Government Education Expenditures in Early and Late Childhood*

Casey Abington†
Kansas State University

William Blankenau‡
Kansas State University

December 29, 2009

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

*We would like to thank Peter Rangazas for helpful comments on an early draft of this paper. Much of the paper was written while Blankenau was a visiting professor in the Department of Economics at Universidad Carlos III de Madrid. He would like to acknowledge the generous hospitality of the department. We also benefited from the insights of seminar participants at the 2009 Public Economic Theory Meeting at National University of Ireland, Universidad Carlos III de Madrid, Universidad del País Vasco, Aarhus University, Fundación de Estudios de Economía Aplicada, Durham University, and the 2008 Southern Economic Association meeting in Washington D.C.

†Department of Economics, 327 Waters Hall, Kansas State University, Manhattan, KS 66506.
‡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

¹See also Carneiro and Heckman (2003), Knudsen, Heckman, Cameron, and Shonkoff (2006), and Currie (2001(b)).
²See, for example, Driskill and Horowitz (2002), Su (2004), Kaganovich (2005), Blankenau (2005), and Restuccia and Urrutia (2004).
³See Jensen (1980) and the discussions in CHLM and Cunha and Heckman (2007).
⁴These numbers are not reported directly. However, table B2.2 of OECD Education at a Glance (2007) states that .4% of GDP 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 day-care, 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 crowds out some early childhood spending. 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. Section 4 demonstrates this point by showing that the results are little changed in a more general case preserving homogeneity. Section 5 considers heterogeneity. Section 6 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. 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 \( j \in J \equiv [0,1] \). 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 heritability of productivity, recent evidence suggests considerable dynastic persistence in relative

\[ \text{In an earlier version of this paper, young adults also made a choice to attend college or not. This proves unimportant for our main points.} \]
earnings. For example, Mazumder (2005) estimates the intergenerational elasticity in earnings to be about .6.\footnote{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.} In our model, inheritance of $a$ is the channel through which such persistence arises.

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,

$$
h_1 = a i_1^{\gamma_1}$$

$$
h_2 = \begin{cases} 
Aa [\gamma_2 i_2^\rho + (1 - \gamma_2) h_1^\rho]^{\frac{1}{\rho}} & \text{if } \rho \neq 0 \\
A a i_1^{\gamma_2} h_1^{1-\gamma_2} & \text{if } \rho = 0
\end{cases}
$$

where $\gamma_1, \gamma_2 \in [0, 1]$ with $\min[\gamma_1, \gamma_2] < 1$, $\rho \leq 1$, 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 $\gamma_1$ and $\gamma_2$ govern the curvature of the functions. The parameter $\rho$ 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

$$
i_k = \begin{cases} 
B [\alpha f_k^\eta + (1 - \alpha) g_k^\eta]^{\frac{1}{\eta}} & \text{if } \eta \neq 0 \\
B f_k^{\alpha} g_k^{1-\alpha} & \text{if } \eta = 0
\end{cases}
$$

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

2.2 The agents’ 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. With an interest rate exogenously given by \( r \), the present value of lifetime income is

\[
I = wh_2 \left( 1 + r^{-1} \right).
\]  

(3)

In modeling education choices, it is common to consider the possibility of borrowing constraints. Such constraints play a key role in a wide variety of recent research. Some examples are Rangazas (2002) and Restuccia and Urratia (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, still, to impose credit constraints for those who spend significantly on children in early childhood. This is likely to be of modest importance.

We will use \( \hat{\cdot} \) 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

\[
U_j = \ln c_3 + \beta \ln c_4 + \xi \ln \hat{I}.
\]  

(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 \( \xi \) 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 \( \tau \) 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.
and
\[ I (1 - \tau) \geq c_3 + \frac{c_4}{\tau} + f_1 + \frac{f_2}{\tau}, \]
\[ c_3, c_4 \geq 0, \]
\[ f_1, f_2 \geq 0, \]
\[ I = I \left( \hat{h}_2 \right), \quad \hat{h}_2 = h_2 (\hat{i}_1, \hat{i}_2), \]
\[ \hat{i}_1 = \hat{i}_1 (f_1, g_1) \quad \text{and} \quad \hat{i}_2 = \hat{i}_2 (f_2, g_2). \]  

\[ (5) \]

2.3 Other entities.

A large number of identical firms employ labor to produce identical consumption goods according to
\[ Y = ZH \]  
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 \( \tau \) 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
\[ G = g_1 + g_2 = ZH \tau \]  
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 \( \psi \) to be the share of \( G \) that is devoted to early childhood education. Thus we set
\[ G = \varsigma Y, \quad \varsigma \psi Y = g_1, \quad \varsigma (1 - \psi) Y = g_2 \]  
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.

The dynamics of the model are simple to trace. 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

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

(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 \( \{I_j, c_{3,j}, c_{4,j}, f_{1,j}, f_{2,j}, h_{1,j}, h_{2,j}, \hat{I}_j, \hat{h}_{1,j}, \hat{h}_{2,j}\} \) \( \forall j \in J \), 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.\(^9\)
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 maintain the following assumption throughout this section:

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

Setting $\eta = 1$, $\alpha = .5$ makes government spending perfectly substitutable with private spending and

$$i_k = f_k + g_k.$$  

(10)

Setting $\rho = 0$ simplifies the human capital expression to

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

(11)

\(^9\)Implications of the model are qualitatively robust to the closed economy case where the goods market clears.
where \( \gamma = (1 - \gamma_2) \gamma_1 \) and \( \bar{A} = A a^{2-\gamma_2} \). Setting \( r = 1 \) 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 and the next, we also assume that all agents are of equal ability. This requires Assumption 2:

**Assumption 2:** \( a_j = a_{j'} \forall j \) and \( j' \in J \).

### 3.1 Equilibria.

We show below that for any choice of parameters in the steady state, a unique equilibrium exists. This equilibrium can be one of four types, depending on family education expenditures. In both early childhood and late childhood, family spending can be zero or positive. To distinguish the types, we use the notation \( f = (f_1^*, f_2^*) \) to indicate that family spending is positive at both stages and \( f = (0, 0) \) to indicate zero spending at both stages. Similarly \( f = (f_1^*, 0) \) means that there is positive family spending only on early childhood while \( f = (0, f_2^*) \) means positive spending only on late childhood. With this, we are ready to state Proposition 1:

**Proposition 1.** If Assumptions 1 and 2 hold, then

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

where \( \beta_1 \equiv 1 + \beta + (\gamma_2 + \gamma) \xi \).

Proposition 1 divides the \( \varsigma \in [0, 1] \times \psi \in [0, 1] \) space into four regions, each permitting exactly one of the four types of equilibria. At the border between any two regions, both types of equilibria are supported but little gained in discussing this knife-edge case and we hereafter omit it. The first line of equation (12) shows that for \( \psi \) sufficiently small, families spend on early childhood but not on late childhood education. The last line shows that for \( \psi \) sufficiently large, families spend on late childhood education but not on early childhood. The second and third lines show that for intermediate values of \( \psi \), families spend at either both or neither level of education. In each case, the cutoff points between equilibria types depend on the level of spending.

Figure 1 serves as an example. Here we show the partition of the \( \varsigma \times \psi \) space for a particular parameterization. We set \( \gamma_2 = .15 \), which is in the range used by Blankenau and Simpson (2004).
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.

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 $\xi = 1 + \beta$.

To see how $\psi$ and $\varsigma$ jointly determine the type of equilibrium, it is useful to consider three values of $\varsigma$. 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 $\psi$ value, $f = (f_1^*, f_2^*)$. Thus when government spending is low, its allocation does not influence the type of equilibrium. Regardless of the allocation of spending, families top-up government spending at both levels. 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 $\psi$ small $f = (f_1^*, 0)$, for $\psi$ large $f = (0, f_2^*)$, and otherwise $f = (f_1^*, f_2^*)$. 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 $f = (0, 0)$. 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.
The analysis with $\varsigma = \varsigma_1$ is valid whenever $\varsigma \leq \min \left[ \xi \gamma_2 \beta_1^{-1}, \xi \gamma \beta_1^{-1} \right]$ and the analysis with $\varsigma = \varsigma_3$ is valid whenever $\varsigma \geq \xi (\gamma + \gamma_2) \beta_1^{-1}$. Otherwise the analysis with $\varsigma = \varsigma_2$ is valid. We can use this to formalize the definitions of high, moderate, and low spending.

**Definition 2.** 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.

Furthermore, we can think of government spending as focused on a stage of education when it fully crowds out private spending at exactly one stage of education. In contrast, when spending is balanced, families spend at both or neither stages, depending on the level of government spending. With this, we can state Corollary 1.

**Corollary 1.** If government spending is low, families always spend on both stages of education. Otherwise, if government spending is focused on one stage of education, families spend only on the other stage. If government spending is balanced, families spend on both stages with moderate government spending and on neither stage with high government spending.

Notice that for moderate levels of spending, the range of $\psi$ values considered balanced spending decreases with $\varsigma$ (i.e. the two bounds are getting closer together). In contrast, when government spending is high, the range is increasing. The intuition for this result is simple. Moderate spending is balanced when both $\varsigma \psi$ and $\varsigma (1 - \psi)$ are small enough that families top-up government spending. Clearly it is easier to satisfy the conditions simultaneously when $\varsigma$ decreases. High spending is balanced when there is sufficient government spending at both levels to fully crowd out both levels of private spending (i.e. when both $\varsigma \psi$ and $\varsigma (1 - \psi)$ are large enough). It is easier to satisfy the conditions simultaneously when $\varsigma$ increases.

### 3.2 Output.

The above discussion clarifies how the equilibrium type depends on the spending level and its allocation. We now consider how these government choices affect output within an equilibrium type. Proposition 2 gives the main result.

**Proposition 2.** Income is related to government policy according to

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

where $\beta_2 \equiv 1 + \beta + \gamma \xi$ and $\beta_3 \equiv 1 + \beta + \gamma_2 \xi$. 

11
The first line of equation (13) corresponds to moderate or high spending which is focused on late childhood education. That is, it considers the case where families spend only on early childhood education. Government spending on early childhood education is \( \zeta \psi \) and output does not depend on this directly. This is because a unit of spending by government offsets a unit that would be spent by the family at this level, leaving total early childhood spending unchanged. Government spending on late childhood is given by \( \zeta (1 - \psi) \). This affects output in two contrasting ways. First, an increase in spending on early childhood increases total education spending as a share of output. If this is accomplished by an increase in \( \zeta \), the part spent on early childhood is offset by a decrease in private spending. However, the part spent on late childhood does not crowd out private spending since family spending at this stage is already zero. If the increase in late childhood spending is instead accomplished by a decrease in \( \psi \), the result is similar. The decrease in government spending in early childhood causes families to spend more at this stage. However, with families already spending zero on late childhood education, there is not a corresponding decrease in private spending.

The rise in spending relative to output has a positive effect on output. However, another effect works counter to this. A higher level of spending can be offset by a less productive mix of spending. When families spend on only one stage of education, it is because that stage is more productive at the margin. Thus a reallocation of spending in the other direction reduces the productivity of a given level of spending. Increasing \( \zeta \) can cause such a reallocation. A higher \( \zeta \), through higher taxation, leaves a smaller share of output with families. This reduces what families allocate to education. Since families are spending only on early childhood, the reduction occurs at this stage. The net result is a shift in overall spending toward late childhood. A smaller \( \psi \) is another, more direct, way to switch the mix toward late childhood.

This interplay of the level and mix of expenditures is reflected in the first line of equation (13). Here \( \zeta (1 - \psi) \) has both a positive and negative effect. The effects offset where \( \psi (1 - \zeta) = \gamma_2 (\gamma + \gamma_2)^{-1} \). When this value of \( \psi \) lies in the region allowing \( f = (f^*_1, 0) \), a local maximum arises at this point.

Figure 2 aids in the discussion. The first panel is equivalent to Figure 1 but further divides the \( \zeta \times \psi \) space into regions where output is increasing, decreasing, and invariant in \( \psi \). The solid curves are as in Figure 1 and thus delineate the four types of equilibria. The dotted lines trace local output maximizing combinations of \( \psi \) and \( \zeta \). The arrows show directions in which output is weakly increasing in \( \psi \) and \( \zeta \). As shown in Figure 1, the lower region to the right of \( \xi \gamma_2 \beta_1^{-1} \) is
Figure 2: Output. The arrows in the first panel show the direction in which \( \varsigma \) and \( \psi \) can be changed to increase output. The intersecting arrows on the far left indicate that output is unchanged in each direction. The second panel shows normalized output over the policy space. Here output is normalized by the value it would take at \( \varsigma = 0 \).

where \( f = (f_1^*, 0) \). The dashed line in this region is where \( \psi (1 - \varsigma) = \gamma_2 (\gamma + \gamma_2)^{-1} \). For smaller values of \( \varsigma \) in the region, the level effect always dominates, so that output is increased by increasing \( \varsigma \) or decreasing \( \psi \). For the larger values, the mix effect dominates and output can be increased by decreasing \( \varsigma \) or increasing \( \psi \). For intermediate values of \( \varsigma \), beginning at \( \varsigma = \gamma_2 (\gamma + \gamma_2)^{-1} \), the effects offset at some point over the range of \( \psi \) supporting the equilibrium, giving a local interior maximum.

The second panel gives similar information from another perspective. This graphs normalized output, \( y \), for all \( \varsigma, \psi \) pairs. Output is normalized by its value at \( \varsigma = 0 \). The points of inflection correspond to the regions delineated in Figure 1 and the first panel of Figure 2. The lower right region corresponds to the lower central region of the first panel and thus again considers the case where \( f = (f_1^*, 0) \). Consider a value of \( \varsigma \) just beyond \( \varsigma = \xi \gamma_2 \beta_1^{-1} \). Starting at \( \psi = 0 \) and moving in the direction of an increase in \( \psi \), we see that output is decreasing in \( \psi \). At a larger value of \( \varsigma \), output initially rises and then falls as we increase \( \psi \) from zero. This will be true for all values of \( \varsigma \) corresponding to those beneath the lower dashed line in the first panel. Beyond this set of \( \varsigma \) values, output is increasing in \( \psi \).
Results are symmetric when spending is focused on early childhood (the fourth line of equation (13)) so that families spend only on late childhood education. Government spending on early childhood is $\zeta \psi$. An increase in this has a level effect and a mix effect analogous to those discussed above. The effects offset where $\zeta \psi = \gamma (\gamma + \gamma_2)^{-1}$. When this value of $\psi$ lies in the region allowing $f = (0, f_{1*}^t)$, a local maximum arises at this point. In the first panel of Figure 2, the upper region to the right of $\xi \gamma \beta_1^{-1}$ corresponds to $f = (0, f_2^*)$. The dotted curve of Figure 2 plots where $\zeta \psi = \gamma (\gamma + \gamma_2)^{-1}$. For the lower values of $\zeta$ in the range, the level effect dominates and lowering $\psi$ or increasing $\zeta$ increases output as indicated by the arrows. For the higher values of $\zeta$ in this space, the mix effect dominates and raising $\psi$ or lowering $\zeta$ increases output. For intermediate values of $\zeta$, beginning at $\zeta = \gamma (\gamma + \gamma_2)^{-1}$, the effects offset.

The above discussion covers focused spending. We now turn our attention to balanced spending. The second line of equation (13) corresponds to the case of low to moderate balanced spending. In this type of equilibrium, output is independent of the mix of spending. Government spending at each level 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 is fully offset by a unit less or more of private spending. Since total spending at each level is unchanged through policy, human capital and hence output are unchanged. In the first panel, the independence of output from policy when $f = (f_{1*}^t, f_{2*}^t)$ is demonstrated by the lack of a partition and by the intersecting arrows. These indicate that output is unchanging in each direction. In the second panel this is demonstrated by the flat area at $y = 1$ for $\zeta$ small or moderate and balanced.

Finally, consider the third line of equation (13) corresponding to the other possibility with balanced spending (the case of moderate to high spending). Here, both forms of private spending are fully crowded out and government is the sole source of education expenditures. In this case an increase in $\zeta$ unambiguously increases output. With government providing all education spending, the mix of expenditures is determined solely by $\psi$ and not through any general equilibrium adjustments. Thus government spending no longer has an effect on the mix of spending. Furthermore, an increase in $\zeta$ cannot crowd out any private spending since there is none.

For a given $\zeta$, output is maximized when each unit of expenditure is put to its highest use. This requires that the marginal quantity of human capital generated should be the same for both levels of expenditure. This occurs where $\psi = \gamma (\gamma + \gamma_2)^{-1}$. Note that so long as $\zeta$ is large and spending is balanced, its optimal allocation is independent of $\zeta$. This is reflected in the first panel of Figure

14
2 by the the horizontal dotted line and in the second by the ridge at \( \psi = \gamma (\gamma + \gamma_2)^{-1} \).

Figure 2 shows that depending on \( \varsigma \) there can be several locally optimal values of \( \psi \). Looking at the far right of the first panel, we see that for \( \varsigma \) large enough, output is always increasing in the direction of \( \psi = \gamma (\gamma + \gamma_2)^{-1} \) so this is a global maximum. Moving to the left, another local maximum arises with \( \psi \) large and further to the left we have yet another local maximum with \( \psi \) small. As we move further to the left, past the dashed lines, these local maxima occur at corners where \( \psi = 1 \) and \( \psi = 0 \).

The second panel gives clues regarding the globally optimal \( \psi \) as a function of \( \varsigma \). As mentioned before, when \( \varsigma \) is small its allocation is unimportant. For a range of \( \varsigma \), some sort of focused spending is always best since balanced spending yields \( f = (f_1^*, f_2^*) \), the lowest possible output. For \( \varsigma \) in a neighborhood of \( \xi \gamma^2 \beta_1^{-1} \) it is best to set \( \psi = 0 \), and for some values of \( \varsigma \) it is best to set \( \psi = 1 \). Only for \( \varsigma \) sufficiently large is it best to have balanced expenditures. Note in particular that when \( f = (f_1^*, f_2^*) \) comes into existence, it is dominated in output by both sorts of focused spending.

Figure 3 provides a more clear summary. The lines show the maximum output attainable in each type of equilibrium over the range of \( \varsigma \) for which the equilibrium type exists. That is, it shows output at each of the local maxima existing at each \( \varsigma \). This output value is denoted by \( y^* \). The solid lines correspond to balanced spending and the dashed lines to focused spending. The increasing portion of the \( f = (f_1^*, 0) \) and \( f = (0, f_2^*) \) curves correspond to cases where output is locally maximized at a corner (\( \psi = 0 \) or \( \psi = 1 \)), and the flat portion is where the local maxima are interior.\(^{10}\)

The brackets indicate which type of equilibrium globally maximizes output at each level of \( \varsigma \). In the bracketed range furthest to the left, output is maximized where families spend at both stages. In the subsequent bracketed range, an equilibrium where families spend only on early childhood is globally optimal. Next, family spending only on late childhood is globally optimal. Over these two ranges, then, focused spending is preferred (in terms of output). For the range furthest to the right, an equilibrium where families spend at neither level is globally optimal. Only in this range is balanced spending preferred to focused spending. This result is general and the precise cutoff points can be found. The result is stated more precisely in Corollary 2.

**Corollary 2.** If \( \gamma > \gamma_2 \) there exists \( \varsigma_y \) and \( \varsigma_y > \varsigma_y \) such that output is globally maximized at \( f = (f_1^*, f_2^*) \) if \( \varsigma \leq \frac{\xi \gamma_2}{\beta_1} \), at \( f = (f_1^*, 0) \) if \( \frac{\xi \gamma_2}{\beta_1} \leq \varsigma \leq \varsigma_y \), at \( f = (0, f_2^*) \) if \( \varsigma_y \leq \varsigma \leq \varsigma_y \), and at \( f = (0, 0) \)

---

\(^{10}\)The downward sloping portion of these curves reflect areas where output is increasing over \( \psi \) over the entire range but the equilibrium exists over a smaller range of \( \psi \).
if $\varsigma \geq \varsigma_y$.

Findings are symmetric if $\gamma \leq \gamma_2^{11}$. Thus regardless of the relationship between $\gamma$ and $\gamma_2$, each type of equilibrium is globally optimal for some value of $\varsigma$. An implication is that focused spending can dominate balanced spending. When focused spending dominates, the corollary 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, the figure shows that this holds only where $\varsigma_y \leq \varsigma \leq \varsigma_y$. 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 output is maximized where families spend on the productive stage indicating that government should fund the unproductive stage of education. If instead government spending exceeds family spending, output is maximized where government spends on the productive stage.

11 For $\gamma < \gamma_2$, the first lower bound is $\frac{\xi}{\mu_1}$ while $f = (f_1^*, 0)$ and $f = (0, f_2^*)$ switch order in the corollary. For $\gamma = \gamma_2^*$, $\varsigma_y = \varsigma_y$ and both are global maxima with $\varsigma_y < \varsigma < \varsigma_y$. 

Figure 3: Maximum output. The figure shows normalized output across values of $\varsigma$ in each type of equilibrium when $\psi$ is chosen to yield a local maximum. The brackets show which type of equilibrium maximizes output globally at the relevant value of $\varsigma$. 

---

16
This provides intuition for the main result but ignores one important consideration. The level of spending by families is higher when they fund the more productive type of education. Thus total education spending depends on where families spend. This effect serves to influence the level of government spending at which government should fund the productive stage. However, it does not overturn the key message. For lower levels of government spending, output is higher when government spends on the unproductive stage and vice versa.

In the figure above, interior solutions are never optimal with focused spending. This is not a general result. For different parameter choices, the intersection of the \( f = (0, 0) \) curve with the \( f = (0, f_2^*) \) curve can occur at the flat portion of the latter. However, this does not change the above discussion in a substantive way. The only difference is that in this case focused spending does not imply a corner solution.

### 3.3 Utility.

The analysis above considers the effect of policy on output. While output is a common concern of policy makers, utility comparisons are needed to understand the full effect of policy. As such, we now consider how government choices affect utility within an equilibrium type. Proposition 3 gives the main result.

**Proposition 3.** Utility is related to government policy according to

\[
\bar{U} = \begin{cases} 
\beta_2^{-1+\beta} (1 - \varsigma (1 - \psi))^{1+\beta} \xi^{1+\beta} & \text{if } f = (f_1^*, 0) \\
\beta_1^{-1+\beta} \xi^{1+\beta} & \text{if } f = (f_1^*, f_2^*) \\
(1 + \beta)^{-1+\beta} (1 - \varsigma)^{1+\beta} \xi^{1+\beta} & \text{if } f = (0, 0) \\
\beta_3^{-1+\beta} (1 - \psi \varsigma)^{1+\beta} \xi^{1+\beta} & \text{if } f = (0, f_2^*)
\end{cases}
\]

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

The results are clearly closely related to those for output. The difference is that except where \( f = (f_1^*, f_2^*) \), utility maximization considers the effect of taxation on consumption. Consider the first line of equation (14). This depends positively on output but is scaled by \((1 - \varsigma (1 - \psi))^{1+\beta}\). This scalar reflects the extent to which consumption is diminished due to the tax burden. While the entire tax burden as a share of income is \( \varsigma \), \( \varsigma \psi \) is spent on early childhood education. Since families also spend at this level, a unit of tax expenditure on early childhood offsets a unit of private expenditure leaving the same share of output available for consumption. Thus the scalar only reflects expenditure on late childhood education. The opposite is true when families spend only on late childhood education as in the fourth line. Here the scalar reflects only government...
Figure 4: Utility. The arrows in the first panel show the direction in which \( \varsigma \) and \( \psi \) can be changed to increase utility. The intersecting arrows on the far left indicate that utility is unchanged in each direction. The second panel shows normalized utility over the policy space.

spending on early childhood. The second line, where families spend at both levels, does not have a scalar since a unit of government expenditure just offsets a unit of private expenditure. In the third line, the scalar reflects total expenditures since there is no crowding out at the margin.

Figure 4 is analogous to Figure 2 but demonstrates normalized utility, \( u \), rather than output. The dotted lines in the first panel correspond to the dashed lines in Figure 2 and are retained to facilitate a comparison between output and utility maximization. The dashed lines show local optima. 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 \( \varsigma \) are to the left of those which maximize output. This reflects that some output goes to taxation rather than consumption.

The bigger difference relative to the output discussion occurs with high balanced spending. Here, the level of \( \psi \) that maximizes output also maximizes utility. In the first panel, this is demonstrated by the horizontal line at the same level as in Figure 2. In the second panel, it is demonstrated by the partially visible ridge at this level of \( \psi \). However, utility is not monotonic in high balanced spending. It is straightforward to show that the utility maximizing level of spending is \( \varsigma = \left( \gamma + \gamma_2 \right) \left( \xi + 1 + \beta \right) \beta_1^{-1} \). Thus in addition to a ridge of local equilibria along the \( \varsigma \) dimension,
Figure 5: Maximum utility. The figure shows normalized utility across values of $\varsigma$ in each type of equilibrium when $\psi$ is chosen to yield a local maximum. The brackets show which type of equilibrium maximizes utility globally at the relevant value of $\varsigma$. There is a ridge of local equilibria along the $\psi$ dimension at this value of $\varsigma$. Utility with high balanced spending is maximized at the intersection of these ridges.

As with Figure 2, we turn to the second panel for insights regarding global maxima as a function of $\varsigma$. Again each equilibrium type is globally optimal for some range of $\varsigma$. This is seen also in Figure 5, which shows the maximum utility attainable in each type of equilibrium over the range of $\varsigma$ for which the equilibrium type exists. Thus it is analogous to Figure 3. As with output, utility is maximized first where families spend on both levels, second where they spend on early childhood, third where they spend on late childhood, and finally where they spend on neither.

The generalization of this result and the intuition are similar to those regarding output. For brevity, these are omitted. A key similarity, however, is that for lower levels of spending beyond $\xi \gamma_2 \beta^{-1}$ utility, like output, is maximized where families spend on the more productive level of education. For a higher level of spending but below a cutoff level, utility is maximized where families spend on the less productive level. Beyond this, it is optimal that families spend on neither level. Thus focused spending is output and utility maximizing for small enough $\varsigma$ and balanced spending is output and utility maximizing for large enough $\varsigma$. With utility, the cutoff points are different and are denoted by $\varsigma_{\mu}$ and $\varsigma_{\bar{\mu}}$. 


4 The general case.

The previous section requires several restrictive assumptions. In this section 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.

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. The key difference is 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 sharp a distinction across the regions. However, each policy pair yields results that are qualitatively similar to the case with perfect substitutability. In particular, for moderate and high government spending, we have local maxima at several values of $\psi$. The global maximum again depends on $\varsigma$ 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 6 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 $\psi$ 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.

5 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.

12 We run a similar experiment to examine changes in $r$. The results change little, with $r$ simply serving to scale the output effect. For this reason, the results are not presented here.
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, $\varsigma$, will represent lower government spending relative to income for high income families than for low income families. To see it, recall that $\varsigma$ is the share of total output that goes to education. With the population of each generation normalized to 1, lifetime government education spending per family is $\varsigma ZH$. Since the income of family $j$ is $Zh^2_j$, government spending as a share of own income for family $j$ is

$$\varsigma_j = \frac{H}{h^2_j}. \quad (15)$$

It is this $\varsigma_j$ value that matters to families, rather than $\varsigma$ alone.

The distribution of $\varsigma_j$ clearly depends on $\varsigma$ and the distribution of output. This latter item maps into the distribution of $a_j$. Stated differently, we can choose the distribution of $\varsigma_j$ through choosing the distribution of $a_j$. The relationship will be such that the smallest $a_j$ is associated with the largest $\varsigma_j$.

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 with common government spending. In the earlier analysis $H = 2h_2$ so differences in $\varsigma_j$ are generated by differences in $\varsigma$. Now we hold $\varsigma$ constant and allow differences in $\varsigma_j$ though
Figure 7: Output and maximum output. The first panel shows normalized output across the income distribution (represented by \( \tilde{\varsigma}_j \)) as a function of \( \psi \). The second panel shows normalized output across the income distribution when \( \psi \) 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 \( \varsigma = 0 \).

heterogeneity in \( h_{2j} \).

With heterogeneity, we must turn to numerical results even with the parameter restrictions in Assumption 1. These results are presented in Figure 7. The first panel is analogous to the second panel of Figure 2. The difference is that the variation in \( \varsigma_j \) is a general equilibrium consequence of variation in ability. Specifically, for this example we assume that ability is uniform over \([0.5, 5]\) and \( \varsigma = 0.02 \). We then find values of \( \{h_{2j}\} \forall j \in J \) and other endogenous items such that the definition of an equilibrium is satisfied. Given \( \{h_{2j}\} \forall j \) we know \( H \) and thus can use equation (15) to find the distribution of \( \varsigma_j \). For ease of comparison, we plot a monotonic transformation of \( \varsigma_j (\tilde{\varsigma}_j) \) against \( \psi \) on the horizontal axes and normalized output along the vertical axis.\(^{13}\) As before, output is normalized by what it would be with \( \varsigma = 0 \).

The first panel of Figure 7 shows that there are again four distinct regions. These correspond to the regions in Figure 2. For \( \tilde{\varsigma}_j \) small (wealthy families), output is independent of \( \psi \). For larger values of \( \tilde{\varsigma}_j \) (less wealthy families), output depends on \( \psi \). In particular, for any level of \( \tilde{\varsigma}_j \) there are

\(^{13}\) Specifically, the axis is \( \tilde{\varsigma}_j = a(j) \in [0.5, 5] \). This allows for easier comparison and provides the same essential information since there is a one-to-one correspondence between \( a(j) \) and \( \varsigma_j \).
up to three local maxima. Which of these is the global maximum depends on $\tilde{\varsigma}_j$.

The second panel of Figure 7 shows how. It is analogous to Figure 3 and shows the global maximum income as a function of $\tilde{\varsigma}_j$. Moving right to left, we see that low income families prefer relatively balanced spending. When government focuses spending on one level, these families spend on the other. With low income however, the private spending level is low, resulting in low human capital and output.

Further to the left, agents prefer focused spending. For the lower income families among these, output is highest when government focuses on the more productive form of education (large $\psi$) and families spend at the other level. This is because family spending is small relative to focused government spending and it is best to have the larger amount of spending allocated to its most productive use. For the next group of families, private spending is large relative to focused government spending. As such, their income is maximized when government spending is focused on the less productive stage. Finally, for the most wealthy agents, government spending at one stage simply displaces private spending so that output is unchanging in $\psi$.

Because the analogy with the homogeneous case is quite strong, this discussion is quite similar to the discussion after Figure 2. There are, however, some differences. The key qualitative difference is that in Figure 2 output is non-monotonic in $\varsigma$ when $f = (0, f^*_2)$ or $f = (f^*_1, 0)$. This is because in Figure 2 an increase in $\varsigma$ requires an increase in taxes which crowds out private spending. In Figure 7, $\varsigma$ is fixed so this effect does not arise. Also in Figure 7, with $f = (f^*_1, f^*_2)$, $y^*$ increases moderately with $\varsigma_j$. 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 relevant $\varsigma$ 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 may benefit some families at the detriment of others.

6 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 concentrated 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 concentrated 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


7 Appendix

Proof of Propositions 1-3. 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:

\[ \mathcal{L} = \ln c_3 + \beta \ln c_4 + \xi \ln \bar{A} w (f_1 + g_1) \gamma (f_2 + g_2) \gamma_2 + \lambda (I (1 - \tau) - c_3 - c_4 - f_1 - f_2). \]

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

\[ c_3 : \frac{1}{c_3} - \lambda = 0 \] (16a)
\[ c_4 : \frac{\beta}{c_4} - \lambda = 0 \] (16b)
\[ f_1 : \frac{\xi\gamma}{f_1 + g_1} - \lambda \leq 0, \quad f_1 \geq 0, \quad \text{and} \quad \left( \frac{\xi\gamma}{f_1 + g_1} - \lambda \right) f_1 = 0 \] (16c)
\[ f_2 : \frac{\xi\gamma_2}{f_2 + g_2} - \lambda \leq 0, \quad f_2 \geq 0, \quad \text{and} \quad \left( \frac{\xi\gamma_2}{f_2 + g_2} - \lambda \right) f_2 = 0 \] (16d)
\[ \lambda : I (1 - \tau) - c_3 - c_4 - f_1 - f_2 = 0. \] (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

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

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

\[ c_3 = \frac{I(1-\tau)+g_1+g_2}{\beta_2}, \quad c_4 = c_3\beta, \quad f_1 = c_3\xi\gamma - g_1, \quad f_2 = c_3\xi\gamma_2 - g_2. \] (18)

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

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

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

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

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

\[ Y = I = wh_2. \] (21)

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

\[ \tau = \varsigma, \quad g_1 = \varsigma\psi I, \quad g_2 = \varsigma (1 - \psi) I. \] (22)

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

\[ h_2 = \begin{cases} \tilde{A} (c_3\xi\gamma) (\varsigma (1 - \psi) I)^{\gamma_2} & \text{if } f = (f_1^*, 0) \\ \tilde{A} (c_3\xi\gamma) (c_3\xi\gamma_2)^{\gamma_2} & \text{if } f = (f_1^*, f_2^*) \\ \tilde{A} (\varsigma\psi I)^{\gamma} (\varsigma (1 - \psi) I)^{\gamma_2} & \text{if } f = (0, 0) \\ \tilde{A} (\varsigma\psi I)^{\gamma} (c_3\xi\gamma_2)^{\gamma_2} & \text{if } f = (0, f_2^*) \end{cases}. \] (23)
Equations (17)-(20) and (22) give

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

(24)

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

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

(25)

Using this in equation (21) and simplifying gives equation (13).

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

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

Solving for \( \psi \), this can be rewritten to give the first line of equation (12).

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 \( \psi \), this can be rewritten to give the second line of equation (12).

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 \( \psi \), this can be rewritten to give the third line of equation (12).
Let \( f = (0, f_2^*) \). Putting equation (16a) into equation (16c), we see that \( f_1 = 0 \) if \( c_3 \leq \frac{\bar{\gamma}}{\xi_2} \). From the fourth item in equation (20), \( f_2 \geq 0 \) requires \( c_3 \geq \frac{\bar{\gamma}}{\xi_2} \). 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 \( \psi \), this can be rewritten to give the fourth line of equation (12). It is straightforward to show that conditions allowing the four cases are mutually exclusive.

Finally, consider utility. From \( \dot{I} = I \), equation (4) and equations (18)-(19), \( U_j = \ln \beta c_3 \frac{1 + \beta}{1 - \gamma} I^\gamma \). Thus equation (14) follows directly from equation (24). \( \bar{U} = \exp (U_j) \beta^{-\beta} \) which is a monotonic transformation.

**Proof of Corollary 1.** Consider the second line of equation (12). The left-hand-side inequality holds for all values of \( \psi \) only if \( \varsigma \leq \xi \gamma_2 \beta_1^{-1} \). The right-hand-side inequality holds for all values of \( \psi \) if and only if \( \varsigma \leq \xi \gamma \beta_1^{-1} \). Thus both always hold only if \( \varsigma \leq \min \left[ \xi \gamma_2 \beta_1^{-1}, \xi \gamma \beta_1^{-1} \right] \). Along with the definition of low spending, this gives the first line of Corollary 1. The second line of the corollary follows from the first and fourth lines of equation (12) where we see that for \( \psi \) small, families spend only on early childhood and with \( \psi \) large, families spend only on late childhood. Finally, note that \( f = (0, 0) \) can exist for some \( \psi \) if and only if \( \frac{\xi \gamma_2 \beta_1}{1 - \varsigma} \leq 1 - \frac{\xi \gamma_2}{1 + \varsigma} \). This requires \( \varsigma \geq \frac{\xi \gamma_2 + \xi \gamma_2}{\beta_1} \). Similarly \( f = (f_1^*, f_2^*) \) can exist for some \( \psi \) if and only if \( 1 - \frac{\xi \gamma_2}{\beta_1 \varsigma} \leq \frac{\xi \gamma}{\beta_1 \varsigma} \). This requires \( \varsigma \leq \frac{\xi \gamma + \xi \gamma_2}{\beta_1} \). Along with the definitions of moderate and high spending, this proves the final line.

**Proof of Corollary 2.** For brevity, we provide only a sketch of the proof. Throughout, we consider \( \dot{I} = (Aw)^{-1} I^{1-\gamma - \gamma_2} \) rather than \( I \) with no loss of generality.

From equation (12), the following equilibria exist for some value of \( \psi \in [0, 1] \) given the values of \( \varsigma \):

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

Define \( \dot{I}_{f_1^*, 0} \) to be output in \( f = (f_1^*, 0) \) when \( \psi \) is chosen to locally maximize output. Stated differently, it is the maximum output over the range of \( \psi \) supporting \( f = (f_1^*, 0) \) given \( \varsigma \). 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 (13). From choosing the output maximizing level of \( \psi \) in equation (13) we find

\[
\dot{I}_{f_1^*, 0} = \begin{cases} 
\left( \frac{\xi \gamma (1 - \varsigma) \beta_2^{-1}}{\gamma \varsigma_2} \right)^\gamma \varsigma_2 & \text{if } \varsigma \leq \frac{\gamma_2}{\gamma + \gamma_2} \\
\left( \frac{\xi \gamma_2}{\gamma + \gamma_2} \beta_2^{-1} \right)^\gamma \left( \frac{\gamma_2}{\gamma + \gamma_2} \right)^\gamma \varsigma_2 & \text{if } \varsigma \geq \frac{\gamma_2}{\gamma + \gamma_2}.
\end{cases}
\]
Similarly

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

(26)

and

\[
\tilde{I}_{0,0} = \left( \frac{\zeta \gamma}{\gamma + \gamma_2} \right)^{\gamma} \left( \frac{\gamma \gamma_2}{\gamma + \gamma_2} \right)^{\gamma_2},
\]  

(27)

Each is continuous. The first two are initially increasing in \(\zeta\) and level out at \(\zeta = \frac{\gamma}{\gamma + \gamma_2}\) and \(\frac{\gamma}{\gamma + \gamma_2}\).

The third is increasing in \(\zeta\) always and the fourth is independent of \(\zeta\).

Consider starting with \(\zeta = 0\) and increasing \(\zeta\). 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 \(\zeta = \frac{\xi \gamma_2}{\beta_1}\), \(\tilde{I}_{f_1^*,0} = \tilde{I}_{f_1^*,f_2^*}\) and the ratio of \(\tilde{I}_{f_1^*,0}\) to \(\tilde{I}_{f_1^*,f_2^*}\) is increasing in \(\zeta\). Thus beginning here, \((f_1^*, 0)\) is optimal and beyond this value of \(\zeta\), \(f = (f_1^*, f_2^*)\) can not be globally optimal.

When \(\tilde{I}_{0,f_2^*}\) comes into existence at \(\zeta = \frac{\xi \gamma_2}{\beta_1}, \tilde{I}_{0,f_2^*} < \tilde{I}_{f_1^*,0}\). 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 \(\zeta\). At their maximum values \(\tilde{I}_{0,f_2^*} > \tilde{I}_{f_1^*,0}\). To see this, put \(\zeta = \frac{\gamma}{\gamma + \gamma_2}\) into the first line of equation (25) and \(\zeta = \frac{\gamma}{\gamma + \gamma_2}\) into the first line of equation (26) and compare. This is sufficient to show that \(\tilde{I}_{0,f_2^*} = \tilde{I}_{f_1^*,0}\) at one value of \(\zeta\). Call it \(\zeta.y\). Beyond \(\zeta.y\), \(f = (f_1^*, 0)\) cannot be globally optimal.

When \(\tilde{I}_{0,0}\) comes into existence at \(\zeta = \frac{\xi(\gamma + \gamma_2)}{\beta_1}, \tilde{I}_{0,0} < \tilde{I}_{0,f_2^*}\). 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 \(\zeta\). At their maximum values \(\tilde{I}_{0,0} > \tilde{I}_{0,f_2^*}\). To see this, put \(\zeta = \frac{\gamma}{\gamma + \gamma_2}\) into the first line of equation (26) and \(\zeta = 1\) into the first line of equation (27) and compare. This is sufficient to show that \(\tilde{I}_{0,0} = \tilde{I}_{0,f_2^*}\) at one value of \(\zeta\). Call it \(\zeta.y\). Beyond \(\zeta.y\), \(f = (0, f_2^*)\) cannot be globally optimal.

We have shown that in the range \(\zeta \in \left( \frac{\xi \gamma_2}{\beta_1}, \zeta.y \right), \tilde{I}_{f_1^*,0} > \tilde{I}_{f_1^*,f_2^*}, \tilde{I}_{0,f_2^*}\). To assure a global maximum, we need to show that \(\tilde{I}_{f_1^*,0} > \tilde{I}_{0,0}\) in this range. Suppose \(\tilde{I}_{f_1^*,0} = \tilde{I}_{0,f_2^*} < \tilde{I}_{0,0}\) at \(\zeta = \zeta.y\). Then \(\tilde{I}_{0,0} = \tilde{I}_{0,f_2^*} (\zeta = \zeta.y)\) must occur at a lower value of \(\zeta\) than \(\tilde{I}_{0,0} = \tilde{I}_{f_1^*,0}\) and before \(\zeta = \zeta.y\). This is because \(\tilde{I}_{0,f_2^*} < \tilde{I}_{f_1^*,0}\) to the left of \(\zeta = \zeta.y\) and their ratio is increasing in \(\zeta\) over this range. At the point where \(\tilde{I}_{0,f_2^*} = \tilde{I}_{0,0}\), it must be that \(\tilde{I}_{f_1^*,0} > \tilde{I}_{0,0}\) so long as \(\zeta < \zeta.y\). \(\tilde{I}_{f_1^*,0} = \tilde{I}_{0,0}\) only at a larger \(\zeta\).

We show that this cannot hold (a contradiction) and that in fact \(\tilde{I}_{0,0} = \tilde{I}_{0,f_2^*}\) occurs at a higher value of \(\zeta\) than \(\tilde{I}_{0,0} = \tilde{I}_{f_1^*,0}\). To show this, we find the values of \(\zeta\) that solve \(\tilde{I}_{0,0} = \tilde{I}_{f_1^*,0}\) and
\( \tilde{I}_{0,0} = \tilde{I}_{0,f_2^*} \) and then determine that \( \tilde{I}_{0,0} = \tilde{I}_{0,f_2^*} \) occurs at a lower value of \( \varsigma \) than \( \tilde{I}_{0,0} = \tilde{I}_{f_1^*,0} \) only if

\[
\frac{1 + \beta + \gamma_2 \xi}{1 + \beta + \gamma \xi} < \left( \frac{\gamma + \gamma_2}{\gamma} \right)^{\frac{\xi}{2}} \left( \frac{\gamma + \gamma_2}{\gamma_2} \right)^{\frac{\gamma_2}{\gamma}}.
\]

The two sides of this are equal at \( \gamma = \gamma_2 \). The left hand side is decreasing in \( \gamma \). The right hand side is increasing in \( \gamma \). This is not obvious but can be shown to hold. Given this, the inequality cannot hold for \( \gamma > \gamma_2 \). Thus \( \tilde{I}_{0,0} = \tilde{I}_{0,f_2^*} \) occurs at a higher value of \( \varsigma \) than \( \tilde{I}_{0,0} = \tilde{I}_{f_1^*,0} \). With this, we know that \( \tilde{I}_{f_1^*,0} > \tilde{I}_{0,0} \) over \( \varsigma \in \left( \frac{\xi \gamma_2}{\beta_1}, \varsigma_y \right) \) and an equilibrium with \( f = (f_1^*, 0) \) is globally optimal in this range.

Since \( \tilde{I}_{f_1^*,0} = \tilde{I}_{0,f_2^*} > \tilde{I}_{0,0} \) at \( \varsigma = \varsigma_y \), by continuity and earlier arguments, we know that in a neighborhood to the right of this \( \tilde{I}_{0,f_2^*} > \tilde{I}_{0,0} \) and \( \tilde{I}_{0,f_2^*} \) is a global optimum. We also know they cross as some point \( \varsigma_y \) and beyond this \( \tilde{I}_{0,0} \) is globally optimal. This completes the sketch of the proof.