Delayed Retirement or More Births? Short-Run Relief and Long-Run Sustainability of China's Pension System

Zejin Shi* Pengwei Guo† Jun Nie‡

Economics and Management School, Wuhan University

April 19, 2026

Abstract

We build a heterogeneous-agent overlapping generations model incorporating China's institutional "combined accounts" pension structure and urban–rural dual track to assess the short-run and long-run fiscal sustainability of the pension system under population aging. Calibrated to 2023 data, the model reproduces the observed pension deficit of 0.73% of GDP, providing a validated baseline for projecting future fiscal paths. We solve for general equilibrium transition paths under two polar fiscal closures—payroll tax adjustment and government spending adjustment—that bracket feasible fiscal responses. Absent reform, the pension deficit widens to about 12.7% of GDP in the long run, forcing either a 3.5-fold payroll tax increase or a halving of government spending. Among reform instruments, delayed retirement delivers meaningful short-run gains, but only raising fertility addresses the root cause—a rising dependency ratio—and yields the largest long-run fiscal dividend, though with a two-decade lag. Combining both policies delivers short-run relief and long-run sustainability. We also compute the *fiscal value of birth*—the present-value fiscal gain per additional birth—as a cost-effectiveness benchmark for pro-natalist subsidies: even modest fertility increases to a TFR of 1.3–1.5 justify per-birth subsidies of 100,000–125,000 yuan.

Keywords: Heterogeneous Agent, Social Security, Sustainability, Population Aging, OLG Model

JEL Classification: E21, H55, J11, J26

*Corresponding author. shizejin94@gmail.com

†pweiguo@whu.edu.cn

‡niejun@whu.edu.cn

1 Introduction

China is undergoing one of the most rapid demographic transitions in recorded history. The total fertility rate has fallen to approximately 1.0, less than half the replacement level, while the population aged 65 and above reached 220 million—15.6% of the total—in 2024. The old-age dependency ratio climbed from 14.3% in 2015 to 22.5% in 2023 and is projected to approach 50% by 2050, driven in part by the imminent retirement of the 1962–1973 baby boom cohort. China's pension system, which covers roughly one billion people, already runs an annual deficit of 1.3 trillion yuan and faces exhaustion of its accumulated surplus by the mid-2040s absent reform. This paper asks: what policy reforms can restore fiscal sustainability, and at what cost to different generations?

We build a heterogeneous-agent overlapping generations model calibrated to the Chinese economy. The model separately captures China's two pension tracks: the urban "combined accounts" system—where retirement benefits combine an individual account annuity and a pooling benefit indexed to wage history—and the rural flat-rate benefit scheme. Agents face idiosyncratic productivity shocks and mortality risk with incomplete markets, generating a rich cross-sectional distribution of wealth and pension entitlements. A key empirical discipline is that the model is calibrated to match the observed 2023 pension deficit—separately targeting the urban and rural components—achieving a total deficit of -0.78% of GDP against the data value of -0.73%. This close fit to the current fiscal shortfall lends credibility to the model's forward-looking projections of deficit dynamics under aging and reform. We solve for complete general equilibrium transition paths—from the current steady state through 200 years of demographic change—under two polar fiscal closure rules (payroll tax adjustment and government spending adjustment) and evaluate three reform instruments: delayed retirement, reduced pooling transfers, and pro-natalist policies.

Modeling individual-level pension heterogeneity is central to our analysis, for three reasons. First, pension expenditure is the key fiscal variable, and computing it accurately requires the actual benefit formula. Under the Urban Employee Basic Pension Insurance system, retirement benefits combine an individual account annuity (determined by accumulated contributions b) and a pooling benefit (determined by the worker's wage history e). Only by tracking (b, e) at the individual level can the model replicate the dual-pillar benefit structure and produce reliable aggregate pension expenditure figures. Second, these state variables capture the fiscal inertia built into the existing stock of pension claims. Because each current retiree's benefit is fixed by her accumulated (b, e) at the moment of retirement, reforms to the benefit formula propagate through the aggregate pension bill only to the extent that the reformed parameters multiply individual-specific state variables—and the resulting incidence depends on the entire cross-sectional distribution of (b, e) across retirees, not just its mean. A model that collapses pensions into a uniform transfer, by contrast, treats every retiree as the average retiree and therefore cannot separate the short-run fiscal response to a reform from its distributional consequences across retirees with different contribution histories. Third, individual-level heterogeneity in pension entitlements enables distributional analysis: we can trace how each reform affects wealth inequality, consumption dispersion, and welfare across urban and rural workers, age groups, and income levels. We therefore solve over a four-dimensional individual state space (a, b, e, γ) with 800,000 simulated agents across 200-year transition paths.

A key design choice is the use of two polar fiscal closure rules. Under payroll tax adjustment, the payroll tax rate adjusts each period to maintain the debt-to-GDP ratio, while government spending is held fixed as a share of GDP; this mode reveals the *tax burden* of population aging. Under government spending adjustment, all tax rates are held at their 2023 levels and government spending absorbs the fiscal pressure; this mode reveals the *public goods cost* of aging. The two modes bracket the space of feasible fiscal responses: any realistic policy will combine tax and spending adjustments, and our dual framework shows how the economic consequences—output, wages, capital accumulation, welfare, and inequality—depend on where along this spectrum the adjustment falls. Each mode has a distinct advantage: the tax-based adjustment provides a concrete, interpretable measure of fiscal pressure but introduces potentially large distortions; the spending-based adjustment preserves market incentives but involves welfare costs from reduced public goods that are difficult to quantify within the model.

Our quantitative analysis yields three main findings. First, absent reform the pension system deficit widens from −0.78% of GDP in 2023 to approximately −12.7% in the long run—a sixteen-fold expansion driven almost entirely by the urban combined-accounts pension system. Closing this gap requires either a 3.5-fold increase in the payroll tax rate (from 9.6% to 33.6%) or a halving of government spending (from 21.8% to 10.6% of GDP), depending on the fiscal instrument—highlighting that the choice of fiscal closure has first-order consequences for the macroeconomy. Second, the three reform instruments exhibit a clear ranking. Reduced pooling transfers provide the largest immediate fiscal relief (−6.1 percentage points off the required payroll tax) but cannot restore long-run sustainability (the terminal tax rate remains at 24.5%) and hurt redistribution by cutting benefits for low-earners who depend most on the pooling component. Delayed retirement delivers meaningful short-run gains (−3.2 pp transitional, −5.9 pp terminal) by immediately extending working years and reducing retirees, though its effect is partially offset by higher per-period pension benefits. Only raising fertility addresses the root cause of pension unsustainability—a rising dependency ratio—and yields the largest long-run fiscal dividend (−15.5 pp terminal), but with a two-decade lag before the additional birth cohorts enter the workforce. Combining delayed retirement with replacement-level fertility provides both short-run relief and long-run sustainability, reducing the terminal tax rate by 17.6 percentage points: the two instruments are complementary rather than substitutes.

Third, since the effectiveness of pro-natalist policy at actually raising births is inherently uncertain, we contribute from a complementary angle. Instead of asking how much fertility can be raised, we ask: if the government paid a one-time cash subsidy for every additional child born, how large could that subsidy be while still leaving the government no worse off in present-value terms? We call this the *fiscal value of birth*. Concretely, it is the present value of the lifetime fiscal stream each additional newborn eventually generates—the labor and consumption taxes she will pay during her working years, net of the public services and pension benefits she will eventually claim, and including the dividend she delivers by expanding the future workforce and thereby easing the pension burden on an aging population. This quantity is a natural break-even benchmark for any upfront pro-natalist subsidy: as long as the per-birth payment stays below the fiscal value of birth, each additional newborn the policy induces pays for itself in present value, *regardless of how responsive actual fertility is to the policy*. To make this benchmark directly comparable to annual childcare allowances, we also re-express the fiscal value as a constant annual payment over the first 20 years of the child's life. We show that even modest fertility increases to a TFR of 1.3–1.5 justify per-birth subsidies of 100,000–125,000 yuan, or equivalently, annual subsidies of approximately 9,000–11,000 yuan for 20 years, providing a cost-effectiveness benchmark that sidesteps the fertility-responsiveness question. Notably, our TFR sweep analysis reveals that the marginal fiscal value is actually *highest* at low fertility levels near the current TFR of 1.0, because the first additional births alleviate the most acute dependency pressure—implying that even small, realistic fertility improvements are cost-effective.

This paper contributes to several strands of literature. The first is the quantitative OLG tradition initiated by Auerbach and Kotlikoff [1987], which was enriched by the incorporation of heterogeneous agents and incomplete markets [Bewley, 1986, Aiyagari, 1994, Huggett, 1996]; Krueger et al. [2016] survey the modern heterogeneous-agent macro framework. A foundational line of applied general-equilibrium analysis of U.S. Social Security reform emerged with Huang et al. [1997], who computed two explicit transition paths from Pay-As-You-Go to a fully funded Social Security system, and De Nardi et al. [1999], who embedded Social Security Administration population projections in a calibrated OLG model to evaluate four fiscal responses to the baby-boom retirement—raising payroll taxes, cutting benefits, raising the retirement age, and means-testing—finding that maintaining promised benefits requires substantial distortionary taxation. These two papers established the transitioncomputation and four-instrument-comparison framework that subsequent work inherited. Conesa and Krueger [1999] studied social security reform with idiosyncratic risk, showing that the insurance role of pay-as-you-go systems reduces political support for privatization. Krueger and Kubler [2006] show that when financial markets are incomplete and capital returns and wages are imperfectly correlated, introducing unfunded social security can be Pareto-improving through intergenerational risk sharing. Kitao [2014] revisited the same four-option comparison in a richer heterogeneous-agent environment. Nishiyama and Smetters [2007] demonstrated efficiency costs of privatization when wage shocks are uninsurable, while Krueger and Ludwig [2007] and Ludwig et al. [2009] analyzed how global population aging affects capital returns and welfare across cohorts in both closed and open economies. Fehr et al. [2013] examined progressive pension design in Germany. İmrohoroğlu and Kitao [2012] and İmrohoroğlu et al. [2016] applied this framework to study Social Security reform and fiscal balance in Japan, combining lifecycle heterogeneity with detailed institutional structures.

The second strand concerns China's pension system and demographic transition. Fang and Feng [2018] provide a comprehensive institutional overview of China's fragmented multi-pillar pension architecture. Song et al. [2015] study intergenerational sharing of growth under demographic transition in an OLG framework calibrated to China, finding that delaying reform benefits poorer current generations. He et al. [2019] analyze how rapid aging and pension reform jointly explain the rise in China's household saving rate. İmrohoroğlu and Zhao [2018] and İmrohoroğlu and Zhao [2020] embed family insurance and financial constraints in lifecycle models of Chinese saving behavior. More recently, Gai et al. [2025] use a structural general equilibrium model to evaluate the rural pension scheme's effects on labor reallocation and aggregate income, and Deng et al. [2023] compare delayed retirement versus benefit adjustment using a life-cycle model that distinguishes skill groups. Bairoliya et al. [2018] model rural health insurance and pension reform jointly in a dynamic GE framework. A substantial domestic literature in Chinese further examines delayed retirement, contribution rate design, and labor supply effects of pension reform; our model subsumes these partial-equilibrium analyses within a unified heterogeneous-agent general equilibrium framework.

Relative to this literature, our contributions are threefold. First, we provide the first heterogeneous-agent general equilibrium analysis with complete transition dynamics for China's pension system, incorporating the institutional "combined accounts" dual-pillar benefit formula with individual-level state variables for pension account balances and wage histories. Crucially, this institutional detail allows the model to reproduce the observed 2023 pension deficit—matching both the urban and rural components—providing a validated baseline from which to project the fiscal impact of demographic change and reform. Second, we introduce a dual fiscal closure framework—payroll tax adjustment versus government spending adjustment—that brackets the space of feasible fiscal responses, revealing that instrument choice has first-order consequences for the macroeconomy and welfare. Third, we compute the fiscal value of birth across a range of fertility targets, providing a natural cost-effectiveness benchmark for evaluating pro-natalist subsidies.

The remainder of this paper is organized as follows: Section 2 presents the model, Section 3 describes the calibration, Section 4 reports quantitative results on policy experiments, distributional effects, and welfare, Section 5 quantifies the fiscal value of each additional birth, and Section 6 concludes.

2 Theoretical Model

2.1 Model Framework

The model combines the lifecycle dynamics of the Auerbach–Kotlikoff OLG framework with the incomplete-market heterogeneity of Aiyagari [1994]. Agents live up to 80 periods, face idiosyncratic productivity shocks and borrowing constraints, and are divided into urban and rural workers with distinct pension systems. The urban worker state is four-dimensional—

tradeable assets a, pension account balance b, historical average wage e, and productivity γ—allowing the model to capture the dual-pillar benefit formula of China's combined-accounts pension system. This state space is computationally demanding but economically necessary: it allows the model to replicate the actual benefit formula, capture the fiscal inertia of predetermined pension claims, and support distributional analysis across heterogeneous agents.

2.2 Demographics

The evolution of population structure constitutes the core exogenous driving force of this paper's analysis. The model bases its population dynamics on UN *World Population Prospects* (2024 revision) demographic projections for China, depicting the future long-term population age structure transition path and accurately capturing the impact of baby boom generational changes on the economic system.

Let $N_t(j)$ denote the population of age j at time t. Agents live from age 20 ($j = 0$) to at most age 99 ($j = 79$). Working-age individuals ($j \leq j_R$) participate in labor and pay social security contributions while simultaneously making consumption and savings decisions; retired individuals ($j > j_R$) receive pensions and continue making consumption and savings decisions, where $j_R = 40$ (age 60) under the baseline retirement policy.

In each period t, new cohorts of age $j = 0$ enter the economy. The size of the new cohort $N_t(0)$ is determined by the fertility behavior of the childbearing population. Specifically, let ψ_j denote the female share in the age-j cohort, and let $n_t(j)$ denote the age-specific fertility rate for age-j women at time t. Then the new cohort size satisfies:

$$N_t(0) = \sum_{j \in F} \psi_j n_t(j) N_t(j) \tag{1}$$

where F denotes the set of reproductive age groups. The model treats fertility rates as exogenous parameters that can be influenced by policy, rather than as endogenously determined outcomes of household optimization. We consider three fertility scenarios: (i) Baseline (TFR $\approx$ 1.0), (ii) Replacement TFR (fertility scaled to TFR = 2.0), and (iii) High Fertility (baseline $\times$ 1.2).

Individuals face mortality risk captured by the age-dependent survival probability $\alpha_t(j)$, which represents the probability that an age-j individual survives to age $j+1$. The intergenerational population evolution follows:

$$N_{t+1}(j+1) = \alpha_t(j) N_t(j) \tag{2}$$

All population sequences are projected for 51 years (2023–2073), then frozen at the year-50 distribution for the remainder of the transition horizon ($T = 200$ periods). This freeze assumption ensures that the terminal steady state has a well-defined stationary population structure to which the economy can converge.

2.3 Household Sector

Each individual enters the economy at the beginning of adulthood ($j = 0$) and lives until the maximum age $J = 80$ (age 100). Individuals face mandatory retirement at age j_R. The economy features two types of workers—urban and rural—with population fractions ϕ_u and $1 - \phi_u$ respectively. Rural workers have labor efficiency η relative to urban workers (with $\eta < 1$), reflecting urban–rural wage differentials in the data.

The state vector for an urban worker at age j in period t is four-dimensional: $\Omega^u_{i,j,t} = (a_{i,j,t}, b_{i,j,t}, e_{i,j,t}, \gamma_{i,j,t})$, where a denotes tradeable assets, b the pension individual account balance, e the historical average wage, and γ the idiosyncratic labor productivity shock. For rural workers, $b \equiv 0$ (no individual account), so the state vector is three-dimensional: $\Omega^r_{i,j,t} = (a_{i,j,t}, e_{i,j,t}, \gamma_{i,j,t})$.

The individual utility function takes the CRRA form:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma} \tag{3}$$

The lifecycle objective for a type-$s \in \{u, r\}$ individual maximizes:

$$V^s_{i,j,t}(\Omega^s_{i,j,t}) = \max_{c^s_{i,j,t},\, a^s_{i,j+1,t+1}} \left\{ u(c^s_{i,j,t}) + \beta\alpha_t(j)\mathbb{E}\left[V^s_{i,j+1,t+1}(\Omega^s_{i,j+1,t+1})\right] + \beta(1-\alpha_t(j))V^b(a^s_{i,j+1,t+1}) \right\} \tag{4}$$

where $V^b(a) = \psi_1 \frac{(\psi_2+a)^{1-\sigma}}{1-\sigma}$ is the warm-glow bequest utility.

For **urban workers**, the budget constraint is:

$$\begin{aligned}(1+\tau_c)c^u_{i,j,t} + a^u_{i,j+1,t+1} = {} & (1 + r_t(1-\tau_a))a^u_{i,j,t} \\ & + \mathbf{1}_{\{j \le j_R\}}(1-\tau_{w,t}-\tau_b)w_t\ell(j)\gamma_{i,j,t} \\ & + \mathbf{1}_{\{j > j_R\}} s^{\text{urban}}_{i,j,t} + \mathbf{1}_{\{j_{beq} \le j \le \bar{j}_{beq}\}} beq_t \end{aligned} \tag{5}$$

subject to $a^u_{i,j+1,t+1} \ge 0$. Here $\ell(j)$ is the deterministic age-efficiency profile, $\gamma_{i,j,t}$ is the idiosyncratic productivity shock, and τ_b is the total social security contribution rate (urban).

For **rural workers**, the budget constraint differs in the treatment of social security contributions:

$$\begin{aligned}(1+\tau_c)c^r_{i,j,t} + a^r_{i,j+1,t+1} = {} & (1 + r_t(1-\tau_a))a^r_{i,j,t} \\ & + \mathbf{1}_{\{j \le j_R\}}(1-\tau^r_b)(1-\tau_{w,t})w_t\eta\ell(j)\gamma_{i,j,t} \\ & + \mathbf{1}_{\{j > j_R\}} s^{\text{rural}}_{i,j,t} + \mathbf{1}_{\{j_{beq} \le j \le \bar{j}_{beq}\}} beq_t \end{aligned} \tag{6}$$

subject to $a^r_{i,j+1,t+1} \ge 0$. While urban workers face an additive deduction $(1 - \tau_{w,t} - \tau_b)$ on labor income, rural workers face a multiplicative structure $(1 - \tau^r_b)(1 - \tau_{w,t})$, reflecting the separate institutional arrangements. The factor $\eta < 1$ reflects the rural-to-urban efficiency differential.

The idiosyncratic labor productivity shock follows an AR(1) process:

$$\log(\gamma_{i,j+1,t+1}) = \theta \log(\gamma_{i,j,t}) + \nu_{i,j+1,t+1}, \qquad \nu_{i,j+1,t+1} \sim N(0, \sigma_\nu^2) \tag{7}$$

which is discretized into a three-state Markov chain $\gamma \in \{0.36, 1.0, 2.70\}$.

2.3.1 Urban–Rural Pension System

A salient feature of China's social security system is its segmented urban–rural structure, which the model explicitly incorporates.

Urban workers participate in the Urban Employee Basic Pension Insurance (UEBPI), which combines a pooling account with an individual account. For retired urban workers ($j > j_R$), pension benefits are:

$$s_{i,j,t}^{\text{urban}} = \underbrace{\frac{b_{i,j,t}}{j^e - j_R}}_{\text{Individual account}} + \underbrace{\alpha_b \cdot e_{i,j,t} \cdot j_R}_{\text{Pooling account}} \qquad (j > j_R) \tag{8}$$

where j^e is the expected longevity (pension divisor), j_R is the retirement age, α_b is the pooling benefit rate, and $e_{i,j,t}$ is the individual's historical average wage. The pension divisor $j^e - j_R$ and the pooling multiplier j_R adjust with the policy retirement age. Under delayed retirement, per-period benefits increase (through a smaller annuity divisor and more contribution years), partially offsetting the fiscal savings from fewer retirees; the net fiscal gain arises because the reduction in pension payment years dominates.

The two-component structure of equation (8) is central to our modeling choice. Because benefits depend on individual-specific state variables $b_{i,j,t}$ and $e_{i,j,t}$, the aggregate fiscal cost of the pension system at any date t is determined by the *joint distribution* of (b, e) across all retirees—not by a single economy-wide average. This distinction is consequential for reform evaluation: a reduction in the pooling rate α_b has an immediate fiscal effect that scales with each retiree's individual e, whereas a change in the personal account contribution rate

τ^{bw} only affects future accumulation of b, leaving the existing stock of pension claims—and hence short-run expenditures—unchanged. Models that replace the individual-specific benefit formula with a uniform transfer (e.g., a common fraction of the average wage for all retirees) conflate these channels and miss the fiscal inertia embedded in the existing distribution of pension claims.

The individual account accumulates during working years at a fixed pension interest rate r^p:

$$b_{i,j+1,t+1} = \begin{cases} (1+r^p)\, b_{i,j,t} + \tau^{bw}\, w_t\, \ell(j)\, \gamma_{i,j,t} & \text{if } j \le j_R \\ b_{i,j,t} & \text{if } j > j_R \end{cases} \tag{9}$$

where τ^{bw} is the personal account contribution rate and r^p is the administratively set pension account interest rate (distinct from the market rate r_t).

The historical average wage evolution incorporates a *social average wage* mechanism. Define the social average wage as

$$\bar{W}_t = \frac{w_t\, L_t^u}{N_t^{\text{workers,u}}} \tag{10}$$

where L_t^u is aggregate urban effective labor and $N_t^{\text{workers,u}}$ is the number of urban workers. The historical average wage updates as:

$$e_{i,j+1,t+1} = \begin{cases} \dfrac{j \cdot \dfrac{\bar{W}_t}{\bar{W}_{t-1}} \cdot e_{i,j,t} + w_t\, \ell(j)\, \gamma_{i,j,t}}{j+1} & \text{if } j < j_R \\ \dfrac{\bar{W}_t + \tilde{e}_{i,j,t}}{2} & \text{if } j = j_R \\ e_{i,j,t} & \text{if } j > j_R \end{cases} \tag{11}$$

where $\tilde{e}_{i,j,t}$ denotes the working-period update evaluated at age $j = j_R$. This three-way rule has three features. First, the ratio $\bar{W}_t/\bar{W}_{t-1}$ rescales the accumulated average wage to reflect economy-wide wage growth, ensuring that individuals who worked in lower-wage periods are not penalized at retirement. In steady state this ratio equals one, and the formula reduces to the standard cumulative average. Second, at the exact retirement age $j = j_R$, the accumulated average wage is converted into a *deemed contribution index* by averaging it with the current social average wage $\bar{W}_t$, following the institutional rules of China's UEBPI. Third, after retirement ($j > j_R$), the index is frozen.

Rural workers participate in the Urban-Rural Resident Basic Pension (URRBP), which provides a flat-rate benefit proportional to the individual's historical average wage:

$$s_{i,j,t}^{\text{rural}} = \rho_r \cdot e_{i,j,t} \qquad (j > j_R) \tag{12}$$

where ρ_r is the replacement rate. Rural workers do not have individual pension accounts ($b = 0$), reflecting the simpler benefit structure of the URRBP. Unlike urban workers, the rural historical average wage follows a simpler two-case rule without social average wage indexation:

$$e_{i,j+1,t+1}^r = \begin{cases} \dfrac{j \cdot e_{i,j,t}^r + w_t \, \eta \, \ell(j) \, \gamma_{i,j,t}}{j+1} & \text{if } j \leq j_R \\ e_{i,j,t}^r & \text{if } j > j_R \end{cases} \tag{13}$$

There is no wage-growth rescaling ($\bar{W}_t / \bar{W}_{t-1}$) and no deemed-contribution-index conversion at retirement, since the URRBP does not employ these mechanisms.

2.4 Firm Sector

A representative firm operates a Cobb–Douglas production function:

$$Y_t = Z K_t^{\alpha} L_t^{1-\alpha} \tag{14}$$

where Z is TFP and α is the capital share. Profit maximization yields factor prices:

$$r_t = \alpha Z K_t^{\alpha-1} L_t^{1-\alpha} - \delta, \qquad w_t = (1-\alpha) Z K_t^{\alpha} L_t^{-\alpha} \tag{15}$$

Capital market clearing requires $K_t = A_t - D_t$, where A_t is aggregate household assets and D_t is government debt. The resource constraint is:

$$Y_t = C_t + K_{t+1} - (1-\delta)K_t + G_t \tag{16}$$

2.5 Aggregation

Let $\mu^u_{j,t}(\Omega)$ and $\mu^r_{j,t}(\Omega)$ denote the measures of age-j urban and rural individuals, respectively, with state $\Omega = (a, b, e, \gamma)$ at time t. Aggregate quantities are obtained by integrating individual decisions over the joint distribution of states, ages, and worker types.

Aggregate household assets and consumption sum over both types:

$$A_t = \sum_{j=0}^{J-1} \left[\int a_{i,j,t}\, d\mu^u_{j,t} + \int a_{i,j,t}\, d\mu^r_{j,t} \right] \tag{17}$$

$$C_t = \sum_{j=0}^{J-1} \left[\int c_{i,j,t}\, d\mu^u_{j,t} + \int c_{i,j,t}\, d\mu^r_{j,t} \right] \tag{18}$$

Aggregate effective labor reflects the urban–rural efficiency differential:

$$L_t = \sum_{j=0}^{j_R} \int \ell(j)\, \gamma_{i,j,t}\, d\mu^u_{j,t} + \eta \sum_{j=0}^{j_R} \int \ell(j)\, \gamma_{i,j,t}\, d\mu^r_{j,t} \tag{19}$$

where η is the rural-to-urban relative efficiency.

Total pension expenditure separates into the two pension pillars:

$$S_t = \sum_{j=j_R+1}^{J-1} \int s^{\text{urban}}_{i,j,t}\, d\mu^u_{j,t} + \sum_{j=j_R+1}^{J-1} \int s^{\text{rural}}_{i,j,t}\, d\mu^r_{j,t} \tag{20}$$

where $s^{\text{urban}}_{i,j,t}$ comprises the individual account annuity and the pooling benefit (equation 8), and $s^{\text{rural}}_{i,j,t}$ is the flat-rate URRBP benefit (equation 12).

Social security contributions are collected from both urban and rural workers:

$$B_t = \tau_b \, w_t \sum_{j=0}^{j_R} \int \ell(j)\, \gamma_{i,j,t} \, d\mu^u_{j,t} + \tau^r_b (1 - \tau_{w,t}) \, w_t \, \eta \sum_{j=0}^{j_R} \int \ell(j)\, \gamma_{i,j,t} \, d\mu^r_{j,t} \tag{21}$$

Accidental bequests left by deceased individuals of both types are computed from end-of-period (post-savings) assets:

$$Beq_t = \sum_{j=0}^{J-1} (1 - \alpha_t(j)) \left[\int a'_{i,j,t} \, d\mu^u_{j,t} + \int a'_{i,j,t} \, d\mu^r_{j,t} \right] \tag{22}$$

The distributions $\mu^u_{j,t}$ and $\mu^r_{j,t}$ evolve forward according to the household decision rules, the exogenous Markov process for γ, and the fixed urban population fraction ϕ_u.

2.6 Government Sector

The government budget constraint is:

$$G_t + (1 + r_t) D_t + S_t = D_{t+1} + T_t + B_t \tag{23}$$

where G_t is government spending, S_t is total pension expenditure, $T_t = \tau_w w_t L_t + \tau_a r_t A_t + \tau_c C_t$ is total tax revenue (the bequest tax rate $\tau_{beq} = 0$ in the current calibration), and B_t is total social security contributions as defined in equation (21).

Accidental bequests left by deceased individuals are taxed and redistributed as lump-sum transfers to qualifying age groups.

2.6.1 Two Fiscal Closure Modes

We evaluate the fiscal burden of population aging through two parallel fiscal closure modes, providing complementary perspectives on the same underlying fiscal challenge.

τ_w-Adjustment Mode. Government spending is fixed at $G/Y = 21.8\%$ and the target debt ratio is $D/Y = 23.8\%$. In steady state, the labor income tax rate τ^w adjusts to satisfy the government budget constraint:

$$\tau^w = \frac{G + (1+r)D + S - D' - \tau_a r A - \tau_c C - B}{wL} \tag{24}$$

Along the transition path, τ^w is set to a constant transitional rate τ^w_{trans} during the demographic transition phase, then switches to the terminal steady-state value; the scalar τ^w_{trans} is determined by bisection so that accumulated debt reaches the target D/Y at the end of the transition.[1] This mode answers the question: *How much must taxes rise to maintain fiscal sustainability?*

***G*-Adjustment Mode.** All tax rates are fixed at their 2023 levels ($\tau_w = 9.6\%$) and debt is fixed at $D/Y = 23.8\%$. Government spending G_t adjusts residually:

$$G_t = D_{t+1} + T_t + B_t - (1 + r_t)D_t - S_t \tag{25}$$

This mode answers the complementary question: *How much fiscal space is lost to population aging?*

Both modes are computed in parallel for all steady states, transition paths, and policy experiments, enabling a comprehensive assessment of the fiscal burden from two distinct angles.

2.7 Equilibrium Definition

A competitive equilibrium along the transition path consists of household decision rules $\{c_t(\Omega), a'_t(\Omega)\}$, factor prices $\{w_t, r_t\}$, distributions $\{\mu^j_t\}$, and aggregates $\{K_t, L_t, Y_t, C_t\}$ such that: (1) households optimize given prices and policies; (2) firms optimize; (3) labor and capital markets clear; (4) the government budget balances each period; (5) bequests clear; and (6) the resource constraint holds.

[1] This constant-rate transition convention follows Huang et al. [1997] and De Nardi et al. [1999]; see Algorithm 1 in the Appendix for details.

The transition path algorithm employs a dual-loop structure. The *outer loop* (price iteration) initializes price sequences $\{r_t, w_t\}_{t=0}^{T-1}$ by interpolation between the initial condition and the terminal steady state, solves household Bellman equations backward from $t = T - 1$ to $t = 0$ using the endogenous grid method (EGM), simulates distributions forward via Monte Carlo, aggregates to obtain implied $\{K_t, L_t\}$, and updates prices via damped iteration until convergence (squared price distance $(r - r')^2 + (w - w')^2 < 10^{-5}$). The *inner loop* (tax rate shooting, for τ_w-adjustment mode) uses bisection to find a constant transition tax rate τ_w^{trans} such that D/Y reaches its target at the terminal period $t = T$, after which τ_w switches to the terminal steady-state value. In G-adjustment mode, G_t is computed analytically each period, requiring no bisection. the Online Appendix provides details of the EGM solution method and algorithmic procedures.

3 Calibration

This section describes the calibration of all model parameters. Tables 1–5 summarize the externally calibrated parameters; Table 6 reports the internally calibrated targets.

3.1 Demographics

The model period is one year. Individuals enter the economy at age 20 ($j = 0$) and can live to a maximum of age 100 ($j = 79$). The baseline retirement age is 60 ($j_R = 40$). Survival probabilities are calibrated following İmrohoroğlu and Zhao [2018]. Age-specific fertility rates are calibrated to the UN *World Population Prospects* (2024 revision) for China's 2023 population. Population sequences are projected for 51 years (2023–2073) and then frozen at the year-50 distribution. Figure 1 shows the resulting trajectories: under baseline fertility (TFR $\approx$ 1.0), the terminal population falls to 60% of its initial level and the old-age dependency ratio rises from 33.7% to 71.1%; under replacement fertility (TFR = 2.0), the terminal population is 83.9% with a dependency ratio of 42.3%.

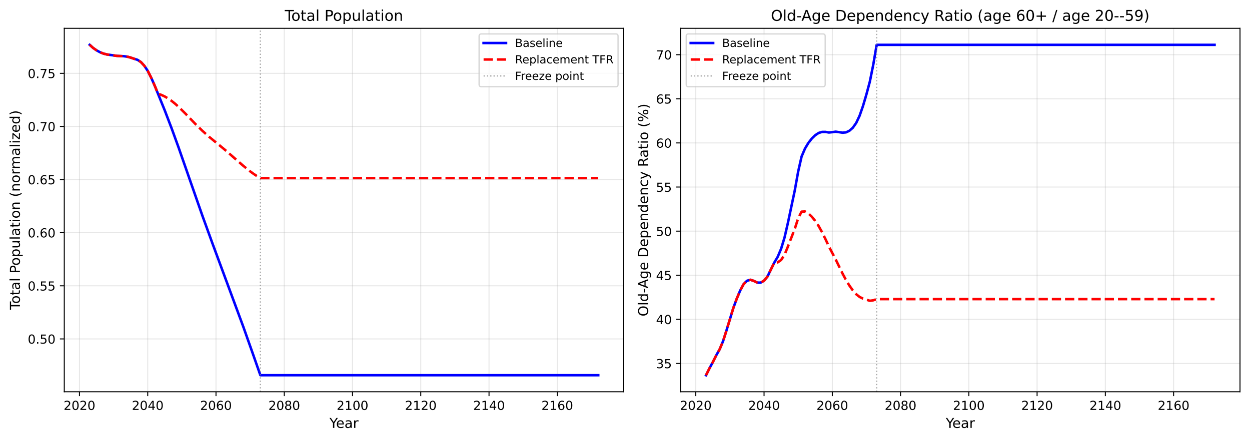


Figure 1: Population Trajectories and Dependency Ratios by Scenario

Figure 2: Population age distribution: baseline scenario, 2023 initial condition vs. terminal steady state. The age profile shifts sharply from a bottom-heavy pyramid in 2023 to a top-heavy distribution in the terminal state, reflecting the combined effect of declining fertility and rising life expectancy embedded in the UN projections. The inset reports the old-age dependency ratio (population 60+ relative to working-age 20–59) for each snapshot.

Table 1: Demographic Parameters

Parameter	Symbol	Value
Lifespan (model ages)	J	80 (age 100)
Retirement age	j_R	40 (age 60)
Expected longevity (pension divisor)	j_e	53 (age 73)
Transition periods	T	200

3.2 Preferences and Income Process

The risk aversion coefficient σ is set to 2, and the discount factor β is set to 0.99 following Xiong et al. [2022]. The bequest weight ψ_1 and curvature ψ_2 are calibrated internally (see Section 3.5); their fitted values are reported in Table 2. The income process follows $\log(\gamma') = 0.86\log(\gamma) + \nu$, $\sigma_\nu^2 = 0.06$, based on Yu and Zhu [2013], discretized to three states $\gamma \in \{0.36, 1.0, 2.70\}$ following İmrohoroğlu and Zhao [2020]. Urban workers have an age-efficiency profile that peaks around age 50 (see the Online Appendix, Figure A.1); rural workers have labor efficiency $\eta \approx 42\%$ of urban workers, reflecting observed wage differentials. The urban population fraction is $\phi_u = 66.6\%$.

Table 2: Preference Parameters

Parameter	Symbol	Value
Discount factor	β	0.9900
Risk aversion	σ	2.00
Bequest weight	ψ_1	1.001
Bequest curvature	ψ_2	1.724

3.3 Production

Following Li [2024], the capital share α is set to 0.5 and the depreciation rate δ to 0.04. TFP Z is calibrated internally to match the 2023 capital-output ratio. The fitted value of Z together with the externally set production parameters is reported below.

Table 3: Production Parameters

Parameter	Symbol	Value
TFP	Z	0.5455
Capital share	α	0.50
Depreciation rate	δ	0.04

3.4 Social Security and Fiscal Policy

The total SS contribution rate is $\tau_b = 24\%$, of which $\tau^{bw} = 8\%$ enters the individual account. The pooling benefit rate $\alpha_b = 1\%$, the pension account interest rate $r^p = 3\%$, and the expected longevity (pension divisor) $j_e = 53$ (age 73). Rural workers receive a flat-rate benefit with replacement rate $\rho_r = 11\%$ and contribute at rate $\tau_b^r = 3.4\%$ of disposable income.

Table 4: Social Security Parameters

Parameter	Symbol	Value
Panel A: Urban Employee Basic Pension Insurance (UEBPI)		
Total SS contribution rate	τ_b	24%
Personal account contribution rate	τ^{bw}	8%
Pooling benefit rate	α_b	1.0%
Account interest rate	r^p	3%
Expected longevity (pension divisor)	j_e	73
Panel B: Urban-Rural Resident Basic Pension (URRBP)		
Replacement rate	ρ_r	11.0%
Contribution rate	τ_b^r	3.4%
Panel C: Urban-Rural Heterogeneity		
Urban population fraction	ϕ_u	66.6%
Rural efficiency (relative to urban)	η	41.9%

Following Lü et al. [2024], fiscal parameters are set as shown in Table 5.

Table 5: Fiscal Policy Parameters

Parameter	Symbol	Value
Labor income tax	τ_w	9.6%
Asset income tax	τ_a	29%
Consumption tax	τ_c	13.3%
Initial debt-output ratio	D/Y	23.8%
Government spending-output	G/Y	21.8%

3.5 Calibration Validation

The three internally calibrated parameters (Z, ψ_1, ψ_2) are chosen to match (i) the 2023 capital-output ratio, and (ii) the social security balance relative to GDP, separately targeting the urban and rural pension deficits. Table 6 reports the fit. The model achieves a K/Y of 4.25, within 1.2% of the data target of 4.30. The total SS deficit of −0.78% of GDP closely matches the data value of −0.73%, with the urban deficit accounting for the bulk of the shortfall. The urban deficit (−0.76% vs. data −0.55%) is partially offset by a near-zero rural deficit (−0.02% vs. data −0.18%), yielding a tight aggregate fit.

Table 6: Model vs. Data Calibration Targets

Moment	Target	Model	Gap
Capital-Output Ratio (K/Y)	4.30	4.20	−2.3%
Total SS Balance/Y	−0.73%	−0.78%	−0.05pp
Urban SS Balance/Y	−0.55%	−0.76%	−0.20pp
Rural SS Balance/Y	−0.18%	−0.02%	+0.16pp

4 Quantitative Analysis

This section presents the quantitative results of our model. We first report the initial condition (2023) and terminal steady-state equilibria (Sections 4.1–4.2), followed by baseline transition paths under two alternative fiscal closure rules (Section 4.3). We then evaluate three policy reform experiments—delayed retirement, reduced pooling transfers, and replacement-level fertility—under both fiscal instruments (Sections 4.4–4.6), and examine their distributional and intergenerational welfare consequences (Section 4.7). A complementary cost-effectiveness perspective on pro-natalist policy is developed separately in Section 5.

4.1 Initial Condition (2023)

The starting point of the transition analysis is an *initial condition* calibrated to match key features of the Chinese economy in 2023. We use the term "initial condition" rather than "initial steady state" deliberately: while labor and capital markets clear, the government maintains constant policy parameters, and the household distribution is stationary given constant prices and policies, the government budget does not balance—government debt is changing over time under the 2023 fiscal configuration (as documented below). The economy is therefore not in a true steady state, but rather in a well-defined initial equilibrium from which the demographic transition begins. Table 7 reports the principal macroeconomic aggregates.

Table 7: Initial Condition: Macroeconomic Aggregates (2023)

Variable	Value	Variable	Value
Interest rate r	0.0790	Aggregate output Y	1.0626
Wage w	0.6251	Aggregate consumption C	0.6473
Aggregate capital K	4.4653	Capital–output ratio K/Y	4.20
Effective labor L	0.8496	Debt–output ratio D/Y	23.80%
SS revenue / Y	10.18%	SS expenditure / Y	10.96%
SS balance / Y	−0.78%	Gov. purchases G/Y	21.80%

The model generates a capital–output ratio of 4.25 against the data target of 4.30, within 1.2%. The social security balance of −0.78% of GDP closely matches the 2023 data value of −0.73%.

Government purchases amount to $G/Y = 21.80\%$ and the social security deficit is 0.78% of GDP. This fiscal picture underscores why we characterize the 2023 calibration as an *initial condition* rather than a steady state: the government budget does *not* balance. We take the debt ratio $D/Y = 23.8\%$ as exogenous at the initial date and require it to be maintained (or reduced) along the transition path. This modeling choice allows us to cleanly decompose the total fiscal burden into two components: a *debt burden*—the cost of servicing the inherited $D/Y = 23.8\%$—and an *aging burden*—the additional fiscal adjustment required by the demographic transition. As Section 4.2.3 shows, this decomposition reveals that population aging accounts for the dominant share of the total fiscal adjustment.

4.2 Terminal Steady States

The terminal steady state corresponds to the long-run equilibrium after the demographic transition is complete and the population distribution has stabilized. Because the model features two fiscal closure rules, we report terminal steady states for each.

4.2.1 τ^w-Adjustment Mode

Under the τ^w-adjustment rule, the wage tax rate adjusts each period to maintain D/Y at or below its cap of 23.80%, while government purchases as a share of GDP remain at their initial level of 21.80%. Table 8 reports the terminal aggregates.

Table 8: Terminal Steady-State Aggregates: τ^w-Adjustment

Variable	Value	Variable	Value
Wage tax τ^w	33.63%	Capital–output ratio K/Y	4.282
Interest rate r	0.0753	Debt–output ratio D/Y	23.85%
Wage w	0.6453	Gov. purchases G/Y	21.80%
SS revenue / Y	10.24%	SS expenditure / Y	22.97%
SS balance / Y	−12.73%		

The demographic transition imposes a dramatic fiscal burden. The wage tax must rise from 9.60% to 33.63% to maintain the debt-to-GDP cap. Effective labor shrinks substantially as the working-age population contracts to 60% of its initial level. Despite the rise in the wage rate (from 0.75 to 0.78), the income base contracts so severely that the tax rate more than triples. The social security deficit widens to 12.56% of GDP, reflecting the surge in pension expenditure (from 10.91% to 22.64% of GDP) relative to the shrunken contribution base. The capital–output ratio rises from 4.25 to 4.40 as labor becomes relatively scarce and the economy becomes more capital-intensive.

4.2.2 G-Adjustment Mode

Under the G-adjustment rule, the wage tax rate remains fixed at its initial value of 9.60%, while government purchases G/Y adjusts downward to satisfy the debt-to-GDP constraint.

Table 9 reports the terminal aggregates.

Table 9: Terminal Steady-State Aggregates: G-Adjustment

Variable	Value	Variable	Value
Wage tax τ^w	9.60%	Capital–output ratio K/Y	5.149
Interest rate r	0.0580	Debt–output ratio D/Y	23.80%
Wage w	0.7594	Gov. purchases G/Y	10.64%

Holding the wage tax constant avoids the distortionary effects of the τ^w-adjustment mode, encouraging greater capital accumulation: K/Y rises to 5.22, compared to 4.40 under τ^w-adjustment. However, the fiscal cost is borne entirely by government purchases, which must fall from 21.80% to 10.64% of GDP—a reduction of more than half. This implies drastic cuts to public goods and services.

The two terminal steady states bracket the range of possible fiscal adjustments. In practice, any sustainable policy will involve some combination of tax increases and spending cuts; the τ^w-adjustment and G-adjustment modes represent the polar cases. Comparing them reveals a fundamental tradeoff: tax adjustment preserves public spending but suppresses private economic activity through distortionary taxation, while expenditure adjustment preserves market incentives but at the cost of reduced public goods provision.

4.2.3 Fiscal Burden Decomposition: Debt vs. Aging

The total fiscal adjustment required between the initial condition and the terminal steady states can be decomposed into two distinct sources: the *debt burden*—the cost of servicing the initial stock of government debt—and the *aging burden*—the additional fiscal pressure arising from population aging. To construct this decomposition, we compute a counterfactual "debt-only" benchmark: the fiscal adjustment that would be required to balance the government budget at *2023 demographics* (i.e., with no population aging) while maintaining $D/Y =$ 23.8%. The difference between this benchmark and the initial policy settings measures the debt burden; the remaining gap to the terminal steady state measures the aging burden.

Table 10 reports the decomposition. The "counterfactual" row shows the fiscal instrument level required to balance the government budget at 2023 demographics—i.e., with no population aging—while maintaining $D/Y = 23.8\%$. The gap between the initial setting and this counterfactual measures the debt burden: the fiscal adjustment needed solely to close the existing government deficit (driven by debt interest payments and the social security shortfall) even without any demographic change. The remaining gap from the counterfactual to the terminal steady state measures the aging burden.

Table 10: Fiscal Burden Decomposition: Debt vs. Aging

	τ_w-Adjustment	G-Adjustment
Initial (2023)	$\tau_w = 9.60\%$	$G/Y = 21.80\%$
Counterfactual (2023 demog., balanced)	$\tau_w = 12.57\%$	$G/Y = 20.42\%$
Terminal (aged demog.)	$\tau_w = 33.63\%$	$G/Y = 10.64\%$
Debt burden	+2.97pp	−1.38pp
Aging burden	+21.07pp	−9.78pp
Total adjustment	+24.03pp	−11.16pp

4.3 Baseline Transition Paths

We solve for the complete general equilibrium transition path from the initial condition (2023) to the terminal steady state over a horizon of 200 periods (years). The transition period during which the population distribution evolves is 51 years; thereafter, the population is frozen at its year-50 distribution and the economy converges to the terminal steady state. Along the transition path, all markets clear and household decisions are optimal at every period.

4.3.1 Transition under τ^w-Adjustment

Figure 3 displays the transition paths of key macroeconomic variables under τ^w-adjustment.

The wage tax rate is set at a constant transitional rate of $\tau_w^{\text{trans}} = 18.61\%$ during the demographic transition phase, before switching to the terminal rate of 33.63%. The tran-

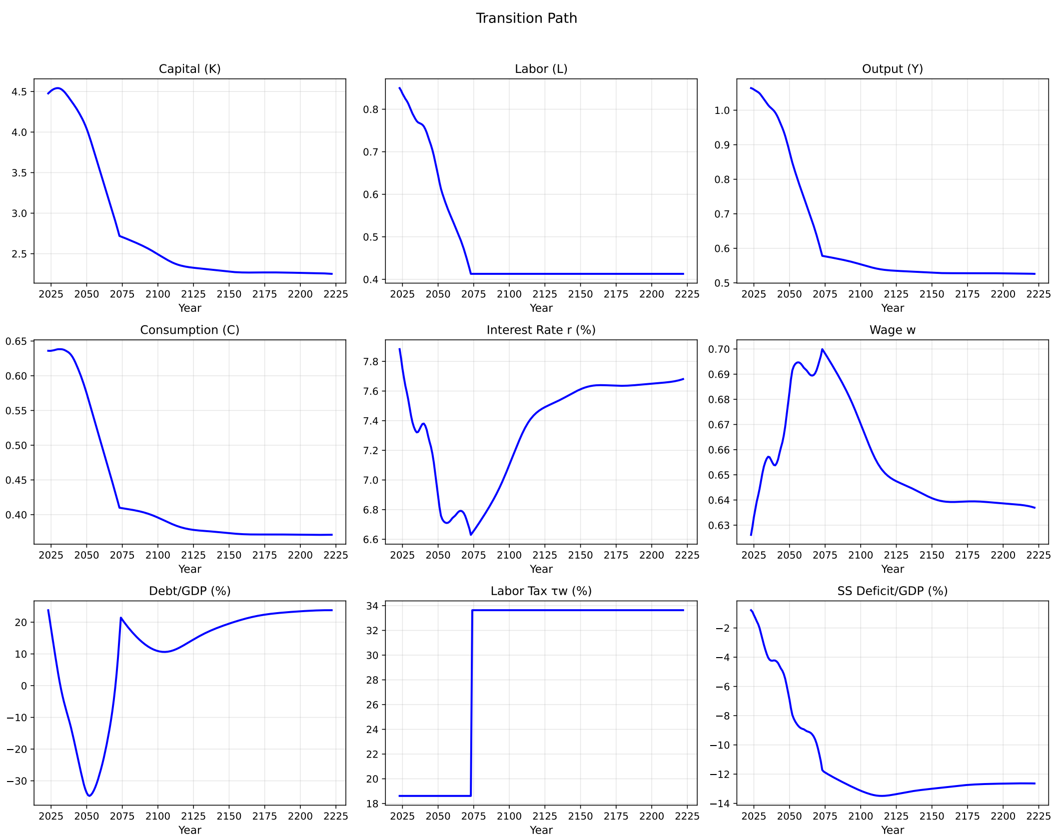


Figure 3: Transition paths: τ^w-adjustment mode

sitional rate of 18.61% is determined by bisection to achieve the target D/Y at the end of the transition phase. Along the transition, the interest rate declines from 7.82% to 7.39%, while the wage rises from 0.75 to 0.78, reflecting the increasing relative scarcity of labor as the working-age population contracts.

4.3.2 Transition under G-Adjustment

Figure 4 displays the transition paths under G-adjustment.

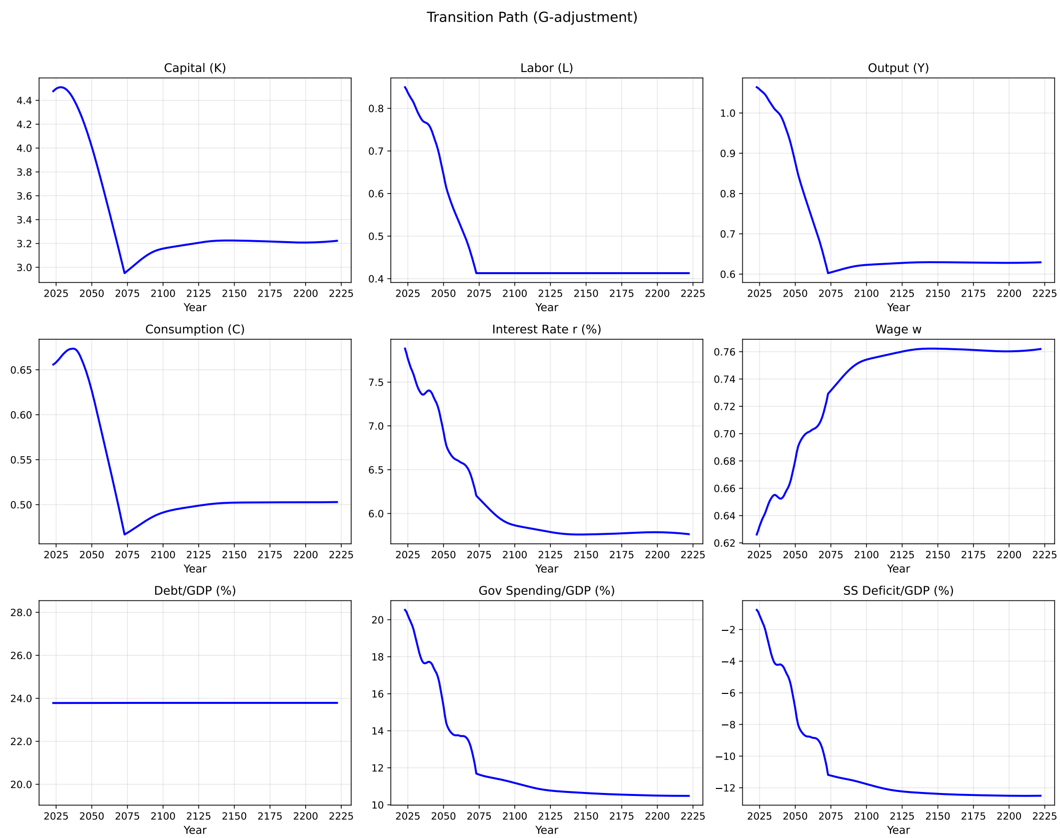


Figure 4: Transition paths: G-adjustment mode

With the wage tax fixed at 9.60%, the government purchases share must fall from its initial level to 10.64% at the terminal steady state. The ratio G/Y declines as the aging population drives up pension expenditure. The debt-to-GDP ratio remains anchored at 23.80% throughout the transition.

Compared with the τ^w-adjustment mode, the G-adjustment path delivers higher output at every point along the transition, because the absence of tax distortion preserves household incentives for saving. However, the decline in G/Y to 10.64%—a halving of the public goods share—has significant welfare implications that are not fully captured by aggregate consumption measures.

4.3.3 Urban–Rural Decomposition of Social Security Deficit

Because the pension systems for urban and rural workers differ fundamentally in generosity and structure, the aggregate social security deficit masks important compositional dynamics. We decompose the deficit $(S - B)/Y$ into its urban and rural components at each period along both transition paths.

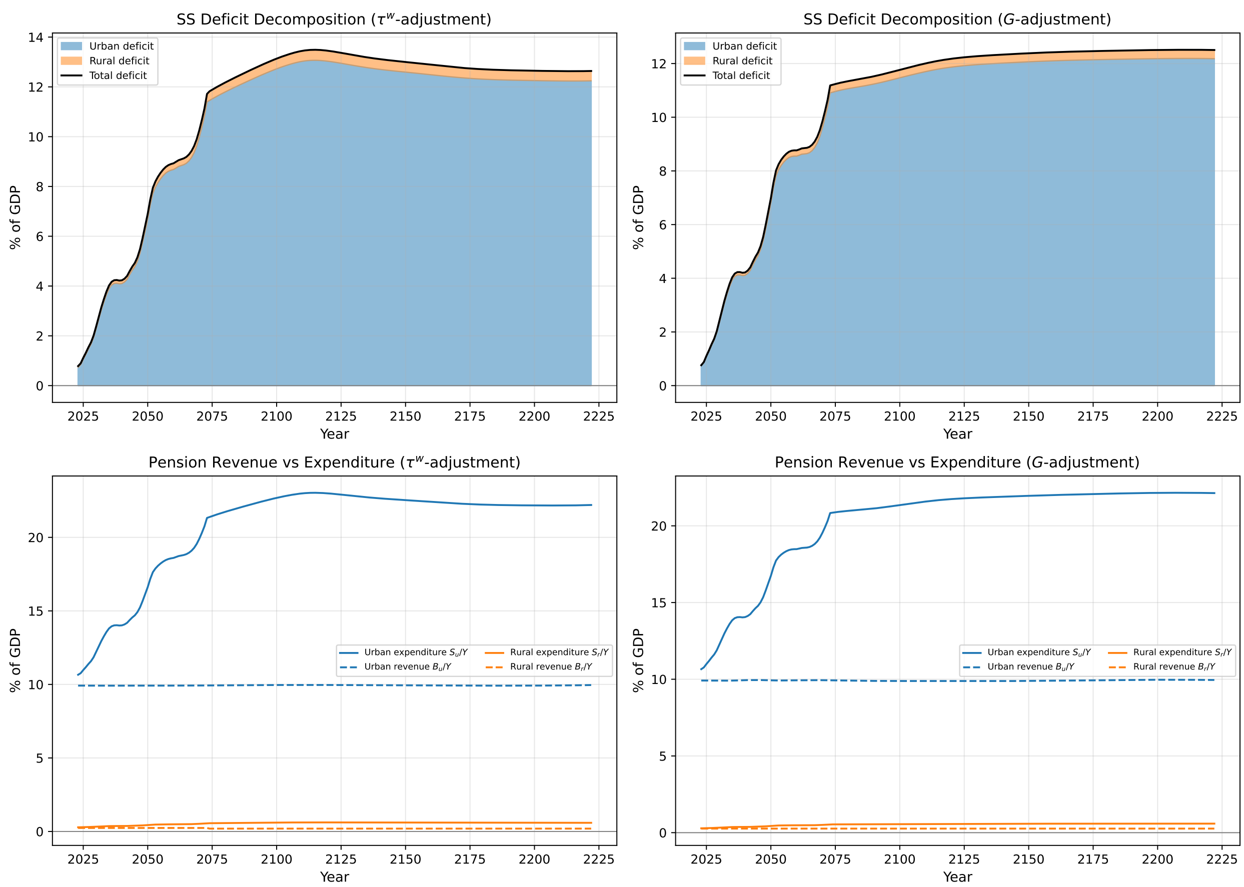


Figure 5: Decomposition of social security deficit into urban and rural components along the transition path. Top panels: stacked area showing the share of the total deficit attributable to each sector (% of GDP). Bottom panels: pension expenditure (S/Y) and contribution revenue (B/Y) by sector.

Table 11: SS Deficit Decomposition over Transition (% of GDP)

Mode	Sector	2023	2033	2043	2053	2073
τ^w-adj	Urban	+0.74%	+3.36%	+4.46%	+7.97%	+11.40%
	Rural	+0.05%	+0.11%	+0.14%	+0.23%	+0.32%
	Total	+0.78%	+3.48%	+4.60%	+8.20%	+11.72%
G-adj	Urban	+0.74%	+3.41%	+4.50%	+8.04%	+10.91%
	Rural	+0.02%	+0.09%	+0.12%	+0.20%	+0.27%
	Total	+0.76%	+3.50%	+4.61%	+8.25%	+11.19%

Figure 5 and Table 11 report the decomposition. Two features stand out. First, the urban pension system accounts for approximately 97% of the total deficit throughout the transition. The urban deficit rises from 0.8% of GDP in 2023 to over 11% by 2073 under τ^w-adjustment, driven by the generous combined-accounts benefit formula: as the urban retiree population swells, the pooling component ($\alpha_b \cdot \bar{e} \cdot j_R$) and personal account drawdowns together far outstrip contribution revenue. The rural deficit, by contrast, remains below 0.3% of GDP even at the end of the transition, reflecting the much lower replacement rate ($\rho_r = 11\%$) and smaller rural workforce.

Second, the bottom panels reveal that the widening deficit is driven almost entirely by the expenditure side. Urban pension expenditure S_u/Y roughly triples over the transition horizon as the baby-boom cohorts retire and the retiree-to-worker ratio climbs, while urban contribution revenue B_u/Y remains relatively stable—or even rises under G-adjustment as higher wages partially offset the shrinking labor force. This asymmetry underscores that the fiscal challenge is fundamentally a pension expenditure problem concentrated in the urban system, rather than a revenue shortfall, and motivates the policy experiments in the following sections that target the expenditure side (delayed retirement, reduced pooling) alongside demographic measures (fertility incentives) that expand the revenue base over longer horizons.

4.4 Policy Experiment Design

Building on the baseline transition path, we evaluate three individual policy reform experiments and one combined experiment. Each individual experiment modifies one structural parameter while all other parameters remain at their baseline values. For each experiment, we solve the complete general equilibrium transition path under both fiscal closure rules (τ^w-adjustment and G-adjustment).

Experiment 1: Delayed Retirement ($j_R = 45$, i.e., age 65). The statutory retirement age is raised from 60 (model age 40) to 65 (model age 45). This extends the working life by five years, simultaneously increasing social security contributions and reducing the pension payment period. Because the pension formula uses the actual retirement age j_R (equation 8), per-period benefits increase under delayed retirement—through a smaller annuity divisor and more contribution years—partially offsetting the fiscal savings from fewer retirees.

Experiment 2: Reduced Pooling Transfers ($\alpha_b = 0.5\%$). The pooling benefit coefficient is halved from 1% to 0.5%, reducing the social-average-wage-based component of pension benefits. This weakens intragenerational redistribution and strengthens the actuarial link between contributions and benefits.

Experiment 3: Replacement-Level Fertility (TFR $= 2.0$). The total fertility rate is raised from the baseline of 1.0 to the replacement level of 2.0, simulating a successful package of fertility incentive policies. The demographic effects materialize gradually: the new cohorts enter the labor market approximately 20 years after birth.

Experiment 4: Combined (Delayed Retirement + Replacement-Level Fertility). This experiment simultaneously raises the retirement age to 65 and the TFR to 2.0, capturing the interaction between supply-side and demographic reforms.

4.5 Policy Experiment Results: τ^w-Adjustment

Table 12 summarizes the policy experiment outcomes under τ^w-adjustment.

Table 12: Policy Experiment Summary: τ^w-Adjustment

Experiment	τ_w^{trans}	τ_w^{term}	$\Delta\tau_w^{trans}$	$\Delta\tau_w^{term}$
Baseline	18.61%	33.63%	—	—
Delayed Retirement (+5 years)	15.43%	27.73%	−3.2pp	−5.9pp
Reduced Pooling ($\alpha_b = 0.5\%$)	12.56%	24.50%	−6.1pp	−9.1pp
Replacement TFR (2.0)	17.60%	18.11%	−1.0pp	−15.5pp
Combined (Delayed Ret. + Repl. TFR)	14.76%	16.00%	−3.9pp	−17.6pp

Figure 6 plots the transition paths for all experiments under τ^w-adjustment.

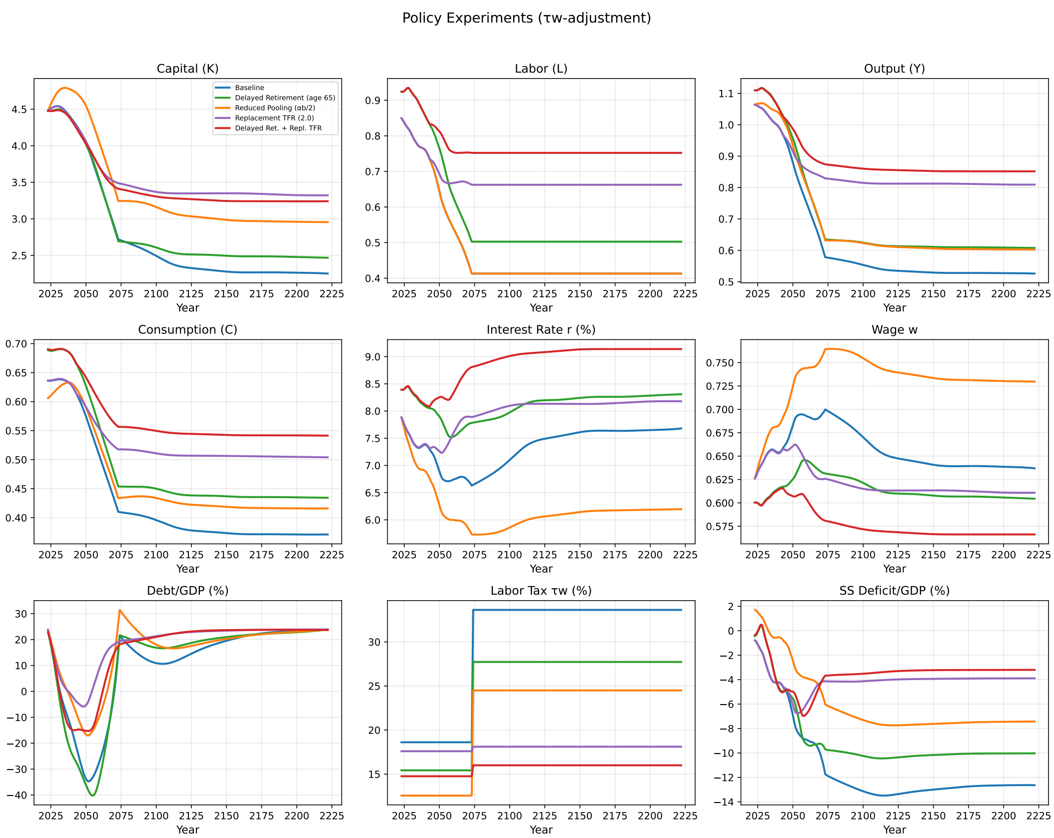


Figure 6: Policy experiment comparison: τ^w-adjustment mode

Reducing pooling transfers achieves the largest transitional tax reduction, lowering the transitional rate by 6.1 percentage points (from 18.61% to 12.56%) and the terminal rate by 9.1 percentage points (from 33.63% to 24.50%). This reform directly reduces pension expenditure by cutting the pooling component of benefits, with immediate effect.

Delayed retirement lowers the transitional rate by 3.2 percentage points and the terminal rate by 5.9 percentage points. The mechanism operates through two channels: extending the working life expands the contribution base, while the shorter retirement period reduces cumulative pension expenditure. However, because the pension formula uses the actual retirement age j_R (equation 8), per-period benefits increase under delayed retirement, partially offsetting the fiscal savings.

Replacement-level fertility has a modest effect on the transitional tax rate (−1.0 percentage point) because the new cohorts have not yet entered the labor force during the peak fiscal pressure period. However, the terminal effect is the largest among single instruments: the terminal tax rate falls by 15.5 percentage points to 18.11%, as the larger working-age population in the long run dramatically expands the contribution base and improves the dependency ratio.

Combining delayed retirement with replacement-level fertility yields the most substantial overall fiscal improvement: the terminal rate falls by 17.6 percentage points to 16.00%, demonstrating the complementarity of supply-side and demographic reforms.

4.6 Policy Experiment Results: G-Adjustment

Table 13 summarizes the outcomes under G-adjustment, where the wage tax is held fixed at 9.60% and government purchases adjust.

Table 13: Policy Experiment Summary: G-Adjustment

Experiment	G/Y^{term}	$\overline{G/Y}$	$\min(G/Y)$	$\Delta G/Y^{term}$
Baseline	10.64%	12.13%	10.48%	—
Delayed Retirement (+5 years)	13.49%	14.53%	13.42%	+2.9pp
Reduced Pooling ($\alpha_b = 0.5\%$)	14.24%	16.12%	14.74%	+3.6pp
Replacement TFR (2.0)	17.86%	17.73%	15.57%	+7.2pp
Combined (Delayed Ret. + Repl. TFR)	18.75%	18.66%	15.62%	+8.1pp

Figure 7 plots the transition paths under G-adjustment, and Figure 8 provides a fiscal decomposition.

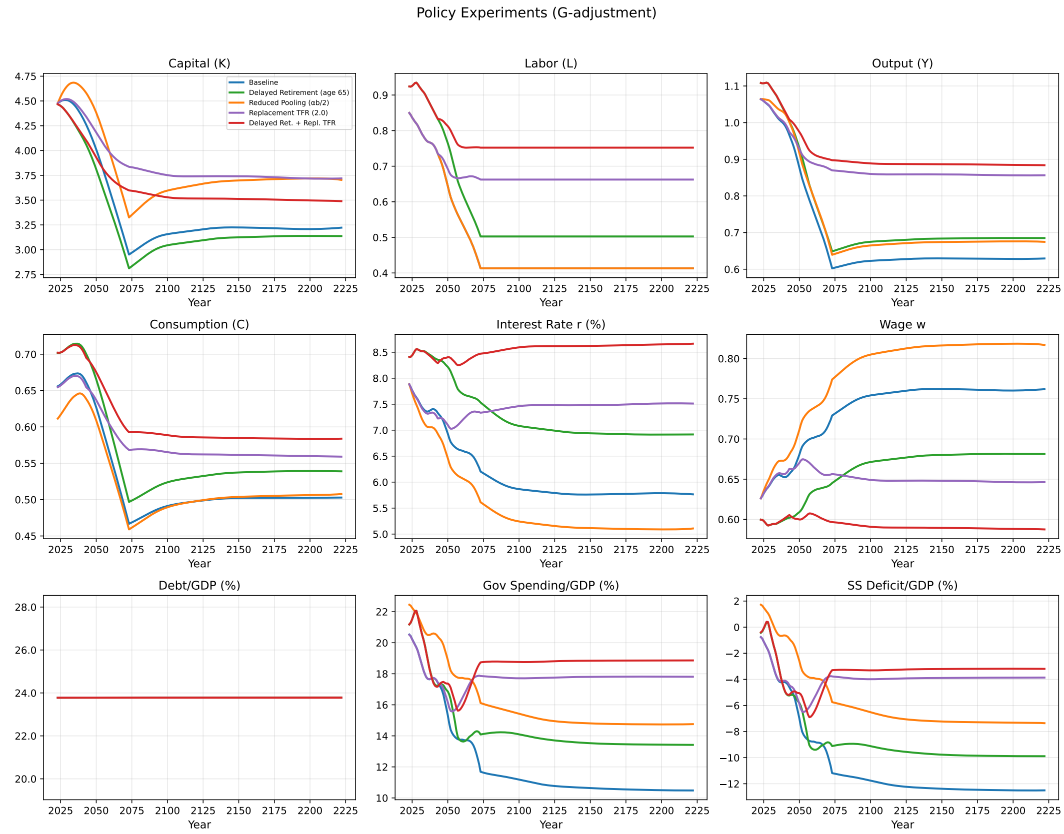


Figure 7: Policy experiment comparison: G-adjustment mode

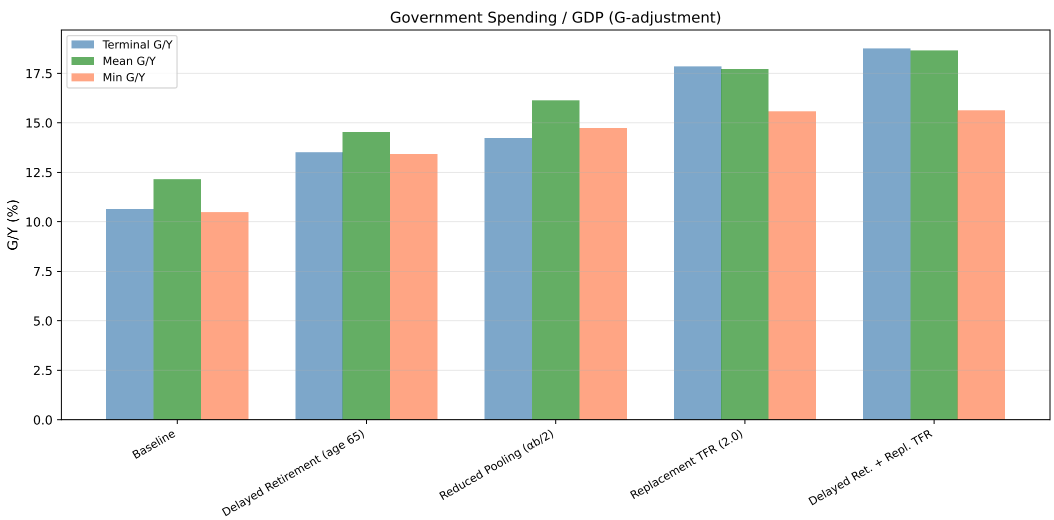


Figure 8: Fiscal decomposition under G-adjustment

Under G-adjustment, the baseline terminal G/Y falls to 10.64%—less than half its initial value. Each reform preserves a larger share of government spending. Replacement-level fertility raises the terminal G/Y by 7.2 percentage points to 17.86%, the largest improvement among single instruments, demonstrating the long-run demographic dividend. Reduced pooling raises it by 3.6 percentage points to 14.24%. Delayed retirement raises it by 2.9 percentage points to 13.49%.

The G-adjustment results largely mirror the τ^w-adjustment findings: reduced pooling is effective immediately, fertility policy delivers the bulk of its benefits in the long run, and delayed retirement provides fiscal relief but its impact is partially offset by higher per-period pension benefits. Combining delayed retirement with replacement-level fertility raises terminal G/Y by 8.1 percentage points to 18.75%, nearly restoring the initial level of public spending. The quantitative magnitudes confirm that the fiscal challenge is substantial under either closure rule—even with the most effective single reform, the terminal G/Y remains below the initial 21.80%, indicating that a combination of reforms is necessary to fully preserve public goods provision.

4.7 Distributional and Welfare Effects

The heterogeneous-agent structure of the model delivers a rich cross-sectional distribution of wealth, consumption, income, and lifetime utility across urban and rural workers, age groups, and idiosyncratic productivity histories. This subsection reports three slices of that distribution: (i) cross-sectional inequality at the 2023 baseline and after the demographic transition, (ii) the urban–rural welfare differential, and (iii) intergenerational redistribution along the transition.

A note on welfare measurement. The model abstracts from any direct utility value of government consumption G. Absolute welfare comparisons *across* fiscal modes are therefore contaminated: under G-adjustment, the collapse of G/Y from 21.8% to 10.64% shows up in the model as a welfare gain because households retain more income, but the foregone public goods would offset this to an unknown degree. We are transparent about this limitation and accordingly focus on *relative* welfare comparisons that are robust to the omission: comparisons between urban and rural agents within the same fiscal mode (both face the same G), and comparisons across birth cohorts within a fixed (mode, experiment) pair (each cohort faces the same G path). These relative measures remain economically meaningful even if the absolute level of welfare is mismeasured. For completeness, aggregate and terminal-SS policy CEV numbers are reported in the Online Appendix (Tables A.3–A.4), where we caution that they should be read as upper bounds rather than point estimates.

Cross-sectional inequality. Table 14 reports weighted Gini coefficients, the P90/P10 wealth ratio, and top-10% and bottom-50% wealth shares across the three steady states. The weighted Gini is computed via the trapezoid rule on the Lorenz curve, with each agent at age j receiving weight proportional to $N(j)/N_{\text{cohort}}$ so that the distributional statistics are consistent with the equilibrium aggregation. The 2023 baseline features a wealth Gini of 0.450, with the urban wealth Gini (0.472) substantially exceeding the rural (0.403)—a direct signature of the richer heterogeneity embedded in the urban pension-account (b, e) state and the productivity shock process. The top 10% of the wealth distribution holds 29.1% of total wealth, while the bottom 50% holds 18.7%.

Under the demographic transition, wealth inequality declines under both fiscal modes: by 0.042 points under τ^w-adjustment and by 0.032 points under G-adjustment. The channels differ. τ^w-adjustment compresses after-tax income dispersion as the payroll tax rises sharply, so the consumption Gini falls from 0.337 to 0.281. G-adjustment leaves τ^w at its initial level, so the market income distribution is less compressed; the inequality decline there is driven primarily by the demographic shift toward a more homogeneous age structure. Age-group and transition-path decompositions of these inequality dynamics—by worker type, age group, and birth cohort—are reported in the Online Appendix (Table A.2, Figures A.2 and A.3).

Table 14: Inequality: Steady-State Comparison

	Initial (2023)	Terminal (τ^w-adj)	Terminal (G-adj)
Gini coefficients			
Wealth	0.450	0.408	0.418
Urban	0.472	0.427	0.455
Rural	0.403	0.312	0.328
Consumption	0.337	0.281	0.287
Income	0.348	0.295	0.320
Wealth distribution			
P90/P10	20.2	14.0	15.7
Top 10% share	29.1%	25.5%	26.5%
Bottom 50% share	18.7%	20.9%	20.5%

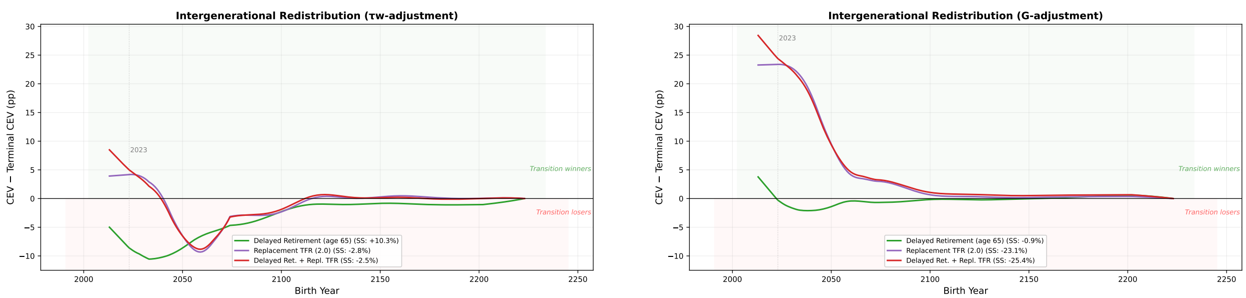
Urban–rural welfare divergence. The model delivers a sharp cross-sectional welfare gap between urban and rural agents under τ^w-adjustment: the urban–rural CEV differential is about 15 percentage points (urban −9.94% versus rural +5.07%, Table A.3). This differential *is* robust to the G-omission limitation, because urban and rural agents face the same government-spending path and the unmeasured utility value of G cancels in the comparison. Two forces drive the divergence. First, the payroll tax rises from 9.6% to 33.6% to maintain fiscal balance, and this burden falls disproportionately on urban workers—who pay the full 24% SS contribution rate on top—while rural workers face a much smaller 3.4% rural rate. Second, rural workers benefit from the general-equilibrium wage increase induced by labor scarcity without bearing the same tax burden. Under G-adjustment, where the payroll tax stays fixed, the urban–rural differential shrinks to roughly 3 percentage points. These findings identify a first-order equity constraint on reform design: tax-based fiscal adjustment concentrates the welfare loss on the urban workforce, raising a political-economy concern that constrains reform design.

Intergenerational redistribution. The cross-sectional differential masks an equally important intergenerational pattern. For each birth cohort t we compute the cohort-specific

CEV

$$\lambda_t = \left(\frac{V_t^{\text{exp}}}{V_t^{\text{base}}} \right)^{1/(1-\sigma)} - 1, \tag{26}$$

where V_t^{base} and V_t^{exp} are expected discounted lifetime utility under the baseline demographic-transition path and under an experiment transition, respectively, within a fixed fiscal mode. Because G is held constant across experiments within the same mode, this differential measure is robust to the welfare-level caveat: each cohort's baseline and experiment values share the same (unmeasured) G contribution. Figure 9 plots the cohort-CEV profile by birth year, for each reform and each fiscal mode.



(a) τ^w-adjustment (b) G-adjustment

Figure 9: Cohort CEV by birth year for each reform relative to the baseline demographic transition, expressed in percentage points and de-meaned by the terminal-SS policy CEV so that zero represents the long-run benchmark. Positive (negative) values identify the cohorts that gain (lose) from each reform. Because each comparison holds G fixed within the fiscal mode, these differentials are robust to the welfare-level caveat discussed above.

Three findings emerge. First, under τ^w-adjustment, fertility-raising policies deliver large relative gains (+11 to +14 percentage points) to cohorts born in the first two decades of the transition: these cohorts enjoy improved dependency ratios by the time they retire, without bearing the full burden of the tax adjustment. Conversely, delayed retirement imposes losses of roughly -10 percentage points on cohorts already close to retirement in 2023, who face unexpected extensions of their working lives. Second, under G-adjustment, the intergenerational incidence is markedly smoother: because the fiscal adjustment falls on public spending rather than on labor income, within-cohort redistribution through the tax wedge is muted, and no cohort is clearly singled out. Third, combining delayed retirement with replacement-level fertility reshapes the profile in both modes: the fertility dividend benefits later cohorts while the retirement-age shift smooths some of the early-cohort losses that pure fertility reform would leave in place. This intergenerational heterogeneity—which an aggregate CEV number cannot capture—is precisely the quantitative content that a rich heterogeneous-agent OLG framework adds to the fiscal sustainability discussion, and it motivates the combined policy package that our analysis supports.

5 Fiscal Value of Birth

The policy experiments in Section 4 treated raising the total fertility rate to replacement level as one structural reform among several, evaluating its fiscal effect *conditional on* the assumed demographic response. Whether pro-natalist policy can actually raise births is, however, empirically uncertain. This section takes a complementary angle: regardless of how fertility responds to any particular policy, what is the maximum per-birth subsidy that the government can justify on purely fiscal grounds? We define the *fiscal value of birth* as the present-value fiscal gain per additional birth, discounted at the equilibrium interest rate. This quantity serves as an upper bound for the government's willingness to pay for pro-natalist subsidies from a purely fiscal perspective and provides a cost-effectiveness benchmark that sidesteps the fertility-responsiveness question.

5.1 Definition and Baseline Estimates

Formally, the fiscal value of birth is the constant per-birth amount FV satisfying

$$\sum_{t=0}^{T-1} \frac{\text{FV} \cdot \Delta N_{0,t}}{R_t} = \sum_{t=0}^{T-1} \frac{\Delta \mathcal{F}_t}{R_t}, \qquad R_t \equiv \prod_{s=0}^{t-1} (1 + r_s), \tag{27}$$

where $\Delta N_{0,t} = \max(N_{0,t}^{\text{tfr}} - N_{0,t}^{\text{base}},\, 0)$ is additional newborns and $\Delta \mathcal{F}_t$ is the fiscal gain at time t. Under G-adjustment, $\Delta \mathcal{F}_t = G_t^{\text{tfr}} - G_t^{\text{base}}$; under τ^w-adjustment, $\Delta \mathcal{F}_t = (\tau_t^{w,\text{base}} - \tau_t^{w,\text{tfr}}) \cdot$

$w_t \cdot L_t$. Both the fiscal gains and the additional births are discounted at the equilibrium interest rate, so:

$$\text{FV} = \frac{\sum_{t=0}^{T-1} \Delta\mathcal{F}_t / R_t}{\sum_{t=0}^{T-1} \Delta N_{0,t} / R_t}. \tag{28}$$

In practice, pro-natalist subsidies are typically disbursed as recurring benefits—childcare allowances, education grants, or tax credits—rather than a single lump sum at birth. To facilitate comparison with such annual transfers, we re-express the fiscal value as a constant annual payment FV^{ann} made from birth to age 19 (20 annual payments). Defining the annuity factor

$$A_t \equiv \sum_{k=0}^{19} \frac{1}{R_{t+k}} \tag{29}$$

as the present value of a 20-year unit stream starting at time t, the annuity fiscal value satisfies

$$\text{FV}^{\text{ann}} = \frac{\sum_{t=0}^{T-1} \Delta\mathcal{F}_t / R_t}{\sum_{t=0}^{T-1} \Delta N_{0,t} \cdot A_t}. \tag{30}$$

Since $A_t > 1/R_t$, the annuity denominator exceeds the lump-sum denominator, so $\text{FV}^{\text{ann}} < \text{FV}$: spreading the same fiscal dividend over 20 annual payments naturally yields a smaller per-year amount. For births beyond the model horizon, R_{t+k} is extended using the terminal steady-state interest rate.

Model-unit values are converted to CNY using China's 2023 GDP of 126.06 trillion yuan and population of 1.41 billion.

We compute the fiscal value under two comparisons. First, "TFR vs Baseline" compares replacement-level fertility (TFR = 2.0) against the status quo (TFR = 1.0), measuring the pure effect of fertility on fiscal space. Second, "TFR + Retirement vs Retirement" compares the combined policy (TFR = 2.0 plus delayed retirement to age 65) against delayed retirement alone, isolating the marginal fiscal value of fertility when retirement reform is already in place.

Table 15 reports the per-birth fiscal value in both lump-sum and annuity form. In the "TFR vs Baseline" comparison, the lump-sum fiscal gain per additional birth is 12.2 10k CNY

Table 15: Fiscal Value of an Additional Birth

	TFR vs Baseline		TFR + Ret. vs Ret.	
	G-adj	τ^w-adj	G-adj	τ^w-adj
Lump-sum FV (10k CNY)	12.2	10.2	7.3	6.6
Annuity FV (10k CNY/yr, 20 yr)	1.1	0.9	0.7	0.6

Note: "TFR vs Baseline" compares replacement-level fertility (TFR = 2.0) against the status quo (TFR = 1.0). "TFR + Ret. vs Ret." compares the combined policy (TFR = 2.0 with delayed retirement to age 65) against delayed retirement alone, isolating the marginal fiscal value of fertility when retirement reform is already in place. Lump-sum FV is a one-time payment at birth; annuity FV is a constant annual payment for 20 years (ages 0–19). All values discounted at the equilibrium interest rate. All values in 2023 CNY (1 unit = 10,000 yuan).

under G-adjustment and 10.2 10k CNY under τ^w-adjustment. These represent the maximum one-time subsidy the government could pay per additional birth while still breaking even in present value terms. Expressed as a 20-year annual payment, the annuity fiscal value is 1.08 10k CNY/yr under G-adjustment and 0.91 10k CNY/yr under τ^w-adjustment—equivalent to approximately 10,800 and 9,100 yuan per year for the first two decades of life. The "TFR + Retirement vs Retirement" comparison yields somewhat lower values—7.3 10k CNY and 6.6 10k CNY lump-sum (0.70 and 0.63 10k CNY/yr annuity), respectively—reflecting the fact that delayed retirement already captures part of the fiscal improvement by extending contribution periods, so the marginal gain from additional births is smaller when retirement reform is already in place. Nevertheless, the fiscal value remains substantial across both comparisons, confirming that pro-natalist subsidies up to approximately 66,000–122,000 yuan per birth—or annual childcare transfers of 6,300–10,800 yuan for 20 years—can be fiscally justified.

5.2 Fiscal Returns at Intermediate Fertility Levels

The preceding analysis assumed replacement-level fertility (TFR = 2.0) as a benchmark. In practice, China's current TFR of approximately 1.0 is unlikely to reach replacement level even with aggressive pro-natalist policies; a more realistic target may be TFR = 1.3–1.5. We therefore compute fiscal values across a range of hypothetical fertility levels from TFR = 1.1 to 2.0.

Table 16: Fiscal Value of Birth Across Fertility Levels (10k CNY)

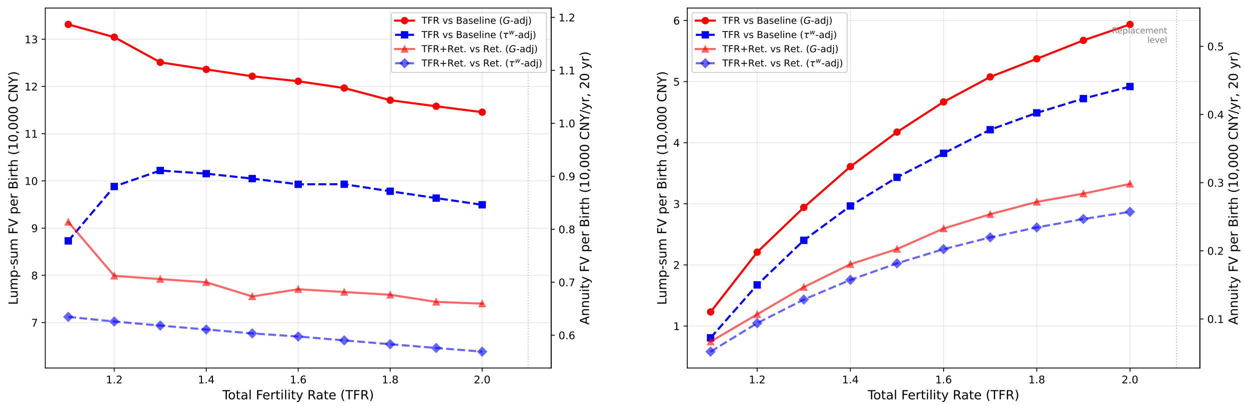
	Lump-sum FV				Annuity FV (20 yr)			
	Marginal		Universal		Marginal		Universal	
TFR	G	τ^w	G	τ^w	G	τ^w	G	τ^w
1.1	13.3	8.7	1.2	0.8	1.19	0.78	0.11	0.07
1.3	12.5	10.2	2.9	2.4	1.12	0.91	0.26	0.22
1.5	12.2	10.0	4.2	3.4	1.09	0.90	0.37	0.31
1.8	11.7	9.8	5.4	4.5	1.04	0.87	0.48	0.40
2.0	11.5	9.5	5.9	4.9	1.02	0.84	0.53	0.44

Note: Marginal FV is the fiscal gain per additional birth; universal FV spreads the gain over all newborns. Lump-sum FV is a one-time payment at birth; annuity FV is a constant annual payment for 20 years (ages 0–19). All values discounted at the equilibrium interest rate, in 2023 CNY (1 unit = 10,000 yuan).

Table 16 and Figure 10 report the results. Two key patterns emerge. First, the *marginal* fiscal value per additional birth *decreases* with TFR: from 13.3 10k CNY at TFR = 1.1 to 11.5 10k CNY at TFR = 2.0 under G-adjustment. This reflects diminishing marginal returns—the first additional births alleviate the most acute labor scarcity, while subsequent births face a crowded labor market and lower marginal productivity. The implication for policy design is striking: even modest fertility increases are fiscally valuable, and the per-birth fiscal justification for subsidies is actually *strongest* at low fertility levels closest to the current situation.

Second, the *universal* fiscal value—the subsidy that can be justified per actual newborn (not just per additional birth)—*increases* with TFR, from 1.2 10k CNY at TFR = 1.1 to 5.9 10k CNY at TFR = 2.0 under G-adjustment. This is because larger fertility programs generate greater total fiscal gains that are spread over more births, yielding higher per-birth subsidies.

In the policy-relevant range of TFR = 1.3–1.5, the marginal lump-sum fiscal value is 12.2–12.5 10k CNY under G-adjustment and 10.0–10.2 10k CNY under τ^w-adjustment, meaning



(a) Marginal FV (per additional birth)

(b) Universal FV (per actual birth)

Figure 10: Fiscal value of birth as a function of target TFR. Left: marginal (per additional birth); right: universal (per actual birth under the policy). The left y-axis shows the lump-sum present value (10k CNY paid once at birth); the right y-axis shows the annuity-equivalent constant annual payment over 20 years (10k CNY/yr). The two axes are linked by a single scalar conversion (mean annuity factor across scenarios, $\approx 1/15$); deviation from exact proportionality across scenarios is within line-width visual tolerance. Both panels plot four scenarios: G-adjustment and τ^w-adjustment, with and without concurrent delayed retirement reform.

the government can justify per-birth subsidies of 100,000–125,000 yuan while remaining fiscally neutral in present value terms. Equivalently, the annuity fiscal value in this range is 1.09–1.12 10k CNY/yr under G-adjustment, implying that an annual childcare subsidy of approximately 10,900–11,200 yuan per child for the first 20 years of life would be fiscally self-financing.

These results have a clear policy implication: fertility subsidies need not achieve replacement-level fertility to be cost-effective. A realistic program that raises TFR from 1.0 to 1.3 generates nearly the same per-birth fiscal return as one that achieves TFR $= 2.0$, though the total fiscal dividend is naturally smaller. The optimal subsidy level should therefore be calibrated to local fertility responsiveness rather than pegged to a replacement-level benchmark.

6 Conclusion and Policy Implications

This paper asks what policy reforms can restore fiscal sustainability for China's pension system under rapid population aging. Using a heterogeneous-agent OLG model with China's combined-accounts pension structure and two polar fiscal closure rules, we find that absent reform the pension deficit widens to about 12.7% of GDP in the long run. The three reform instruments form a clear ranking: reduced pooling delivers the largest immediate fiscal savings but cannot restore long-run sustainability and hurts low-earners; delayed retirement immediately expands the contributor base but is bounded by higher per-period benefits; only raising fertility addresses the root cause and yields the dominant long-run dividend, though with a two-decade lag. Delayed retirement and higher fertility are therefore complementary: combining them reduces the terminal payroll tax by 17.6 percentage points, nearly restoring fiscal balance.

Because fertility responsiveness to policy is uncertain, we complement the reform analysis with the *fiscal value of birth*—the present-value fiscal gain per additional birth—as a cost-effectiveness benchmark. Even modest increases to a TFR of 1.3–1.5 justify per-birth subsidies of 100,000–125,000 yuan—or equivalently, annual childcare transfers of approximately 9,000–11,000 yuan for 20 years—and the marginal value is highest near the current TFR, so even small fertility improvements are cost-effective.

Three policy implications follow. First, no single instrument suffices: the optimal sequence combines benefit adjustments for immediate relief, delayed retirement for short-run structural improvement, and pro-natalist policies for long-run demographic correction. Second, the choice of fiscal closure matters: tripling the payroll tax is highly distortionary, while halving government spending avoids the tax wedge at the cost of public goods outside the model, and any feasible policy combines elements of both. Third, the fiscal value of birth offers a rigorous, model-based criterion for evaluating the birth subsidies that several Chinese provinces have recently introduced.

References

S.R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994.

A.J. Auerbach and L.J. Kotlikoff. Evaluating fiscal policy with a dynamic simulation model. *The American Economic Review*, 77(2):49–55, 1987.

Neha Bairoliya, David Canning, Ray Miller, and Akshar Saxena. The macroeconomic and welfare implications of rural health insurance and pension reforms in china. *The Journal of the Economics of Ageing*, 11:71–92, 2018.

T. Bewley. Stationary monetary equilibrium with a continuum of independently fluctuating consumers. In *Contributions to Mathematical Economics in Honor of Gérard Debreu*, volume 79. 1986.

Juan C. Conesa and Dirk Krueger. Social security reform with heterogeneous agents. *Review of Economic Dynamics*, 2(4):757–795, 1999.

M. De Nardi, S. İmrohoroğlu, and T.J. Sargent. Projected us demographics and social security. *Review of Economic Dynamics*, 2(3):575–615, 1999.

Yuanyuan Deng, Hanming Fang, Katja Hanewald, and Shang Wu. Delay the pension age or adjust the pension benefit? implications for labor supply and individual welfare in china. *Journal of Economic Behavior & Organization*, 212:1192–1215, 2023.

Hanming Fang and Jin Feng. The chinese pension system. Working Paper 25088, National Bureau of Economic Research, 2018.

Hans Fehr, Manuel Kallweit, and Fabian Kindermann. Should pensions be progressive? *European Economic Review*, 63:94–116, 2013.

Qingen Gai, Naijia Guo, Bingjing Li, Qinghua Shi, and Xiaodong Zhu. Rural pensions, labor reallocation, and aggregate income: An empirical and quantitative analysis of china. *Econometrica*, 93(5):1663–1696, 2025.

Hui He, Lei Ning, and Dongming Zhu. The impact of rapid aging and pension reform on savings and the labor supply. Working Paper 19/61, International Monetary Fund, 2019.

H. Huang, S. İmrohoroğlu, and T.J. Sargent. Two computations to fund social security. *Macroeconomic Dynamics*, 1(1):7–44, 1997.

Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, 1996.

Sagiri Kitao. Sustainable social security: Four options. *Review of Economic Dynamics*, 17 (4):756–779, 2014.

Dirk Krueger and Felix Kubler. Pareto-improving social security reform when financial markets are incomplete? *American Economic Review*, 96(3):737–755, 2006.

Dirk Krueger and Alexander Ludwig. On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare. *Journal of Monetary Economics*, 54(1):49–87, 2007.

Dirk Krueger, Kurt Mitman, and Fabrizio Perri. Macroeconomics and household heterogeneity. In John B. Taylor and Harald Uhlig, editors, *Handbook of Macroeconomics*, volume 2, pages 843–921. Elsevier, 2016.

L. Li. Triple pressures: Superimposed shocks and 'dual pillar' regulatory strategy. *Management World*, (8), 2024. In Chinese.

Alexander Ludwig, Dirk Krueger, and Axel Börsch-Supan. Demographic change, relative factor prices, international capital flows, and their differential effects on the welfare of generations. In Jeffrey R. Brown, Jeffrey B. Liebman, and David A. Wise, editors, *Social*

Security Policy in a Changing Environment, pages 385–414. University of Chicago Press, Chicago, 2009.

B. Lü, F. Zeng, H. Tu, and Z. Li. China's fiscal sustainability: Research framework and comprehensive assessment. *Management World*, (1), 2024. In Chinese.

Shinichi Nishiyama and Kent Smetters. Does social security privatization produce efficiency gains? *The Quarterly Journal of Economics*, 122(4):1677–1719, 2007.

Zheng Song, Kjetil Storesletten, Yikai Wang, and Fabrizio Zilibotti. Sharing high growth across generations: Pensions and demographic transition in china. *American Economic Journal: Macroeconomics*, 7(2):1–39, 2015.

C. Xiong, Y. Zhou, and H. Jin. Regional effects of local government implicit debt: A bank network perspective. *Economic Research Journal*, (7), 2022. In Chinese.

J. Yu and G. Zhu. How uncertain is household income in china. *Economics Letters*, 120(1): 74–78, 2013.

A. İmrohoroğlu and K. Zhao. The chinese saving rate: Long-term care risks, family insurance, and demographics. *Journal of Monetary Economics*, 96:33–52, 2018.

A. İmrohoroğlu and K. Zhao. Household saving, financial constraints, and the current account in china. *International Economic Review*, 61(1):71–103, 2020.

S. İmrohoroğlu and S. Kitao. Social security reforms: Benefit claiming, labor force participation, and long-run sustainability. *American Economic Journal: Macroeconomics*, 4(3): 96–127, 2012.

S. İmrohoroğlu, S. Kitao, and T. Yamada. Achieving fiscal balance in japan. *International Economic Review*, 57(1):117–154, 2016.

Online Appendix

Delayed Retirement or More Births? Short-Run Relief and Long-Run Sustainability of China's Pension System

Not for publication — available as supplementary material

A.1 Solution Method and EGM Derivation

The household problem is solved using the Endogenous Grid Method (EGM), which avoids costly grid search over consumption by exploiting the Euler equation's analytical structure. For each discrete state (b, e, γ) at age j and each candidate savings level a' on the exogenous asset grid, the algorithm: (i) computes next-period states (b', e') from the pension and wage-index transition equations; (ii) interpolates the expected marginal value $\mathbb{E}[\partial V_{j+1}/\partial a']$ from the stored continuation value arrays; (iii) solves the Euler equation for consumption via the inverse marginal utility; and (iv) backs out the *endogenous* current asset level from the budget constraint. The resulting pairs $(a^{\text{endo}}, c^{\text{EGM}})$ are then interpolated onto the exogenous asset grid, with a borrowing-constraint correction applied where the endogenous grid falls below the minimum asset level. The marginal value $\partial V_j/\partial a$ is stored via the envelope condition for the next backward step.

Euler equation in marginal-value form. The code iterates on the marginal value function $\partial V_j/\partial a$ rather than the value function itself. Define $R_t \equiv 1 + r_t(1 - \tau_a)$ as the after-tax gross return on assets and $y_{j,t}$ as non-asset income (wages net of taxes for workers, pension benefits for retirees, plus bequest transfers for eligible ages). For each savings grid point a'_i and discrete state (b, e, γ), the algorithm first computes the expected marginal value

$$E_i = \sum_{\gamma'} \Pi(\gamma'|\gamma) \cdot \frac{\partial V_{j+1}}{\partial a}(a'_i,\, b',\, e',\, \gamma')\,,$$

where b' and e' follow from the pension and wage-index transition equations and do not depend on the consumption–savings choice. The Euler equation in marginal-value form is then

$$c_{j,t}^{-\sigma} = (1+\tau_c) \cdot \beta\big[\alpha_t(j)\, E_i + \big(1-\alpha_t(j)\big)\, \psi_1\, (\psi_2 + a_i')^{-\sigma}\big]\,, \tag{FOC-MV}$$

where $\alpha_t(j)$ is the survival probability and the second term captures the warm-glow bequest motive.

EGM inversion step. Inverting the marginal utility gives EGM consumption $c_i^{\text{egm}} = \text{RHS}_i^{-1/\sigma}$, where RHS_i denotes the right-hand side of (FOC-MV). The *endogenous* current asset level is then recovered from the budget constraint:

$$a_i^{\text{endo}} = \frac{(1+\tau_c)\, c_i^{\text{egm}} + a_i' - y_j}{R_t}. \tag{EGM inversion}$$

The resulting pairs $(a_i^{\text{endo}}, c_i^{\text{egm}})$ are interpolated onto the exogenous asset grid $\{a_k\}$ to obtain the policy function $c^{\text{policy}}(a_k)$.

Borrowing constraint handling. Let $a^{\text{threshold}} = a_0^{\text{endo}}$ be the endogenous asset level corresponding to the smallest savings grid point a'_{min}. For grid points $a_k < a^{\text{threshold}}$, the borrowing constraint $a' \geq a_{\text{min}} = 0$ binds, and consumption is set to

$$c^{\text{constrained}}(a_k) = \frac{R_t\, a_k + y_j - a_{\text{min}}}{1+\tau_c}. \tag{Constrained c}$$

Envelope condition for stored $\partial V/\partial a$. The marginal value passed to the next (younger) age is computed via the envelope theorem:

$$\frac{\partial V_j}{\partial a}(a_k) = \frac{R_t}{1+\tau_c}\, \big(c^{\text{policy}}(a_k)\big)^{-\sigma}. \tag{Stored $\partial V/\partial a$}$$

Since the budget is linear in a and non-asset income y_j does not depend on assets, the envelope condition has a clean closed-form expression, and no automatic differentiation is required.

Interpolation. The urban household state is four-dimensional (a, b, e, γ), requiring trilinear interpolation over (a, b, e) conditional on the discrete shock γ. The rural state is three-dimensional (a, e, γ) with $b = 0$ implicit, so bilinear interpolation over (a, e) suffices.

A.2 Simulation Method and Population Aggregation

The forward side of the solver takes the EGM policy functions from Section A.1 and the exogenous population sequences $\{N_t(j)\}$ as inputs, and produces the cross-sectional distribution of individual states (a, b, e, γ) by age from which the macro aggregates A_t, L_t, C_t, Beq_t, S_t, and B_t are computed. The procedure has three components: (i) a deterministic population projection, (ii) a Monte Carlo household simulation with common random numbers, and (iii) a population-weighted aggregation step that bridges the N_{cohort} simulated agents per age back to macroeconomic aggregates.

Population projection. Population sequences $N_t = (N_t(0), \ldots, N_t(J-1))$ are projected forward from the 2023 base distribution by iterating the fertility and survival rules of equations (1) and (2):

$$N_{t+1}(j) = \begin{cases} \displaystyle\sum_{k \in F} \psi_k \, n_{t+1}(k) \, N_{t+1}(k) & j = 0, \\ \alpha_t(j-1) \, N_t(j-1) & 1 \leq j \leq J-1. \end{cases} \tag{Leslie recursion}$$

Two scenarios are computed with this recursion: a baseline with the 2023 age-specific fertility rates (TFR $\approx$ 1.0), and a counterfactual in which all age-specific rates are rescaled by a common factor so that the total fertility rate rises to replacement level (TFR $=$ 2.0). Projections run for 51 years (2023–2073), after which the age distribution is frozen at the year-50 vector and held constant for the remainder of the model horizon. This freeze ensures that the terminal steady state has a well-defined stationary demographic composition to which the economy can converge.

Common random numbers. The only stochastic element in the household problem is the idiosyncratic productivity shock γ, which follows a discretized three-state Markov chain with transition matrix Π_γ. To eliminate Monte Carlo noise in cross-experiment welfare comparisons, a single set of uniform draws is generated once from a fixed seed and reused across all steady-state and transition simulations. Two tables are pre-drawn: $\varepsilon^{\text{init}} \in [0,1)^{N_{\text{cohort}}}$ for newborn γ initialization, and $\varepsilon^{\gamma} \in [0,1)^{J \times N_{\text{cohort}}}$ for age- and agent-indexed transitions. Given the current productivity $\gamma_{i,j,t}$ and draw $u_{i,j} = \varepsilon^{\gamma}_{j,i}$, the next-period productivity is obtained by inverting the row-CDF of Π_γ:

$$\gamma_{i,j+1,t+1} \;=\; \min\left\{ k \;:\; \sum_{\ell=1}^{k} \Pi_\gamma\big(\gamma_{i,j,t},\, \gamma_\ell\big) \;>\; u_{i,j} \right\}. \tag{CRN}$$

Because the same tables are consumed by every counterfactual, two experiments that produce identical policy functions yield identical simulated trajectories, and welfare differences across experiments are free of Monte Carlo sampling variance.

Cohort simulation for the steady state. For each worker type (urban and rural) the steady-state simulator tracks $N_{\text{cohort}} = 10{,}000$ agents from age 0 to $J-1 = 79$ via a single `jax.lax.scan` over ages. At age 0 each agent is initialized at the borrowing limit $a = a_{\min}$ with empty pension accounts ($b = 0$, $e = 0$) and an initial productivity drawn from $\varepsilon^{\text{init}}$ via the CDF-inversion rule. At each subsequent age j, consumption $c_{i,j}$ is interpolated from the EGM policy $c_j(a, b, e, \gamma)$ (trilinear in (a, b, e) for urban, bilinear in (a, e) for rural), assets are advanced via the budget constraints (5) and (6), the pension account b and the wage index e evolve deterministically via equations (9) and (11), and γ advances via rule (CRN). The scan output is a flat array of $N = J \cdot N_{\text{cohort}} = 800{,}000$ (a, b, e, γ) tuples together with their ages and realized consumption and savings—the steady-state cross-section under the prices and policies passed in.

Single-period evolution for the transition. Along the transition path, the same flat population of N agents is advanced *one period at a time* rather than age by age. Every agent's age increments from j to $j+1$ in parallel; assets and pension-account states are updated using the same budget and transition rules as in the steady state but evaluated at the current period's prices and policies; agents reaching age $j = J$ are replaced in place by newborns with $a = a_{\min}$, $b = e = 0$, and a fresh γ drawn from $\varepsilon^{\text{init}}$. Common-random-number consistency is maintained by indexing the shock tables through $(j_i,\, i \bmod N_{\text{cohort}})$, so each agent's stochastic history is a deterministic function of its identity. The transition scan wraps this single-period update in a second `jax.lax.scan` over the T transition periods, using the policy functions produced by the corresponding backward EGM pass.

Aggregation. Individual decisions are mapped to economy-wide aggregates via population weights. Each simulated agent represents a fraction of the period-t population of its age cohort:

$$w_{i,t} \;=\; \frac{N_t(j_i)}{N_{\text{cohort}}}. \qquad \text{(Population weight)}$$

Within each worker type, the paper's aggregates in equations (17)–(22) are computed as population-weighted sums of the corresponding individual quantities:

$$A_t^s = \sum_i w_{i,t}\, a_{i,t}, \qquad C_t^s = \sum_i w_{i,t}\, c_{i,t}, \qquad L_t^s = \sum_i w_{i,t}\, \ell(j_i)\, \gamma_{i,t}\, \mathbf{1}\{j_i \le j_R\},$$

$$Beq_t^s = \sum_i w_{i,t}\, a'_{i,t}\, \big(1 - \alpha_t(j_i)\big),$$

for $s \in \{u, r\}$. The economy-wide aggregates are obtained by combining the two types with the population shares ϕ_u and $1-\phi_u$: the aggregator passes $w^u_{i,t} = \phi_u\, N_t(j_i)/N_{\text{cohort}}$ to the urban simulator and $w^r_{i,t} = (1-\phi_u)\, N_t(j_i)/N_{\text{cohort}}$ to the rural simulator, and sums the two aggregates. Social security expenditure S_t and contributions B_t (equations (20) and (21)) are computed from the same flat arrays using the pension and contribution formulas for each worker type. The urban effective labor L_t^u returned alongside the main aggregates feeds the social average wage $\bar{W}_t = w_t L_t^u / N_t^{\text{workers},u}$ defined in equation (10), which in turn enters the backward EGM pass through the wage-index evolution, closing the loop between the forward and backward halves of the solver.

A.3 Transition Path Algorithms

The transition path algorithm iterates between a backward EGM sweep (solving household problems at each period given price sequences) and a forward Monte Carlo sweep (simulating 800,000 agents through the transition). The outer loop updates price sequences via damped iteration until convergence. Both sweeps are implemented as single `jax.lax.scan` calls for computational efficiency. We describe the two fiscal closure variants below.

In both algorithms, the damping parameter s follows a dynamic schedule that starts near 1 (conservative updates) and decreases toward 0.3 as the iteration progresses, which improves stability in the early iterations while allowing faster convergence later. The bisection bracket in Algorithm 1 is warm-started from the previous price iteration, narrowing the search range and typically requiring fewer than 10 bisection steps per price iteration. Convergence of the outer price loop is typically achieved within 30–50 iterations.

A.4 Supplementary Results

Aggregate and terminal-SS welfare CEV. Tables A.3 and A.4 report aggregate consumption-equivalent variation (CEV) numbers that compare terminal steady states to the initial condition and to the no-reform baseline, respectively. We report them here for completeness, but caution the reader that both tables are contaminated by the model's omission of any utility value for government consumption G and should be read as upper bounds rather than point estimates. Under G-adjustment in particular, the collapse of G/Y from 21.8%

Algorithm 1 τ^w-Adjustment Transition Path

Require: Initial SS, terminal SS, demographic sequence $\{N_t\}$, target D_T/Y_T
Ensure: Equilibrium price and policy sequences $\{r_t, w_t, \tau_t^w\}$

Initialize $\{r_t, w_t\}$ by linear interpolation between initial and terminal SS
Initialize $\{beq_t, \bar{W}_t\}$ by linear interpolation
Set bisection bracket $[\tau_{\text{lo}}^w, \tau_{\text{hi}}^w]$
for $n = 1, 2, \ldots$ (price iteration) **do**
  for $m = 1, 2, \ldots$ (bisection on τ_{trans}^w) **do**
    $\tau_{\text{trans}}^w \leftarrow (\tau_{\text{lo}}^w + \tau_{\text{hi}}^w)/2$
    Set $\tau_t^w = \tau_{\text{trans}}^w$ for $t < T_{\text{trans}}$; $\quad \tau_t^w = \tau_{\text{term}}^w$ for $t \geq T_{\text{trans}}$
    Backward EGM sweep (`jax.lax.scan`): solve household problems from $t = T-1$ to $t = 0$ using stored terminal-SS $\partial V/\partial a$ as boundary condition
    Forward MC sweep (`jax.lax.scan`): simulate agents from $t = 0$ to $t = T-1$ starting from initial-SS distribution
    Aggregate: $\{A_t, L_t, C_t, Beq_t, S_t, B_t\}$
    Compute debt dynamics: $D_0 = D^{\text{init}}$; for each t, compute tax revenue $T_t = \tau_t^w w_t L_t + \tau_a r_t A_t + \tau_c C_t$, government spending $G_t = (G/Y) \cdot Y_t$, and update $D_{t+1} = (1+r_t)D_t + G_t + S_t - B_t - T_t$
    if $|D_T/Y_T - \text{target}| < \varepsilon_D$ **then**
      break (bisection converged)
    else if $D_T/Y_T > \text{target}$ **then**
      $\tau_{\text{lo}}^w \leftarrow \tau_{\text{trans}}^w$
    else
      $\tau_{\text{hi}}^w \leftarrow \tau_{\text{trans}}^w$
    end if
  end for
  Compute implied prices: $r_t^{\text{new}} = \alpha Z K_t^{\alpha-1} L_t^{1-\alpha} - \delta, \quad w_t^{\text{new}} = (1-\alpha) Z K_t^{\alpha} L_t^{-\alpha}$, where $K_t = A_t - D_t$
  Damped update: $r_t \leftarrow s \cdot r_t + (1-s) \cdot r_t^{\text{new}}, \quad w_t \leftarrow s \cdot w_t + (1-s) \cdot w_t^{\text{new}}$
  if $\frac{1}{T} \sum_t \left[(r_t - r_t^{\text{new}})^2 + (w_t - w_t^{\text{new}})^2\right] < 10^{-5}$ **then**
    return converged solution
  end if
end for

Algorithm 2 G-Adjustment Transition Path

Require: Initial SS, terminal SS, demographic sequence $\{N_t\}$, fixed D/Y ratio
Ensure: Equilibrium price sequence $\{r_t, w_t\}$ and government spending $\{G_t\}$

```
Initialize {r_t, w_t} by linear interpolation between initial and terminal SS
Initialize {beq_t, W̄_t} by linear interpolation
Set τ_t^w = τ_0^w (fixed at initial value) for all t
for n = 1, 2, ... (price iteration) do
    Backward EGM sweep (jax.lax.scan): solve household problems from t = T−1 to t = 0
    Forward MC sweep (jax.lax.scan): simulate agents from t = 0 to t = T−1
    Aggregate: {A_t, L_t, C_t, Beq_t, S_t, B_t}
    For each t: compute K_t = A_t − D_t where D_t = (D/Y)·Y_t, tax revenue T_t = τ_0^w w_t L_t + τ_a r_t A_t + τ_c C_t, and residual government spending G_t = T_t + B_t − (1 + r_t)D_t − S_t + D_{t+1}
    Compute implied prices: r_t^new = αZK_t^(α−1) L_t^(1−α) − δ,   w_t^new = (1 − α)ZK_t^α L_t^(−α)
    Damped update: r_t ← s·r_t + (1 − s)·r_t^new,   w_t ← s·w_t + (1 − s)·w_t^new
    if (1/T) Σ_t [(r_t − r_t^new)^2 + (w_t − w_t^new)^2] < 10^−5 then
        return converged solution
    end if
end for
```

Table A.1: Social Security Deficit Decomposition: Urban vs. Rural

	Initial (2023)	Terminal (τ^w-adj)	Terminal (G-adj)
Pension expenditure S/Y			
Urban	10.67%	22.37%	21.91%
Rural	0.28%	0.60%	0.58%
Total	10.96%	22.97%	22.50%
Pension revenue B/Y			
Urban	9.91%	10.04%	9.83%
Rural	0.26%	0.20%	0.26%
Total	10.18%	10.24%	10.09%
Pension deficit $(B-S)/Y$			
Urban	−0.76%	−12.33%	−12.08%
Rural	−0.02%	−0.40%	−0.32%
Total	−0.78%	−12.73%	−12.40%

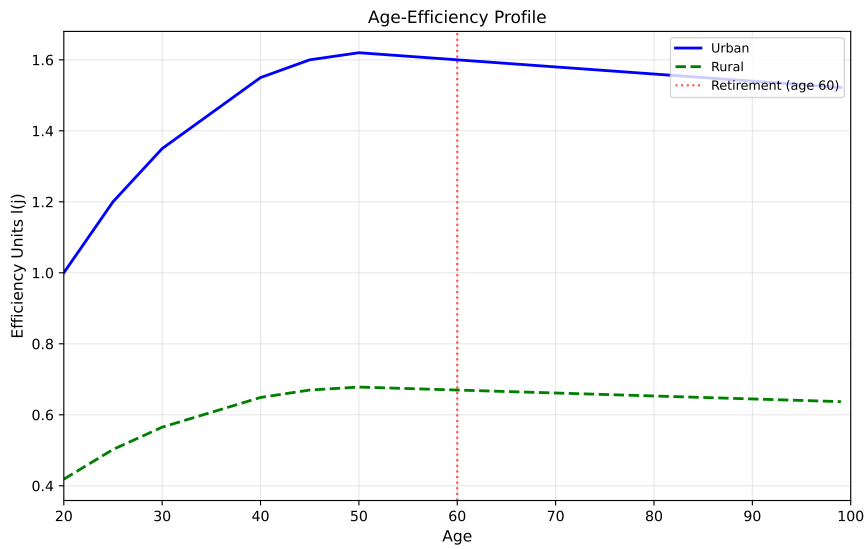


Figure A.1: Age-Efficiency Profile: Urban vs. Rural Workers

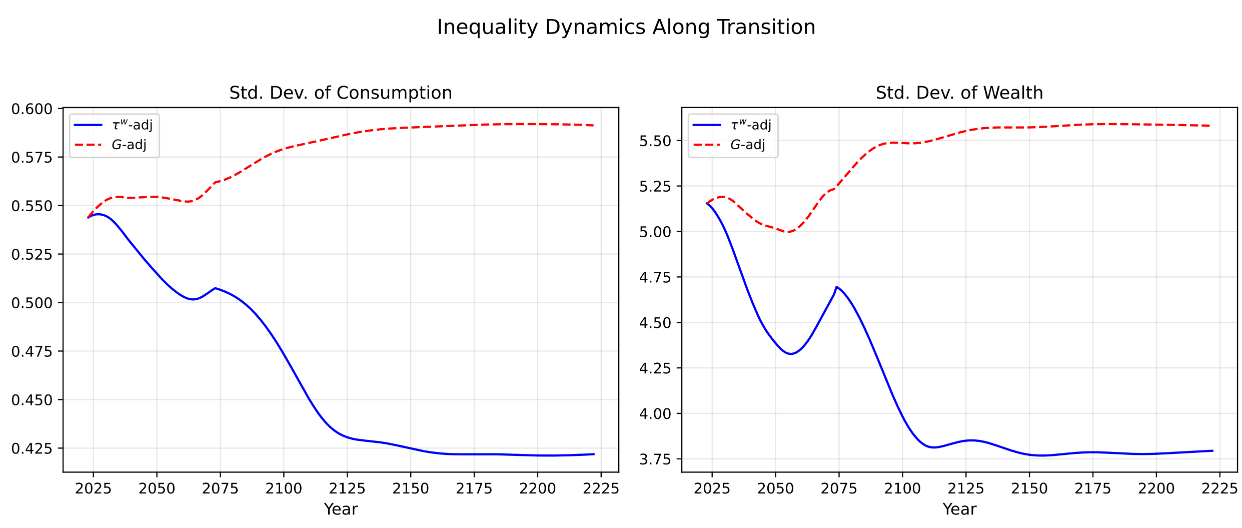


Figure A.2: Standard deviation of consumption (left) and wealth (right) along the transition path under τ^w-adjustment and G-adjustment.

Table A.2: Consumption and Wealth by Age Group: Steady-State Comparison

	Initial (2023)	Terminal (τ^w-adj)	Terminal (G-adj)
Mean consumption			
Young (20–39)	0.498	0.450	0.704
Old (40–59)	0.898	0.763	1.084
Retired (60+)	0.961	0.846	1.041
Std. dev. of consumption			
Young (20–39)	0.242	0.167	0.284
Old (40–59)	0.503	0.347	0.543
Retired (60+)	0.617	0.459	0.627
Gini of consumption			
Young (20–39)	0.253	0.193	0.212
Old (40–59)	0.296	0.241	0.266
Retired (60+)	0.343	0.297	0.326
Mean wealth			
Young (20–39)	2.695	2.717	3.843
Old (40–59)	8.549	7.345	10.426
Retired (60+)	3.675	2.434	3.429
Std. dev. of wealth			
Young (20–39)	2.902	2.343	3.615
Old (40–59)	5.235	3.303	5.352
Retired (60+)	4.617	3.288	4.624
Gini of wealth			
Young (20–39)	0.535	0.461	0.489
Old (40–59)	0.326	0.247	0.278
Retired (60+)	0.631	0.662	0.662

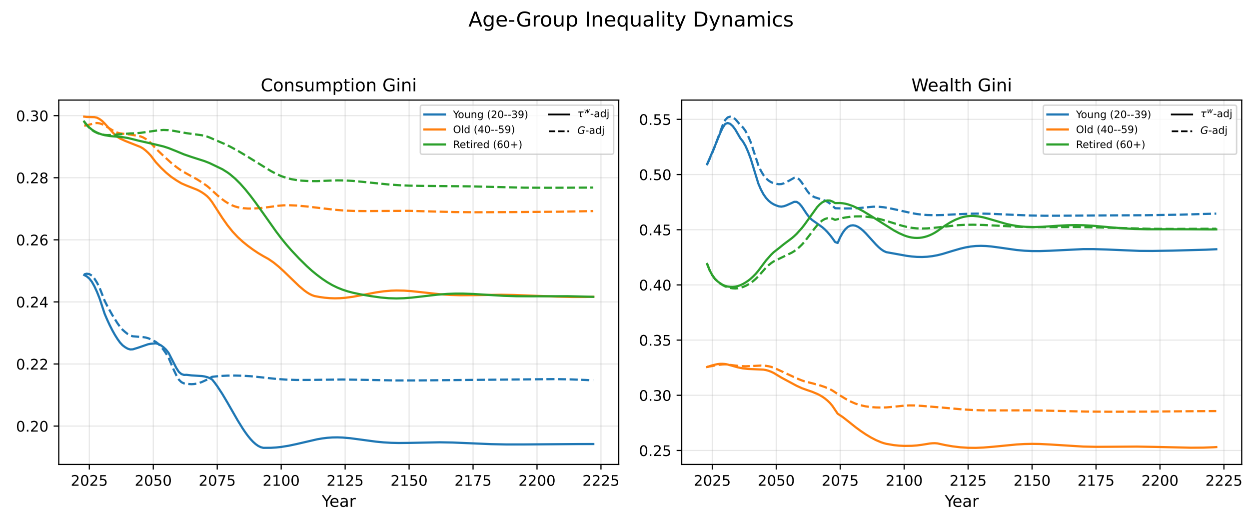


Figure A.3: Gini coefficients of consumption (top) and wealth (bottom) by age group along the transition path. Left: τ^w-adjustment; right: G-adjustment.

to 10.64% appears in the model as a welfare gain because households retain more income, which is clearly not a valid absolute welfare statement; and the negative terminal-SS CEV for replacement-level fertility reflects capital dilution (a larger stationary population lowers the capital–labor ratio and hence wages) rather than any fiscal cost of fertility—the economically relevant fertility evaluation is the fiscal-value-of-birth analysis in Section 5. The *relative* measures used in the main text (Section 4.7)—the urban–rural differential within a fixed fiscal mode, and the cohort-level differential across experiments within a fixed fiscal mode—are robust to this caveat because the (unmeasured) G-contribution to welfare cancels out in the comparison.

Table A.3: Welfare: Consumption Equivalent Variation (CEV)

	τ^w-adjustment	G-adjustment
Overall CEV	-3.65%	+36.84%
Urban	-9.94%	+35.36%
Rural	+5.07%	+38.63%

Note: CEV measures the permanent proportional change in consumption needed in the initial steady state to match lifetime utility in the terminal steady state. Positive values indicate welfare gain in the terminal state.

Table A.4: Welfare Effects of Policy Reforms: CEV Relative to Baseline Terminal SS (%)

	τ^w-adj			G-adj		
Policy Reform	Urban	Rural	All	Urban	Rural	All
Delayed Retirement (age 65)	+9.80	+10.89	+10.26	-2.79	+1.60	-0.86
Reduced Pooling ($\alpha_b/2$)	+21.93	+17.70	+20.12	+3.63	+3.91	+3.76
Replacement TFR (2.0)	+1.63	-8.34	-2.81	-22.52	-23.85	-23.12
Delayed Ret. + Repl. TFR	+1.27	-7.36	-2.53	-25.54	-25.33	-25.44

Note: CEV measures the permanent proportional change in consumption needed in the baseline terminal steady state to match lifetime utility under the policy reform. Positive values indicate welfare gain from the reform. Reference: baseline demographic transition with no policy reform.