University of Kent

School of Economics Discussion Papers

Intergenerational Mobility, Education, and The Skill Premium: Investigating Optimal Government Cash Transfers in China

Wei Jiang, Jingwei Wu

April 2026

KDPE 2602

Intergenerational Mobility, Education, and The Skill Premium: Investigating Optimal Government Cash Transfers in China*

Wei Jiang‡ Jingwei Wu§

April 3, 2026

Abstract

We examine optimal government education expenditure policies in the form of cash transfers to school-aged children, with the objectives of enhancing intergenerational income mobility and reducing the skill premium. Using a calibrated overlapping generations (OLG) model tailored to recent empirical data from the Chinese economy, we analyse the welfare-maximizing policy design of this policy. Our findings indicate that optimal policy involves directing greater cash transfers toward children with lower initial ability, as their educational attainment requires higher relative effort. This aligns with China's nationwide higher education expansion initiative launched in 1999. Quantitatively, the optimal policy enhances upward income mobility by 56%, reduces the skill premium by 58%, and increases aggregate welfare by approximately 13%.

Keywords: Chinese economy, intergenerational income mobility, skill premium, government cash transfers

JEL codes: E24, E62, I25, I28

*We would like to thank, Professor Miguel León-Ledesma, Dr Christian Siegel for their insightful comments. We are also grateful for the helpful feedback received from the participants of Annual PhD Workshop at School of Economics, University of Kent, 30th CEF Conference, Nanyang Technological University. The remaining errors are ours.

†School of Economics, University of Kent, Canterbury, CT2 7PE. E-mail: W.Jiang@kent.ac.uk.

‡Corresponding author: Wei Jiang, E-mail: w.jiang@kent.ac.uk

§School of Economics, Shenzhen Polytechnic University, China. E-mail: jingweiwu@szpu.edu.cn

1 Introduction

Income inequality has been central to the economic and political debate in the past few decades. In countries that have undergone a fast process of development such as China, inequality has increased significantly over time. A large literature studies the reasons behind these trends, see e.g. Benhabib and Bisin (2018), Dabla-Norris, Kochhar, Suphaphiphat, Ricka, and Tsounta (2015), and Nolan, Richiardi, and Valenzuela (2019). Empirical evidence also suggests that intergenerational income mobility is closely linked to income inequality. Countries with higher intergenerational income mobility have more even income distributions (the so-called "Great Gatsby Curve").[1] Societies characterised by high concentration of income and wealth ('inequality of outcomes') tend to be politically acceptable only when social mobility ('equality of opportunity') is high (see Kanbur and Stiglitz (2016)). Not surprisingly, the design of education investment policies to promote intergenerational mobility has been an important focus of policymakers.

We study optimal cash transfer policies in the context of an Overlapping Generations (OLG) model with heterogeneity in skills and abilities calibrated to the recent experience of China. In the model, households' education investment decisions are constrained by a lack of access to credit markets. In this context, cash transfers can alleviate financial constraints and affect education investment decisions. Our study builds on Maoz and Moav (2001) and Murayama (2019), where we incorporate a government cash transfer program based on children's abilities. Specifically, children with varying abilities may receive different amounts of cash transfers from the government. The relative amount they receive will depend on what we call a political parameter which determines the government's bias towards subsidising the low vs. high ability children. The cash transfer influences their education investment decisions and, hence, the share of educated workers. The cash transfer provided by the government is unconditional, meaning that all children, regardless of their parents' type, are eligible to receive the transfer which can be used for either consumption or education investment.[2].

The model with exogenous policy instruments is calibrated so that its steady state reflects the main empirical characteristics of the Chinese economy with particular focus on its population shares of educated and non-educated workers, and the wage premium between the two groups. We then use the calibrated model to analyse welfare maximising policies under commitment.

The main findings can be summarised as follows. The optimal policy suggests that

[1] See e.g. Fisher, Johnson, Latner, Smeeding, and Thompson (2016) and Corak (2020).

[2] In China, the government cash transfer program is one of the policies enacted to improve education attainment. There are, however, many other policies such as the Compulsory Education Policies started in 1986.

the government should bias cash transfers in favor of children with lower ability as their educational cost (effort) is higher. The numerical results show that the policy can improve upward income mobility by about 56% and, therefore, reduce the skill premium by about 58%. Under the optimal policy, there are substantial welfare gains amounting to approximately 13%.

Beyond the aforementioned studies by Maoz and Moav (2001) and Murayama (2019), our paper is related to a literature that studies the effect of education policies on intergenerational mobility. Hanushek, Leung, and Yilmaz (2014) used a three-period overlapping generations model to investigate how the government's different college aid schemes can influence individuals' decisions on receiving education and, therefore, intergenerational mobility. They found that a needs-based approach can help promote intergenerational mobility and achieve a more equal income distribution. Different from their approach, we focus here solely on the design of simple unconditional cash transfers that are easily implementable, especially in the context of middle income economies. Hanushek, Yuan Wang, and Zhang (2023) measured the returns to skills in China over the 2007-2018 period and find that the college wage premium has decreased over time. This suggests the dominant influence of the surge in the supply of college-educated workers, emphasising the importance of higher education attainment in reducing the wage premium in labour markets. A similar mechanism operates in our context, but this will be crucially influenced by the government's policy choices which, in our context, are optimally determined. Recently, Krueger, Ludwig, and Popova (2025) study the distributional and welfare consequences of alternative government education policies to encourage college completion, such as making college free and improving funding for public schooling, using a rich OLG setting calibrated to the US. They find that it is optimal to combine tuition subsidies to make college affordable and better schools to increase human capital. As before, our paper's focus is on unconditional transfer policies to families and their effect on intergenerational mobility in the context of China.

Government cash transfer programs have been evaluated in African, Latin American and Caribbean countries showing overall positive impacts on higher education attainment, see e.g. Patel-Campillo and García (2022), Barrera-Osorio, Linden, and Saavedra (2019). Similar findings have been confirmed in the US by Herrington (2015) and in Norway by Jerrim and Macmillan (2015). These cash transfers also play an important role in the Chinese economy since late 1990s. China witnessed a dramatic surge in the supply of college-educated workers due to China's nationwide higher education expansion program implemented in 1999. Since the expansion of college enrollment, China's higher education has developed considerably. As of end of 2018, over 28 million students were enlisted in 2,663 colleges and universities. As a result of its World Trade Organization access in 2001 and trends in technology, China has also experienced increases in the demand for college-educated workers. This is the so-called "race between education and technology"[3](Goldin and Katz 2008). Government cash transfer policies can reduce the investment gap between poor and rich parents in their children's education. This can then play a crucial role in generating upward intergenerational mobility, see e.g. X. Huang, S. Huang, and Shui (2021). J. Yang and Qiu (2016) used a calibrated model based on Chinese data and found that direct subsidies from the government to poor parents are the most efficient policies for alleviating budget constraints on investments in children's early education. Tang, Sun, and W. Yang (2021) also found that government policies are crucial for increasing the social mobility of poorer households and improving equality of opportunity in China. Our model formalises this framework in the context of optimal policies with commitment allowing for heterogenous children abilities.

The rest of the paper is organised as follows: Section 2 present the model. Section 3 discusses the calibration and steady state of the model given exogenous policy. Section 4 studies the Ramsey problem of the government and the optimal policy in the steady-state. Section 5 concludes.

2 Model

We employ an overlapping generations model with two periods. In each period firms produce a single homogeneous good using educated and non-educated labor inputs, which can be allocated to either consumption or investment in education by both educated and non-educated workers. Individuals inelastically supply one unit of time in a competitive labor market and we assume they pay labour income taxes based on their labour types. The number of educated workers is endogenously determined by education investment decisions. Following Maoz and Moav (2001) and Owen and Weil (1998), and consistent with the Chinese context, we rule out educational loans. In other words, individuals do not have access to credit markets to finance education investments.

2.1 Production and factor prices

Aggregate output Y in period t is given by constant returns to scale production function:

$$Y_t = A_t (n_t^e)^{1-\alpha} (n_t^u)^{\alpha}, \tag{2.1}$$

[3]Human capital investment increases the supply of educated workers. When the relative demand for college-educated labour increases faster than the relative supply, the wage gap between college- and high school-educated labour widens.

where n_t^e is the number of educated workers, n_t^u is the number of uneducated workers, A_t is total factor productivity, α represents the elasticity of output to uneducated workers. Since $n_t^e + n_t^u = 1$, we re-write the production function as : $Y_t = A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}$.

Because the economy is competitive, production factors are paid at their marginal products. We define w_t^e and w_t^u as the wages of an educated and an uneducated worker respectively in period t[4]:

$$\begin{aligned} w_t^e &= A_t(1-\alpha)(n_t^e)^{-\alpha}(n_t^u)^{\alpha}; \\ w_t^u &= A_t\alpha(n_t^e)^{1-\alpha}(n_t^u)^{\alpha-1}. \end{aligned} \tag{2.2}$$

Then, we assume the tax rates of educated and uneducated workers are τ_t^e and τ_t^u respectively, and $\tau_t^e > \tau_t^u$. Therefore, the after tax wages in period $t+1$ are as follow:

$$\widetilde{w}_{t+1}^e = (1-\tau_{t+1}^e)w_{t+1}^e; \tag{2.3a}$$

$$\widetilde{w}_{t+1}^u = (1-\tau_{t+1}^u)w_{t+1}^u. \tag{2.3b}$$

2.2 Individuals

The economy is made up of two periods overlapping generations of individuals, each with a single parent and a single child. In the first period, the individual does not work and decide whether to acquire education based on the transfers they receive from their parents and the government. The transfer from their parents is a bequest, while the transfer from the government is a cash transfer to children according to their abilities funded from the revenues collected through taxes. In the second period, individuals work and allocate their wealth between consumption and a transfer to their child. An individual's labour type is by default uneducated at the beginning, if they receive education in the first period, then their type becomes educated.

The decision to pursue education or not may vary from person to person because of the individual's abilities and transfers received. The budget constraints of individual i are:

$$c_t^i + \delta^i h_t^i = x_t^i + s_t^i, \tag{2.4a}$$

$$c_{t+1}^i + x_{t+1}^i = \widetilde{w}_{t+1}^i, \tag{2.4b}$$

[4] $w_t^u < w_t^e$ iff $E_t < 1-\alpha$.

where

$$\delta^i = \begin{cases} 1, & \text{individual } i \text{ acquires education} \\ 0, & \text{otherwise} \end{cases},$$

hence,

$$\widetilde{w}^i_{t+1} = \begin{cases} \widetilde{w}^e_{t+1}, & \delta^i = 1 \\ \widetilde{w}^u_{t+1}, & \delta^i = 0. \end{cases}$$

x^i_t is the transfer that individual i of generation t (i.e. born in period t) receives from their parent, s^i_t is the cash transfer from the government, and x^i_{t+1} is the transfer the agent makes to their children in the second period of their life ($t+1$).

In this model, differences in abilities are expressed as differences in individual education costs. h^i_t denotes the education cost of individual i who was born in period t and we further assume the higher is h^i_t, the lower is i's ability[5], and the government is able to observe the individual's ability through their non-tertiary education outcomes.[6] It is worth emphasizing that the cash transfer here is unconditional, that is, it is not conditional on the children's predetermined investment in education. Regardless of the parents' income level, all children can get the cash transfer and use it for consumption and education (Murayama 2019). The cash transfer s^i_t is assumed to have the following form:

$$s^i_t = (1-\eta_t)(h^i_t - h^{0.5}_t) + \tilde{s}_t, \tag{2.5}$$

where $\tilde{s}_t$ is the budget for cash transfers defined as $\tilde{s}_t = \tau^e_t w^e_t n^e_t + \tau^u_t w^u_t n^u_t - g_t$, $h^{0.5}_t$ is the mean of h^i_t, and η_t is a political parameter. In the spirit of Sano and Tomoda (2010), if $\eta = 1$, then $s^i_t = \tilde{s}_t$, meaning all children receive the same cash transfer regardless of their abilities. If $\eta < 1$, individuals with lower (high) ability receives larger (smaller) transfers than $\tilde{s}_t$, while if $\eta > 1$ individuals with higher (lower) ability receive larger (smaller) transfers than $\tilde{s}_t$. That is, parameter η controls the degree to which government cash transfers are biased in favor of lower or higher ability children.

Individuals gain utility from consumption in both periods and from transfer to their children. Thus, we have the utility function:

$$u^i_t = \log c^i_t + \log c^i_{t+1} + \log x^i_{t+1}, \tag{2.6}$$

where c^i_t and c^i_{t+1} are the consumption of an individual born in period t in the two

[5]Higher-ability individuals might enter the education system with a stronger foundational knowledge or learning efficiency, enabling them to complete education faster or at lower costs (e.g., by earning credits in high school or placing out of introductory classes). This leads to a reduced cost of obtaining higher education compared to individuals who need more time or additional support (e.g., private tutoring) to reach the same level.

[6]China's college entrance examination scores are a good proxy for abilities of individuals.

periods of his life and x^i_{t+1} is the transfer to his offspring.

Since we assume the capital market is imperfect and the utility function is separable, the individual optimization problem boils down to deciding whether to purchase education after deciding how to allocate income between consumption and bequests in the second period. Therefore, the allocation of second period income involves solving:

$$z(\widetilde{w}^i_{t+1}) \equiv \max(\log c^i_{t+1} + \log x^i_{t+1})$$
$$s.t.(2.4b).$$

The maximization gives the optimal allocations:

$$\mathcal{L} = \log c^i_{t+1} + \log x^i_{t+1} - \lambda^i_{t+1}[c^i_{t+1} + x^i_{t+1} - \widetilde{w}^i_{t+1}]$$
$$\frac{\partial \mathcal{L}}{\partial c^i_{t+1}} = \frac{1}{c^i_{t+1}} - \lambda^i_{t+1} = 0 \Rightarrow \lambda^i_{t+1} = \frac{1}{c^i_{t+1}} \tag{2.7a}$$
$$\frac{\partial \mathcal{L}}{\partial x^i_{t+1}} = \frac{1}{x^i_{t+1}} - \lambda^i_{t+1} = 0 \Rightarrow \lambda^i_{t+1} = \frac{1}{x^i_{t+1}} \tag{2.7b}$$
$$\frac{\partial \mathcal{L}}{\partial \lambda^i_{t+1}} = c^i_{t+1} + x^i_{t+1} - \widetilde{w}^i_{t+1} = 0 \Rightarrow c^i_{t+1} + x^i_{t+1} = \widetilde{w}^i_{t+1} \tag{2.7c}$$

From (2.7a) and (2.7b), we can get

$$\frac{1}{c^i_{t+1}} = \frac{1}{x^i_{t+1}} \tag{2.8}$$

Together with (2.7c), we can get

$$x^i_{t+1} = \widetilde{w}^i_{t+1}\big/2. \tag{2.9}$$

From (2.4b) we can get $c^i_{t+1} = \widetilde{w}^i_{t+1} - x^i_{t+1}$, then substituting it into $z(\widetilde{w}^i_{t+1})$, we get

$$z(\widetilde{w}^i_{t+1}) \equiv max[\log(\widetilde{w}^i_{t+1} - x^i_{t+1}) + \log x^i_{t+1}] \tag{2.10}$$

Then substitute (2.9) into (2.10), we get

$$\begin{aligned} z(\widetilde{w}^i_{t+1}) &= \log\left(\widetilde{w}^i_{t+1} - \frac{\widetilde{w}^i_{t+1}}{2}\right) + \log\frac{\widetilde{w}^i_{t+1}}{2} \\ &= \log\frac{\widetilde{w}^i_{t+1}}{2} + \log\frac{\widetilde{w}^i_{t+1}}{2} \\ &= \log\widetilde{w}^i_{t+1} - \log 2 + \log\widetilde{w}^i_{t+1} - \log 2 \\ &= 2\log\widetilde{w}^i_{t+1} - 2\log 2. \end{aligned} \tag{2.11}$$

Hence, individual will choose to invest in human capital if and only if:

$$\log(x_t^i + s_t^i - h_t^i) + z(\widetilde{w}_{t+1}^e) \geq \log(x_t^i + s_t^i) + z(\widetilde{w}_{t+1}^u). \tag{2.12}$$

From (2.12) it follows that, given the transfer from parents, x_t^i, and cash transfers received from the government, s_t^i, they will invest in education if the cost, h_t^i, is small enough.

Substituting (2.11) into (2.12), we get

$$\log(x_t^i + s_t^i - h_t^i) + 2\log\widetilde{w}_{t+1}^e - 2\log 2 \geq \log(x_t^i + s_t^i) + 2\log\widetilde{w}_{t+1}^u - 2\log 2$$

$$\log(x_t^i + s_t^i - h_t^i) - \log(x_t^i + s_t^i) \geq 2\log\widetilde{w}_{t+1}^u - 2\log\widetilde{w}_{t+1}^e$$

$$\log\left(\frac{x_t^i + s_t^i - h_t^i}{x_t^i + s_t^i}\right) \geq \log\left(\widetilde{w}_{t+1}^u\right)^2 - \log\left(\widetilde{w}_{t+1}^e\right)^2$$

$$\log\left(1 - \frac{h_t^i}{x_t^i + s_t^i}\right) \geq \log\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2$$

$$1 - \frac{h_t^i}{x_t^i + s_t^i} \geq \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2$$

$$h_t^i \leq (x_t^i + s_t^i)\left[1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right] \tag{2.13}$$

Substituting (2.5) into (2.13), we can get

$$h_t^i \leq \left[x_t^i + (1-\eta_t)(h_t^i - h_t^{0.5}) + \tilde{s}_t\right]\left[1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right], \tag{2.14}$$

Rearranging it we get

$$h_t^i \leq \frac{1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1 - (1-\eta_t)\left[1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right]}\left[x_t^i - (1-\eta_t)h_t^{0.5} + \tilde{s}_t\right] \tag{2.15}$$

Let $\hat{h}_t^i$ be the critical value of the education cost for individual i so that they will invest in education iff: $h_t^i \leq \hat{h}_t^i$. $\hat{h}_t^i$ is calculated as the largest value of h_t^i where (2.12) and (2.4) hold, and it can be expressed as a function of future wages and transfers received by individual i:

$$\hat{h}_t^i = \frac{1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1 - (1-\eta_t)\left[1 - \left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right]}\left[x_t^i - (1-\eta_t)h_t^{0.5} + \tilde{s}_t\right] \tag{2.16}$$

It follows that $\frac{\partial \hat{h}_t^i}{\partial\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)} < 0$, $\frac{\partial \hat{h}_t^i}{\partial x_t^i} > 0$, and $\frac{\partial \hat{h}_t^i}{\partial \tilde{s}_t} > 0$. The negative sign of the derivative of $\hat{h}_t^i$ with respect to $\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}$ shows that higher wage inequality affects education cost threshold negatively. The positive signs of the derivative of $\hat{h}_t^i$ in terms of parent transfers and government cash transfers suggest that the more transfers received, the more likely an individual is to invest in education.

From equation (2.4b) we know that workers belonging to the same group (educated or uneducated) have the same second-period income. So, their consumption and transfer to their children, and also their children's critical values are the same. Therefore, the individual index i in the transfer notation can be replaced by an index describing the parent type: e or u. Therefore, from (2.16):

$$\hat{h}_t^e = \frac{1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1-(1-\eta_t)\left[1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right]}\left[\frac{\widetilde{w}_t^e}{2}-(1-\eta_t)h_t^{0.5}+\tilde{s}_t\right]; \tag{2.17a}$$

$$\hat{h}_t^u = \frac{1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1-(1-\eta_t)\left[1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\right]}\left[\frac{\widetilde{w}_t^u}{2}-(1-\eta_t)h_t^{0.5}+\tilde{s}_t\right], \tag{2.17b}$$

where

$$\tilde{s}_t = \tau_t^e w_t^e n_t^e + \tau_t^u w_t^u n_t^u - g_t \tag{2.18}$$

Following Maoz and Moav (2001), we further assume education cost of individual i at period t is:

$$h_t^i = \theta_t^i(a + b\overline{w}_t), \tag{2.19}$$

where $a \geq 0$, $b \in [0,1]$, $\overline{w}_t$ denotes the average wage in period t, θ_t^i is the individual i's education cost parameter and the higher is i's ability, the lower is θ_t^i. The average wage in period t is a weighted average of educated and uneducated wages which can be expressed as: $\overline{w}_t = n_t^e \widetilde{w}_t^e + n_t^u \widetilde{w}_t^u$. With regards to the education cost parameter, θ_t^i, is independent of the parent's labour type and is uniformly distributed over the interval $(\underline{\theta}, \overline{\theta})$ with $\underline{\theta} \geq 0$. Hence, h_t^i is uniformly distributed over the interval $(\underline{h}_t, \overline{h}_t)$, where $\underline{h}_t = \underline{\theta}(a + b\overline{w}_t)$, $\overline{h}_t = \overline{\theta}(a + b\overline{w}_t)$.

2.3 Government

The government finances its stream of purchases $\{g_t\}_{t=0}^{\infty}$ by levying a time-varying flat-rate on earnings from educated workers τ_t^e and uneducated workers τ_t^u. The government also conducts the cash transfer program for the educated s_t^e and uneducated workers s_t^u. The government's budget constraint is

$$g_t = \tau_t^e w_t^e n_t^e + \tau_t^u w_t^u n_t^u - s_t^e n_t^e - s_t^u n_t^u, \tag{2.20}$$

where $s_t^e = (1-\eta_t)(\hat{h}_t^e - h_t^{0.5}) + \tilde{s}_t$ and $s_t^u = (1-\eta_t)(\hat{h}_t^u - h_t^{0.5}) + \tilde{s}_t$.

2.4 Resource constraint

The aggregate resource constraint implies that:

$$Y_t = c_t^e n_t^e + c_t^u n_t^u + g_t. \tag{2.21}$$

3 Calibration and steady state for the exogenous policy

The decentralized competitive equilibrium (DCE) consists of the budget constraint for the educated and uneducated workers, i.e. BC_{t+1}^e, BC_{t+1}^u; the first order conditions, i.e. FOC_{t+1}^e, FOC_{t+1}^u; government budget constraint, i.e. GBC_t, and the aggregate resource constraint, i.e. ARC_t.[7]

Table 1 reports the values of the model parameters with Chinese data, indicating the source used for calibration.

Parameter	Value	Definition	Source
$A > 0$	1.20	total factor productivity	calibration
$0 < \alpha < 1$	0.39	elasticity of uneducated workers	estimate
$a \geq 0$	0.05	intercept of education cost function	calibration
$0 \leq b \leq 1$	0.20	coefficient of education cost function	estimate
$\overline{\theta} > 0$	0.02	the highest bound of education cost parameter	calibration
$\underline{\theta} > 0$	0.01	the lowest bound of education cost parameter	calibration
$\eta > 0$	1.00	political parameter	calibration
$0 < \tau^e < 1$	0.1028	tax rate of educated workers	data
$0 < \tau^u < 1$	0.0625	tax rate of uneducated workers	data

Table 1: Model calibration

A is calibrated to match the number of educated workers in the data. To parameterize α we use the results in Fleisher, Hu, H. Li, and Kim (2011), who investigate the role of education on worker productivity and a firm's total factor productivity using a panel of firm-level data from China for the period of 1998 - 2000.[8].

[7]The full DCE conditions are provided in Appendix A.

[8]Fleisher, Hu, H. Li, and Kim (2011) specify the value-added production function: $Y_{it} =$

The average wage coefficient in the education cost function, b, is estimated using the China Institute for Educational Finance Research-Household Survey (CHIEFR-HS) database in 2017. We extract two survey variables from it: one is the total education expenses of one of the interviewed household's children in elementary, middle, and high schools, the other is the total household income. After filtering the data, we run a linear regression between education expenditure and total income (log(education expenditure) $\sim$ log(total income)) with a sample size of 111.[9] From this procedure, we obtain the estimated coefficient, $b = 0.20$. We use China Household Finance Survey (CHFS) database to calculate the tax rates of educated and uneducated workers. First, we choose three variables from the survey: educational level, individual tax amount, and after-tax wages. According to the measurement of educational level, we define people who have bachelor's, master's, or doctorate degree as educated workers, while people who hold college/vocational degree and below are uneducated workers. Secondly, the tax rates are calculated through using following equation: individual tax/(individual tax + after-tax money wages) controlling the education and considering the weight. Since CHFS has five survey waves which are 2011, 2013, 2015, 2017, and 2019, finally we take the average of these five tax rates for the educated and uneducated workers respectively, and I get $\tau^e = 0.1025$, $\tau^u = 0.0628$.

Variable	Data average	Model
C^e/Y	0.2587	0.2737
C^u/Y	0.2377	0.1828
C/Y	0.4964	0.4564
G/Y	0.5036	0.5435
N^e	0.3835	0.3850
$(w^e * N^e)/Y$	0.6297	0.6100
w^e/w^u	2.4748	2.4985

Table 2: Data Averages and Model's Steady State Values

We now compare model outcomes with averages from the data for a set of relevant data targets. By making use of the data from the China Household Income Project (CHIP) in 2013 and 2018[10], we calculate the consumption rate of educated and une-

$A_i K_{it}^{\beta_k} L_{sit}^{\beta_s} L_{pit}^{\beta_p} e^{u_{it}}$, where Y is output measured by value-added, K is capital, L_s is the number of highly educated workers, L_p is the number of workers with less education, and u is a disturbance term for firms $i = 1, 2, \ldots, n$ and from year $t = 1, 2, \ldots, T$. The parameters β_k, β_s, and β_p are the output elasticizes of the corresponding inputs. Given the estimation of the production function, they get $\beta_s = 0.538$, $\beta_p = 0.344$. Then, we re-calculate the elasticity of uneducated workers: $\alpha = \frac{0.344}{0.538+0.344} = 0.390$.

[9] We match the education expenditure and total income through the household ID to get the sample size.

[10] The data quality of previous survey waves is not considered good enough.

ducated people[11]. Since the consumption variable in this project is household level, we filter the samples with only one household member which is the household head, so the consumption of this household just belongs to this household head[12]. After controlling the age between 16 and 64, education level, and matching the GDP per capita of different provinces, then we can get the average consumption over GDP of educated and uneducated individuals.[13] Because our model does not have separate investment and net export terms, they are added to government spending. Therefore, we can get $G/Y = 1 - C^e/Y - C^u/Y = 1 - 0.2587 - 0.2377 = 0.5036$ from the above data average. Following W. Li, Yanwu Wang, and Chen (2019), after they revise and correct the data on the Chinese consumption rate in terms of both residents' own housing consumption and grey consumption, the average consumption rate between 2010 and 2012 is 0.4303, which aligns closely with the our result: $C/Y = C^e/Y + C^u/Y = 0.2736 + 0.1828 = 0.4564$. With regards to N^e, we choose the gross enrollment rate of higher education[14] from 2010 to 2019, then the average is 0.3835. w^e is calculated from five waves of CHFS, we get $w^e = 86027.222$, then the output Y is the average GDP of those years [15], $Y = 71,853,252,000,000$, therefore we can get the educated labor income share $(w^e * N^e)/Y = 0.6297$. The skill premium $w^e/w^u = 2.4748$ is calculated from the CHIP in 2018[16].

Table 2 compares the data with the calibrated results and we can see the model's steady solution matches most of the data averages well.

4 Optimal policy with commitment

4.1 Ramsey problem

In the commitment framework, the government considers that both educated and uneducated individuals will act in their own best interest, taking all variables as given. Each applicable cash transfer program leads to a feasible equilibrium allocation that fully reflects the optimal behavioral response to resource distribution. Given a welfare criterion, the government's optimization problem is to select the best cash transfer

[11]Since the Chinese Bureau of Statistics does not have data on consumption by education level, and the existing literature does not have such data, we can only use the survey data to calculate it.

[12]We exclude the housing expenditure from the consumption expenses.

[13]The consumption over GDP of educated and uneducated people from CHIP is 0.6747 and 0.3855 respectively, then we multiply the share of educated and uneducated individuals, 0.3835 and 0.6165, finally we get $C^e/Y = 0.2587$, $C^u/Y = 0.2377$

[14]According to the "China Education Monitoring and Evaluation Statistical Index System", the formula for calculating the gross enrollment rate of higher education is: gross enrollment rate of higher education (%) = The total number of students in higher education / the population in the age group of 18 to 22 * 100%.

[15]China Statistical Yearbook 2020, average GDP of 2011, 2013, 2015, 2017, 2019, keep consistent with the waves of CHFS.

[16]Average the rural and urban data.

program that can produce an equilibrium allocation, thereby yielding the highest aggregate welfare. To avoid the general issue of time inconsistency in policy-making, it is assumed that the government commits to a once-and-for-all cash transfer program, which is announced during the initial period and never re-optimized. This problem is commonly referred to as the Ramsey problem of government under commitment.

The government now optimally chooses the policy instrument while simultaneously determining the allocation of individuals. This method is known as the dual approach to the Ramsey problem. The government's objective is to maximize the present discounted value of a weighted average of the welfare of both educated and uneducated workers:

$$\sum_{t=0}^{\infty} \beta^t [n^e_{t+1}(\log c^e_{t+1} + \log x^e_{t+1}) + (1 - n^e_{t+1})(\log c^u_{t+1} + \log x^u_{t+1})] \tag{4.1}$$

The optimal policy approach emphasises the constraints under which the government must operate. These are summarised in the DCE conditions. In order to simplify the optimization problem of the government - it is necessary to reduce the number of choice variables for the government, we substitute out, s^e_t, s^u_t, by making use of the expressions of $\hat{h}^e_t$, $\hat{h}^u_t$, $h^{0.5}_t$, $\tilde{s}_t$[17]. To summarize, the choice variables for the government are six allocation variables, $\{c^e_{t+1}, x^e_{t+1}, c^u_{t+1}, x^u_{t+1}, n^e_t, g_t\}_{t=0}^{\infty}$ and the policy parameter $\{\eta_t\}_{t=0}^{\infty}$. The optimization problem can thus be summarized as follows:

$$\max_{\{c^e_{t+1}, x^e_{t+1}, c^u_{t+1}, x^u_{t+1}, n^e_t, g_t, \eta_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t [n^e_{t+1}(\log c^e_{t+1} + \log x^e_{t+1}) + (1 - n^e_{t+1})(\log c^u_{t+1} + \log x^u_{t+1})] \tag{4.2}$$

[17]The simplified government budget constraint is provided in the Appendix B.

subject to the DCE conditions of

$$c^e_{t+1} + x^e_{t+1} = (1-\tau^e_{t+1})(1-\alpha)A_{t+1}(n^e_{t+1})^{-\alpha}(1-n^e_{t+1})^{\alpha} \tag{4.3}$$

$$c^u_{t+1} + x^u_{t+1} = (1-\tau^u_{t+1})\alpha A_{t+1}(n^e_{t+1})^{1-\alpha}(1-n^e_{t+1})^{\alpha-1} \tag{4.4}$$

$$\frac{1}{c^e_{t+1}} = \frac{1}{x^e_{t+1}} \tag{4.5}$$

$$\frac{1}{c^u_{t+1}} = \frac{1}{x^u_{t+1}} \tag{4.6}$$

$$\begin{aligned}
g_t &= A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}T \\
&- n^e_t\left\{(1-\eta_t)\left(\frac{1-\Phi\left(\frac{n^e_{t+1}}{1-n^e_{t+1}}\right)^2}{(1-\eta_t)\Phi\left(\frac{n^e_{t+1}}{1-n^e_{t+1}}\right)^2+\eta_t}\left\{\frac{1}{2}(1-\tau^e_t)A_t(1-\alpha)(n^e_t)^{-\alpha}(1-n^e_t)^{\alpha} - (1-\eta_t)\Theta a\right.\right.\right. \\
&\left.\left. - [\Theta b(1-\eta_t)(1-T)-T]A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - g_t\right\} - \Theta\left[a + b(1-T)A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}\right]\right) \\
&\left. + A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}T - g_t\right\} \\
&- (1-n^e_t)\left\{(1-\eta_t)\left(\frac{1-\Phi\left(\frac{n^e_{t+1}}{1-n^e_{t+1}}\right)^2}{(1-\eta_t)\Phi\left(\frac{n^e_{t+1}}{1-n^e_{t+1}}\right)^2+\eta_t}\left\{\frac{1}{2}(1-\tau^u_t)A_t\alpha(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha-1} - (1-\eta_t)\Theta a\right.\right.\right. \\
&\left.\left. - [\Theta b(1-\eta_t)(1-T)-T]A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - g_t\right\} - \Theta\left[a + b(1-T)A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}\right]\right) \\
&\left. + A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}T - g_t\right\}
\end{aligned} \tag{4.7}$$

$$g_t = A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - c^e_t n^e_t - c^u_t(1-n^e_t) \tag{4.8}$$

where $\Theta = 0.5(\underline{\theta}_t + \overline{\theta}_t)$, $T = \tau^e_t - \alpha(\tau^e_t - \tau^u_t)$, $\Phi = \left[\frac{\alpha(1-\tau^u_t)}{(1-\alpha)(1-\tau^e_t)}\right]^2$.[18]

4.2 Optimal policy

Table 3 presents the steady-state of the optimal policy along with the exogenous policy. Under the exogenous policy, the government makes the same cash transfers to all the children regardless of their ability level. However, under the optimal policy, when the government aims to maximise aggregate welfare, it chooses to bias towards the children with low abilities by offering more cash transfers to them. As a result, the educated population share has increased to 60.18% from 38.50%. This suggests that the optimal policy can generate upward mobility by about 56%. It also has a significant influence on the wage premium between those two groups. We can see

[18]The Lagrangian function of government in the Ramsey problem is provided in the Appendix C.

that, by increasing the supply of educated workers, the wage premium falls sharply by about 58%. That is, the government's optimal policy reduces wage inequality. Both uneducated and aggregate welfare improve under this policy. The welfare of the educated population worsens due to the reduced wage income as a result of the increase in educated labour supply. The optimal policy resembles the 1999 nationwide higher education expansion policy in China which enhanced the supply of educated labour.

Variable	Exogenous Policy	Optimal Policy
C^e	0.3943	0.2795
C^u	0.1648	0.2822
G	0.3014	0.3341
η	1.0000	0.9220
N^e	0.3850	0.6018
$(w^e * N^e)/Y$	0.6100	0.6100
w^e/w^u	2.4985	1.0351
U^e	-1.8612	-2.5492
U^u	-3.6060	-2.5303
U	-2.9343	-2.5417

Table 3: Model's Steady State Values and Optimal Policy

5 Conclusions

We build an overlapping generations model to examine the optimal government expenditure in the form of unconditional cash transfers to children and its impact on their educational decision-making. The model incorporates heterogeneous labor inputs by skill level and heterogeneous child abilities in the form of educational costs. To analyse the quantitative results, the model is calibrated to the Chinese economy. We show that the model does a good job at matching key steady state moments of the Chinese economy.

Our findings indicate that, in order to maximise aggregate welfare, it is optimal for the government to offer enhanced cash transfers to children with lower abilities. This policy encourages more of these children to pursue higher education and, subsequently, become educated workers when they enter the labour force. Quantitative results show that this policy can achieve approximately 56% upward mobility within the population. As a result, the a larger share of the population becomes educated, significantly reducing the wage premium and income gap between educated and uneducated individuals. Overall, this policy enhances aggregate welfare by about 13%.

References

Barrera-Osorio, Felipe, Leigh L. Linden, and Juan E. Saavedra (July 2019). "Medium- and Long-Term Educational Consequences of Alternative Conditional Cash Transfer Designs: Experimental Evidence from Colombia". In: *American Economic Journal: Applied Economics* 11.3, pp. 54–91.

Benhabib, Jess and Alberto Bisin (2018). "Skewed wealth distributions: Theory and empirics". In: *Journal of Economic Literature* 56.4, pp. 1261–1291.

Corak, Miles (June 2020). "Intergenerational Mobility: What Do We Care About? What Should We Care About?" In: *Australian Economic Review* 53.2, pp. 230–240.

Dabla-Norris, Ms Era, Ms Kalpana Kochhar, Mrs Nujin Suphaphiphat, Mr Frantisek Ricka, and Ms Evridiki Tsounta (2015). *Causes and consequences of income inequality: A global perspective*. International Monetary Fund.

Fisher, Jonathan, David Johnson, Jonathan P. Latner, Timothy Smeeding, and Jeffrey Thompson (2016). "Inequality and Mobility Using Income, Consumption, and Wealth for the Same Individuals". In: *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2.6, pp. 44–58.

Fleisher, Belton M, Yifan Hu, Haizheng Li, and Seonghoon Kim (2011). "Economic transition, higher education and worker productivity in China". In: *Journal of Development Economics* 94.1, pp. 86–94.

Goldin, Claudia Dale and Lawrence F Katz (2008). *The race between education and technology*. Harvard University Press.

Hanushek, Eric A., Charles Ka Yui Leung, and Kuzey Yilmaz (2014). "Borrowing Constraints, College Aid, and Intergenerational Mobility". In: *Journal of Human Capital* 8.1, pp. 1–41.

Hanushek, Eric A., Yuan Wang, and Lei Zhang (June 2023). "Understanding Trends in Chinese Skill Premiums, 2007-2018". In: *National Bureau of Economic Research*.

Herrington, Christopher M. (2015). "Education Financing, Earnings Inequality, Intergerenational Mobility". In: *Review of Economic Dynamics* 18.4, pp. 822–842.

Huang, Xiao, Shoujun Huang, and Ailun Shui (2021). "Government spending and intergenerational income mobility: Evidence from China". In: *Journal of Economic Behavior & Organization* 191.C, pp. 387–414.

Jerrim, John and Lindsey Macmillan (2015). "Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve: Is Education the Key?" In: *Social Forces* 94, pp. 505–533.

Kanbur, Ravi and Joseph Stiglitz (2016). "Dynastic inequality, mobility and equality of opportunity". In: *The Journal of Economic Inequality* 14.4, pp. 419–434.

Krueger, Dirk, Alexander Ludwig, and Irina Popova (2025). "Shaping inequality and intergenerational persistence of poverty: Free college or better schools?" In: *Journal* *of Monetary Economics* 150. Labor Market Dynamics: Drivers and Consequences, p. 103694.

Li, Wenpu, Yanwu Wang, and Tingting Chen (2019). "Whether dweller consumption can be the main driving force that pulls our economy growth at present stage (in Chinese)". In: *Review of Economic Research* 1, pp. 5–19.

Maoz, Y. D. and Omer Moav (Dec. 2001). "Intergenerational Mobility and the Process of Development". In: *The Economic Journal* 109.458, pp. 677–697. eprint: https://academic.oup.com/ej/article-pdf/109/458/677/26639669/ej0677.pdf.

Murayama, Yu (2019). "Cash transfers, intergenerational mobility, and the process of development". In: *Bulletin of Economic Research* 71.3, pp. 209–218.

Nolan, Brian, Matteo G. Richiardi, and Luis Valenzuela (2019). "The Drivers of Income Inequality in Rich Countries". In: *Journal of Economic Surveys* 33.4, pp. 1285–1324. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12328.

Owen, Ann L. and David N. Weil (1998). "Intergenerational earnings mobility, inequality and growth". In: *Journal of Monetary Economics* 41.1, pp. 71–104.

Patel-Campillo, Anouk and V.B. Salas García (2022). "Breaking the poverty cycle? Conditional cash transfers and higher education attainment". In: *International Journal of Educational Development* 92.C.

Sano, Koichiro and Yasunobu Tomoda (2010). "Optimal public education policy in a two sector model". In: *Economic Modelling* 27.5, pp. 991–995.

Tang, Le, Shiyu Sun, and Weiguo Yang (2021). "Does government education expenditure boost intergenerational mobility? Evidence from China". In: *International Review of Economics & Finance* 74, pp. 13–22.

Yang, Juan and Muyuan Qiu (2016). "The impact of education on income inequality and intergenerational mobility". In: *China Economic Review* 37.C, pp. 110–125.

A The DCE conditions

The DCE conditions:

$$\begin{aligned}
BC^e_{t+1} &: c^e_{t+1} + x^e_{t+1} = \tilde{w}^e_{t+1} \\
BC^u_{t+1} &: c^u_{t+1} + x^u_{t+1} = \tilde{w}^u_{t+1} \\
FOC^e_{t+1} &: \frac{1}{c^e_{t+1}} = \frac{1}{x^e_{t+1}} \\
FOC^u_{t+1} &: \frac{1}{c^u_{t+1}} = \frac{1}{x^u_{t+1}} \\
GBC_t &: g_t = \tau^e_t w^e_t n^e_t + \tau^u_t w^u_t (1-n^e_t) - s^e_t n^e_t - s^u_t (1-n^e_t) \\
ARC_t &: g_t = Y_t - c^e_t n^e_t - c^u_t (1-n^e_t)
\end{aligned}$$

B The government budget constraint

From (2.17) we know the expression of $\hat{h}^e_t$ and $\hat{h}^u_t$:

$$\begin{aligned}
\hat{h}^e_t &= \frac{1-\left(\frac{\widetilde{w}^u_{t+1}}{\widetilde{w}^e_{t+1}}\right)^2}{1-(1-\eta_t)\left[1-\left(\frac{\widetilde{w}^u_{t+1}}{\widetilde{w}^e_{t+1}}\right)^2\right]}\left[\frac{\widetilde{w}^e_t}{2} - (1-\eta_t)h^{0.5}_t + \tilde{s}_t\right], \\
\hat{h}^u_t &= \frac{1-\left(\frac{\widetilde{w}^u_{t+1}}{\widetilde{w}^e_{t+1}}\right)^2}{1-(1-\eta_t)\left[1-\left(\frac{\widetilde{w}^u_{t+1}}{\widetilde{w}^e_{t+1}}\right)^2\right]}\left[\frac{\widetilde{w}^u_t}{2} - (1-\eta_t)h^{0.5}_t + \tilde{s}_t\right].
\end{aligned}$$

where

$$\widetilde{w}^e_t = (1-\tau^e_t)w^e_t = (1-\tau^e_t)A_t(1-\alpha)(n^e_t)^{-\alpha}(1-n^e_t)^{\alpha}, \tag{B.1}$$

$$\widetilde{w}^u_t = (1-\tau^u_t)w^u_t = (1-\tau^u_t)A_t\alpha(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha-1}, \tag{B.2}$$

$$\frac{\widetilde{w}^u_{t+1}}{\widetilde{w}^e_{t+1}} = \frac{n^e_{t+1}}{1-n^e_{t+1}}\left[\frac{(1-\tau^u_t)\alpha}{(1-\tau^e_t)(1-\alpha)}\right], \tag{B.3}$$

$$\begin{aligned}
h^{0.5}_t &= 0.5(\underline{h}_t + \overline{h}_t) = 0.5(\underline{\theta}_t + \overline{\theta}_t)[a + b(n^e_t\tilde{w}^e_t + n^u_t\tilde{w}^u_t)] \\
&= 0.5(\underline{\theta}_t + \overline{\theta}_t)\Big[a + b[A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}(1-\tau^e_t + \alpha(\tau^e_t - \tau^u_t))]\Big] \\
&= \Theta\Big[a + b(1-T)A_t(n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha}\Big]
\end{aligned} \tag{B.4}$$

where $\Theta = 0.5(\underline{\theta}_t + \overline{\theta}_t)$, $T = \tau^e_t - \alpha(\tau^e_t - \tau^u_t)$.

$$
\begin{aligned}
\tilde{s}_t &= \tau_t^e w_t^e n_t^e + \tau_t^u w_t^u n_t^u - g_t \\
&= \tau_t^e n_t^e (1-\alpha) A_t (n_t^e)^{-\alpha} (1-n_t^e)^{\alpha} + \tau_t^u (1-n_t^e) \alpha A_t (n_t^e)^{1-\alpha} (1-n_t^e)^{\alpha-1} - g_t \\
&= (1-\alpha)\tau_t^e A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} + \alpha \tau_t^u A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t \\
&= (1-\alpha)\tau_t^e A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} + \alpha \tau_t^u A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t \\
&= A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}[\tau_t^e - \alpha(\tau_t^e - \tau_t^u)] - g_t \\
&= A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} T - g_t
\end{aligned} \tag{B.5}
$$

$$
\begin{aligned}
\hat{h_t^e} &= \frac{1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1-(1-\eta_t)\Big[1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\Big]}\Big[\frac{\widetilde{w}_t^e}{2} - (1-\eta_t) h_t^{0.5} + \tilde{s}_t\Big] \\
&= \frac{1-\Big[\frac{\alpha(1-\tau_t^u) n_{t+1}^e}{(1-\alpha)(1-\tau_t^e)(1-n_{t+1}^e)}\Big]^2}{(1-\eta_t)\Big[\frac{\alpha(1-\tau_t^u) n_{t+1}^e}{(1-\alpha)(1-\tau_t^e)(1-n_{t+1}^e)}\Big]^2 + \eta_t}\Big\{\frac{1}{2}(1-\tau_t^e) A_t (1-\alpha)(n_t^e)^{-\alpha}(1-n_t^e)^{\alpha} \\
&\quad - (1-\eta_t)\Theta\Big(a + b(1-T) A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big) + A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} T - g_t\Big\} \\
&= \frac{1-\Phi\Big(\frac{n_{t+1}^e}{1-n_{t+1}^e}\Big)^2}{(1-\eta_t)\Phi\Big(\frac{n_{t+1}^e}{1-n_{t+1}^e}\Big)^2 + \eta_t}\Big\{\frac{1}{2}(1-\tau_t^e) A_t (1-\alpha)(n_t^e)^{-\alpha}(1-n_t^e)^{\alpha} - (1-\eta_t)\Theta a \\
&\quad - [\Theta b(1-T)(1-\eta_t) - T] A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Big\},
\end{aligned} \tag{B.6}
$$

where $\Phi = \Big[\frac{\alpha(1-\tau_t^u)}{(1-\alpha)(1-\tau_t^e)}\Big]^2$.

$$
\begin{aligned}
\hat{h_t^u} &= \frac{1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2}{1-(1-\eta_t)\Big[1-\left(\frac{\widetilde{w}_{t+1}^u}{\widetilde{w}_{t+1}^e}\right)^2\Big]}\Big[\frac{\widetilde{w}_t^u}{2} - (1-\eta_t) h_t^{0.5} + \tilde{s}_t\Big] \\
&= \frac{1-\Big[\frac{\alpha(1-\tau_t^u) n_{t+1}^e}{(1-\alpha)(1-\tau_t^e)(1-n_{t+1}^e)}\Big]^2}{(1-\eta_t)\Big[\frac{\alpha(1-\tau_t^u) n_{t+1}^e}{(1-\alpha)(1-\tau_t^e)(1-n_{t+1}^e)}\Big]^2 + \eta_t}\Big\{\frac{1}{2}(1-\tau_t^u) A_t \alpha (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha-1} \\
&\quad - (1-\eta_t)\Theta\Big(a + b(1-T) A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big) + A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} T - g_t\Big\} \\
&= \frac{1-\Phi\Big(\frac{n_{t+1}^e}{1-n_{t+1}^e}\Big)^2}{(1-\eta_t)\Phi\Big(\frac{n_{t+1}^e}{1-n_{t+1}^e}\Big)^2 + \eta_t}\Big\{\frac{1}{2}(1-\tau_t^u) A_t \alpha (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha-1} - (1-\eta_t)\Theta a \\
&\quad - [\Theta b(1-T)(1-\eta_t) - T] A_t (n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Big\},
\end{aligned} \tag{B.7}
$$

Then substituting them into s_t^e and s_t^u, so we can express them as:

$$
\begin{aligned}
s_t^e &= (1-\eta_t)(\hat{h}_t^e - h_t^{0.5}) + \tilde{s}_t \\
&= (1-\eta_t)\Bigg(\frac{1-\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2}{(1-\eta_t)\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2+\eta_t}\Bigg\{\frac{1}{2}(1-\tau_t^e)A_t(1-\alpha)(n_t^e)^{-\alpha}(1-n_t^e)^{\alpha} - (1-\lambda_t)\Theta a \\
&- [\Theta b(1-\eta_t)(1-T)-T]A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Bigg\} - \Theta\Big[a + b(1-T)A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big]\Bigg) \\
&+ A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}T - g_t
\end{aligned} \tag{B.8}
$$

$$
\begin{aligned}
s_t^u &= (1-\eta_t)(\hat{h}_t^u - h_t^{0.5}) + \tilde{s}_t \\
&= (1-\eta_t)\Bigg(\frac{1-\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2}{(1-\eta_t)\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2+\eta_t}\Bigg\{\frac{1}{2}(1-\tau_t^u)A_t\alpha(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha-1} - (1-\eta_t)\Theta a \\
&- [\Theta b(1-\eta_t)(1-T)-T]A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Bigg\} - \Theta\Big[a + b(1-T)A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big]\Bigg) \\
&+ A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}T - g_t
\end{aligned} \tag{B.9}
$$

At the end, the government's budget constraint can be expressed as:

$$
\begin{aligned}
g_t &= \tau_t^e w_t^e n_t^e + \tau_t^u w_t^u (1-n_t^e) - s_t^e n_t^e - s_t^u(1-n_t^e) \\
&= A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}T \\
&- n_t^e\Bigg\{(1-\eta_t)\Bigg(\frac{1-\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2}{(1-\eta_t)\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2+\eta_t}\Bigg\{\frac{1}{2}(1-\tau_t^e)A_t(1-\alpha)(n_t^e)^{-\alpha}(1-n_t^e)^{\alpha} - (1-_t)\Theta a \\
&- [\Theta b(1-\eta_t)(1-T)-T]A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Bigg\} - \Theta\Big[a + b(1-T)A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big]\Bigg) \\
&+ A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}T - g_t \\
&- (1-n_t^e)\Bigg\{(1-\eta_t)\Bigg(\frac{1-\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2}{(1-\eta_t)\Phi\left(\frac{n_{t+1}^e}{1-n_{t+1}^e}\right)^2+\eta_t}\Bigg\{\frac{1}{2}(1-\tau_t^u)A_t\alpha(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha-1} - (1-\eta_t)\Theta a \\
&- [\Theta b(1-\eta_t)(1-T)-T]A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha} - g_t\Bigg\} - \Theta\Big[a + b(1-T)A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}\Big]\Bigg) \\
&+ A_t(n_t^e)^{1-\alpha}(1-n_t^e)^{\alpha}T - g_t\Bigg\}
\end{aligned} \tag{B.10}
$$

C The Lagrangian function the government in Ramsey problem

The Lagrangian function of the government can be written as:

$$
\begin{aligned}
\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Bigg\{ & n^e_{t+1}(\log c^e_{t+1} + \log x^e_{t+1}) + (1 - n^e_{t+1})(\log c^u_{t+1} + \log x^u_{t+1}) \\
& + \lambda^1_t \Big\{ (1 - \tau^e_{t+1})(1-\alpha) A_{t+1} (n^e_{t+1})^{-\alpha} (1 - n^e_{t+1})^{\alpha} - c^e_{t+1} - x^e_{t+1} \Big\} \\
& + \lambda^2_t \Big\{ (1 - \tau^u_{t+1})\alpha A_{t+1} (n^e_{t+1})^{1-\alpha} (1 - n^e_{t+1})^{\alpha-1} - c^u_{t+1} - x^u_{t+1} \Big\} \\
& + \lambda^3_t \left\{ \frac{1}{x^e_{t+1}} - \frac{1}{c^e_{t+1}} \right\} \\
& + \lambda^4_t \left\{ \frac{1}{x^u_{t+1}} - \frac{1}{c^u_{t+1}} \right\} \\
& + \lambda^5_t \Bigg\{ A_t (n^e_t)^{1-\alpha} (1 - n^e_t)^{\alpha} T \\
& - n^e_t \Bigg\{ (1-\eta_t) \Bigg(\frac{1 - \Phi\Big(\frac{n^e_{t+1}}{1-n^e_{t+1}}\Big)^2}{(1-\eta_t)\Phi\Big(\frac{n^e_{t+1}}{1-n^e_{t+1}}\Big)^2 + \eta_t} \Bigg\{ \frac{1}{2}(1-\tau^e_t) A_t (1-\alpha)(n^e_t)^{-\alpha}(1-n^e_t)^{\alpha} - (1-\eta_t)\Theta a \\
& - [\Theta b (1-\eta_t)(1-T) - T] A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - g_t \Bigg\} - \Theta\Big[a + b(1-T) A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} \Big] \Bigg) \\
& + A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} T - g_t \Bigg\} \\
& - (1-n^e_t) \Bigg\{ (1-\eta_t) \Bigg(\frac{1 - \Phi\Big(\frac{n^e_{t+1}}{1-n^e_{t+1}}\Big)^2}{(1-\eta_t)\Phi\Big(\frac{n^e_{t+1}}{1-n^e_{t+1}}\Big)^2 + \eta_t} \Bigg\{ \frac{1}{2}(1-\tau^u_t) A_t \alpha (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha-1} - (1-\eta_t)\Theta a \\
& - [\Theta b (1-\eta_t)(1-T) - T] A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - g_t \Bigg\} - \Theta\Big[a + b(1-T) A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} \Big] \Bigg) \\
& + A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} T - g_t \Bigg\} - g_t \Bigg\} \\
& + \lambda^6_t \Big\{ A_t (n^e_t)^{1-\alpha}(1-n^e_t)^{\alpha} - c^e_t n^e_t - c^u_t (1-n^e_t) - g_t \Big\} \Bigg\},
\end{aligned}
$$

where λ^i_t (for $i = 1, 2, ..., 6$) represents the multiplier associated with each constraint. The constraints in the Lagrangian function have been rearranged to ensure that all multipliers are non-negative at the steady state. Additionally, the FOCs of the government should also include the constraints to the Ramsey problem.[19]

[19] The derivation of FOCs of the Ramsey are available unpon request.