Rochester Center for Economic Research

**Higher Education Subsidies and Heterogeneity
A Dynamic Analysis**

Elizabeth M. Caucutt and Krishna B. Kumar

Working Paper No. 472
May 2000

UNIVERSITY OF
ROCHESTER

Higher Education Subsidies and Heterogeneity

A Dynamic Analysis[*]

Elizabeth M. Caucutt

Department of Economics,

University of Rochester.

Krishna B. Kumar

Marshall School of Business,

University of Southern California.

June 08, 1999

Abstract

In this paper, we develop a simple dynamic general equilibrium framework that can be used to study issues in higher education policy. The model features heterogeneity in income of parents and academic ability of students. Liquidity constraints create persistence in educational attainment even when ability is independently distributed. A unique steady state with a positive fraction of college educated workers, or one with a development trap, or multiple steady states which feature both of these, can result in equilibrium. We add a government that is equipped with a simple tax and subsidy scheme and calibrate the model to the US economy to get a benchmark for our policy analysis. The government can design a tax and subsidy scheme that guarantees equality of opportunity, but only at the expense of a decrease in the efficiency of utilization of education resources; the welfare gain is minimal. A policy that aims to maximize the fraction of college-educated labor, by sending as many children as possible to college, results in a big drop in the above-mentioned efficiency with little or no welfare gain. If the government has the political will to use any available signal on ability and provide merit-based aid, it can increase this efficiency with little decrease in welfare. Education subsidies may be a potent tool for countries that are caught in a development trap; a sufficient level of subsidy can cause the economy to emerge from the trap.

1 Introduction

The case for increasing the level of higher education subsidies is repeatedly made by politicians and policy makers.[1] The current level of subsidies is already quite substantial. McPherson and Schapiro (1991) report that $7,839 of the instructional cost of $8,760 in public universities is subsidized. In private institutions, $7,224 of the cost of $12,669 is subsidized. Most of the work to date to study the economics of higher education subsidies, has been either descriptive or empirical.[2] The contributions of this paper are twofold. First, we develop and analyze a simple dynamic general equilibrium framework that models a parent's decision to send a child to college. The model is general enough to be useful in other contexts. Second, we calibrate our model to the US economy, and conduct several experiments to study the effect of education policy on this decision. We view our dynamic modeling and quantitative approach as being complementary to most existing work in this area.

The case for subsidization is usually based on equality of opportunity (improving the access of poor people to college education in the face of increasing college premiums) and on growth externalities (the notion that economic growth depends on the average education level of the entire workforce, thus justifying subsidization by the government to get human capital investment to efficient levels).[3] We leave the issue of economic growth for a future extension of our framework, and focus here on the efficacy of higher education subsidies when the government is unable to fully observe the potential of students to successfully complete college.

We develop a simple overlapping generation model with two period lived agents. A liquidity constrained parent makes the decision of whether or not to incur the cost of sending a child to college, taking into account the child's academic ability.[4] The probability that a child who is sent to college will successfully graduate depends positively on this ability. The parent is herself either a school educated or a college educated worker. The two types of workers are imperfect substitutes in production. We study in detail the parental decision and the resulting dynamics, and prove the existence and uniqueness of equilibrium. This gives us confidence in the robustness of our setup and allows us to devise strategies to compute and calibrate the model to the US economy; we then compare outcomes under different policy regimes relative to this benchmark.

The liquidity constraint leads to persistence in higher education attainment; children whose parents are college educated are themselves more likely to be college educated. This occurs even though the academic ability of children is independently and identically distributed across both types of parents.

---

[1] The senate recently approved a "higher education bill that will cut student loan rates to the lowest level in nearly 20 years and make grants available to low-income students." (Reuters news report, September 30, 1998.) "This legislation marks an important step forward in my effort to help more Americans enter the doors of college," President Clinton remarked. A senator called it the "financial key to unlock the door to higher education."

The same report places at $48.5 billion, the financial aid provided by Washington for 8.5 million students. This is in addition to $216 million spent directly for colleges.

[2] See, for instance, McPherson and Schapiro (1991), which is a comprehensive descriptive and empirical work, and the references therein.

[3] See McPherson and Schapiro (1991) and Heckman and Klenow (1997) for an elaboration of each of these cases.

[4] We defend our assumption of liquidity constraints in detail in the next section.

The model gives rise to three types of dynamic equilibria depending on the curvature parameter of the utility function (or alternately, the cost of education): 1. a unique steady state with a positive fraction of college educated workers, 2. one with a development trap, and 3. multiple steady states which feature both of these. The model thus has the potential to explain the broad diversity in educational attainment that is observed across countries.

A development trap occurs when the initial fraction of college educated people in the workforce is too low. The wages of the school educated workers are too low for them to find it profitable to send their children to college. This results in a decrease in the fraction of college educated people next period, which further decreases the wages of school educated workers, and reinforces the above-mentioned behavior.

We then add a government that is equipped with a simple tax and subsidy scheme, and conduct several policy experiments. The liquidity constraint, and the ensuing persistence in attainment, provides a natural justification for subsidies. We assume that the government, unlike parents, is unable to observe student ability. An alternate interpretation of the model is that the government can observe student ability, but is forced to give need-based, rather than merit-based, aid because of political considerations. Thus, college subsidies are given only to children of school-educated parents, and the level of subsidies is common to all.

The government can design a tax and subsidy scheme that guarantees equality of opportunity, but only at the expense of a decrease in the efficiency of utilization of education resources. Increased subsidies attract inframarginal students and cause an increase in the dropout rate. The welfare gain that comes from increasing the fraction of people with higher marginal productivity is minimal, once the negative effect of taxes during the transition is taken into account. A policy that aims to maximize the fraction of college-educated labor by sending as many children as possible to college results in a big drop in the above-mentioned efficiency with little or no welfare gain. If the government has the political will to use any available signal on ability and provide merit-based aid, it can increase this efficiency with little decrease in welfare.

Education subsidies may be a potent tool for countries that are caught in a development trap. A sufficient level of subsidy can cause the economy to emerge from the trap and also allow the gains to be realized in a timely manner.

Even though the model is used here to analyze higher education subsidies, it is general enough to apply to other contexts. For instance, the heterogeneity can be interpreted as participation in formal and informal financial sectors, where an initial fixed investment is required to participate in the formal sector. Output is determined by the credit intermediated by both sectors, with the formal sector having an edge over the informal one in expected productivity. The model's dynamics can then be used to examine the evolution of financial intermediation in an economy and its level of development.[5]

In an abstract sense, there is no reason why the education modeled here should be confined to higher education; it can capture any process of costly human capital accumulation. In fact, in the context

---

[5]For an analysis of similar issues see, for example, Greenwood and Jovanovic (1990).

of development traps the education modeled here is better interpreted as primary and secondary education. That is, liquidity constrained parents in very poor economies do not send their children to school, let alone college. However, the empirical anchor to the model is the interpretation of education as higher education. The model is calibrated to US data on college premium and college attainment. For the US, and for other developed countries, the process of education modeled here is better interpreted as higher education and the production function described below is consistent with that interpretation.

As mentioned earlier, there are few analytical studies of higher education subsidies in the literature. Hanushek, Cheah, and Leung (1997) use a static model to compare educational subsidies to other tax-transfer schemes in achieving distributional objectives. They conclude that there is a role for education subsidies only when there are externalities in production. Without externalities, redistribution through education is inferior to direct wage subsidies. We are more interested in the interaction between tuition subsidies and parental incomes (which is an important motivation in McPherson and Schapiro (1991) as well), than in redistribution *per se.* In our model, education subsidies are in kind and cannot be directly used for consumption. The motive for redistribution is not increasing consumption for the poor as much as it is to encourage investment in higher education, a necessarily dynamic issue which cannot be analyzed in the static setup of Hanushek, Cheah, and Leung (1997).

Our model bears an analytical similarity to Andrade (1998), who uses a model with two types of agents to study issues on growth and inequality. We study a different set of issues, and assume that educational expenditure is fixed as opposed to varying. He does not model heterogeneity in ability, a feature that is central to our analysis.

Heckman and Klenow (1997) address the issue of empirically assessing the degree of externalities in production and human capital accumulation that are needed to justify current levels of subsidies. They find that the social marginal product of human capital gained from college education has to be 30% higher than the private one. We leave the theoretical analysis of subsidies under externalities and growth to future work and study a more basic issue of how subsidies affect higher educational attainment when there is heterogeneity in income and ability.

An early influence on modeling general educational investment with liquidity constraints and heterogeneity in income is the work by Loury (1981). In his model, ability matters only for production, and even parents do not know the child's ability while investing in her education. In our basic model, ability matters only for college completion, and parents are aware of the child's academic ability while deciding whether to send her to college. This difference has important policy implications. In Loury's model it is possible to design redistributive policies that make *all* parents better off in an *ex ante* sense. Since parents are unaware of the child's ability while investing, redistribution can provide insurance – policies that decrease the spread of income distribution will be favored by all parents, rich or poor. Providing insurance is not the primary motive for subsidies have in our setup; the aim is to induce participation through an income effect. Rich parents become worse off under redistribution, even when an aggregate measure of welfare increases. In fact, when there are opposite effects on the welfare of the rich and the poor, it is not clear what the right measure of aggregate welfare is; this is one of the reasons we also consider a positive measure of education efficiency. Even if working ability were

modeled in our framework, we would expect it to be strongly correlated with academic ability. To then presume that parents have no knowledge of a child's working ability while making the investment decision would not be reasonable.

Analytically, our model differs from Loury's in its fixed educational investment and limited heterogeneity, which makes it simpler on some dimensions as well as more readily applicable to issues in higher education subsidies and the college wage premium. Our model is also readily amenable to transitional analysis. In his model, educational investment varies with income, and there is a continuous distribution of earnings; the focus is on the stationary distribution that arises from a Markovian transition function.[6]

The rest of the paper is organized as follows. In section 2, we present the basic model without the government. The model is characterized in section 3. The government is introduced in section 4, where we also discuss the calibration strategy and the policy experiments. Section 5 concludes the paper.

2   The Model Economy

2.1   Demographics

The economy is populated by a continuum of two-period lived agents in an overlapping generations setup. The size (measure) of each generation is normalized to one. Agents are children in the first period and parents in the second. Children complete schooling and some of them will go on to attend college. We do not model the schooling decision. However, the parental decision to send a schooled child to college is central to the model. We thus take the stance that the education decision has a strong intergenerational aspect to it. Such a stance seems empirically valid.[7] This is in contrast to a strategy where there is no intergenerational connection and students born without endowments finance their own education in a perfect credit market.

School as well as college educated children enter the labor pool in the second period of their lives. At this point each adult worker becomes a parent, has a schooled child, and the economy continues. A parent derives utility from her consumption and from her child's utility. Therefore, altruism provides the intergenerational linkage. We will use "rich" and "college educated" interchangeably, as we will "poor" and "school educated", throughout the paper.

2.2   Parents' Education Problem

Each parent faces a decision of whether to send her child to college. Children differ in their ability to complete college. We assume that, conditional on being sent to college, a child with ability a completes college with probability $\pi(a)$. With probability $(1 - \pi(a))$, the child drops out of college

---

[6] However, it should be possible to use the computational techniques of Krusell and Smith (1998) to conduct transitional analysis on Loury's model.

[7] See, for instance, Table 5-7 in McPherson and Schapiro (1991). For a particular sample they analyze, parents' price over gross cost ranged from 19.4% to 88.2%. As a percentage of the family cost, which nets out scholarships, grants, and student loans from the gross cost, the figure is not less than 65% and for some income groups is as high as 95%.

and enters the schooled labor force irrespective of the number of years she has attended college. Partial college education is useless in this model. This all-or-nothing aspect to college education limits the heterogeneity in education levels, and therefore income, to two types.[8]

Assumption 1: $a \in [0,1]$. $0 \leqslant \pi(a) \leqslant 1$, $\pi'(a) > 0$, $\forall a$. $\pi(0) = 0$. The ability of the child is fully observed by the parent, but not others (including the government, which we will introduce in Section 4). $\square$

The above informational assumption captures the idea that a parent, by spending more time with her child than anyone else does, is in a better position to observe the child's ability. Let $F(\cdot)$ denote the distribution function for ability on the support $[0,1]$, and $f(\cdot)$ the corresponding density function. The distribution is identical across types and within parents of the same type. All ability draws are independent of each other. In this section, we assume that the cost of college education is the same for all children. Sending a child to college involves a real cost of e units of consumption.

Assumption 2: A parent cannot borrow to finance her child's college education.[9] $\square$

Incompleteness in the human capital market, on account of the moral hazard inherent in the process of accumulation and problems of enforcement, is a widely accepted assumption. See, for example, the discussion in Becker (1991), pp. 247-48. However, Heckman and Klenow (1997) cite evidence from Cameron and Heckman (1998) as reason to be skeptical on the prevalence of liquidity constraints. Therefore, our reasons for continuing to assume that parents are liquidity constrained are worth noting. First, Cameron and Heckman find that the average derivative of probabilities of schooling transition with respect to family income decreases in magnitude when a proxy for ability, the AFQT score, is controlled for. Based on this evidence, they suggest that long term family characteristics that affect the AFQT scores are more important than short term liquidity constraints. In most of the equilibrium outcomes of our model with liquidity constraints, the most able of the poor children attend college, just as the most able of the rich children. In other words, conditioning on ability lowers the dependence on income of the enrollment decision even in our model, which is qualitatively consistent with the evidence in Cameron and Heckman. Second, in spite of controlling for AFQT scores and family characteristics other than income, they find positive and significant values for the above-mentioned average derivatives. Third, it seems highly unlikely that the family characteristics are uncorrelated with income. As long as there is incomplete biological transmission of family characteristics, and there is a dearth of socio-economic models in which family characteristics and income are jointly determined,

---

[8]The assumption that not all students who enter college complete it successfully is empirically justified. OECD (1997) reports that tertiary non-completion rates in OECD countries ranged from 6% to 64% of those who were enrolled in college (pp. 93). Even when the high levels of subsidies in these countries are taken into account, these drop-out rates seem significant. In fact, in section 4 we calibrate the model to be broadly consistent with US dropout rates.

[9]With this assumption, the private information aspect in Assumption 1 does not have any bite in the private sector economy. It will matter when we introduce the government.

Also, we do not explicitly model the moral hazard involved in the process of human capital accumulation that causes the incompleteness of markets in the first place. This has to be borne in mind while interpreting welfare measures. The process of completion of the market by the government can increase the moral hazard and decrease welfare.

it seems premature to entirely dismiss models with liquidity constraints. Fourth, in spite of liquidity constraints, our model retains the policy implications of Cameron and Heckman's study that policies that raise family income attract lower ability students into college, and ignoring heterogeneity in abilities is likely to overstate the gains of such policies.[10] Finally, as Heckman and Klenow note, "the absence of any convincing evidence on liquidity constraints may be a consequence of the generosity of existing programs." Modeling liquidity constraints can help one to structurally evaluate the force of these constraints in the face of subsidies, though, admittedly, assuming that they always apply (as we do) is a bit extreme.

The goods cost ("tuition") of e is the only cost of higher education. Ignoring the opportunity cost of foregone earnings, an important component of the cost of education, is not likely to be a serious omission for purposes of studying the efficacy of higher education subsidies. As noted by Heckman and Klenow (1998), the government subsidizes only the direct cost of college, not the opportunity cost faced by students.[11]

Let the fraction (measure) of college educated people entering the labor force at any time be denoted by n_c. This is the only aggregate state variable in this economy. Let $w_c(n_c)$ denote the wage earnings of a college educated parent as a function of the aggregate state n_c, and let $w_s(n_c)$ denote the wage of a school educated parent. Consider a parent of type i, $(i = c, s)$, who has the good (ill) fortune of having a child of ability a. This liquidity constrained parent can decide to further educate the (already schooled) child in college by expending e, or not educate the child any further and incur no cost. Let $V_i(a; n_c)$ be the value of this parent who optimally decides whether or not to send her child to college. We have included the aggregate state n_c, in addition to the individual state a, as an argument of the value function, to highlight the role it plays in parental decisions. This parent's Bellman equation is:

$$V_i(a; n_c) = max_{c,s}\{u(w_i(n_c) - e) + \beta\left[\pi(a)E_{a'}V_c(a'; n_c') + (1 - \pi(a))E_{a'}V_s(a'; n_c')\right],$$
$$u(w_i(n_c)) + \beta E_{a'}V_s(a'; n_c')\}, \ i = c, s. \tag{1}$$

Here, $E_{a'}V_i(a'; n_c') = \int_0^1 V_i(a'; n_c')dF\left(a'\right)$, $i = c, s$, is the child's expected utility, which depends on whether the child enters adulthood as a college educated or a school educated worker. Henceforth, we will denote this quantity by $EV_i\left(n_c'\right)$. We take β to be an intergenerational discount (altruism) factor. The decision is between sending the child to college (c) and letting the child remain school educated (s). The aggregate state that will prevail when the child enters the labor force is denoted by n_c'. We assume that all parents posit that the aggregate state follows the law of motion:

$$n_c' = \Phi(n_c), \tag{2}$$

which they assume to be outside their control. Since there are no aggregate shocks in this economy,

---

[10] See their discussion in pp. 312-313.

[11] A trap is more likely when e (or σ) is high. The poorest countries are likely to have a high cost of education relative to income. Higher education, especially, is a capital intensive activity (in both human capital and physical capital such as laboratory equipment, computers, and structures); the availability of cheap labor is therefore unlikely to bring down the cost of education in poor countries.

this is a deterministic function. Parents thus take the current and future wages as given in solving their optimization problem. We assume a standard utility function, with $u' > 0$, and $u'' < 0$.

Before proceeding to describe production in this economy, we note two aspects of intergenerational transmission as we have modeled it. We have assumed that a child's ability is independent of her parent's ability. This is the reason we can write the *ex ante* expected utility as a function of the aggregate state alone. We are thus focusing on the intrinsic ability that would help children complete college and are ignoring any biological transmission of ability from parent to child. When we calibrate the model and conduct policy experiments, we use different probability functions for the children of college and school educated parents, $\pi_c(a) > \pi_s(a), \forall a \epsilon [0,1]$. This allows us to model the advantages that children of rich parents have from higher quality primary and secondary schooling, and captures one dimension of persistence. Second, in our setup only educational bequests are possible. We speculate on the effects of adding physical capital bequests in the final section.

2.3 Production and Wages

In the earlier subsection we assumed that parents took current and future wages of college and school educated labor as given.

Assumption 3: We assume that $w_c(n_c)$ is strictly decreasing, $w_s(n_c)$ is strictly increasing, and $w_c(n_c) > w_s(n_c), \forall n_c \epsilon [0,1]$. □

The results do not materially depend on $w_c > w_s$ holding uniformly in the interval $[0,1]$; the assumption is made purely for convenience. Even if this inequality does not hold in the entire interval, equilibrium considerations demand that we focus attention on the interval $[0,\overline{n_c}), \overline{n_c} < 1$, in which $w_c > w_s$. Since college education uses real resources, parents will send their children to college only if they anticipate college-educated wage to be higher. If at any time $w_c < w_s$ is anticipated, no one will be sent to college, which will make the next period college wage very high and induce college enrollment.

We now motivate one production structure that yields wage functions with the above properties. Suppose firms in the economy manufacture a single consumption good using school and college educated labor as distinct inputs. The production function is constant returns to scale. Therefore, we assume a single production firm that manufactures the good according to the CES specification:

$$Y = A \left[\theta \left(N_c + \gamma N_s\right)^\nu + (1 - \theta)(N_s + \varepsilon N_c)^\nu\right]^{\frac{1}{\nu}}, \tag{3}$$

where $0 < \varepsilon, \nu < 1$, and $\gamma \ll \epsilon$. The first term within the square brackets can be thought of as "brain" and the second term as "brawn".[12] N_c is the number of college educated workers employed by the firm, while N_s is the number of school educated workers employed. College educated workers are the primary suppliers of "brain" (the weight of school educated workers in this factor, γ, is close to zero, and is included mainly for the technical reason of keeping wages bounded). Both types of workers contribute toward "brawn". Since going to college can take away time that one could have spent developing one's biceps, each unit of college educated labor is assumed to contribute less to "brawn" than does a unit of

---

[12]The terminology is inspired by Stokey (1996).

school educated labor – hence the factor $\varepsilon < 1$ that multiplies N_c in the second term of the production function. This production function captures the notion that college educated workers can do (almost) all that school educated workers can do, and more. Note that the mere employment of N_c workers contributes to "brain" and "brawn" – there is no decision involved on the part of these workers in devoting time to one activity versus the other. Simple manipulation of the marginal products yields a sufficient condition that guarantees higher wages for college-educated workers in the entire interval $[0,1]$: $\varepsilon > \left(\frac{1-\theta}{(1-\gamma)\theta}\right)^{\frac{1}{1-\upsilon}}$.[13] The coefficient of skilled labor in the production of "brawn" should be high enough to guarantee that skilled wage is always higher. Since the marginal product of college educated labor is always higher than that of school educated labor, output is maximized when n_c is maximized.

This completes the description of the economy.

3   Characterization of the Model

3.1   Parents' Decision Rules

We begin by characterizing the behavior of parents. Each type of parent has to make the decision of whether or not to send her child to college, given the child's ability a (individual state) and n_c the aggregate state. Let $C_i(n_c)$ be the set of intervals such that a parent of type i sends her child to college if a belongs to any of the set's elements. We assume that if a parent is indifferent between the two options, she will choose the college option. In the appendix, where all results are derived or proved, we show the following result.

Lemma 1 *In any equilibrium in which parents behave optimally, for any $n_c \in [0,1]$, $EV_c(n_c) > EV_s(n_c)$.*

The above intuitive lemma asserts that college educated parents cannot get lower *ex ante* utility than the school educated parents on average. From the parents' problem in (1), it is clear that the option of not sending the child to college yields a value that is independent of the child's ability, a. Given assumption 1 and the above lemma, it is also clear that the value of sending the child to college is increasing in a. It is therefore intuitive to expect the parent's decision to exhibit a threshold (or reservation) ability behavior. This intuition is formalized in the following lemma.

Lemma 2 *For any given n_c, there exist unique $a_i^*(n_c) \in (0,1]$, $i = c, s$, such that a parent of type i sends her child to college if $a \geqslant a_i^*(n_c)$, and does not otherwise. That is, $C_i(n_c)$ is the singleton set $\{[a_i^*(n_c),1]\}$.*

When we discuss the parents' problems, we really need to write $a_i^*(n_c; \Phi)$. Since Φ is the law of motion for n_c that parents posit, their decision rule would depend on this function. However, for the

---

[13]The earning function in Loury (1981), where earnings positively depend on educational investment and ability, can be viewed as a reduced form of the production function we consider.

sake of notational convenience we will be explicit about this only when absolutely necessary. For a parent whose child is at the threshold ability, we can examine the two options of (1) and write:

$$\beta\pi(a_i^*(n_c))\left(EV_c\left(n_c'\right) - EV_s\left(n_c'\right)\right) \leqslant u(w_i(n_c)) - u(w_i(n_c) - e), \tag{4}$$

with equality if $a_i^*(n_c) < 1$. This expression says that for the marginally able child, the discounted benefit of going to college is exactly equal to the utility cost of college education. Sending any child who is more able will result in the parent benefiting by more than the utility cost. It is not worth the parent's utility cost to send any child who is less able. If the above holds as an inequality even when $a_i^*(n_c) = 1$, even the most able child will not be sent to college. That is, *no* parent of type i sends her child to college. It is convenient to introduce new notation for a few expressions that we will frequently encounter. Define:

$$\Lambda(n_c) \equiv EV_c(n_c) - EV_s(n_c) \tag{5}$$

$$g_i(n_c) \equiv u(w_i(n_c)) - u(w_i(n_c) - e). \tag{6}$$

$\Lambda(n_c)$ is the benefit of sending a child to college as a function of the aggregate state (and hence common to both parent types), and $g_i(n_c), i = c, s$, is the utility cost of a parent of type i of sending a child to college.[14] From the production function (3) and the concavity of u, we obtain $g_c'(n_c) > 0$ and $g_s'(n_c) < 0$. Expression (4) can now be written as $\beta\pi(a_i^*(n_c))\Lambda(\Phi(n_c)) \leqslant g_i(n_c)$.

Given assumption 3, it seems reasonable to expect that the richer college educated parents can afford to send even children of lower ability to college while the poorer school educated parents can afford to send only higher ability children to college. We therefore have the following result.

Lemma 3 *In any equilibrium in which parents behave optimally and some college education occurs, $a_c^*(n_c) < a_s^*(n_c)$, for $n_c \in [0,1]$. It also follows that $a_c^*(n_c) < 1$.*

This lemma can be viewed as a result on persistence. That is, children whose parents are college-educated are themselves more likely to be college educated, a piece of evidence that is often mentioned. We get this result even though ability is i.i.d., on account of the liquidity constraint parents face in financing college education and curvature of the utility function.[15]

3.2   Definition of Equilibrium

We are now in a position to define an equilibrium.

---

[14]Λ can be thought of as the "costate" of this dynamic system. Given the discrete nature of the heterogeneity, it is convenient to analyze the system in terms of this "costate" even though the problem is stated in recursive rather than sequential form.

[15]This result is not unexpected. Loury's (1981) assumption that education is a normal good implies persistence in income in his model.

Definition 1 *A **competitive equilibrium** is a collection of functions $w_i(n_c)$, $a_i^*(n_c)$, $i = c, s$, $\Lambda(n_c)$, and $\Phi(n_c)$, on $[0, 1]$, such that:*

- *Given $w_i(n_c)$, $\Lambda(n_c)$, and $\Phi(n_c)$, $a_i^*(n_c)$ solves the college decision problem of parent i. That is, (4) is satisfied.*

- *Given wages, the firm maximizes profits. That is, both types of workers are paid their marginal products.*

- *The labor market clears. That is, $N_c = n_c$, and $N_s = 1 - n_c$.*

- $\Lambda(n_c)$ *is consistent with parental decisions. That is, $EV_c(n_c)$ and $EV_s(n_c)$ satisfy the equation system:*

$$EV_c(n_c) = \int_0^{a_c^*(n_c)} \left\{ u(w_c(n_c)) + \beta EV_s(n_c') \right\} dF(a) +$$

$$\int_{a_c^*(n_c)}^1 \left\{ u(w_c(n_c) - e) + \beta\pi(a)EV_c(n_c') + \beta(1 - \pi(a))EV_s(n_c') \right\} dF(a),$$

$$EV_s(n_c) = \int_0^{a_s^*(n_c)} \left\{ u(w_s(n_c)) + \beta EV_s(n_c') \right\} dF(a) +$$

$$\int_{a_s^*(n_c)}^1 \left\{ u(w_s(n_c) - e) + \beta\pi(a)EV_c(n_c') + \beta(1 - \pi(a))EV_s(n_c') \right\} dF(a), \tag{7}$$

where, n_c' is given by $\Phi(n_c)$.

- $\Phi(n_c)$ *is consistent with parental decisions. That is, the law of motion for n_c (the transition function) satisfies:*

$$\Phi(n_c) = n_c \int_{a_c^*(n_c;\Phi)}^1 \pi(a)\, dF(a) + (1 - n_c) \int_{a_s^*(n_c;\Phi)}^1 \pi(a)\, dF(a). \tag{8}$$

Definition 2 *A **steady state** is a competitive equilibrium with $n_c = n_c^* \in [0, 1]$, which satisfies $\Phi(n_c^*) = n_c^*$.*

On a steady state, the wages, reservation abilities, expected utilities, and the fraction (measure) of college educated parents are all constant over time. Analytically, a steady state is fixed point of the transition function Φ.

3.3 Dynamics

In the appendix, we derive the following equation by manipulating the equation system (7):

$$\Lambda(n_c) = x(n_c) + \left[\beta \int_{a_c^*(n_c)}^{a_s^*(n_c)} \pi(a)\, dF(a) \right] \Lambda(\Phi(n_c)), \tag{9}$$

where,

$$x(n_c) \equiv [F(a_c^*(n_c))u(w_c(n_c)) + (1 - F(a_c^*(n_c)))u(w_c(n_c) - e)] -$$
$$[F(a_s^*(n_c))u(w_s(n_c)) + (1 - F(a_s^*(n_c)))u(w_s(n_c) - e)]. \tag{10}$$

Equation (9) has an equilibrium Euler equation interpretation, and is the central relationship characterizing the model. Here, x is extra contemporaneous (*ex ante*, expected) utility a college educated parent gets, taking into account the endogenous response of the higher wage earning college parent having a higher probability of sending a child to college (recall Lemma 3). The value of being college educated has two components – a contemporaneous utility gain and a discounted future value. If utility is linear, the utility cost for both types of parents is the same. Equation (4) then implies that $a_s^* = a_c^*$, and the second term on the right hand side of (9) vanishes. That is, the only advantage of being college educated is the greater utility resulting from higher w_c. But for a concave utility function, the second term is positive – there is an additional gain that results from the parent being able to send more of her children to college, more than a school educated parent will be able to. The discounted value from these additional college educated children feeds into today's value. This is what makes the parent's problem truly dynamic.

We repeat the equations resulting from (4):

$$\beta\pi(a_c^*(n_c))\Lambda(\Phi(n_c)) = g_c(n_c) \tag{11}$$
$$\beta\pi(a_s^*(n_c))\Lambda(\Phi(n_c)) \leqslant g_s(n_c), \; w.e.i \; a_s^*(n_c) < 1. \tag{12}$$

Equations (8) through (12) are four *functional* equations in the four functions Λ, Φ, a_c^*, and a_s^*. We can characterize the dynamics of the economy by characterizing these four functions. When n_c is replaced by n_c^*, we get four equations in four steady state *quantities*.

The equation (9) is an equilibrium expression, with the ability thresholds and the transition function being endogenous quantities. There is no obvious reason why the functional equation should have a solution. Since we cannot relate this system of functional equations to any existing setup, we prove (in the appendix), by providing reasonable conditions, that a unique solution exists to this system. To simplify the algebra involved, we specialize to the case of a linear probability function ($\pi(a) = a$) and uniform distribution of ability ($F(a) = a$). These simplifying assumptions are also consistent with agnosticism regarding these directly unobservable functions.[16]

Since $w_c(n_c)$ is a decreasing function, it seems reasonable to expect the extra value obtained by being college educated, Λ, to also be a decreasing function (provided $\Phi(n_c)$ is increasing and the endogenous response of a_c^* and a_s^* do not cause perverse effects in the contemporaneous extra utility x). We, therefore, work in the space of weakly decreasing functions and study the mapping T defined by:

$$T\Lambda(n_c) = x(n_c) + \left[\beta\int_{a_c^*(n_c)}^{a_s^*(n_c)} \pi(a)\,dF(a)\right]\Lambda(\Phi(n_c)),$$

---

[16] These assumptions also imply that the fraction of college enrollees who are successful is bounded by $1/2$. (This follows from the fact, $n_c\int_0^1 a\,da + (1 - n_c)\int_0^1 a\,da = 1/2$.) This is clearly unrealistic – for the calibration we use more general probability functions and also assume $\pi_c(a) > \pi_s(a)$, $\forall\, a \, \epsilon \, [0, 1]$.

in conjunction with (8), (11), and (12).

When n_c increases, it can be seen from (11) that $a_c^*(n_c)$ increases. The utility cost of sending a child to college (the right hand side) increases with n_c, while the benefit Λ decreases (provided the anticipated state tomorrow, $\Phi(n_c)$ is increasing). This in turn causes the reservation ability for the marginal parent to increase. However, for the school educated parent, the two effects work in opposite directions. When n_c increases, the utility cost decreases, exerting a downward pressure on $a_s^*(n_c)$ (the income effect); Λ decreases, which causes the parent to shift away from the child's college education into personal consumption, exerting an upward pressure on $a_s^*(n_c)$ (the substitution effect). In the appendix, we provide sufficient conditions to ensure that the income effect is strong enough relative to the substitution effect, the transition function increases, and the above-mentioned perverse effect in x is avoided. These conditions allow us to prove the following theorem.

Theorem 1 *The equilibrium mapping, T, is a contraction. A unique solution to the functional equations (8) through (12), and hence a unique competitive equilibrium exists.*

To provide a sharp characterization of the dynamics, we specialize to the isoelastic (CRRA) utility function:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \ \sigma \geqslant 0,$$

with the $\sigma = 1$ case interpreted as $\log(c)$. We want to argue that the curvature parameter, σ, plays an important role in the types of equilibria that we can obtain. It is, therefore, useful to consider the linear utility case, $u(c) = c$, first. As mentioned above, in this case $a_s^* = a_c^*$, and there is no real dynamics.

Proposition 1 *(Linear Utility): When $\sigma = 0$, there is a unique steady state which is reached in a single step.*

Once we prove the existence and uniqueness of equilibrium, we rely on computations to characterize the types of dynamic equilibria that obtain – that is, the shape of the transition function that results. The effect of the assumptions made on the income and substitution effects is such that $\Lambda(\Phi(n_c))$ decreases at a slower rate than does $g_s(n_c)$. If for any $\widehat{n}_c$, $a_s^*(\widehat{n}_c) = 1$, that is, $\beta\Lambda(\Phi(\widehat{n}_c)) < g_s(\widehat{n}_c)$, then it follows that $a_s^*(n_c) = 1$, $\forall \ n_c < \widehat{n}_c$. In the interval where no school-educated parent sends a child to college, the transition function is:

$$\Phi(n_c) = n_c \int_{a_c^*(n_c;\Phi)}^{1} \pi(a)\, dF(a) < n_c.$$

Graphically, the transition function is a straight line that lies below the 45° line. Any economy that starts out in this region will asymptotically reach the origin. This fact is useful in discussing the behavior of the economy in the neighborhood of $n_c = 0$.

It is helpful to distinguish among three cases for the curvature parameter σ, "high", "low", and "medium" while discussing the types of dynamic equilibria.

- **"Low"** σ**:** For a given set of production parameters, when σ is low the utility cost of school educated parents is low even when $n_c = 0$ (note that the wages of school educated parents are highest when $n_c = 1$ and lowest when $n_c = 0$). School educated parents will always send their children to college, $\Phi(n_c)$ is concave and intersects the 45 degree line once. A unique steady state with a positive n_c^* obtains.

- **"Medium"** σ**:** For intermediate values of σ, $a_s^*(n_c) = 1$, for low values of n_c. Once the school educated parent's wage becomes high enough (that is, n_c becomes large enough), $a_s^*(n_c) < 1$, $\Phi(n_c)$ increases at a rapid rate to cross the 45 degree line, and then becomes concave again to re-intersect the 45 degree line. There are three steady states, zero and two positive steady states, with the lower of the two positive steady states being unstable.

- **"High"** σ: When σ is high, the utility cost of school educated parents is very high. School educated parents will either never send their children to college or start sending them to college only for high wages. The increase in the fraction of college educated labor is low enough to cause $\Phi(n_c) < n_c$ in the entire interval in this case. The only possible steady state then is $n_c^* = 0$. That is, a *development trap* obtains.

The dynamic behavior of the economy around the origin is governed mainly by the utility cost of school educated rather than college educated parents. Rich parents always send a positive fraction of their children to college and especially so when their wages are very high ($n_c \rightarrow 0$). But given that they are a very small fraction of the labor force when n_c is in the vicinity of 0, their behavior matters little to the dynamics of the economy. Whether the fraction of college educated workers continues to grow in the vicinity of $n_c = 0$, and if so whether it grows at a rate that can sustain a long run equilibrium with a positive fraction of such workers, depend on the behavior of the poor parents.

These three types of equilibria can also be illustrated with constant σ but varying cost e. Increasing either would have the effect of increasing the utility cost $g_s(n_c)$, especially in the vicinity of $n_c = 0$. It is clear that the fixed nature of the goods cost of education is responsible for the trap. One could envision an alternate setup in which the time cost of an old agent (the "teacher") is the cost of education. If the teacher is the parent of the child herself, the cost of educating the poor would be very low when the wages of the poor are low, and a trap is less likely obtain. However, if the cost of educating the child is the time cost of a *skilled* worker, which seems a lot more realistic, a trap would continue to obtain. In fact, the situation would be exacerbated since the wages of the skilled are highest when the wages of the unskilled are at their lowest.

In Section 4, we explore whether a simple tax and subsidy policy can nudge an economy out of a trap equilibrium.

3.4 Computation

The equations that describe the dynamics of the system are highly non-linear, which limits the extent of analytical characterization we can provide. We use numerical algorithms to compute and more

completely characterize the model. A detailed outline of the algorithm used for computation is provided in the appendix. In this subsection we provide only a synopsis.

As mentioned earlier, equations (8) through (12) are four functional equations in the four functions Λ, Φ, a_c^*, and a_s^*. The computation involves an iteration scheme to find the fixed points of the transition function Φ, and the costate Λ. Suppose parents take as given these functions. Then for any given state n_c, they can solve for $a_s^*(n_c)$ and $a_c^*(n_c)$. (When (12) holds with inequality, $a_s^*(n_c)$ is set to 1.) These reservation abilities can be used in (8) to update Φ, and in (9) to update Λ to be used in the next iteration. In practice, we use Chebyshev approximation and interpolation and compute the functions Φ and Λ at a discrete set of grid points in $[0,1]$.[17] See Judd (1998) for details on this method.

3.5 Numerical Examples

We postpone issues on calibration and parameters chosen until the next section (where we discuss the subsidies that are necessary to match US data), and present numerical examples here with the sole purpose of illustrating the three types of equilibrium mentioned above. The lack of subsidies explains the higher than realistic skill premiums. We keep the production and probability functions fixed and vary only σ, the curvature of the utility function. The resulting steady states are shown in the following Table 1.

<div align="center">

Table 1

Steady states with different σ

σ	n_c^*	a_c^*	a_s^*	w_c^*/w_s^*	Y
2.0 ("Low")	.244	.049	.762	2.06	.474
2.65 ("Medium")	.108	.005	.919	3.45	.398
2.65	.054	.001	.963	4.89	.348
2.65	.000	.000	1.000	10.22	.264
4.0 ("High")	.000	.000	1.000	10.22	.264

</div>

---

[17] There are cases where $a_s^* = 1$ for an interval of n_c. This causes Φ and Λ to be kinked at the point where a_s^* becomes < 1. The Chebyshev method has a tendency to smooth out functions and miss out on such kinks. So, we used a lot more grid points than is normally used with this method. FORTRAN programs used to compute numerical solutions are availabe from the authors upon request.

Figures 1 - 4

$\Phi(n_c)$ when σ=2

a_c and a_s when σ=2

w_c and w_s when σ=2

$\Lambda(n_c)$ when σ=2
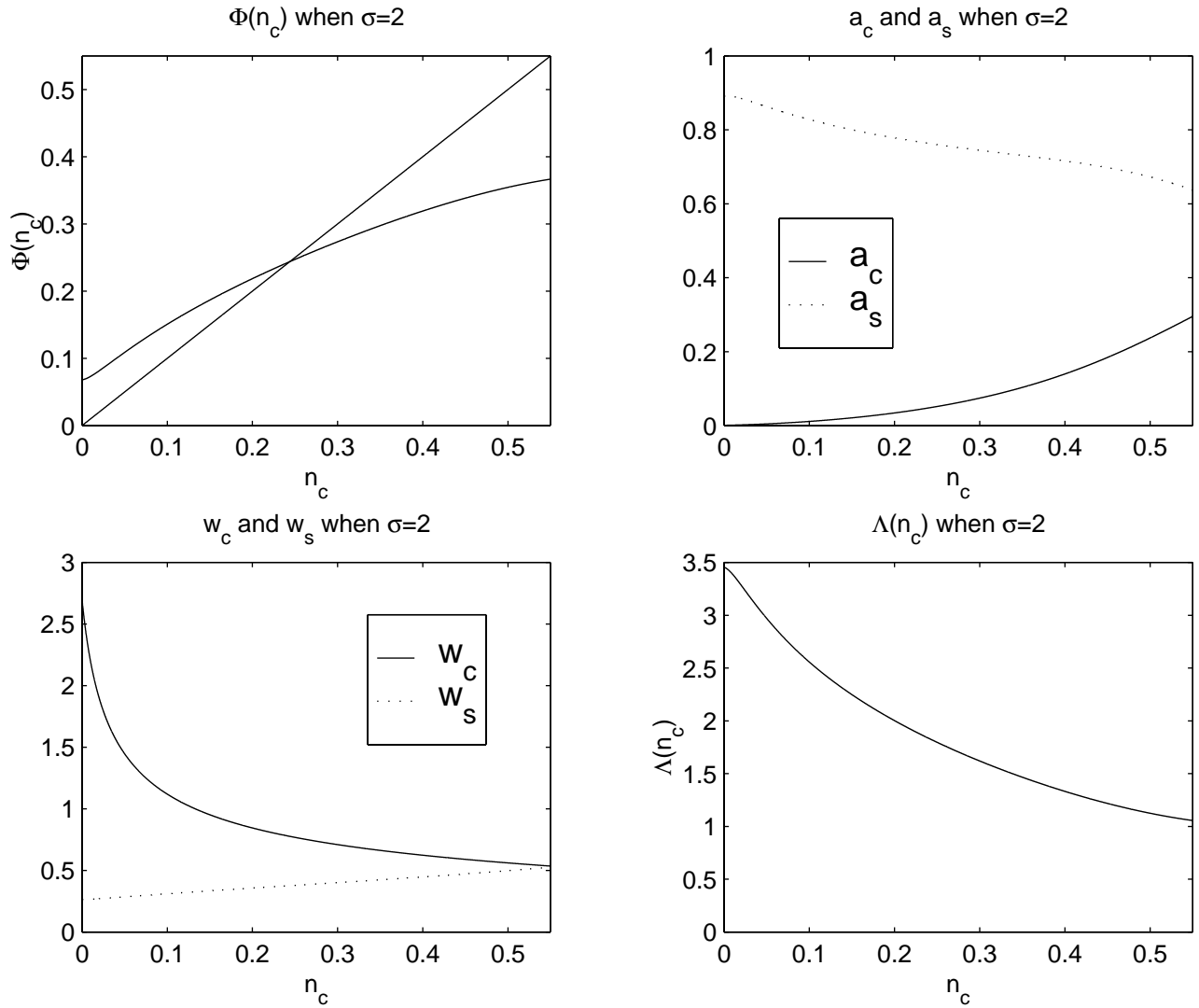
When $\sigma = 2$ ("low" σ), we get a unique positive steady state. Figures 1 through 4 graph Φ, a_c^*, a_s^*, w_c, w_s and Λ, as functions of n_c for this case. Here, in the long run about 24% of the labor force is college educated, the rich send around 95% of their children to college and the poor send close to 24% of their children to college.

Figures 5 - 8

Φ(n_c) when σ=2.65

a_c and a_s when σ=2.65

w_c and w_s when σ=2.65

Λ(n_c) when σ=2.65
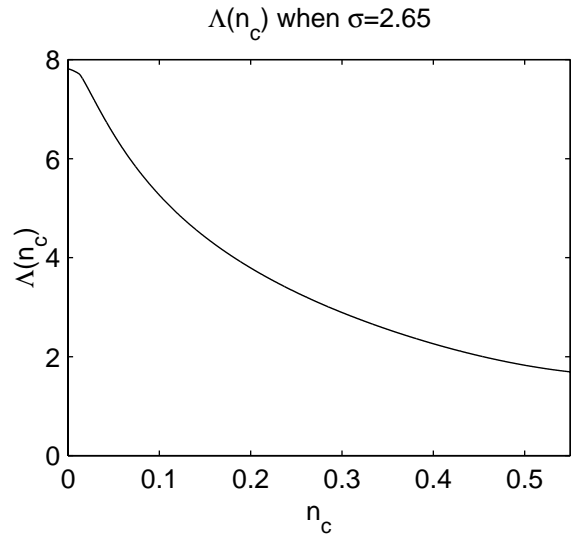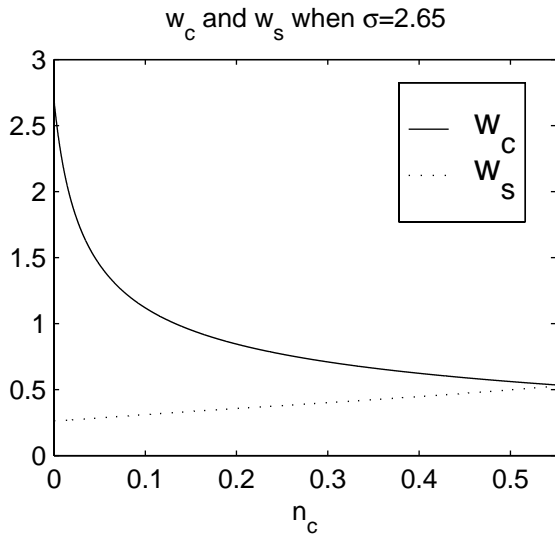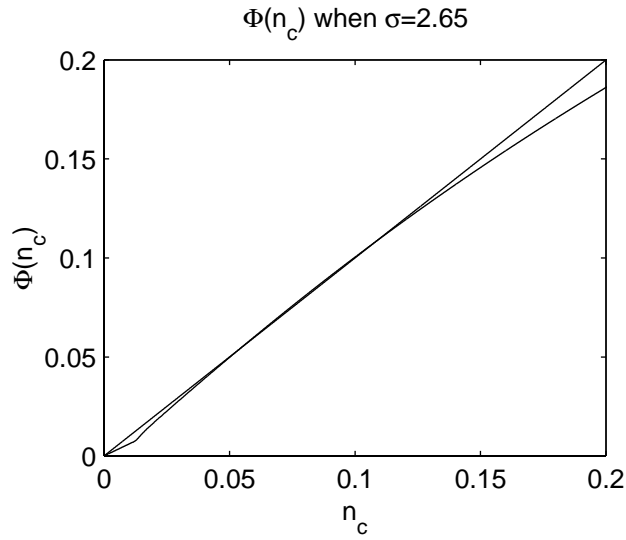
When $\sigma = 2.65$ ("medium" σ), we get the second type of equilibrium, with a low ($n_c^* = 0$) as well as a high ($n_c^* = 0.10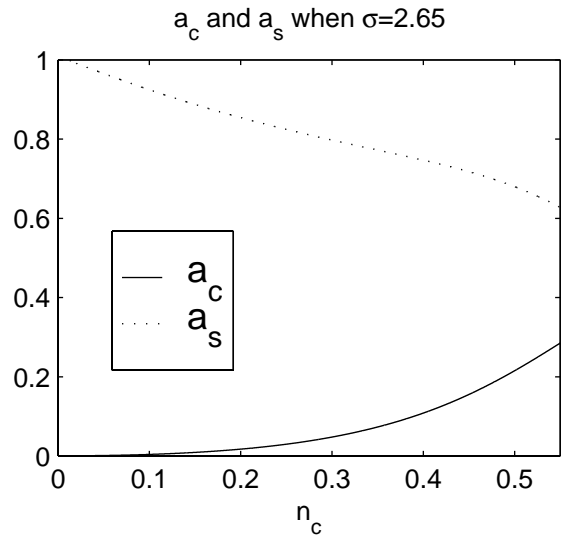8$) steady state. The middle ($\widetilde{n}_c = 0.054$) steady state is an unstable one. If an economy starts with an $n_c < \widetilde{n}_c$, the low steady state is reached. For any other starting value for n_c, the higher steady state is reached. Note that a_s^* stays at 1 for a brief interval near the origin. Figures 5 through 8 show the functions for this case.

Φ(n_c) when σ=4

a_c and a_s when σ=4

w_c and w_s when σ=4

Λ(n_c) when σ=4
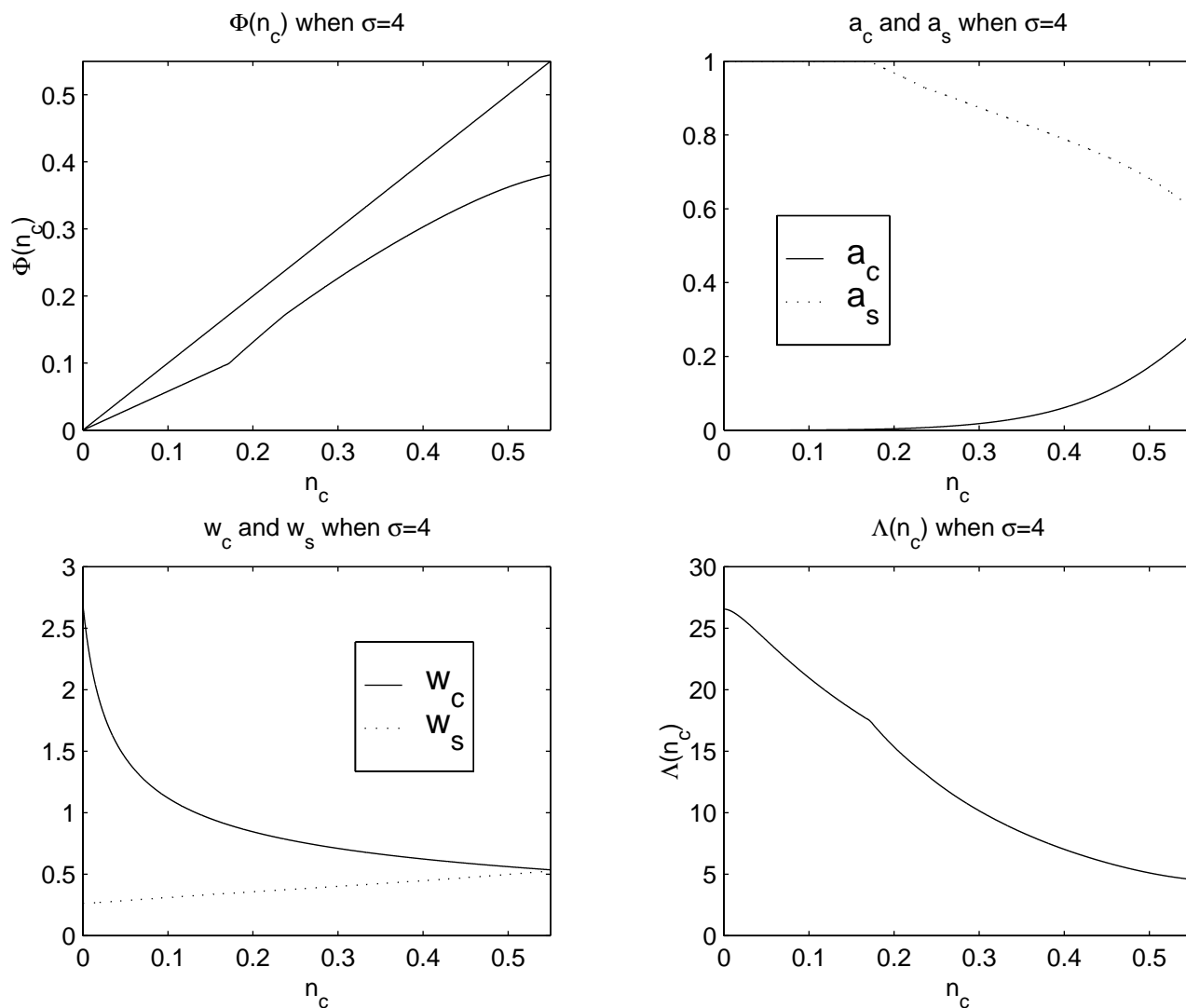
Finally, when $\sigma = 4$ (high σ), we get the third type of equilibrium. The only steady state is a "trap", with $n_c^* = 0$, to which the economy converges, no matter where it starts (Figures 9 through 12.). The interval in which $a_s^* = 1$ is much wider.

4   The Government and Higher Education Policy

Thus far, we have studied the behavior of an economy with only private agents. This characterization is essential to understanding the policy analysis that is conducted in this section. As mentioned earlier, the assumption on private information about a child's ability has no bearing on the private economy, since the parent is liquidity constrained anyway. No trade or exchange is affected by the private information. But it will matter when the government is introduced.

In this model, the rich parents are more likely to send the child to college than the poor one (see Lemma 3). Therefore, in a liquidity constrained economy, less able rich children are sent to college,

while more able poor children do not get the opportunity to go to college. For instance, in the first row of Table 1, the marginal poor child has a probability of success of 50%, while the rich child has only a 11% chance of graduating. This will certainly be the case for $\pi_c = \pi_s = \pi$, as assumed in the previous section. Even if $\pi_c(a) > \pi_s(a)$, provided the difference is not too large, $\pi_c(a_c^*) < \pi_s(a_s^*)$ will obtain.[18] If aggregate educational efficiency is measured as the amount of resources invested per student who successfully completes college, the fully private economy might not provide an efficient outcome. That is, by redirecting resources from educating the marginally able rich children to the more able poor children, it would seem that educational efficiency (though not necessarily, aggregate welfare) would be improved. This measure is also likely to improve equality of opportunity. Thus a natural policy that the government could consider in this setup is the subsidization of the poor children.

This is where the informational assumption has an effect. Since the government cannot observe the ability of poor children, it will have to subsidize *all* poor children to the same extent – it cannot index the subsidy to the child's ability. That is, all poor parents face a lower cost of college, which might cause them to send more as well as less able children to college. This is a countervailing force to the one mentioned above, and can actually *decrease* educational efficiency. In our setup, given the higher response of the school educated parents to policy, the countervailing effect dominates and education efficiency decreases in most cases. The monetary rate of return would likewise decrease.

Similar considerations apply if the government is under political or other constraints to provide need-based, rather than merit-based aid.[19] For the sake of completeness, we also consider the case where the government can observe ability and subsidize students according to their ability.

Labor is inelastically supplied in this model. Therefore, we do not have the loss in productive efficiency that is usually found in models where taxes distort labor supply decisions. Since the portion of income taxes that go toward higher education financing is likely to be very small, it seems that we are not missing out on first order effects by assuming inelastic labor supply. However, there is another dimension in which productive efficiency matters. Production depends on skilled and unskilled labor, and for an arbitrary production function there is likely to be an optimal mix of these two types of labor. A social planner would therefore choose a subsidy scheme that yields this optimal mix in equilibrium. With assumption 3, this optimum is a corner solution (that is, $n_c = \overline{n_c}$, the highest possible equilibrium

---

[18]Likewise, even if ability is transmitted from the parent to the child, as long as the transmission is not perfect, there will be an element of this similarity in potential that is independent of parentage.

[19]The model in this paper is not set up to consider the design of an incentive compatible scheme that would allow people to reveal their true abilities and get commensurate subsidies. For instance, if the government gives sudsidized loans, it can consider schemes where the loan amount and the interest rate were specified together. However, "limited liability" (no repayment in the case of failure at college) can confound the institution of such a scheme. This issue deserves attention in the future.

On a related note, Fernandez and Gali (1997) compare the outcomes of markets and tournaments in assigning students of different income and ability to schools of different quality, when there are perfect credit markets and when there are not. Markets assign students based on their willingness to pay, and tournaments assign students based on their performance in a contest. When there are credit constraints, they find that tournaments dominate markets in terms of aggregate output and at times in terms of aggregate consumption as well. It will be useful to extend our model by having colleges of different qualities and studying the assignment problem that Fernandez and Gali study.

n_c, maximizes output). This is expensive and even physically impossible to achieve, but the planner would want to subsidize education and increase n_c if the aim is to maximize steady state production. As long as the return to college education is high enough, this will maximize an aggregate measure of consumption as well.

We can also study the effect of subsidies on a weighted sum of welfare. If the government values all citizens equally, the welfare function at steady state will be given by:

$$n_c^* EV_c(n_c^*) + (1 - n_c^*) EV_s(n_c^*) = EV_s(n_c^*) + n_c^* \Lambda(n_c^*).$$

Any subsidization policy that increases the steady state fraction of skilled people will cause a higher fraction of the population to be in the higher utility state, but will lower the extra utility gained; the weight n_c^* increases, but $\Lambda(n_c^*)$ decreases. Since all parents are taxed but only the poor college attendees get a subsidy, the tax burden will also have a negative effect on both EV_s and Λ. Steady state welfare can therefore increase or decrease with any policy that increases n_c^*.

The policy that is efficient in the long run is not necessarily optimal when the transition is also considered. The sacrifices made in getting to the steady state that is better in an aggregate sense might not be worth the ultimate benefit of staying there in the long run, especially for the currently rich who will have to make most of the sacrifices. For this reason, we study transitional as well as steady state effects on welfare while comparing different policy regimes.

One reasonable question at this stage is, "Why do we care about the effect of subsidies on the positive efficiency measure, especially in those cases where the aggregate welfare increases?" There are at least three reasons why one might want to study the positive measure. First, there can be little controversy on the definition of such a measure. While we chose a symmetrically weighted welfare measure because it seemed least ad-hoc, there is no reason to expect it to be the universally accepted measure. This is especially true when there are opposite effects on the welfare of the rich and the poor that arise from the redistributive policies considered. Second, it is useful to know if the effect on the positive measure differs from the one on the normative measure, which it does in our case. If we envision a "Director of Education" who makes education budgeting decisions, this person has a stronger incentive to be concerned about the efficiency of the education sector than about aggregate welfare. Finally, we have not explicitly modeled the moral hazard involved in the process of human capital accumulation that causes the incompleteness of markets in the first place. The process of completion of the market by the government can increase the moral hazard and decrease welfare.

Even this preliminary discussion has uncovered several issues regarding government subsidization of poor children. In this section we explore some of these issues in greater detail.

4.1   A Simple Tax and Subsidy Scheme

Throughout this section, we will study only one kind of policy. All workers will have their income taxed at the rate τ. All school educated workers will get a college subsidy of s, *if* they send their children to college. That is, they will face an effective college tuition of $e - s$. Since the poor receive

a much higher level of subsidy in practice, such a policy assumption is empirically relevant.[20] From a calibration point of view, this would mean that we assign all subsidies seen in practice to the poor children.

The structure of the problem remains the same as in (1), but the current-period utility terms differ. College educated parents get $u((1-\tau)w_c(n_c))$ if they do not send their children to college, and $u((1-\tau)w_c(n_c) - e)$ otherwise. The corresponding utilities for the school educated parents are $u((1-\tau)w_s(n_c))$ and $u((1-\tau)w_s(n_c) - (e-s))$. In transitions, to impose some discipline on the experiments, we will consider a subsidy that is a constant percentage of total expenditure, but will vary τ over time to balance the government's budget. Therefore, $\tau \equiv \tau(n_c)$.

Since, the structure of the parent's problem is the same, we will get the same reservation ability behavior, but reservation abilities will now be functions of the policy parameters as well. They can be written as $a_c^*(n_c, \tau(n_c), s)$, $a_s^*(n_c, \tau(n_c), s)$. We require the government's budget to be balanced every period. The government's budget constraint is:

$$(1-n_c)(1 - F(a_s^*(n_c, \tau(n_c), s))) s = (n_c w_c(n_c) + (1-n_c)w_s(n_c)) \tau(n_c). \tag{13}$$

Here, $(1-n_c)(1 - F(a_s^*(n_c, \tau, s)))$ is the number of children of school-educated parents sent to college, and hence the left hand side is the total amount of subsidies paid by the government in equilibrium. These expenses have to be borne even for children who drop out, and hence π does not appear anywhere in the expression. The right hand side is the total amount of taxes collected from all parents.

4.2 Efficiency of the Education Sector

We provide a simple measure of efficiency for the education sector – a ratio of the number (measure) of students graduating from college to the total resources expended in educating them, *irrespective* of who spends this, the parent or the government. We define:

$$eff = \frac{n_c'}{[n_c(1 - F(a_c^*)) + (1-n_c)(1 - F(a_s^*))] e}.$$

As in (13), π does not appear anywhere in the expression, since expenses have to be borne even for children who drop out. The success rate (conditional probability of success), for the rich children is $S_c \equiv \frac{\int_{a_c^*}^1 \pi_c(a)\, dF(a)}{1 - F(a_c^*)}$, and for the poor children is $S_s \equiv \frac{\int_{a_s^*}^1 \pi_c(a)\, dF(a)}{1 - F(a_s^*)}$. Using (8) for n_c', and the above success rates, the formula for efficiency can be written more revealingly (with the state variable argument suppressed for simplicity) as:

$$eff = \left\{ \frac{n_c(1 - F(a_c^*))}{n_c(1 - F(a_c^*)) + (1-n_c)(1 - F(a_s^*))} S_c + \frac{(1-n_c)(1 - F(a_s^*))}{n_c(1 - F(a_c^*)) + (1-n_c)(1 - F(a_s^*))} S_s \right\} \frac{1}{e}. \tag{14}$$

The efficiency depends on a weighted sum of the success rates of the two types, where the weight is the share of each type of enrollees. It is easy to show that each of the success rates is increasing in

---

[20]In 1992-93, the students in the lowest income bracket received aid that was 48% of total costs, while those in the highest income bracket received 12% of total costs as aid. (The Condition of Education, 1996, Supplemental Table 13-1.) If one adjusts for the progressivity of taxes, the effective difference is likely to be even higher.

the relevant a^*, since the quality of the enrolling pool is higher. So any policy that redistributes funds from the rich to the poor would have the effect of increasing S_c and decreasing S_s. As mentioned in the introduction to this section, since the poor are located in the more responsive part of the utility function the increase in their enrollment causes the negative effect on the efficiency to dominate.[21]

An alternate quantity to use would be the monetary rate of return. At an individual level this rate of return is $\pi(a)\left(w_c(n_c') - w_s(n_c')\right)/e$. It is easy to show that at the aggregate level, this rate of return is just $eff \cdot \left(w_c(n_c') - w_s(n_c')\right)$. Since the wage gap is decreasing in n_c', the return will drop even more steeply when a subsidization policy decreases efficiency as defined above; so, we present details only on the efficiency variable in all policy experiments.

4.3   A Welfare Measure

In addition to the above-mentioned efficiency measure, we also consider a symmetric welfare measure by summing up welfare over all parents in the economy. Such an equal-weighted measure seems to be the least *ad hoc* one. Thus, the welfare in each period is the measure-weighted sum of utilities over college educated parents who send their children to college and those who do not, and school educated parents who do and do not:

$$
\begin{aligned}
W_t(n_c) &= n_c[(1 - F(a_c^*))u(w_c(n_c) - e_c) + F(a_c^*)u(w_c(n_c))] + \\
&\quad (1 - n_c)[(1 - F(a_s^*))u(w_s(n_c) - e_s) + F(a_s^*)u(w_s(n_c))],
\end{aligned}
$$

where e_c and e_s are the tax or subsidy adjusted education expenditures. The welfare at a steady state is the discounted sum of the welfare in each period: $W_{ss}(n_c^*) = \frac{1}{1-\beta}W_t(n_c^*)$. If we want to calculate the welfare of going from one steady state to another, we calculate the discounted sum of period welfare along the transition and then add the appropriately discounted steady state welfare, $W_{transition} = \sum_{i=0}^{T} \beta^i W_i(n_c) + \beta^{T+1} W_{ss}(n_c^*)$, where T is the number of periods after which we are "very close" to the steady state.

We can translate these welfare numbers into a "compensating equivalence" consumption measure. Suppose we want to evaluate a policy change. We can calculate, ω, the percentage by which consumption in *each* period needs to be increased (decreased) in order to be as well off under the benchmark as under the policy change. In other words, we calculate ω such that the $W_{transition}$ given above is related to the benchmark welfare according to the expression:

$$
\begin{aligned}
W_{transition} &= \frac{1}{1-\beta}\{n_{c,b}^*[(1 - F(a_{c,b}^*))u(\omega(w_c(n_{c,b}^*) - e_{c,b})) + F(a_{c,b}^*)u(\omega w_c(n_{c,b}^*))] \\
&\quad + (1 - n_{c,b}^*)[(1 - F(a_{s,b}^*))u(\omega(w_s(n_{c,b}^*) - e_{s,b})) + F(a_{s,b}^*)u(\omega w_s(n_{c,b}^*))]\},
\end{aligned}
$$

where, the subscript b on the right hand side steady-state quantities refer to the benchmark case. Given the homogeneity of the utility function, it is easy to see that, $\omega = \left(\frac{W_{transition}}{W_{benchmark}}\right)^{\frac{1}{1-\sigma}}$. Even though they

---

[21] The weights will also be altered by such redistributive policies. But if the elasticity of a_c^* is not too high, the weights will move in the same direction as the success rates.

are dropped for notational simplicity, the relevant tax rates are applied to the wages to get the welfare measure.

4.4 Calibration

The difficulties of calibrating a two-period model to the actual economy are obvious. However, there are some mitigating factors in this setup. First, even if we were to expand our model to a several-period OLG model, for any given family there will be little happening during most of the periods. This is not a model of continuous asset trading or portfolio choice – parents have a single, discrete decision to make. Second, the steady state predictions are interesting in their own right, and it is not clear that a several-period OLG model will lead to significantly different results in this regard. Finally, the various effects in the model are a lot more transparent in this simple setup, and we can therefore provide some analytical characterization. So, we set aside the task of computing a more realistic large scale OLG model for future research, and present our attempt at calibrating the model to US data. The aim is to get a set of benchmark parameters with which it is sensible to conduct policy analysis.

We start by assuming values for certain parameters that are commonly used in the literature. The generational discount factor is set at $\beta = 0.55$, which corresponds to a yearly discount factor of 0.98 compounded over 30 years. We set $\nu = 0.35$, which corresponds to an elasticity of substitution between college and high school equivalents of 1.54. Autor, Katz, and Krueger (1998) report that the emerging consensus on this elasticity is approximately 1.4 to 1.5. From the *Condition of Education 1996* which has data for the 1992-93 academic year, we approximate total college yearly expenditure of $15,000. From the Digest of Education Statistics 1998, we get the median family income in 1992 to be $36,573. We calculate the present value of educational expenses over 4 years and yearly income over 30 years, at a discount rate of 8% to get e/Y to be around 0.12. Using a $Y = 0.5$ from preliminary computations of the model, we estimate the expenditure parameter e to be 0.06.[22] From the *Condition of Education 1996,* we also get a subsidy of 48% for the low income group (enrollment-weighted total aid divided by total cost at private and public institutions). Therefore, given the e of 0.06, we set s to be 0.03. We use a fairly standard value for σ of 2. These are the parameters for which we have independent evidence.

We assume a uniform ability distribution in $[0, 1]$; that is, $F(a) = a$. We then choose the remaining production parameters $(A, \theta, \epsilon, \text{and } \gamma)$ and the probability functions $(\pi_c(a), \pi_s(a))$ to broadly match enrollment and dropout rates, as well as college attainment and the skill premium.

The benchmark parameters used are summarized below:

$$\text{Production} \quad : \quad A = 1, \ \theta = 0.5, \ \nu = 0.35, \ \epsilon = 0.1, \ \gamma = 0.02,$$
$$\text{Preference} \quad : \quad \beta = 0.55, \ \sigma = 2,$$
$$\text{Education} \quad : \quad e = 0.06, \ s = 0.03, F(a) = a, \ \pi_c(a) = a^{0.74}, \ \pi_s(a) = 0.66a^{0.9}.$$

---

[22]Note that this is the discount rate that is used *within* a period of the model to map the data into the model's framework, as opposed to the intergenerational rate implicit in the β that is built into the model. Discount rates in the range of 5% to 8% all yielded e/Y measures in the range of 0.10 to 0.12. Since the Y used is from preliminary computations, we verify that the e/Y that results in the benchmark equilibrium (0.118) is close to that assumed.

We compare the steady state outcome of the model to US data in the following table.

Table 2

Comparison of Model Outcomes with US Data

Empirical Quantity	Model variable	US data	Year	Source	Model Outcome
% labor force college educated	n_c^*	38.6% (full-time equiv.)	1990	Autor, Katz, Krueger (AKK) (1998)	35.7%
Enrollment rate for rich (%)	$1 - F\left(a_c^*\right)$	76.6% ("high" income)	1990	The Condition of Education (1998)	80.0%
Enrollment rate for poor (%)	$1 - F\left(a_s^*\right)$	46.7% ("low" income)	1990	The Condition of Education (1998)	50.2%
Dropout rate for rich (%)	$\frac{\int_{a_c^*}^1 (1-\pi_c(a))\, dF(a)}{1-F(a_c^*)}$	30.2% to 43.0% (2 yr) 15.6% to 19.6% (4 yr)	1989-90 (starting year)	Digest of Education Statistics (1998)	32.5%
Dropout rate for poor (%)	$\frac{\int_{a_s^*}^1 (1-\pi_c(a))\, dF(a)}{1-F(a_s^*)}$	43.4% to 45.1% (2 yr) 25.5% to 29.3% (4 yr)	1989-90 (starting year)	Digest of Education Statistics (1998)	49.2%
Skill premium	w_c^*/w_s^*	1.58 $1.66 - 1.73$	1989 (avg.) 1990	Murphy, Welch (1992) AKK (1998)	1.54
Govt. exp. as % of GDP	$\frac{(1-F(a_s^*))s}{Y}$	0.55%	1989	McPherson and Schapiro (1991)	1.9%

The model's outcomes are broadly consistent with the corresponding quantities seen in US data. The percentage of population that is college educated is very close to the 38.6% full-time college equivalents reported by Autor, Katz, Krueger (1998). It is also in between the range of 27.3% who have completed higher education and 45.2% who have some higher education as reported by Barro and Lee (1996). The enrollment rates are consistent with the enrollment rate at income brackets at either end of the distribution as reported by the *Condition of Education* (1998).[23] The enrollment gap of about 30% is almost the same as the one seen in data. The model dropout rates are higher, especially if the four-year rates are considered. Data is reported by income quartiles, and the lower ranges given in the table correspond to the higher income quartiles.[24]

---

[23]Reallocation of the 54.4% enrollment rate for the "middle" income group reported for the US to the "low" and the "high" group is likely to increase the rate for the former and decrease it for the latter. The model's outcome would then be closer to the data for the enrollment of the poor, but higher than what is found in the data for the rich.

[24]The data shown corresponds to the "no degree, not enrolled" column of the *Digest of Education Statistics* table. There is a separate column titled, "no degree, still enrolled," which corresponds to students who have not graduated on time. Given the high likelihood that some of these students will also eventually drop out, the model outcomes are likely to be closer to the data if a portion of these students are also counted as dropouts.

The skill premium is very close to the Murphy and Welch (1992) figure and not too far from the Autor, Katz, and Krueger (1998) estimates. The government subsidy expenditure as a fraction of GDP is higher than but in the same order of magnitude as the one calculated for US data.[25]

Given this close match between the model outcomes and data, the computational results in the next section can be viewed as a numerical illustration of the concepts discussed at the introduction to this section as well as outcomes of higher education policy experiments conducted on the US economy.

4.5 Policy Experiments

In this section we consider the effects of the tax and subsidy policy on the transitional and long run outcome of the economy. The social planner's problems that we solve are of the "Ramsey" type. That is, the planner takes as given the actions of individual agents in response to policy parameters, and chooses the policy vector to maximize some objective function. The firm's actions are taken as given too, since the wages are given by the MPLs. This is in contrast with a scheme in which the planner collects all resources and allocates them as consumption and education expenditures of both types of parents, and organizes production as well. In other words, the planner's (government's) policy is purely an educational one and *not* used for general purpose redistribution. The latter is clearly outside the scope of this paper. We will refer to the following Table 3 when we discuss the outcomes of the policy experiments.

Table 3

Policy changes when $\sigma = 2$

Policy	n_c^*	a_c^*	a_s^*	$1 - S_c$	$1 - S_s$	w_c^*/w_s^*	Y	τ	s	sub	eff	ω	ω_{ss}
Benchmark	.357	.200	.498	.325	.492	1.54	.509	.019	.030	0.0097	9.637	1.0000	1.0000
Equal opp.	.382	.311	.311	.275	.550	1.45	.515	.036	.043	0.0185	9.245	1.0012	1.0112
Max. n_c^*	.389	.395	.102	.239	.618	1.43	.519	.058	.055	0.0300	8.257	0.9931	1.0051
Eliminate sub.	.252	.054	.754	.396	.414	2.02	.477	0	0	0	9.934	.9738	.9086

Here, $1 - S_c$ and $1 - S_s$ are the dropout rates for the two types, and *sub* is the amount of subsidies given at the steady state and is equal to $(1 - F(a_s^*))(1 - n_c^*) s$. The outcome with the benchmark parameters mentioned in the previous subsection is given in the first row. Outcomes of new policies will be compared to the benchmark outcome.

4.5.1 Equality of Opportunity

Equality of educational opportunity in this model is achieved when $a_c^* = a_s^*$. A child of ability a will either attend or not attend college based on a common threshold ability – the type of her parent

---

[25]The US figure is computed using the nearly $32 Billion reported by McPherson and Schapiro (1991) as total federal, state, and institutional aid for higher education in 1988-89 divided by the GDP figure of about $5,200 Billion in 1989. In making this comparison we are assuming that the ratio holds within any sub-period of the model period.

will not matter. Since $\pi_c > \pi_s$, another possibility is to search for a policy that gives a_c^*, a_s^*, such that the likelihood of success of the marginal enrollee is the same irrespective of parent type; that is $\pi_c(a_c^*) = \pi_s(a_s^*)$. In the benchmark economy $\pi_c(a_c^*)$ is 0.3 and $\pi_s(a_s^*)$ is 0.62. However, discussions on equality of opportunity are rarely framed in terms of likelihood of success, so we examine the policy that yields $a_c^* = a_s^*$.[26] The drop in the efficiency measure is likely to be more pronounced in this policy experiment when compared to the one where the probabilities of success of the marginal enrollees are equalized.

We can use the analogue of equations (8) through (12) for the case where $\pi_c \neq \pi_s$ to derive conditions that will have to be satisfied in a steady state with equality of opportunity. We use the following notation to make the resulting expressions more succinct: $u_c \equiv u((1-\tau)w_c(n_c^*))$, $u_{c-e} \equiv u((1-\tau)w_c(n_c^*) - e)$, $u_s \equiv u((1-\tau)w_s(n_c^*))$, and $u_{s-e} \equiv u((1-\tau)w_s(n_c^*) - e + s)$. The conditions required to have $a_c^* = a_s^* = a^*$ can then be derived to be:

$$
\begin{aligned}
\pi_s(a^*)(u_c - u_{c-e}) &= \pi_c(a^*)(u_s - u_{s-e}), \\
u_c - u_{c-e} &= \beta\pi_c(a^*)\frac{[F(a^*)(u_c - u_s) + (1 - F(a^*))(u_{c-e} - u_{s-e})]}{\left[1 - \beta\int_{a^*}^{1}(\pi_c(a) - \pi_s(a))\,dF(a)\right]}, \\
n_c^* &= n_c^*\int_{a^*}^{1}\pi_c(a)\,dF(a) + (1 - n_c^*)\int_{a^*}^{1}\pi_s(a)\,dF(a), \\
(1 - n_c^*)(1 - F(a^*))\,s &= \tau\left(n_c^*w_c(n_c^*) + (1 - n_c^*)w_s(n_c^*)\right).
\end{aligned}
$$

These are four equations in the four unknowns n_c^*, a^*, s, and τ, which we solve numerically.

As can be seen from Table 3, such a policy increases the measure of people with college attainment in the long run, relative to the benchmark case. College educated workers get taxed, but do not get any subsidies – the decrease in their net income as well as their anticipation that the college educated workforce is going to increase will cause them to send fewer of their children to college. The school educated parents are taxed but also get education subsidies, and as long as the net income effect is positive, they will send more of their children to college. The effect of school educated parents, who are larger in measure, sending their children to college swamps out the effect of college educated parents sending fewer children. A decrease in college premium (a measure of "inequality" in this setup) is an immediate consequence – a premium of 1.45 in the subsidized economy, compared to the 1.54 in the benchmark case.

While it is true that only the more able of the rich children now go to college, a larger fraction of lower ability poor children go to college. The increase in efficiency coming from the increased graduation rate of the rich children is more than offset by the decrease in efficiency caused by the decreasing graduation rate of the poor. In terms of (14), the effect of the drop in S_s dominates the increase in S_c. The efficiency goes down throughout the transition, since a_c^* and a_s^* monotonically adjust from the old steady state to the new one.

---

[26] In fact, the usual argument is that students with the same intrinsic ability should get the same opportunity irrespective of any early advantage that being born in a high income family brings (which is captured in our model, by π_c being greater than π_s).

Given the production function in which more college educated people is better, a higher level of output Y also obtains under this policy. The utility gain for the poor outweighs the utility loss for the rich – there are more of the former and they are located in the more curved part of the utility function. Therefore, the welfare under the new system is greater. All households would need to be given 1.1% more consumption every period when there are benchmark subsidies, to get the same welfare that obtains in the steady state with the equal opportunity subsidy. Once the tax burden during the transition to the equal opportunity subsidy level is taken into account, this figure decreases to 0.12%. About 72% of the college expenses need to be subsidized for the poor to get equality of opportunity. An income tax rate of about 3.6% needs to be levied on all workers in the long run to finance these subsidies, compared to the 1.9% rate that needs to be levied to meet the subsidy expenses in the benchmark. Close to 1.9% of output is given as subsidies.

Figures 13 - 16
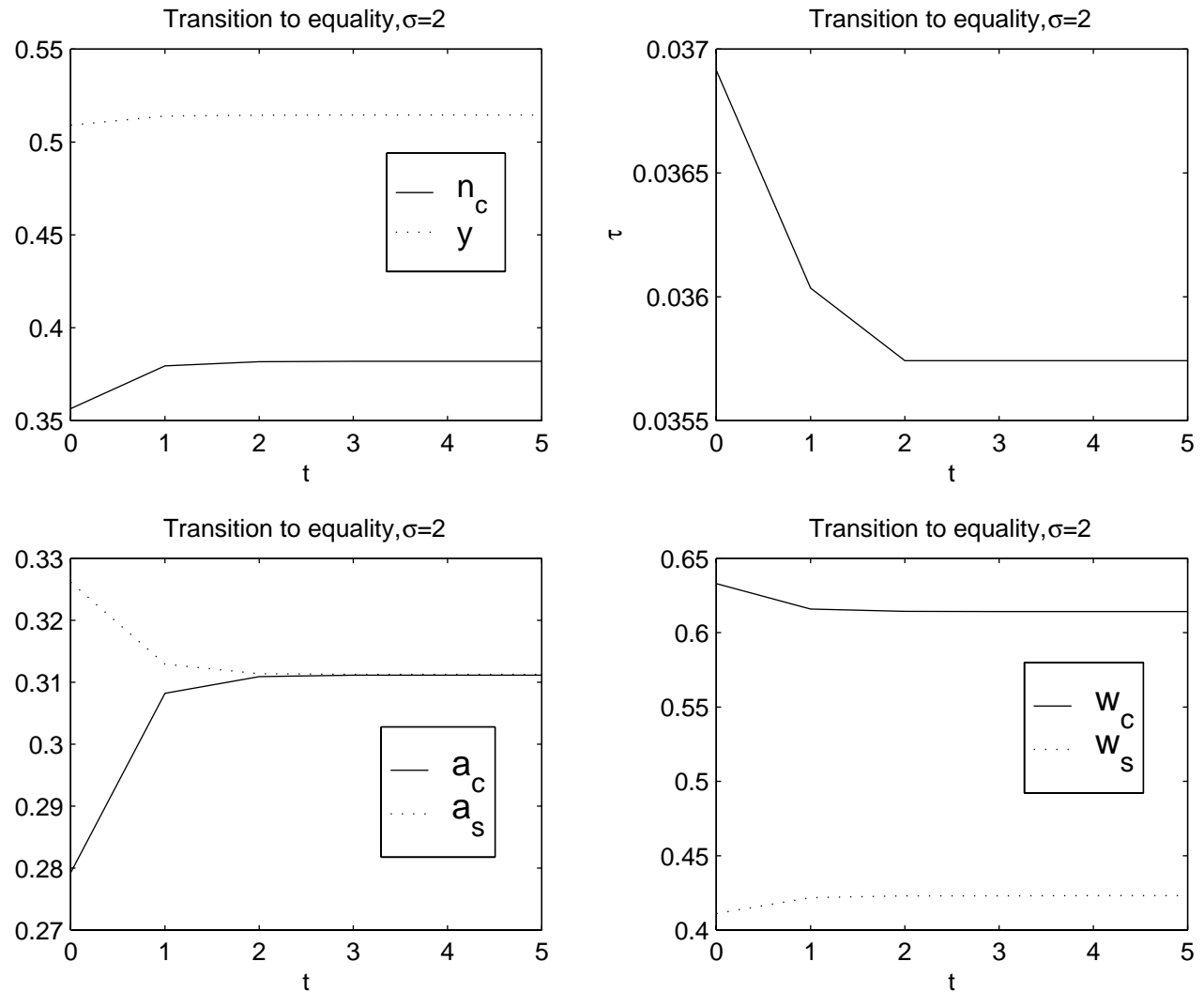
Figures 13 through 16 show the transition paths for n_c, Y, τ, a_c^*, a_s^*, w_c, and w_s in going from the benchmark steady state to equality of opportunity in the long run. As can be seen from these graphs all variables monotonically adjust to their new levels within a generation or two.

4.5.2 Maximization of the Skilled Labor Force

Politicians often call for policies that would enable as many students to attend college as possible. What would the effects of such a policy be? As mentioned earlier, equilibrium conditions demand that we focus on the region in which $w_c > w_s$. This condition will hold in any steady state as well. Steady state output therefore increases with the steady state fraction of college educated labor, n_c^*. If $(w_c - w_s) > e$ in the steady state, (the differences in marginal products is greater than the consumption cost of education), it can be seen that aggregate consumption would also increase with n_c^*.[27] Given that e is a very small fraction of lifetime wages, this is a fairly weak condition. As mentioned in the introduction to this section, aggregate steady state welfare could either increase or decrease with n_c^*. We study the problem:

$$\max_{\tau, s} n_c^*,$$

subject to the constraints:

$$n_c^* \int_{a_c^*(n_c^*, \tau, s)}^1 \pi_c(a) \, dF(a) + (1 - n_c^*) \int_{a_s^*(n_c^*, \tau, s)}^1 \pi_s(a) \, dF(a) = n_c^*,$$

$$\tau \left(n_c^* w_c(n_c^*) + (1 - n_c^*) w_s(n_c^*) \right) = (1 - n_c^*)(1 - F(a_s^*(n_c^*, \tau, s))) s,$$

where, the planner takes as given, the behavior of parents as characterized by the functions $a_c^*(n_c^*, \tau, s)$ and $a_s^*(n_c^*, \tau, s)$. The planner would really like to set policies such that $a_c^* = a_s^* = 0$ since this maximizes n_c^*. This is clearly infeasible when the only instrument available is an education tax and subsidy – the wages for both types would have to be simultaneously high to get both a^*s to be zero. So, as can be seen in Table 3, the planner heavily subsidizes the poor parents, who have a high response to subsidies. Since n_c^* is the highest it can be under the constraints mentioned (close to 39%), the college premium is the lowest, at 1.43. About 91% of the education expenses of the poor is subsidized, and income tax rate of 5.8% is required to pay for it; 5.8% of output has to be given out as subsidies in the steady state. Output is about 2% higher than in the benchmark.

The drop in education efficiency is high, corresponding to the steep drop in a_s^*. The dropout rate for the poor children goes from 49% in the benchmark to nearly 62%, and the effect of a decrease in the dropout rate of rich children is completely swamped out. While steady state welfare increases relative to the benchmark case, the increase is less than the one with a subsidy that guarantees equality of opportunity. The high tax rate and the decline in Λ coming from a high n_c^* are the likely reasons. When the transitional losses of the rich are factored in, welfare is lower than the benchmark value. Agents in the benchmark would be willing to pay up to 0.7% of their steady state benchmark consumption to avoid the new policy.

This experiment highlights the difficulty with proposals of the sort quoted in the footnote in the introduction. Increased subsidies do help more "enter the doors of college," but by decreasing the

---

[27]Aggregate consumption is given by $Y - [n_c^* (1 - F(a_c^*)) + (1 - n_c^*)(1 - F(a_s^*))] e$.

marginal ability of entrants causes a decline in the efficiency of the education sector.[28] The high tax rates necessary to support the subsidies can also decrease aggregate welfare. These losses are likely to be even higher if ability also matters for production, and labor supply decisions are endogenous.

4.5.3 Elimination of Subsidies

For sake of completeness, we also present the results when subsidies are eliminated completely. Attainment drops by 10 percentage points in the long run, and the "enrollment gap" is 70%. As expected, the premium is higher, at 2.02 and the education efficiency is higher. It is welfare decreasing to eliminate subsidies. Agents in the benchmark would be willing to pay up to 2.6% of their steady state benchmark consumption to avoid elimination of subsidies.

In the range of zero subsidies to about the 80% that is needed to get enrollment equalized, whether equity increases or decreases the efficiency depends on the measure considered – the positive measure decreases with the subsidy level, while the normative measure increases with the subsidy level. However, for the high level of subsidy required to maximize attainment, both measures decrease.

4.5.4 Ability-based Subsidies

Thus far we have assumed that information on ability a is unobserved by the government, and therefore all poor students get the same amount of subsidy. As mentioned above, by reducing the effective cost of college attendance for all poor students including the less able ones, such a scheme reduces the efficiency of the education sector. One interpretation of the private information assumption is that college aid is need based, rather than merit based. In the US, need-based aid constitutes an overwhelming fraction of total aid.[29] It is nevertheless interesting to examine a scheme where the government is able to observe a and base the subsidy on it.

When the aid is just need based, the decision of whether or not to send the child to college depends on the child's ability only through the probability of completion. When the aid is merit based, ability affects this decision by also altering the amount of subsidy the child receives. The problem of a college educated parent is not directly affected; however, a_s^* for the school educated parent is given implicitly now by:

$$\beta \pi_s \left(a_s^*(n_c) \right) \Lambda(\Phi(n_c)) \leqslant u(w_s(n_c)) - u(w_s(n_c) - e + s(a_s^*(n_c))),$$

where $s\left(\cdot\right)$ is an increasing function, and the expression holds with equality when a_s^* is interior. For simplicity, we assume the linear subsidy function $s(a) = sa$. We know that a solution exists because the left hand side of the equation is increasing in a, and the right hand side is decreasing in a. The Bellman equation changes in a straightforward way. The merit based subsidy scheme does not affect

---

[28]A similar point is made in Cameron and Heckman (1998).

[29]Federal aid is currently almost entirely need based. Even if we assume that all state and institutional aid is merit based, the ratio of need based aid to total aid was about 75% in 1988-89. (See Table 2-4 in McPherson and Schapiro (1991).)

Φ, except through it's effect on a_s^*. The government budget constraint (13) changes – the amount the government spends depends on the ability of the child.

$$(1 - n_c) \int_{a_s^*}^1 sa \ dF(a) = \tau(n_c w_c(n_c) + (1 - n_c)w_s(n_c)).$$

The expression for aggregate welfare would change appropriately. The following tables illustrate the effect of ability based subsidies.

Table 4

Flat vs ability based subsidy

Policy	n_c^*	a_c^*	a_s^*	$1 - S_c$	$1 - S_s$	w_c^*/w_s^*	Y	τ	sub	eff	ω	ω^{ss}
$s = 0.03$ flat	.357	.200	.498	.325	.492	1.54	.509	.019	.0097	9.637	1.000	1.000
$s = .03 * a$	.312	.230	.608	.312	.458	1.72	.497	.013	.0066	10.179	0.990	0.967
$s = .042 * a$	.324	.272	.560	.292	.473	1.67	.501	.019	.0097	10.113	0.993	0.978

The first row is the benchmark case, repeated here for convenience. The second row considers the case where the maximum subsidy given is the same as the flat subsidy given, 0.03. The role of ability based subsidies in raising the quality of poor college entrants is clear. Education efficiency improves relative to the flat subsidy scheme. However, the flat subsidy case retains a slight edge in terms of welfare, because aggregate welfare is increasing with n_c^* in that region. The transitional welfare is greater than the steady state one (but still lower than the one in the benchmark case), since the tax rate is lower in the new scheme and considering only the steady state would not include these transitional gains. In the aggregate, households would be willing to pay up to 1% of their steady state benchmark consumption to avoid ability based subsidies. The skill premium increases to 1.72.

An alternate comparison between the two schemes is when the steady state aggregate *amount* of subsidization, *sub*, is equalized (a "revenue neutral" experiment). The third row considers the case where this amount 0.0097 in both subsidy schemes. The above-mentioned effects are seen here as well – the ability based scheme provides better education efficiency, while the flat subsidy scheme provides higher steady state welfare. Once transition is taken into account, welfare is a bit lower than the benchmark one.

The results of this experiment indicate that if the government has the political will to use any available signal on ability, it can increase educational efficiency with little change in welfare by targeting aid to the more able students.

4.5.5 Productive Efficiency

The production function (3) does not take into account ability in production. It seems unrealistic that a person who has a higher probability of completing college *ex ante*, does not also have a higher *ex post* probability of completing a task at work successfully. In order to study the robustness of the

above results, as well as to gain independent insight, we consider the case with productive efficiency. We make the "brain" term depend on $\overline{a}$, the average ability of college educated workers. We therefore consider the production function:

$$Y = A\left[\theta g\left(\overline{a}\right)\left(N_c + \gamma N_s\right)^{\upsilon} + (1-\theta)(N_s + \varepsilon N_c)^{\nu}\right]^{\frac{1}{\nu}},$$

where g is an increasing function. In particular, we use $g\left(\overline{a}\right) = \kappa\overline{a}$, where κ is a positive constant. For simplicity, the same a matters not only for college completion but also for work ability.[30] There is also some theoretical precedence to this approach. Lucas (1988) uses a specification where the average level of human capital positively affects production. The production structure in Kremer (1993), where an assortative matching of workers with the same human capital occurs, is similar in flavor to the Lucas (1988) specification.

By making the effect purely external, we do not add an extra state variable in the *individual's* problem. However, there will be an extra *aggregate* state variable, $\overline{a}$. The functions w_c, a_c, etc. are now dependent on this second aggregate state variable, though the analysis of the individual's problem is mostly unchanged. The law of motion for $\overline{a}$, is given by:

$$\overline{a}'\left(n_c\right) = \frac{n_c\int_{a_c^*(n_c,\overline{a})}^{1} a\pi_c(a)\,dF(a) + (1-n_c)\int_{a_s^*(n_c,\overline{a})}^{1} a\pi_s(a)\,dF(a)}{n_c\int_{a_c^*(n_c,\overline{a})}^{1} \pi_c(a)\,dF(a) + (1-n_c)\int_{a_s^*(n_c,\overline{a})}^{1} \pi_s(a)\,dF(a)}.$$

This formula shows the possibility that any policy that decreases a_s^* more than it increases a_c^*, can decrease the average ability and thus output.[31]

When κ is 1.3, and all other parameters are at their benchmark levels, we get outcomes that are not too different from the earlier benchmark outcomes. We therefore use that value of κ, and $s = 0.03$ as our benchmark for this specification.

Table 5

Ability in the production function (s=0.03 is benchmark)

s	n_c^*	$\bar{a}$	a_c^*	a_s^*	$1-S_c$	$1-S_s$	w_c^*	w_s^*	w_c^*/w_s^*	Y	τ	eff	ω	ω^{ss}
0.00	.237	.748	.052	.775	.397	.407	.750	.365	2.05	.456	.000	9.974	.976	.933
0.02	.309	.734	.144	.638	.352	.449	.646	.390	1.66	.469	.011	10.017	.997	.986
0.03	.337	.730	.225	.544	.314	.478	.615	.400	1.54	.472	.019	9.969	1.000	1.000
0.04	.357	.725	.331	.409	.266	.520	.593	.407	1.46	.473	.032	9.635	.996	1.002
0.05	.364	.716	.457	.224	.213	.579	.574	.405	1.42	.466	.053	8.767	.979	.980

---

[30] The externality assumption is purely for convenience and not to rationalize subsidization. As will be evident from the results, subsidies in this case have the potential to reduce not only education efficiency, but also output and welfare.

[31] In a footnote, Autor, Katz, and Krueger (1998) note: "The large increases in educational attainment of the U.S. workforce since 1940 may overstate increases in the relative supply of "more-skilled" workers to the extent that the "unobserved" quality of more-educated workers declines with the "relabeling" of "lower productivity" workers into higher education categories." The decrease in average ability in our model is closely related to the point they make.

We can see that the average ability of college educated people in the workforce decreases with the level of subsidy. The educational efficiency initially increases with the subsidy level due to the decrease in dropout rates for the rich, but the increase in dropout for the poor eventually decreases education efficiency. As before, as the subsidy level increases, steady state welfare improves for a while but with very high levels of subsidy, the tax rate is high enough to decrease welfare. When the transition is taken into account, welfare is never higher than the one in the benchmark.

A major difference relative to the figures in Table 2 is that there is a *decrease* in per capita output when subsidies are increased beyond the 80% level. An increase in college educated people is not necessarily a process that increases per capita output if the new graduates have lower ability than before. At these high levels of subsidies, *both* w_c^* and w_s^* fall, even though n_c^* increases, since aggregate output falls. The skill premium continues to decrease with subsidies.

In summary, the earlier conclusions that any increase in subsidies relative to the benchmark will decrease education efficiency with no appreciable increase in welfare is robust to the production specification with ability. In fact, as conjectured earlier, any steady state welfare increases are lower and transitional welfare always decreases.

4.5.6   Reviving a "Trapped" Economy

As we saw in the previous section, even for empirically plausible values of σ, the economy can be trapped in a low output equilibrium, where the fraction of college educated people tends to zero. As discussed there, the behavior of the poor parents is responsible for the occurrence of the trap. If they form a large fraction of the population, their wages are low and they tend to send none of their children to college. The number of workers who are only schooled further increases in the next period, their wages further decreases, and the above behavior gets reinforced. It seems natural to ask if an education subsidy given to the poor parents can induce them to start sending their children to college and break this dynamically reinforcing behavior.

For this experiment we depart from the strategy of comparing the new policy to the US benchmark for obvious reasons. Instead, we consider the $\sigma = 4$ case, where without a subsidy, a trap is the only possibility. Figure 17 shows how the transition function Φ changes with subsidy. We can see how a subsidy level that is sufficient can get an economy out of a trap equilibrium and move it to one where a high steady state obtains. For instance, a 17% or a 33% subsidization ($s = 0.01$ or 0.02) is not enough to revive the trapped economy.[32] But with a 50% subsidization ($s = 0.03$), the economy emerges from a trap toward a higher steady state. These are the subsidies that are *offered*, rather than *spent*; in the first two cases, the equilibrium expenditure on subsidies is zero given that no poor parent takes up the offer.

---

[32] A skill premium of about 10 does not seem unreasonable for the poorest countries.

Table 6

Reviving a trapped economy when $\sigma = 4$

Policy (s)	n_c^*	a_c^*	a_s^*	$1 - S_c$	$1 - S_s$	w_c^*/w_s^*	Y	τ	sub	eff	ω	ω^{ss}
0.00, 0.01, 0.02	.000	.000	1.000	.425	N/A	10.22	.264	.000	0	N/A	1.000	1.000
0.03	.309	.047	.657	.400	.443	1.73	.497	.014	0.007	9.677	1.037	1.611
0.04	.369	.137	.486	.355	.496	1.50	.512	.025	0.013	9.564	1.090	1.705

Figure 17

$\Phi(n_c)$ for s=.01,.02,.03, and .04, when σ=4
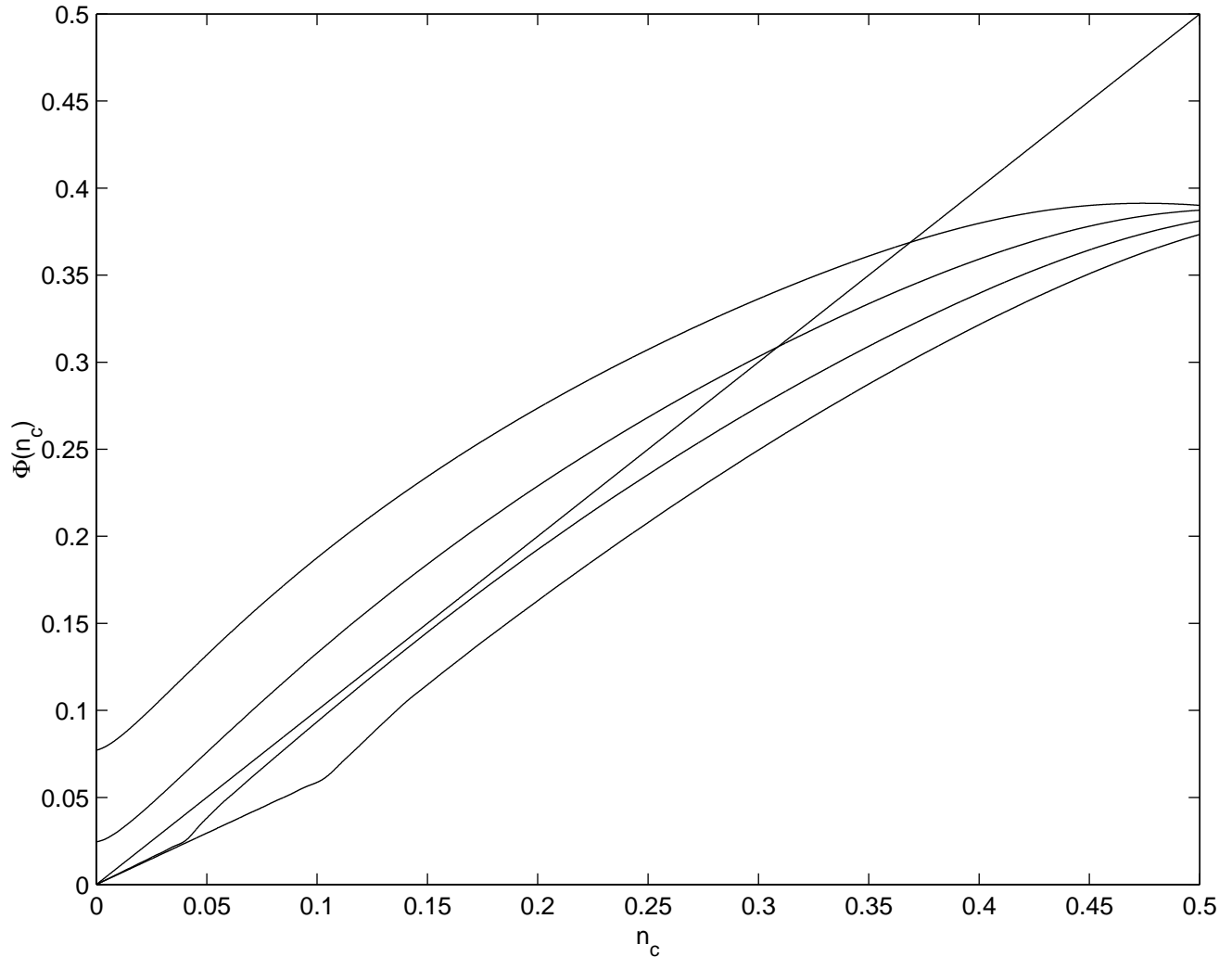
Table 6 highlights the various subsidy schemes when compared to the case with no subsidy, which corresponds to a trap for the economy. For $s = 0.03$, which is sufficient to get the economy out of the trap, the amount of subsidy given in the steady state is just 0.7% of the output. From the last two columns in this table, we can see that comparing steady state welfare alone (ω^{ss}) can be misleading

– once the transitional effect of decreasing welfare due to taxation is taken into account, the overall welfare gain (ω) is lower. In spite of this, the welfare gain for a trapped economy is substantial – when $s = 0.03$, the compensating consumption is 3.7% of the benchmark value, and when $s = 0.04$ it is 9%. The differences in the speed of transition is responsible for the differences in welfare between the two cases. Figure 18 shows the transition path to the higher steady state under the two subsidy levels. The transition is faster when $s = 0.04$, allowing welfare gains to be realized in a "timely" manner.

Figure 18

Welfare over the transition when s=.03 and s=.04, σ=4
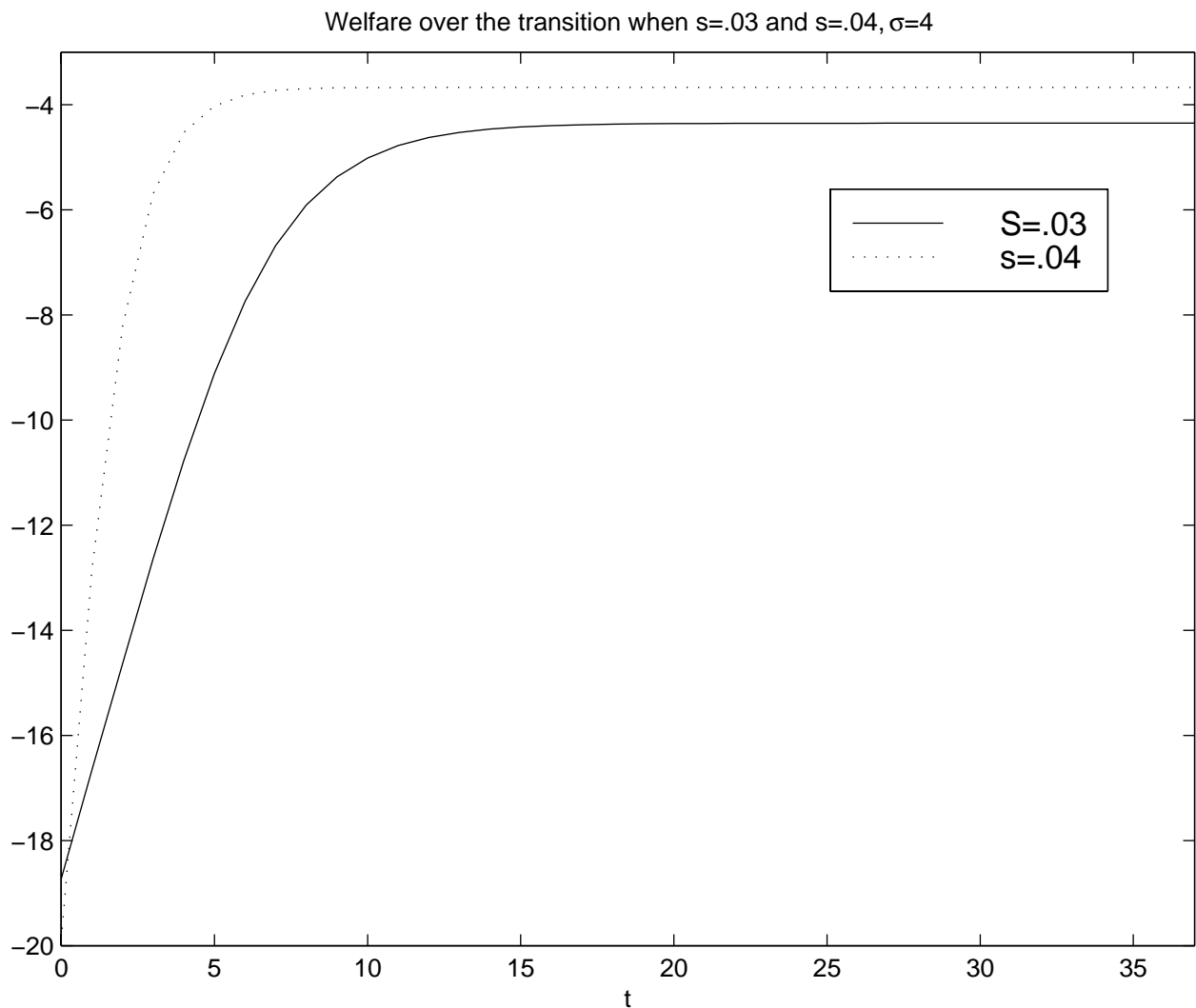
Figures 19 through 22 plot the transition of an economy when a system with $s = 0.04$ (a 67% subsidy) is introduced in a trapped economy. We can see that n_c and Y increase monotonically to a "developed" steady state. They are accompanied by a drop in a_s^* and w_c and a rise in a_c^* and w_s. The adjustment takes a few generations, but a precipitous drop in w_c is immediate. The educated elite in the developing countries, who often occupy key policy making positions, are blamed for paying undue attention to higher education – the supposition is that they benefit more from government involvement

in higher education than do the poor. This model illustrates a countervailing incentive that such an elite might have to *not* subsidize higher education if the poor also benefit, and preserve the monopoly they enjoy for their children who are more likely to be college educated. If this incentive effectively causes subsidization to be blocked, the economy will remain in a trap.

Figures 19 - 22
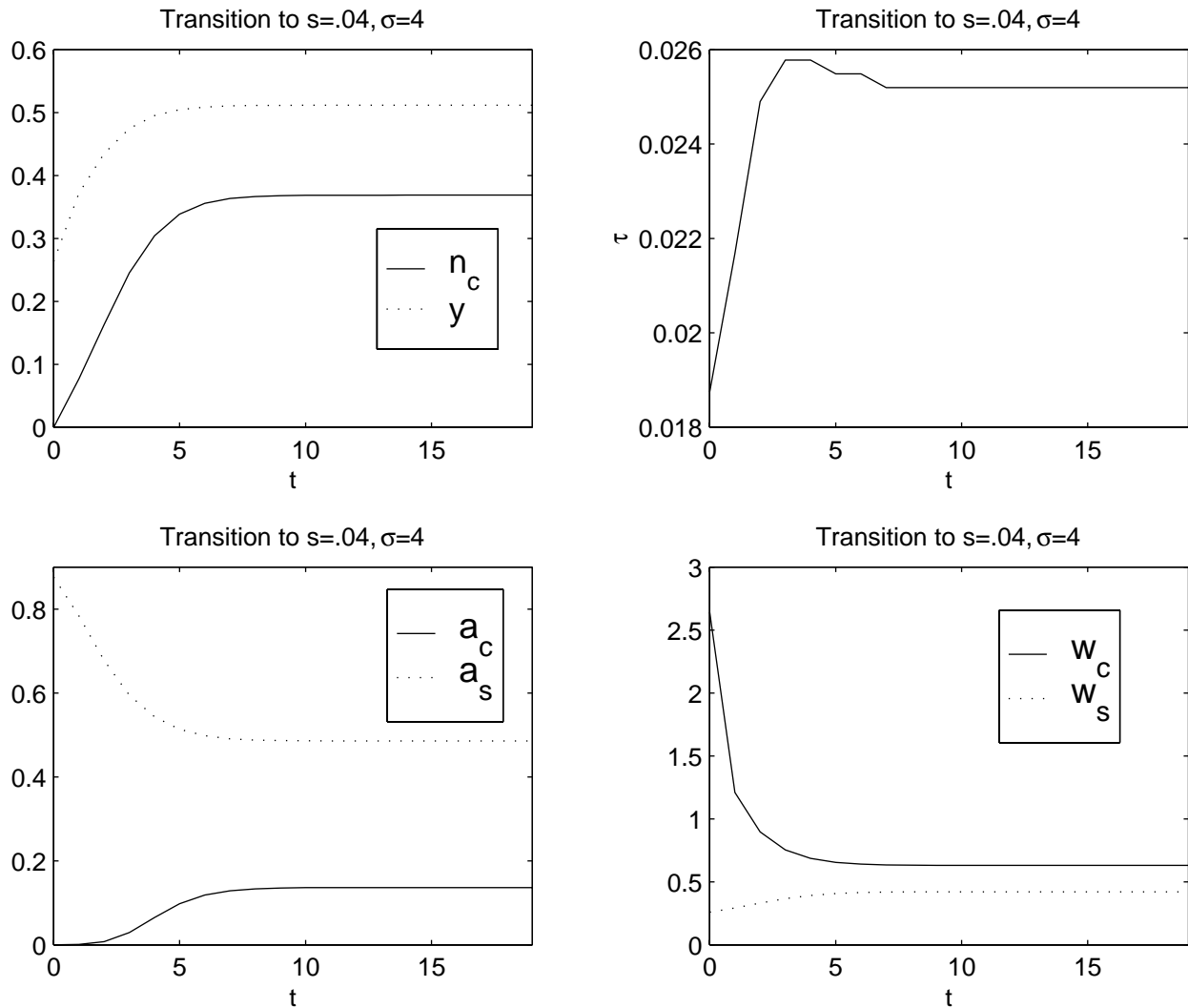
5   Conclusions

In this paper we have developed a dynamic framework to analyze higher education subsidies when there is heterogeneity in parental income and student ability. The model is calibrated to the US economy to provide us with a sensible benchmark for comparing outcomes from different policies. The government can design a tax and subsidy scheme that guarantees equality of opportunity, but only at the expense of a decrease in the efficiency of utilization of education resources. The welfare gain that comes from increasing the fraction of people with higher marginal productivity is minimal. A policy that aims

to maximize the fraction of college-educated labor by sending as many children as possible to college results in a big drop in the above-mentioned efficiency with little or no welfare gain. If the government has the political will to use any available signal on ability and provide merit-based aid, it can increase this efficiency with little decrease in welfare. These conclusions appear robust to the addition of ability in the production function. This analysis indicates that any further increases in higher education subsidies to bridge the "enrollment gap" in the US may not be warranted.

However, education subsidies may be a potent tool for countries that are caught in a development trap. A sufficient level of subsidy can cause the economy to emerge from the trap and also allow the gains to be realized in a timely manner. It would be useful to calibrate the model to a poor economy which has been stagnant for several years and study this issue in greater detail.

Since poor parents are situated in the more curved part of the utility function and are more responsive to flat subsidies, as well as for simplicity, we focus on a simple policy where everyone is taxed at the same rate and only poor children get subsidies. It would be useful to see how robust our conclusions are to alternate policy specifications.

The case for subsidies is clearly overstated if true college completion ability is *not* i.i.d. This could be implemented by making the distribution function F a Markov distribution. Having a continuum of abilities with persistence across generations will add to the technical complexity, but might yield interesting insights. The assumption we have made for our numerical experiments, that children of college educated parents systematically outperform children of school educated parents $(\pi_c > \pi_s)$, is only one step in this direction. A high ability person who due to bad luck fails to graduate will not be able to transmit any advantage in our setup, but will be able to in one where there is true persistence.

We do not have physical capital in this framework. It will also be useful to study the robustness of our results when an alternate form of bequest is available to parents. Low ability children, especially the rich ones, can be left with capital instead of being sent to college. It will also be interesting to model state run universities, where subsidization is indirectly available to *all* students who are admitted instead of just the poor.

As mentioned in the introduction, we have abstracted from economic growth. One of the most common reasons mentioned for subsidizing higher education is to promote growth; if there are growth externalities, subsidies might be viewed in a more favorable light than they are in our model. In future work, we plan to incorporate growth in our framework, both with and without externalities, and conduct the experiments similar to the ones we have conducted here. The growth can result from neutral technical change or a skill biased one; modeling the latter type of change would allow us to analyze policies that have been suggested to alleviate increasing inequality in the US in recent times.

A   Appendix

A.1   Lemma 1: $EV_c\left(n_c\right) > EV_s\left(n_c\right)$

Suppose not; i.e. $EV_c\left(n'_c\right) \leq EV_s\left(n'_c\right)$, for any given n'_c that parents take as given. Given $0 \leq \pi(a) \leq 1$, we have:

$$\beta\pi(a)EV_c\left(n'_c\right) + \beta(1-\pi(a))EV_s\left(n'_c\right) \leq \beta EV_s\left(n'_c\right),$$

$\forall a$. Given $e > 0$, it follows that for any given state n_c:

$$u(w_s\left(n_c\right) - e) + \beta\pi(a)EV_c\left(n'_c\right) + \beta(1-\pi(a))EV_s\left(n'_c\right) < u(w_s\left(n_c\right)) + \beta EV_s\left(n'_c\right).$$

So, no school-educated parent sends her child to college. Likewise one can show that no college-educated parent sends her child to college. Therefore, $C, S = \phi$. Since a college educated parent does not send her child to college,

$$V_c(a; n_c) = u(w_c\left(n_c\right)) + \beta EV_s\left(n'_c\right), \ \forall a.$$

Since ability a does not appear on the right side of this expression, it follows that $EV_c\left(n_c\right) = u(w_c\left(n_c\right)) + \beta EV_s\left(n'_c\right)$. Likewise, one can show that $EV_s\left(n_c\right) = u(w_s\left(n_c\right)) + \beta EV_s\left(n'_c\right)$. Since Assumption 3 implies $w_c\left(n_c\right) > w_s\left(n_c\right)$, it follows that $EV_c\left(n_c\right) > EV_s\left(n_c\right)$. This is true for any $n_c \epsilon \left[0,1\right]$. In particular, it is true for n'_c, which contradicts the hypothesis that $EV_c\left(n'_c\right) \leq EV_s\left(n'_c\right)$. $\bullet$

A.2   Lemma 2: There exist unique $a_c^*\left(n_c\right), a_s^*\left(n_c\right) \in (0,1]$

Consider any $n_c, n'_c \epsilon \left[0,1\right]$. The school portion of the Bellman equation, $u(w_i\left(n_c\right)) + \beta EV_s\left(n'_c\right)$ is independent of a. Therefore, when plotted against a it is a horizontal line. The college portion of the Bellman equation (1) is $u(w_i\left(n_c\right) - e) + \beta EV_s\left(n'_c\right) + \beta\pi(a)(EV_c\left(n'_c\right) - EV_s\left(n'_c\right))$. When evaluated at 0, it is $u(w_i\left(n_c\right) - e) + \beta EV_s\left(n'_c\right)$, (assuming $\pi(0) = 0$), which is less than the school portion. The college portion is strictly increasing in a, assuming $\pi'(a) > 0$ and given the above result that $EV_c\left(n'_c\right) > EV_s\left(n'_c\right)$. These two curves either intersect at a unique $a_i^* > 0$, or do not intersect at all, in which case $a_i^* = 1$.

These imply that there exist threshold ("reservation") abilities $a_i^* \in (0,1]$, $i = s, c$, such that a parent of type i finds it profitable to send her child to college only if the child's ability is $\geqslant a_i^*$. $\bullet$

A.3   Lemma 3: $a_c^*\left(n_c\right) < a_s^*\left(n_c\right)$

Write the Euler-like equations (4) for the marginal children of both parent types, for a given combination of aggregate state n_c and anticipated next-period state n'_c:

$$\beta\pi(a_s^*\left(n_c\right))\Lambda\left(n'_c\right) \ \leqslant \ g_s\left(n_c\right),$$
$$\beta\pi(a_c^*\left(n_c\right))\Lambda\left(n'_c\right) \ \leqslant \ g_c\left(n_c\right),$$

with equality if the relevant threshold ability is < 1. We have used the notation introduced in the main text, where the right hand side is the utility cost of college education, and the left hand side is the discounted expected benefit of sending the child to college, weighted by the probability of college completion.

The key to proving this lemma is the fact that $g_c(n_c) < g_s(n_c)$, which follows from the strict concavity of the utility function and $w_c(n_c) > w_s(n_c)$. That is college-educated parents always have a lower utility cost of sending their children to college. We first rule out the possibility that $a_c^* = 1$, and $a_s^* < 1$. Suppose this were true, we have $\beta\pi(1)\Lambda\left(n_c'\right) \leqslant g_c(n_c)$ and $\beta\pi(a_s^*(n_c))\Lambda\left(n_c'\right) = g_s(n_c)$. These imply the following inequality:

$$\frac{\pi(1)}{\pi(a_s^*(n_c))} < \frac{g_c(n_c)}{g_s(n_c)}.$$

Given that $\pi(\cdot)$ is strictly increasing and the above fact on relative utility costs, this cannot be true – the left side is > 1 while the right side is < 1. Since we are considering only equilibria in which some college education takes place, we do not analyze the case $a_c^*(n_c) = a_s^*(n_c) = 1$. (At this stage we are analyzing the parents' problem and we have to explicitly rule out an equilibrium of this type. Once we consider the whole equilibrium, it will become obvious that such a situation cannot arise. Only $n_c' = 0$ is consistent with this situation, and the anticipated gain to college education will be the highest in this case, which will induce some college educated parents to send their children to college.)

Now consider the case where a_c^*, a_s^* are interior. From the Euler-like equations, we get:

$$\frac{\pi(a_c^*(n_c))}{\pi(a_s^*(n_c))} = \frac{g_c(n_c)}{g_s(n_c)} < 1.$$

Given that $\pi(\cdot)$ is strictly increasing, it follows that $a_c^* < a_s^*$.

Since the maximum a_s^* can be is 1, it is immediately apparent that $a_c^* < 1$. •

A.4 Derivation of equation (9)

Simplify the equation system (7) by first integrating out terms that do not depend on a to get:

$$
\begin{aligned}
EV_c(n_c) &= F(a_c^*(n_c))\left[u(w_c(n_c)) + \beta EV_s\left(n_c'\right)\right] \\
&\quad + (1 - F(a_c^*(n_c)))\left[u(w_c(n_c)) - e) + \beta EV_s\left(n_c'\right)\right] \\
&\quad + \beta\left[EV_c\left(n_c'\right) - EV_s\left(n_c'\right)\right]\int_{a_c^*(n_c)}^{1}\pi(a)dF(a) \\
EV_s(n_c) &= F(a_s^*(n_c))\left[u(w_s(n_c)) + \beta EV_s\left(n_c'\right)\right] \\
&\quad + (1 - F(a_s^*(n_c)))\left[u(w_s(n_c)) - e) + \beta EV_s\left(n_c'\right)\right] \\
&\quad + \beta\left[EV_c\left(n_c'\right) - EV_s\left(n_c'\right)\right]\int_{a_s^*(n_c)}^{1}\pi(a)dF(a).
\end{aligned}
$$

Subtract one from the other to get:

$$(EV_c - EV_s)(n_c) =$$

$$F(a_c^*(n_c))\, u(w_c(n_c)) + (1 - F(a_c^*(n_c)))\, u(w_c(n_c) - e) -$$

$$[F(a_s^*(n_c))\, u(w_s(n_c)) + (1 - F(a_s^*(n_c)))\, u(w_s(n_c) - e)] +$$

$$\beta\left[(EV_c - EV_s)\left(n_c^{'}\right)\right]\left[\int_{a_c^*(n_c)}^{1}\pi(a)dF(a) - \int_{a_s^*(n_c)}^{1}\pi(a)dF(a)\right].$$

Using the notational definitions for x and Λ defined in the main text, and the transition function Φ, we get equation (9). $\bullet$

A.5   Theorem 1 (Existence and Uniqueness)

We first write equations (8) through (12) for the specialization $\pi(a) = a$, $F(a) = a$. We define the ratio of utility costs, $J(n_c) \equiv \frac{g_s(n_c)}{g_c(n_c)}$. Given the concavity of the utility function and the fact $w_c^{'}(n_c) < 0$, and $w_s^{'}(n_c) > 0$, we get $J(\cdot)$ to be strictly decreasing and bounded below by 1. It also follows that $g_c(n_c) < g_s(n_c)$, $g_c^{'}(n_c) > 0$, and $g_s^{'}(n_c) > 0$. We will use these facts repeatedly in the proof. The relevant equations then are (note that we have dropped the "$*$" from the notation for threshold abilities for the sake of convenience):

$$\beta a_c(n_c)\Lambda\left(n_c^{'}\right) = g_c(n_c) \tag{15}$$

$$a_s(n_c) = J(n_c)a_c(n_c) \tag{16}$$

$$n_c^{'} = \frac{1}{2} - \frac{1}{2}\left[n_c a_c(n_c)^2 + (1 - n_c)a_s(n_c)^2\right] \tag{17}$$

$$\Lambda(n_c) = x(n_c) + \frac{\beta}{2}\left[a_s(n_c)^2 - a_c(n_c)^2\right]\Lambda(n_c^{'}) \tag{18}$$

$$x(n_c) \equiv [a_c(n_c)u(w_c(n_c)) + (1 - a_c(n_c))u(w_c(n_c) - e)] - \tag{19}$$
$$[a_s(n_c)u(w_s(n_c)) + (1 - a_s(n_c))u(w_s(n_c) - e)].$$

where $n_c^{'} \equiv \Phi(n_c)$. We will derive a set of sufficient conditions that will guarantee existence and uniqueness as we proceed through the various steps of the proof, and also provide an example of a set of parameters that satisfy the sufficient conditions. The sufficient conditions will be stringent given our strategy of using the contraction mapping argument on a single function, Λ, in a system of four functional equations. However, any other strategy, for instance using the contraction argument on a pair of functions, would make the proof much more complicated. The assumptions made are on the utility-wage composite functions g_s, g_c, J, and hence and are likely to be satisfied for a wide variety of individual utility and production functions. These assumptions are, therefore, not too restrictive. The main thrust of these assumptions are to make the strength of the income effect mentioned in the main text, as captured by $\frac{J^{'}}{J}$, strong enough. Equivalently, the assumptions set a bound on the elasticity of a_c, which is an indication of the strength of the substitution effect.

Step 1: Consider the mapping T defined by:

$$T\Lambda(n_c) = x(n_c) + \frac{\beta}{2}\left[a_s(n_c)^2 - a_c(n_c)^2\right]\Lambda(\Phi(n_c)), \tag{20}$$

in conjunction with (15), (16), (17), and (19). Let $C[0,1]$ be the space of bounded, continuous, non increasing functions $\Lambda : [0,1] \to R$, with the sup norm. T maps $C[0,1]$ into itself, $T : C[0,1] \to C[0,1]$.

The strategy will be to start with a weakly decreasing Λ, and find sufficient conditions on the utility and production function composite for each of the three components of the right hand side of the mapping to be decreasing in n_c.

Sub-step (i): $a_c(n_c)$ is increasing. That is, if $n_{c1} < n_{c2}$, $a_c(n_{c1}) < a_c(n_{c2})$.

We can use (15) and (17) to write:

$$a_c(n_c) = \frac{g_c(n_c)}{\beta\Lambda\left\{\frac{1}{2} - \frac{1}{2}\left[n_c + (1-n_c)J(n_c)^2\right]a_c(n_c)^2\right\}}.$$

For a given n_c, we can plot the two sides of this equation as functions of a_c. The left side is the 45° line. The right side is a weakly decreasing function, given the hypothesis of a weakly decreasing Λ, that starts at a positive number. Therefore a unique $a_c^*(n_c)$ exists. Suppose n_c increases, say from n_{c1} to n_{c2}. The left side is unchanged. The quantity $\left[n_c + (1-n_c)J(n_c)^2\right]$ is a convex combination of 1 and $J(n_c)^2 > 1$; $J(n_c)^2$ is decreasing in n_c; moreover, a greater weight is placed on the smaller quantity, 1, as n_c increases. Therefore, the entire term within the square brackets is decreasing in n_c. Using this fact, and the hypothesis on Λ, we can see that the right side shifts out. Even with a weakly decreasing Λ, we therefore get $a_c(n_{c1}) < a_c(n_{c2})$. $\bullet$

Sub-step (ii): Suppose $\left[n_c g_c^2(n_c) + (1-n_c)g_s^2(n_c)\right]$ is strictly decreasing in n_c. Then $\Phi(n_c)$ (that is, n_c') is increasing. That is, if $n_{c1} < n_{c2}$, $n_{c1}' < n_{c2}'$.

Suppose not. That is, $n_{c2}' < n_{c1}'$. The hypothesis on Λ implies $\Lambda\left(n_{c2}'\right) \geqslant \Lambda\left(n_{c1}'\right)$. It also follows from (17) that $\left[n_{c1} + (1-n_{c1})J_1^2\right]a_{c1}^2 < \left[n_{c2} + (1-n_{c2})J_2^2\right]a_{c2}^2$. We will use subscripts 1 and 2 to denote the values of a function evaluated at n_{c1} and n_{c2}. Use (15) to then get:

$$
\begin{aligned}
\left[n_{c1} + (1-n_{c1})J_1^2\right]g_{c1}^2 &\leqslant \frac{\Lambda\left(n_{c1}'\right)}{\Lambda\left(n_{c2}'\right)}\left[n_{c2} + (1-n_{c2})J_2^2\right]g_{c2}^2 \\
&\leqslant \left[n_{c2} + (1-n_{c2})J_2^2\right]g_{c2}^2,
\end{aligned}
$$

which implies $\left[n_{c1}g_{c1}^2 + (1-n_{c1})g_{s1}^2\right] \leqslant \left[n_{c2}g_{c2}^2 + (1-n_{c2})g_{s2}^2\right]$, contradicting the assumption that this quantity is strictly decreasing. This quantity is a sum of the square of the utility costs weighted by the fraction of people that constitute each type– the first term is increasing and the second decreasing. The assumption guarantees that the decrease in the weighted square utility costs for the poor parents dominates the increase for the rich, a condition satisfied for any reasonable specification of the production and utility functions. $\bullet$

Sub-step (iii): Suppose $\left[n_c g_c^2(n_c) + (1-n_c)g_s^2(n_c)\right] / \left(g_s^2(n_c) - g_c^2(n_c)\right)$ is strictly increasing in n_c. Then $\left(a_s^2(n_c) - a_c^2(n_c)\right)$ is strictly decreasing.

Use (16) to write $a_s^2 - a_c^2 = \left(J^2 - 1\right)a_c^2$. Therefore, we need to show for $n_{c1} < n_{c2}$:

$$\frac{a_{c1}^2}{a_{c2}^2} > \frac{J_2^2 - 1}{J_1^2 - 1}. \tag{21}$$

We know from the previous sub-step that:

$$\frac{a_{c1}^2}{a_{c2}^2} > \frac{\left[n_{c2} + (1 - n_{c2}) J_2^2\right]}{\left[n_{c1} + (1 - n_{c1}) J_1^2\right]}.$$

A sufficient condition for (21) to be true is:

$$\frac{\left[n_{c2} + (1 - n_{c2}) J_2^2\right]}{\left[n_{c1} + (1 - n_{c1}) J_1^2\right]} > \frac{J_2^2 - 1}{J_1^2 - 1},$$

that is, when $\left[n_c + (1 - n_c) J^2\right] / \left(J^2 - 1\right)$ is strictly increasing, which is what is assumed. The previous sub-step assumed that the numerator is decreasing. Since g_c is increasing, and g_s is decreasing, the above assumption holds when the difference in the square of utility costs is decreasing faster. It is easy to verify that this condition is satisfied when $-\frac{J'}{J} > \frac{1}{2} \left(\frac{J^2 - 1}{J}\right)^2$, and as mentioned earlier is an assumption on the strength of the income effect. The result that $\left(a_s^2 - a_c^2\right)$ is decreasing is weaker than claiming that a_s is decreasing, which we find in all our computations. (We know from sub-step (i) that a_c is increasing). However, it does capture the a flavor of the intuitive result that the difference in educational investment between the rich and the poor decreases with n_c. ●

Sub-step (iv): Suppose $u(w_c) - u(w_s) - J(g_s - g_c)$ is strictly decreasing in n_c. Then $x(n_c)$ is strictly decreasing.

Use (16) and the definition of x in (10) to see that $x(n_{c2}) < x(n_{c1})$, for $n_{c2} > n_{c1}$ if:

$$[u(w_{c1} - e) - u(w_{s1} - e)] - [u(w_{c2} - e) - u(w_{s2} - e)] >$$
$$a_{c1} \left(\frac{g_{s1}^2 - g_{c1}^2}{g_{c1}}\right) - a_{c2} \left(\frac{g_{s2}^2 - g_{c2}^2}{g_{c2}}\right).$$

Since, from sub-step (i), $a_{c2} > a_{c1}$, a sufficient condition can be obtained by seeing that the right hand side is bounded by $a_{c1} \left(\frac{g_{s1}^2 - g_{c1}^2}{g_{c1}} - \frac{g_{s2}^2 - g_{c2}^2}{g_{c2}}\right)$, which in turn is bounded by $\left(\frac{g_{s1}^2 - g_{c1}^2}{g_{c1}} - \frac{g_{s2}^2 - g_{c2}^2}{g_{c2}}\right)$. Therefore, we need:

$$[u(w_{c1} - e) - u(w_{s1} - e)] - \left(\frac{g_{s1}^2 - g_{c1}^2}{g_{c1}}\right) > [u(w_{c2} - e) - u(w_{s2} - e)] - \left(\frac{g_{s2}^2 - g_{c2}^2}{g_{c2}}\right).$$

A few algebraic steps will show that this is exactly the sufficient condition assumed. The assumption guarantees that the unweighted utility difference between the rich and the poor decrease fast enough with n_c, so that there are no perverse effects of a wage decrease for the rich that would cause them to send far fewer children to college and actually get a greater weighted utility x. In other words, this is an assumption on the upper bound on the increase in a_c. ●

Sub-step (v): If $n_{c1} < n_{c2}$, and Λ is weakly decreasing, $T\Lambda(n_{c2}) \leqslant T\Lambda(n_{c1})$. That is, T maps C into itself.

Consider the mapping T defined in (20). Sub-step (ii) provided conditions under which $\Phi(n_c)$ increases. Given the weakly decreasing Λ, we can see that $\Lambda(\Phi(n_c))$ is weakly decreasing. Sub-step (iii) provided conditions under which $\left(a_s^2 - a_c^2\right)$ decreases, and sub-step (iv) provided conditions under which $x(n_c)$ decreases.

Therefore every component in T decreases, two of them strictly. Therefore, T maps C into itself. (In fact, T maps C into the space of strictly decreasing continuous functions.) ●

An example of a set of parameters that satisfies the three sufficient conditions listed above is: $A = 1$, $\theta = 0.525$, $v = 0.5$, $\epsilon = 0.25, \gamma = 0.02$, $e = 0.05$, and $\sigma = 1$ (log utility).

Step 2: T satisfies the monotonicity property. That is, if $\Lambda_2(n_c) \geqslant \Lambda_1(n_c)$, $\forall\ n_c \epsilon [0,1]$, $T\Lambda_2(n_c) \geqslant T\Lambda_1(n_c)$.

Fix n_c. Use (15) and (17) to write:

$$a_c(n_c) = \frac{g_c(n_c)}{\beta\Lambda\left\{\frac{1}{2} - \frac{1}{2}\left[n_c + (1-n_c)J(n_c)^2\right]a_c(n_c)^2\right\}}$$

The left hand side is the 45 degree line when plotted against a_c and given the nonincreasing property of Λ, the right hand side is a decreasing function of a_c that starts at a value greater than zero. There is a unique intersection that yields $a_c(n_c)$. When the entire function Λ increases, the right hand curve shifts inward, causing $a_c(n_c)$ to decrease. Therefore, $a_{c,2}(n_c) < a_{c,1}(n_c)$, where the additional subscript now refers to the relevant function Λ (as opposed to step 1, where the subscripts referred to different n_c's for the same function). Use (15) and (18) to write:

$$\begin{aligned}(T\Lambda_2 - T\Lambda_1)(n_c) &= x_2(n_c) - x_1(n_c)\\ &+ \frac{\left(J(n_c)^2 - 1\right)}{2}g_c(n_c)(a_{c,2}(n_c) - a_{c,1}(n_c)).\end{aligned}$$

Using the definition of x in (19), $F(a) = a$, and (16) we can write:

$$x_2(n_c) - x_1(n_c) = (a_{c,2}(n_c) - a_{c,1}(n_c))(g_c(n_c) - J(n_c)g_s(n_c)).$$

Using this in the above equation, we get:

$$(T\Lambda_2 - T\Lambda_1)(n_c) = (a_{c,2}(n_c) - a_{c,1}(n_c))\left(\frac{\left(J(n_c)^2 - 1\right)}{2}g_c(n_c) - J(n_c)g_s(n_c)\right).$$

We showed that the first factor on the right side is negative. It is easy to show that the second factor is negative if $g_c(n_c)^2 < g_s(n_c)^2$, which is true for any n_c under the assumptions already made in the main text on the utility and production functions. Therefore, $(T\Lambda_2 - T\Lambda_1)(n_c) > 0$. That is, the mapping T satisfies the monotonicity property. ●

Step 3: T satisfies the discounting property. If we consider $(\Lambda + \delta)(n_c) \equiv \Lambda(n_c) + \delta, \forall n_c$, it follows that $T(\Lambda + \delta)(n_c) \leqslant T\Lambda(n_c) + \rho\delta$, for some $0 < \rho < 1$.

Again, fix n_c. Let the Λ_2 considered in step 2 now refer to $\Lambda + \delta$, and let Λ_1 refer to Λ. We saw in step 2 that $a_{c,2}(n_c) < a_{c,1}(n_c)$, but in this step we need to set a bound on this change. From this fact and (17),

we can deduce that $n'_{c,2} > n'_{c,1}$, which given the space $C[0,1]$ assumed, implies $\Lambda\left(n'_{c,2}\right) \leqslant \Lambda\left(n'_{c,1}\right)$, $\Rightarrow$ $-\frac{1}{\Lambda\left(n'_{c,2}\right)} \leqslant -\frac{1}{\Lambda\left(n'_{c,1}\right)}$.

Use this fact, and (15) for the two functions to get:

$$
\begin{aligned}
a_{c,1}\left(n_c\right) - a_{c,2}\left(n_c\right) &= \frac{g_c\left(n_c\right)}{\beta} \left\{ \frac{1}{\Lambda\left(n'_{c,1}\right)} - \frac{1}{\Lambda\left(n'_{c,2}\right) + \delta} \right\} \\
&\leqslant \frac{g_c\left(n_c\right)}{\beta} \left\{ \frac{1}{\Lambda\left(n'_{c,1}\right)} - \frac{1}{\Lambda\left(n'_{c,1}\right) + \delta} \right\} \\
&= \frac{g_c\left(n_c\right)}{\beta} \frac{\delta}{\Lambda\left(n'_{c,1}\right)\left[\Lambda\left(n'_{c,1}\right) + \delta\right]} \\
&= \frac{\delta\beta a_{c,1}\left(n_c\right)^2}{g_c\left(n_c\right) + \delta\beta a_{c,1}\left(n_c\right)}.
\end{aligned}
$$

Using the definition of $J\left(n_c\right)$, we can get:

$$
J\left(n_c\right) g_s\left(n_c\right) - \frac{\left(J\left(n_c\right)^2 - 1\right)}{2} = \frac{g_s\left(n_c\right)^2 - g_c\left(n_c\right)^2}{2g_c\left(n_c\right)}.
$$

Using the expression for $T\Lambda_2 - T\Lambda_1$ from the previous step, and the above two expressions, we can write:

$$
T\left(\Lambda + \delta\right)\left(n_c\right) - T\Lambda\left(n_c\right) \leqslant \delta\left(\frac{\beta a_{c,1}\left(n_c\right)^2}{g_c\left(n_c\right) + \delta\beta a_{c,1}\left(n_c\right)}\right)\left(\frac{g_s\left(n_c\right)^2 - g_c\left(n_c\right)^2}{2g_c\left(n_c\right)}\right).
$$

We would like the product of the last two terms to be < 1. A sufficient condition is

$$
\left(\frac{\beta.1}{g_c\left(n_c\right) + \delta\beta.0}\right)\left(\frac{g_s\left(n_c\right)^2 - g_c\left(n_c\right)^2}{2g_c\left(n_c\right)}\right) < 1,
$$

which is satisfied when $J\left(n_c\right)^2 < 1 + \frac{2}{\beta}$, which we assume holds. That is, by making β, the intergenerational discount factor, small enough, we can ensure that discounting holds. $\bullet$

Step 4: T thus satisfies Blackwell's sufficient conditions for a contraction. The contraction mapping theorem (see Stokey, Lucas, and Prescott (1989)) guarantees the existence of a unique function $\Lambda(n_c) \in C[0,1]$ that satisfies the system of functional equations. $\bullet$

Step 5: From step 1 it is clear that T actually maps $C[0,1]$ into the space of strictly decreasing functions, $C'[0,1]$. That is $T(C) \subseteq C' \subseteq C$. Using Corollary 1, (page 52), in Stokey, Lucas, and Prescott (1989), we get that the fixed point of $T \in C'$. In other words, the equilibrium $\Lambda(n_c)$ is strictly decreasing. From (15) through (18), and the monotonicity of $g_s\left(n_c\right)$ and $g_c\left(n_c\right)$, we get that the other functions that are part of the equilibrium, $a_c\left(n_c\right)$, $a_s\left(n_c\right)$, and $\Phi(n_c)$ exist and are uniquely determined once $\Lambda(n_c)$ is uniquely determined. $\bullet$

A.6 Proposition 1 (Linear Utility)

The fact that a unique steady state is reached in a single step with linear utility can be proved for any π and F. With linear utility, the utility cost of college education is the same for both types – $g_s(n_c) = g_c(n_c) = e$, $\forall\ n_c \epsilon [0, 1]$. It can then be seen from equations (11) and (12) that $a_s^*(n_c) = a_c^*(n_c) \equiv a^*(n_c)$. From (10), $x(n_c) = w_c(n_c) - w_s(n_c)$. It follows that $\Lambda'(n_c) < 0$. The law of motion (8) reduces to:

$$\Phi(n_c) = \int_{a^*(n_c)}^1 \pi(a)\, dF(a).$$

We can rewrite (12) for the linear utility case can be written as:

$$\beta\pi(a^*(n_c))\left[w_c(\Phi(n_c)) - w_s(\Phi(n_c))\right] = e.$$

Suppose a^* increases with n_c, which implies that Φ decreases with n_c. These imply that the left side is increasing in n_c, while the right side is constant. Thus it is not possible to have a^* increasing in n_c. Similarly, if a^* decreases with n_c, Φ increases with n_c, and the left side is decreasing in n_c, while the right side is constant. Thus the possibility of a^* decreasing in n_c is also ruled out. The only possibility is that a^* is independent of n_c – that is, $a^*(n_c) \equiv a^*$. The unique a^*, which is common to both types of parents, is given by:

$$\beta\pi(a^*)\left[w_c\left(\int_{a^*}^1 \pi(a)\, dF(a)\right) - w_s\left(\int_{a^*}^1 \pi(a)\, dF(a)\right)\right] = e,$$

with n_c^* then given by $\int_{a^*}^1 \pi(a)\, dF(a)$. Thus, when utility is linear, the economy ends up at this unique steady state in one step. •

A.7 Computation

The computational strategy that we use involves approximating the two functions Φ and Λ, defined in equations (8) and (9). The function Φ depends on $a_i^*(n_c; \Phi)$, which in turn depends on the function Λ (refer to equations 11 and 12). Consequently, we need to first approximate Λ. We approach this by nesting the estimation of Λ within the estimation of Φ. The algorithm that we use is a Chebyshev interpolation algorithm. (See Judd (1998) for details on constructing the Chebyshev nodes and coefficients.)

1. We start by selecting N Chebyshev interpolation nodes, $n_c^k,\ k = 1, ..., N$. These nodes will be used in approximating both Φ and Λ. We can do this because both are functions of the same variable.

2. Next, we choose an initial guess for $\Phi(n_c^k),\ \forall k$.

3. We can then compute the Chebyshev coefficients, bp_i. We can then construct an approximation of Φ at any point n_c: $\Phi(n_c) = \sum_1^N bp_i T_i(n_c)$.

    i. This is the inner loop devoted to approximating Λ. We start here by choosing an initial guess for $\Lambda(n_c^k),\ \forall k$.

    ii. We construct the Chebyshev coefficients bl_i in the same manner as we constructed the bp_i.

iii. We then approximate Λ at $\Phi(n_c^k)$, $\forall k : \Lambda(\Phi(n_c^k)) = \sum_1^N bl_i T_i(\Phi(n_c^k))$.

iv. Given these approximations we can construct a_c^{k*} $\forall k$ and a_s^{k*} $\forall k$, following equations (11) and (12). Then we can update $\Lambda(n_c^k)$, $\forall k$ following Equation (9).

v. Lastly we recalculate the bl_i and check for convergence. If the bl_i have converged, we continue on to step 4. If the bl_i have not converged, we return to step iii, with the updated bl_i.

4. Given the approximation of Λ, we can update Φ at all k nodes, following equation (8).

5. Lastly we recalculate the bp_i and check for convergence. If the bp_i have converged, we are finished. If the bp_i have not converged, we return to step i, with the updated Φ, and reapproximate Λ.

6. Newton's algorithm can then be used on $\Phi(n_c^*) = n_c^*$ to compute the steady state.

References

[1] Andrade, Eduardo De Carvalho. (1998). "Growth, Distribution, and School Policy," Ph.D. dissertation, University of Chicago.

[2] Autor, David H., Lawrence F. Katz, and Alan B. Krueger (1998). "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, November 1998, 1169-1213.

[3] Barro, Robert J. and Jong-Wha Lee. (1996). "International Measures of Schooling Years and Schooling Quality," *AEA Papers and Proceedings*, vol 86 no 2, 218-223.

[4] Becker, Gary S. (1993). *A Treatise on the Family*, (Enlarged Edition) Cambridge, MA: Harvard University Press.

[5] Cameron, Stephen V., and James J. Heckman. (1998). "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, vol 106 no 2, 262-333.

[6] Fernandez, Raquel and Jordi Gali. (1997). "To Each According To ...? Markets, Tournaments, and the Matching Problem with Borrowing Constraints," *CEPR Discussion Paper*, no 1627, May 1997.

[7] Greenwood, Jeremy and Boyan Jovanovic. (1990). "Financial Development, Growth, and the Distribution of Income," *Journal of Political Economy*, vol 98 no 5, 1076-1107.

[8] Heckman, James J. and Peter J. Klenow. (1997). "Human Capital Policy," Working Paper, University of Chicago.

[9] Hanushek, Eric A., Ban Chuan Cheah, and Charles Ka Yui Leung. (1997). "Redistribution through Education and Other Transfer Mechanisms," Working Paper, University of Rochester.

[10] Judd, Kenneth L. (1998). *Numerical Methods in Economics*, Cambridge, MA: The MIT Press.

[11] Kremer, Michael. (1993). "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, August 1993, 551-575.

[12] Krusell, Per, and Anthony A. Smith, Jr. (1998). "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, vol 106 no 5, 867-896.

[13] Lucas, Robert E., Jr. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics*, vol 22, 3-42.

[14] Loury, Glenn C. (1981). "Intergenerational Transfers and the Distribution of Earnings." *Econometrica*, vol 49 no 4, 843-867.

[15] McPherson, Michael S. and Morton Owen Schapiro. (1991). *Keeping College Affordable: Government and Educational Opportunity*, Washington, D.C.: The Brookings Institution.

[16] Murphy, Kevin M. and Finis Welch. (1992). "The Structure of Wages," *The Quarterly Journal of Economics*, February 1992, 285-326.

[17] OECD. (1997). *Education Policy Analysis, 1997*, Paris, France: OECD Press.

[18] Stokey, Nancy L. (1996). "Free Trade, Factor Returns, and Factor Accumulation," *Journal of Economic Growth*, vol 2 no 1, 421-447.

[19] Stokey, Nancy L., Robert E. Lucas, Jr., and Edward C. Prescott. (1989) *Recursive Methods in Economic Dynamics*, Cambridge, MA: Harvard University Press.