

An adjustment cost model of social mobility

Parantap Basu^{*,a}, Yoseph Getachew^b

^a*Durham University Business School, Durham University, Mill Hill Lane, DH1 3LB, Durham, UK, e-mail: parantap.basu@durham.ac.uk, tel.: +447872040275*

^b*Department of Economics, University of Pretoria, 0028, Pretoria, South Africa, e-mail: yoseph.getachew@up.ac.za, tel.: +27760722472*

Abstract

We analyze herein the effects of the human capital adjustment cost on social mobility. Such an adjustment cost is modeled as a rising marginal cost schedule for augmenting human capital. We use a general human capital technology, which disentangles the adjustment cost from the depreciation cost of the human capital. Missing credit markets prevent individuals from equalizing the initial differences in the human capital. We find that a higher adjustment cost for human capital acquisition slows down the social mobility and results in a persistent inequality across generations. On the other hand, a higher rate of human capital depreciation could increase mobility via a positive effect on new investment. The quantitative analysis of our model suggests that the human capital adjustment cost is nontrivial to account for the observed persistence of inequality and social mobility. In addition, we find that the government redistribution policy could account for the large observed variation in estimates of social mobility.

Key words:

Adjustment cost of capital, inequality persistence, intergenerational mobility

JEL Classification: D24, D31, E13, O41

^{*}Corresponding author

1. Introduction

It is an open question whether the son of a poor farmer will become a highly paid executive manager. The evidence during the last two decades is mixed. A large body of literature documenting intergenerational income elasticity provides disparate evidence of social mobility. Although Machin (2004) and Clark and Cummins (2012) argue that there is considerable persistence in the wealth status of households in England from 1800 to 2012, other papers such as Grawe and Corak (2004) document considerably faster social mobility. Social mobility is an important issue in macro-development literature because it is inextricably connected to the intergenerational persistence of the inequality or dispersion of wealth. If income inequality is persistent, social or intergenerational mobility is likely to be slower.¹ The seminal paper of Becker and Tomes (1979) draws the conclusion that a stable distribution of income could be explained by individual and market luck. Their crucial assumption is that the credit market is perfect, implying that individuals with low wealth and a high marginal product of capital could borrow from individuals with the opposite trait. This tends to equalize the differences in wealth. The residual inequality is then mostly attributed to luck. Since then a considerable body of literature (e.g., Lounry, 1981, Banerjee and Newman, 1993, Galore and Zeira, 1993, Benabou, 1996, Mulligan, 1997, Bandyopadhyay and Basu, 2005, Bandyopadhyay and Tang, 2011) has evolved emphasizing the role of credit market imperfection in perpetuating the inequality.

In this paper we explore the role of the adjustment cost, a relatively ignored feature of human capital production, on social mobility. The human capital adjustment cost is modeled as a rising marginal cost of augmenting human capital that is measured in terms of foregone consumption. Such an adjustment cost can arise for a number of reasons. First, there could be basic human inertia to respond to change and adjust to new opportunities or a new environment. An example of such inertia is where adults find a better job opportunity with higher pay in a region that is remote

¹Although social mobility is a broader notion of a change in social status, including occupation, we use this term narrowly to indicate the intergenerational mobility of wealth where the wealth is primarily intangible human capital.

from their home town but due to friends and family ties they are reluctant to move (Alesina and Giuliano, 2010). Similar sluggish and varied responses to opportunity are found in the study of Katz et al. (2001) where adults in high poverty inner cities in the United States respond rather sluggishly to a subsidized move to low poverty regions; this is known as the Moving to Opportunity for Fair Housing (MTO) program. Second, a similar adjustment cost could be attributed to market based factors such as the higher cost of an advanced education compared to primary schooling or a higher employment adjustment cost as in Hansen and Sargent (1980). We show that the presence of such an adjustment cost of human capital could impede the process of social mobility.

We develop a model with missing credit and insurance markets as in Loury (1981), Galor and Zeira (1993), and Benabou (2000, 2002). Individuals differ in terms of the initial distribution of human capital and productivities, which cannot be hedged using credit and insurance markets. In this environment, a higher adjustment cost impedes social mobility through the following transmission channel. When the credit market is missing, the investment opportunities facing individuals (investment in human capital or education in our model) are limited to the resources that they have in hand. Given a production function with private diminishing returns to reproducible human capital, poor people with lower human capital have a higher marginal product than rich people. Thus, relative growth potential of poor people is higher. However, if a human capital adjustment cost is present, this growth of the poor will be impeded because they face a higher marginal cost of investment when they try to grow. Thus, the adjustment cost will slow the process of social mobility leading to a higher persistence of human capital inequality in the aggregate. The central point of our paper is to demonstrate that a society facing such a costly adjustment of human capital could experience persistent inequality and low social mobility measured by the intergenerational elasticity of income. Although an adjustment cost is a standard feature in any physical capital production function, quite surprisingly, its role has been ignored in the inequality and social mobility literature.

Our functional form for the human capital technology is more general compared to the extant literature (e.g., Benabou, 2000, 2002). We parametrize two major costs

of human capital acquisition, namely the adjustment cost and the depreciation cost. The depreciation of human capital could be attributed to the inherent obsolescence of skills. Using such a general functional form for human capital production, we also show that a higher depreciation of human capital could expedite social mobility. This is because a high depreciation of human capital triggers a greater investment in human capital to replenish the human capital lost due to obsolescence. Incomplete depreciation of human capital makes social mobility dependent on the history of inequality. It also enables us to understand intergenerational knowledge transfer in the spirit of Mankiw et al. (1992). To the best of our knowledge, a human capital production function allowing for an adjustment cost as well as the incomplete depreciation of human capital in determining social mobility has not been explored in the literature.²

In our model, in the spirit of Galor and Zeira (1993), adults receive a warm-glow utility from investing in their child's education. As in Loury (1981), human capital is the only form of reproducible capital in the economy and the credit market is missing. Idiosyncratic productivity shocks, together with an initial difference in human capital and missing credit markets, give rise to a cross-sectional inequality that transmits from one generation to another. The absence of credit and insurance markets prevents agents from mitigating negative idiosyncratic shocks. Unlucky agents experiencing a negative productivity shock invest less resources in their child's education, which means that the child inherits less human capital. How quickly the offspring overcomes this disadvantage depends on how costly it is to adjust the human capital. We develop a novel closed form analytical solution for the endogenous law of motion of inequality. The key theoretical result is that inequality is persistent and social mobility is less in economies with a higher human capital adjustment cost or a lower depreciation cost of capital. In addition, social mobility is less in more unequal societies. Since inequality is history dependent in our model, social mobility

²Ben-Porath (1967) uses a human capital production function that has a similar tenor to ours. He explores the implications of the rising marginal cost of investment in human capital. However, he does not explore the implications of such a rising marginal cost for social mobility, which is the main theme in our paper.

also shares the same property.

In our quantitative exercise, we explore how the adjustment cost and depreciation costs alone can explain the observed social mobility. We find that the human capital adjustment cost has nontrivial effects on social mobility and long-run inequality. On the other hand, the depreciation cost has quantitatively minor effects on these variables. Our calibrated social mobility parameter is in accordance with the slow social mobility predicted by Mazumder (2005) and Clark and Cummins (2012). Adding a redistributive government policy with public funding of education significantly alters the calibrated social mobility. A pro-poor public service program expedites intergenerational mobility and brings the estimate of intergenerational earnings elasticity on par with the majority of studies.

Using our adjustment cost model, we also calibrate an adult's human capital response to luck. In the absence of any other estimate of such a response, we target the response rate of adults in the well-known MTO program (reported by Katz et al., 2001) where adults are given the opportunity to move away from a high poverty region. Our adjustment cost model is in accordance with the response rate of adults to the MTO program. The resulting impulse responses of human capital with respect to initial luck differences suggest that the social mobility is slower in economies with higher adjustment costs. These are all consistent with our key theoretical results.³

The paper is organized as follows. Section 2 presents the model and the dynamics and equilibrium of individual wealth accumulation. Section 3 provides the quantitative analysis of the model. Section 4 studies the distributional dynamics in an environment with public funding and Section 5 concludes.

³Our key result that higher human capital adjustment costs slow down social mobility is robust in a more general model environment in an infinite horizon, with physical investment and time allocation decisions. The details are available as a supplementary appendix from the authors.

2. The model

2.1. Preference and technology

Consider a continuum of heterogeneous households $i \in [0, 1]$ in overlapping generations. Each household i consists of an adult of generation t attached to a child. A child only inherits human capital from its parents and does not make any decisions as the child's consumption is already included in that of the parents. An adult, at date t , employs a unit of raw labor and human capital into the production process that translates into h_{it} efficiency units for the production of final goods and services to earn income (y_{it}) using the following Cobb-Douglas production function:

$$y_{it} = a_1 \varphi_{it} (h_t)^{1-\alpha} (h_{it})^\alpha \quad (1)$$

where $a_1 > 0$ is simply an exogenous productivity parameter, $\alpha \in (0, 1)$, h_t represents the aggregate stock of knowledge in the spirit of Arrow (1962) and Romer (1986).⁴ Individuals are subject to i.i.d idiosyncratic productivity shocks (φ_{it}) that drive their total marginal productivity. The idiosyncratic shock φ_{it} follows the process: $\ln \varphi_{it} \sim N(-v^2/2, v^2)$. The child at date t behaves as an adult at $t + 1$.

2.1.1. Utility function and budget constraint

Agents care about their own consumption and receive a "joy of giving" from the human capital stock of their children. In other words, the utility of the adult at date t is given by:⁵

$$u(c_{it}, h_{it+1}) = \ln c_{it} + \beta \ln h_{it+1} \quad (2)$$

where $0 < \beta < 1$ is the degree of parental altruism; h_{it+1} represents the human capital of the offspring of agent i . At the end of the period, the parents allocate income

⁴Such a technology basically means that there are private diminishing returns but social constant returns to human capital.

⁵The choice of a logarithmic utility function and altruistic agents with a "joy of giving" motive is merely for simplicity. Also see Glomm and Ravikumar (1992), Galor and Zeira (1993), Saint-Paul and Verdier (1993), and Benabou (2000) for similar settings.

between current consumption (c_{it}) and saving (s_{it}). The latter is used for investment in the human capital accumulation of the offspring. The budget constraint is thus given by:

$$c_{it} + s_{it} = y_{it} \tag{3}$$

2.1.2. Technology of human capital production

The human capital is the only form of reproducible input in our model. The stock of human capital inherited by the current adults from their predecessors determines their state of technological knowledge, which can be modified to advance the technological frontier. This can be done by investing in education. The production of the next period human capital (h_{it+1}) takes place with the aid of the parent's time (l_{it}), the parent's human capital (h_{it}) and the investment in schooling (s_{it}):

$$h_{it+1} = a_2 (l_{it}h_{it})^{1-\theta} ((1-\delta)h_{it} + s_{it})^\theta \tag{4}$$

where $\theta \in (0, 1)$, $\delta \in (0, 1)$ and $a_2 > 0$.

The human capital production function is in the spirit of Glomm and Ravikumar (1992) and Benabou (2002) except for the inclusion of the depreciation parameter δ . The term $l_{it}h_{it}$ may be used to capture home schooling in quality time as knowledgeable parents are better equipped to promote the learning of their children. Without any loss of generality, we normalize l_{it} to unity.⁶ In contrast, $(1-\delta)h_{it}$ is used to capture some inherited component of human capital, which represents the amount of workable human capital that a child inherits from its parents in the absence of any new investment. If adults undertake no investment in their child's education, for instance, unlike Benabou (2000), the child still inherits some workable human capital in proportion to $(1-\delta)h_{it}$. Viewed from this perspective, one may think of $1-\delta$ as the degree of intergenerational spillover of knowledge as in Mankiw et al. (1992) and Bandyopadhyay and Basu (2005).

⁶It is straightforward to generalize the model to endogenize time allocation. We do this in an infinite horizon model, which is available in a supplementary appendix.

The parameter θ in the human capital production function (4) is of central interest in this paper. It determines the curvature of the marginal return to investment ($\partial h_{it+1}/\partial s_{it}$), which we ascribe to a convex human capital adjustment cost. The marginal return to investment based on (4) is given by:

$$\partial h_{it+1}/\partial s_{it} = a_2\theta / (1 - \delta + s_{it}/h_{it})^{1-\theta} \quad (5)$$

Figure 1 plots (5), $\partial h_{it+1}/\partial s_{it}$ against s_{it}/h_{it} for $\theta = 0.8$ (our baseline value in the calibration later) and $\theta = 0.6$.⁷ Lower θ makes the investment return schedule shift downward with a steeper curvature. This steep decrease in marginal return to investment due to lower θ is ascribed to a higher adjustment cost of human capital. If θ reaches the upper bound of unity, there is zero adjustment cost and the investment technology reverts to a standard linear depreciation rule. This notion of θ as the degree of the human capital adjustment cost is borrowed from the standard capital adjustment cost technology used in Lucas and Prescott (1971), Basu (1987), and Basu et al. (2012).

To see more clearly the close connection between θ and the adjustment cost, take the reciprocal of $\partial h_{it+1}/\partial s_{it}$ in (5) to get the marginal cost of individual investment that characterizes the individual specific marginal Tobin's q of human capital (q_{it}). How q_{it} responds to the i th agent's effort to augment human capital is a measure of the adjustment cost. Using (4), q_{it} can be rewritten as:

$$q_{it} = (h_{it+1}/h_{it})^{(1-\theta)/\theta} / (\theta a_2) \quad (6)$$

The elasticity of q_{it} (with respect to h_{it+1}/h_{it}) is $(1 - \theta)/\theta$. A lower θ makes the Tobin's q rise more steeply in response to the agent's attempt to augment human capital. In view of this, hereafter a lower value of θ will be interpreted as a higher adjustment cost of investing in human capital.

⁷The other two parameters a_2 and δ are fixed at their baseline levels of 1.655 and 0.03, respectively (see Section 3).

Figure 1: Effect of a change in θ on the marginal return to investment

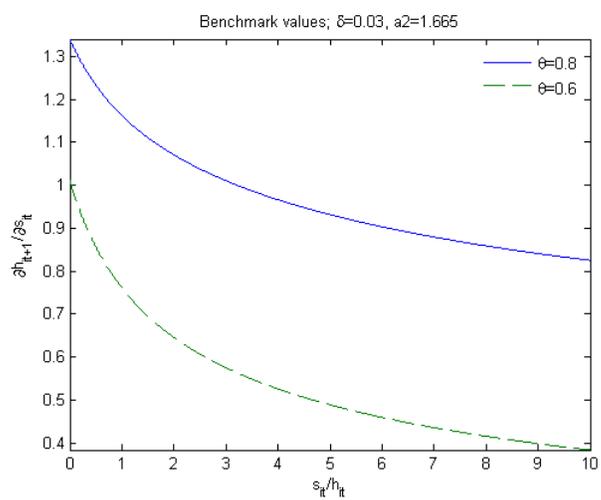
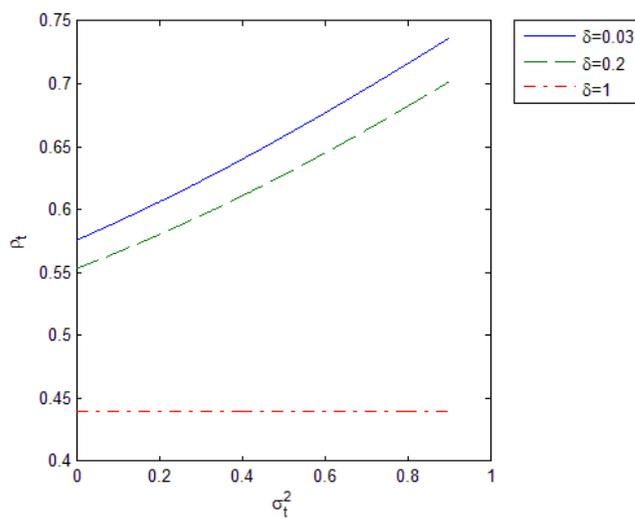


Figure 2: Social mobility versus inequality



Although a lower θ drives down the return to investment via the adjustment cost channel, a lower rate of depreciation of human capital makes the capital long-lasting, which contributes to a lower marginal return on investment (5). A lower depreciation cost makes the current generation inherit more human capital from their ancestors (along with a "joy of giving" bequest), which, as will be seen later, contributes to slower convergence.

2.2. Initial distribution of human capital

At the beginning, each adult of the initial generation is endowed with human capital h_{i0} . The distribution of h_{i0} takes a known probability distribution,

$$\ln h_{i0} \sim N(\mu_0, \sigma_0^2) \quad (7)$$

and it evolves over time along an equilibrium trajectory.

2.3. Equilibrium

In equilibrium, all individuals behave optimally and the aggregate consistency conditions hold.

Optimality: An adult of cohort t solves the following problem, obtained by substituting (2) and (4) into (3),

$$\max_{s_{it}} \left\{ \ln(y_{it} - s_{it}) + \beta \ln((1 - \delta)h_{it} + s_{it})^\theta \right\} \quad (8)$$

taking h_{it} as given. The optimization yields the following investment function,

$$s_{it} = (y_{it}\theta\beta - (1 - \delta)h_{it}) / (1 + \theta\beta) \quad (9)$$

An adult's optimal investment decision constitutes new investment plus a replacement of depreciated capital. A lower rate of depreciation depresses current investment because the adult carries forward human capital from the previous generation.

Note that, in general, such models allow disinvestment ($s_{it} \leq 0$), in which case some adults could consume more than their income at the expense of depleting the

human capital of their children.⁸ Nevertheless, the optimal human capital of the offsprings always remains positive in our model.⁹

Aggregate Consistency: (i) $c_t \equiv \int c_{it} di$, $s_t \equiv \int s_{it} di$, $y_t \equiv \int y_{it} di$, $h_t \equiv \int h_{it} di$ where the left-hand side variable in each of them means the aggregate. (ii) The aggregate budget constraint is thus given by:¹⁰

$$c_t + s_t = y_t \tag{10}$$

2.4. Individual optimal human capital accumulation

From (1), (4), and (9), the i th individual optimal human capital accumulation is given by,

$$h_{it+1} = \phi h_{it} \left(1 - \delta + a_1 \varphi_{it} h_t^{1-\alpha} h_{it}^{\alpha-1}\right)^\theta \tag{11}$$

where

$$\phi \equiv a_2 (\theta\beta / (1 + \theta\beta))^\theta$$

Thus, the i th individual offspring's optimal human capital, which is always positive, is determined by both the depreciation and adjustment cost of human capital and the parental income.

2.5. Incomplete depreciation, adjustment cost, and social mobility

Loglinearizing (11) around a balanced growth path where all agents are identical in terms of luck ($\ln \varphi_{it} = \ln \varphi = 0$), one gets a clear picture about the evolution of

⁸Real life examples of human capital disinvestment include child abuse or sending one's offspring to work as child labor in less developed countries without undertaking any investment in schooling. This could arise in the model if adults are very myopic (low β) or if the total factor productivity (TFP) is too low.

⁹See eq. (11). Also, in our calibration exercise (Section 3), the extreme scenario of disinvestment in children does not arise for plausible parameter values.

¹⁰We use the operators \int and E interchangeably in the text to denote aggregation across individuals.

individual human capital.¹¹

$$\ln h_{it+1} \simeq \varrho \ln h_{it} + \chi \ln \varphi_{it} \quad (12)$$

where

$$\varrho \equiv 1 - \chi(1 - \alpha) \in (0, 1) \quad (13)$$

$$\chi \equiv \theta a_1 / (1 - \delta + a_1) \in (0, 1) \quad (14)$$

Eq. (12) implies that, around a balanced growth path, children inherit human capital related traits from their parents to the extent of ϱ . The inverse of ϱ is often considered as the measure of social mobility (e.g., Benabou, 2002). A larger value of ϱ indicates a slower social mobility. A lower depreciation rate (lower δ) and a higher adjustment cost (lower θ) raise ϱ and thus imply a slower social mobility.

2.6. Distributional dynamics

We are now ready to characterize the dynamics of the cross sectional variance of human capital:

Proposition 1. *Given the initial cross sectional inequality characterized by (7) and (11), the dynamics of inequality and growth are given by the following laws of motion, respectively,*

$$\sigma_{t+1}^2 = \theta^2 \ln \frac{\kappa^2 \exp(\theta^{-2} \sigma_t^2) + (a_1)^2 \exp((0.5\omega + \lambda^2) \sigma_t^2 + v^2) + 2\kappa a_1 \exp((0.5\omega + \lambda/\theta) \sigma_t^2)}{(\kappa + a_1 \exp(0.5\omega \sigma_t^2))^2} \quad (15)$$

and

$$\gamma_{t+1} = \ln \phi + 0.5(1/\theta - 1)(\sigma_t^2 - \sigma_{t+1}^2) + \theta \ln(\kappa + a_1 \exp(0.5\omega \sigma_t^2)) \quad (16)$$

¹¹Algebraic derivation is omitted for brevity and is available from the authors upon request.

where

$$\begin{aligned}\gamma_{t+1} &\equiv \ln h_{t+1} - \ln h_t, \quad \sigma_t^2 = \text{var}(\ln h_{it}) \\ \kappa &\equiv 1 - \delta, \quad \lambda \equiv 1/\theta + \alpha - 1 > 0 \\ \omega &\equiv (\alpha - 1)(2/\theta + \alpha - 2) < 0\end{aligned}$$

Proof. See the Appendix. ■

The dynamics of inequality is thus determined by its own history as seen from (15). Although inequality is not influenced by the growth rate of the economy, growth depends on the current and past inequality. This is evident from the fact that σ_{t+1}^2 is a function of σ_t^2 alone whereas γ_{t+1} depends on σ_{t+1}^2 and σ_t^2 .

2.6.1. History dependent social mobility

The social mobility based on the exact solution is the inverse of the gradient of (15). This gradient is given by, (see the Appendix for details of the derivation),

$$\varrho_t^2 \equiv \partial\sigma_{t+1}^2/\partial\sigma_t^2 = f(\sigma_t^2) \tag{17}$$

The exact solution for social mobility (17) reveals a path dependence property. It depends on the current state of inequality, σ_t^2 , which is history dependent as seen in (15). Figure 2 plots ϱ_t against σ_t^2 for alternative values of the depreciation parameter δ . In line with Clark's (2013) empirical finding, social mobility is less in a more unequal society.¹² Incomplete depreciation discourages new investment in human capital and prevents the new generation from overcoming the initial deficiency of human capital. Thus, initial inequality slows down current social mobility. Lower depreciation impedes mobility for all inequality states as seen by the comparison (when $\delta = 0.2$, $\delta = 0.03$ and $\delta = 1$). It is noteworthy that for full depreciation ($\delta = 1$), this mobility loses its history dependence property. In this case, ϱ_t reduces to a constant, $\varrho = 1 - \theta(1 - \alpha)$. Because (11) is loglinear when $\delta = 1$, the loglinearization and the actual solution converge.

¹²See Figures 1 and 2 of Clark (2013).

Taking the variance of the loglinear version in (12), one clearly sees the roles θ and δ play on the dynamics of inequality of the economy:¹³

$$\sigma_{t+1}^2 = \varrho^2 \sigma_t^2 + \chi^2 v^2 \quad (18)$$

Eq. (17) reduces to (18) when $\delta = 1$. Combining (1) and (18) gives the following dynamics of income inequality ($\sigma_{y,t}^2$):

$$\sigma_{y,t+1}^2 = \varrho^2 \sigma_{y,t}^2 + v^2 (1 - \varrho^2 + \alpha^2 \chi^2) \quad (19)$$

The individual capital share parameter, α , and the adjustment cost parameter, θ , have opposing effects on the persistence of inequality. When α is higher, the relative growth potential of the poor with respect to the rich (due to the poor's higher marginal product) is dampened, which means that the process of convergence between the rich and the poor will be slower. This explains why the initial inequality tends to persist when α is higher.¹⁴ On the other hand, a higher adjustment cost (lower θ) will make the process of convergence between the poor and the rich slower because it is costly for the poor to invest. This technological disadvantage imposed by the human capital adjustment cost is compounded by the credit market imperfection that makes the social mobility slower and inequality more persistent.¹⁵ These effects of θ and α on the social mobility are remarkably robust to alternative environments.¹⁶

On the other hand, a lower depreciation rate makes the inequality process more

¹³Because of its highly nonlinear nature, it is difficult to ascertain the comparative statics effects of the four underlying parameters. We study them numerically in the next section. The numerical comparative results are in accordance with the loglinearized version.

¹⁴The inverse relationship between the rate of convergence and the capital share parameter is well known in the convergence literature (see for example, Benabou, 2002).

¹⁵When α is close to unity, the adjustment cost ceases to play any role in determining the inequality persistence because the poor do not have any relative advantage in terms of higher marginal product.

¹⁶In a supplementary appendix (available from the authors upon request) we have worked out a model with dynastic altruism, as in Barro (1974) and derive the same results.

persistent by depressing current investment because the adult carries forward human capital from the previous generation, as shown in eq. (9). A greater depreciation or obsolescence of human capital promotes more investment in schooling.¹⁷ Since investment is the only vehicle for social mobility, lower investment in human capital caused by low depreciation slows this process of social mobility and makes the inequality more persistent. The following proposition summarizes the results.

Proposition 2. *A higher degree of adjustment cost (lower θ), lower depreciation cost (δ), and a higher capital share α make the social mobility slower and the inequality process more persistent.*

2.7. Long-run inequality and growth

The steady-state inequality and growth based on the closed form solutions (15) and (16) are given by the following expressions, respectively:

$$\sigma^2 = \theta^2 \ln \frac{\kappa^2 \exp(\theta^{-2} \sigma^2) + (a_1)^2 \exp((0.5\omega + \lambda^2) \sigma^2 + v^2) + 2\kappa a_1 \exp((0.5\omega + \lambda/\theta) \sigma^2)}{(\kappa + a_1 \exp(0.5\omega \sigma^2))^2} \quad (20)$$

and

$$\gamma = \ln \phi + \theta \ln \{ \kappa + a_1 \exp(0.5\omega \sigma^2) \} \quad (21)$$

The steady-state equivalents based on the log-linearized version of the model have simpler expressions. Considering (18),

$$\sigma^2 = \chi^2 v^2 / (1 - \varrho^2) \quad (22)$$

The steady-state income inequality (σ_y^2) based on (19) is:

$$\sigma_y^2 = v^2 (1 - \varrho^2 + \alpha^2 \chi^2) / (1 - \varrho^2) \quad (23)$$

¹⁷There are two principal drivers of human capital depreciation: (i) technical obsolescence and (ii) economic obsolescence. See Rosen (1975), de Grip and van Loo (2002), and de Grip (2006) for a discussion on human capital obsolescence. The former is due to changes that originate in individuals' personal circumstances such as ageing, illness, injury while technological progress accounts for the latter.

Note that when $\delta = 1$, all of the loglinearized and the exact solutions converge.¹⁸ Thus, inequality in the long-run is mainly the result of individuals' differences in human capital investment decisions in response to differences in luck.

3. Quantitative analysis

In this section, we evaluate the quantitative effects of θ and δ on social mobility and inequality dynamics based on our model. We first construct parameter values that reasonably reflect actual economies. Assuming a psychological discount factor of 0.96, we set $\beta = 0.96^{30} \approx 0.3$, in a period of 30 years (de la Croix and Michel, 2002, p.255).¹⁹ The choices of $v^2 = 0.42$, $a_1 = 1.96$ and $a_2 = 1.655$ are made to target the variance of log of income (σ_y^2) equal to 0.4386 and a long-run annual average growth rate of about 1.94 percent for the US economy.²⁰ Regarding θ , we take Glomm's (1997) estimate of 0.8 as a baseline. The baseline value of δ is taken from Mankiw et al. (1992). The intergenerational wealth or earning elasticity (ϱ) is very sensitive to the choice of α . We fix a value of α equal to 0.3, which gives a baseline calibrated estimate of social mobility of 0.6254. This estimate falls within the range of the Mazumder (2005) and Clark and Cummins (2012) estimates.²¹ This estimate is still on the low side but we take it as our initial baseline estimate. In the next section we introduce a redistributive government policy that would bring the social mobility estimate more into line with the median estimate in the literature.

Our calibrated model also has implications for the human capital elasticity with respect to luck, χ , which is an indicator of an agent's response to luck or opportu-

¹⁸For instance, the steady-state inequality in both (20) and (22) will reduce to $\sigma^2 = v^2\theta / ((1 - \alpha)(2 - (1 - \alpha)\theta))$.

¹⁹A psychological discount factor of 0.96 matches a 4.17 percent rate of time preference ρ in an infinite horizon. That is, $\beta = 1 / (1 + \rho) = 1 / (1 + .0417) = 0.96$.

²⁰Assuming a lognormal distribution of income, the mean-median ratio implies a 0.44 average log-income variance for the United States for the years 1991, 1994, 1997, and 2000 based on the Luxembourg Income Study (UNU-WIDER, 2007).

²¹Mazumder (2005) estimates the intergenerational elasticity for the United States at about 0.6. He argues that the earlier estimates of intergenerational elasticity are downward biased by more than 30% due to persistent transitory fluctuations. On the other hand, Clark and Cummins (2012) get ϱ estimates between 0.7 and 0.8 for the United Kingdom.

nity. We use the response rate of households from the well known MTO program as reported by Katz et al. (2001) as a proxy for the agent’s response to luck. Katz et al. (2001) report that about 48% to 62% of households living in high poverty regions in Boston move through the MTO program.²² Using the above calibrated values, we get an estimate for this response to luck of around 0.54, which is in the range of Katz et al.’s (2001) study. This provides some additional credence to our calibrated structural parameters. Table 1 summarizes the baseline parameter values.

Table 1: Benchmark values

Preference and technology parameters	$\beta = 0.3, a_1 = 1.96, a_2 = 1.655$
Production and policy parameters	$\alpha = 0.3, \theta = 0.8, \delta = 0.03$
Inequality parameter	$v^2 = 0.42$

Tables 2 and 3 report the sensitivity analysis of social mobility ϱ with respect to changes in θ and δ around the baseline values. A higher adjustment cost and lower depreciation cost (lower θ and δ , respectively) slow down social mobility (leading to higher ϱ).

²²The MTO program in Boston was a housing mobility programme in which low-income families in the inner city in Boston were given assistance (vouchers) to move to less segregated regions and low poverty regions. Katz et al. (2001) report the results of a randomized control trial experiment. The share of MTO-Boston families in their sample who used the voucher (which they refer to as the "take-up rate") ranged from 48% to 62%. We take this response as a proxy for an agent’s response to luck. A higher response is deemed to be an indicator of greater mobility.

Table 2: Effects of adjustment cost on inequality, mobility, and growth

Adjustment cost (θ)	ϱ
0.9	0.5786
0.85	0.6020
0.8	0.6254
0.75	0.6488
0.7	0.6722
0.6	0.7190

Table 3: Effects of depreciation cost on inequality, mobility, and growth

Depreciation cost (δ)	ϱ
0.2	0.6023
0.15	0.6094
0.13	0.6122
0.10	0.6162
0.05	0.6228
0.03	0.6254
0.01	0.6279

Figure 3 examines the inequality dynamics based on the closed form solution (15) with respect to the key parameters θ and δ . It demonstrates the effects of changes in adjustment and depreciation costs on the distributional dynamics. Given the baseline values of other parameters, a higher adjustment cost (or a decrease in θ from 0.8 to 0.6) slows down the social mobility by about 4 generations. A lower rate of depreciation (by 2%) slows down the convergence but its quantitative effect is rather small.

Figure 3: Adjustment and depreciation costs, with inequality dynamics.

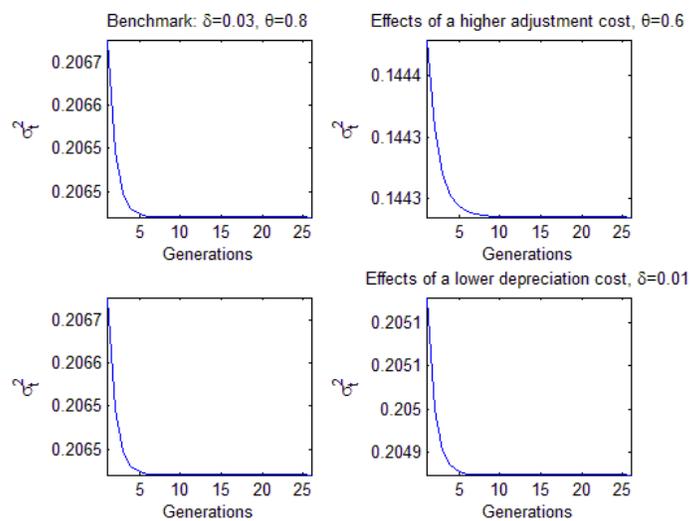
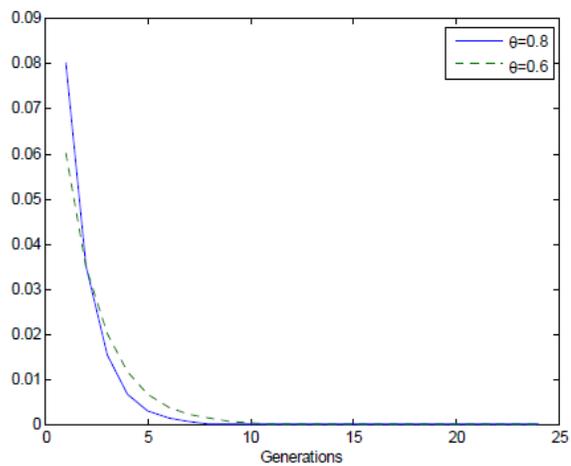


Figure 4: Effect of a difference in luck on human capital when $\theta=0.6$ and $\theta=0.8$



state variables as follows:

$$v(h_{it}, z_{it}) = \pi_0 + \pi_1 \ln h_{it} + \pi_2 \ln z_{it} \quad (\text{B.32})$$

where $z_{it} \equiv f\varphi_{it}h_t^{1-\alpha}$. Plugging (B.32) into the value function (B.31),

$$\begin{aligned} & \pi_0 + \pi_1 \ln h_{it} + \pi_2 \ln z_{it} \quad (\text{B.33}) \\ &= \max_{h_{it+1}} \left[\ln \left\{ y_{it} - \frac{(h_{it+1})^{1/\theta}}{l_{h,it} (h_{it})^{(1-\theta)/\theta}} \right\} + b \ln(1 - l - l_{h,it}) + \beta \{ \pi_0 + \pi_1 \ln h_{it+1} + \ln z_{it+1} \} \right] \end{aligned}$$

We will use the method of undetermined coefficients to solve for π_i which only matters for determining the decision rule of investment. Conjecture that $l_{h,it}$ is time invariant and is equal to $l_{h,i}$.

Differentiating with respect to h_{it+1} and rearranging terms one gets:

$$h_{it+1} = (\pi_1 \beta \theta / (1 + \pi_1 \beta \theta))^\theta (h_{it})^{\alpha\theta+1-\theta} (l_{h,it})^\theta (z_{it})^\theta \quad (\text{B.34})$$

Plugging (B.34) into (B.33) and comparing the left hand side and right hand side coefficients of the value function we can uniquely solve π_1 as follows:

$$\pi_1 = \frac{\alpha}{1 - \beta(\alpha\theta + 1 - \theta)}$$

which after plugging into (B.34) we get,

$$h_{it+1} = (\alpha\beta\theta / (1 - \beta(1 - \theta)))^\theta (h_{it})^{\alpha\theta+1-\theta} (l_{h,i})^\theta (z_{it})^\theta \quad (\text{B.35})$$

Next solve $l_{h,i}$ by noting the fact that

$$l_{h,i} = \arg \max [b \ln(1 - l - l_{h,i}) + \beta\pi_1\theta \ln l_{h,i}]$$

which gives

$$l_{h,i} = \beta\pi_1\theta(1-l)/(b + \beta\pi_1\theta) = \frac{\beta\alpha\theta(1-l)}{b - b\beta(\alpha\theta + 1 - \theta) + \beta\alpha\theta} \quad (\text{B.36})$$

Time devoted to education is thus a constant confirming our conjecture. Note that $l_{h,i}$ is also increasing in θ .

B.1. Distributional dynamics

Based on (B.35) the social mobility equation is given by,

$$\ln h_{it+1} = \bar{p} + (\alpha\theta + 1 - \theta) \ln h_{it} + \theta(\ln z_{it}) \quad (\text{B.37})$$

where

$$\bar{p} \equiv \theta \ln \left(\frac{\beta\alpha\theta(1-l)}{b - b\beta(\alpha\theta + 1 - \theta) + \beta\alpha\theta} \right) + \theta \ln (\alpha\beta\theta / (1 - \beta(1 - \theta)))$$

Note that (B.37) takes the same form as (12) where the social mobility parameter is the same as ρ when $\delta = 1$. The dynamics of cross sectional variance of wealth based on (B.37) is given by,

$$\sigma_{t+1}^2 = ((\alpha - 1)\theta + 1)^2 \sigma_t^2 + \theta^2 v^2 \quad (\text{B.38})$$

which display similar properties as in our baseline "joy of giving" utility function. Higher adjustment cost (lower θ) and a higher human capital share parameter (α) slows down social mobility and raises the persistence of inequality as before.

3.1. Luck, initial inequality, and social mobility

Are children poor due to bad luck or poor parents? Both adjustment and depreciation costs play important roles in determining how inequality transmits across generations through past episodes of luck and initial conditions. To see this, do the same as (12) for the j th individual and subtract it from (12) to get

$$\Delta h_{it+1} = \varrho (\Delta h_{it}) + \chi (\Delta \varphi_{it}) \quad (24)$$

where $\Delta h_{it+1} \equiv \ln h_{it+1} - \ln h_{jt+1}$, $\Delta h_{it} \equiv \ln h_{it} - \ln h_{jt}$ and $\Delta \varphi_{it} \equiv \ln \varphi_{it} - \ln \varphi_{jt}$. Eq. (24) reflects the relative human capital evolution of the i th and j th households.

When the capital market is incomplete, the difference in initial human capital of the first generation as well as the difference in luck play a central role in transmitting initial inequality through generations. The first term in (24) shows the effect of a difference in the human capital of the parents while the second term captures the effect of the parents' luck on the wealth inequality of their children. The differences in both initial human capital and luck transmit through the generations. How they impact future generations depends on the parameters ϱ and χ , which in turn depend on the structural parameters. The initial difference in wealth has a decaying effect on the wealth difference of successive generations. The rate of decay is determined by $\varrho \in (0, 1)$. A larger ϱ makes the initial inequality have a persistent effect and thus slows down social mobility. It is easy to verify that ϱ is larger if the adjustment cost is higher (lower θ) or the depreciation cost (δ) is lower.

On the other hand, the parents' luck has an immediate impact on the difference in wealth of their children through the optimal investment function (9). This inequality transmits to the future generations through the adjustment of the human capital stock. If it is costly to adjust the human capital, the luck effect of the parents could persist over generations. To see this, we start from a steady-state where all agents are identical in terms of human capital and let the initial generations experience a difference in luck (say, the i th family enjoys some good luck and the j th family suffers from some bad luck). Figure 5 plots (24) the impulse responses of a 0.1 standard deviation difference in such luck on the time path of dynastic inequality for two values

of the adjustment cost parameters ($\theta = 0.8$ and $\theta = 0.6$). When $\theta = 0.8$ convergence occurs after 7 generations whereas it takes about 10 generations to converge when $\theta = 0.6$. An adult reacts 1.3 times more in response to luck when $\theta = 0.8$ as opposed to $\theta = 0.6$. This reinforces our earlier result that a higher adjustment cost slows down social mobility because of a sluggish response to luck by adults.

4. Redistribution policy, adjustment cost, and social mobility

Our baseline estimate of social mobility in Table 2 suggests that the social mobility is rather slow. Tamura et al. (2014) present a wide range of estimates of intergenerational elasticity of earnings based on the extant literature. In this section, we show that a redistribution policy could significantly expedite intergenerational mobility.

Assume that the human capital production is partly funded by the government. This is in fact more consistent with the empirical evidence that in most industrialized countries, education funding comprises public and private resources. For instance, in the United Kingdom, until recently, more than a third of universities' total funding came from the government.²³ Let the government levy a flat tax rate of τ on the total output and provide a g_t bundle of public services to households. Suppose that the use and efficiency of the public resource is not identical among households. Depending on its type, a given public service could benefit certain households more than proportionally. It could disproportionately benefit the poor due to their lack of access to its private substitutes or it may benefit the rich more because of their access to complementary inputs for human capital production. For example, Internet provision by the public sector may disproportionately benefit those individuals who own a laptop computer. On the other hand, a provision of public transport or free meals in school may benefit the poor more than proportionally because they lack these basic inputs.

²³BBC, June 27, 2011.

Based on these premises, the i th individual's human capital production function (4) is formulated as follows:

$$h_{it+1} = h_{it}^{1-\theta} \left(g_{it}^{(1-\xi)} s_{it}^{\xi} \right)^{\theta} \quad (25)$$

We assume here the complete depreciation of human capital, $\delta = 1$, without loss of generality.²⁴

The parameter $1 - \xi$ represents the government intensity in the education sector, which we call government bias in education following Basu and Bhattacharai (2012).²⁵ For example, if $\xi = 1$, the government plays no role in the education sector, which means that the government bias is zero in which case (25) reverts to (4). As in Getachew (2010), we assume that the public good technology takes the following form:

$$g_{it} = \frac{g_t}{h_{it}^{\kappa}} \quad (26)$$

The sign of the parameter κ features the redistributive nature of the public service. If $\kappa > 0$ ($\kappa < 0$), such a public good technology basically means that for a given g_t an individual with low (high) h_{it} receives greater public good. In other words, the government program is pro-poor (pro-rich), meaning that public funding benefits the poor more than the rich. The case $\kappa = 0$ implies a neutral public policy. In this case, the provision of a public service has equal effects on both the rich and the poor because all receive the same g_t . The size of κ characterizes the extent of the redistribution.

The government has a balanced budget, which means that the government imposes a flat rate (τ) tax on aggregate output to finance its public spending. In other words,

$$g_t = \tau y_t \quad (27)$$

²⁴Recall that depreciation has a minor quantitative effect on social mobility. Also, we let $a_1 = a_2 = 1$ for simplicity without any loss of generality.

²⁵Basu and Bhattacharai (2012) employ a slightly different human capital technology that involves schooling time, whereas in our present model there is no time allocation.

Given that we have a heterogenous agent economy where individuals differ in terms of productivity and human capital, we need to aggregate individual income levels to arrive at y_t . From (1), using the lognormal distributional property, it is easy to verify that

$$y_t = h_t e^{\frac{1}{2}\sigma_t^2(\alpha-1)\alpha} \quad (28)$$

Given that $0 < \alpha < 1$ and $\partial \ln y_t / \partial \sigma_t^2 < 0$, inequality and income negatively covary. Due to the diminishing returns to capital (1), the poor have a relatively higher marginal product of capital than the rich. Due to the missing credit market, they cannot borrow from the rich to undertake Pareto optimal investment. Therefore, a higher inequality leads to a greater inefficiency as it reflects missing productive opportunities.

Plugging (28) into (27) we get,

$$g_t = \tau h_t e^{\frac{1}{2}\sigma_t^2(\alpha-1)\alpha} \quad (29)$$

Since inequality reduces aggregate income (y_t), the public fund (g_t) goes down proportionally in a more unequal economy.

4.1. The optimal tax rate

We next solve the optimal tax rate for this economy. Households solve optimization problems in two steps. In the first step, they solve the optimal investment and consumption assuming the public spending is given exogenously. In the next step, they solve for their preferred tax rate.

Substituting (26) into (25), we obtain the individuals' human capital accumulation function that takes into account the disproportionate government funding of education,

$$h_{it+1} = h_{it}^{(1-\theta(1+(1-\xi)\kappa))} g_t^{\theta(1-\xi)} s_{it}^{\xi\theta} \quad (30)$$

Thus, in the first step, the i th household's optimization problem based on (2), (3), and (30) is given by,

$$\max_{s_{it}, \tau} \left\{ \ln (y_{it} (1 - \tau) - s_{it}) + \beta \ln \left(h_{it}^{(1-\theta)(1+(1-\xi)\kappa)} g_t^{\theta(1-\xi)} s_{it}^{\xi\theta} \right) \right\} \quad (31)$$

The first order condition is,

$$s_{it} = y_{it} \frac{\xi\theta\beta(1-\tau)}{1+\xi\theta\beta} \quad (32)$$

Because of the public funding of education, private investment in schooling decreases with government bias $1 - \xi$ and τ .

In the second step, the adult solves the preferred tax rate as follows:

$$\max_{\tau} \left\{ \ln (1 - \tau)^{\beta\xi\theta+1} + \ln \tau^{((1-\xi)\theta\beta)} \right\} \quad (33)$$

which yields the optimal tax rate,

$$\tau = \frac{\theta\beta(1-\xi)}{1+\theta\beta} \quad (34)$$

Given similar preferences, each individual chooses the same tax rate.²⁶ Not surprisingly τ is proportional to the "government bias" $(1 - \xi)$. On the other hand, the preferred tax rate is lower if the adjustment cost is high (i.e., θ is small). Since the government spending complements the human capital production and the higher adjustment cost is a tax imposed by mother nature on human capital production, all individuals desire that the government should tax them less for the provision of public goods to aid human capital.

From (1), (29), (30), (32), and (34), the optimal human capital accumulation under the public funding that is associated with the i th individual is given by,

$$h_{it+1} = \phi_1 (h_{it})^{\theta_1} (\varphi_{it})^{\chi_1} (h_t)^{\zeta_1} e^{\frac{1}{2}\sigma_t^2\zeta_2} \quad (35)$$

where

²⁶See also de la Croix and Michel (2002, p. 264) for a similar optimal tax result where all individuals choose the same tax rate.

$$\phi_1 \equiv \left(\frac{\xi\theta\beta(1-\tau)}{1+\xi\theta\beta} \right)^{\xi\theta} \tau^{\theta-\xi\theta} \quad (36)$$

and

$$\varrho_1 \equiv 1 - \theta(1 - \xi\alpha + \kappa(1 - \xi)) \quad (37)$$

$$\chi_1 \equiv \xi\theta \quad (38)$$

$$\zeta_1 \equiv \theta(1 - \xi) + (1 - \alpha)\xi\theta \quad (39)$$

$$\zeta_2 \equiv (\alpha - 1)\alpha\theta(1 - \xi) < 0 \quad (40)$$

where τ is given by (34).

4.2. Estimates of social mobility in the presence of a redistribution policy

The intergenerational elasticity of earnings is now modified to ϱ_1 in (37), which incorporates the redistributive policy (κ) and government bias ($1 - \xi$). According to Tamura et al. (2014), Table 3, ϱ_1 ranges from 0.158 to 0.535. However, Tamura et al. (2014) do not include the estimates of social mobility of Clark and Cummins (2012) and Mazumder (2005). If we include their estimates, the range of intergenerational earning elasticity widens to (0.158, 0.8). This wide zone of disagreement about social mobility could be due to alternative sampling designs, estimation methodology, and various other factors. However, in the context of our adjustment cost model with a redistribution policy, a reasonable question to ask is whether a redistribution policy could provide any additional insights into the observed differences in social mobility.

We use the Basu and Bhattacharai (2012) estimate of the mean government bias $1 - \xi$, which is equal to 0.07 based on 166 countries. We take this as a baseline estimate of government bias and fix $\theta = 0.8$ and $\alpha = 0.3$ as in our earlier calibration. Next, we compute a range for the redistributive parameter κ that could reproduce the ϱ_1 estimate in the range (0.158 to 0.8). Table 4 reports the results. For a value of κ equal to 1.5, we come close to the median estimate of intergenerational earning elasticity used by Tamura et al. (2014). As expected, a larger κ means greater social mobility. The value of κ ranges from 4.7 to -7 to reproduce the empirical range of

social mobility. In other words, our model predicts that with everything else being equal, greater social mobility may be due to a pro-poor redistribution policy.

Table 4: Effects of redistributive policy on social mobility

Redistributive Policy (κ)	ϱ_1
4.7	0.16
3	0.25
2	0.31
1.5	0.34
0	0.42
-1	0.47
-3	0.59
-5	0.7
-7	0.81

This numerical exercise has to be interpreted with caution. Given the stylized nature of the model, it may not account for all the factors that are important for social mobility. Notwithstanding this caveat, our model identifies a set of structural parameters that could provide some insights into mobility. The central focus of the model is the adjustment cost parameter in determining the degree of social mobility, which is novel in the literature. The model also explores the roles played by the depreciation rate and the redistribution policy parameters in reproducing an empirically plausible social mobility.

5. Conclusion

In this paper, we analyze the effects of the human capital adjustment cost on social mobility using a model with incomplete markets and heterogeneous agents. Although the adjustment cost is a standard feature in a physical capital production function, its role is not well explored in growth models with human capital. The adjustment cost of human capital is important because of the inherent inertia of

agents to respond to new opportunity, which is well documented in the literature. We use a more general functional form for human capital production that identifies the adjustment cost and the depreciation cost of human capital. Our human capital production function allows for the incomplete depreciation of human capital, which is ignored in the extant literature. A higher adjustment cost of human capital slows down social mobility by escalating the marginal cost of investment. Slower depreciation of human capital also contributes to sluggish mobility because it discourages new investment in schooling and through this channel it makes social mobility history dependent. The quantitative analysis of our model suggests that the human capital adjustment cost has nontrivial effects on social mobility and inequality. Our adjustment cost model, in conjunction with a pro-poor redistribution policy, gives rise to a calibrated social mobility on par with that shown in the literature. As mentioned previously, the key result that a higher human capital adjustment cost slows down social mobility is robust in a more general model environment. A future extension of this work would be to bring in endogenous occupational choice and explore the implications on the adjustment cost for occupational mobility.

Appendix

Outline of the derivation of (15) and (16)

This appendix provides an outline of the derivation of these two key equations. More technical details can be found in a supplementary appendix available from the authors upon request.

Rewrite (11) as

$$(h_{it+1})^\varsigma = \phi^\varsigma (h_{it}^\varsigma \kappa + \epsilon_t \varphi_{it} h_{it}^{\varkappa+\varsigma}) \quad (\text{A.1})$$

where $\varsigma \equiv 1/\theta$, $\varkappa \equiv \alpha - 1$, $\kappa \equiv 1 - \delta$, and $\epsilon_t \equiv a_1 h_t^{1-\alpha}$.

Given the lognormality assumption of φ_{it} and h_{it} ,

$$\ln \varphi_{it} \sim N(-v^2/2, v^2) \quad (\text{A.2})$$

$$\ln h_{it} \sim N(\mu_t, \sigma_t^2) \quad (\text{A.3})$$

for any constant x , note that

$$\ln (h_{it})^x \sim N(x\mu_t, x^2\sigma_t^2) \quad (\text{A.4})$$

This also implies,

$$\text{E}[(h_{it})^x] = (h_t)^x e^{0.5\sigma_t^2 x(x-1)} \quad (\text{A.5})$$

$$\text{var}[(h_{it})^x] = (h_t)^{2x} e^{\sigma_t^2 x(x-1)} (e^{x^2\sigma_t^2} - 1) \quad (\text{A.6})$$

Applying (A.5) and (A.6), aggregate (A.1) as follows:

$$h_{t+1}^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_{t+1}^2} = \phi^\varsigma h_t^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_t^2} \left\{ \kappa + a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} \right\} \quad (\text{A.7})$$

One can easily compute (16) now from (A.7).

To derive the distributional dynamics, take the variance from both sides of (A.1),

$$\text{var} [(h_{i,t+1})^\varsigma] = \phi^{2\varsigma} [\kappa^2 \text{var} [h_{it}^\varsigma] + \epsilon_t^2 \text{var} [\varphi_{it} h_{it}^{\varsigma+\varkappa}] + 2\kappa\epsilon_t \text{cov} (h_{it}^\varsigma, \varphi_{it} h_{it}^{\varsigma+\varkappa})] \quad (\text{A.8})$$

which, after further manipulations, and using (A.5), (A.6), and (A.7) yields:

$$e^{\theta^{-2}\sigma_{t+1}^2} = \frac{\kappa^2 e^{\theta^{-2}\sigma_t^2} + (a_1)^2 \left(e^{(\omega+\lambda^2)\sigma_t^2+v^2} \right) + 2\kappa a_1 \left(e^{(0.5\omega+\lambda/\theta)\sigma_t^2} \right)}{(\kappa + a_1 e^{0.5\omega\sigma_t^2})^2} \quad (\text{A.9})$$

where

$$\omega \equiv (\alpha - 1) (2/\theta + \alpha - 2) < 0, \quad \lambda \equiv 1/\theta + \alpha - 1 > 0$$

as given by (15).

The exact social mobility parameter, ϱ_t , is derived by simply taking the first derivative of (A.9):

$$\begin{aligned} \varrho_t^2 &\equiv \partial\sigma_{t+1}^2 / \partial\sigma_t^2 \\ &= \left(\frac{\kappa^2 \theta^{-2} \exp(\theta^{-2}\sigma_t^2) + (a_1)^2 b_1 b_2 \exp(b_2\sigma_t^2) + 2\kappa a_1 b_3 \exp(b_3\sigma_t^2)}{\kappa^2 \exp(\theta^{-2}\sigma_t^2) + (a_1)^2 b_1 \exp(b_2\sigma_t^2) + 2\kappa a_1 \exp(b_3\sigma_t^2)} - \frac{a_1 \omega \exp(0.5\omega\sigma_t^2)}{\kappa + a_1 \exp(0.5\omega\sigma_t^2)} \right) \theta^2 \end{aligned}$$

where

$$b_1 \equiv \exp(v^2), \quad b_2 \equiv \omega + \lambda^2, \quad b_3 \equiv 0.5\omega + \lambda/\theta.$$

Acknowledgements

We benefitted from the comments of the participants of the growth workshop at the University of St Andrews in 2011, particularly Charles Nolan. Thanks are also due to the participants in a seminar at Queens University, 2012, Belfast, Macro Money and Finance conference, 2012, Trinity College Dublin, and the Growth Conference in the Indian Statistical Institute in New Delhi. Elisa Keller is gratefully acknowledged for her insightful comments. The usual disclaimer applies.

References

- Alesina, A., Giuliano, P., 2010. The power of the family. *Journal of Economic Growth* 15 (2), 93–125.
- Arrow, K. J., 1962. The economic implications of learning by doing. *Review of Economic Studies* 29 (June), 155–173.
- Bandyopadhyay, D., Basu, P., 2005. What drives the cross-country growth and inequality correlation? *The Canadian Journal of Economics / Revue canadienne d'Economique* 38 (4), 1272–1297.
- Bandyopadhyay, D., Tang, X., 2011. Parental nurturing and adverse effects of redistribution. *Journal of Economic Growth* 16 (1), 71–98.
- Banerjee, A. V., Newman, A. F., 1993. Occupational choice and the process of development. *The Journal of Political Economy* 101 (2), 274–298.
- Barro, R. J., 1974. Are government bonds net wealth? *Journal of Political Economy* 82 (6), 1095–1117.
- Basu, P., 1987. An adjustment cost model of asset pricing. *International Economic Review* 28 (3), 609–621.
- Basu, P., Bhattarai, J., July 2012. Government Bias in Education, Schooling Attainment, and Long-Run Growth. *Southern Economic Journal* 79 (1), 127–143.

- Basu, P., Gillman, M., Pearlman, J., 2012. Inflation, human capital and tobin's q. *Journal of Economic Dynamics and Control* 36 (7), 1057–1074.
- Becker, G. S., Tomes, N., 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *The Journal of Political Economy* 87 (6), 1153–1189.
- Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75 (4), pp. 352–365.
- Benabou, R., 1996. *Inequality and Growth*. NBER Macroeconomics Annual, Volume 11, MIT Press.
- Benabou, R., 2000. Unequal societies: Income distribution and the social contract. *The American Economic Review* 90 (1), 96–129.
- Benabou, R., 2002. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica* 70 (2), 481–517.
- Clark, G., 2013. What is the True Rate of Social Mobility? Evidence from the Information Content of Surnames. Working Paper, <http://www.econ.ucdavis.edu/faculty/gclark/research.html>.
- Clark, G., Cummins, N., 2012. What is the True Rate of Social Mobility? Surnames and Social Mobility, England, 1800-2012. Working Paper, <http://www.econ.ucdavis.edu/faculty/gclark/research.html>.
- de Grip, A., 2006. *Evaluating Human Capital Obsolescence*. ROA Working Paper 001, Maastricht University, Research Centre for Education and the Labour Market (ROA).
- de Grip, A., van Loo, J., 2002. The economics of skill obsolescence: a review. In: de Grip, A., van Loo, J., Mayhew, K. (Eds.), *The Economics Of Skills Obsolescence (Research in Labor Economics)*. Vol. 38. Emerald Group Publishing Limited, pp. 1–26.

- de la Croix, D., Michel, P., 2002. *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*. Cambridge: Cambridge University Press.
- Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. *The Review of Economic Studies* 60 (1), 35–52.
- Getachew, Y. Y., 2010. Public capital and distributional dynamics in a two-sector growth model. *Journal of Macroeconomics* 32 (2), 606–616.
- Glomm, G., 1997. Parental choice of human capital investment. *Journal of Development Economics* 53 (1), 99–114.
- Glomm, G., Ravikumar, B., 1992. Public versus private investment in human capital: Endogenous growth and income inequality. *The Journal of Political Economy* 100 (4), 818–834.
- Grawe, N. D., Corak, M., 2004. *Intergenerational Mobility for Whom? The Experience of High- and Low-Earning Sons in International Perspective*. Cambridge; New York and Melbourne: Cambridge University Press, pp. 58–89.
- Hansen, L. P., Sargent, T. J., 1980. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control* 2 (0), 7–46.
- Katz, L. F., Kling, J. R., Liebman, J. B., 2001. Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics* 116 (2), 607–654.
- Loury, G. C., 1981. Intergenerational transfers and the distribution of earnings. *Econometrica* 49 (4), 843–867.
- Lucas, Robert E., J., Prescott, E. C., 1971. Investment under uncertainty. *Econometrica* 39 (5), 659–681.
- Machin, S., 2004. *Education Systems and Intergenerational Mobility*. CESifo PEPG Conference, Munich.

- Mankiw, N. G., Romer, D., Weil, D. N., 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107 (2), 407–437.
- Mazumder, B., 2005. Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *The Review of Economics and Statistics* 87 (2), 235–255.
- Mulligan, C. B., 1997. Parental priorities and economic inequality. University of Chicago Press, Chicago.
- Romer, P. M., 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94 (5), 1002–1037.
- Rosen, S., 1975. Measuring the obsolescence of knowledge. In: Juster, F. (Ed.), *Education, Income and Human Behavior*. New York: Carnegie Foundation for the Advancement of Teaching & National Bureau of Economic Research, pp. 199–232.
- Saint-Paul, G., Verdier, T., 1993. Education, democracy and growth. *Journal of Development Economics* 42, 399–407.
- Tamura, R., Dwyer, G. P., Devereux, J., Baier, S., 2014. Economic growth in the long run. Working paper.
- UNU-WIDER, 2007. World Income Inequality Database, Version 2.0b. UNU-WIDER.

Supplementary appendix

A. Aggregation and distribution dynamics

In this section we derive (15) from (11). We can also rewrite (11) as

$$(h_{it+1})^\varsigma = \phi^\varsigma (h_{it}^\varsigma \kappa + \epsilon_t \varphi_{it} h_{it}^{\varkappa+\varsigma}) \quad (\text{A.1})$$

where $\varsigma \equiv 1/\theta$, $\varkappa \equiv \alpha - 1$, $\kappa \equiv 1 - \delta$ and $\epsilon_t \equiv a_1 h_t^{1-\alpha}$.

Recall that first φ_{it} and h_{it} are assumed to have lognormal distributions:

$$\ln \varphi_{it} \sim N(-v^2/2, v^2) \quad (\text{A.2})$$

$$\ln h_{it} \sim N(\mu_t, \sigma_t^2) \quad (\text{A.3})$$

And, from a normal-lognormal relationship, we have:

$$\text{E}[h_{it}] \equiv h_t = e^{\mu_t + 0.5\sigma_t^2} \quad (\text{A.4})$$

$$\text{var}[h_{it}] = \left(e^{\sigma_t^2} - 1\right) e^{2\mu_t + \sigma_t^2} \quad (\text{A.5})$$

If h_{it} , then $(h_{it})^x$ is also lognormal for any constant x .

$$\ln (h_{it})^x \sim N(x\mu_t, x^2\sigma_t^2) \quad (\text{A.6})$$

Thus, considering (A.6), (A.4) and (A.5), we have:

$$\text{E}[(h_{it})^x] = (h_t)^x e^{0.5x^2\sigma_t^2} \quad (\text{A.7})$$

$$\text{var}[(h_{it})^x] = (h_t)^{2x} e^{x^2\sigma_t^2} \left(e^{x^2\sigma_t^2} - 1\right) \quad (\text{A.8})$$

We now apply (A.7) and (A.8) to derive the following important relations that we use later on:

$$\mathbf{E} [(h_{it+1})^\varsigma] = h_{t+1}^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_{t+1}^2} \quad (\text{A.9})$$

$$\mathbf{E} [h_{it}^\varsigma] = h_t^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_t^2} \quad (\text{A.10})$$

$$\mathbf{E} [h_{it}^{\varsigma+\varkappa}] = h_t^{\varsigma+\varkappa} e^{0.5(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \quad (\text{A.11})$$

$$\mathbf{E} [h_{it}^{2\varsigma+\varkappa}] = h_t^{2\varsigma+\varkappa} e^{0.5(2\varsigma+\varkappa)(2\varsigma+\varkappa-1)\sigma_t^2} \quad (\text{A.12})$$

$$\mathbf{E} [\varphi_{it}] = 1 \quad (\text{A.13})$$

$$\text{var} [h_{it+1}^\varsigma] = h_{t+1}^{2\varsigma} e^{\varsigma(\varsigma-1)\sigma_{t+1}^2} \left(e^{\varsigma^2\sigma_{t+1}^2} - 1 \right) \quad (\text{A.14})$$

$$\text{var} [\varphi_{it}] = \left(e^{v^2} - 1 \right) \quad (\text{A.15})$$

$$\text{var} [h_{it}^\varsigma] = h_t^{2\varsigma} e^{\varsigma(\varsigma-1)\sigma_t^2} \left(e^{\varsigma^2\sigma_t^2} - 1 \right) \quad (\text{A.16})$$

$$\text{var} [h_{it}^{\varsigma+\varkappa}] = h_t^{2(\varsigma+\varkappa)} e^{(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \left(e^{(\varsigma+\varkappa)^2\sigma_t^2} - 1 \right) \quad (\text{A.17})$$

Then, aggregate (A.1) from both side to derive the aggregate human capital:

$$\mathbf{E} [(h_{it+1})^\varsigma] = \phi^{1/\theta} \mathbf{E} [h_{it}^\varsigma \kappa + \epsilon_t \varphi_{it} h_{it}^{\varsigma+\varkappa}] = \phi^{1/\theta} \left\{ \kappa \mathbf{E} [h_{it}^\varsigma] + \epsilon_t \mathbf{E} [h_{it}^{\varsigma+\varkappa}] \right\}$$

Substituting (A.9), (A.10) and (A.11) into the above,

$$\begin{aligned} h_{t+1}^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_{t+1}^2} &= \phi^\varsigma \left\{ \kappa h_t^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_t^2} + \epsilon_t h_t^{\varsigma+\varkappa} e^{0.5(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \right\} \\ &= \phi^\varsigma h_t^\varsigma \left\{ \kappa e^{0.5\varsigma(\varsigma-1)\sigma_t^2} + a_1 e^{0.5[\varsigma(\varsigma-1)+\varsigma\varkappa+\varkappa(\varsigma+\varkappa-1)]\sigma_t^2} \right\} \end{aligned}$$

Thus, the aggregate human capital is given by:

$$h_{t+1}^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_{t+1}^2} = \phi^\varsigma h_t^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_t^2} \left\{ \kappa + a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} \right\} \quad (\text{A.18})$$

To derive the distributional dynamics, take the variance from both sides of (A.1),

$$\begin{aligned}\text{var} [(h_{i,t+1})^\varsigma] &= \phi^{2\varsigma} \text{var} [h_{it}^\varsigma \kappa + \epsilon_t \varphi_{it} h_{it}^{\varsigma+\varkappa}] \\ &= \phi^{2\varsigma} [\kappa^2 \text{var} [h_{it}^\varsigma] + \epsilon_t^2 \text{var} [\varphi_{it} h_{it}^{\varsigma+\varkappa}] + 2\kappa\epsilon_t \text{cov} (h_{it}^\varsigma, \varphi_{it} h_{it}^{\varsigma+\varkappa})] \quad (\text{A.19})\end{aligned}$$

Using (A.10), (A.11), (A.12) and (A.13), the cov term is computed as follows:

$$\begin{aligned}\text{cov} (h_{it}^\varsigma, \varphi_{it} h_{it}^{\varsigma+\varkappa}) &= \text{E} [h_{it}^\varsigma \varphi_{it} h_{it}^{\varsigma+\varkappa}] - \text{E} [h_{it}^\varsigma] \text{E} [\varphi_{it} h_{it}^{\varsigma+\varkappa}] \\ &= \text{E} [h_{it}^{2\varsigma+\varkappa}] - \text{E} [h_{it}^\varsigma] \text{E} [h_{it}^{\varsigma+\varkappa}] \\ &= h_t^{2\varsigma+\varkappa} e^{0.5(2\varsigma+\varkappa)(2\varsigma+\varkappa-1)\sigma_t^2} - h_t^\varsigma e^{0.5\varsigma(\varsigma-1)\sigma_t^2} h_t^{\varsigma+\varkappa} e^{0.5(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \\ &= h_t^{2\varsigma+\varkappa} e^{0.5(\varkappa(2\varsigma+\varkappa-1)+2\varsigma(\varsigma-1))\sigma_t^2} \left(e^{\varsigma(\varsigma+\varkappa)\sigma_t^2} - 1 \right) \quad (\text{A.20})\end{aligned}$$

The second term in the right hand side of (A.19) is computed as,¹

$$\begin{aligned}\text{var} [\varphi_{it} h_{it}^{\varsigma+\varkappa}] &= (\text{E} [\varphi_{it}])^2 \text{var} [h_{it}^{\varsigma+\varkappa}] + \text{var} [\varphi_{it}] \left((\text{E} [h_{it}^{\varsigma+\varkappa}])^2 + \text{var} [h_{it}^{\varsigma+\varkappa}] \right) \\ &= \text{var} [h_{it}^{\varsigma+\varkappa}] (1 + \text{var} [\varphi_{it}]) + \text{var} [\varphi_{it}] (\text{E} [h_{it}^{\varsigma+\varkappa}])^2\end{aligned}$$

since $\text{E} [\varphi_{it}] = 1$.

Substituting (A.11), (A.15) and (A.17) into the above yields:

¹If x and y are independent, the variance of their product is:

$$\text{var} [xy] = (\text{E} [x])^2 \text{var} [y] + (\text{E} [y])^2 \text{var} [x] + \text{var} [y] \text{var} [x]$$

$$\begin{aligned}
\text{var} [\varphi_{it} h_{it}^{\varsigma+\varkappa}] &= h_t^{2(\varsigma+\varkappa)} e^{(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \left(e^{(\varsigma+\varkappa)^2\sigma_t^2} - 1 \right) \left(1 + \left(e^{v^2} - 1 \right) \right) \\
&\quad + \left(e^{v^2} - 1 \right) h_t^{2(\varsigma+\varkappa)} e^{(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \\
&= h_t^{2(\varsigma+\varkappa)} e^{(\varsigma+\varkappa)(\varsigma+\varkappa-1)\sigma_t^2} \left(\left(e^{(\varsigma+\varkappa)^2\sigma_t^2} - 1 \right) e^{v^2} + \left(e^{v^2} - 1 \right) \right) \\
&= h_t^{2(\varsigma+\varkappa)} e^{\varkappa(2\varsigma+\varkappa-1)+\varsigma(\varsigma-1)\sigma_t^2} \left(e^{(\varsigma+\varkappa)^2\sigma_t^2} e^{v^2} - 1 \right) \tag{A.21}
\end{aligned}$$

Then, substituting, (A.14), (A.16), (A.20) and (A.21) into (A.19) yields:

$$\begin{aligned}
&h_{t+1}^{2\varsigma} e^{\varsigma(\varsigma-1)\sigma_{t+1}^2} \left(e^{\varsigma^2\sigma_{t+1}^2} - 1 \right) \\
&= \phi^{2\varsigma} \left[\begin{aligned} &h_t^{2\varsigma\kappa^2} e^{\varsigma(\varsigma-1)\sigma_t^2} \left(e^{\varsigma^2\sigma_t^2} - 1 \right) \\ &+ \epsilon_t^2 \left\{ h_t^{2(\varsigma+\varkappa)} e^{[\varkappa(2\varsigma+\varkappa-1)+\varsigma(\varsigma-1)]\sigma_t^2} \left(e^{(\varsigma+\varkappa)^2\sigma_t^2} e^{v^2} - 1 \right) \right\} \\ &+ 2\kappa\epsilon_t \left\{ h_t^{2\varsigma+\varkappa} e^{0.5[\varkappa(2\varsigma+\varkappa-1)+2\varsigma(\varsigma-1)]\sigma_t^2} \left(e^{\varsigma(\varsigma+\varkappa)\sigma_t^2} - 1 \right) \right\} \end{aligned} \right] \tag{A.22}
\end{aligned}$$

Finally, substituting (A.18) into the above, we get :

$$\begin{aligned}
&\phi^{2\varsigma} h_t^{2\varsigma} e^{\varsigma(\varsigma-1)\sigma_t^2} \left\{ \kappa + a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} \right\}^2 \left(e^{\varsigma^2\sigma_{t+1}^2} - 1 \right) \\
&= \phi^{2\varsigma} h_t^{2\varsigma} \left[\begin{aligned} &\kappa^2 e^{\varsigma(\varsigma-1)\sigma_t^2} \left(e^{\varsigma^2\sigma_t^2} - 1 \right) \\ &+ (a_1)^2 \left\{ e^{[\varkappa(2\varsigma+\varkappa-1)+\varsigma(\varsigma-1)]\sigma_t^2} \left(e^{(\varsigma+\varkappa)^2\sigma_t^2} e^{v^2} - 1 \right) \right\} \\ &+ 2\kappa a_1 \left\{ e^{0.5[\varkappa(2\varsigma+\varkappa-1)+2\varsigma(\varsigma-1)]\sigma_t^2} \left(e^{\varsigma\varkappa\sigma_t^2} - 1 \right) \right\} \end{aligned} \right] \tag{A.23}
\end{aligned}$$

since $\epsilon_t \equiv a_1 h_t^{-\varkappa}$.

Considering,

$$\left(\kappa + a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} \right)^2 = \kappa^2 + 2\kappa a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} + (a_1)^2 e^{\varkappa(2\varsigma+\varkappa-1)\sigma_t^2}$$

further simplifying (A.23) gives

$$e^{\varsigma^2 \sigma_{t+1}^2} = \frac{\kappa^2 e^{\varsigma^2 \sigma_t^2} + (a_1)^2 \left(e^{\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} e^{(\varsigma+\varkappa)^2 \sigma_t^2} e^{v^2} \right) + 2\kappa a_1 \left(e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2} e^{\varsigma(\varsigma+\varkappa)\sigma_t^2} \right)}{(\kappa + a_1 e^{0.5\varkappa(2\varsigma+\varkappa-1)\sigma_t^2})^2}$$

Alternatively,

$$e^{\theta^{-2} \sigma_{t+1}^2} = \frac{\kappa^2 e^{\theta^{-2} \sigma_t^2} + (a_1)^2 \left(e^{[(\alpha-1)(2/\theta+\alpha-2)+(1/\theta+\alpha-1)^2]\sigma_t^2 + v^2} \right) + 2\kappa a_1 \left(e^{[0.5(\alpha-1)(2/\theta+\alpha-2)+(1/\theta+\alpha-1)/\theta]\sigma_t^2} \right)}{(\kappa + a_1 e^{0.5(\alpha-1)(2/\theta+\alpha-2)\sigma_t^2})^2}$$

after substituting $\varsigma \equiv 1/\theta$, $\varkappa \equiv \alpha - 1$. Or,

$$e^{\theta^{-2} \sigma_{t+1}^2} = \frac{\kappa^2 e^{\theta^{-2} \sigma_t^2} + (a_1)^2 \left(e^{(\omega+\lambda^2)\sigma_t^2 + v^2} \right) + 2\kappa a_1 \left(e^{(0.5\omega+\lambda/\theta)\sigma_t^2} \right)}{(\kappa + a_1 e^{0.5\omega\sigma_t^2})^2} \quad (\text{A.24})$$

where

$$\omega \equiv (\alpha - 1)(2/\theta + \alpha - 2) < 0, \quad \lambda \equiv 1/\theta + \alpha - 1 > 0$$

as given by (15).

A.1. Social mobility

The exact social mobility parameter, ϱ_t , is derived by simply taking the first derivative of (A.24):

$$\begin{aligned} \varrho_t^2 &\equiv \partial \sigma_{t+1}^2 / \partial \sigma_t^2 \\ &= \left(\frac{\kappa^2 \theta^{-2} \exp(\theta^{-2} \sigma_t^2) + (a_1)^2 b_1 b_2 \exp(b_2 \sigma_t^2) + 2\kappa a_1 b_3 \exp(b_3 \sigma_t^2)}{\kappa^2 \exp(\theta^{-2} \sigma_t^2) + (a_1)^2 b_1 \exp(b_2 \sigma_t^2) + 2\kappa a_1 \exp(b_3 \sigma_t^2)} - \frac{a_1 \omega \exp(0.5\omega \sigma_t^2)}{\kappa + a_1 \exp(0.5\omega \sigma_t^2)} \right) \theta^2 \end{aligned}$$

where

$$b_1 \equiv \exp(v^2), \quad b_2 \equiv \omega + \lambda^2, \quad b_3 \equiv 0.5\omega + \lambda/\theta$$

B. A model of dynastic altruism with labour and capital

In this appendix, we show that the key result that a higher adjustment cost slows down social mobility continues to hold in a model with dynastic altruism as in Barro (1974) with labour and physical capital. Each generation lives one period and discounts the future generation's utility by β . The i th agent born at date t has the utility function:

$$v_{it} = \ln c_{it} + b \ln(1 - l_{g,it} - l_{h,it}) + \beta E_t v_{it+1} \quad (\text{B.25})$$

where $l_{g,it}$ is labour time spend on good production, $l_{h,it}$ is labour time spent on child's education and $b \in (0, 1)$ is the relative importance of leisure in utility. The total time is normalized at unity. We assume for simplicity that the technology of goods production requires a fixed amount of raw labour time (say 8 hours a day) and the adult has no choice to allocate more or less time to it. Thus fix $l_{g,it} = l$. The remaining time can be allocated freely between leisure and child's education. The goods production function is thus:

$$x_{it} = \vartheta_{it} (l h_{it})^\varpi (k_{it})^\eta (h_t)^\varepsilon \quad (\text{B.26})$$

where ϑ_{it} is idiosyncratic shock; k_{it} is physical capital and the production function obeys constant returns to scale property meaning $\varpi + \eta + \varepsilon = 1$.

The human capital production function is specified as (assuming $\delta = 1$).

$$h_{it+1} = h_{it}^{1-\theta} (s_{it} l_{h,it})^\theta \quad (\text{B.27})$$

The effective investment in human capital is the raw labour ($l_{h,it}$) spent on children times resources (s_{it}) spent on schooling.

Although parents cannot borrow against their offspring's human capital because of the immutable moral hazard and adverse selection issues, they have an access to an international credit market to finance their purchase of physical capital, k_{it} at a fixed interest rate r^* . The adult fully pays off the loan with interest before the end of their life. All physical capital is used up in the production process and nothing is

left for the upcoming generation. The optimal purchase of capital is thus given by the equality between the marginal product of physical capital and the user cost of capital, which means,

$$\partial x_{it} / \partial k_{it} = r^* + \delta_k \quad (\text{B.28})$$

where δ_k is the rate of depreciation of physical capital.² The adult's value added (y_{it}) after paying off the loan servicing and the user costs of capital is given by,

$$y_{it} = x_{it} - (r^* + \delta_k)k_{it} \quad (\text{B.29})$$

Substituting out k_{it} using (B.28) and (B.26), equation (B.29) can be rewritten as:

$$y_{it} = f \varphi_{it} (h_{it})^\alpha (h_t)^{1-\alpha}$$

where $\varphi_{it} \equiv (\vartheta_{it})^{1/(1-\eta)}$, $f \equiv l^{\varpi/(1-\eta)} (\eta / (r^* + \delta_k))^{\eta/(1-\eta)} (1-\eta)$ and $\alpha \equiv \varpi / (1-\eta)$.

The i th adult is thus subject to the budget constraint:

$$c_{it} + s_{it} = y_{it} \quad (\text{B.30})$$

and she maximizes (B.25), subject to (B.27) and (B.30), taking φ_{it} , l and h_t as given.

The value function for this problem can be written as:

$$v(h_{it}, z_{it}) = \max_{h_{it+1}} \left[\ln \left\{ y_{it} - \frac{(h_{it+1})^{1/\theta}}{l_{h,it} (h_{it})^{(1-\theta)/\theta}} \right\} + b \ln(1 - l - l_{h,it}) + \beta \text{E}t v(h_{it+1}, z_{it+1}) \right] \quad (\text{B.31})$$

The proof mimics Basu (1987). Conjecture that the value function is loglinear in

²To see this arbitrage condition, note that the adult's choice of physical capital solves the static maximization problem:

$$\max_{k_{it}} [x_{it} + (1 - \delta_k)k_{it} - (1 + r^*)k_{it}]$$

state variables as follows:

$$v(h_{it}, z_{it}) = \pi_0 + \pi_1 \ln h_{it} + \pi_2 \ln z_{it} \quad (\text{B.32})$$

where $z_{it} \equiv f\varphi_{it}h_t^{1-\alpha}$. Plugging (B.32) into the value function (B.31),

$$\begin{aligned} & \pi_0 + \pi_1 \ln h_{it} + \pi_2 \ln z_{it} \quad (\text{B.33}) \\ & = \max_{h_{it+1}} \left[\ln \left\{ y_{it} - \frac{(h_{it+1})^{1/\theta}}{l_{h,it} (h_{it})^{(1-\theta)/\theta}} \right\} + b \ln(1 - l - l_{h,it}) + \beta \{ \pi_0 + \pi_1 \ln h_{it+1} + \ln z_{it+1} \} \right] \end{aligned}$$

We will use the method of undetermined coefficients to solve for π_i which only matters for determining the decision rule of investment. Conjecture that $l_{h,it}$ is time invariant and is equal to $l_{h,i}$.

Differentiating with respect to h_{it+1} and rearranging terms one gets:

$$h_{it+1} = (\pi_1 \beta \theta / (1 + \pi_1 \beta \theta))^\theta (h_{it})^{\alpha\theta+1-\theta} (l_{h,it})^\theta (z_{it})^\theta \quad (\text{B.34})$$

Plugging (B.34) into (B.33) and comparing the left hand side and right hand side coefficients of the value function we can uniquely solve π_1 as follows:

$$\pi_1 = \frac{\alpha}{1 - \beta(\alpha\theta + 1 - \theta)}$$

which after plugging into (B.34) we get,

$$h_{it+1} = (\alpha\beta\theta / (1 - \beta(1 - \theta)))^\theta (h_{it})^{\alpha\theta+1-\theta} (l_{h,i})^\theta (z_{it})^\theta \quad (\text{B.35})$$

Next solve $l_{h,i}$ by noting the fact that

$$l_{h,i} = \arg \max [b \ln(1 - l - l_{h,i}) + \beta\pi_1\theta \ln l_{h,i}]$$

which gives

$$l_{h,i} = \beta\pi_1\theta(1-l)/(b + \beta\pi_1\theta) = \frac{\beta\alpha\theta(1-l)}{b - b\beta(\alpha\theta + 1 - \theta) + \beta\alpha\theta} \quad (\text{B.36})$$

Time devoted to education is thus a constant confirming our conjecture. Note that $l_{h,i}$ is also increasing in θ .

B.1. Distributional dynamics

Based on (B.35) the social mobility equation is given by,

$$\ln h_{it+1} = \bar{p} + (\alpha\theta + 1 - \theta) \ln h_{it} + \theta(\ln z_{it}) \quad (\text{B.37})$$

where

$$\bar{p} \equiv \theta \ln \left(\frac{\beta\alpha\theta(1-l)}{b - b\beta(\alpha\theta + 1 - \theta) + \beta\alpha\theta} \right) + \theta \ln (\alpha\beta\theta / (1 - \beta(1 - \theta)))$$

Note that (B.37) takes the same form as (12) where the social mobility parameter is the same as ρ when $\delta = 1$. The dynamics of cross sectional variance of wealth based on (B.37) is given by,

$$\sigma_{t+1}^2 = ((\alpha - 1)\theta + 1)^2 \sigma_t^2 + \theta^2 v^2 \quad (\text{B.38})$$

which display similar properties as in our baseline "joy of giving" utility function. Higher adjustment cost (lower θ) and a higher human capital share parameter (α) slows down social mobility and raises the persistence of inequality as before.