects 50% value-added tax.

within a province, including taxes, non-tax revenue, extra-budgetary revenue, and land sales. Appendix A.1 provides details on data and measurement. Appendix A.2 provides an alternative measurement approach where the net central tax is a function of GDP per capita, producing similar but more robust estimates.

Let x_i denote per capita pre-redistribution fiscal revenue. Let $\tilde{x}_i$ denote per capita post-redistribution local fiscal revenue, defined as pre-redistribution revenue x minus per capita central taxes plus per capita net transfers from the central government. Then, the per capita net taxes paid to the central government are defined as

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

where ϕ determines the degree of progressivity and λ sets the average central tax level. $1 - \phi$ defines the elasticity of post-redistribution fiscal revenue with respect to pre-redistribution fiscal revenue. Higher values of ϕ indicate greater progressivity. In two extreme cases, when $\phi = 0$, the net central tax is linear in the pre-redistribution revenue, and when $\phi = 1$, post-redistribution revenues are uniform across regions. It is easy to check that for $\phi > 0$, the marginal tax rate $T'(x)$ is increasing in x.

We can estimate the tax function given by Equation (1) using the following regression:

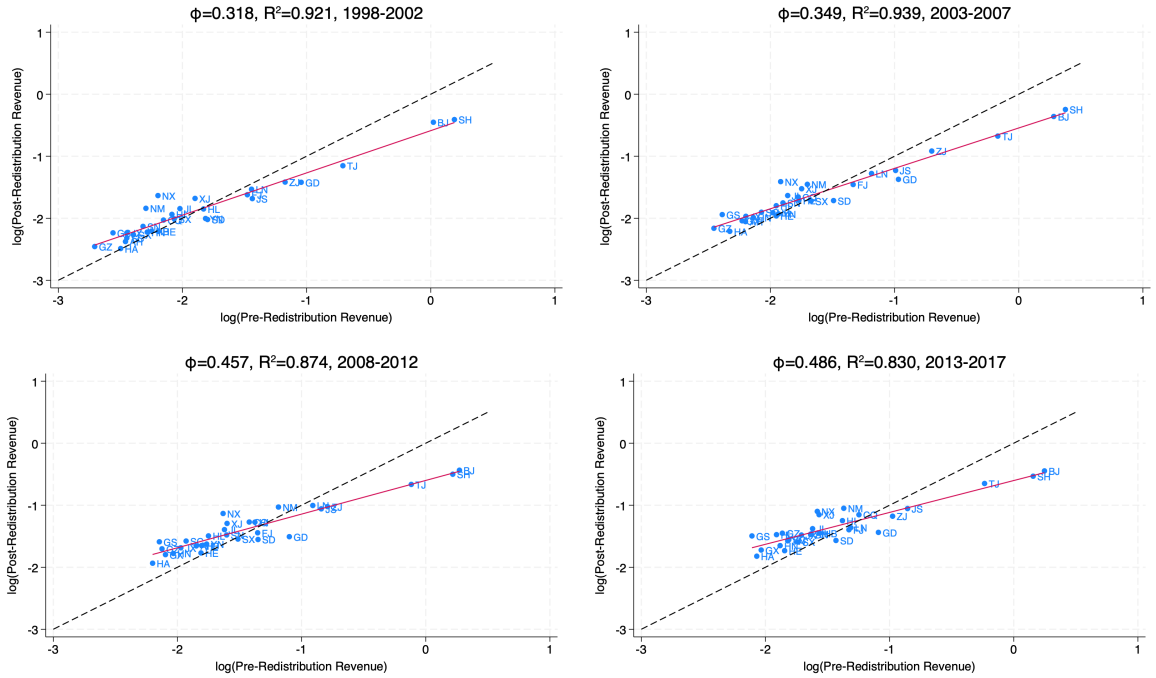
$$\log \tilde{x}_{it} = \log \lambda + (1 - \phi) \log x_{it} + \epsilon_{it} \tag{2}$$

where ϵ_i is the error term. Tibet and Qinghai are excluded from the estimation because they receive substantial ethnic-policy-related subsidies. When Tibet and Qinghai are included, the estimated ϕ is much higher. Pre- and post-redistribution revenues are normalized annually by dividing by the annual national GDP per capita to ensure comparability.

Figure 4 displays the scatter plots of the log of post-redistribution local fiscal revenue against the log of pre-redistribution fiscal revenue, along with the fitted regression line. The dashed 45-degree line indicates points where pre-redistribution revenue equals post-redistribution revenue. Points above (below) the 45-degree line represent provinces receiving net subsidies from (paying net taxes to) the central government. Over time, the fitted lines become flatter, clearly illustrating the increase in progressivity. Wealth-

ier provinces contribute more taxes to the central government, while poorer provinces receive greater subsidies. The R^2 is approximately 0.9, indicating a good fit of the tax function to the actual redistribution system.

Figure 4: The Scatter Plot of log Post-Redistribution Local Government Revenue Against Pre-Redistribution Local Government Revenue.



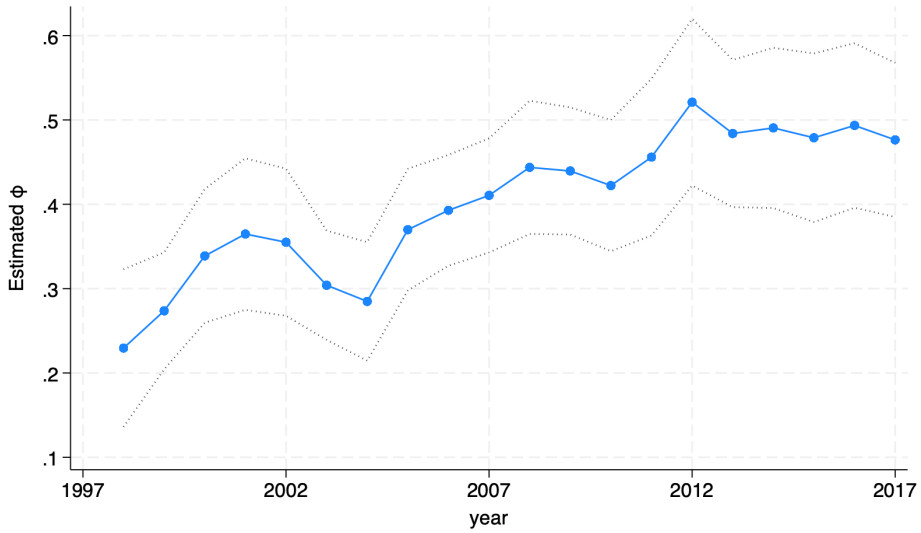
Note: These figures plot the log of post-redistribution revenue per capita against the log of pre-redistribution revenue per capita across different periods. To ensure annual comparability, both pre and post-redistribution are normalized by dividing by national GDP per capita and averaged within each period. Tibet and Qinghai are excluded.

I estimate progressivity using Equation (2) by year and plot the results in Figure 5. The dashed lines represent the 95% confidence interval. The estimated ϕ increased from 0.23 in 1998 to 0.48 in 2017, more than doubling over twenty years.

It is noticeable that progressivity did not rise immediately after the 2002 tax reform. To maintain fiscal stability, during the reform in 2002 and 2003, the central government temporarily increased transfers to the provinces that lost significant fiscal revenue due to the reform. Consequently, we observe a temporary decline in progressivity in these two years. Following the reform, the central government gradually increased (decreased) transfers to poorer (richer) provinces, resulting in a steady rise in progressivity after 2004. The progressivity becomes stable near 0.5 after 2013. Therefore, in the following quantitative analysis in Section 8, I treat the period from 2004 to 2013 as the transitional

dynamics and the period from 2013 to 2017 as the new steady state.

Figure 5: The Estimated Progressivity Over Time



Note: Each dot presents the OLS estimation of ϕ using Equation (2) for each year. The dotted line represents the 95% confidence interval. We excluded Tibet and Qinghai in our sample.

I provide an alternative estimation of progressivity in Appendix A.2, assuming that the net central tax is a function of GDP per capita. GDP per capita is a more robust measurement compared to pre-redistribution government revenue, as fiscal revenue includes volatile components such as non-tax revenues and land sales. This alternative approach reveals a similar but more stable pattern. Corresponding results of Figures 4 and 5 are presented in Figures A1 and A2. In Figure A1, the fitted line becomes increasingly flat over time and the R^2 is close to 1. In Figure A2, estimated progressivity increased from 0.05 in 1998 to 0.26 in 2017. Figure A2 also shows quite stable progressivity around the 2002 tax reform and the increase in progressivity starting from 2004.

2.3 Local Fiscal Policy

How does the change in progressivity over time affect local fiscal policy? This subsection presents some illustrative time-series evidence regarding local fiscal policy. Figure 6 displays the evolution of local fiscal revenue and spending per capita, normalized by dividing by annual national GDP per capita to ensure comparability across years.

When the central government redistributes fiscal revenue from richer to poorer provinces, the most direct observable outcome should be the reduced disparity in post-redistribution government revenues. The top panels of Figure 6 compare pre- and post-redistribution fiscal revenue over time. While the substantial differences between richer and poorer provinces persist in pre-redistribution revenues, the gap significantly narrows after redistribution, indicating enhanced fiscal equality across regions from redistribution. Figure A3 further shows these patterns expressed as a share of local GDP. Rich provinces initially exhibit higher ratios of fiscal revenue to local GDP due to greater tax capacity and more profitable industries. After redistribution, however, low-income regions receive transfers that result in higher revenue-to-GDP ratios than richer provinces.

How does the local government fiscal spending respond to the tax reform? I broadly classify local government spending into government consumption and infrastructure investment.[1] Local governments typically invest in infrastructure to promote regional economic growth, anticipating higher future tax revenues. However, increased progressivity raises the marginal tax rate and reduces the marginal return to investment, particularly for rich provinces. Consequently, we expect local governments in wealthier regions to allocate more fiscal resources to government consumption and less to investment.

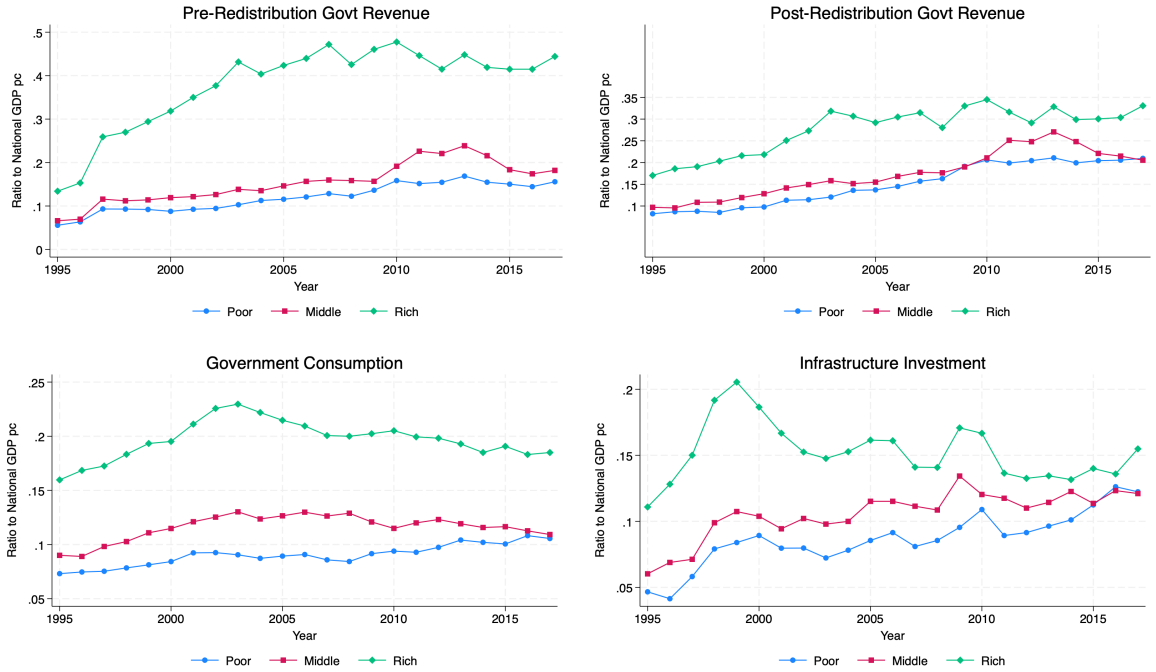
The bottom panel of Figure 6 displays the trends in local government consumption and infrastructure investment.[2] While significant differences in government consumption remain between richer and poorer provinces, the disparity in infrastructure investment decreases over time and nearly vanishes by the end of 2017. With increased progressivity, local governments in wealthier provinces face higher marginal tax rates

[1]There is no official data on Chinese government investment or capital outlay as defined in National Income and Product Accounts (NIPA). Following Song and Xiong (2023), I approximate government investment using infrastructure investment. Further details on data and measurement are provided in Appendix A.1.

[2]The sum of government consumption and infrastructure investment may not match exactly the post-redistribution government revenues. Possible reasons include: (1) Official NBS data does not separate central and local government consumption clearly, potentially overstating local government consumption. The data I use here is regional government consumption from regional GDP expenditure-approach table. (2) Infrastructure investment used here as a proxy for government investment may include investments by SOEs, potentially overestimating actual government investment. (3) Prior to 2009, although local governments were officially prohibited from borrowing or issuing debt, they launched local government financial vehicles (LGFVs) to finance infrastructures investment. The difference between fiscal revenue and calculated expenditure partly reflects government debt.

and thus prefer government consumption over infrastructure investment due to the reduced returns.

Figure 6: Local Government Fiscal Revenue and Spending per capita, Grouped by GDP per capita



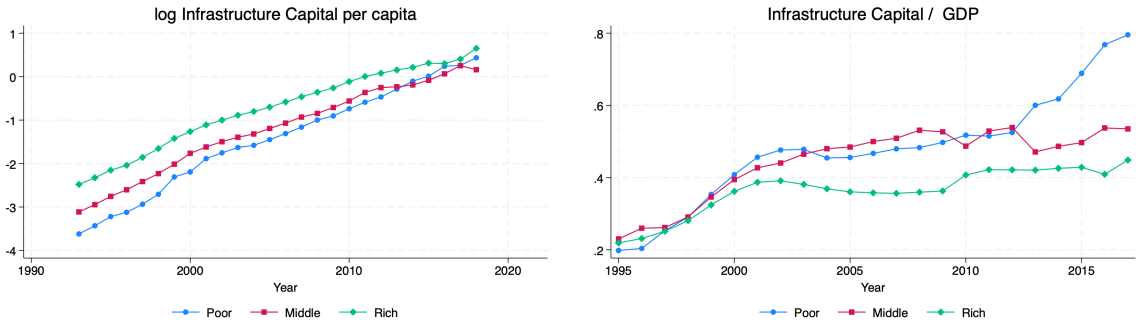
Note: These figures display the local government pre-redistribution revenue (top-left), post-redistribution revenue (top-right), government consumption (bottom-right), and infrastructure investment (bottom-right). Provinces are categorized into three groups based on GDP per capita: poorest (circles), middle (squares), and richest (diamonds). The y-axis represent the ratio to national GDP per capita.

2.4 Uneven Growth in Infrastructure Capital

Under the increasingly progressive redistribution scheme, local governments shifted spending away from infrastructure investment toward government consumption, especially in rich provinces. A direct consequence is uneven growth in infrastructure capital across regions. Figures 7 show the evolution of infrastructure capital, measured both in log per capita terms and as a share of GDP, grouped by income levels.

In the early 1990s, per capita infrastructure capital in rich provinces was roughly three times that in poor provinces. As high-income regions gradually reduced infrastructure investment, this gap narrowed over time and nearly disappeared by 2015. In the spirit of Hsieh and Klenow (2009), we can assess the efficiency of capital allocation

Figure 7: Infrastructure Capital: Log of Per Capita Level and Share of GDP



using the dispersion of capital-output ratios across regions. Initially, capital-output ratios were similar across income groups, suggesting an efficient allocation. Over time, however, poorer regions experienced rapid growth in capital-output ratios due to substantial transfers from the central government, while richer regions underinvested. This rising dispersion indicates growing capital misallocation.

2.5 Economic Consequences

What are the potential economic consequences of changes in local government fiscal policies? The shift in spending away from infrastructure investment, especially in high-income regions, can have important implications for both the cross-sectional and aggregate performance of the economy. Figures 1 present two salient empirical patterns: rapid regional economic convergence and a pronounced slowdown in aggregate TFP growth following the tax reform.

The top panels show that the variance of log provincial GDP per capita declined from 0.26 in 2005 to 0.16 by 2020. This convergence is so strong that we can observe σ-convergence. Using a Barro-style convergence regression (Barro, 1991), I estimate an annual convergence rate of 2.4% during the transitional period of 2004 to 2013, suggesting that it takes less than 30 years to reduce GDP differences by half for any two provinces.

The bottom panels plot the trajectory of aggregate TFP, measured as the Solow residual. Two alternative TFP measures are shown: one that distinguishes between infrastructure and non-infrastructure capital (circles), and one that does not (diamonds). Both display a similar trend: TFP growth decelerated sharply after the mid-2000s. The

annual growth rate falls from above 4% around 2005 to below 2% after 2010.

3 The Economic Effects of the 2002 Tax Reform

So far, I have documented the sharp rise in progressivity, accompanied by a decline in local fiscal spending on infrastructure investment and narrowing regional disparities. While these trends coincide, time-series patterns alone do not establish causality. The main goal of this section is to build the causal links among regional redistribution, local fiscal policy, and regional economic growth. I start by describing the event-study empirical strategy. Then, I present the main results on the causal effects of the 2002 tax reform on growth in infrastructure capital and output. The robustness check is provided in Appendix B

3.1 Empirical Strategy

Although the 2002 tax reform applied nationwide, the intensity of its impact varied across provinces. As the central government started to take 50% of corporate income taxes and personal income taxes, the pre-reform share of income taxes in total tax revenue determined the exposure to the reform. Provinces with higher shares of income taxes had a larger portion of their revenue reallocated to the central government, thereby weakening their fiscal incentives to invest.

Although the reform was officially implemented in 2002, discussions began and details were announced in 2001. The central government also committed to compensating local governments based on their 2001 income tax collections. This led to strategic anticipation by local governments starting in 2001. For example, corporate income tax collection spiked in the fourth quarter of 2001 as local governments front-loaded taxes to raise the baseline for future transfers (Lan, 2021). To avoid this anticipation effect, I use 2000 as the baseline year. Formally, I define the treatment variable Z_i as:

$$Z_i = \frac{\text{Corporate Income Tax}_{it_0} + \text{Personal Income Tax}_{it_0}}{\text{Total Tax Revenue}_{it_0}} \tag{3}$$

where $t_0 = 2000$.

I conduct a difference-in-differences (DiD) strategy with continuous treatment. The reduced-form DiD equation is:

$$\text{outcome}_{it} = \zeta_i + \nu_t + \sum_{\iota \neq t_0} \beta_\iota Z_i \times \mathbb{1}\{t = \iota\} + \varepsilon_{it} \tag{4}$$

where outcome_{it} is the outcome variable (e.g. log of output per capita and infrastructure capital stock per capita) of province i in period t. The outcome variable is a function of regional fixed effects ζ_i, time fixed effects ν_t, the interaction of the treatment variable Z_i with time dummies, and the error term ε_{it}. The time dummy for period t_0 is omitted. Each coefficient β_ι measures how the relationship between the outcome variable and the treatment variable Z_i at year ι differs from that in year t_0.

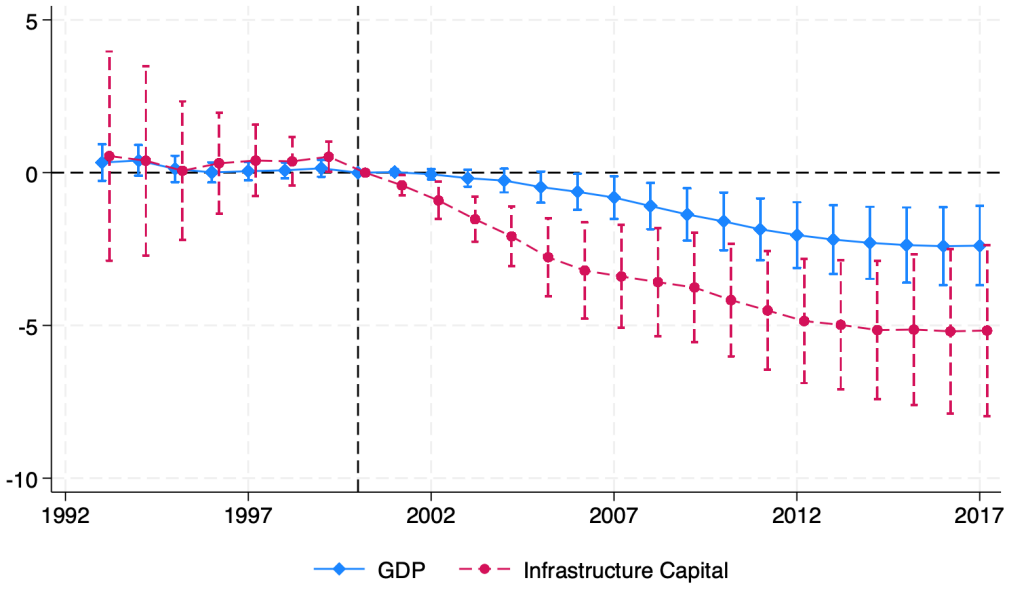
3.2 Main Results

This subsection presents the main results of the event study. Figure 8 plots the estimated coefficients β_ι from Equation (4) with their 95% confidence intervals. Diamonds (solid line) represent results for log GDP, while circles (dashed line) represent results for log infrastructure capital. The vertical dashed line marks the baseline year $t_0 = 2000$.

The pre-trends are flat and statistically indistinguishable from zero for both outcome variables, supporting the assumption of parallel trends. Following the reform, the estimated coefficients become significantly negative, consistent with the time-series evidence that the reform reduced local governments' incentives to invest in infrastructure and, in turn, slowed down economic growth. The static two-way fixed-effects coefficients are -3.90 (s.e. 1.19) for infrastructure capital and -1.44 (s.e. 0.42) for output. One standard deviation increase in the pre-reform share of income taxes will causally reduce infrastructure capital by 17% and output by 6%.

Figure B4 presents the effects on local fiscal revenue and government consumption. While the reform significantly reduces post-redistribution government revenue in high-share provinces, it did not significantly affect pre-redistribution revenue. Likewise, it has no significant effect on government consumption. These findings are consistent with the time-series evidence in Figure 6 that the gap in pre-redistribution revenue and government consumption between rich and poor regions persists, while the disparity in

Figure 8: The Treatment Effects of the Tax Reform on Regional GDP and Infrastructure Capital, Event Study.



Note: Each dot represents the estimated coefficient β_ι from Equation (4), with 95% confidence intervals. Coefficients capture the deviation in the correlation between the outcome variables and the income tax share from the base year (2000). Diamonds (solid line) correspond to log GDP; circles (dashed line) correspond to log infrastructure capital. The sample includes 30 provinces from 1993 to 2017. Standard errors are clustered by provinces.

infrastructure investment nearly disappears.

Appendix B reports additional results of robustness checks. Table B1 shows that the pre-reform share is uncorrelated to a wide range of confounders. Figure B4 confirms that the reform does not significantly influence other variables. Figure B5 presents the placebo test confirming that the treatment effects arise from the targeted tax categories rather than others.

In sum, the event study established the causal links among regional redistribution, local fiscal policy, and regional economic growth. The 2002 tax reform significantly depressed local governments' incentive to invest in infrastructure and thereby diminished regional economic growth. These empirical findings motivate the theoretical framework in the following sections.

An important by-product of this analysis is the justification for using the pre-reform share of income taxes as an instrument to estimate the output elasticity of infrastructure capital. As discussed above, the tax-sharing rule between central and local governments affects local economic activity exclusively through local fiscal policy. The results of the event study support the assumption of parallel trends, lending credibility to the exclusion restriction that the pre-reform share is uncorrelated with the structural error term of growth in infrastructure and output. In Section 7.1, I use the pre-reform share of income taxes as an IV to estimate the output elasticity of infrastructure capital.

4 Economic Environment

This section introduces the economic environment. The model is static. The economy consists of a unit measure of regions, indexed by $i \in [0, 1]$. Each region has a representative household, a representative firm, and a local government. There is one central government for the whole economy. Regions differ with respect to regional productivity A_i. At this point, productivity is the only source of heterogeneity. In the quantitative analysis, I will include wedges on the firm's capital rent aligning with the data.[3]

[3]I formulate the model with capital wedges on public capital and private capital in Appendix C.

4.1 Timeline

The interaction between the central government and local governments is modeled as a Stackelberg game, where the central government is the leader that designs the redistribution policy, and local governments are followers. The timing is as follows:

1. The central government announces the redistribution policy (ϕ, λ), the output tax rate τ_y, and central government consumption G^C. For convenience, I define g as the ratio of central government consumption to aggregate output. The output tax rate is uniform across regions.

2. Given the central government policies $(\phi, \lambda, \tau_y, g)$, each local government chooses its own consumption G_i and public capital k_{Gi}.

3. Given the central and local government policies, the firm rents private capital k_i and hires labor n_i to produce, and pays taxes $\tau_y y_i$. The household inelastically supply unit labor.

4. The local government collects taxes from the firm, pays taxes to or receives subsidies from the central government, and allocates post-redistribution revenue to local government consumption and public capital rental payments Rk_{Gi}.

4.2 Firm and Production

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

$$y_i = A_i k_{Gi}^{\alpha} k_i^{\theta} n_i^{1-\alpha-\theta} \tag{5}$$

where α and θ are output elasticities of public and private capital, respectively. I further assume $a_i = \log A_i \sim \mathcal{N}(-\sigma_a^2/2, \sigma_a^2)$. The firm's problem is given by

$$\max_{k_i, n_i} \pi_i = (1 - \tau_y) y_i - R k_i - w_i n_i \tag{6}$$

subject to Equation (5). The firm chooses private capital k_i and labor n_i to maximize the profit. The profit of the local firm belongs to the representative household.

First-order conditions with respect to n_i and k_i and the labor market clearing condition $n_i = 1$ yield the local firm's production plan and local wage rate:

$$\begin{aligned} k_i &= \left(\frac{\theta(1-\tau_y)}{R}\right)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}} \\ y_i &= \left(\frac{\theta(1-\tau_y)}{R}\right)^{\frac{\theta}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}} \\ w_i &= (1-\alpha-\theta)\left(\frac{\theta}{R}\right)^{\frac{\theta}{1-\theta}} (1-\tau_y)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}} \\ \pi_i &= \alpha\left(\frac{\theta}{R}\right)^{\frac{\theta}{1-\theta}} (1-\tau_y)^{\frac{1}{1-\theta}} A_i^{\frac{1}{1-\theta}} k_{Gi}^{\frac{\alpha}{1-\theta}} \end{aligned} \tag{7}$$

The redistribution policy (ϕ, λ) and central government consumption g do not directly appear in the firm's decisions. Although these factors potentially influence the firm's production through public capital k_G, the firm does not internalize the effects of redistribution. The intuition is that the firm does not distinguish between central and local taxes as long as the tax rate τ_y is constant.

Corollary 1 (Firm's production)**.** *The redistribution policy (ϕ, λ) does not directly affect the firm's production plan.*

4.3 Household

The representative household supplies unit labor inelastically and earns wages. The household also claims the profit of the local firm. I assume the logarithmic utility function. The household problem is given by

$$\begin{aligned} \max_{c_i} \quad & u(c_i, G_i, G^C) = \log c_i + \chi^L \log G_i + \chi^C \log G^C \\ & \text{s.t.} \quad c_i = w_i + \pi_i \end{aligned} \tag{8}$$

where w_i and π_i are determined in the equilibrium. The household derives utility from household consumption c_i, local government consumption G_i, and central government consumption G^C. χ^L and χ^C are parameters controlling for the household's relative preference over local and central government consumption, respectively. Government consumption G_i and G^C are exogenous to the household. In a static setting, the house-

hold problem is straightforward and is solved by $c_i = w_i + \pi_i = (1-\theta)(1-\tau_y)y_i$.

4.4 Local Government

The object of the local government is to maximize the local household utility. The local government problem is given by

$$\begin{aligned} \max_{G_i, k_{Gi}} \quad & \log c_i + \chi^L \log G_i + \chi^C \log G^C \\ \text{s.t.} \quad & G_i + R k_{Gi} = \lambda (\tau_y y_i)^{1-\phi} \end{aligned} \tag{9}$$

where y_i solves the firm's problem defined in (6) and c_i solves the household problem defined in (8). The local government allocates post-redistribution revenue $\lambda(\tau_y y_i)^{1-\phi}$ to local government consumption G_i and rent on public capital Rk_{Gi}. G^C is exogenously chosen by the central government.

4.5 Central Government

The object of the central government is to maximize household welfare across all regions. The central government problem is given by

$$\begin{aligned} \max_{\phi, \tau_y, g} \quad & \mathcal{W}(\phi, \tau_y, g) = \int W(a_i) u(c_i, G_i, G^C) dF_a(a) \\ \text{s.t.} \quad & \tau_y \int y_i dF_a(a) = \lambda \int (\tau_y y_i)^{1-\phi} dF_a(a) + g \int y_i dF_a(a) \end{aligned} \tag{10}$$

where c_i, y_i, and G_i solve the local government problem defined by (12), $W(a)$ is the Pareto weight function, and $F_a(a)$ is the cumulative density function of productivity. λ is implicitly chosen to equate the budget constraint. The central government's budget constraint ensures that total tax revenue equals post-redistribution local government revenue plus central government consumption.

I follow Heathcote and Tsujiyama (2021) and assume that the Pareto weight function takes the form of

$$W(a; \gamma) = \frac{\exp(-\gamma a)}{\int \exp(-\gamma a) dF_a(a)} \tag{11}$$

where γ is the parameter controlling for the central government's taste for redistribution. With a negative (positive) γ, the central government puts relatively higher weight

on more (less) productive regions. The utilitarian social welfare function corresponds to $\gamma = 0$.

4.6 Discussion

Before characterizing the equilibrium allocation, I discuss some pivotal features of the model.

First, the model integrates the HSV tax function into the Mandarin model developed by Song and Xiong (2023) to analyze how local governments adjust fiscal policy in response to the central government's redistribution scheme. My framework distinguishes between public goods (government consumption in our model) and government investment (modeled as public capital). Government consumption directly enters household utility but does not affect output, while infrastructure investment enhances regional productivity. Local governments invest in infrastructure to stimulate economic growth and increase tax revenue. However, when a larger share of tax revenue is taken by the central government, the marginal return to investment declines. By explicitly separating government consumption from infrastructure investment, the model captures a central distortion of redistribution: rising progressivity shifts spending away from productive public capital toward unproductive consumption, thereby reducing aggregate output.

Second, the model explicitly distinguishes between central government consumption and local government consumption. In reality, central and local government consumption serve functionally distinct roles.[4] Theoretically, modeling central government consumption features a classic free-riding problem in the public finance literature. Local governments do not internalize the positive externality that their investment in public capital can improve local household welfare by enhancing aggregate output and central government consumption. Therefore, incentivizing local public capital through the redistribution system is necessary.

[4]In China, roughly one-third of central government expenditures are devoted to the military, while other expenditures on R&D, education, and healthcare benefit households nationwide. In contrast, local governments primarily finance expenditures on local education, healthcare, social security, and related services.

4.7 Problem of the Representative Region

It is useful to solve a special case of the local government problem. When $\sigma_a^2 = 0$, there is no dispersion in regional productivity. The economy collapses to a representative region model. The representative local government's problem is given by:

$$\begin{aligned} \max_{G,k_G} \quad & \log C + \chi^L \log G + \chi^C \log G^C \\ \text{s.t.} \quad & G + RK_G = \lambda(\tau_y Y)^{1-\phi} \end{aligned} \tag{12}$$

where C solves the household problem and Y solves the firm's problem. Taking the policy parameter $(\phi, \lambda, \tau_y, g)$ as given and treating λ as determined in the equilibrium, the optimal choices for a representative local government are given by:

$$\begin{aligned} \log K_G^R(\phi, \tau_y, g) = & \frac{1-\theta}{1-\theta-\alpha} \log\left(\frac{\alpha(1+\chi(1-\phi))}{\alpha+\chi(1-\theta)}\right) \\ & + \frac{1}{1-\theta-\alpha}\left((1-\theta)\log(\tau_y - g) - \log R + \theta \log(\theta(1-\tau_y))\right) \end{aligned} \tag{13}$$

$$\begin{aligned} \log Y^R(\phi, \tau_y, g) = & \frac{\alpha}{1-\theta-\alpha} \log\left(\frac{\alpha(1+\chi(1-\phi))}{\alpha+\chi(1-\theta)}\right) \\ & + \frac{1}{1-\theta-\alpha}(\alpha \log(\tau_y - g) - (\alpha+\theta)\log R + \theta \log(\theta(1-\tau_y))) \end{aligned} \tag{14}$$

Both public capital and output are decreasing in the degree of progressivity ϕ. Progressivity increases the marginal tax rate on fiscal revenue and reduces the marginal return to renting capital, which motivates local governments to shift spending away from capital rent toward government consumption. This distortionary effect is uniform across regions. The second line in each equation captures the effects of the output tax rate τ_y, central government expenditure g, and the interest rate R. Higher output tax rates levy more taxes for local governments to finance capital rent but also distort the firm's production plan. Higher central government consumption decreases transfers towards local governments, thereby decreasing capital rent and output.

5 Equilibrium

This section characterizes the equilibrium allocations under the redistribution policy. I begin by formally defining the equilibrium of this economy.

Definition 1 (Equilibrium)**.** *The equilibrium consists of central government policy* $(\phi, \lambda, g, \tau_y)$*, local government fiscal policy* $\{k_{Gi}, G_i\}_{i\in[0,1]}$*, household consumption* $\{c_i\}_{i\in[0,1]}$*, the firm's production plan* $\{y_i, k_i, n_i, \pi_i\}_{i\in[0,1]}$*, and the local wage rate* $\{w_i\}_{i\in[0,1]}$*, such that*

1. *Given the local government fiscal policy* (k_{Gi}, G_i) *and wage rate* w_i*,* (y_i, k_i, n_i, π_i) *solves the firm's problem defined in* (6)*.*

2. *Given the wage rate* w_i *and profit* π_i*,* c_i *solves the household problem defined in* (8)*.*

3. *Given the central government policy* $(\phi, \lambda, g, \tau_y)$*, the firm's production plan* y_i*, and the household consumption* c_i*,* (k_{Gi}, G_i) *solves the local government problem defined in* (12)*.*

4. *Given the local government best response* (k_{Gi}, G_i)*,* $(\phi, \lambda, g, \tau_y)$ *solves the central government problem defined in* (10)*.*

5. w_i *clears the local labor market,* $n_i = 1$*.*

Substituting the expression of the firm's production plan (7) into the local government problem yields optimal local government policy k_{Gi} and G_i as a function of productivity A_i and the central government policy $(\phi, \lambda, g, \tau_y)$. Plugging k_{Gi} back into y_i in Equation (7) and using the central government's budget constraint in (10) can pin down the parameter λ. Details and proofs are presented in Appendix C. Proposition 1 characterizes the equilibrium allocation k_{Gi} and y_i as a function of productivity A_i and central government policy (ϕ, g, τ_y).

Proposition 1 (Equilibrium allocation)**.** *Given the central government policy* ϕ*, the equilibrium local public capital is given by*

$$\log k_G(A; \phi, \tau_y, g) = \log K_G^R(\phi, \tau_y, g) + \frac{1-\phi}{1-\theta-\alpha(1-\phi)} \log A_i + \log \Omega^{k_G}(\sigma_a^2; \phi) \tag{15}$$

and the equilibrium regional output is given by

$$\log y(A;\phi,\tau_y,g) = \log Y^R(\phi,\tau_y,g) + \frac{1}{1-\theta-\alpha(1-\phi)}\log A_i + \log \Omega^y(\sigma_a^2;\phi) \tag{16}$$

where $\log K_G^R$ *and* $\log Y^R$ *are given by Equations* (13) *and* (14). Ω^{k_G} *and* Ω^y *are constants and defined in Appendix C.1*

The public capital and output are composed of three terms. The first term reflects the allocation of the representative region. As explained before, higher progressivity ϕ reduces the marginal return to renting capital, motivating local governments to reallocate spending from renting capital toward government consumption. This leads to lower levels of public capital and output uniformly across regions. The second term captures the role of regional heterogeneity in productivity A_i. Public capital and output are increasing in productivity. As ϕ increases, the elasticity of k_G and y with respect to productivity declines. This captures the redistributive effects that progressivity shrinks the pass-through from regional productivity to regional economic outcomes, which is central to our analysis. While high-productivity regions generate higher returns to renting capital, a progressive scheme reallocates fiscal resources away from these regions, weakening their ability to convert productivity advantages into higher output. The third term reflects the influence of the distribution of productivity. With the mean-preserving spread, larger productivity dispersion increases aggregate productivity and output by motivating high-productivity regions to rent more capital. Progressivity weakens this channel by dampening the efficient response of capital rent to productivity.

Proposition 1 further sheds light on the relationships among the economic variables in the equilibrium. Corollary 2 delivers the ratio of government capital rental payments to local government consumption. This ratio is uniform across regions and is decreasing in ϕ as progressivity reduces the marginal return to public capital, thereby shifting local government spending from capital rental towards public goods provision. This ratio is also decreasing in χ^L and increasing in α. That is, local governments should rent more capital and reduce government consumption if public capital is more productive and/or if public goods are less valued by households.

Corollary 2 (Local government spending)**.** *The ratio of local government capital rent*

to consumption is given by

$$\frac{Rk_{Gi}}{G_i} = \frac{\alpha(1+\chi^L(1-\phi))}{\chi^L(1-\theta-\alpha(1-\phi))}$$

Following Hsieh and Klenow (2009), I assess capital misallocation by the marginal return to capital. In a non-redistribution economy, the marginal return to capital should be uniform across regions. Under the Cobb-Douglas specification, the marginal return to capital is inversely proportional to the capital-output ratio. Corollary 3 derives the capital-output ratio. When $\phi > 0$, the public capital-output ratio is decreasing in productivity, indicating that redistribution redirects public capital from high- to low-productivity regions. Notably, the private capital-output ratio is independent of progressivity and is uniform across regions because firms do not internalize the effects of redistribution as shown in Corollary 1.

Corollary 3 (Capital misallocation)**.** *The public capital-output ratio is given by*

$$\frac{k_{Gi}}{y_i} \propto A_i^{\frac{-\phi}{1-\theta-\alpha(1-\phi)}} \tag{17}$$

while the private capital-output ratio k_i/y_i is uniform across regions.

Corollary 4 characterizes how the redistribution policy shapes the distribution of regional output. Higher progressivity compresses the dispersion in regional output. This follows immediately from Equation (16) that progressivity lowers the elasticity of output with respect to productivity. Holding everything else constant, an increase in ϕ over time leads to regional economic convergence.

Corollary 4 (Output dispersion)**.** *The variance of log regional output is given by*

$$\text{var}(\log y(a;\phi,\tau_y,g)) = \frac{\sigma_a^2}{(1-\theta-\alpha(1-\phi))^2}$$

I visualize the effects of progressivity on local government fiscal conditions, fiscal policy, and economic consequences in Figure 9, corresponding to the stylized facts introduced in Section 2. In the top panel, while the pre-redistribution government revenue does not change too much, the post-redistribution revenue in high-productivity regions

is largely reduced by the increase in progressivity. The higher degree of progressivity achieves a more equalized local government fiscal revenue after redistribution.

In the middle panel, while the increase in progressivity decreases the rent on public capital and government consumption in high-productivity regions simultaneously, the decline in public capital rent exceeds the decline in government consumption. This reflects that local governments shift fiscal spending away from renting public capital towards purchasing public goods, especially in rich regions. Meanwhile, public capital and government consumption in low-productivity regions increase at the same time, implying the improvement in regional equality.

In the bottom panel, the increase in progressivity redistributes capital from high- to low-productivity regions, thereby increasing (decreasing) the output-capital ratio in low-productivity (high-productivity) regions and exacerbating the capital misallocation. Meanwhile, as capital flows out, output in high-productivity regions decreases.

Next, Proposition 2 derives the closed-form expression of aggregate quantities. Let $Y = \int y_i di$, $K_G = \int k_{Gi} di$, and $K = \int k_i di$ denote the aggregate quantities. Aggregate TFP is calculated as the Solow residual: $\log Y - \alpha \log K_G - \theta \log K$.

Proposition 2 (Aggregate quantities)**.** *The aggregate output is given by*

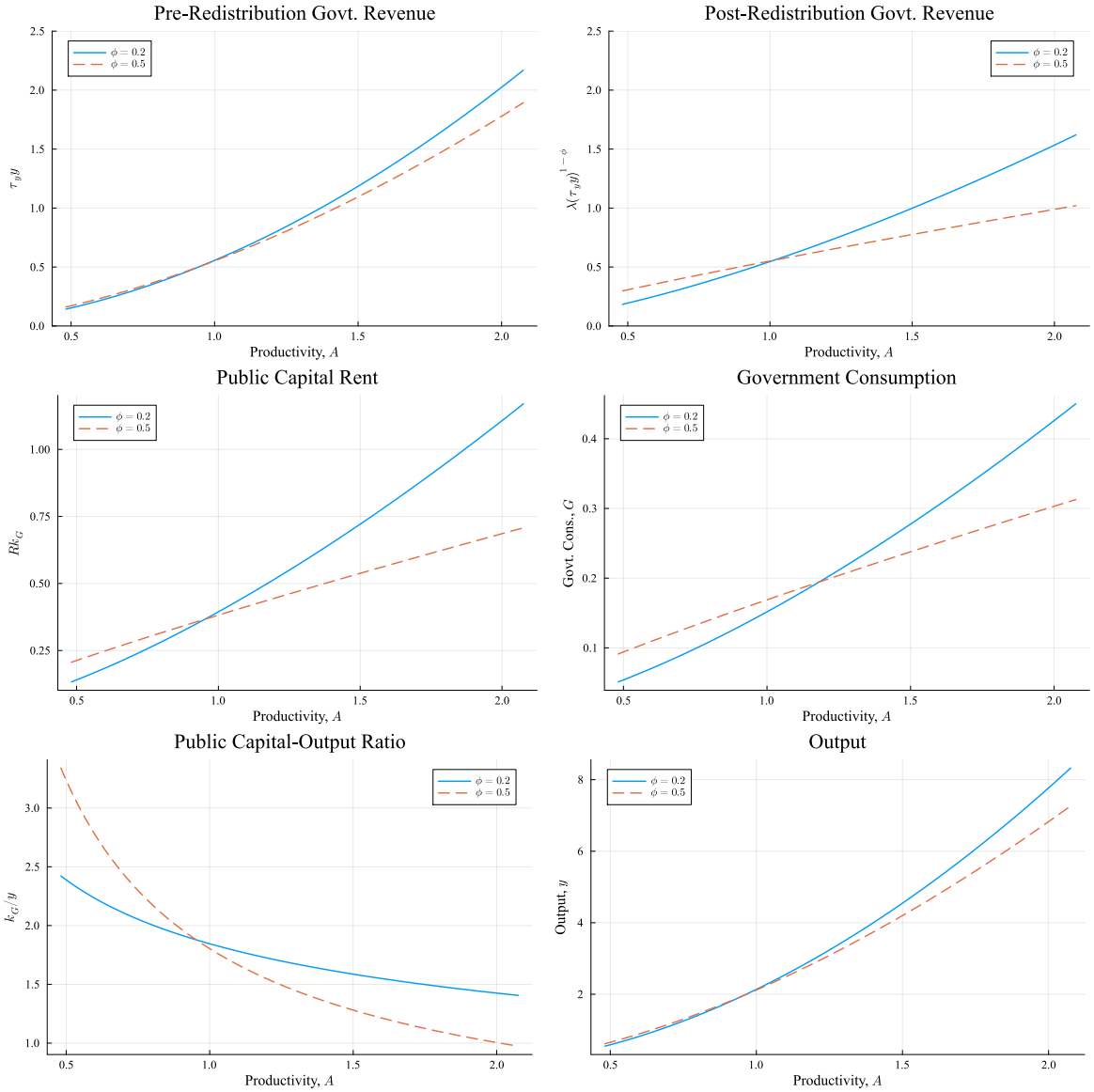
$$\log Y(\sigma_a^2; \phi, \tau_y, g) = \log Y^R(\phi, \tau_y, g) + \frac{1}{1-\theta-\alpha} \log \text{TFP}$$

where $\log Y^R(\phi, \tau_y, g)$ *is output in the representative region given by Equation* (14) *and* $\log \text{TFP}$ *is the aggregate TFP given by:*

$$\log \text{TFP}(\sigma_a^2; \phi) = \frac{\theta+\alpha}{1-\theta-\alpha} \frac{\sigma_a^2}{2} - \underbrace{\frac{\alpha(1-\theta)\phi^2}{(1-\theta-\alpha(1-\phi))^2(1-\theta-\alpha)} \frac{\sigma_a^2}{2}}_{\text{Distortion}} \tag{18}$$

Aggregate output is composed of two main components. The first term is output in the representative region, which, as discussed before, is decreasing in progressivity ϕ as progressivity shifts local fiscal spending away from renting public capital. The second term is aggregate productivity scaled by the technology parameters. Aggregate TFP is composed of two terms. The first term captures the effects of the productivity dispersion in a *non-redistribution* economy. When $\phi = 0$, aggregate efficiency is increasing

Figure 9: Local Government's Fiscal Conditions, Policy Function, and Economic Consequences



Note: These figures plot pre- and post-redistribution local government revenue, local government capital rent and consumption, public capital-output ratio, and output as functions of regional TFP. Solid and Dashed lines represent policy functions for $\phi = 0.2$ and $\phi = 0.5$, respectively.

in productivity dispersion. This occurs because both output and capital are convex in productivity. A mean-preserving spread in productivity causes capital to reallocate disproportionately toward high-productivity regions, resulting in the increase in aggregate efficiency. The second term captures the distortionary effects of redistribution policy on aggregate efficiency. Any deviation from $\phi = 0$ will cause the efficiency loss. When $\phi > 0$, high-productivity regions obtain excessive capital while low-productivity regions get insufficient capital compared to the optimal level. Conversely, under regressive redistribution ($\phi < 0$), high-productivity regions receive insufficient capital while low-productivity regions have excessive capital. Both situations generate capital misallocation and reduce aggregate efficiency.

Corollaries 5 and 6 further summarize the effects of progressivity on aggregate output and efficiency.

Corollary 5 (Progressivity and output)**.** *When $\phi \geq 0$, aggregate output is strictly decreasing in ϕ*

Corollary 5 states that increasing progressivity always reduces aggregate output. This result arises through two main channels. First, progressivity distorts the resource allocation by shifting capital from high- to low-productivity regions, reducing aggregate efficiency. Second, progressivity encourages a shift away from productive public capital towards non-productive government consumption, lowering output as well. Notably, decreasing ϕ does not always increase output. Although regressivity encourages local governments to rent more public capital, it causes misallocation in public capital and reduces aggregate efficiency.

Corollary 6 (Progressivity and efficiency)**.** $\phi = 0$ *maximizes aggregate TFP. When* $\chi^C = 0$, $\phi = 0$ *solves the central government's problem when* $\gamma = -1/(1-\alpha-\theta)$, *corresponding to the laissez-faire case.*

Corollary 6 summarizes the relationship between redistribution and efficiency. $\phi = 0$ maximizes aggregate efficiency as any deviation from $\phi = 0$ will cause capital misallocation. Additionally, when $\chi^C = 0$, the model does not feature the free-riding problem for central government consumption. Then, $\phi = 0$ is efficient and solves the central government problem with the Pareto weight proportional to $\exp[1/(1-\alpha-\theta)a]$. This

corresponds to the laissez-faire case, where the Pareto weights are inversely proportional to the equilibrium marginal utility in the absence of redistribution. The underlying intuition is that in a non-redistribution economy, more productive regions naturally enjoy higher levels of household and government consumption, leading to lower marginal utility. Therefore, if it is optimal not to implement redistribution, the social planner must place relatively higher Pareto weights on these regions, which should be inversely proportional to their marginal utility.

6 Regional Redistribution and Social Welfare

This section investigates the welfare implications of redistribution and explores the determinants of the optimal progressivity. I start by deriving the closed-form expression of social welfare and explaining the economic interpretation of each term. Then, I derive the condition for optimal progressivity and discuss the economic forces shaping optimal progressivity. Following Heathcote and Tsujiyama (2021), this framework can infer the central government's social preference from the observed redistribution policy.

6.1 Social Welfare Function

Proposition 3 presents the closed-form expression of social welfare as a function of central government policy tools (ϕ, τ_y, g) while λ is set to balance the central government's budget constraint in the equilibrium. I omit terms that are constant and independent of policy variables.

Proposition 3 (Closed-form social welfare function)**.** *With the social welfare function*

defined in Equation (10), the social welfare of implementing policy (ϕ, τ_y, g) is given by

$$\mathcal{W}(\phi, \tau_y, g) =$$

$$\begin{cases} \dfrac{1}{1-\theta-\alpha}[(\alpha(1+\chi^C)+\chi^L(1-\theta))\log(\tau_y-g)+(1-\alpha+\theta(\chi^L+\chi^C))\log(1-\tau_y)]+\chi^C\log g & \text{(a)} \\ +\dfrac{(1-\theta)\chi^L+\alpha(1+\chi^C)}{1-\theta-\alpha}\log\left(\dfrac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)+\chi^L\log\left(\dfrac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha(1+\chi^L(1-\phi))}\right) & \text{(b)} \\ +\dfrac{1+\chi^L+\chi^C}{1-\theta-\alpha}\left[\dfrac{\theta+\alpha}{1-\theta-\alpha}\dfrac{\sigma_a^2}{2}-\dfrac{\alpha(1-\theta)\phi^2}{(1-\theta-\alpha)(1-\theta-\alpha(1-\phi))^2}\dfrac{\sigma_a^2}{2}\right] & \text{(c)} \\ -\dfrac{1}{(1-\theta-\alpha(1-\phi))^2}\dfrac{\sigma_a^2}{2}-\dfrac{\chi^L(1-\phi)^2}{(1-\theta-\alpha(1-\phi))^2}\dfrac{\sigma_a^2}{2} & \text{(d)} \\ -\dfrac{(1+\chi^L(1-\phi))}{1-\theta-\alpha(1-\phi)}\gamma\sigma_a^2 & \text{(e)} \end{cases} \tag{19}$$

Equation 19 expresses the social welfare as a function of preference, technology, productivity dispersion, and policy parameters. There are two important properties of the social welfare function.

First, in Appendix C.3, I show that the twice differentiations of the social welfare function with respect to τ_y and g are negative, establishing global concavity in these two arguments. For ϕ, I show that the twice differentiation is negative for all components except for Line (c), which becomes negative when $1-\alpha-\theta > 2\alpha\phi$. This condition is satisfied in the calibration. For a concave function, first-order conditions are sufficient to deliver the optimal solution.

Corollary 7 (Concavity of the social welfare function)**.** *The social welfare function is globally concave in τ_y and g, and, if $1-\alpha-\theta > 2\alpha\phi$, is globally concave in ϕ.*

Second, in Equation (19), ϕ and (τ_y, g) enter the social welfare function separately and do not interact with each other. It implies that the optimal degree of progressivity must be irrelevant to any choice of (τ_y, g) even if it is not optimal. Therefore, we can derive the optimal tax rate and central government consumption ratio by taking the first-order conditions of the social welfare function with respect to τ_y and g without solving for the optimal value of ϕ. Corollary 8 formulates the optimal choice of output tax rate and central government consumption.

Corollary 8 (Optimal government size)**.** *The optimal output tax rate τ_y and the central government consumption to output ratio g are given by*

$$\tau_y^* = \frac{\alpha + (1-\theta)(\chi^L + \chi^C)}{1 + \chi^L + \chi^C}, \quad g^* = \frac{(1-\alpha-\theta)\chi^C}{1 + \chi^L + \chi^C} \tag{20}$$

The optimal output tax rate is increasing in α, χ^L, and χ^C. The central government should raise the tax rate if public capital is more productive so that local governments can collect more taxes and rent more public capital, and/or if households value government consumption more so that both central and local governments can purchase more public goods. The optimal output tax rate is decreasing θ because the output tax distorts the firm's production plan.

The optimal ratio of central government consumption to aggregate output is increasing in χ^C. The central government should expand spending when households value it more. Meanwhile, the optimal ratio is decreasing in α, θ, and χ^L. Greater output elasticities and stronger preferences for local government consumption indicate that the central government should leave more fiscal revenue for local governments to finance public goods and capital rental payments.

6.2 Decomposition of Social Welfare Function

In this subsection, I interpret each term of Equation (19) and explore the economic forces that determine the optimal progressivity. I decompose the social welfare into the welfare of the representative region and the welfare from heterogeneity, with the latter consisting of the welfare gain from efficiency and the welfare cost from inequality.

6.2.1 Welfare of the Representative Region

Lines (a) and (b) in Equation (19) correspond to the welfare of the representative region:

$$\begin{aligned}
&\mathcal{W}^R(\phi, \tau_y, g) \\
&= \frac{1}{1-\alpha-\theta}[(\alpha(1+\chi^C) + \chi^L(1-\theta))\log(\tau_y - g) + (1 - \alpha + \theta(\chi^L + \chi^C))\log(1-\tau_y)] + \chi^C \log g \\
&+ \frac{(1-\theta)\chi^L + \alpha(1+\chi^C)}{1-\alpha-\theta} \log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha + \chi^L(1-\theta)}\right) + \chi^L \log\left(\frac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha(1+\chi^L(1-\phi))}\right)
\end{aligned} \tag{21}$$

1. Welfare from the central fiscal policy (τ_y, g). This first line of Equation (21) represents the welfare effects of central government policy (τ_y, g). Increasing the output tax rate raises fiscal capacity and can enhance welfare through public spending. But it also distorts firms' production plans, reducing the equilibrium wage rate and output. Similarly, the increase in g can improve household utility from central government consumption but reduces funds available for local public goods and public capital. The optimal choice of the output tax rate and the central government consumption share is given by Equation (20).

2. Welfare from the local government fiscal policy. The second line of Equation (21) represents the welfare from local fiscal policy:

$$\frac{Rk_{Gi}}{\lambda(\tau_y y_i)^{1-\phi}} = \frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)} \tag{22}$$

$$\frac{G_i}{Rk_{Gi}} = \frac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha(1+\chi^L(1-\phi))} \tag{23}$$

Equation (22) gives the ratio of local government spending on capital rent to post-redistribution revenue, which is decreasing in ϕ. This ratio appears in the first term of Line (b) in Equation (19) and is scaled by the preference parameters, reflecting that progressivity lowers the equilibrium household and government consumption by reducing public capital. Equation (23) gives the ratio of local government consumption to the rent on public capital, which is increasing in ϕ. This ratio appears in the second term of Line (b) and is scaled by χ^L, capturing the household utility from local government consumption.

These two terms capture a major motivation of the redistribution policy: financing central government consumption through central taxes. The central government consumption is exogenously determined for the household and local governments. $\chi^C > 0$ raises a free-rider problem: local governments do not internalize that increasing public capital improves household utility additionally through increasing central taxes and central government consumption. As a result, extra incentives should be provided to encourage local governments to rent more public capital through a regressive tax scheme. Proposition 4 derives the optimal ϕ in a representative region model.

Proposition 4 (Representative region model)**.** *In a representative region economy, i.e.*

$\sigma_a^2 = 0$, *the welfare-maximizing* ϕ^{R*} *should be negative and is given by*

$$\phi^{R*} = -\frac{(1-\theta-\alpha)\chi^C}{\alpha(1+\chi^C)+(1-\theta)\chi^L} \tag{24}$$

With τ_y^* *and* g^* *given by Equation* (20)*, the marginal tax rate is 0 and the average tax rate is* g^*/τ_y^*.

When heterogeneity is eliminated, the goal of redistribution to reduce inequality vanishes. The only objective is to finance central government consumption. In a representative-region economy, with ϕ^{R*} given by Equation (24) and (τ_y^*, g^*) given by Equation (20), the optimal marginal tax rate (MTR) on local fiscal revenue is exactly 0, which replicates a lump-sum transfer and introduces no distortions. Meanwhile, the average tax rate (ATR) is g/τ_y, which is exactly sufficient to finance central government consumption. When ATR > MRT, the system is regressive.

6.2.2 Welfare from Heterogeneity

Lines (c), (d), and (e) in Equation (19) correspond to welfare from heterogeneity.

1. Welfare from efficiency. Line (c) of Equation (19) captures the effects of heterogeneity on aggregate efficiency interacted with the redistribution policy as

$$\log \text{TFP} = \frac{\theta+\alpha}{1-\theta-\alpha}\frac{\sigma_a^2}{2} - \frac{\alpha(1-\theta)\phi^2}{(1-\theta-\alpha(1-\phi))^2(1-\theta-\alpha)}\frac{\sigma_a^2}{2}$$

Line (d) is the log of aggregate TFP scaled by the preference parameters. As discussed in Section 5, while the dispersion in productivity improves aggregate efficiency, redistribution, both progressivity and regressivity, diminishes the efficiency gain. This negative effect is minimized precisely at $\phi = 0$.

2. Welfare cost from inequality. Line (d) captures the welfare cost from inequality in household consumption and local government consumption:

$$\text{var}(\log c_i) = \frac{1}{(1-\theta-\alpha(1-\phi))^2}\sigma_a^2$$

$$\text{var}(\log G_i) = \frac{(1-\phi)^2}{(1-\theta-\alpha(1-\phi))^2}\sigma_a^2$$

The terms in Line (d) are the variance of log household consumption and the variance of log local government consumption scaled by the preference parameters. Due to the concavity of the log utility function, inequality is costly. Line (d) is strictly increasing in ϕ, indicating that a more progressive redistribution policy reduces regional inequality and thereby improves social welfare.

3. The central government's social preference. Finally, Line (e) captures the central government's attitude towards innate inequality. With a larger γ, the central government puts relatively higher Pareto weights on less productive regions. This increases the welfare cost from inequality due to productivity dispersion and thus strengthens the central government's desire to redistribute. Line (e) is also strictly increasing in ϕ when $\gamma > 0$.

6.3 Condition for Optimal Progressivity

When should the tax scheme be progressive? From Proposition 4, in the representative region model, the optimal tax scheme should be regressive and the optimal ϕ should be negative. Once heterogeneity is introduced, progressivity can improve social welfare by reducing social inequality. Since the social welfare function is continuous in ϕ, there must exist a threshold beyond which progressivity improves social welfare. To identify this threshold, I evaluate the differentiation of the social welfare function with respect to ϕ at $\phi = 0$. Proposition 5 presents the sufficient and necessary condition under which progressivity improves social welfare.

Proposition 5 (Condition for optimal progressivity)**.** *The optimal level of ϕ should be strictly positive if and only if*

$$\underbrace{\frac{\chi^L(1-\theta)+\alpha}{(1-\theta-\alpha)^2}\sigma_a^2}_{\text{(a) less inequality due to TFP dispersion}} + \underbrace{\frac{\chi^L(1-\theta)+\alpha}{1-\theta-\alpha}\gamma\sigma_a^2}_{\text{(b) more redistribution}} + \underbrace{\alpha\chi^L}_{\text{(c) more local govt. cons.}} > \underbrace{\frac{\alpha\chi^L}{1+\chi^L}(1+\chi^C+\chi^L)}_{\text{(d) less hh. and govt. cons. due to less } k_G} \tag{25}$$

The left-hand side of Inequality (25) captures the marginal benefit of increasing pro-

gressivity at $\phi = 0$. Progressivity can improve social welfare by (a) reducing regional inequality due to productivity dispersion, (b) redistributing towards less productive regions when the central government assigns higher Pareto weights to them, and (c) motivating local governments to purchase more public goods. The right-hand side captures the marginal cost of increasing progressivity at $\phi = 0$. Specifically, higher progressivity reduces local governments' incentives to rent public capital, thereby lowering aggregate output, equilibrium wage rate, and government consumption. [5]

What economic forces determine progressivity? First, it is straightforward to verify that Inequality (25) always holds when $\chi^L = 0$ and/or $\chi^C = 0$. When $\chi^C = 0$, there is no free-rider problem for central government consumption. Therefore, the need for correcting it via the regressive scheme disappears. When $\chi^L = 0$, local governments spend fiscal revenue exclusively on renting public capital. Hence, the distortion on local governments' fiscal decisions disappears. On the contrary, larger χ^C makes the condition for $\phi^* > 0$ more difficult to satisfy, as the welfare loss from less central government consumption will be enlarged. The effects of higher χ^L are ambiguous. Both the welfare loss from less local government consumption and the welfare gain from reducing inequality in local government consumption are amplified by higher χ^L.

Second, higher dispersion in productivity makes the condition for $\phi^* > 0$ easier to satisfy, as the marginal welfare gain from reducing inequality in household and local government consumption is amplified. Meanwhile, a larger γ increases the marginal value of redistribution, thereby making a positive ϕ more desirable.

The role of output elasticity of public capital α is unambiguous. Both the welfare gain from reducing inequality and the distortionary effects are exaggerated when α is higher. In contrast, higher output elasticity of private capital θ will make the condition easier to satisfy. Higher θ will also exaggerate the inequality due to productivity dispersion. However, since the firm does not internalize the effects of redistribution, there will not be distortionary effects related to θ.

[5] Notably, Inequality (25) does not account for the welfare gain from efficiency. This effect is captured by Line (c) in Equation (19), the differentiation of which with respect to ϕ is 0 at $\phi = 0$. That is, the negative effect of redistribution on aggregate efficiency is minimized exactly at $\phi = 0$. Therefore, this term does not exist in the condition.

6.4 Social Preference

So far, I have examined economic forces that determine the optimal degree of progressivity. Among these factors, the central government's taste for redistribution plays a critical role. A higher γ reflects a stronger desire for redistribution and motivates the central government to implement a more progressive tax scheme. Conversely, it is tempting to recover the central government's social preference that makes the observed policy optimal, like Bourguignon and Spadaro (2012); Lockwood and Weinzierl (2016); Heathcote and Tsujiyama (2021). Proposition 6 provides the mapping from the observed progressivity $\hat{\phi}$ to the corresponding value of $\hat{\gamma}$ that rationalizes the observed policy.

Proposition 6 (Social preference)**.** *The social preference parameter $\hat{\gamma}$, which rationalizes the redistribution policy $\hat{\phi}$, is a solution to the following linear equation*

$$\begin{aligned}\frac{\chi^L(1-\theta)+\alpha}{(1-\theta-\alpha(1-\hat{\phi}))^2}\sigma_a^2\hat{\gamma} = &\frac{\alpha(1+\chi^L+\chi^C)}{1-\theta-\alpha}\frac{\chi^L}{1+\chi^L(1-\hat{\phi})} - \frac{\alpha\chi^L}{1-\theta-\alpha(1-\hat{\phi})} \\ &+ \frac{(1+\chi^L+\chi^C)\alpha(1-\theta)}{1-\theta-\alpha}\frac{\hat{\phi}}{(1-\theta-\alpha(1-\hat{\phi}))^3}\sigma_a^2 \\ &- \frac{\alpha+\chi^L(1-\hat{\phi})(1-\theta)}{(1-\theta-\alpha(1-\hat{\phi}))^3}\sigma_a^2\end{aligned} \tag{26}$$

where $\hat{\phi}$ is the observed progressivity.

Equation (26) is just the first-order condition of the social welfare function given by Equation (19) with respect to ϕ. If the observed progressivity $\hat{\phi}$ solves the Ramsey problem defined by (10), then $\hat{\gamma}$ and $\hat{\phi}$ must satisfy Equation (26). As Equation (26) is linear in $\hat{\gamma}$, it has a unique solution.

7 Parameterization

This section discusses the strategy of model parameterization. The main mechanism of our framework is that the redistribution policy influences regional growth by distorting local fiscal spending on infrastructure capital. Therefore, it is vital to estimate the output elasticity. The main goal of this section is to estimate the output elasticity

using the instrumental variable (IV) approach. I also introduce the calibration of other parameters and briefly discuss the evaluation of the model's performance.

7.1 Output Elasticity

This subsection estimates the output elasticity using the pre-reform share of income tax as an IV. As discussed in Section 3, this share measures the local government's exposure to the 2002 tax reform. Higher-share provinces experienced significantly lower post-reform growth in infrastructure capital and output. More importantly, how central and local governments share the tax revenue influences local fiscal policy exclusively, without directly affecting the household's or the firm's decisions. This is consistent with the model's prediction summarized in Corollary 1.

For any variable x_{it}, define the average growth in x from the base year t_0 to year t as $\Delta x_{it} = (x_{it} - x_{it_0})/(t - t_0)$. Then, the reduced form of IV identification is specified as:

$$\Delta \log y_{it} = \alpha_0 + \alpha \Delta \log k_{Git} + \epsilon_{it}$$

with the first stage given as:

$$\Delta \log k_{Git} = \rho_0 + \rho Z_i + \xi_{it}$$

where y_{it} and k_{Git} are output and infrastructure capital, the same as their theoretical counterparts in Section 4. Z_i is the pre-form share of income tax defined by Equation (3). The exclusion restriction requires that Z_i is uncorrelated to both ϵ_{it} and ξ_{it}. Although not directly testable, the parallel pre-trends in Figure 8 imply that the pre-reform share is uncorrelated to the growth in output and infrastructure capital before the reform. Table B1 shows that Z_i is uncorrelated to a wide range of confounders, and Figure B4 shows that Z_i affects local fiscal policy exclusively and has no significant effects on other economic variables.[6] All these empirical findings support the exclusion restriction. For the relevance condition, we should expect the first-stage coefficient ρ to

[6]In particular, the lack of an association between the pre-reform share and other policy variables implies that the IV Z_i helps rule out confounding from other fiscal tools, such as tax cuts, government consumption, and government debt.

be significantly negative. Under a valid instrument, the two-stage least squares estimate $\hat{\alpha}$ recovers the causal effect of infrastructure capital on output.

Notably, in Figure B4, after 2010, high-share provinces experience a decline in non-infrastructure capital. This likely reflects the crowd-in effects of infrastructure investment on private capital in the long run. While this constitutes an indirect channel through which infrastructure improves output, it may upwardly bias the estimated output elasticity. Additionally, in Figure 8, the standard errors and the confidence intervals become increasingly large as they move further from the baseline year. This may reflect the diminishing effects of the tax reform over time. As a result, the pre-reform income tax share becomes less predictive for infrastructure growth, which may weaken the precision of the IV estimation. To address this concern, the IV estimation restricts the sample period to 2001–2010.

Table 2 summarizes the IV estimates. All regressions control for year fixed effects, and standard errors are clustered by 30 provinces. Columns (1) and (2) use pre-reform data (1993–2000) to regress changes in log infrastructure capital and log GDP on the pre-reform share of income taxes. Consistent with the parallel trends observed in the event study, the coefficients are insignificant, supporting the exclusion restriction.

Columns (3)–(10) use post-reform data (2001–2010). Columns (3) and (4) present the OLS estimates of the output elasticity of infrastructure capital. The estimated elasticity is approximately 0.10.

Columns (5) and (6) show the first stage of the IV estimation. The coefficient on the share of income taxes is significantly negative: a one percentage point increase in the pre-reform share of income taxes will reduce the growth rate of infrastructure capital by roughly 0.5 percentage points. The coefficient remains stable when controlling for growth in non-infrastructure capital. The first-stage F-statistic exceeds 20, alleviating concerns about weak instruments. Columns (7) and (8) report reduced-form regressions of GDP growth on the tax share, yielding significantly negative coefficients.

Column (9) and (10) present the main IV results. The estimated output elasticity of infrastructure capital is around 0.185–0.189 and statistically significant. Adding non-infrastructure capital as a control variable improves precision without altering the magnitude.

The magnitude of the IV estimate compares well with the existing literature. The meta-analysis by Bom and Ligthart (2014) reports an average elasticity of 0.106. In subsamples that more closely match our context—regional or local government investment in core infrastructure over a long-run horizon—the average rises to 0.193. Our estimates align closely with this benchmark.

Table 2: Output Elasticity of Infrastructure Capital, IV Estimation

	1993-2000		2001-2010							
	OLS		OLS		OLS First Stage		OLS Second Stage		IV	
	$\Delta \log k_G$	$\Delta \log y$	$\Delta \log y$		$\Delta \log k_G$		$\Delta \log y$		$\Delta \log y$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Z_i	-0.153	-0.048			-0.476***	-0.473***	-0.090**	-0.087***		
	(0.200)	(0.050)			(0.105)	(0.104)	(0.041)	(0.032)		
$\Delta \log k_G$			0.117**	0.095**					0.189**	0.185***
			(0.051)	(0.044)					(0.080)	(0.065)
$\Delta \log k$				0.167***		0.184		0.183***		0.149***
				(0.044)		(0.127)		(0.050)		(0.049)
Constant	0.233***	0.088***	0.080***	0.065***	0.217***	0.196***	0.111***	0.091***	0.051***	0.041***
	(0.040)	(0.010)	(0.007)	(0.008)	(0.023)	(0.023)	(0.008)	(0.008)	(0.014)	(0.011)
R^2	0.026	0.078	0.316	0.475	0.307	0.328	0.268	0.467	0.271	0.407
Observations	210	210	300	300	300	300	300	300	300	300
F-stat					20.33	20.69				

Note: y, k_G, and k denote GDP per capita, infrastructure capital per capita, and non-infrastructure capital per capita, respectively. Z_i is the share of income taxes in 2000. All regressions include year fixed effects. Standard errors are clustered by 30 provinces.

The IV estimate is nearly twice the OLS estimate, suggesting a downward bias in OLS due to omitted variables. In particular, high-productivity provinces where local governments face higher distortions due to fiscal transfers may have weaker infrastructure incentives despite higher output. Meanwhile, the central government provides preferential policies and subsidies for investment in undeveloped regions. This introduces a negative correlation between unobserved distortions and infrastructure capital, biasing OLS estimates downward.

The IV estimate also exceeds the elasticity of 0.05 implied by infrastructure capital share, suggesting underinvestment relative to the social optimum. Despite frequent concerns of overinvestment by Chinese governments, infrastructure investment in China may remain suboptimal due to fiscal constraints. In 2010, total tax revenue amounted to only 14% of GDP, half of which was retained by the central government. Before 2009, local governments were legally forbidden to issue debt and borrow money from banks, forcing them to rely on land sales and local government financing vehicles (LGFVs)

to finance infrastructure investment. The central redistribution scheme further disincentivizes investment in richer provinces. In such settings, estimates based on capital shares likely understate true output elasticities.

Additional results of the robustness check are presented in Appendix B.3. Table B2 uses the fraction of unoccupied urban land as an additional instrument introduced by Su (2023) in the context of land finance. The estimates are robust and the overidentification test cannot reject joint validity. Figure B6 uses the Bartik-style instruments to jointly estimate the output elasticity of non-infrastructure. It supports the assumption of constant returns to scale.

Based on the IV estimation, I set $\alpha = 0.185$ in the quantitative analysis. $\theta = 0.315$ is set to match the 50% labor share in China.

7.2 Other Parameters

Thanks to the analytical tractability of the model, most of the structural parameters can be identified by linking the analytical expressions of the equilibrium allocation to the moments of their empirical counterparts. The calibration of the remaining parameters relies on data from 2013 to 2017, the most recent five-year data available from the NBS website. In Figures 1 and 5, output dispersion and the degree of progressivity become stable after 2013. Therefore, I treat this period as a new steady state. All variables are deflated to the same price level and averaged over this period to reduce the influence of transitory fluctuations and measurement error.

Policy parameters are estimated directly from the data. Specifically, the degree of progressivity ϕ is estimated using Equation (2) for every period, and the results are displayed in Figure 5. The output tax rate τ_y is computed as the ratio of aggregate government revenue to aggregate GDP. I assume that net central taxes are fully used to finance central government consumption. Hence, the share of central government consumption in GDP, g, equals the share of central taxes net of transfers to local governments in aggregate GDP. I abstract from the fact that a fraction of central government spending is financed by government debt.

Preferences over government consumption are inferred from the model's optimality condition. I assume that the central government optimally chooses the output tax

rate τ_y and the size of central government consumption g. Under this assumption, the preference weights $\chi^L = 0.230$ (for local government consumption) and $\chi^C = 0.050$ (for central government consumption) are inferred from Equation (20).

The region-specific productivity for each period is calculated as the Solow residual:

$$\log A_{it} = \log y_{it} - \alpha \log k_{Git} - \theta \log k_{it}$$

where y_{it}, k_{Git}, and k_{it} are per capita measurements.

In the quantitative analysis, I assume that firms face region-specific distortions on capital rent. The firm's problem (6) is modified as:

$$\max_{k_i, n_i} \pi_i = (1 - \tau_y) y_i - (1 + \tau_{ki}) R k_i - w_i n_i$$

This setting is consistent with previous literature on China's capital misallocation, such as Hsieh and Klenow (2009); Song, Storesletten and Zilibotti (2011); Brandt, Kambourov and Storesletten (2025). The capital wedge $(1 + \tau_{kit})$ summarizes distortions on capital rent, such as differential lending rates, collateral constraints, and the dominance of large state-owned enterprises (SOEs). This helps to reconcile capital allocation and the private capital-output ratio in the data. The capital wedge can be identified through the first-order condition:

$$\log(1 + \tau_{kit}) = \log\left(\frac{\theta(1 - \tau_y)}{R}\right) + \log\left(\frac{y_{it}}{k_{it}}\right)$$

I set $R = r + \delta = 0.18$ where I have in mind that the depreciation rate $\delta = 0.13$ from Song and Xiong (2023) and the real interest rate $r = 0.05$.

Finally, the parameter governing social preferences, γ, is inferred using Equation (26), under the assumption that the observed degree of progressivity reflects an optimal policy choice by the central government. The inferred γ is -0.279, indicating that the central government puts relatively higher Pareto weight on high-productivity regions. Table 3 lists all parameter values for the baseline model during the period from 2013 to 2017.

Table 3: Baseline Parameterization, 2013-2017

Parameter	Target	Value
A. Policy		
ϕ	Estimated using the HSV tax function (2)	0.485
τ_y	Share of aggregate fiscal revenue in GDP	0.294
g	Share of central taxes net of transfers in GDP	0.020
B. Technology		
α	Estimated using the pre-reform income share as IV	0.185
θ	50% labor share in China	0.315
C. Preference		
χ^L	Assume τ_y and g are optimally chosen	0.230
χ^C		0.050
γ	Assume ϕ is optimally chosen	-0.279
D. Distribution		
σ_a^2	Solow residual	0.099

7.3 Model Evaluation

I evaluate model performance by comparing untargeted moments generated from the model to their data counterparts. The model is designed to capture distortions in local fiscal policy and their implications for regional economic performance. I assess three types of cross-sectional moments. First, as shown in Proposition 1, progressivity attenuates the local fiscal response to productivity. I evaluate this using the cross-sectional relationships among output, infrastructure capital, and productivity. Second, I examine how redistribution narrows regional disparities and promotes regional convergence, using the distributional moments of output and capital. Third, I quantify effects on aggregate efficiency using aggregate moments. In particular, I calculate aggregate TFP relative to its theoretical efficient level. I treat the period from 2013 to 2017–when progressivity becomes stable–as the new stead state and compare moments during this period. I also assess moments during the post-reform transitional dynamics from 2004 to 2013.

I do not model distortions to local fiscal policy other than regional redistribution. Local governments face additional distortions affecting infrastructure investment. For example, Figure A3 documents cross-regional heterogeneity in the ratio of local fiscal revenue to GDP, even though statutory tax rates are uniform across regions (both in

fact and in the model). Before 2009, local governments were prohibited from issuing debt or borrowing from banks. In response, many created local government financing vehicles (LGFVs) to issue debt, using land as collateral to access credit (Chen, He and Liu, 2020). Meanwhile, the central government systematically subsidized inland and western regions to redirect capital toward underdeveloped areas and infrastructure sectors, particularly through the Great Western Development Program around 2000 and the Four-Trillion Stimulus in 2009 (Bai, Hsieh and Song, 2016). These policies introduced large distortions into local fiscal decisions. In addition, Song and Xiong (2023) document heterogeneity in local governors' career incentives that influenced infrastructure investment and growth. Accordingly, the model captures only a fraction of the distortions in local fiscal policy.

8 Model Implications for Fiscal Distortion, Output Dispersion, and Aggregate Productivity

This section documents the performance of the model in generating untargeted cross-sectional moments, resembling those observed in the data. I construct the average of provincial-level variables during the period of 2013 to 2017 and compute various statistics as data moments. To assess the role of progressivity in the new steady state, I compute the counterfactual moments with $\phi = 0$. Table 4 lists the main results. For robustness, I report the same statistics averaged over the whole sample period from 1998 to 2017. Results are displayed in Table D3 and the model continues to exhibit good performance. Additionally, I track the evolution of output dispersion and aggregate TFP growth during the transitional period from 2004 to 2013 in Figure 10, mirroring the dynamic patterns in Figure 1. Table 5 further reports the aggregate responses during the transitional dynamics.

8.1 Cross-Sectional Moments

Panel A of Table 4 focuses on the distortionary effects of progressivity by assessing the cross-sectional relationships among output, public capital, and productivity. Proposition 1 indicates that progressivity reduces the elasticities of output and public capital

Table 4: Untargeted Moments: Data, Model, and Counterfactual (2013-2017)

		Data	Model	Counterfactual $\phi = 0$
A. Cross-sectional moments				
1.	Coefficient of $\log y$ on $\log A$	1.679	1.696	2
2.	Coefficient of $\log k_G$ on $\log A$	0.813	0.873	2
3.	Coefficient of $\log(k_G/y)$ on $\log A$	-0.866	-0.822	0
4.	Coefficient of $\log y$ on $\log k_G$	0.434	0.515	1
B. Distributional moments				
5.	$\text{var}(\log y)$	0.171	0.188	0.262
6.	$\text{var}(\log k)$	0.200	0.168	0.219
7.	$\text{var}(\log k_G)$	0.196	0.050	0.262
C. Aggregate moments				
8.	$\log Y$	0.777	0.973	1.019
9.	$\log K_G$	0.269	0.893	1.033
10.	$\log \text{TFP}$	0.281	0.308	0.318
11.	$\log(\text{TFP}/\text{TFP}^*)$	-0.063	-0.036	-0.027

with respect to productivity. The data column of Panel A reports the coefficients from two main regressions. Rows 1 to 3 display the results of the following regression:

$$\text{outcome}_i = \beta_0 + \beta_a \log A_i + \beta_{\tau_k} \log(1 + \tau_{ki}) + u_i$$

The outcome variables are log output (Row 1), log infrastructure capital (Row 2), and log infrastructure capital-output ratio (Row 3). $\log(1 + \tau_{ki})$ is to control for the distortionary effects of the private capital wedge.

In Row 1, the elasticity of output with respect to productivity estimated from the data is 1.68, while the elasticity implied by the model, $1/(1-\theta-\alpha(1-\phi))$, is calibrated to be 1.70. In Row 2, the elasticity of infrastructure capital estimated from the data is 0.81, while the elasticity implied by the model, $(1-\phi)/(1-\theta-\alpha(1-\phi))$, is calibrated to be 0.87. The model effectively matches the responsiveness of output and infrastructure to productivity. With $\phi = 0$, both elasticities increase to 2.

In Row 3, the elasticity of the capital-output ratio with respect to productivity estimated from the data is -0.87. High-productivity regions have lower public capital-output ratios, indicating capital misallocation. The elasticity implied by the model from Equation (17) is $-\phi/(1 - \theta - \alpha(1 - \phi))$, which is calibrated to be -0.82. The

model effectively captures that redistribution distorts capital flow into low-productivity regions. When setting $\phi = 0$, this elasticity should be 0, in which case the capital-output ratio is uniform across regions.

Estimates in Rows 1 to 3 depend on the choice of the technology parameter values and the measurement of productivity. A more direct way is to examine the correlation between output and public capital. Row 4 reports the coefficient from the regression of log output on log infrastructure capital without controlling for capital wedges. The coefficient estimated from the data is 0.43, while the coefficient implied by the model, $1 - \phi$, is calibrated to be 0.52. The model effectively captures the cross-sectional relationship between output and infrastructure capital. When setting $\phi = 0$, this coefficient increases to 1.

8.2 Distributional Moments

Panel B of Table 4 focuses on regional inequality by assessing the distributional moments of output and capital. In Rows 5 and 6, the variances of log output and log private capital estimated from the data are 0.17 and 0.20, respectively. The model counterparts are 0.19 and 0.17, closely aligning with the data. These results are not by construction as I only include productivity and private capital wedges as heterogeneity and largely ignore distortions in fiscal policy. Consistent with this omission, the model understates the dispersion of public capital. While the variance of log public capital estimated from the data is 0.20, the model counterpart is 0.05. Nevertheless, the model still generates about one quarter of the observed dispersion.

In the new steady state, when eliminating progressivity, output dispersion will increase dramatically to 0.26. This implies that progressivity near 0.5 reduces output dispersion by more than one quarter. In the non-redistribution counterfactual scenario, public capital dispersion should be equal to output dispersion in theory, since no other fiscal distortions are modeled.

8.3 Aggregate Moments

Panel C of Table 4 examines the aggregate moments. The model generates overstated aggregate quantities relative to the data, especially for public capital. In Rows 8 and

9, the log of aggregate output and public capital estimated from the data are 0.78 and 0.27, respectively, while the model counterparts are 1.09 and 1.00. A key reason for this deviation is that the model omits fiscal distortions other than regional redistribution. Prior work and empirical evidence above indicate tight local budget constraints and other fiscal frictions that potentially depress aggregate public capital. Under the calibration, the public capital-output ratio should be close to 1 in an efficient allocation, whereas its data counterpart is only 0.6, indicating underinvestment.

Rows 10 and 11 present aggregate TFP. In Row 10, aggregate TFP is 0.281 in the data and 0.31 in the model. Again, as the model excludes other fiscal distortions, it likely underestimates capital misallocation and thus overstates aggregate efficiency. Removing progressivity increases aggregate TFP by only 1%. Row 11 compares the actual TFP to its efficient level, TFP^*. The efficient TFP is calculated from Equation (18) by setting $\phi = 0$. Efficiency loss in the model arises exclusively from cross-regional capital allocation. In the data, the aggregate efficiency is 6.3% lower than its efficient level, whereas the model accounts for over half of this gap. Eliminating progressivity reduces the efficiency loss by 0.9 percentage points.

8.4 Transitional Dynamics

Having shown that the model exhibits success in generating cross-sectional moments, I next evaluate its performance during the transitional dynamics. Figure 10 compares the evolution of the post-reform output dispersion and TFP growth in the data with the path simulated from the model, resembling the empirical motivation in Figure 1. In the top-left panel, the model matches both the level and the downward trend of output dispersion. In the top-right panel, the model simulates an annual convergence rate of 2.7%, closely aligning with the data estimate of 2.4%. In the bottom panel, the model reproduces the slowdown in TFP growth, although it somewhat overestimates aggregate TFP.

Table 5 summarizes the aggregate responses. The first column presents changes in key statistics between 2004 and 2013 in the data. The second column shows the model counterparts. The third column reports the counterfactual counterparts with progressivity fixed at the initial level 0.286.

Row 1 reports the change in the coefficient from a regression of log output on log infrastructure capital. In the data, the correlation between output and infrastructure capital becomes significantly less pronounced over time. The coefficient declines by 0.374. The model generates a smaller decline of 0.196, indicating that it can capture over half of the weakening correlation observed in the data. In the counterfactual scenario where progressivity remains fixed at its initial level, the model predicts no change in this correlation.

Rows 2 and 3 present results of regional convergence. In Row 1, the variance of log output falls by 0.09 during the transition, and the model replicates this decline. Holding progressivity at the initial level reduces the decline to 0.06. In Row 2, as shown before, the annual convergence rate is 2.4% in the data and 2.7% in the model. Fixing progressivity at the initial level lowers it to 2.1%. Taken together, the rise in progressivity explains one quarter of the rapid regional convergence after the tax reform.

Rows 4 and 5 present results regarding aggregate productivity. In Row 3, the aggregate TFP increases by 24% in the data and 25% in the model. Fixing progressivity at its initial period changes the overall TFP growth little. Row 4 reports that aggregate TFP relative to its efficient level declines by 2.4 percentage points in the data and by 1.7 percentage points in the model. Although omitting fiscal distortions, the model captures over 60% of worsening cross-regional capital misallocation. Fixing progressivity at its initial level does not alter the decline very much.

Rows 6 and 7 further report the change in aggregate output and public capital stock. Although the model does not perform well in matching the aggregate level of output and capital stock, it effectively tracks the change. Both output and public capital roughly doubled over the transitional period, and the model generates similar growth rates. Fixing progressivity at the initial level will increase annual output growth by 0.2 percentage points and infrastructure growth by 0.7 percentage points.

Overall, the quantitative analysis in this section indicates that the rise in progressivity largely compresses regional disparities and stimulates economic convergence at a modest efficiency cost. This provides an economic rationale for setting progressivity as high as 0.5. Based on the quantitative implications, the next section will calculate optimal progressivity. Appendix D extends the model by incorporating public capital

wedges to exactly match the data moments and replicates the counterfactual exercises. Conclusions are robust.

Figure 10: Output Dispersion, Regional Convergence and TFP Growth During the Transitional Dynamics between 2004 and 2013

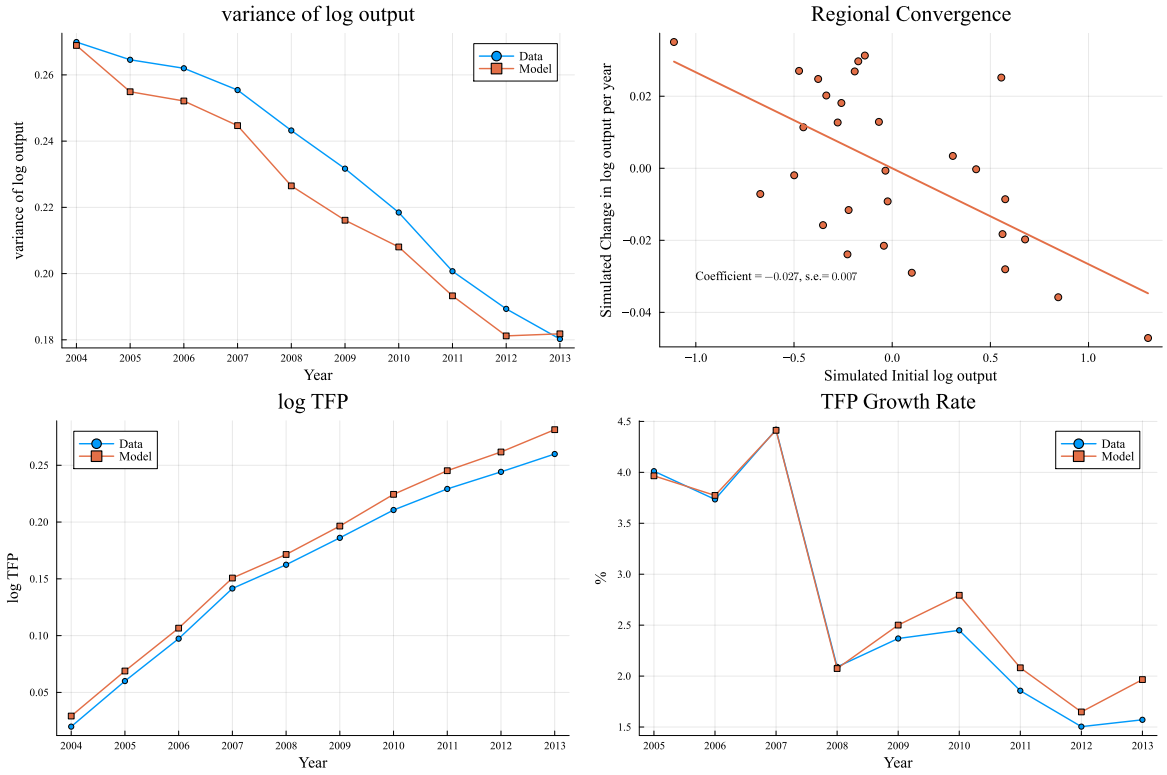


Table 5: Aggregate Responses during Transitional Dynamics: Data, Model, and Counterfactual (2004–2013)

	Data	Model	Counterfactual $\phi = \phi_{2004}$
1. Δ Coefficient of $\log y$ on $\log k_G$	-0.374	-0.196	0
2. $\Delta \text{var}(\log y)$	-0.090	-0.087	-0.063
3. Annual convergence rate	-2.446	-2.665	-2.112
4. $\Delta \log \text{TFP}$	0.240	0.252	0.256
5. $\Delta \log(\text{TFP}/\text{TFP}^*)$	-0.029	-0.017	-0.013
6. $\Delta \log Y$	0.864	0.885	0.906
7. $\Delta \log K_G$	1.096	1.128	1.188

9 Optimal Progressivity

Having quantified that the rise in progressivity substantially reduces regional disparities at a modest efficiency cost, this section studies the optimal degree of progressivity.

I start by decomposing the social welfare function to separate different motivations that discipline redistribution. I then compute optimal progressivity in various model scenarios to clarify the underlying mechanisms.

9.1 Decompose Social Welfare Function

Progressivity in recent years is approaching 0.5. What drives such a high degree of progressivity? This subsection answers this question by decomposing the social welfare. After substituting the optimality condition of g^* and τ_y^* into Equation (19), the social welfare $\mathcal{W}(\phi)$ can be expressed as a single-valued function of the degree of progressivity ϕ.

The top panel of Figure 11 plots the components of the social welfare as a function of progressivity. The y-axis represents the welfare change relative to the current redistribution system, measured as the percentage of consumption equivalents. The decomposition proceeds in three steps: starting from the representative region model in (1), I sequentially incorporate heterogeneity in regional productivity in (2), which is analogous to the utilitarian case, and the central government's social preference in (3), which rationalizes the current redistribution system.

In the representative region model (1), the only motivation for the redistribution system is to finance central government consumption. Given the relatively small calibrated value of the preference weight on central government consumption χ^C, the planner optimally chooses a mildly regressive scheme, with $\phi^{R*} = -0.071$. This outcome reflects weak incentives to collect central taxes when the welfare gains from central provision are limited. This scenario highlights the distortionary effect of progressivity that motivates local governments to shift fiscal spending from capital rent toward government consumption.

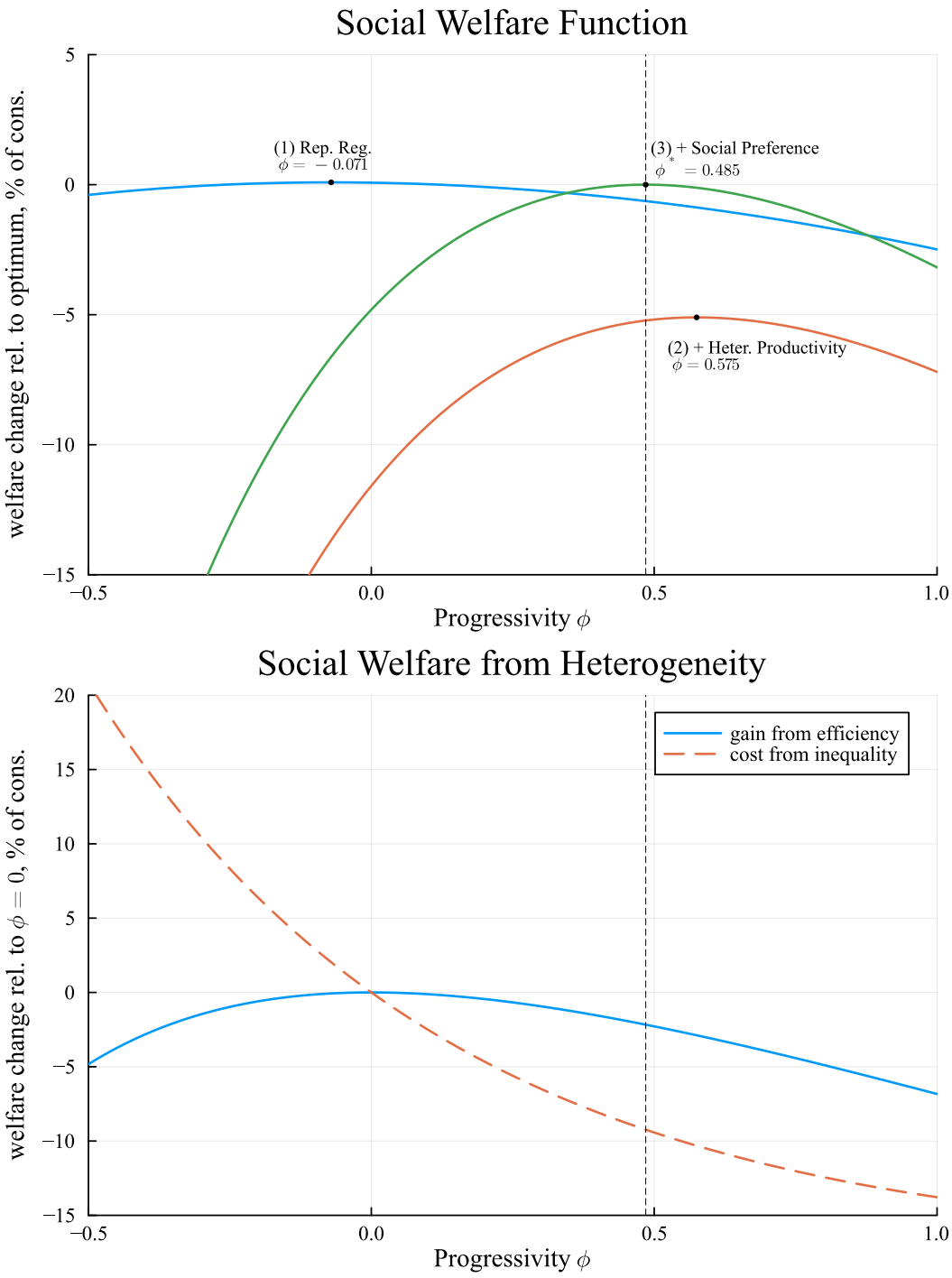
In scenario (2), introducing heterogeneity in regional productivity substantially raises the optimal progressivity to $\phi = 0.575$. Productivity is the only source of heterogeneity in the baseline environment, so the sharp increase reflects sizable welfare gains from narrowing regional disparities in fiscal capacity and output. To see this, the bottom panel of Figure 11 compares the welfare gain from efficiency and the welfare cost of inequality, relative to the non-redistribution scenario ($\phi = 0$). The welfare gain

from efficiency is Line (d) from Equation (19), maximized precisely at $\phi = 0$. The welfare cost of inequality is Line (e), which falls with progressivity. As a function of progressivity, the welfare gain curve is flatter than the cost curve, implying that raising progressivity from 0 can largely improve welfare by reducing regional inequality at a modest efficiency loss. Compared to the no-redistribution benchmark, reducing inequality delivers a welfare gain equivalent to 9.2% of consumption, while the associated welfare loss from inefficiency amounts to 2.2%. This scenario captures the welfare effects of reallocating resources from high- to low-productivity regions.

A utilitarian social planner will optimally choose a progressivity of 0.575, slightly above the actual progressivity of the current system, 0.485. This suggests that the Chinese central government places slightly higher Pareto weight on more productive regions. One possible justification is that more productive regions have larger populations, and thus receive greater Pareto weights. Even under the utilitarian assumption, the current system is nearly optimal in the sense that moving from the current system to the utilitarian optimum delivers negligible welfare gains, equivalent to 0.11% of consumption. Relative to the non-redistribution scenario, the current system yields a welfare gain equivalent to 4.8% of consumption.

Such a high degree of optimal progressivity in the utilitarian case is driven by several facts. First, central government spending is quite small in the new steady state. This is reflected both in Figure 2 and the calibrated value of χ^C. Under this circumstance, the central government has less motivation to collect central taxes, and therefore optimal progressivity in the representative-region economy is close to 0. Second, although output dispersion is declining over time, productivity dispersion–the only source of heterogeneity in the baseline model–is rising. This largely amplifies the welfare gain from reducing regional inequality and pushes the optimal progressivity up. Third, although the IV approach estimates a higher output elasticity of infrastructure capital, it remains relatively low compared to the output elasticity of non-infrastructure capital. The smaller elasticity mitigates the distortionary effects of progressivity. The conclusions are robust when capital wedges are added in Appendix D. The next subsection continues to explore these mechanisms by calculating optimal progressivity in various model scenarios.

Figure 11: Social Welfare as a Function of ϕ and Welfare from Heterogeneity



Note: The vertical dashed line marked progressivity of the current redistribution system.

9.2 Optimal Progressivity Under Different Parameter Values

This subsection explores optimal progressivity under different parameter values. The specification of the Pareto weight function given by Equation (11) nests several classic specifications. $\theta = 0$ corresponds to the utilitarian planner with equal weights on each region. $\theta \to \infty$ corresponds to the Rawlsian planner who seeks to maximize the welfare of the least productive region. $\theta = -1/(1-\theta-\alpha)$ corresponds to the laissez-faire planner with weights inversely proportional to the marginal utility from productivity in the non-redistribution economy.

Table 6 Column (1) summarizes the main results. Row 1 corresponds to the baseline model where the current redistribution system is rationalized by setting $\gamma = -0.279$. Row 2 corresponds to the utilitarian case with $\gamma = 0$. Row 3 corresponds to the Rawlsian case $\gamma \to \infty$. In order to maximize the welfare of the least productive region, the social planner should always set $\phi = 1$ to achieve the maximum redistribution. Row 4 corresponds to the laissez-faire case where the social planner does not want to redistribute. Therefore, the optimal ϕ is around 0. In Row 5, I set $\gamma = 0.5$ corresponding to a planner who puts relatively higher weights on less productive regions. The optimal progressivity is 0.741, significantly higher than that in the baseline and the utilitarian case. In Row 6, I set $\gamma = -0.5$, corresponding to a planner who puts higher weights on less productive regions, which reduces optimal progressivity to 0.416.

I also consider the optimal progressivity under alternative parameter values. The baseline model nests several special cases. $\chi^L = 0$ corresponds to the case where local government consumption is not valued by households. In this situation, local governments' fiscal decisions will not be distorted, as they spend all disposable revenue on renting capital. $\chi^C = 0$ corresponds to the case where central government spending is exogenous. The model with $\chi^C = 0$ no longer features the free-rider problem, in which case the central government does not need to provide extra motivation for local public capital anymore.

Columns (2)-(6) summarize the optimal progressivity for alternative parameter values. Compared to the baseline model, both $\chi^L = 0$ and $\chi^C = 0$ cases yield higher optimal progressivity. As discussed before, when $\chi^L = 0$, progressivity does not distort local government fiscal decisions. When $\chi^C = 0$, the central government does not have

to offer additional incentives for public capital. Both cases reduce the distortionary effects and increase optimal progressivity. In Column (4), larger productivity dispersion raises optimal progressivity, as it increases the marginal welfare gain from reducing regional inequality. Columns (5) and (6) display optimal progressivity when $\alpha = 0.05$ and $\alpha = 0.3$. I adjust θ accordingly to always make $1 - \alpha - \theta = 0.5$. $\alpha = 0.05$ yields higher optimal progressivity as it reduces distortionary effects. In an extreme case where $\alpha = 0$, all models with $\gamma > 0$ yield the optimal progressivity of 1. Conversely, optimal progressivity decreases when $\alpha = 0.3$, as distortionary effects are enlarged when public capital is more productive. Moreover, as Corollary 6 indicates, $\phi = 0$ can be rationalized in the laissez-faire case with $\gamma = -1/(1 - \theta - \alpha)$ and $\chi^C = 0$, which is precisely the result in Column (3) Row 4.

Table 6: Optimal Progressivity Under Alternative Values for Different Pareto Weights

γ	(1) Baseline	(2) $\chi^L = 0$	(3) $\chi^C = 0$	(4) Larger σ_a^2	(5) $\alpha = 0.05$	(6) $\alpha = 0.3$
1. -0.279 (Baseline)	0.485	0.577	0.508	0.531	0.714	0.379
2. 0 (Utilitarian)	0.575	0.695	0.601	0.633	0.838	0.450
3. ∞(Rawlsian)	1.000	1.000	1.000	1.000	1.000	1.000
4. $-\frac{1}{1-\alpha-\theta}$(Laissez-faire)	-0.007	0.000	0.000	-0.003	-0.005	-0.006
5. 0.5	0.741	0.928	0.774	0.827	1.000	0.582
6. -0.5	0.416	0.490	0.436	0.453	0.617	0.324

10 Conclusion

This paper examins the role of regional redistribution in shaping regional economic growth following China's tax reforms. The first major contribution is to provide novel empirical evidence on the evolution of the redistribution system, local fiscal responses, and their economic consequences. I model regional redistribution as a progressive tax-transfer scheme and document a dramatic rise in progressivity. Exploiting the 2002 tax reform as a natural experiment, I conduct an event study and causally identify that local governments responded to the central-local tax-sharing rule by reducing fiscal spending on infrastructure investment, thereby slowing economic growth.

Motivated by the empirical findings, the second contribution is to develop an analyt-

ically tractable general equilibrium framework to systematically study the implications of regional redistribution for economic growth and social welfare. The model captures two key distortions: a shift in fiscal spending from productive public capital to government consumption, and a reallocation of resources from high- to low-productivity regions. The calibrated model matches most of the key moments both in the new steady state and during the transitional dynamics. Quantitatively, the rise in progressivity largely reduces regional disparities and stimulates regional convergence at a mild cost of inefficiency. This provides a rationale for high progressivity in recent years.

Consistent with Song and Xiong (2023), this paper highlights the pivotal role of local governments in regional economic growth. I contribute by identifying regional redistribution as a key fiscal distortion and by causally estimating the output elasticity of infrastructure capital using an IV approach. The baseline model can generate over one quarter of the dispersion in infrastructure capital and over half of the efficiency loss. A more comprehensive analysis of local fiscal distortions remains an important avenue for future research.

The framework developed in this paper lends itself to several extensions. First, the analysis abstracts from household-level tax progressivity and focuses on intergovernmental redistribution. A natural next step is to compare redistribution across households and across regions within a unified framework. Second, the model's static structure facilitates analytical solutions but limits its ability to capture dynamic trade-offs. Embedding regional redistribution into a dynamic neoclassical growth model would enable richer quantitative analysis. Moreover, by equalizing infrastructure and public services across space, regional redistribution may reshape migration and agglomeration patterns. Future research can extend the framework to incorporate these spatial dynamics and study how fiscal transfers influence the geographic distribution of economic activity.

References

Adamopoulos, Tasso, Loren Brandt, Chaoran Chen, Diego Restuccia, and Xiaoyun Wei. 2024. "Land security and mobility frictions." *The Quarterly Journal of Economics*, 139(3): 1941–1987.

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia. 2022. "Misallocation, selection, and productivity: A quantitative analysis with panel data from China." *Econometrica*, 90(3): 1261–1282.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2024. "The intertemporal keynesian cross." *Journal of Political Economy*, 132(12): 4068–4121.

Autor, David H, David Dorn, and Gordon H Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American economic review*, 103(6): 2121–2168.

Bai, Chong-En, Chang-Tai Hsieh, and Yingyi Qian. 2006. "The return to capital in China."

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Michael Song. 2016. "The long shadow of a fiscal expansion." National Bureau of Economic Research.

Bai, Yan, Keyu Jin, and Dan Lu. 2024. "Misallocation under trade liberalization." *American Economic Review*, 114(7): 1949–1985.

Barro, Robert J. 1991. "Economic growth in a cross section of countries." *The quarterly journal of economics*, 106(2): 407–443.

Barro, Robert J. 1996. "Determinants of economic growth: A cross-country empirical study."

Barro, Robert J, and Xavier Sala-i Martin. 1992. "Convergence." *Journal of political Economy*, 100(2): 223–251.

Bartik, Timothy J. 1991. "Who benefits from state and local economic development policies?"

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

Boar, Corina, and Virgiliu Midrigan. 2022. "Efficient redistribution." *Journal of Monetary Economics*, 131: 78–91.

Bom, Pedro RD, and Jenny E Ligthart. 2014. "What have we learned from three decades of research on the productivity of public capital?" *Journal of economic surveys*, 28(5): 889–916.

Borella, Margherita, Mariacristina De Nardi, Michael Pak, Nicolo Russo, and Fang Yang. 2023. "The Importance of Modeling Income Taxes over Time: U.S. Reforms and Outcomes." *Journal of the European Economic Association*, 21(6): 2237–2286.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2025. "A practical guide to shift-share instruments." *Journal of Economic Perspectives*, 39(1): 181–204.

Bourguignon, François, and Amedeo Spadaro. 2012. "Tax–benefit revealed social preferences." *The Journal of Economic Inequality*, 10: 75–108.

Brandt, Loren, and Xiaodong Zhu. 2010. "Accounting for China's growth." IZA Discussion Papers.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2025. "Barriers to entry and regional economic growth in China." *Review of Economic Studies*, rdaf029.

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.

Caselli, Francesco, and Wilbur John Coleman II. 2001. "The US structural transformation and regional convergence: A reinterpretation." *Journal of political Economy*, 109(3): 584–616.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan. 2007. "Business cycle accounting." *Econometrica*, 75(3): 781–836.

Chen, Zhuo, Zhiguo He, and Chun Liu. 2020. "The financing of local government in China: Stimulus loan wanes and shadow banking waxes." *Journal of Financial Economics*, 137(1): 42–71.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski.

2024. "The political development cycle: The right and the left in people's republic of china from 1953." *American Economic Review*, 114(4): 1107–1139.

David, Joel M, and Venky Venkateswaran. 2019. "The sources of capital misallocation." *American Economic Review*, 109(7): 2531–2567.

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.

Ferriere, Axelle, Philipp Grübener, Gaston Navarro, and Oliko Vardishvili. 2023. "On the optimal design of transfers and income tax progressivity." *Journal of Political Economy Macroeconomics*, 1(2): 276–333.

Fleck, Johannes, Jonathan Heathcote, Kjetil Storesletten, and Giovanni L Violante. 2025. "Fiscal Progressivity of the US Federal and State Governments." National Bureau of Economic Research.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez De Silanes, and Andrei Shleifer. 2014. "Growth in regions." *Journal of Economic growth*, 19: 259–309.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review*, 110(8): 2586–2624.

Guerreiro, Joao, Sergio Rebelo, and Pedro Teles. 2022. "Should robots be taxed?" *The Review of Economic Studies*, 89(1): 279–311.

Hall, Robert E, and Charles I Jones. 1999. "Why do some countries produce so much more output per worker than others?" *The quarterly journal of economics*, 114(1): 83–116.

Hanushek, Eric A, and Ludger Woessmann. 2008. "The role of cognitive skills in economic development." *Journal of economic literature*, 46(3): 607–668.

Hao, Tongtong, Ruiqi Sun, Trevor Tombe, and Xiaodong Zhu. 2020. "The effect of migration policy on growth, structural change, and regional inequality in China." *Journal of Monetary Economics*, 113: 112–134.

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.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics*, 124(4): 1403–1448.

Jones, Charles I. 2016. "The facts of economic growth." In *Handbook of macroeconomics*. Vol. 2, 3–69. Elsevier.

Kremer, Michael, Jack Willis, and Yang You. 2022. "Converging to convergence." *NBER macroeconomics annual*, 36(1): 337–412.

Lan, Xiaohuan. 2021. *Inside the System: China's Government and Economic Development.* Beijing: Century Wenjing Publishing Company.

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.

Liu, Kegu, and Kang Jia, ed. 2008. *Thirty Years of China's Fiscal and Tax Reform: Personal Experience and Review.* Beijing:Economic Science Press.

Lockwood, Benjamin B, and Matthew Weinzierl. 2016. "Positive and normative judgments implicit in US tax policy, and the costs of unequal growth and recessions." *Journal of Monetary Economics*, 77: 30–47.

Mankiw, N Gregory, David Romer, and David N Weil. 1992. "A contribution to the empirics of economic growth." *The quarterly journal of economics*, 107(2): 407–437.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The economic*

journal, 37(145): 47–61.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic dynamics*, 11(4): 707–720.

Rodrik, Dani. 2013. "Unconditional convergence in manufacturing." *The quarterly journal of economics*, 128(1): 165–204.

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. 2011. "Growing like china." *American economic review*, 101(1): 196–233.

Su, Yang. 2023. *Local government financial constraint and spending multiplier in China.* The University of Chicago.

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, migration, and productivity: A quantitative analysis of China." *American Economic Review*, 109(5): 1843–1872.

Villamil, Anne, Xiaobing Wang, and Ning Xue. 2021. "A political foundation of public investment and welfare spending." *Journal of public economic theory*, 23(4): 660–690.

Zhou, Feizhou. 2012. *Fiscal Relations and Local Government Behavior.* Shanghai:Shanghai Sanlian Bookstore.

A Data, Measurement, and Additional Stylized Facts

A.1 Data and Measurement

All data used in this paper come from the published statistical yearbooks and the official website of the National Bureau of Statistics (NBS), https://data.stats.gov.cn/easyquery.htm?cn=C01. The sample period ranges from 1997 to 2017.

Data on national and regional GDP, population, price index, government budgetary revenue and budgetary spending, tax categories, the fixed capital formation, and government consumption come from the NBS website.

Data on central and local taxes by region comes from the Taxation Yearbook of China.

Data on transfers from the central government, non-budgetary revenue comes from the Finance Yearbook of China.

Data on investment comes from the Statistical Yearbook of the Chinese Investment in Fixed Assets (SYCIFA) and the China Statistical Yearbook (CSY).

In order to measure progressivity using Equation (2), x represents the pre-redistribution government revenue. I follow Song and Xiong (2023) and define pre-redistribution revenue as the sum of taxes, non-tax revenue, non-budgetary revenue, and land sales. $\tilde{x}$ represents the post-redistribution government revenue, defined as x minus central taxes plus net transfers from the central government. Both x and $\tilde{x}$ are normalized by population size and by national GDP per capita.

I follow Song and Xiong (2023) to construct the regional infrastructure investment and infrastructure capital. Infrastructure investment includes investment in (1) Production and supply of electricity, gas, and water; (2) Management of water conservancy, environment, and public facilities; and (3) Transportation. SYCIFA and CSY only provide provincial and sectoral information on investment in fixed assets instead of fixed capital formation. Therefore, we assume the share of each sectoral capital formation is the same as the share in investment in fixed assets. In 1992-1995 and 1999-2001, CYS and SYFAI only provide investment in capital construction and innovation. We assume the share of the corresponding sector is constant. See Song and Xiong (2023) for details.

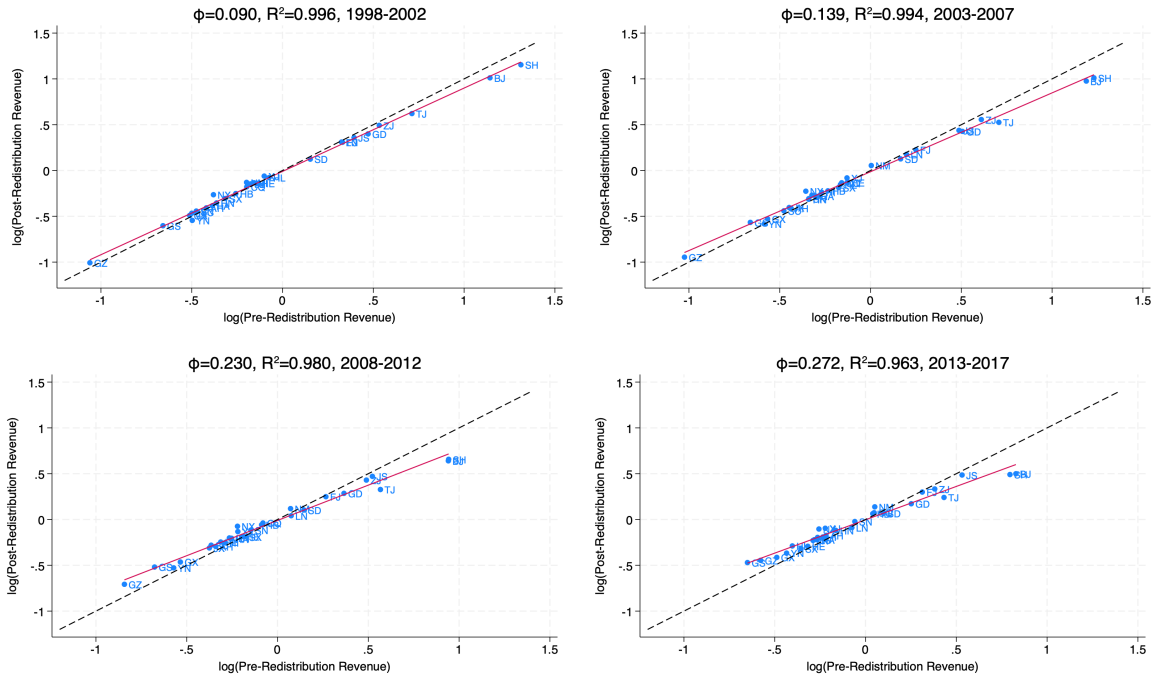
A.2 Measuring Progressivity

I alternatively measure progressivity using Equation (2). x is defined as GDP per capita. $\tilde{x}$ is defined as GDP per capita central taxes per capita plus net transfers from the central government per capita. Again, both x and $\tilde{x}$ are normalized by national GDP per capita to ensure comparability. We display the scatter plots in Figures A1 and the estimated results in A2. Results are more robust when measured using GDP.

A.3 Additional Empirical Evidence on Local Fiscal Conditions

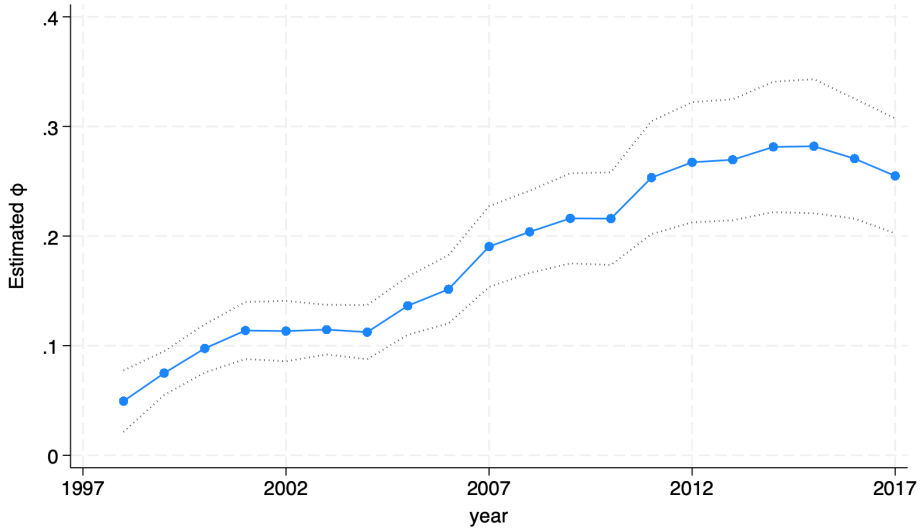
Here, I display pre- and post-reform local fiscal revenue in the share of regional GDP in Figure A3.

Figure A1: The Scatter Plot of log Post-Redistribution Local Government Revenue Against Pre-Redistribution Local Government Revenue.



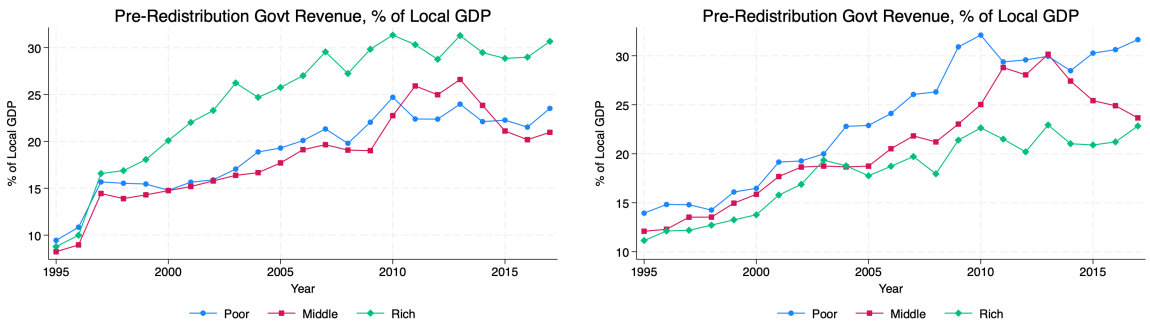
Note: These figures plot the log of (GDP - central taxes + net transfers) capita against the log of GDP per capita in different periods. To make the data comparable annually, we divided both variables by national GDP per capita and calculate the average within each period. We exclude Tibet and Qinghai in our sample when estimating ϕ.

Figure A2: The Estimated Progressivity Over Time



Note: Each dot presents the OLS estimation of ϕ using Equation (2) for each year. The dependent variable is log of (GDP - central taxes + net transfers) capita and the log of GDP per capita. The dotted line represents the 95% confidence interval. We excluded Tibet and Qinghai in our sample.

Figure A3: Pre- and Post-Redistribution Fiscal Revenue in the Share of Local GDP, Grouped by GDP per capita



B Additional Empirical Results of the Event Study and IV

B.1 Correlates of the Share of Income Taxes

I begin by examining whether the share of income taxes prior to the reform is systematically related to other observable local characteristics. Table B1 reports the results from a series of cross-sectional regressions, where the dependent variable is the share of income taxes in 2000 and the covariates are measured before the reform.

Column (1) regresses the instrument on provincial levels of GDP, infrastructure capital, non-infrastructure capital, and employment. None of these variables are significantly correlated with the instrument, suggesting that the level of economic development does not predict exposure to the reform.

Column (2) examines the sectoral composition of GDP. The value-added shares of the secondary and tertiary sectors are not significantly related to the instrument, implying that cross-province differences in industrial structure are unlikely to violate the exclusion restriction.

Column (3) considers fiscal capacity by including government revenue and government consumption. Again, no evidence shows that these variables are correlated with the share of income taxes.

Column (4) adds indicators of economic openness, including FDI and total imports and exports. These are also uncorrelated with the instrument.

To further explore potential spillover concerns, we test whether a province's income tax share is correlated with the characteristics of its neighbors. In Columns (5)–(8), we define the nearest province (NP) using two criteria: the province sharing the longest land border (Columns 5–6), and the province with the smallest centroid distance (Columns 7–8). In both cases, we test whether the share of income taxes is related to either the NP's instrument or its level of economic development. The results consistently show that neither the instrument nor the development level of the nearest province predicts a region's instrument value. These findings suggest that spatial spillovers are unlikely to drive the identification.

Taken together, the evidence in Table B1 indicates that our instrument is uncor-

related with a wide range of potential confounders. This supports the plausibility of the exclusion restriction and strengthens the case for using income tax share as an exogenous source of variation in local government investment incentives.

Table B1: Correlates between the Share of Corporate Income Taxes and Characteristics before the Tax Reform

	Dependent Variable: the Share of Income Taxes in 2000, Z_i							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	0.022							
	(0.050)							
Infra. capital	0.013							
	(0.019)							
Non-infra. capital	-0.009							
	(0.034)							
Employment	-0.004							
	(0.012)							
Value-added Share: Secondary		0.029						
		(0.110)						
Value-added Share: Tertiary		0.322						
		(0.267)						
Govt Revenue			0.012					
			(0.034)					
Govt Consumption			0.019					
			(0.019)					
FDI				0.007				
				(0.010)				
Import & Export				-0.002				
				(0.011)				
Z_j, NP					-0.006	-0.065	0.022	-0.038
					(0.342)	(0.360)	(0.149)	(0.171)
Infra. Capital, NP						0.005		-0.003
						(0.020)		(0.022)
GDP, NP						0.008		0.015
						(0.026)		(0.023)
Constant	0.257*	0.045	0.271***	0.157**	0.169***	0.208***	0.165***	0.188***
	(0.132)	(0.115)	(0.098)	(0.059)	(0.050)	(0.074)	(0.023)	(0.054)
R^2	0.119	0.184	0.102	0.038	0.000	0.009	0.001	0.019

Note: Each column reports results of a single regression of the IV on regional characteristics before the tax reform. The sample period covers 1993 to 2000. Years fixed effects are controlled. Standard errors are clustered by 30 provinces. *** p<0.01, ** p<0.05, * p<0.1.

B.2 Additional Results of the Event Study

B.2.1 Potential Effects of the Tax Reform on Other Economic Variables

One pitfall of the DID estimation is that the observed treatment effects on GDP and infrastructure capital are confounded by concurrent changes in other variables. To address this, I estimate Equation (4) using a range of alternative outcomes. Figures B4 summarize the results.

A primary concern is that the tax reform may have affected other fiscal instruments, thereby violating the exclusion restriction. The top two rows of Figures B4 examine the local fiscal response. I define pre-redistribution government revenue as total fiscal revenue collected within the province (including both local and central taxes), and post-redistribution revenue as pre-redistribution revenue net of central taxes plus transfers from the central government. While the reform significantly reduces post-redistribution government revenue (consistent with the central revenue reallocation), it does not affect pre-redistribution revenue, suggesting that pro-growth policies like tax cuts are unlikely to have changed.

The second row shows no significant changes in local government consumption or debt. Facing tighter budget constraints, local governments reduced investment and reallocated spending toward consumption. As a result, government consumption remained stable despite declining revenue. This helps to disentangle the effects of public capital from government consumption, which are often conflated in the fiscal multiplier estimates. A similar offsetting mechanism explains the stability in debt. While tighter constraints could have increased borrowing, the drop in investment demand mitigated the pressure. These patterns support the exclusion restriction by indicating that the reform primarily affected fiscal behavior through investment, with limited spillovers to other channels.

A further concern is that the tax reform may have affected other production inputs. The third row shows no immediate post-reform effects on non-infrastructure capital or employment.[7] However, after 2010, provinces with higher income tax shares experience a decline in non-infrastructure capital, likely reflecting the crowd-in effects of infrastructure investment on private capital in the long run. While this constitutes an indirect channel through which infrastructure affects output, it may upwardly bias the estimated output elasticity. To address this, in Section **??**, I restrict the estimation sample to 2001–2010 when non-infrastructure capital is unaffected, and later use

[7]The employment data published by the NBS are considered less reliable, partly due to political sensitivity surrounding unemployment. For instance, reported unemployment rates remain unusually stable even during periods of economic recession. In addition, detailed provincial-level employment data are sometimes missing. Therefore, we use per capita measures instead of per worker to account for differences population size and employment. This data unreliability may also contribute to the unusual pattern in which infrastructure capital and output decrease with the income tax share, while reported employment rises, though the increase is not statistically significant.

a Bartik-style instrument to estimate the output elasticities of both capital types over the full sample.

Finally, the bottom panel shows no significant effects on FDI, imports, or exports.

B.2.2 Placebo test

Having established the treatment effects of the 2002 tax reform, I further assess whether these effects are specific to the targeted tax categories. To do so, I conduct a placebo test by replacing the treatment variable Z_i in Equation (4) with the shares of other major tax categories. Figure B5 presents the results.

First, the shares of value-added taxes, sales taxes, and excise taxes show no significant effects on infrastructure capital during the reform period. By contrast, both corporate and personal income tax shares have significant negative effects post-reform. While the magnitudes are similar, the effect of personal income taxes is statistically weaker. This is expected, as personal income tax—though included in the 2002 reform—was much smaller in scale than corporate income tax, as illustrated in Figure 3. Thus, the variation in the corporate income tax share remains the primary source of identification.

B.3 Additional Results of IV Estimation

B.3.1 Robustness Check: Comparison with Su (2023)

A recent study Su (2023) proposes an alternative instrument for local government spending: the initial share of unoccupied land in urban areas prior to the expansion of land sales. In the 2000s, as the central government claimed a large share of tax revenue, local governments became increasingly budget constrained. To compensate, the central government permitted them to finance expenditures by selling land-use rights. Land sales therefore became a key source of local public finance in China, especially for infrastructure. When land is unoccupied, local governments avoid paying compensation to occupants, making land sales less costly and more profitable. A higher fraction of unoccupied land thus relaxes fiscal constraints and facilitates higher spending. Therefore, Su (2023) uses this instrument to estimate the local fiscal multiplier.

Figure B4: The Treatment Effects of the Tax Reform on Other Economic Variables, DID Estimation

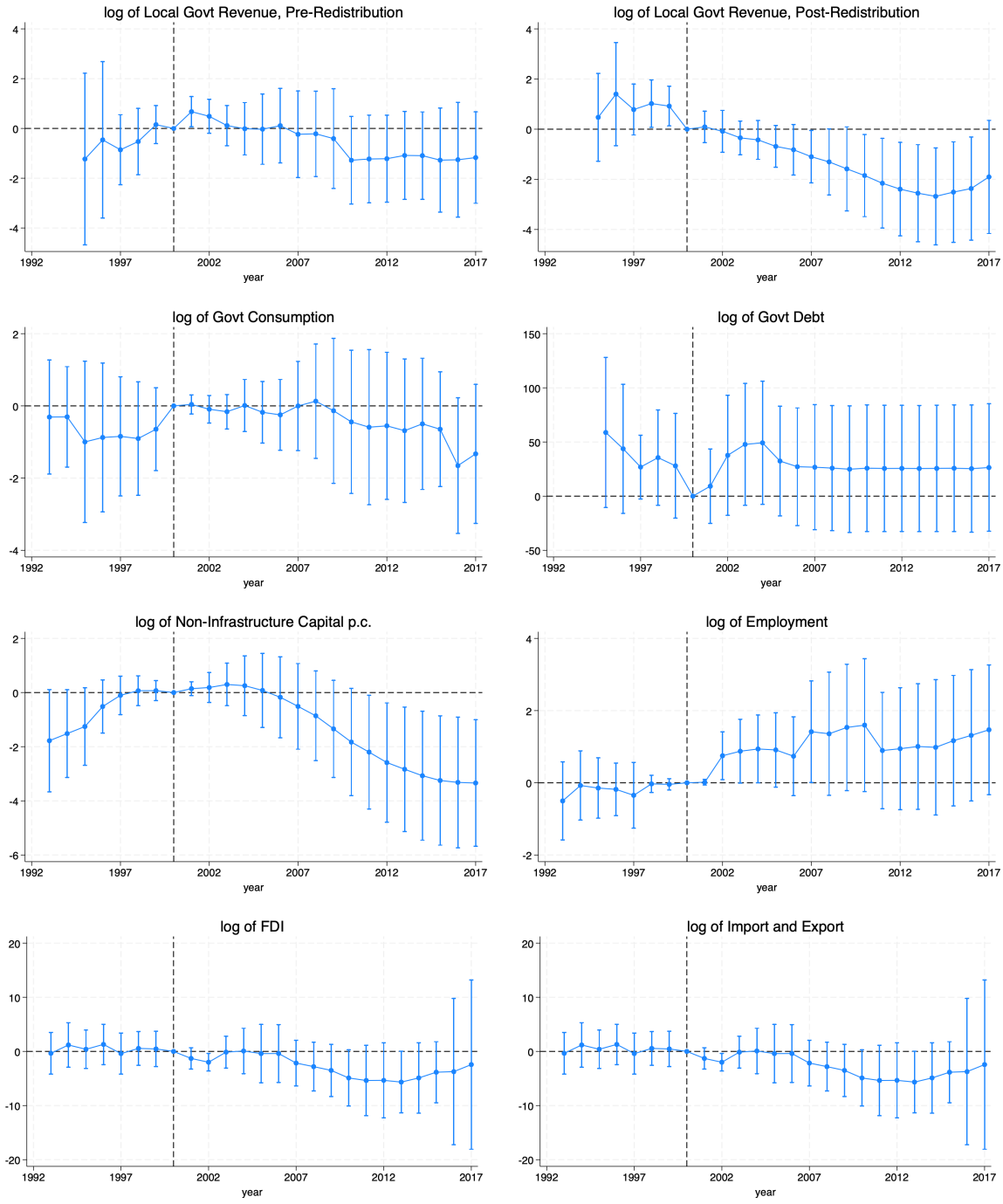
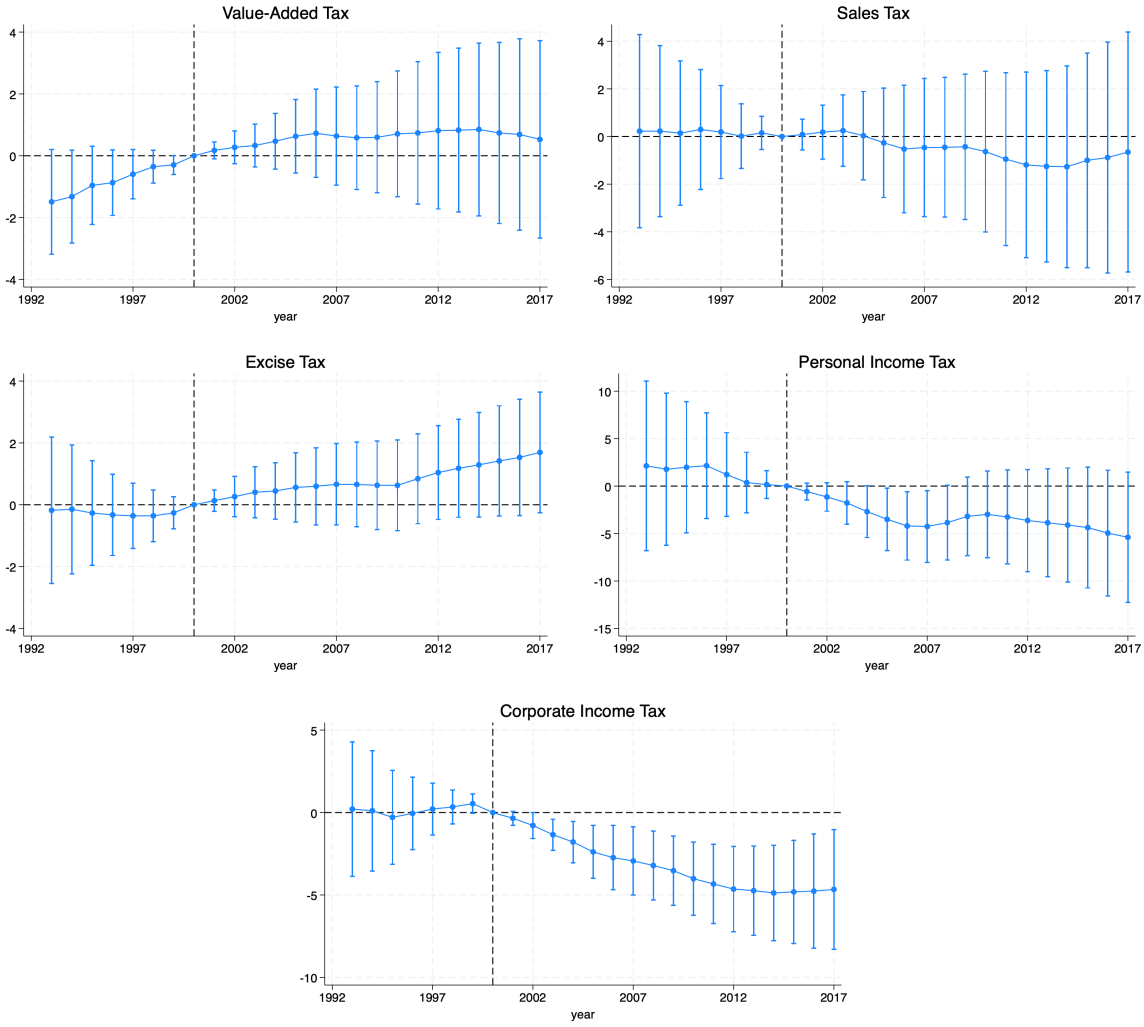


Figure B5: Placebo Test: the Treatment Effects of the Share of Other Tax Categories on Infrastructure Capital, DID Estimation



Note: These figures display the estimated coefficients and the 95% confidence intervals from Equation (4), using log infrastructure capital as the outcome variable. The treatment variable Z_i is replaced by the share of each respective tax category.

Importantly, however, budget data show that net profits from land sales are used almost exclusively for infrastructure construction.[8] Land sale revenues are generally not used to finance purchasing public goods. Therefore, I argue that the fraction of unoccupied land is more appropriately viewed as an instrument for infrastructure investment rather than for broader government spending.

I include the pre-reform fraction of unoccupied land as an additional instrument to test robustness. Table B2 presents the results. As a first step, Columns (1) and (2) verify the exclusion restriction. The fraction is uncorrelated with changes in government consumption and non-infrastructure capital, consistent with its narrow channel of influence through infrastructure spending.[9]

Columns (3) and (4) report the first-stage regressions. The instrument is strongly predictive of infrastructure capital growth: a one percentage point increase in the fraction of unoccupied land raises infrastructure capital growth by 0.118 percentage points. Adding our original instrument, the share of income taxes, does not alter the coefficient or its significance. However, the first-stage F-statistics (11.6 and 18.4) are somewhat lower than in the baseline specification, suggesting the new instrument is weaker than our original instrument.

Column (5) and (6) report the results of IV estimation. I use only the fraction of unoccupied land in Column (5) and use both instruments in Column (6). The estimated coefficients are ranging from 0.168 to 0.181, which is very close to our estimate of 0.185 in Table 2. As we have two instruments in hand, we can use the overidentification test. The Hansen J statistic is 0.029 (p=0.864), indicating that we cannot reject the hypothesis of valid instruments. This supports the joint use of both instruments in identifying the causal effect of infrastructure capital on output.

B.3.2 Returns to Scale

A number of studies assume that production functions with infrastructure capital exhibit increasing returns to scale.[10] We assess whether this assumption holds empirically.

[8] See "The Budget and Final Account of Local Government-Managed Fund" in Finance Yearbook of China (2010–2014). These are the only years in which the disaggregated land fund account is available.

[9] For additional analysis and results on this instrument, see Su (2023).

[10] See, for example, **??**Leeper, Walker and Yang (2010); **?**); **?**); Song and Xiong (2023). It is common to assume constant returns to scale in private inputs (labor and non-infrastructure capital), with

Table B2: Robustness Check: Estimating the Output Elasticity Using the Pre-reform Fraction of Unoccupied Land as an Additional Instrument Variable, 2001-2010

	OLS		OLS First Stage		IV	
	$\Delta \log G$	$\Delta \log K$	$\Delta \log K_G$		$\Delta \log Y$	
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of Unoccupied Land	0.051	-0.017	0.115***	0.090***		
	(0.030)	(0.023)	(0.034)	(0.028)		
Share of Income Taxes				-0.383***		
				(0.089)		
$\Delta \log K_G$					0.168*	0.181***
					(0.092)	(0.059)
$\Delta \log K$			0.240	0.222*	0.153***	0.149***
			(0.148)	(0.127)	(0.052)	(0.048)
Constant	0.092***	0.119***	0.046	0.126***	0.044***	0.042***
	(0.017)	(0.011)	(0.029)	(0.034)	(0.015)	(0.010)
R^2	0.255	0.291	0.311	0.437	0.431	0.413
First Stage F-stat			11.645	18.391		
Overidentification Test p-value						0.864

Note: This table reports robustness checks using the initial share of unoccupied urban land as an instrument. Y, K_G, G, and K denote GDP, infrastructure capital, government consumption, and non-infrastructure capital, respectively. Column (5) uses only the unoccupied land share as an instrument; Column (6) includes both the land share and the income tax share. Observations are 30 provinces for 10 years.

To test returns to scale, we must estimate the output elasticity of non-infrastructure capital alongside that of infrastructure capital, which requires additional instruments.

I construct Bartik-style instruments for non-infrastructure capital growth. Total investment comprises components such as agriculture, mining, manufacturing, construction, real estate, transportation, and others. I use the earliest available data—sectoral investment shares in 1996—to construct a vector of initial exposure shares. This approach follows Bartik (1991), Autor, Dorn and Hanson (2013), and Goldsmith-Pinkham, Sorkin and Swift (2020).[11] The key identifying assumption is that the 1996 investment structure is exogenous and predetermined, reflecting long-standing regional characteristics rather than endogenous responses. National investment patterns after 2000 were driven by some sector-specific shocks. For instance, the WTO accession spurred export-oriented manufacturing, while the 2008–2009 stimulus disproportionately boosted infrastructure and real estate. Provinces with different initial sectoral structures thus

infrastructure capital entering the production function through TFP with diminishing returns.

[11] Goldsmith-Pinkham, Sorkin and Swift (2020) and Borusyak, Hull and Jaravel (2025) show that the 2SLS estimator using a Bartik instrument is numerically equivalent to a GMM estimator with the initial shares as instruments and a weight matrix based on national sectoral growth. Identification comes from the exogeneity of the initial shares, which reflect differential exposure to common shocks.

experienced varying exposure to these national shocks, validating the use of sector shares as instruments.

I estimate the elasticities year by year and report results in Figures B6.[12] The output elasticity of non-infrastructure capital is slightly larger than that of infrastructure capital. Once we control for both inputs, the estimated elasticity of infrastructure capital no longer rises above 0.4, as seen in Figure **??**. On average, the elasticity of infrastructure capital is around 0.2, while that of non-infrastructure capital is slightly below 0.3.

Existing studies often set the labor elasticity at 0.5 to match the 50% labor share in China (Bai, Hsieh and Qian, 2006; Song, Storesletten and Zilibotti, 2011). Thus, we can test for increasing returns by examining whether the sum of capital elasticities exceeds 0.5. The bottom panel of Figures B6 shows that, during 2008–2015, the sum is statistically indistinguishable from 0.5, suggesting constant rather than increasing returns to scale.

C Proof

This Appendix section provides all the results and proofs of the main part of this paper. I formulate the complete expression for the quantitative model with capital wedges. The basline model in Sections 4, 5, and 6 is a special case when $\tau_{gi} = \tau_{ki} = 0$ for all i. Also, $\sigma_g^2 = \sigma_k^2 = \sigma_{ag} = \sigma_{ak} = \sigma_{kg} = 0$ accordingly.

C.1 Proofs of Proposition 1, Corollaries 1, 2, 3, 4

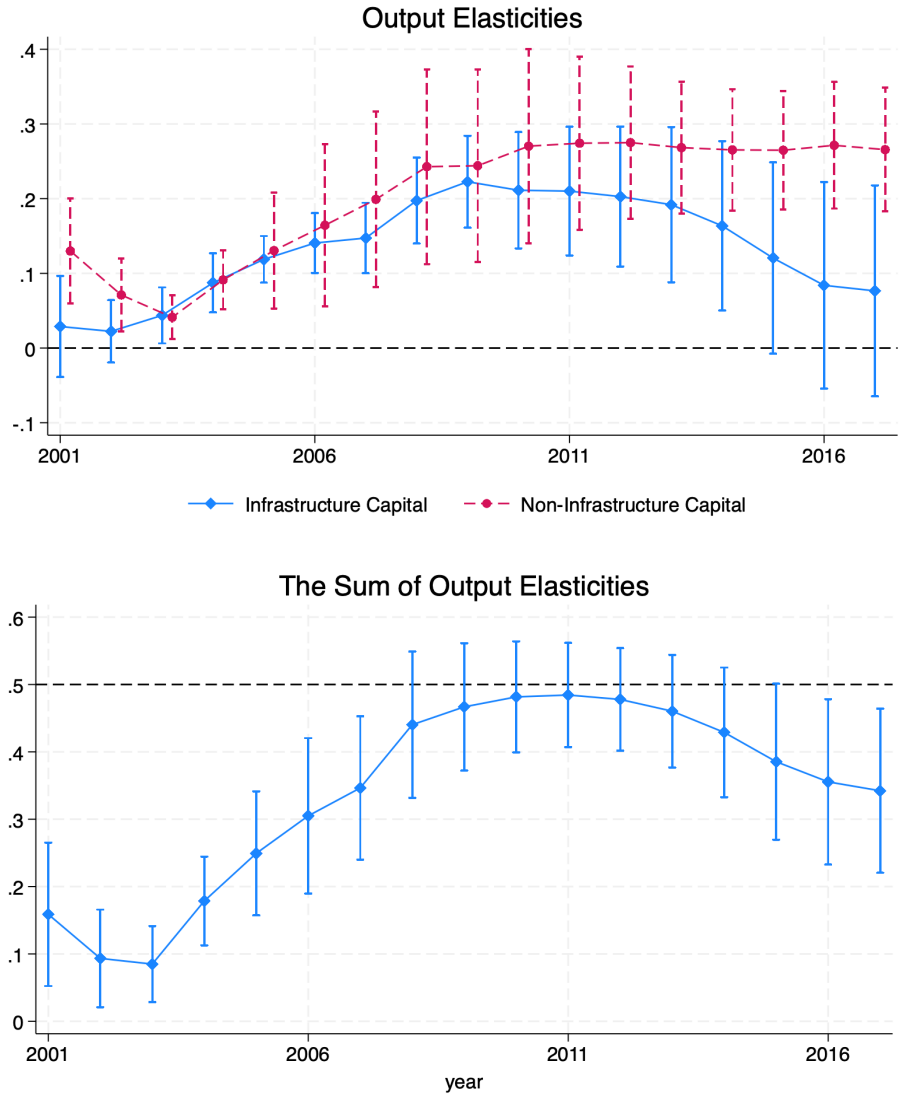
Before we formally prove Proposition 1, it is necessary to solve the firm's problem.

Plugging the production function Equation (5) into the firm's problem Equation (6) and taking first-order conditions with respect to k_i and n_i yields

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

[12] The full-sample regression yields a first-stage F-statistic of 17.74. The Hansen J statistic is 15.47 (p = 0.28), indicating that the overidentifying restrictions are not rejected.

Figure B6: Output Elasticities of Infrastructure and Non-infrastructure Capital



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

$$\frac{\theta}{1-\alpha-\theta}\frac{n_i}{k_i} = \frac{(1+\tau_{ki})R}{w_i} \quad \Rightarrow \quad k_i = \frac{\theta}{1-\alpha-\theta}\frac{w_i}{(1+\tau_{ki})R} \tag{C1}$$

Plugging (C1) 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

$$\begin{aligned}
k_i &= \left(\frac{\theta(1-\tau_y)}{(1+\tau_{ki})R}\right)^{\frac{1}{1-\theta}}A_i^{\frac{1}{1-\theta}}k_{Gi}^{\frac{\alpha}{1-\theta}} \\
y_i &= \left(\frac{\theta(1-\tau_y)}{(1+\tau_{ki})R}\right)^{\frac{\theta}{1-\theta}}A_i^{\frac{1}{1-\theta}}k_{Gi}^{\frac{\alpha}{1-\theta}} \\
\pi_i &= \alpha\left(\frac{\theta}{(1+\tau_{ki})R}\right)^{\frac{\theta}{1-\theta}}(1-\tau_y)^{\frac{1}{1-\theta}}A_i^{\frac{1}{1-\theta}}k_{Gi}^{\frac{\alpha}{1-\theta}}
\end{aligned}$$

Corollary 1 is obvious as ϕ and λ do not directly appear in the firm's allocation given by Equation (7).

Plugging $c_i = w_i + \pi_i$ and y_i given by Equation (7) into the local government's problem, replacing G_i using the budget constraint, and taking the first-order condition with respect to k_{Gi} yields

$$k_{Gi} = \left(\frac{\alpha(1+\chi^L(1-\phi))\lambda}{(\alpha+\chi^L(1-\theta))(1+\tau_{gi})}\right)^{\frac{1-\theta}{1-\theta-\alpha(1-\phi)}}\left(\frac{1}{R}\right)^{\frac{1-\theta\phi}{1-\theta-\alpha(1-\phi)}}\left(\frac{\theta(1-\tau_y)}{(1+\tau_{ki})}\right)^{\frac{\theta(1-\phi)}{1-\theta-\alpha(1-\phi)}}\tau_y^{\frac{(1-\phi)(1-\theta)}{1-\theta-\alpha(1-\phi)}}A_i^{\frac{1-\phi}{1-\theta-\alpha(1-\phi)}} \tag{C2}$$

which also gives the local output

$$y_i = \left(\frac{\alpha(1+\chi^L(1-\phi))\lambda}{(\alpha+\chi^L(1-\theta))(1+\tau_{gi})}\right)^{\frac{\alpha}{1-\theta-\alpha(1-\phi)}}\left(\frac{1}{R}\right)^{\frac{\alpha+\theta}{1-\theta-\alpha(1-\phi)}}\left(\frac{\theta(1-\tau_y)}{(1+\tau_{ki})}\right)^{\frac{\theta}{1-\theta-\alpha(1-\phi)}}\tau_y^{\frac{(1-\phi)\alpha}{1-\theta-\alpha(1-\phi)}}A_i^{\frac{1}{1-\theta-\alpha(1-\phi)}} \tag{C3}$$

Since A_i, $(1+\tau_{ki})$, and $(1+\tau_{gi})$ are log-normally distributed, y_i is also log-normally distributed.

$$\begin{aligned}
\mathbb{E}[\log y_i] &= \frac{1}{1-\theta-\alpha(1-\phi)}\left(\frac{\alpha}{2}\sigma_g^2 + \frac{\theta}{2}\sigma_k^2 - \frac{\sigma_a^2}{2}\right) + \log\tilde{\Omega}^y \\
\mathrm{var}(\log y_i) &= \frac{1}{(1-\theta-\alpha(1-\phi))^2}\left(\alpha^2\sigma_g^2 + \theta^2\sigma_k^2 + \sigma_a^2\right)
\end{aligned}$$

According to the property of the log-normal distribution, we have

$$\log(\mathbb{E}[y]) = \mathbb{E}[\log y_i] + \frac{1}{2}\text{var}(\log y_i)$$

Then, using the central government's budget constraint, we can calculate λ in the equilibrium. Define

$$\begin{aligned}\Omega^\lambda = &\frac{\phi\left[(\theta+(1+\alpha)(1-\phi))\sigma_a^2+\alpha(1-\theta+\alpha)\sigma_g^2+\theta(1+(\theta-\alpha)(1-\phi))\sigma_k^2\right]}{2(1-\theta-\alpha)(1-\theta-\alpha(1-\phi))}\\ &+\frac{\phi(2-\phi)}{(1-\theta-\alpha)(1-\theta-\alpha(1-\phi))}(-\alpha\sigma_{ag}-\theta\sigma_{ak}+\alpha\theta\sigma_{gk})\end{aligned}$$

We can solve for λ as

$$\lambda = (\tau_y-g)^{\frac{1-\theta-\alpha(1-\phi)}{1-\theta-\alpha}}\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)^{\frac{\alpha\phi}{1-\theta-\alpha}}\left(\frac{1}{R}\right)^{\frac{(\alpha+\theta)\phi}{1-\theta-\alpha}}(\theta(1-\tau_y))^{\frac{\theta\phi}{1-\theta-\alpha}}\tau_y^{-(1-\phi)}\exp(\Omega^\lambda)$$

Plugging λ into k_{Gi} and y_i given by Equations (C2) and (C3) yields Equations (15) and (16) in Proposition 1. Plugging y_i and k_{Gi} into the local government's budget constraint can deliver G_i given by Equation (**??**). Plugging k_{Gi} into Equations (7) can deliver k_i.

Corollaries 2, 3, and 4 can be obtained directly using the equilibrium allocations.

The complete version of equilibrium allocations is given by

$$\begin{aligned}\log k_G(a,\tau_g,\tau_k) = &\frac{1}{1-\theta-\alpha(1-\phi)}\left((1-\phi)a-(1-\theta)\log(1+\tau_g)-\theta(1-\phi)\log(1+\tau_k)\right)\\ &+\frac{1}{1-\theta-\alpha}\left((1-\theta)\log(\tau_y-g)-\log R+\theta\log(\theta(1-\tau_y))\right)\\ &+\frac{1-\theta}{1-\theta-\alpha}\log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)+\frac{1-\theta}{1-\theta-\alpha(1-\phi)}\Omega^\lambda\end{aligned} \tag{C4}$$

$$\begin{aligned}\log y(a,\tau_g,\tau_k) = &\frac{1}{1-\theta-\alpha(1-\phi)}\left(a-\alpha\log(1+\tau_g)-\theta\log(1+\tau_k)\right)\\ &+\frac{1}{1-\theta-\alpha}(\alpha\log(\tau_y-g)-(\alpha+\theta)\log R+\theta\log(\theta(1-\tau_y)))\\ &+\frac{\alpha}{1-\theta-\alpha}\log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)+\frac{\alpha}{(1-\theta-\alpha(1-\phi))}\Omega^\lambda\end{aligned} \tag{C5}$$

$$\begin{aligned}\log k(a,\tau_g,\tau_k) = &\frac{1}{1-\theta-\alpha(1-\phi)}\left(a-\alpha\log(1+\tau_g)-(1-\alpha(1-\phi))\log(1+\tau_k)\right)\\&+\frac{1}{1-\theta-\alpha}\left(\alpha\log(\tau_y-g)-\log R+(1-\alpha)\log(\theta(1-\tau_y))\right)\\&+\frac{\alpha}{1-\theta-\alpha}\log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)+\frac{\alpha}{(1-\theta-\alpha(1-\phi))}\Omega^\lambda\end{aligned} \tag{C6}$$

$$\begin{aligned}\log G(a,\tau_g,\tau_k) = &\frac{1-\phi}{1-\theta-\alpha(1-\phi)}\left(a-\alpha\log(1+\tau_g)-\theta\log(1+\tau_k)\right)\\&+\frac{1}{1-\theta-\alpha}\left((1-\theta)\log(\tau_y-g)-(\theta+\alpha)\log R+\theta\log(\theta(1-\tau_y))\right)\\&+\frac{1-\theta}{1-\theta-\alpha}\log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right)+\log\left(\frac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha(1+\chi^L(1-\phi))}\right)\\&+\frac{1-\theta}{1-\theta-\alpha(1-\phi)}\Omega^\lambda\end{aligned} \tag{C7}$$

$\tau_g=\tau_k=0$ yields the equilibrium allocations displayed in Proposition 1.

C.2 Proofs of Proposition 2, Corollaries 5, 6

Lemma 1. *If a random variable X is log-normally distributed, $\log X\sim\mathcal{N}(\mu,\sigma_a^2)$, we have*

$$\log(\mathbb{E}[X])=\mu+\frac{\sigma_a^2}{2}$$

As k_{Gi}, y_i, and k_i given by Equations (C4), (C5), and (C6) are log-normally distributed, it is straightforward to get the expectations following Lemma 1 and calculate TFP as the Solow residual. When calculating the Solow residual, all terms in the second, third, and fourth lines will be canceled out.

When $\sigma_g^2=\sigma_k^2=\sigma_{ag}=\sigma_{ak}=\sigma_{kg}=0$, taking the derivatives of $\log Y$ with respect to ϕ yields

$$\frac{\partial\log Y}{\partial\phi}=-\frac{\alpha(1-\theta)}{1-\theta-\alpha}\frac{2\phi}{(1-\theta-\alpha(1-\phi))^3}\frac{\sigma_a^2}{2}-\frac{\alpha}{1-\theta-\alpha}\frac{\chi^L}{1+\chi^L(1-\phi)}$$

$\partial\log Y/\partial\phi<0$ when $\phi\geq 0$. This completes the proof of Corollary 5. Corollary 5 still holds when σ_g^2 and σ_k^2 not equal to 0.

When $\sigma_g^2=\sigma_k^2=0$, $\log\text{TFP}<0$ for all $\phi\neq 0$ and $\log\text{TFP}=0$ for $\phi=0$. This

makes $\phi = 0$ a global maximizer and establishes the proof of the first half of Corollary 6. We will prove the rest of the corollary in the next part.

It is convenient to derive $\mathbb{E}[G]/\mathbb{E}[y]$ using Equations (C5) and (C7). We can also prove it intuitively. By Corollary 2, we know that the ratio of local government consumption to the rent on public capital is constant. Therefore, the ratio of local government consumption to the post-redistribution local government revenue should be constant as well. The post-redistribution local government revenue equals total tax revenue less central government consumption, which is $(\tau_y - g)\mathbb{E}[y]$. This leads to the expression of $\mathbb{E}[G]/\mathbb{E}[y]$ as:

$$\frac{\mathbb{E}[G_i]}{\mathbb{E}[y_i]} = (\tau_y - g) \times \frac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha + \chi^L(1-\theta)} \tag{C8}$$

C.3 Proofs of Propositions 3, 4, 5, Corollaries 7, 8

Before we derive the analytical expression of the social welfare function, we introduce some useful lemmas.

Lemma 2. *If a random variable X is log-normally distributed, $\log X \sim \mathcal{N}(\mu, \sigma_a^2)$, we have*

$$\mathbb{E}[\log(X)] = \log \mathbb{E}[X] - \frac{1}{2}\text{var}(\log X)$$

Proof. This follows directly from the property of the log-normal distribution.

$$\begin{aligned} & \mathbb{E}[X] = \exp(\mu + \frac{\sigma^2}{2}) \\ \Rightarrow \quad & \log \mathbb{E}[X] = \mu + \frac{\sigma^2}{2} = \mathbb{E}[\log X] + \frac{1}{2}\text{var}(\log X) \end{aligned}$$

□

Lemma 3. *If $X \sim \mathcal{N}(\mu, \sigma^2)$,*

$$\mathbb{E}[X \exp(tX)] = (\mu + t\sigma^2)\mathbb{E}[\exp(tX)]$$

Proof.

$$\begin{aligned}
\mathbb{E}[X\exp(tX)] &= \int X\exp(tX)\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{(X-\mu)^2}{2\sigma^2}\right)dX \\
&= \int X\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\mu-t\sigma^2)^2-2\mu t\sigma^2-t^2\sigma^4}{2\sigma^2}\right)dX \\
&= \int X\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\mu-t\sigma^2)^2}{2\sigma^2}\right)dX\exp\left(\frac{2\mu t\sigma^2+t^2\sigma^4}{2\sigma^2}\right) \\
&= (\mu+t\sigma^2)\exp\left(\mu t+\frac{t^2\sigma^2}{2}\right) \\
&= (\mu+t\sigma^2)\mathbb{E}[\exp(tX)]
\end{aligned}$$

□

Lemma 4 (Law of Iteration)**.**

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY|X]] = \mathbb{E}[X\mathbb{E}[Y|X]]$$

Lemma 5 (Conditional Distribution)**.** *If a and b are jointly normally distributed, i.e.*

$$(X,Y) \sim \mathcal{N}\left(\begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \sigma_{XY}\\ \sigma_{XY} & \sigma_Y^2\end{bmatrix}\right)$$

Then,

$$Y|X \sim \mathcal{N}\left(\mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(X-\mu_X), \frac{\sigma_X^2\sigma_Y^2-\sigma_{XY}^2}{\sigma_X^2}\right)$$

Now if X and Y are both normally distributed and $\operatorname{cov}(X,Y)=0$, we show how to calculate $\mathbb{E}[(X+Y)\exp(tX)]$. First,

$$\begin{aligned}
\mathbb{E}[Y\exp(tX)] &= \mathbb{E}[\mathbb{E}[Y\exp(tX)|X]] \\
&= \mathbb{E}[\mathbb{E}[Y|X]\exp(tX)] \\
&= \mathbb{E}\left[\left(\mu_Y+\frac{\sigma_{XY}}{\sigma_X^2}(X-\mu_X)\right)\exp(tX)\right] \\
&= (\mu_Y+t\sigma_{XY})\mathbb{E}[\exp(tX)]
\end{aligned}$$

The first line follows Lemma 4. The second line follows the property of the conditional expectation. The third line follows Lemma 5. The fourth line follows the assumption that $\text{cov}(X, Y) = 0$.

Therefore, we have

$$\begin{aligned}\mathbb{E}[(X+Y)\exp(tX)] &= \mathbb{E}[X\exp(tX)] + \mathbb{E}[Y\exp(tX)] \\ &= (\mu_X + t\sigma_X^2)\mathbb{E}[\exp(tX)] + (\mu_Y + t\sigma_{XY})\mathbb{E}[\exp(tX)]\end{aligned}$$

Take $\int \exp(a_i)\log y_i di$ for example, for the first line of Equation (C5), we have

$$\begin{aligned}&\frac{1}{\int \exp(-\gamma a)dF_a(a)} \int \frac{\exp(-\gamma a)}{1-\theta-\alpha(1-\phi)} \left(a - \alpha\log(1+\tau_g) - \theta\log(1+\tau_k)\right) dF(a, \tau_y, \tau_g) \\ &= \frac{\gamma(-\sigma_a^2 + \alpha\sigma_{ag} + \theta\sigma_{ak})}{1-\theta-\alpha(1-\phi)} + \frac{1}{1-\theta-\alpha(1-\phi)}\left(-\frac{\sigma_a^2}{2} + \frac{\alpha\sigma_g^2}{2} + \frac{\theta\sigma_k^2}{2}\right)\end{aligned}$$

Therefore, we have

$$\begin{aligned}&\frac{1}{\int \exp(-\gamma a)dF_a(a)} \int \exp(-\gamma a)\log y(a, \tau_y, \tau_g) dF(a, \tau_y, \tau_g) \\ &= \mathbb{E}[\log y] + \frac{\gamma(-\sigma_a^2 + \alpha\sigma_{ag} + \theta\sigma_{ak})}{1-\theta-\alpha(1-\phi)} \\ &= \log(\mathbb{E}[y]) - \frac{1}{2}\text{var}(\log y) + \frac{\gamma(-\sigma_a^2 + \alpha\sigma_{ag} + \theta\sigma_{ak})}{1-\theta-\alpha(1-\phi)}\end{aligned}$$

The last line follows Lemma 2.

With $\log c_i = \log(w_i + \pi_i) = \log((1-\theta)(1-\tau_y)) + \log y_i$, $\log G^C = \log(g\mathbb{E}[y])$, and the aggregate local government consumption to output ratio given by Equation (C8), the welfare from aggregate output and the welfare cost from inequality:

$$\begin{aligned}\mathcal{W}(\phi, \tau_y, g) = (1 + \chi^L + \chi^C)\log\mathbb{E}[y_i] - \frac{1}{2}\text{var}(\log c_i) - \frac{\chi^L}{2}\text{var}(\log G_i) - \tilde{\gamma}\sigma_a^2 \\ + \log((1-\tau_y)(1-\theta)) + \chi^C \log g + \chi^L\left(\log\left(\frac{\chi^L(1-\theta-\alpha(1-\phi))}{\alpha + \chi^L(1-\theta)}\right) + \log(\tau_y - g)\right)\end{aligned}$$

Plugging the expression of $\log Y$ given by Equation 2 into Equation C.3 yields the expression of Equation 19 in Proposition 3. This expression with capital wedges is given by:

$$
\begin{aligned}
&\mathcal{W}(\phi, \tau_y, g) = \\
&\Bigg\{ \frac{1}{1-\theta-\alpha}((\alpha(1+\chi^C)+\chi^L(1-\theta))\log(\tau_y - g) + (1-\alpha+\theta(\chi^L+\chi^C))\log(1-\tau_y)) + \chi^C \log g \\
&+ \frac{(1-\theta)\chi^L + \alpha(1+\chi^C)}{1-\theta-\alpha} \log\left(\frac{\alpha(1+\chi^L(1-\phi))}{\alpha+\chi^L(1-\theta)}\right) + \chi^L \log\left(\frac{\chi^L(1-\theta-\alpha(1-\phi)))}{\alpha(1+\chi^L(1-\phi))}\right) \\
&+ \frac{1+\chi^L+\chi^C}{2(1-\theta-\alpha)^2}\left[(\theta+\alpha)\sigma_a^2 + \alpha(1-\theta)\sigma_g^2 + \theta(1-\alpha)\sigma_k^2 + 2\left(-\alpha\sigma_{ag} - \theta\sigma_{ak} + \alpha\theta\sigma_{kg}\right)\right] \\
&- \frac{(1+\chi^L+\chi^C)\alpha(1-\theta)\phi^2}{2(1-\theta-\alpha)^2(1-\theta-\alpha(1-\phi))^2}\left(\sigma_a^2 + \alpha^2\sigma_g^2 + \theta^2\sigma_k^2 + 2\left(-\alpha\sigma_{ag} - \theta\sigma_{ak} + \alpha\theta\sigma_{kg}\right)\right) \\
&- \frac{1}{2(1-\theta-\alpha(1-\phi))^2}\left(\sigma_a^2 + \alpha^2\sigma_g^2 + \theta^2\sigma_k^2 + 2\left(-\alpha\sigma_{ag} - \theta\sigma_{ak} + \alpha\theta\sigma_{kg}\right)\right) \\
&- \frac{\chi^L(1-\phi)^2}{2(1-\theta-\alpha(1-\phi))^2}\left(\sigma_a^2 + \alpha^2\sigma_g^2 + \theta^2\sigma_k^2 + 2\left(-\alpha\sigma_{ag} - \theta\sigma_{ak} + \alpha\theta\sigma_{kg}\right)\right) \\
&- \frac{\gamma(1+\chi^L(1-\phi))}{1-\theta-\alpha(1-\phi)}(\sigma_a^2 + \alpha\sigma_{ag} + \theta\sigma_{ak})
\end{aligned}
$$

Given the social welfare function in Equation C.3, we can derive the first and the twice differentiations with respect to ϕ, τ_y, and g, respectively.

$$
\begin{aligned}
\frac{\partial \mathcal{W}(\phi, \tau_y, g)}{\partial g} &= -\frac{\alpha(1+\chi^C)+\chi^L(1-\theta)}{1-\theta-\alpha}\frac{1}{\tau_y - g} + \frac{\chi^C}{g} \\
\frac{\partial^2 \mathcal{W}(\phi, \tau_y, g)}{\partial g^2} &= -\frac{\alpha(1+\chi^C)+\chi^L(1-\theta)}{1-\theta-\alpha}\frac{1}{(\tau_y - g)^2} - \frac{\chi^C}{g^2} < 0
\end{aligned} \tag{C9}
$$

$$
\frac{\partial \mathcal{W}(\phi, \tau_y, g)}{\partial \tau_y} = \frac{\alpha(1+\chi^C)+\chi^L(1-\theta)}{1-\theta-\alpha}\frac{1}{\tau_y - g} - \frac{1-\alpha+\theta(\chi^L+\chi^C)}{1-\theta-\alpha}\frac{1}{1-\tau_y} \tag{C10}
$$

$$
\frac{\partial^2 \mathcal{W}(\phi, \tau_y, g)}{\partial \tau_y^2} = -\frac{\alpha(1+\chi^C)+\chi^L(1-\theta)}{1-\theta-\alpha}\frac{1}{(\tau_y - g)^2} - \frac{1-\alpha+\theta(\chi^L+\chi^C)}{1-\theta-\alpha}\frac{1}{(1-\tau_y)^2} < 0
$$

$$\begin{aligned}\frac{\partial\widetilde{\mathcal{W}}(\phi,\tau_y,g)}{\partial\phi} &= -\frac{\alpha(1+\chi^L+\chi^C)}{1-\theta-\alpha}\frac{\chi^L}{1+\chi^L(1-\phi)}+\frac{\alpha\chi^L}{1-\theta-\alpha(1-\phi)}\\ &\quad-\frac{(1+\chi^L+\chi^C)\alpha(1-\theta)}{1-\theta-\alpha}\frac{\phi}{(1-\theta-\alpha(1-\phi))^3}\left(\sigma_a^2+\alpha^2\sigma_g^2+\theta^2\sigma_k^2-\alpha\sigma_{ag}-\theta\sigma_{ak}+\alpha\theta\sigma_{kg}\right)\\ &\quad+\frac{\alpha+\chi^L(1-\phi)(1-\theta)}{(1-\theta-\alpha(1-\phi))^3}\left(\sigma_a^2+\alpha^2\sigma_g^2+\theta^2\sigma_k^2-\alpha\sigma_{ag}-\theta\sigma_{ak}+\alpha\theta\sigma_{kg}\right)\\ &\quad+\frac{\chi^L(1-\theta)+\alpha}{(1-\theta-\alpha(1-\phi))^2}\gamma(\sigma_a^2-\alpha\sigma_{ag}-\theta\sigma_{ak})\end{aligned} \tag{C11}$$

$$\begin{aligned}\frac{\partial^2\widetilde{\mathcal{W}}(\phi,\tau_y,g)}{\partial\phi^2} &= -\frac{\alpha(1+\chi^L+\chi^C)}{1-\theta-\alpha}\frac{(\chi^L)^2}{(1+\chi^L(1-\phi))^2}-\frac{\alpha^2\chi^L}{(1-\theta-\alpha(1-\phi))^2}\\ &\quad-\frac{(1+\chi^L+\chi^C)\alpha(1-\theta)}{1-\theta-\alpha}\frac{1-\theta-\alpha-2\alpha\phi}{(1-\theta-\alpha(1-\phi))^4}\left(\sigma_a^2+\alpha^2\sigma_g^2+\theta^2\sigma_k^2-\alpha\sigma_{ag}-\theta\sigma_{ak}+\alpha\theta\sigma_{kg}\right)\\ &\quad-\frac{(1-\theta-\alpha(1-\phi))\chi^L(1-\theta)+3(\alpha+\chi^L(1-\phi)(1-\theta))}{(1-\theta-\alpha(1-\phi))^4}(\sigma_a^2+\alpha^2\sigma_g^2+\theta^2\sigma_k^2\\ &\qquad-\alpha\sigma_{ag}-\theta\sigma_{ak}+\alpha\theta\sigma_{kg})\\ &\quad-\frac{2(\chi^L(1-\theta)+\alpha)\alpha}{(1-\theta-\alpha(1-\phi))^3}\gamma(\sigma_a^2-\alpha\sigma_{ag}-\theta\sigma_{ak})\end{aligned} \tag{C12}$$

In Equation (C12), the first, third, and the last lines are negative. The second line is negative if $1-\theta-\alpha > 2\alpha\phi$ and $\sigma_a^2 > \alpha\sigma_{ag}+\theta\sigma_{ak}$. This completes the proof of Corollary 7 when $\sigma_g^2=\sigma_k^2=0$.

With the concavity, the first-order condition is sufficient to deliver the maximizer. For (C11), when $\sigma_g^2=\sigma_k^2=\chi^C=0$, the sufficient and necessary condition for $\partial\mathcal{W}(\phi,\tau_y,g)/\partial\phi|_{\phi=0}=0$ is $\gamma=1/(1-\theta-\alpha)$. This completes the proof of the second half of Corollary 6.

Corollary 8 is derived from letting $\partial\mathcal{W}(\phi,\tau_y,g)/\partial g=0$ and $\partial\mathcal{W}(\phi,\tau_y,g)/\partial\tau_y=0$ using Equations (C9) and (C10).

For Proposition 4, ϕ^{R*} is derived from letting the first line of Equation (C11) equal to 0. To calculate the marginal tax rate and the average tax rate under the scheme (ϕ^{R*},τ_y^*,g^*), recall that in the representative model, λ is calculated as

$$(\tau_y-g)y=\lambda(\tau_y y)^{1-\phi} \quad\Rightarrow\quad \lambda=\frac{\tau_y-g}{\tau_y^{1-\phi}}y^{\phi}$$

Then, the marginal tax rate and the average tax rate are given by

$$MTR = 1 - \lambda(1-\phi)(\tau_y y)^{-\phi} = 1 - (\tau_y - g)\tau_y^{-1}(1-\phi)$$

$$ATR = 1 - \lambda(\tau_y y)^{-\phi} = 1 - (\tau_y - g)\tau_y^{-1}$$

From Equations (20) and (24), we can calculate

$$\tau_y^* - g^* = \frac{\alpha(1+\chi^C) + (1-\theta)\chi^L}{1+\chi^L+\chi^C}$$

$$(\tau_y^*)^{-1} = \frac{1+\chi^L+\chi^C}{\alpha + (1-\theta)(\chi^L+\chi^C)}$$

$$1 - \phi^{R*} = \frac{\alpha + (1-\theta)(\chi^C+\chi^L)}{\alpha(1+\chi^C) + (1-\theta)\chi^L}$$

These above three equations yield that $MTR = 0$ and

$$ATR = 1 - (\tau_y^* - g^*)(\tau_y^*)^{-1} = \frac{(1-\alpha-\theta)\chi^C}{\alpha + (1-\theta)(\chi^L+\chi^C)} = \frac{g^*}{\tau_y^*}$$

This completes the proof of Proposition 4.

Inequality (25) in Proposition 5 is derived from evaluating Equation (C11) at $\phi = 0$ and letting it be strictly larger than 0.

Moreover, $\partial \mathcal{W}(\phi, \tau_y, g)/\partial\phi = 0$ leads to Equation (26) in Proposition 6.

D Additional Quantitative Results

Table D3: Cross-sectional Moments: Data, Baseline Model, and Counterfactual, 1998-2017

		Data	Baseline Model	Counterfactual $\phi = 0$
A. Cross-sectional moments				
1.	Coefficient of $\log y$ on $\log A$	1.763	1.743	2
2.	Coefficient of $\log k_G$ on $\log A$	1.123	1.048	2
3.	Coefficient of $\log(k_G/y)$ on $\log A$	-0.64	-0.695	0
4.	Coefficient of $\log y$ on $\log k_G$	0.729	0.599	1
B. Distributional moments				
5	$\text{var}(\log y)$	0.23	0.219	0.288
6	$\text{var}(\log k)$	0.319	0.281	0.346
7	$\text{var}(\log k_G)$	0.234	0.082	0.288
C. Aggregate moments				
8	$\log Y$	0.108	0.301	0.340
9	$\log K_G$	-0.624	0.041	0.16
10	$\log \text{TFP}$	0.102	0.116	0.121
11	$\log(\text{TFP}/\text{TFP}^*)$	-0.045	-0.022	-0.017

Table D4: Cross-sectional Moments: Model with Public Capital Wedges and Counterfactual

	2013-2017		1998-2017	
	Model	Counterfactual $\phi = 0$	Model	Counterfactual $\phi = 0$
$\text{var}(\log y)$	0.171	0.238	0.230	0.301
$\text{var}(\log k)$	0.200	0.253	0.319	0.391
$\text{var}(\log k_G)$	0.196	0.385	0.234	0.480
$\log \text{TFP}$	0.281	0.287	0.102	0.101
$\log(\text{TFP}/\text{TFP}^*)$	-0.063	-0.057	-0.036	-0.037

Note: the model should match the moments in data by construction.

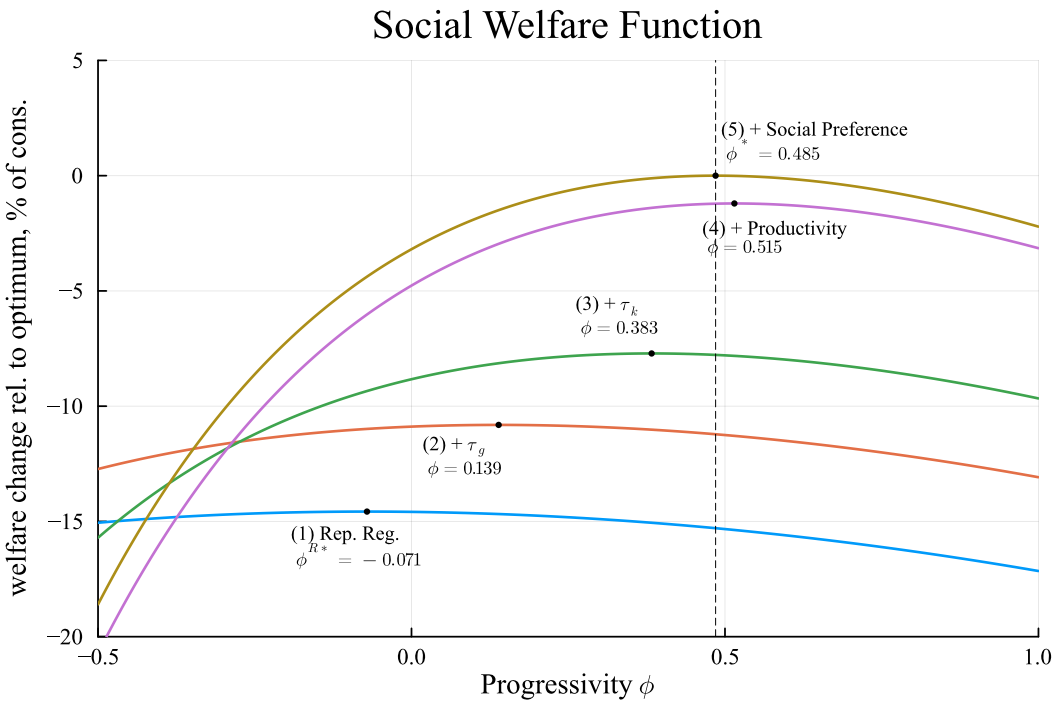
Table D5: Aggregate Responses in the Model with Public Capital Wedges and Counterfactual, 2004-2013

	Model	Counterfactual $\phi = \phi_{2004}$
$\Delta\text{var}(\log y)$	-0.09	-0.065
Convergence rate	-2.446	-1.878
$\Delta \log \text{TFP}$	0.24	0.242
$\Delta \log(\text{TFP}/\text{TFP}^*)$	-0.029	-0.027

Table D6: Parameterization in the Model with Capital Wedges, 2013-2017

γ	σ_a^2	σ_g^2	σ_k^2	σ_{ag}	σ_{ak}	σ_{kg}
-0.096	0.099	0.154	0.104	0.055	0.069	0.078

Figure D7: Decomposition of the Social Welfare Function with Capital Wedges



Note: γ in the model with capital wedges is calibrated to be -0.096.