This is the author's final, peer-reviewed manuscript as accepted for publication. The publisher-formatted version may be available through the publisher's web site or your institution's library.

Government education expenditures in early and late childhood

Casey Abington, William Blankenau

How to cite this manuscript

If you make reference to this version of the manuscript, use the following information:

Abington, C., & Blankenau, W. (2013). Government education expenditures in early and late childhood. Retrieved from http://krex.ksu.edu

Published Version Information

Government Education Expenditures in Early and Late Childhood[*]

Casey Abington[†]
Northwest Missouri State University

William Blankenau[‡]
Kansas State University

June 5, 2012

Abstract

Human capital investment in early childhood can lead to large and persistent gains. Beyond this window of opportunity, human capital accumulation is more costly. Despite compelling evidence in support of this notion, government education spending is allocated disproportionately toward late childhood and young adulthood. We consider the consequences of a reallocation using an overlapping generations model with private and public spending on early and late childhood education. Taking as given the higher returns to early childhood investment, we find that the current allocation may nonetheless be appropriate. When we consider a homogeneous population, this can hold for moderate levels of government spending. With heterogeneity, this can hold for middle income workers. Lower income workers, by contrast, may benefit from a reallocation.

JEL Classification: E62, I22, H52, J24
Keywords: Government education expenditures, human capital, heterogeneous agents, life-cycle model

[†]Department of Accounting, Economics, and Finance, 2147 Colden Hall, Northwest Missouri State University, Maryville, MO 64468

[‡]Corresponding author: Department of Economics, 327 Waters Hall, Kansas State University, Manhattan, KS 66506, (785) 532-6340, fax:(785) 532-6919, email: blankenw@k-state.edu.

1 Introduction.

Research by education specialists, psychologists, and economists is bringing into sharp focus a fundamental feature of human capital accumulation. Human capital investment in early childhood can lead to large and persistent gains while investment beyond this window of opportunity yields diminished returns. Recent work by Cunha, Heckman, Lochner, and Masterov (CHLM, 2007) provides a comprehensive overview of work in the field.[1]

One conclusion of their overview is that the process of human capital accumulation is best modeled as a hierarchical process wherein early childhood education sets the stage for productive education in late childhood. Skills attained early in life leave a learner better prepared to take advantage of later opportunities to develop more refined skills. Similarly, late childhood investment reinforces investment in early childhood. Without follow-up investment, early investment is unproductive over the longer term.

This complementarity is often neglected when economists model human capital accumulation. While it is becoming more common to think about a hierarchical education process, this is typically to distinguish between K-12 and college education.[2] CHLM argue that the more meaningful distinction is between human capital investment during and after "critical" periods for the acquisition of particular skills. Perhaps the most straightforward example is the critical period for developing IQ. By age 10, the IQ of a child is essentially set. Before that time it is more malleable.[3] Low investment in the first 10 years leaves IQ lower and later investment less productive. At the same time, low investment later in life fails to exploit the potential to turn IQ into specific life skills.

Since government is a ubiquitous presence in funding human capital production, the nature of the process might suggest that government should allocate resources disproportionately toward early childhood education. Presently, it does not. In 2004, about .3% of GDP was spent by government on pre-primary education in educational institutions for students aged 3-6 while 4% of GDP was spent by government on K-12 education.[4] With the duration of K-12 around six times that of pre-primary education, this suggests that on a per capita basis government spending on

---

[1] See also Carneiro and Heckman (2003), Knudsen, Heckman, Cameron, and Shonkoff (2006), and Currie (2001(b)).

[2] See, for example, Driskill and Horowitz (2002), Su (2004), Kaganovich (2005), Blankenau (2005), and Restuccia and Urrutia (2004).

[3] See Jensen (1980) and the discussions in CHLM and Cunha and Heckman (2007).

[4] These numbers are not reported directly. However, table B2.2 of OECD Education at a Glance (2007) states that .4% of GPD is spent in total pre-primary education and 4.4% is spent on K-12 education. Table B3.2a of this same publication indicates that about 75% of pre-primary and 90% of K-12 funding is provided by government.

K-12 education is more than 2.2 times that on pre-primary education. On a per student basis, the difference is less pronounced as pre-primary enrollment is lower. Still, pre-primary per student expenditures are only 63% as large as upper secondary expenditures. Within K-12 education, spending is again weighted toward the later years. Per student spending on primary education is about 84% of upper secondary spending.[5]

Human capital spending is more than just education spending. In addition, government affects spending beyond its direct payments. A fuller analysis of relative spending levels would consider health care expenditures, tax breaks for daycare, after school programs, and a variety of related issues. While a complete accounting is a useful endeavor for later work, the conclusion that government does not spend disproportionately on human capital in early childhood is likely robust to any fuller analysis.

With spending concentrated in later years and development opportunities arising early, the allocation of government spending may have important implications. This paper considers the general equilibrium effects of allocating government expenditures across early and late childhood. We build a heterogeneous agent overlapping generations model where general human capital is generated in a two-stage hierarchical education system. The first period generates early human capital. An agent's endowment of early human capital depends on an exogenous family effect, first stage family spending, and first stage government spending. The second stage generates general human capital as a function of early human capital and second stage spending by the family and government.

Families value consumption and the lifetime income of their offspring. They allocate income across consumption spending and education spending at the two stages. Government interacts with households through taxation and provision of education inputs at each stage of childhood. The provision of education inputs has two consequences. There is a direct effect as inputs increase but also general equilibrium effects as private education spending adjusts in response. Two questions dominate the analysis. Is it best for government to concentrate its spending on one stage of education or to balance expenditures across the two stages? Secondly, if more concentrated expenditures are best, which level should be the focus of government expenditures?

The intuition is most clear when family and government inputs are perfectly substitutable, so

---

[5] Table B1.1a of Education at a Glance (2007) provides expenditures per student for pre-primary education, primary education, lower secondary education, and upper secondary education. The figures are arrived at by taking the ratio and, in the case of pre-primary, weighting it by the relative shares funded by government found in table B3.2a.

we focus on this case. Roughly speaking, a family prefers balanced spending only if government spending is high relative to personal income. When overall spending as a share of income is in a high intermediate range, an agent's income is maximized with government spending concentrated on early childhood education. When overall spending is in a low intermediate range, an agent's income is maximized with government spending concentrated on late childhood. Below some threshold level, the allocation of spending is irrelevant.

Results stem from the nature of human capital development and the crowding out of private spending by public spending. At high levels of spending relative to income, private spending is fully crowded out so the level of expenditure is dictated by government choices. In this case, productivity of public expenditures is key to output. Productivity is highest with a more balanced allocation. Since public spending is high in relation to the lowest incomes in the economy, this suggests that low income families are better off with more balanced government expenditures. At low levels of government spending relative to income, public spending simply displaces private spending, leaving total spending at each stage unchanged. This suggests that high income families may be unaffected by the allocation.

Between the relative extremes is the case where one type of spending is fully (or largely) crowded out and the other is not. When the allocation favors late childhood education, family spending at this stage is fully crowded out by government spending and family spending remains positive at the early childhood stage. This matches the situation in the U.S. where more than 90% of K-12 education spending is provided by government. With a disproportionate level of private K-12 spending by higher income families, this implies that some share of the population spends little or nothing privately on K-12 education. For these agents, an allocation toward early childhood education by government crowds out some early childhood spending by families. Since later spending is zero, there is no offsetting 'crowding in' of later spending. While the mix of spending may be more productive, total education spending decreases. This effect can dominate, leading to lower output. Hence, concentrated spending can maximize the income of middle income families.

After establishing that middle income families might prefer concentrated expenditure, we show that the preferred stage of concentration depends on family income. While the lower income workers in this group would prefer government spending concentrated on early childhood education, the rest prefer a focus on late childhood education. In essence, the larger of the expenditures (public or private) should be allocated to the most productive stage. For some middle income workers,

public spending exceeds private. It is best to allocate this to early childhood. For the more wealthy, the opposite holds. All told, the current concentration of government spending on late childhood education can be optimal for some income levels. At other income levels it may not be optimal but still preferred to more balanced spending. With the most wealthy indifferent, this leaves only the most poor to benefit from a reallocation.

We present the model in Section 2 and consider a special case in Section 3. Here agents are homogeneous and private and public spending are perfect substitutes. Much of the intuition is captured by this special case. The final part of Section 3 demonstrates this point by showing that the results are little changed in a more general case preserving homogeneity. Section 4 considers heterogeneity. Section 5 summarizes, provides some more speculative insights on policy implications, and concludes the paper.

2 The model.

2.1 The technology of education.

We consider an overlapping generations economy where agents live four periods. In each period, a mass of new agents, normalized to one, enters the economy and passes through early childhood. In the subsequent period, these agents are in late childhood. Throughout childhood, agents are passive economic agents. They receive endowments of human capital in each period but make no decisions of their own. Agents enter early adulthood in their third period. This is an active period where agents allocate income as specified below.[6] In addition, young adults each have one child. Thus the young adults in period t are parents to the new agents in that period. The fourth period of life is late adulthood where agents face a separate allocation decision and are parents to the late childhood generation.

The agents born in each period may be heterogeneous and are indexed by $j \in \mathcal{J} \equiv [0,1]$. A productivity parameter is related to the index through the function $a_j = a(j)$ where a_j is the productivity of agent j and $0 < a_j \leq a_{j'} < \infty$ for all $j < j'$. If the middle inequality is strict for at least one j, j' pair, there is heterogeneity in productivity. Though not modeled, we assume that through nature and nurture a child inherits the productivity of her parents. While this overstates the intergenerational persistence of productivity, recent evidence suggests considerable dynastic

---

[6] In an earlier version of this paper, young adults also made a choice to attend college or not. This proves unimportant for our main points.

4

persistence in relative earnings. For example, Mazumder (2005) estimates the intergenerational elasticity in earnings to be about .6.[7] In our model, inheritance of a is the root cause of such persistence.

Agent j in early childhood is endowed with $h_{1j(t)}$ units of early childhood human capital, which indicates that the endowment is time and agent specific. We hereafter compromise on precision in favor of aesthetics by suppressing the j and t notation when no confusion arises. The endowment is a function of ability and resources invested on behalf of the agent in her first period, i_1. In late childhood, the agent is endowed with general human capital. The size of this endowment depends on ability, early childhood human capital, and resources invested on behalf of the agent in her second period, i_2. Specifically,

$$
\begin{aligned}
h_1 &= a i_1^{\gamma_1} \\
h_2 &= Aa \left[\gamma_2 i_2^\rho + (1 - \gamma_2) h_1^\rho \right]^{\frac{1}{\rho}}
\end{aligned}
\tag{1}
$$

where $\gamma_1, \gamma_2 \in (0,1)$, $\rho \le 1$, and $A > 0$ are common across agents and fixed through time while other items are agent and time specific. The parameter A serves as a scalar in the production of human capital while γ_1 and γ_2 govern the curvature of the functions. The parameter ρ governs the substitutability of early childhood investment and late childhood investment in creating human capital. This specification is similar to Cunha and Heckman (2007).

Education investments, i_1 and i_2, depend on spending by parents and government. We expect that spending by government and families are largely substitutable as inputs into the production of human capital. For example, the productivity of otherwise identical books and teachers does not differ according to the means of finance, and students may learn as much from school field trips as from family outings. On the other hand, parents may provide some inputs that do not substitute well for government inputs. For example, a family may live in a more costly neighborhood in order to gain educational or peer-effect advantages for the child. To accommodate possible imperfect substitutability, we specify

$$
i_k = B \left[\alpha f_k^\eta + (1 - \alpha) g_k^\eta \right]^{\frac{1}{\eta}}
\tag{2}
$$

for $k \in \{1, 2\}$ where f_1 and g_1 are family and government resources devoted to early childhood education while f_2 and g_2 are resources devoted to late childhood education. The specification

---

[7] In the U.S., recent estimates are .4 or greater. See, for example, Solon (1999). Solon (2002) provides a review of elasticity estimates across nations.

requires $\eta \leq 1$. With $\eta = 0$, this is the Cobb-Douglas specification used (for example) by Blankenau (2005), and with $\eta = 1$, this is the specification used by Glomm and Kaganovich (2003).

2.2 The agent's problem.

Each agent is endowed with one unit of time in each period. Agents receive an income of wh_2 in each period of adulthood.[8] Here w is the wage per unit of human capital. Agents in our model are not borrowing constrained. In contrast, borrowing constraints play a key role in a wide variety of recent research. Some examples are Rangazas (2002) and Restuccia and Urrutia (2004). We exclude such considerations for two reasons. First, we show below that for low income agents most or all education expenditures are made by government. Thus low income agents, for whom constraints are most likely to bind, are not interested in borrowing. Secondly, recent work by Carneiro and Heckman (2002) indicates that few families are credit constrained in making education decisions later in life. It would be reasonable to impose credit constraints for those who spend significantly on children in early childhood but this is likely to be of modest importance. With no borrowing constraints, we can ignore the timing of income and focus on the present value of lifetime income. With an interest rate exogenously given by r, this is

$$I = wh_2 \left(1 + \frac{1}{r} \right). \tag{3}$$

We will use $\,\hat{}\,$ notation to indicate items that relate to the children of the generation being considered. For example, while I is the income of the generation being considered, $\hat{I}$ is the income of the offspring. Each agent has preferences given by

$$U_j = \ln c_3 + \beta \ln c_4 + \xi \ln \hat{I}. \tag{4}$$

Here c_3 and c_4 denote consumption in the third and fourth periods of life, and $\beta < 1$ discounts the future. Aside from own consumption, the agent cares about the lifetime income of her children where the term ξ scales the importance of progeny income. Parents can effect progeny income through spending on human capital in the first and second periods of childhood. Combining period budget constraints and defining τ to be the tax rate on income, the agent's allocation problem is to choose c_3, c_4, f_1, and f_2 to maximize equation (4) subject to the relationships in equation (1)

---

[8]It is simple to allow for human capital to be gained also through experience so that income rises through the life cycle. As this serves only to scale our results, it is omitted.

6

and

$$I\left(1-\tau\right) \geq c_3 + \frac{c_4}{r} + f_1 + \frac{f_2}{r},$$
$$c_3, c_4. f_1, f_2 \geq 0,$$
$$\hat{I} = \hat{I}\left(\hat{h}_2\right), \ \hat{h}_2 = \hat{h}_2\left(\hat{i}_1, \hat{i}_2\right),$$
$$\hat{i}_1 = \hat{i}_1\left(f_1, g_1\right), \ \text{and} \ \hat{i}_2 = \hat{i}_2\left(f_2, g_2\right). \tag{5}$$

2.3 Other entities.

A large number of identical firms employ labor to produce identical consumption goods according to

$$Y = ZH \tag{6}$$

where $Z > 0$ is a scalar, Y is output, and H is the human capital adjusted labor input of a representative firm. Since all markets are competitive $w = Z$ will hold in equilibrium.

We assume that government taxes all labor income at the common rate τ and uses the revenue to fund early and late childhood education. Furthermore, government spends equally on all children over their lifetime. Given this and the normalization of the generation size to one, a balanced budget requires that

$$G = g_1 + g_2 = ZH\tau \tag{7}$$

where G is total government spending in period t.

It is convenient to scale spending to the size of the economy. We do this by making total spending in any period proportional to output. Furthermore, we define ψ to be the share of G that is devoted to early childhood education. Thus we set

$$G = \varsigma Y, \quad \varsigma \psi Y = g_1, \quad \varsigma\left(1 - \psi\right) Y = g_2 \tag{8}$$

where $\varsigma \in [0, 1]$ is the share of output devoted to government education spending.

To complete the model, we assume that agents can borrow and lend in an international market. Here a unit of the consumption good today purchases a claim to r units in the subsequent period. This makes the interest rate exogenous as required for analytical tractability.

2.4 Equilibrium.

While output is a constant returns to scale technology in human capital (equation (6)) and human capital is constant returns to scale in i_2 and h_1 (equation (1)), the model does not exhibit long run positive growth. This is due to the assumption that $\gamma_1 < 1$ which assures that h_1 exhibits

decreasing returns to scale in i_1. Because of this, h_2 exhibits decreasing returns to scale in i_1 and i_2 and the economy converges to a steady state.[9] Thus we focus on the level rather than growth effects of policy. With heterogeneity, this provides a simpler setting to consider how policy implications depend on a family's place in the income distribution.

In Appendix 2, we develop the dynamic model and demonstrate convergence of the economy to its steady state. However, our concern is with comparative statics and as such we focus on a steady-state equilibrium. In this case, the total amount of labor available in each period, H_2 is

$$H_2 = 2 \int_{j=0}^{1} h_{2j} dj \tag{9}$$

where the 2 reflects that two generations are at work in each period.

Definition 1. A steady-state competitive equilibrium in this economy is a wage w, income, allocations and educational outcomes $\left\{ I_j, c_{3,j}, c_{4,j}, f_{1,j}, f_{2,j}, h_{1,j}, h_{2,j}, \hat{I}_j, \hat{h}_{1,j}, \hat{h}_{2,j} \right\} \ \forall j \in \mathcal{J}$, labor supply and demand $\{H_2, H\}$, and fiscal instruments $\{\tau, \varsigma, \psi, g_1, g_2\}$ such that

1. Human capital allocations satisfy equation (1).

2. Each agent takes $h_{1,j}, h_{2,j}$, fiscal instruments, and the choices of others as given and chooses $c_{3,j}, c_{4,j}, f_{1,j}, f_{2,j}$ to satisfy equation (4) subject to the constraints in equation (5).

3. The firms choose labor inputs to maximize profits, $w = Z$.

4. Government spending satisfies equation (7).

5. The labor market clears, $H_2 = H$.

6. Surpluses and shortages in the goods market are accommodated by the international bond market.[10]

7. $h_{2,j} = \hat{h}_{2,j}$ and similarly other generation specific variables are constant.

3   A special case.

The model generally requires numerical solutions but insights can be gained by first looking at a special case. For this purpose we make the following assumption:

---

[9]In Appendix 2 we demonstrate convergence numerically in the general case. The key to convergence is decreasing returns to scale in the reproducible inputs, i_1 and i_2, which is assured by $\gamma_1, \gamma_2 \in (0, 1)$.

[10]Implications of the model are qualitatively robust to the closed economy case where the goods market clears.

Assumption 1: $\eta = r = 1$, $\rho = 0$, $\alpha = .5$.

Setting $\eta = 1$ makes government spending perfectly substitutable with private spending. The value of α gauges the relative productivity of government and family spending in producing human capital. By setting $\alpha = .5$ these two types of expenditures are equally productive. With these two assumptions

$$i_k = f_k + g_k \tag{10}$$

so that investment at a stage of education is simply the sum of private and public spending.[11] Setting $\rho = 0$ causes the elasticity of substitution across expenditure in the two stages to be 1. In this case the production function for h_2 becomes

$$h_2 = \bar{A} i_1^{\gamma} i_2^{\gamma_2} \tag{11}$$

where $\gamma = (1 - \gamma_2) \gamma_1$ and $\bar{A} = A a^{2-\gamma_2}$. Setting the return to savings, r, equal to one is an algebraic convenience with little consequence for any of our results. Using a different r serves only to scale some of our later findings.

For this section we also assume that agents are of equal ability. This requires:

Assumption 2: $a_j = a_{j'} \; \forall j$ and $j' \in \mathcal{J}$.

3.1 Equilibria.

We show in Appendix 1 that a unique equilibrium exists for any choice of government policy and that families may spend on one, both, or neither stage of education. The family spending pattern is influenced by government policy. To facilitate our discussion of this relationship, it is useful to define several regions of the level of government spending and its allocation.

Definition 2. The level of government education spending is *low* when $\varsigma < \min \left[\xi \gamma_2 \beta_1^{-1}, \xi \gamma \beta_1^{-1} \right]$, *high* when $\varsigma > \xi (\gamma + \gamma_2) \beta_1^{-1}$, and *moderate* otherwise. The allocation of government education spending is *focused on late childhood* when $\psi \leq \min \left[1 - \xi \gamma_2 (\beta_1 \varsigma)^{-1}, \; \xi \gamma (1 - \varsigma) ((1 + \beta) \varsigma)^{-1} \right]$, *focused on early childhood* when $\psi \geq \max \left[\xi \gamma (\beta_1 \varsigma)^{-1}, \; 1 - \xi \gamma_2 (1 - \varsigma) ((1 + \beta) \varsigma)^{-1} \right]$, and *balanced* otherwise.

Here $\beta_1 \equiv 1 + \beta + (\gamma_2 + \gamma) \xi$. The precise values in the definition are less important than the notion that they delineate distinct regions of the policy parameter space. Critical values in the definition depend on preference parameters and also on the parameters that gauge the curvature

---

[11] More specifically $i_k = .5 B (f_k + g_k)$ but there is no loss of generality in setting $B = 2$.

of the function for accumulating human capital. For example, we state below that spending is low when families spend on both stages of education regardless of how government spending is allocated. The cutoff level of government spending for this depends on how productive spending is at each stage and how much families value human capital relative to consumption. More generally, these same parameters influence family spending patterns and define regions where this is positive or not at each stage.

Definition 2 maps these precise values into several descriptive titles which suffice for the remaining discussion. Spending is low, moderate, or high with no overlap or gaps across regions. The allocation of spending is focused on early childhood, focused on late childhood, or balanced. Again there are no overlaps or gaps across regions.[12] Proposition 1 uses these terms to describe how the level and allocation of government spending jointly determine family spending patterns.

Proposition 1. Let Assumptions 1 and 2 hold.

   i. If the level of government spending is low, families spend on both stages of education.

   ii. If government spending is moderate and focused on one stage of education, families spend only on the other stage. If government spending is moderate and balanced, families spend on both stages.

   iii. If government spending is high and focused on one stage of education, families again spend only on the other stage. If government spending is high and balanced, families spend on neither stage.

Proofs to all propositions appear in Appendix 1. Item (i) of Proposition 1 shows that regardless of how a low level of spending is allocated, families spend at both stages of education. For example, even if government spends exclusively on early childhood, families devote additional resources to this stage of education. We use the notation $f = (f_1^*, f_2^*)$ to indicate this spending pattern. Here f_1^* indicates a positive optimal family expenditure on early childhood and f_2^* indicates a positive optimal family expenditure on late childhood.

Item (ii) shows that when government education expenditures are moderate, at most one level of family spending can be fully crowded out. The first sentence indicates that when this level of spending is focused on early childhood, families spend nothing at this stage and direct expenditures

---

[12]It is straightforward to show that regions do not overlap and span the entire domain. This is demonstrated in Figure 1.
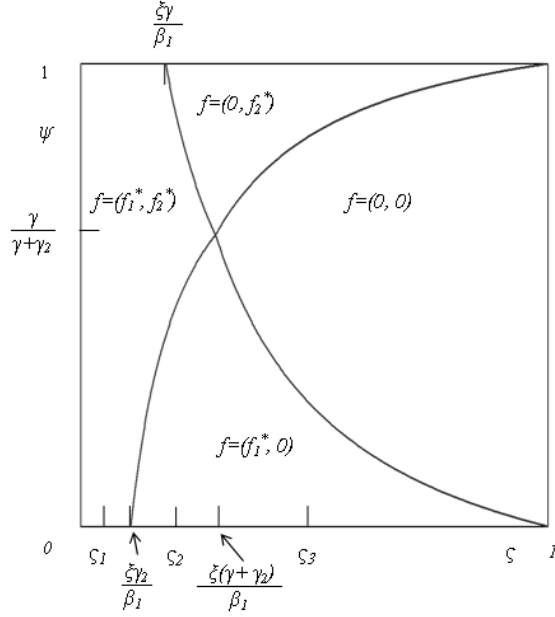
Figure 1: Equilibria. The curves divide the $\varsigma \times \psi$ space into four regions. Where $f = (f_1^*, f_2^*)$, families spend on both levels of education and where $f = (0,0)$, they spend on neither. Otherwise they spend on one level of education. Where $f = (0, f_2^*)$, they spend on late childhood and where $f = (f_1^*, 0)$, they spend on early childhood.

instead to late childhood. We use the notation $f = (0, f_2^*)$ to indicate positive family expenditures on late childhood and no expenditures on early childhood. If instead government spending is focused on late childhood, families spend only on early childhood. We denote this case by $f = (f_1^*, 0)$. The second sentence in item (ii) states that when the same level of expenditures is balanced across the two uses, families again spend at both levels so $f = (f_1^*, f_2^*)$.

Item (iii) considers the case of high government spending. When this is focused, the results are the same as with moderate spending; one level is fully crowded out and the other is not. However, when high government spending is balanced, family spending at both stages is fully crowded out. We denote this case by $f = (0,0)$.

Figure 1 illustrates the relationship between government spending and the pattern of family spending. Here we show the partition of the $\varsigma \times \psi$ space for a particular parameterization. The regions do not overlap and are delineated by the critical values described in Definition 2[13].

To see how ψ and ς jointly determine the pattern of family expenditures, it is useful to consider

---

[13] We set $\gamma_2 = .15$, which is in the range used by Blankenau and Simpson (2004). To reflect a higher productivity for expenditures in early childhood we set $\gamma_1 = .3$. This gives $\gamma = .225$. We set $\beta = .63$ to reflect an annual discount rate of .97 over 15 years and set $\xi = 1 + \beta$.

three values of ς. First, consider $\varsigma = \varsigma_1$ as an example of a low level of government spending. Tracing a line from $\psi = 0$ to $\psi = 1$ at $\varsigma = \varsigma_1$ in Figure 1, we see that for every ψ value, $f = (f_1^*, f_2^*)$. Thus when government spending is low, its allocation does not influence the type of equilibrium.

Next consider $\varsigma = \varsigma_2$ as an example of a moderate level of spending. Tracing a line from $\psi = 0$ to $\psi = 1$, we see that for ψ small, families spend only on early childhood, for ψ large families spend only on late childhood, and otherwise families spend at both stages. That is, when this level of spending is sufficiently focused on one stage of education, families spend only on the other stage. When it is split more equally, the dilution results in private spending at both stages.

Finally consider $\varsigma = \varsigma_3$ as an example of a high level of spending. With focused spending at this level, families again spend only on the stage neglected by government. However, now with more balanced spending, families do not spend at all. That is, when spending is high enough, government spending diluted across the two levels is still sufficiently high at both stages to eliminate private spending.

Appendix 1 shows that throughout the parameter space family spending at each stage depends on at least one policy parameter so that the level of family spending, when positive, varies throughout the region. However the equilibria in the different regions have in common that one, both, or neither level of spending is fully crowded out.

For moderate levels of spending, the range of ψ values considered balanced spending decreases with ς; i.e. the two bounds are getting closer together as ς increases. In contrast, when government spending is high, the range is increasing. The intuition for this result is simple. Moderate spending is balanced when total government spending at each stage, as gauged by $\varsigma\psi$ and $\varsigma(1-\psi)$, is small enough that families top-up government spending at each stage. It is easier to satisfy the conditions simultaneously when government spends less in aggregate. High spending is balanced when there is sufficient government spending at both levels to fully crowd out both levels of private spending. That is, total government spending at each stage must be large enough. It is easier to satisfy the conditions simultaneously when government spends more in aggregate.

3.2 Output.

The above discussion clarifies how the equilibrium family spending pattern depends on the government spending level and its allocation. We now consider how these government choices affect output. We first discuss the effect of policy within the different regions and identify policies that

maximize output locally. We then compare these local maxima to identify the policies that maximize output globally. We fully specify the relationship between output and policy in Appendix 1. Proposition 2 gives some of the important features of this relationship.

Proposition 2.

i. In the regions where government spending is either low or both moderate and balanced, output is independent of the level and allocation of spending and families spend at each stage of education.

ii. In the region where government spending is high and balanced, output is increasing in the level of spending and maximized where $\psi = \frac{\gamma}{\gamma + \gamma_2}$. Family spending is fully crowded out.

iii. In the region where government spending is focused on late childhood, output is maximized where $\psi < \frac{\gamma}{\gamma + \gamma_2}$. Family spending on late childhood is fully crowded out. For a range of ς values, output in the region is maximized with government expenditures allocated exclusively to late childhood.

iv. In the region where government spending is focused on early childhood, output is maximized where $\psi > \frac{\gamma}{\gamma + \gamma_2}$. Family spending on early childhood is fully crowded out. For a range of ς values, output in the region is maximized with government expenditures allocated exclusively to early childhood.

We have shown that when spending is either low or both moderate and balanced, families spend at both stages of education. Item (i) in Proposition 2 shows that in this case output is independent of both the level and mix of spending. Regardless of its allocation, government spending at each stage falls below what the family would choose and thus is topped-up with private spending. Since private and public spending are perfect substitutes, a unit more or less of government spending at any stage is fully offset by a unit less or more of private spending. With total spending at each stage unchanged through policy, human capital and hence output are unchanged. The effect of any policy change within this region is fully negated by crowding out.

Item (ii) in the proposition considers the case where expenditures are high and balanced. In this region, both stages of private spending are fully crowded out and government is the sole source of education expenditures. With private expenditures fully crowed out, there is no effect on private spending when government spending rises. For a fixed allocation of government spending, a further

increment to its level unambiguously increases total human capital expenditures at each stage. Consequently, more spending yields more output. For a given level of government spending, output is maximized when the marginal quantity of human capital generated is the same for both levels of expenditure. This occurs where $\psi = \gamma \left(\gamma + \gamma_2 \right)^{-1}$. An implication is that with early childhood spending relatively productive, output is maximized with resources spent disproportionately on early childhood.

Item (iii) examines the case where government focuses its expenditures on late childhood and families consequently spend only on early childhood. In this case, an increase in government spending on early childhood simply crowds out a unit of family spending. This suggests that for purposes of output maximization in the region, government spending should be allocated exclusively to late childhood. Any deviation from this in the direction of early childhood spending decreases government spending on late childhood only to crowd out family spending in early childhood. The level of total spending falls.

This 'level effect' is indeed at play. However, another mechanism works counter to this. Since an increment to government spending on early childhood crowds out private spending at this stage, income net of education spending is increased. The increased income is allocated across all items valued by the family. In particular, part is allocated to family spending on early childhood. The equilibrium consequence is that family spending does not decrease one-for-one with government spending after this general equilibrium effect is considered.

All told, increasing government expenditures on early childhood leaves total government spending as a share of output unchanged and reduces family spending. However, the mix of total expenditures shifts in favor of early childhood. When families spend on only early childhood, it is because that stage is more productive at the margin. A reallocation of total spending in the direction of early childhood, then, means that remaining expenditures are more productive. We call this the 'mix effect'.

These effects offset when $\psi = 1 - \frac{\gamma_2}{(\gamma + \gamma_2)\varsigma} < \frac{\gamma}{\gamma + \gamma_2}$. Depending in part on the level of government spending this value may be less than zero, may occur in the range of ψ values considered to be focused expenditures, and may occur when ψ is too large to be considered focused on late childhood. If it occurs where ψ is less than zero, output over the region is maximized where expenditures are allocated fully to late childhood. That is, the level effect always dominates and output in the region is highest with all government spending devoted to late childhood. If instead the equality

holds where expenditures are focused, there is a local maximum in the interior of this region; the mix effect dominates initially and the level effect dominates for larger values of ψ. Otherwise, the mix effect always dominates and output increases when more resources are devoted to early childhood. Output in the region may be maximized with government resources spent exclusively on late childhood but in any case, it will be maximized with spending focused more on late childhood than is optimal with full crowding out.

Item (iv) examines the case where government focuses its expenditures on early childhood and families spend only on late childhood. This is a mirror image of the case above. In moving from $\psi = 1$ to smaller values, crowding out causes a level effect as the share of output directed to education falls. A mix effect occurs as the share of resources devoted to late childhood rises. Output in the region may be maximized with government resources spent exclusively on early childhood but in any case, it will be maximized with spending focused more on early childhood than is optimal with full crowding out.

Comparing item (ii) with items (iii) and (iv), note that in the first, family spending on education is fully crowded out and higher government spending is not offset by reductions in family spending. In the others potential crowding out must be considered. An implication, then, is that when crowding out is operative, output is maximized with expenditures more focused; i.e. with the share of expenditures allocated to early childhood closer to zero or one.

Figure 2 aids in the discussion. The first panel is equivalent to Figure 1 but further divides the space into regions where output is increasing, decreasing, and invariant the share of spending going to early childhood. The solid curves are as in Figure 1 and thus delineate the four types of family spending patterns. The arrows show directions in which output is increasing or invariant in the level and mix of spending. The dotted lines trace local output maximizing combinations of ψ and ς. The second panel gives similar information from another perspective. This graphs normalized output, y, for all combinations of expenditures and its allocation. Output is normalized by its value with no government expenditures. The points of inflection correspond to the regions delineated in the first panel.

Item (i) in Proposition 2 corresponds to the left-hand region in each panel ($f = (f_1^*, f_2^*)$). The independence of output from policy is demonstrated by the intersecting arrows in the first panel; regardless of direction, a change in the level or allocation of spending has no effect on output in this region. In the second panel, this is demonstrated by the flat area at $y = 1$ over this region.
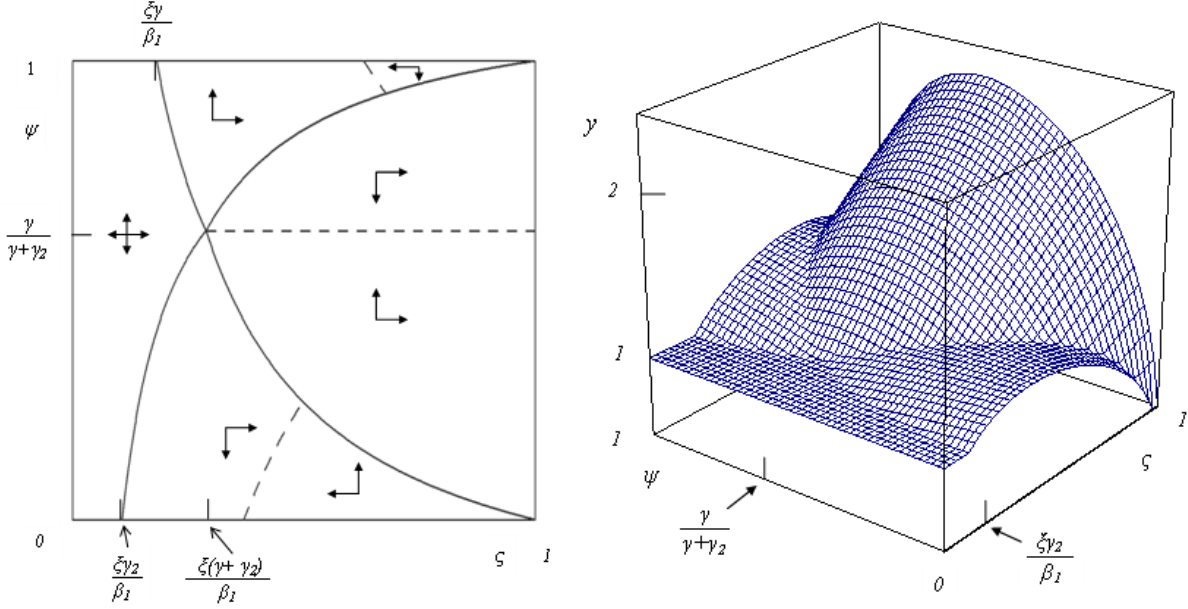
15

Figure 2: Output. The arrows in the first panel show the direction in which output is increasing or invariant in ς and ψ. The second panel shows normalized output. Here output is normalized by its value at $\varsigma = 0$.

Output with government spending is the same as output without government spending. Item (ii) corresponds to the right-hand region between the solid lines ($f = (0,0)$). In the first panel, the arrows indicate that output in this region can be increased by changing ψ in the direction of the dashed line or by increasing the level of expenditures. This is reflected in the second panel by the ridge at $\psi = \gamma \left(\gamma + \gamma_2\right)^{-1}$.

Item (iii) corresponds to the lower region between the solid lines ($f = (f_1^*, 0)$) in the first panel. The arrows indicate that for smaller values of ς, output is decreasing in ψ; for intermediate values, there is a local maximum at the dashed line; and for larger values, output increases in ψ. The contour of this region in the second panel reiterates this relationship. Thus it may be best to focus expenditures fully on late childhood and it may be best to allocate some amount to early childhood. Either way though, the local output maximizing ψ is smaller in this region. Finally, item (iv) corresponds to the upper region between the solid lines. For smaller levels of government spending, output is increasing in ψ and thus maximized with spending allocated exclusively to early childhood. For intermediate values, there is a local maximum at the dashed line and for larger values, output decreases in ψ. Regardless of the level of spending, the local output maximizing ψ is larger in this region.

The above discussion points out that for any level of government spending, there may be several different allocations that locally maximize output. We now consider which of these local maxima is a global maximum for a given level of government spending. Figure 3 shows output at each of the local maxima so that a global maximum, y^*, is easily identified. The horizontal line at $y^* = 1$ shows the local maximum in the region where families spend at both stages. When government spending is low, only this equilibrium exists regardless of its allocation. This level of output, then, is necessarily the global maximum. The bracket below the horizontal axis and furthest to the left encompasses the space of spending levels where this holds.

An equilibrium with family spending at both stages exists beyond this bracketed region. This is indicated by the extension of the horizontal line further to the right. However, immediately beyond the first bracketed region, families spend at both stages only if government expenditures are balanced. With spending focused on late childhood, families will now spend only on early childhood. Depending on the allocation of expenditures, then, two types of equilibria are possible. The dashed line originating furthest to the left in Figure 3 shows output under this spending pattern and with ψ chosen to maximize output. That is, it shows a second local maximum. The figure shows that immediately beyond the first bracketed region, output is globally maximized when government focuses its spending on late childhood.

For yet larger values of ς, a third type of equilibrium is possible. With spending focused on early childhood, families will now spend only on late childhood. The dashed line originating furthest to the right in Figure 3 shows output when families spend on late childhood education and with ψ chosen to maximize output. Thus it shows a third local maximum. Output at this local maximum is initially below that where families spend on early childhood education. However, immediately beyond $\varsigma = \underline{\varsigma_y}$ this local maximum is the global maximum.

Similarly for larger values of ς, an equilibrium with no family education spending arises. The solid upward sloping curve gives locally maximized output in this region. Beyond $\varsigma = \overline{\varsigma_y}$ this local maximum is the global maximum.

Note that focusing government spending on late childhood, focusing spending on early childhood, and balancing spending across these two uses can each be globally optimal for some range of government spending. The brackets indicate regions of spending where each of these circumstances holds. This turns out to be a general result and we state it more precisely as Proposition 3.

Proposition 3. Let $\gamma > \gamma_2$. The space $\varsigma \in [0, 1]$ can be divided into four regions such that in
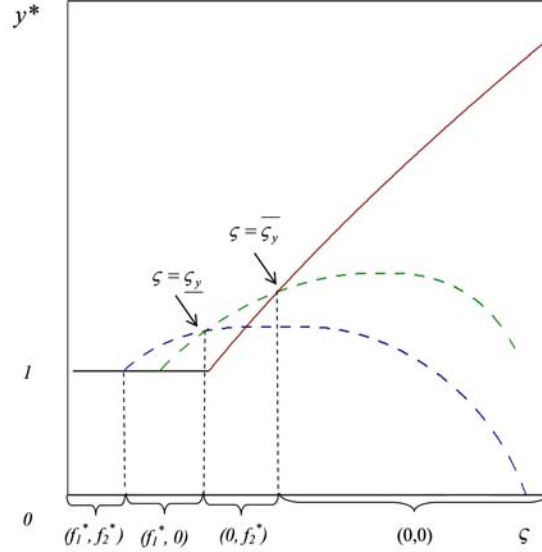
Figure 3: Maximum output. The flat and increasing solid curves give maximum output in the regions where $f = (f_1^*, f_2^*)$ and $f = (0,0)$. The dashed curve originating further to the left shows the maximum level of output in the region where $f = (f_1^*, 0)$. The other dashed curve shows maximum output in the region where $f = (0, f_2^*)$. The brackets show which type of equilibrium maximizes output globally at the relevant value of ς.

the first (lowest) region, output is globally maximized at any allocation of expenditures, in the second highest region, output is globally maximized with government expenditures focused on late childhood, in the third region output is globally maximized with government expenditures focused on early childhood, and for the highest region output is globally maximized when spending is balanced.

An implication of Proposition 3 is that focused spending can dominate balanced spending.[14] When focused spending dominates, the proposition also shows which level of education should receive the lion's share of funding. One might expect that with $\gamma > \gamma_2$, education spending should be focused on early childhood where it is more productive. However, this holds only where ς is between these critical values. For smaller values, it should be focused on the less productive form of education.

To see why, note that we are considering cases where government spends on one stage of education and families spend on the other. The key is to apply the largest block of funds to its most productive use. Suppose that family spending is higher than government spending. Then out-

---

[14]Findings are symmetric if $\gamma \leq \gamma_2$ though $f = (f_1^*, 0)$ and $f = (0, f_2^*)$ switch order along the horizontal axis in Figure 3.

18

put is maximized where families spend on the more productive stage indicating that government should fund the less productive stage of education. If instead government spending exceeds family spending, output is maximized where government spends on the more productive stage.

The discussion provides a potential validation of the pattern of education spending in the U.S. While mounting evidence supports the importance of early childhood education, government focuses expenditures on older students. This pattern may be optimal as it minimizes crowding out. Government spending on late childhood motivates private spending on early childhood. A reallocation toward early childhood may result in increased crowding out of private expenditure and lower output. We find that it is best to balance expenditures only when government spending relative to income is sufficiently high.

3.3   Utility.

We now show that to a large extent, the qualitative arguments above hold when we instead consider utility. As with output, for any level of government spending there may be several locally optimal ways to allocate expenditures. Each of these is globally optimal for some level of spending. In particular, the current U.S. focus on late childhood may maximize both output and utility over some range of expenditures.

Precise expressions for utility are given in Appendix 1. We present some of the key implications of this in Figure 4. The first panel is analogous to the second panel of Figure 2 but graphs normalized utility, u, rather than output. When spending is low, output and utility maximization are equivalent since government spending just offsets private spending, leaving all allocations unchanged. With focused spending, the welfare maximizing levels of ς are lower than those which maximize output. The bigger difference relative to the output discussion occurs with high balanced spending. While output with balanced spending always increases as more is spent on education, utility does not. These differences reflect that increased expenditures increase the tax burden and lower consumption through this channel.

The second panel of Figure 4 is analogous to Figure 3. It shows the maximum utility attainable in each type of equilibrium over the range of spending for which the equilibrium type exists. As with output, the utility maximizing mix of expenditures depends on the level of expenditures. While the cutoff points ($\underline{\varsigma_u}$ and $\overline{\varsigma_u}$) for the various regions are different, other results are similar. In the bracketed range furthest to the left, utility is maximized where families spend at both stages. In the
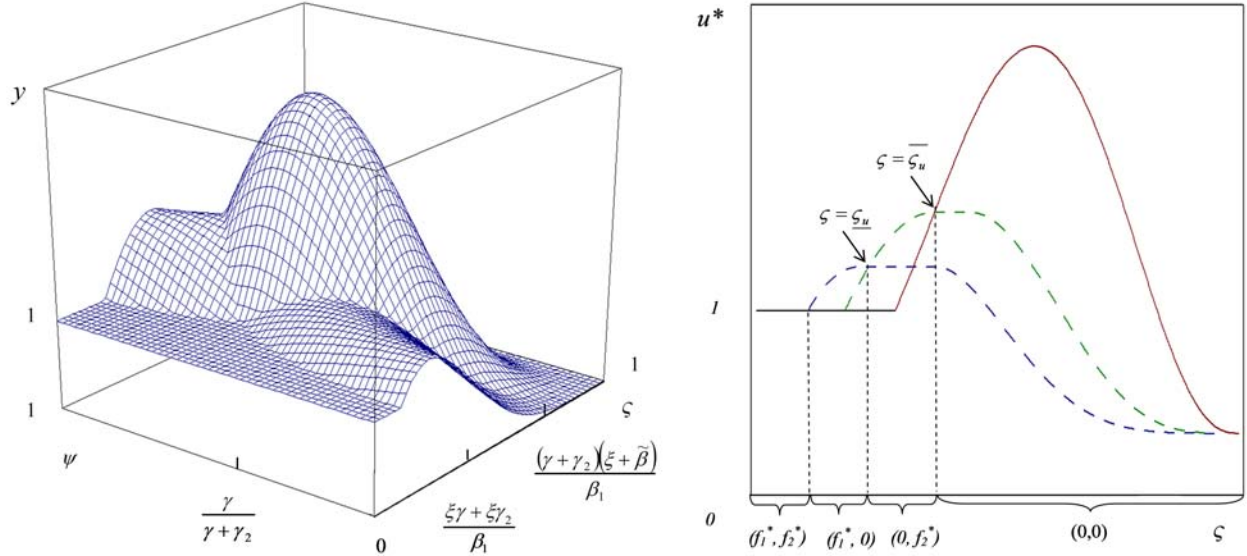
19

Figure 4: Utility. The first panel shows normalized utility over the policy space. The second shows normalized utility across values of ς in each type of equilibrium when ψ is chosen to yield a local maximum. The brackets show which type of equilibrium maximizes utility globally at the relevant value of ς.

subsequent two ranges, utility is maximized with government spending focused on late childhood and then early childhood. For larger spending levels, balanced expenditures maximize utility.

3.4 The general case.

The previous subsections require several restrictive assumptions. Here we relax several items of Assumption 1 and demonstrate that the restrictive model captures much of the key intuition arising in the more general model.

Relaxing any of the assumptions requires solving the model numerically. The first order conditions for the more general problem are straightforward extensions of those in the proof to Proposition 1 and are not presented here. For brevity we hereafter focus on output. From the preceding section it is clear that results regarding utility are similar. These results and a more full sensitivity analysis are available from the authors upon request.

In the first panel, we set $\eta = .95$ so that private and government spending are imperfectly substitutable in the production of human capital. Results are similar to the second panel of Figure 2 but there is now a smoothing of the surface between the different regions. With imperfect substitutability, family spending in either category will never go to zero. Thus we no longer have as
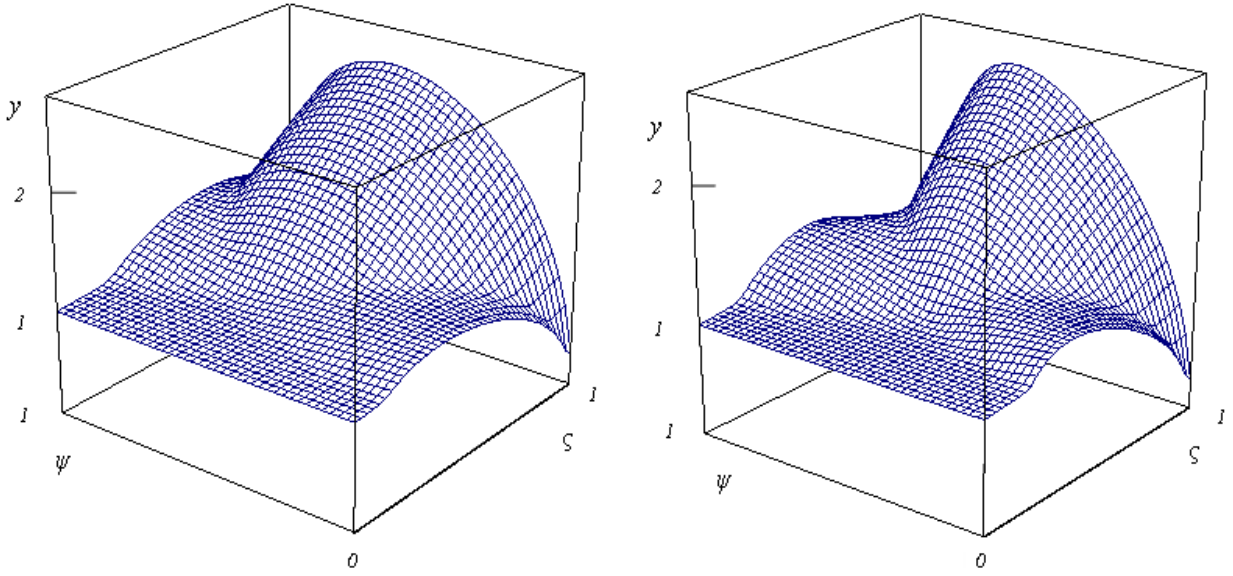
20

Figure 5: Sensitivity. The first panel sets $\eta = .95$. The second panel additionaly sets $\rho = -1$.

sharp a distinction across the regions. However, each policy pair yields results familiar from the case with perfect substitutability. In particular, for moderate and high government spending, we have local maxima at several values of ψ. The global maximum again depends on the level of spending and in the same way as before. However, it is straightforward to show that when human capital is a Cobb-Douglas combination of private and government spending, output is always maximized when resources are split relatively equally. From this we conclude that concentrated public spending can maximize output only in the case where private and public spending are relatively close substitutes.

In the second panel of Figure 5, we additionally set $\rho = -1$ so that early and late childhood expenditures are more complementary than in the Cobb-Douglas case. One difference is that the output maximizing level of early childhood spending shifts to the left (when not a corner). This is because early childhood spending now has a larger positive effect on the productivity of later spending. Still, the results mirror those in Figure 2 and the intuition above still serves to understand the results.

4   Heterogeneity.

We now consider the impact of policy across a heterogeneous population. As stated in Section 2, heterogeneity is expressed by different levels of a_j. There are strong similarities between the

heterogeneous family economy and the one family economy discussed above. Since the higher indexed families will have a higher value of a_j, in equilibrium they will also have higher income. This is due both to the differences in ability and to the resulting differences in family education spending. With heterogeneity, the common level of government expenditure for each family will represent different ratios of government spending to individual income. In particular, a common level of government education spending will represent lower government spending relative to income for high income families than for low income families. To see it, recall that ς is the share of total output that goes to education. With the population of each generation normalized to one, lifetime government education spending per family is ςZH. Since the income of family j is Zh_{2j}, government spending as a share of own income for family j is

$$\varsigma_j = \varsigma \frac{H}{h_{2j}}. \tag{12}$$

It is this value that matters to families, rather than ς alone.

The distribution of ς_j clearly depends on ς and the distribution of output (i.e. the distribution of h_{2j}). This latter item, maps into the distribution of a_j. Stated differently, when we choose an exogenous distribution of a_j, this generates an equilibrium endogenous distribution of ς_j. In this way there is a relationship between the distributions of a_j and ς_j. The relationship is such that the smallest a_j is associated with the largest ς_j. For wealthier, high ability families government spending is small in relation to income. For poorer, low ability families this same government spending is large relative to income.

With a few caveats we can provide a different interpretation of our earlier findings. Rather than considering a representative family at different levels of spending, we can consider different families where the common absolute government spending means different relative spending for the different families. In the earlier analysis $H = 2h_2$ so differences in ς_j are generated by differences in ς. Now we hold the level of expenditures constant and allow differences in ς_j though heterogeneity in h_{2j}.

From equation (9) and the final line of Definition 1, both total human capital and its distribution must now be fixed in a steady state. No additional assumptions are required to assure the existence of this steady state. This is shown numerically in Appendix 2.[15] For any level of per capita government spending and tax rate, we can independently find h_{2j} for each family in a steady state

[15] Appendix 2 solves the model for a particular parameterization. The results hold over wide range of parameter values and details are available upon request.
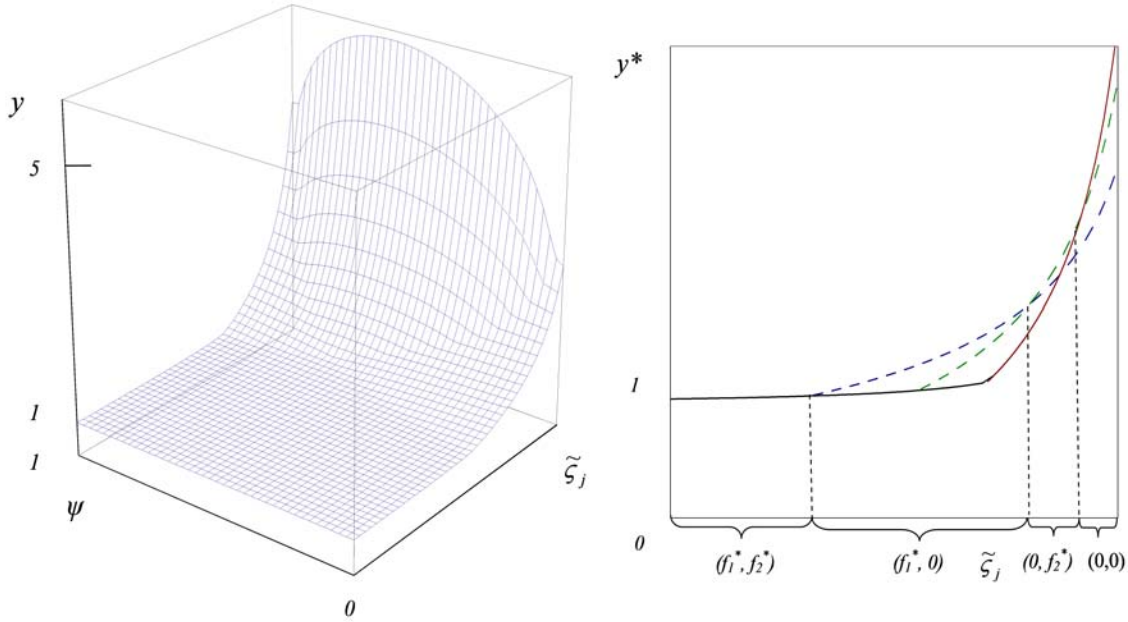
Figure 6: Output and maximum output. The first panel shows normalized output across the income distribution (represented by $\tilde{\varsigma}_j$) as a function or ψ. The second panel shows normalized output across the income distribution when ψ is chosen to yield a local maximum. The brackets show which type of equilibrium maximizes output globally at the relevant point in the income distribution. For each agent, output is normalized by output for that agent when $\varsigma = 0$.

where h_{2j} will be function of a_j. That is, the agent's problem is similar to the special case in Section 3. It differs in that government spending need not equal the tax burden for a particular agent. While this forces numerical solutions for each agent, it does not change the nature of the agent's problem and a steady state arises where h_{2j} is unchanging in a family through time. By aggregating these we can find H and from equation (12) we can find the distribution of ς_j. Since this can be done for any level of spending, we can iterate on ς until additionally the government budget constraint is satisfied. Whatever level of spending holds at this point, there will be a steady state distribution of human capital.

With heterogeneity, we must turn to numerical results even with the parameter restrictions in Assumption 1. These results are presented in Figure 6. The first panel is analogous to the second panel of Figure 2. The difference is that the variation in ς_j is a general equilibrium consequence of variation in ability. Specifically, for this example we assume that ability is uniform over $[.5, 5]$ and $\varsigma = .03$. We then find values of $\{h_{2,j}\} \ \forall j \in \mathcal{J}$ and other endogenous items such that the definition

23

of an equilibrium is satisfied. Given the level of human capital for each family, we know H and thus can use equation (12) to find the distribution of ς_j. For ease of comparison, we plot a monotonic transformation of ς_j ($\tilde{\varsigma}_j$) against ψ on the horizontal axis and normalized output along the vertical axis.[16] As before, output is normalized by what it would be with $\varsigma = 0$.

The first panel of Figure 6 shows that there are again four distinct regions. These correspond to the regions in Figure 2. To facilitate comparison with the special case in Section 3, we provide the following summary of Figure 6 which corresponds closely to Proposition 2. While a formal proof is not available in this general case and statements are less precise, the similarity of the results is clear.

    i. For families where government spending is low relative to income or moderate but balanced, output is independent of the allocation of spending and families spend at each stage of education.

    ii. For families where government spending is high relative to income and balanced, proportional output gains from spending are increasing in the relative level of spending and maximized where $\psi \in (0, 1)$. Family spending is fully crowded out.

    iii. In the region where government spending is focused on late childhood, output is maximized where ψ is small. Family spending on late childhood is fully crowded out. For a range of $\tilde{\varsigma}_j$ values, output in the region is maximized with government expenditures allocated exclusively to late childhood.

    iv. In the region where government spending is focused on early childhood, output is maximized where ψ is large. Family spending on early childhood is fully crowded out. For a range of $\tilde{\varsigma}_j$ values, output in the region is maximized with government expenditures allocated exclusively to early childhood.

The second panel of Figure 6 is analogous to Figure 3 and shows the global maximum income as a function of $\tilde{\varsigma}_j$. To again facilitate comparing the heterogeneous case with the special case in Section 3, we provide the following summary which corresponds closely to Proposition 3: The space $\tilde{\varsigma} \in [0, 1]$ can be divided into four regions such that in the first region (low $\tilde{\varsigma}$ and high income),

---

[16]Specifically, the axis is $\tilde{\varsigma}_j = a(j) \in [.5, 5]$. This allows for easier comparison and provides the same essential information since there is a one-to-one correspondence between $a(j)$ and ς_j.

output is globally maximized at any allocation of expenditures, in the second region (higher $\tilde{\varsigma}$ and lower income), output is globally maximized with government expenditures focused on late childhood, in the third region output is globally maximized with government expenditures focused on early childhood, and for the lowest income (highest $\tilde{\varsigma}$) region output is globally maximized when spending is balanced. Stated differently, the lowest income families prefer balanced spending, lower middle income families prefer spending focussed on early childhood, upper middle income families prefer spending focussed on late childhood, and the most wealthy are indifferent. The intuition for these findings is analogous to that in the special case.

While the analogy with the homogeneous case is quite strong, there are some differences. In Figure 2 output is non-monotonic in government spending levels when families spend on only one level of education. In Figure 6 it is monotonic. This is because in Figure 2 an increase in ς requires an increase in taxes which crowds out private spending. In Figure 6, ς is fixed so this effect does not arise. Also in Figure 6, when families spend on both levels of education, y^* increases moderately with ς_j. This reflects that the income tax to finance education is more onerous for those with larger incomes.

Despite these minor differences, we can by and large take the discussion regarding output and utility in the above sections and generalize it to the case where families differ in ability. We need only to recognize that a level of government spending signifies a different relative level of spending for the different families. In general, when there are substantial differences in income, there will be differences in preferred policies. In particular, focusing expenditures on late childhood, as is done in the U.S., may be the best allocation for agents over some region of the income distribution whereas others would benefit from a reallocation toward early childhood.

5   Conclusion.

Early childhood education builds a foundation of knowledge and habits that makes later education more productive. Later education gives this foundation value through a realization of potential. Most prior work abstracts from this hierarchical structure of human capital accumulation. This paper contributes to a nascent literature that instead makes this structure the focal point of its investigations. Our purpose is to evaluate the structure of government education spending in a model of hierarchical human capital accumulation. Currently, government spending favors late childhood over early childhood. We explore whether a reallocation toward early childhood would

be beneficial.

Our general equilibrium environment accounts for crowding out of private spending by public spending. In our baseline model, private and public spending are perfectly substitutable so that a unit of government spending offsets a unit of private spending. Only when private spending on at least one stage of education is driven to zero can policy affect output. We show that for low levels of funding, government maximizes output by funding only the less productive type of education. For intermediate levels of funding, government should finance only the more productive type of education. Only when the total level of funding is above a threshold should it fund both.

The first results are derived in a highly stylized setting. This has the advantage of analytical tractability. The stylized model also proves sufficient for demonstrating the key implications of the model. Through sensitivity analyses, we demonstrate that relaxing this strict structure leaves the most interesting results qualitatively unchanged. An exception is the perfect substitutability of private and public resources. When we make these inputs relatively substitutable, but not perfectly so, results are largely unchanged. When the inputs are relatively complementary, output is no longer maximized by focused spending.

The final part of the paper shows that these results can be easily generalized to the case of heterogeneous agents. The different levels of spending in earlier sections correspond to different income levels in the final section. With a common level of education spending across agents, there will be agents who privately spend at both stages, one stage, or no stage. The analysis shows that focused spending can be best for some part of the population while inappropriate for the lower income agents.

Our concern is the theoretical implications of allocating government education expenditures in a hierarchical education system. To maintain focus, even our more general model abstracts from many important considerations. As such, we do not attempt to quantify our findings through a careful calibration. Such a quantitative investigation would be a useful next step. There are a number of issues that might prove interesting in a fuller model. Our model has no physical capital in production. Thus there is no worry of taxation lowering the capital stock. Our model has no credit constraints despite their central role in many other studies of education. We do not consider imperfect inheritance of ability. These omissions could be remedied in a fuller, empirical investigation. However, we expect that the key intuition developed above will continue to hold and thus aid in our understanding of the implications of government education spending.

A more complete analysis might also consider a fuller set of policy options. For brevity, we have restricted attention to the experiments described above. The model, however, is suggestive of other policy implications. Rather than considering spending policies which are symmetric across the population, we could consider the effects of progressive spending where government spends more on those with lower income. This is more reflective of the well-known Perry Preschool Project, the Abecedarian Project (see CHLM (2007)), and Head Start (see Currie (2001b)). Each of these has targeted low income families and has arguably been highly beneficial to the targeted population. In our setup, we would expect to see expenditures at these levels have the largest impact due to diminished crowding out and a higher marginal benefit to an increment in total spending for low income households. This would be consistent with the conclusion by Currie (2001a) that "priority should be given to expanding Head Start rather than funding universal preschool" since children of the lower income parents are more in need of quality preschool. Furthermore, progressive spending may have additional economy-wide benefits when different levels of skill are complements in production. A potentially fruitful direction for future policy analysis, then, is the exploration of optimal spending allocation across the income distribution.

References

[1] Blankenau, W., 2005. Public schooling, college subsidies and growth. Journal of Economic Dynamics and Control 29, 487-507.

[2] Blankenau, W., 2006. A simple economic theory of skill accumulation and schooling decisions. Review of Economic Dynamics 9, 93-115.

[3] Carneiro, P., Heckman, J., 2003. Human capital policy. National Bureau of Economic Research, Cambridge, working paper No. 9495.

[4] Carnerio, P., Heckman, J., 2002. The evidence on credit constraints in post-secondary schooling. Economic Journal 112, 705-734.

[5] Cunha, F., Heckman, J., 2007. The technology of skill formation. American Economic Review 97, 31-47.

[6] Cunha, F., Heckman, J., Lochner, L., Masterov, D., 2006. Interpreting the evidence on life cycle skill formation, in: Hanushek, E., Welch, F. (Eds.), Handbook of the Economics of Education, Vol 1. Elsevier, Amsterdam, pp. 697-812.

[7] Currie, J., 2001. A fresh start for head start. The Brookings Institutions Children's Roundtable. http://www.brookings.edu/~/media/research/files/papers/2001/3/education%20currie/issue5.pdf.

[8] Currie, J., 2001. Early childhood intervention programs: What do we know? Journal of Economic Perspectives 15, 213-238.

[9] Driskill, R., Horowitz, A., 2002. Investment in hierarchical human capital. Review of Development Economics 6, 48-58.

[10] Glomm, G., Kaganovich M., 2003. Distributional effects of public education in an economy with public pensions. International Economic Review 44, 917-937.

[11] Jensen, A., 1980. Bias in Mental Testing. Free Press, New York.

[12] Kaganovich, M., 2005. Education reform and access to public education in Russia. Indiana University, working paper.

[13] Knudsen, E., Heckman, J., Cameron, J., Shonkoff, J., 2006. Economic, neurobiological, and behavioral perspectives on building America's future workforce. National Bureau of Economic Research, Cambridge, working paper No. 12298.

[14] Mazumder, B., 2005. Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. Review of Economics and Statistics 87, 235-255.

[15] Organization for Economic Cooperation and Development, 2007. Education at a Glance 2007. OECD, Paris.

[16] Rangazas, P., 2000. Schooling and economic growth: A King-Rebelo experiment with human capital. Journal of Monetary Economics 46, 397-416.

[17] Restuccia, D., Urrutia C., 2004. Intergenerational persistence of earnings: The role of early and college education. American Economic Review 94, 1354-1378.

[18] Solon, G., 1999. Intergenerational mobility in the labor market, in: Ashenfelter O., Card D., (Eds.), Handbook of Labor Economics, Vo1. 3. Elsevier, Amsterdam, pp. 1761-1800.

[19] Solon, G., 2002. Cross-country differences in intergenerational earnings mobility. Journal of Economic Perspectives 16, 59-66.

[20] Su, X., 2004. The allocation of public funds in a hierarchical educational system. Journal of Economic Dynamics and Control 28, 2485-2510.

6 Appendix 1

Proof of Propositions 1 and 2 and some results regarding utility. We first give more precise statements regarding the existence of equilibria, output, and utility. Next we argue that these more precise statements are consistent with the propositions and with the claims regarding utility. Finally, we prove these statements.

Equilibria are related to government policy according to

$$f = \begin{cases} (f_1^*, 0) & \text{if} & \psi \leq \min\left[1 - \xi\gamma_2\left(\beta_1\varsigma\right)^{-1}, \ \xi\gamma\left(1-\varsigma\right)\left((1+\beta)\varsigma\right)^{-1}\right] \\ (f_1^*, f_2^*) & \text{if} & 1 - \xi\gamma_2\left(\beta_1\varsigma\right)^{-1} \leq \psi \leq \xi\gamma\left(\beta_1\varsigma\right)^{-1} \\ (0, 0) & \text{if} & \xi\gamma\left(1-\varsigma\right)\left((1+\beta)\varsigma\right)^{-1} \leq \psi \leq 1 - \xi\gamma_2\left(1-\varsigma\right)\left((1+\beta)\varsigma\right)^{-1} \\ (0, f_2^*) & \text{if} & \psi \geq \max\left[\xi\gamma\left(\beta_1\varsigma\right)^{-1}, \ 1 - \xi\gamma_2\left(1-\varsigma\right)\left((1+\beta)\varsigma\right)^{-1}\right] \end{cases} \tag{13}$$

where $\beta_1 \equiv 1 + \beta + (\gamma_2 + \gamma)\xi$, income is related to government policy according to

$$
I^{1-\gamma-\gamma_2} = \begin{cases}
\bar{A}2w \left(\xi\gamma \left(1 - \varsigma \left(1 - \psi\right)\right) \beta_2^{-1}\right)^\gamma \left(\varsigma \left(1 - \psi\right)\right)^{\gamma_2} & \text{if} \quad f = (f_1^*, 0) \\
\bar{A}2w \left(\xi\gamma\right)^\gamma \left(\xi\gamma_2\right)^{\gamma_2} \beta_1^{-(\gamma+\gamma_2)} & \text{if} \quad f = (f_1^*, f_2^*) \\
\bar{A}2w \left(\varsigma\psi\right)^\gamma \left(\varsigma \left(1 - \psi\right)\right)^{\gamma_2} & \text{if} \quad f = (0, 0) \\
\bar{A}2w \left(\varsigma\psi\right)^\gamma \left(\left(1 - \varsigma\psi\right) \beta_3^{-1}\xi\gamma_2\right)^{\gamma_2} & \text{if} \quad f = (0, f_2^*)
\end{cases}
\tag{14}
$$

where $\beta_2 \equiv 1 + \beta + \gamma\xi$ and $\beta_3 \equiv 1 + \beta + \gamma_2\xi$, and utility is related to government policy according to

$$
\bar{U} = \begin{cases}
\beta_2^{-(1+\beta)} \left(1 - \varsigma \left(1 - \psi\right)\right)^{1+\beta} I^{\xi+1+\beta} & \text{if} \quad f = (f_1^*, 0) \\
\beta_1^{-(1+\beta)} I^{\xi+1+\beta} & \text{if} \quad f = (f_1^*, f_2^*) \\
\left(1 + \beta\right)^{-(1+\beta)} \left(1 - \varsigma\right)^{1+\beta} I^{\xi+1+\beta} & \text{if} \quad f = (0, 0) \\
\beta_3^{-(1+\beta)} \left(1 - \psi\varsigma\right)^{1+\beta} I^{\xi+1+\beta} & \text{if} \quad f = (0, f_2^*)
\end{cases}
\tag{15}
$$

where $\bar{U}$ is a monotonic transformation of U.

Item (i) of Proposition 1 follows from the second line of equation (13) and the definition of low spending. To show it, note that when ς satisfies the definition of low spending, both inequalities in the second line are satisfied for all values of $\psi \in [0, 1]$ and inequalities in lines 1, 3, and 4, are not satisfied for any values of ψ.

The first sentence in item (ii) follows from the first and fourth lines of equation (13) and the definitions of moderate and focused spending. When ς satisfies the definition of moderate spending, and ψ satisfies the definition of spending focused on late childhood, only the inequality in line one is satisfied. When ς satisfies the definition of moderate spending, and ψ satisfies the definition of spending focused on early childhood, only the inequality in line four is satisfied. In each case, then, families are spending only on the stage of education which is not the focus of government spending. The second sentence follows from the second line and the definitions of moderate spending and balanced spending.

The first sentence in item (iii) follows from the first and fourth lines of equation (13) and the definitions of high spending and focused spending. When ς satisfies the definition of high spending, and ψ satisfies the definition of spending focused on late childhood, only the inequality in line 1 is satisfied. When ς satisfies the definition of high spending, and ψ satisfies the definition of spending focused on early childhood, only the inequality in line four is satisfied. The second sentence follows from the third line and the definitions of high spending and balanced spending.

Item (i) in Proposition 2 follows from the second lines of equations (13) and (14). Item (ii) in Proposition 2 follows from the third lines of equations (13) and (14). It is straightforward to show that $\psi = \frac{\gamma}{\gamma+\gamma_2}$ maximizes output in this region. Item (iii) in Proposition 2 follows from the first

lines of equations (13) and (14). It is straightforward to show that $\psi = \max\left[0, 1 - \frac{\gamma_2}{\varsigma(\gamma+\gamma_2)}\right] < \frac{\gamma}{\gamma+\gamma_2}$ maximizes output in this region, and that for some range in the region, $0 > 1 - \frac{\gamma_2}{\varsigma(\gamma+\gamma_2)}$. Item (iv) in Proposition 2 follows from the fourth lines of equations (13) and (14). It is straightforward to show that $\psi = \min\left[1, \frac{\gamma}{\varsigma(\gamma+\gamma_2)}\right] > \frac{\gamma}{\gamma+\gamma_2}$ maximizes output in this region and that for some range in the region $1 < \frac{\gamma}{\varsigma(\gamma+\gamma_2)}$.

Claims regarding utility in Section 3.3 are supported by equation (15). In particular, it is straightforward to find levels of ψ and ς that maximize utility and compare them to those which maximize output.

To derive equations (13) - (15) note that the agent's problem is to maximize equation (4) subject to the constraints in equation (5) and the relationships in equations (1) and (2). We impose the last two lines of equation (5) to arrive at the following Lagrangian:

$$\begin{aligned}
\mathcal{L} & = \ln c_3 + \beta \ln c_4 + \xi \ln \bar{A} w \left(f_1 + g_1\right)^\gamma \left(f_2 + g_2\right)^{\gamma_2} \\
& \quad + \lambda \left(I\left(1 - \tau\right) - c_3 - c_4 - f_1 - f_2\right).
\end{aligned}$$

The structure of the problem assures that the first line of equation (5) will hold with equality and that the non-negativity constraints in the second line of equation (5) will not bind in equilibrium. However, the non-negativity constraints in the third line may bind so we write the Kuhn-Tucker conditions as

$$c_3 \quad : \quad \frac{1}{c_3} - \lambda = 0 \tag{16a}$$

$$c_4 \quad : \quad \frac{\beta}{c_4} - \lambda = 0 \tag{16b}$$

$$f_1 \quad : \quad \frac{\xi\gamma}{f_1+g_1} - \lambda \leq 0, \quad f_1 \geq 0, \quad \text{and} \quad \left(\frac{\xi\gamma}{f_1+g_1} - \lambda\right) f_1 = 0 \tag{16c}$$

$$f_2 \quad : \quad \frac{\xi\gamma_2}{f_2+g_2} - \lambda \leq 0, \quad f_2 \geq 0, \quad \text{and} \quad \left(\frac{\xi\gamma_2}{f_2+g_2} - \lambda\right) f_2 = 0 \tag{16d}$$

$$\lambda \quad : \quad I\left(1 - \tau\right) - c_3 - c_4 - f_1 - f_2 = 0. \tag{16e}$$

There are four cases to consider.

Let $f = (f_1^*, 0)$. Equations (16a)-(16c) into equation (16e) and the assumption $f_2 = 0$ give

$$c_3 = \frac{I(1-\tau)+g_1}{\beta_2}, \quad c_4 = c_3 \beta, \quad f_1 = c_3 \xi \gamma - g_1, \quad f_2 = 0. \tag{17}$$

Let $f = (f_1^*, f_2^*)$. Equations (16a)-(16d) into equation (16e) gives

$$c_3 = \frac{I(1-\tau)+g_1+g_2}{\beta_1}, \quad c_4 = c_3 \beta, \quad f_1 = c_3 \xi \gamma - g_1, \quad f_2 = c_3 \xi \gamma_2 - g_2. \tag{18}$$

31

Let $f = (0,0)$. Equations (16a) and (16b) into equation (16e) and the assumption $f_1 = f_2 = 0$ give

$$c_3 = \frac{I(1-\tau)}{1+\beta}, \quad c_4 = c_3\beta, \quad f_1 = 0, \quad f_2 = 0. \tag{19}$$

Let $f = (0, f_2^*)$. Equations (16a), (16b), and (16d) into equation (16e) and the assumption $f_1 = 0$ give

$$c_3 = \frac{I(1-\tau)+g_2}{\beta_3}, \quad c_4 = c_3\beta, \quad f_1 = 0, \quad f_2 = c_3\xi\gamma_2 - g_2. \tag{20}$$

With $r = 1$, from equations (3), (6), and (9) and the equilibrium conditions that $H = H_2$ and $w = Z$ we have

$$Y = I = 2wh_2. \tag{21}$$

From equations (6), (7), (8), and (21) we have

$$\tau = \varsigma, \quad g_1 = \varsigma\psi I, \quad g_2 = \varsigma(1-\psi)I . \tag{22}$$

Next, using the third and fourth items in equations (17)-(20) along with equations (10) and (22) in equation (11) gives

$$h_2 = \begin{cases} \bar{A}(c_3\xi\gamma)^\gamma (\varsigma(1-\psi)I)^{\gamma_2} & \text{if} \quad f = (f_1^*, 0) \\ \bar{A}(c_3\xi\gamma)^\gamma (c_3\xi\gamma_2)^{\gamma_2} & \text{if} \quad f = (f_1^*, f_2^*) \\ \bar{A}(\varsigma\psi I)^\gamma (\varsigma(1-\psi)I)^{\gamma_2} & \text{if} \quad f = (0, 0) \\ \bar{A}(\varsigma\psi I)^\gamma (c_3\xi\gamma_2)^{\gamma_2} & \text{if} \quad f = (0, f_2^*). \end{cases} \tag{23}$$

Equations (17)-(20) and (22) give

$$c_3 = \begin{cases} I(1-\varsigma(1-\psi))\beta_2^{-1} & \text{if} \quad f = (f_1^*, 0) \\ I\beta_1^{-1} & \text{if} \quad f = (f_1^*, f_2^*) \\ I(1-\varsigma)(1+\beta)^{-1} & \text{if} \quad f = (0, 0) \\ I(1-\varsigma\psi)\beta_3^{-1} & \text{if} \quad f = (0, f_2^*). \end{cases} \tag{24}$$

Using equations (21) and (24) in equation (23) gives

$$h_2 = \begin{cases} \left(\bar{A}(2w)^{\gamma+\gamma_2} \left(\xi\gamma(1-\varsigma(1-\psi))\beta_2^{-1}\right)^\gamma (\varsigma(1-\psi))^{\gamma_2}\right)^{\frac{1}{1-\gamma-\gamma_2}} & \text{if} \quad f = (f_1^*, 0) \\ \left(\bar{A}(2w)^{\gamma+\gamma_2} (\xi\gamma)^\gamma (\xi\gamma_2)^{\gamma_2} \beta_1^{-(\gamma+\gamma_2)}\right) & \text{if} \quad f = (f_1^*, f_2^*) \\ \left(\bar{A}(2w)^{\gamma+\gamma_2} (\varsigma\psi)^\gamma (\varsigma(1-\psi))^{\gamma_2}\right)^{\frac{1}{1-\gamma-\gamma_2}} & \text{if} \quad f = (0, 0) \\ \left(\bar{A}(2w)^{\gamma+\gamma_2} (\varsigma\psi)^\gamma \left((1-\varsigma\psi)\beta_3^{-1}\xi\gamma_2\right)^{\gamma_2}\right)^{\frac{1}{1-\gamma-\gamma_2}} & \text{if} \quad f = (0, f_2^*). \end{cases} \tag{25}$$

Using this in equation (21) and simplifying gives equation (14).

Consider circumstances under which equilibrium types exist.

Let $f = (f_1^*, 0)$. Putting equation (16a) into equation (16d), we see that $f_2 = 0$ if $c_3 \leq \frac{g_2}{\xi \gamma_2}$. From the third item in equation (17), $f_1 \geq 0$ requires $c_3 \geq \frac{g_1}{\xi \gamma}$. Using equation (8) and the first line of equation (24) along with $Y = I$, these constraints can be written as

$$\frac{\varsigma \psi I}{\xi \gamma} \leq \frac{I(1 - \varsigma) + I \varsigma \psi}{\beta_2} \leq \frac{\varsigma (1 - \psi) I}{\xi \gamma_2}.$$

Solving for ψ, this can be rewritten to give the first line of equation (13).

Let $f = (f_1^*, f_2^*)$. From equation (18), $f_1 \geq 0$ and $f_2 \geq 0$ requires

$$c_3 \geq \max \left(\frac{g_1}{\xi \gamma}, \frac{g_2}{\xi \gamma_2} \right)$$

and using equation (8) and the second line of equation (24) this is

$$I \beta_1^{-1} \geq \max \left(\frac{\varsigma \psi I}{\xi \gamma}, \frac{\varsigma (1 - \psi) I}{\xi \gamma_2} \right).$$

Solving for ψ, this can be rewritten to give the second line of equation (13).

Let $f = (0, 0)$. Putting equation (16a) into equations (16c) and (16d), we see that $f_2 = 0$ if $c_3 \leq \frac{g_2}{\xi \gamma_2}$ and $f_1 = 0$ if $c_3 \leq \frac{g_1}{\xi \gamma}$. Using equation (8) and the third line of equation (24), these constraints can be written as

$$\frac{I(1 - \varsigma)}{1 + \beta} \leq \min \left[\frac{\varsigma \psi I}{\xi \gamma}, \frac{\varsigma (1 - \psi) I}{\xi \gamma_2} \right].$$

Solving for ψ, this can be rewritten to give the third line of equation (13).

Let $f = (0, f_2^*)$. Putting equation (16a) into equation (16c), we see that $f_1 = 0$ if $c_3 \leq \frac{g_1}{\xi \gamma}$. From the fourth item in equation (20), $f_2 \geq 0$ requires $c_3 \geq \frac{g_2}{\xi \gamma_2}$. Using equation (8) and the fourth line of equation (24), these constraints can be written as

$$\frac{\varsigma (1 - \psi) I}{\xi \gamma_2} \leq I(1 - \varsigma \psi) \beta_3^{-1} \leq \frac{\varsigma \psi I}{\xi \gamma}.$$

Solving for ψ, this can be rewritten to give the fourth line of equation (13). It is straightforward to show that conditions allowing the four cases are mutually exclusive.

Finally, consider utility. From $\hat{I} = I$, equation (4) and equations (18)-(19), $U_j = \ln \beta^\beta c_3^{1+\beta} I^\xi$. Thus equation (15) follows directly from equation (24). $\bar{U} = \exp(U_j) \beta^{-\beta}$ which is a monotonic transformation.

Proof of Proposition 3. For brevity, we provide only a sketch of the proof. Throughout, we consider $\tilde{I} = (\bar{A}w)^{-1} I^{1 - \gamma - \gamma_2}$ rather than I with no loss of generality. From equation (13), the

following equilibria exist for some value of $\psi \in [0,1]$ given the values of ς :

$$
f = \begin{cases}
(f_1^*, f_2^*) & \text{if} & \varsigma \leq \frac{\xi\gamma_2}{\beta_1} \\
(f_1^*, f_2^*), (f_1^*, 0) & \text{if} & \frac{\xi\gamma_2}{\beta_1} \leq \varsigma \leq \frac{\xi\gamma}{\beta_1} \\
(f_1^*, f_2^*), (f_1^*, 0), (0, f_2^*) & \text{if} & \frac{\xi\gamma}{\beta_1} \leq \varsigma \leq \frac{\xi(\gamma+\gamma_2)}{\beta_1} \\
(0, 0), (f_1^*, 0), (0, f_2^*) & \text{if} & \varsigma \geq \frac{\xi(\gamma+\gamma_2)}{\beta_1}.
\end{cases}
$$

Define $\tilde{I}_{f_1^*,0}$ to be output in $f = (f_1^*, 0)$ when ψ is chosen to locally maximize output. Stated differently, it is the maximum output over the range of ψ supporting $f = (f_1^*, 0)$ given ς. Output is maximized over this range either at $\psi = 0$ or where $\frac{\partial I}{\partial \psi} = 0$, with I given by the first line of equation (14). From choosing the output maximizing level of ψ in equation (14) we find

$$
\tilde{I}_{f_1^*,0} = \begin{cases}
\left(\xi\gamma(1-\varsigma)\beta_2^{-1} \right)^\gamma \varsigma^{\gamma_2} & \text{if} \quad \varsigma \leq \frac{\gamma_2}{\gamma+\gamma_2} \\
\left(\xi\frac{\gamma^2}{\gamma+\gamma_2}\beta_2^{-1} \right)^\gamma \left(\frac{\gamma_2}{\gamma+\gamma_2} \right)^{\gamma_2} & \text{if} \quad \varsigma \geq \frac{\gamma_2}{\gamma+\gamma_2}.
\end{cases} \tag{26}
$$

Similarly

$$
\tilde{I}_{0,f_2^*} = \begin{cases}
\varsigma^\gamma \left((1-\varsigma)\beta_3^{-1}\xi\gamma_2 \right)^{\gamma_2} & \text{if} \quad \varsigma \leq \frac{\gamma}{\gamma+\gamma_2} \\
\left(\frac{\gamma}{\gamma+\gamma_2} \right)^\gamma \left(\frac{\gamma_2^2}{\gamma+\gamma_2}\beta_3^{-1}\xi \right)^{\gamma_2} & \text{if} \quad \varsigma \geq \frac{\gamma}{\gamma+\gamma_2},
\end{cases} \tag{27}
$$

$$
\tilde{I}_{0,0} = \left(\frac{\varsigma\gamma}{\gamma+\gamma_2} \right)^\gamma \left(\frac{\varsigma\gamma_2}{\gamma+\gamma_2} \right)^{\gamma_2}, \tag{28}
$$

and

$$
\tilde{I}_{f_1^*,f_2^*} = (\xi\gamma)^\gamma (\xi\gamma_2)^{\gamma_2} \beta_1^{-(\gamma+\gamma_2)}. \tag{29}
$$

Each is continuous. The first two are initially increasing in ς and level out at $\varsigma = \frac{\gamma_2}{\gamma+\gamma_2}$ and $\frac{\gamma}{\gamma+\gamma_2}$.

The third is increasing in ς always and the fourth is independent of ς.

Consider starting with $\varsigma = 0$ and increasing ς. Initially output is globally maximized at $f = (f_1^*, f_2^*)$ since only this equilibrium exists. When $\tilde{I}_{f_1^*,0}$ comes into existence at $\varsigma = \frac{\xi\gamma_2}{\beta_1}$, $\tilde{I}_{f_1^*,0} = \tilde{I}_{f_1^*,f_2^*}$ and the ratio of $\tilde{I}_{f_1^*,0}$ to $\tilde{I}_{f_1^*,f_2^*}$ is increasing in ς. Thus beginning here, $f = (f_1^*, 0)$ is optimal and beyond this value of ς, $f = (f_1^*, f_2^*)$ can not be globally optimal.

When $\tilde{I}_{0,f_2^*}$ comes into existence at $\varsigma = \frac{\xi\gamma}{\beta_1}$, $\tilde{I}_{0,f_2^*} < \tilde{I}_{f_1^*,0}$. This is because at this value $\tilde{I}_{0,f_2^*} = \tilde{I}_{f_1^*,f_2^*}$ and $\tilde{I}_{f_1^*,0} > \tilde{I}_{f_1^*,f_2^*}$. Also, the ratio of $\tilde{I}_{0,f_2^*}$ to $\tilde{I}_{f_1^*,0}$ is increasing in ς. At their maximum values $\tilde{I}_{0,f_2^*} > \tilde{I}_{f_1^*,0}$. To see this, put $\varsigma = \frac{\gamma_2}{\gamma+\gamma_2}$ into the first line of equation (26) and $\varsigma = \frac{\gamma}{\gamma+\gamma_2}$ into the first line of equation (27) and compare. This is sufficient to show that $\tilde{I}_{0,f_2^*} = \tilde{I}_{f_1^*,0}$ at one value of ς. Call it $\underline{\varsigma_y}$. Beyond $\underline{\varsigma_y}$, $f = (f_1^*, 0)$ cannot be globally optimal.

When $\tilde{I}_{0,0}$ comes into existence at $\varsigma = \frac{\xi(\gamma+\gamma_2)}{\beta_1}$, $\tilde{I}_{0,0} < \tilde{I}_{0,f_2^*}$. This is because at this value $\tilde{I}_{0,0} = \tilde{I}_{f_1^*,f_2^*}$ and $\tilde{I}_{0,f_2^*} > \tilde{I}_{f_1^*,f_2^*}$. Also, the ratio of $\tilde{I}_{0,0}$ to $\tilde{I}_{0,f_2^*}$ is increasing in ς. At their maximum

values $\tilde{I}_{0,0} > \tilde{I}_{0,f_2^*}$. To see this, put $\varsigma = \frac{\gamma}{\gamma+\gamma_2}$ into the first line of equation (27) and $\varsigma = 1$ into the first line of equation (28) and compare. This is sufficient to show that $\tilde{I}_{0,0} = \tilde{I}_{0,f_2^*}$ at one value of ς. Call it $\overline{\varsigma_y}$. Beyond $\overline{\varsigma_y}$, $f = (0, f_2^*)$ cannot be globally optimal.

We have shown that in the range $\varsigma \in \left(\frac{\xi\gamma_2}{\beta_1}, \underline{\varsigma_y} \right)$, $\tilde{I}_{f_1^*,0} > \tilde{I}_{f_1^*,f_2^*}, \tilde{I}_{0,f_2^*}$. To assure a global maximum, we need to show that $\tilde{I}_{f_1^*,0} > \tilde{I}_{0,0}$ in this range. This is done in a supplement available from the authors. This completes the sketch of the proof.

7 Appendix 2

In this appendix we present the dynamics of the model. We then show numerically that the steady state analysis of the paper is relevant in the sense that the economy converges to this unique steady state equilibrium.

7.1 The economy at $t = 0$.

Consider the initial period $t = 0$. The economy is populated by original old workers denoted by subscript $t - 3$, initial young workers denoted by subscript $t - 2$, initial children in late childhood denoted by subscript $t - 1$, and initial children in early childhood denoted by subscript $t = 0$. We write t rather than 0 to facilitate a succinct introduction of subsequent time periods.

Each member of generation $t - 3$ has an endowment of human capital and is the parent of a member of generation $t - 1$. The child is in late childhood with the same ability index as the parent and with an endowment of early childhood human capital. For initial old worker j, the optimization problem is

$$\mathcal{L} = \beta \ln c_{t,t-3,j} + \xi \ln I_{t+1,t-1,j} + \lambda \left(wh_{t,t-3,j} (1 - \tau) - c_{t,t-3,j} - f_{t,t-1,j} \right). \tag{30}$$

Here $c_{t,t-3,j}$ is consumption in period t by agent j of generation $t - 3$. Similarly $I_{t+1,t-1,j}$ is the present discounted lifetime income measured at period $t + 1$ of agent j of generation $t - 1$. With $t = 0$, it is the income of the agent j who is in late childhood as the economy begins. The human capital endowment of agent j of generation $t - 3$ is $h_{t,t-3,j}$ so $wh_{t,t-3,j}$ is income. Our notation on education items refers to the recipients rather than the providers. Hence $f_{t,t-1,j}$ is family expenditure in period t on agent j of generation $t - 1$. With $t = 0$, this is expenditure by original old worker j on her offspring when the offspring is in late childhood.

From the first order conditions and the budget constraint, we arrive at

$$I_{t+1,t-1,j}\beta = \xi c_{t,t-3,j} \frac{\partial I_{t+1,t-1,j}}{\partial h_{t+1,t-1,j}} \frac{\partial h_{t+1,t-1,j}}{\partial i_{t,t-1,j}} \frac{\partial i_{t,t-1,j}}{\partial f_{t,t-1,j}}$$

$$c_{t,t-3,j} = wh_{t,t-3,j}(1-\tau) - f_{t,t-1,j} \tag{31}$$

where

$$I_{t+1,t-1,j} = wh_{t+1,t-1,j}\left(1+r^{-1}\right)$$

$$h_{t+1,t-1,j} = Aa_j\left[\gamma_2 i_{t,t-1,j}^{\rho} + (1-\gamma_2)h_{t,t-1,j}^{\rho}\right]^{\frac{1}{\rho}} \tag{32}$$

$$i_{t,t-1,j} = B\left[\alpha f_{t,t-1,j}^{\eta} + (1-\alpha)g_{t,t-1}^{\eta}\right]^{\frac{1}{\eta}}.$$

The derivatives follow directly from equation (32).[17] The human capital of the initial late childhood student j upon entering the labor market in the subsequent period $(t+1)$ is given by $h_{t+1,t-1,j}$. It is a combination of investment on her behalf in period t, $i_{t,t-1,j}$, and the human capital this person had as the period began, $h_{t,t-1,j}$. For the original late childhood agents, $h_{t,t-1,j}$ is an exogenous endowment. In turn, $i_{t,t-1,j}$ is a combination of family spending on behalf of the student in late childhood and government spending on behalf of the agent in late childhood, $g_{t,t-1}$. Government spending is common across all agents and so has no j subscript. Knowledge of this is required for the initial old workers to solve equation (30).

The initial young workers and old workers both pay labor taxes at rate τ to determine government spending. As with the original old workers, their human capital, $h_{t,t-2,j}$, is an exogenous endowment. Given a discrete number of agents, M, in each generation, we can calculate $g_{t,t-1}$ by

$$g_{t,t-1} = \frac{\varsigma(1-\psi)}{M}\left(\sum_{j=1}^{M}wh_{t,t-3,j} + \sum_{j=1}^{M}wh_{t,t-2,j}\right). \tag{33}$$

Since M can be set arbitrarily large, this is a good approximation to the continuum of agents in the body of the paper.

It is straightforward to numerically solve the problem of the original old. We choose parameters and generate initial vectors of $h_{t,t-1,j}$, $h_{t,t-3,j}$, and $h_{t,t-2,j}$ with no zero values. We use equation (33) to find $g_{t,t-1}$. For each original old, we take a guess at $f_{t,t-1,j}$, use the second line of equation (31) to find $c_{t,t-3,j}$, and iterate on $f_{t,t-1,j}$ until the first line of equation (31) holds. For this we use the optmum procedure in Gauss. We then use equation (32) to arrive at vectors of $I_{t+1,t-1,j}$ and $h_{t+1,t-1,j}$.

Consider now the problem for the initial young workers who have children in early childhood

---

[17]With $\eta = 1$, we need to consider the cases where some families spend nothing on early or late childhood. For brevity, we do not develop that case here.

education. Their optimization problem is

$$\mathcal{L} = \ln c_{t,t-2,j} + \beta \ln c_{t+1,t-2,j} + \xi \ln I_{t+2,t,j}$$
$$+ \lambda \left(I_{t,t-2,j} \left(1 - \tau \right) - c_{t,t-2,j} - c_{t+1,t-2,j} r^{-1} - f_{t,t,j} - f_{t+1,t,j} r^{-1} \right). \tag{34}$$

Items $c_{t,t-2,j}$ and $c_{t+1,t-2,j}$ are consumption as a young and old worker, $I_{t,t-2,j}$ and $I_{t+2,t,j}$ are the lifetime income of the agent and her offspring, and $f_{t,t,j}$ and $f_{t+1,t,j}$ are spending by the agent on offspring in early and late childhood. First order conditions and the budget constraint reduce to

$$\frac{\partial h_{t+2,t,j}}{\partial i_{t+1,t,j}} \frac{\partial i_{t+1,t,j}}{\partial f_{t+1,t,j}} r = \frac{\partial h_{t+2,t,j}}{\partial h_{t+1,t,j}} \frac{\partial h_{t+1,t,j}}{\partial i_{t,t,j}} \frac{\partial i_{t,t,j}}{\partial f_{t,t,j}}$$
$$I_{t+2,t,j} = c_{t,t-2,j} \xi r w \left(1 + r^{-1} \right) \frac{\partial h_{t+2,t,j}}{\partial i_{t+1,t,j}} \frac{\partial i_{t+1,t,j}}{\partial f_{t+1,t,j}} \tag{35}$$
$$c_{t+1,t-2,j} = \left(1 + \beta \right)^{-1} \left(I_{t,t-2,j} \left(1 - \tau \right) - f_{t,t,j} - f_{t+1,t,j} r^{-1} \right)$$

where $I_{t+2,t,j}$, $h_{t+2,t,j}$, and $i_{t+1,t,j}$ are defined by adjusting items in equation (32) forward one period. Additionally

$$h_{t+1,t,j} = a_j i_{t,t,j}^{\gamma_1}$$
$$i_{t,t,j} = B \left[\alpha f_{t,t,j}^{\eta} + \left(1 - \alpha \right) g_{t,t}^{\eta} \right]^{\frac{1}{\eta}} \tag{36}$$
$$g_{t,t} = \frac{\varsigma \psi}{M} \left(\sum_{j=1}^{M} w h_{t,t-3,j} + \sum_{j=1}^{M} w h_{t,t-2,j} \right).$$

From updating the third line of equation (32) we see that the original young worker needs to know $g_{t+1,t}$ in order to make decisions. That is, the original young worker needs to know per student spending on late childhood education in the next period when her offspring is in late childhood. Since we solved above for human capital investment by the original old, agents know what next period's income will be for those originally in late childhood. Agents also know their own income next period and so can find

$$g_{t+1,t} = \frac{\varsigma \left(1 - \psi \right)}{M} \left(\sum_{j=1}^{M} w h_{t+1,t-2,j} + \sum_{j=1}^{M} w h_{t+1,t-1,j} \right). \tag{37}$$

It is therefore straightforward to solve the problem of the original young workers. We have parameters and generated vectors of $h_{t,t-1,j}$, $h_{t,t-3,j}$, and $h_{t,t-2,j}$ with no zero values and have solved for $h_{t+1,t-1,j}$. For each original young worker, we take a guess at $f_{t,t,j}$ and $f_{t+1,t,j}$, use the third line of (35) to find $c_{t+1,t-2,j}$, and iterate on $f_{t,t,j}$ and $f_{t+1,t,j}$, until the first and second lines of (35) hold. For this we use the optmum procedure in Gauss. We then use equations (32) and (36) to arrive at vectors of the other endogenous items.

7.2 The economy at $t > 0$.

At time $t = 0$ the initial old and young workers each optimize. In subsequent periods we can think of young workers making decisions and old workers simply implementing the decisions made in the

prior period. Thus in each period we solve a problem analogous to equation (34) for each young worker. The only difference is that above the human capital of decision makers is an exogenous endowment. In subsequent periods, it is an endogenous response to the circumstances faced by the agent's parents. Since it is has been solved for, however, the math proceeds as above. In each period, the young worker needs to know both current government education spending and spending in the next period. This requires knowing both own income in the next period and the income next period of today's children in late childhood. As above, both of these are determined at the time of the agent's decision.

A couple of things regarding notation can be confusing and are thus worth noting. First, each dynasty is making decisions every other period. That is, the young workers are deciding on a spending pattern taking income and government spending as given. They are spending on their children in early childhood. The older workers are also spending on their children who are in late childhood. However, their optimization problem was solved in the preceding period and is simply carried out in the current period. While dynasties differ in the timing of their decisions, the ability distribution of the two types of dynasties are equivalent. Thus we use the same notation for the two types of generations and in particular an agent with the ability associated with index j is born in each period.

Second, the notation $h_{t+2,t,j}$ indicates the human capital of an agent as she enters the labor market. It is human capital on hand in the agent's third period of life. In the initial period, we have $h_{t,t-3,j}$ for the initial old. The two items in the subscript of $h_{t+2,t,j}$ are separated by two periods. The two items in the subscript of $h_{t,t-3,j}$ are separated by three periods. This can be reconciled by noting that the agent's human capital is the same across the two periods of working so $h_{t+3,t,j} = h_{t+2,t,j}$.

7.3 Convergence to the steady state.

To demonstrate existence and uniqueness of the equilibrium in the general case, we show numerically that the economy converges to a steady state for an arbitrary initial distribution $h_{t,t-1}$, $h_{t,t-3}$, and $h_{t,t-2}$. The lack of j subscripts here indicates that we are referring to $M \times 1$ vectors. Setting $t = 0$ in the initial period, these are $h_{0,-1}$, $h_{0,-2}$, and $h_{0,-3}$. We set parameters consistent with Section 4: $\varsigma = .03$, $\gamma_1 = .3$, $\gamma_2 = .15$, $\beta = \xi = \eta = r = a = \omega = b = 1$, $\alpha = \psi = .5$, $\rho = 0$, and a_j is uniform over $[.5, 5]$.

We let each element of initial vectors $h_{0,-1}$, $h_{0,-2}$, and $h_{0,-3}$ be an independent random draw from a uniform distribution from zero to 1. We then follow the algorithm above to find $h_{1,-1}$, and $h_{2,0}$ in period 0. With these, we can find $h_{3,1}$, and then $h_{4,2}$, etc. We continue in this fashion until $h_{t+2,t} = h_{t'+2,t'}$ for all $t' > t$. From these vectors we calculate the time series of output, $I_{t+2,t}$. We then repeat the exercise with $\varsigma = 0$. For comparison with the first panel of Figure 6, we normalize output with $\varsigma = .03$ in each period by the value it takes with $\varsigma = 0$ upon convergence.

Figure 7 shows the results. As in Figure 6 the horizontal axis is descending values of $a(j) \in [.5, 5]$. That is, the higher income dynasties for whom ς_j is smaller are represented further to the left. The dotted curve represents normalized $I_{0,-2}$. Each item of $I_{0,-2}$ is a function of levels of human capital which were drawn from the uniform distribution. The upward trend is due to the normalization by ever smaller steady state incomes. The short dashed curve shows normalized $I_{5,3}$; i.e., it shows the distribution after five periods. The long dashed curve and solid curve show $I_{10,8}$ and $I_{20,18}$. Beyond this, the distribution has converged.

The program is written with sufficient generality to handle each case specified in the paper. We choose the above parametrization so that upon convergence, the curve aligns with that in the first panel of Figure 6 for $\psi = .5$. With different parameters, results change in an intuitive way consistent with the main text of the paper and convergence continues to hold. The general result of convergence is robust. We conduct a thorough sensitivity analysis of this result (available upon request), varying each of the parameters considerably and find convergence in each case.
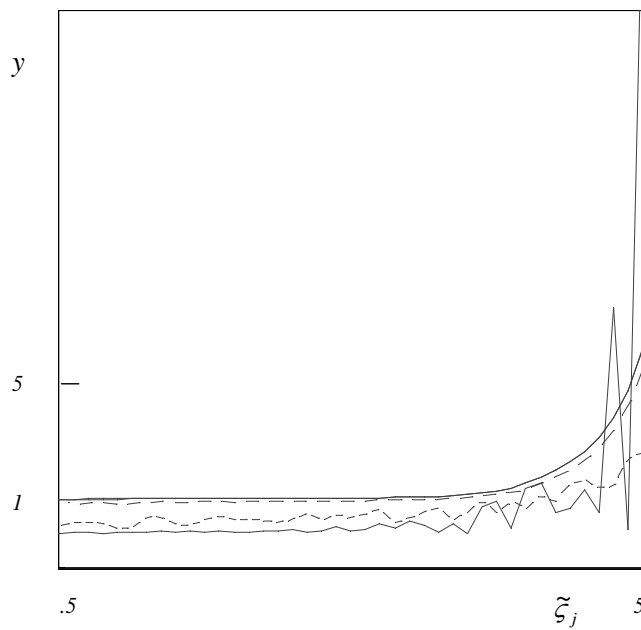
Figure 7: The jagged solid curve shows the initial distribution of normalized human capital. The short dashed curve shows the distribution after five periods. The long dashed curve shows the distribution after ten periods and the smoother solid curve shows the distribution after twenty or more periods.