

Does Democracy Engender Equality?

By

John E. Roemer

September 2001

COWLES FOUNDATION DISCUSSION PAPER NO. 1281R



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

<http://cowles.econ.yale.edu/>

Abstract “ Does democracy engender equality?”

John E. Roemer
Dept of Political Science
Yale University
PO Box 208301
New Haven CT 06520

Many suppose that democracy is an ethos which includes, inter alia, a degree of economic equality among citizens. In contrast, we conceive of democracy as ruthless political competition between groups of citizens, organized into parties. We inquire whether such competition, which we assume to be concerned with distributive matters, will engender economic equality in the long run.

The society consists of an infinite sequence of generations, each comprised of adults and their children. Adults care about household consumption, and the future wages of their children, which are determined by educational policy. A given generation is characterized by the distribution of wages earned by its adults. Parties form and propose policies to redistribute income among households, and to invest in the education of children; the educational policy that is victorious determines the distribution of wages in the next generation of adults.

A political equilibrium concept is proposed which determines two parties endogenously, and their proposed policies in political competition. One party wins the election (stochastically). This process determines a sequence of wage distributions across the generations, and we ask: Under what conditions does the wage distribution tend to one of equality?

We show that, under a technological assumption that appears to hold empirically, there is no assurance that wage equality is eventually achieved, but if certain ‘social norms’ hold, which restrict the space of acceptable political policies, then equality is eventually achieved. We suggest, moreover, that the social norms in question will tend to hold, the more technologically developed the democracy is.

JEL numbers

“Does democracy engender equality?”

by

John E. Roemer*

§1 Introduction

Among types of political system, the one most identified in contemporary western society with the production of justice is democracy. Even on the political left, democracy has largely replaced socialism as the regime desideratum. Just as those socialists who were dissatisfied with aspects of Soviet society claimed that the Soviet regime was not *real* socialism, so those who continue to be dissatisfied with, for example, the American system, now argue that it is not an instance of real democracy. Real democracy is thought to be a political system in which genuine representation of all citizens – and even justice – is achieved.

The identification of democracy with justice is not simply a practice of many political theorists: perhaps the most important aspect of political transformation in the world in the last fifty years has been the toppling of authoritarian regimes, and their replacement with democratic ones. Just as socialism was a powerful movement in the first half of the twentieth century – by 1950, fully one-third of the world’s peoples lived under regimes that described themselves as socialist – so democracy has been the

• Departments of Political Science and Economics, Yale University. This paper originated in a series of discussions with Ignacio Ortuno-Ortin. His ideas and critique have been immensely valuable. I am also indebted to Roger Howe and Herbert Scarf for mathematical discussions, and to John Geanakoplos for valuable critique at a later stage.

massively appealing political doctrine in, let us say, the period since 1960. And as it was an error of socialists to identify socialism with All Good Things, so now it is an error of democrats to identify democracy with All Good Things¹. The most common example of this fallacy is when some say that regime X cannot be a democracy, because it sustains Bad Thing Y (oppression of women, abrogation of civil rights, etc.). If democracy is defined as a set of political institutions, rather than as an ethos, then the correct approach is to study what those institutions entail. Perhaps, for example, both the oppression of women and its absence can co-exist with democracy.

In this article, we undertake a study of this kind: we ask whether democracy, understood as a system of political competition between parties that represent different coalitions of citizens, will engender justice, or – as we here interpret justice, equality. Of course, we cannot answer that broad question generally, and so we narrow it down to something manageable. In particular we focus upon the role of public education as an instrument for reducing the differentials in human capital that would otherwise obtain, and we ask whether democracy will entail the equalization of human capital through political decisions concerning educational investment.

We model the following society, one which reproduces itself over many generations. At the initial date, there are households led by adults characterized by a distribution of human capital, that is, capacities to produce income. Each adult has one child. The human capital the child will have, when next period, she has become an adult,

is a monotone increasing function of her parent's level of human capital and the amount that was invested in her education. This relationship is (until section 6) deterministic, and describes the educational production function for all children. Thus, it requires more investment to bring a child from a poor (low human capital) family up to a given level of human capital than a child from a richer family. All parents have the same utility function: a parent cares about her household's consumption (that is, her after-tax income), and the earning power her child will have, as an adult. We will, for simplicity, assume that adults do not value leisure.

Educational finance is, until section 5, purely public. The polity of adults, at each date, must make four political decisions: how much to tax themselves, how to split the tax revenues between a redistributive component for households' consumption and the educational budget, how to partition the budget for redistribution among adults, and how to partition the educational budget as investment in particular children, according to their type (that is, their parental human capital). Once these political decisions are implemented, a distribution of human capital is determined for the next generation. When the present children become adults, characterized by that distribution of human capital, they face the same four political decisions. We wish to study the asymptotic distribution of human capital of this dynamic process.

In the society we have described, a child is characterized by the family (household) into which he is born, for his capacity to transform educational investment

¹ There are many people who identify democracy with justice. For instance, Adolfo Perez Esquivel, a Nobel Peace Prize laureate, recently said, "The vote does not define democracy. Democracy means justice and equality." (*The Daily Journal* [Caracas], July 12, 2001)

into future earning power is determined by his family background, proxied by his parent's human capital. We imagine that the transmission of 'culture' to the child is indicated by the parent's human capital endowment. We view the child's capacity successfully to absorb educational investment, and transform it into human capital, as a circumstance beyond his control, and so a society of this kind that wished rapidly to equalize opportunities for all children would compensate children from poorer families with more educational investment. Equality of opportunity would be achieved when all adults have the same human capital, for that means, as children in the previous generation, the compensation for disadvantageous circumstances was complete. (See Roemer (1998) for the theory of equality of opportunity, based upon social compensation for disadvantageous circumstances.) In the real world, equality of opportunity does not require equalizing outcomes in this way, because people may remain responsible for some aspect of their condition, even after the necessary compensation for disadvantage has been made. But in our model there is no such element, and so, if we take equality of opportunity as our conception of justice, then justice will have been achieved exactly when the wage-earning capacities of all adults are equal.

One might object that it is sufficient to equalize (post-tax) *incomes* for justice. But it may well be the case that individuals derive welfare not only from consumption, but from their human capital, and so we insist that this more demanding condition of human-capital equality is the one of interest. Indeed, if one's human capital is an enabler of self-realization, then it is surely the case that justice would require a concern with levels of human capital in a society, not simply income levels.

We will stipulate a democratic process for solving society's political problems, at each generation. Our question becomes: How close will the asymptotic distribution of human capital engendered by this political process be to an equal distribution?

The focus of our model will be on that democratic process. We employ a concept of democratic political equilibrium that takes as data the distribution of preferences of the polity over a given policy space, and produces as its output an endogenous partition of the polity into two political parties, a policy proposal by each party, and a probability that each party will win the election. We suppose that an election occurs, and the policy of the victorious party is implemented. Our procedure will be to begin with a distribution of adult human capital at date 0, which will determine the distribution of adult preferences at date 0. The dynamic process is thus initialized.

Although we have described the political choice as consisting of four independent decisions, we will in fact model the political problem as one on an infinite dimensional policy space. That policy space, denoted T , will consist of pairs of functions (ψ, r) where $\psi(h)$ is the after-tax household income of an adult with human capital h , and $r(h)$ is the public educational investment in a child from an h -family. The only restrictions on these functions are that they be continuous, jointly satisfy a budget constraint, and satisfy two constraints that we call *social norms*. Thus, the present analysis marks a substantial technical advance over analyses in political economy that must limit their scope to unidimensional policy spaces, or policy spaces of small dimension. But the advance is not merely technical. It is surely artificial to restrict a democratic polity's choice of policies to ones with simple mathematical properties, such as linearity. Our

ability to solve the political problem with no such restrictions means that we are able to model the democratic struggle as *ruthlessly competitive*: no holds, in the sense of unmotivated restrictions on the nature of policy proposals, are barred, except (importantly), those precluded by the social norms.

That political equilibrium concept is ‘party unanimity Nash equilibrium with endogenous parties.’ In two recent articles, I introduced the concept of ‘party unanimity Nash equilibrium (PUNE), (Roemer [1999, 1998]). The extension to ‘PUNE with endogenous parties’ is introduced in Roemer (2001, Chapter 13). The endogenous-party aspect is grafted from a model of Baron (1993).

It is probably fair to say that most articles in political economy propose a relatively sophisticated model of the economy, and a trivial model of politics (standardly, political equilibrium consists in both parties’ proposing the median voter’s ideal point, or, more generally, a Condorcet winner in the policy space). Our approach here is just the opposite: the economy is very simple, but the politics are quite complex. Our first justification for the complex politics is that it enables us to solve the problem of political equilibrium with multi- and even infinite dimensional policy spaces, when Condorcet winners do not exist. Our second justification, for the problem at hand, is that our focus is upon the workings of democracy, and therefore, a careful articulation of democratic institutions is appropriate. Of course, a more highly articulated model of the economy would also be desirable, if tractability were not sacrificed.

In section 2, the definition of political equilibrium that we will use, and a companion concept of quasi-equilibrium, are presented. In section 3, we characterize the

policies in the political equilibria of the model. Section 4 does the dynamics. Section 5 relaxes the assumption that all educational investment is public. Section 6 introduces a stochastic element in the determination of the human capital of the next generation, and section 7 concludes with a brief discussion of the implications of our results for democratic theory.

§2 Party unanimity Nash equilibrium with endogenous parties (PUNEED)

In this section, I define PUNEED and a related concept.

Let H be a set of voter types, where $h \in H$ is distributed according to a probability measure \mathbf{F} in the society in question. Let T be a set of policies. There is a function $v: T \times H \rightarrow \mathbf{R}$ which represents the preferences of types over policies; thus $v(\cdot, h)$ is the utility function of type h on T . For each h , we assume that $v(\cdot, h)$ is a von Neumann-Morgenstern utility function for lotteries on T .

Let $t^1, t^2 \in T$ be two policies; we define $\pi(t^1, t^2)$, the probability that policy t^1 defeats policy t^2 . Our datum is a function $\pi^*: [0,1] \rightarrow [0,1]$, such that $\pi^*(0) = 0, \pi^*(1) = 1$, and π^* is strictly increasing on $[0,1]$.

Let $\Omega(t^1, t^2)$ be the set of types who prefer t^1 to t^2 and $I(t^1, t^2)$ be the set of types who are indifferent between t^1 and t^2 . Then we define, *pro tem*²:

$$\pi(t^1, t^2) = \pi^*(\mathbf{F}(\Omega(t^1, t^2)) + \frac{1}{2}\mathbf{F}(I(t^1, t^2))). \quad (2.1)$$

² To be modified below.

In other words, $\mathbf{F}(\Omega(t^1, t^2)) + \frac{1}{2}\mathbf{F}(I(t^1, t^2))$ is the mass of voters who in principle will vote for t^1 — but perhaps some voters will make mistakes or perhaps \mathbf{F} is measured imperfectly. Equation (2.1) says that the probability that t^1 defeats t^2 is an increasing function of the ‘expected’ vote for t^1 .

A *party structure* is a partition of H into two elements. We specialize, now, to the case $H = \mathbf{R}_+$, and further specialize by requiring that both elements in a party structure be intervals: thus a party structure is characterized by a pivotal type h^* , with $L = [0, h^*)$ and $R = [h^*, \infty)$. We call the two parties Left (L) and Right (R).

Associated with a party is a utility function, which is the average of its members utility functions.

Thus

$$\left. \begin{aligned} v_L(t) &= \int_0^{h^*} v(t, h) d\mathbf{F}(h) \\ v_R(t) &= \int_{h^*}^{\infty} v(t, h) d\mathbf{F}(h) \end{aligned} \right\} \quad (2.2)$$

(We drop a multiplicative constant.) The utility functions $v(\cdot, h)$ are assumed to be cardinally measurable and unit comparable (CUC), so that averaging them makes sense.

All parties contain three factions: opportunists, reformists, and militants. (These factions are *not* to be identified with particular citizen types.) Each faction possesses a real-valued *payoff function* defined on $T \times T$. The payoff functions of the three factions in Left are defined by:

$$\left. \begin{aligned} {}^L\Pi^{\text{Opp}}(t^1, t^2) &= \pi(t^1, t^2) \\ {}^L\Pi^{\text{Ref}}(t^1, t^2) &= \pi(t^1, t^2)v_L(t^1) + (1 - \pi(t^1, t^2))v_L(t^2) \\ {}^L\Pi^{\text{Mil}}(t^1, t^2) &= v_L(t^1) \end{aligned} \right\} \quad (2.3)$$

with an analogous definition for Right's three factions. The three factions are interested, respectively, in winning (opportunists), party-member welfare (reformists), and publicity (militants). For elaboration, the reader is referred to Roemer (1999, 2001).

Definition 1. A party unanimity Nash equilibrium with endogenous parties (PUNEED) is a party structure (L, R) given by $L = [0, h^*)$ and $R = [h^*, \infty)$ with $h^* > 0$, and a pair of policies $t^L, t^R \in T$ such that

(A) there is no policy $t \in T$ such that

$${}^L\Pi^J(t, t^R) \geq {}^L\Pi^J(t^L, t^R), \text{ for } J = O, R, M$$

with at least one of these inequalities strict;

(B) there is no policy $t \in T$ such that

$${}^R\Pi^J(t^L, t) \geq {}^R\Pi^J(t^L, t^R), \text{ for } J = O, R, M$$

with at least one of these inequalities strict;

(C) $h \in L \Rightarrow v(t^L, h) \geq v(t^R, h)$
 $h \in R \Rightarrow v(t^R, h) \geq v(t^L, h)$.

The three payoff functions of a parties' factions each represent a complete order on $T \times T$. We may view their intersection as representing a quasi-order on $T \times T$. Then a PUNEED is a Nash equilibrium of the game played by these two quasi-orders, with the

additional requirement (C). Requirement (C) was initially proposed by Baron (1993) as modeling the stability of a party structure.

Remark 1. It is easily shown that the reformists are gratuitous in definition 1. That is, if we eliminate the reformist factions, we do not alter the set of equilibria. But notice, once this is done, we never need mention expected utility, since only the reformists calculate that. It thus suffices that $\{v(\cdot, h) | h \in H\}$ be a profile of CUC utility functions (i.e., they need not represent preferences over lotteries).

Remark 2. It is now convenient to alter the convention on how indifferent voters vote, in the presence of parties. When parties are present, we will say that a voter who is indifferent between policies votes for the policy of his party. Thus, formally, we now *revise* the definition of π to:

$$\pi(t^L, t^R) = \pi^*(\mathbf{F}(\Omega(t^L, t^R)) + \mathbf{F}(L \cap I(t^L, t^R))). \quad (2.1')$$

Remark 3. In Roemer (2001, Chapter 8), it is shown that if sufficient convexity is present, then every PUNEEP can be viewed as the outcome of generalized Nash *bargaining* between the militant and opportunist factions of each party, given the other party's proposal. There is, in general, a two dimensional manifold of PUNEEP. Each one is characterized by specifying the relative bargaining strengths of the two active factions in each party – thus, two positive numbers. Thus, parties compete with each other à la Nash *equilibrium*, while internal factions bargain with each other à la Nash *bargaining*. The PUNEEP concept thus owes its origins doubly to John Nash.

We now further specialize to the case that \mathbf{F} has a continuous, strictly increasing distribution function, F , on \mathbf{R}_+ .

We next define an auxiliary notion that is useful in the analysis.

Definition 2. A *quasi-PUNE* is an ordered pair $(h^*, y) \in H \times \mathbf{R}$ and a pair of policies

$t^L, t^R \in T$, such that $v(t^L, h^*) = y = v(t^R, h^*)$ and:

2A. t^L solves

$$\max \int_0^{h^*} v(t, h) d\mathbf{F}(h)$$

subject to

$$t \in T$$

$$h \in [0, h^*) \Rightarrow v(t, h) \geq v(t^R, h) \quad (\text{L0})$$

$$v(t, h^*) \geq y \quad (\text{L1})$$

2B. t^R solves

$$\max \int_{h^*}^{\infty} v(t, h) d\mathbf{F}(h)$$

subject to

$$t \in T$$

$$h \in [h^*, \infty) \Rightarrow v(t, h) \geq v(t^L, h) \quad (\text{R0})$$

$$v(t, h^*) \geq y \quad (\text{R1})$$

2C. Constraints (L1) and (R1) bind at t^L and t^R respectively.

We have:

Proposition 1. Let v be continuous in h . If (t^L, t^R, h^*) is a PUNE, then (t^L, t^R, h^*, y) is a quasi-PUNE, with $y = v(t^L, h^*)$.

Proof:

Let (t^L, t^R, h^*) be a PUNEEP with $h^* > 0$, $L = [0, h^*)$, and $R = [h^*, \infty)$. By Remark 2, $\pi(t^L, t^R) = \pi^*(F(L))$ and $0 < \pi^*(F(L)) < 1$ by definition of π^* and the fact that F is strictly increasing on \mathbf{R}_+ . By Condition 1A of PUNEEP, there is no policy t that gives Left's militants a higher payoff than they receive at t^L and gives a higher probability of victory against t^R . In particular, there is no policy t that gives Left's militants a higher payoff than at t^L and such that

$$h \in [0, h^*) \Rightarrow v(t, h) \geq v(t^R, h),$$

and

$$v(t, h^*) > y,$$

for if there were, then, by continuity of v in h there would be an interval $[h^*, h^* + \varepsilon)$ such that

$$h \in [h^*, h^* + \varepsilon) \Rightarrow v(t, h) > v(t^R, h).$$

It would then follow that at least the set of voters $L \cup [h^*, h^* + \varepsilon)$ would favor t and so a higher probability of victory could be achieved for Left at no cost to her militants.

It therefore follows that statement 2A of definition 2 is true, and that (L1) binds.

In like manner, statement 2B of definition 2 is true and (R1) binds, which concludes the proof.

The converse of Proposition 1 is not true: there may be quasi-PUNEs that are not PUNEES. For if (t^L, t^R, h^*, y) is a quasi-PUNE, it is possible that there exists a policy t

which improves the payoff of both Left's militants and opportunists, by assembling a set of voters who favor t over t^R that is disconnected and does not contain h^* .

We can now give a preview of our strategy. In our politico-economic environment, we can fully characterize the set of quasi-PUNEs: the nice fact is that no recourse to fixed point theorems is required, only to optimization methods. We will further note that the set of PUNEEDs is a non-empty subset of the set of quasi-PUNEs. We then conduct our dynamic analysis assuming that each generation's political equilibrium is some quasi-PUNE. Whatever we conclude will hold *a fortiori* for societies whose political equilibria are genuine ones, that is, PUNEEDs. In this manner we avoid ever having to solve the very difficult problem of characterizing precisely the set of PUNEEDs.

§3 The manifold of quasi-PUNEEDs

Throughout this section, we analyze the society at one date.

1. The politico-economic environment

(i) Preferences

A typical society, in our problem, consists of a continuum of adult types, each characterized by his/her human capital h , where h is distributed according to a probability measure \mathbf{F} , whose mean is denoted μ . Each adult has one child. Adults care about their own consumption, and their child's (future) human capital.

We assume:

$$u(x, h') = \log(x - C_0) + \gamma \log h', \quad (3.1)$$

where x is the household's consumption, or after-tax income, and h' is the child's (future) human capital. C_0 is to be interpreted as the level of household income below which the parent would rather send the child to work than to school. We think of zero consumption as minimal household consumption. If the household's income is less than C_0 , then it would like to devote it solely to consumption, and none to education. Note there is no preference for leisure³.

(ii) Technology

If r is invested in the education of a child whose parent is of type h then the child's future human capital will be

$$h' = \alpha h^b r^c \quad (3.2)$$

where α, b, c are positive constants. We assume that

A1. $0 < b + c < 1$.

b is the elasticity of child's human capital w.r.t. parental human capital and c is the elasticity of child's human capital w.r.t. educational investment. The influence of the parent's human capital on the child's human capital we think of as operating through family culture, or perhaps neighborhood effects (if neighborhoods are income-segregated). Later on, we will insert a stochastic term into (3.2), but for the present, note that the

³ Formally, we can represent the ordinal preferences of the adult over the complete consumption set by the continuous utility function:

$$U(x, h') = \begin{cases} (x - C_0)h'^\gamma + C_0, & \text{if } x \geq C_0 \\ x, & \text{if } x < C_0 \end{cases}$$

human capital determination process is deterministic. Bénabou (in press) uses a relationship like (3.2), and notes that, according to the empirical literature, assumption A1 holds.

If an adult of type h works at her full potential then her (pre-tax) earnings are h . Thus human capital is measured in units of income-earning capacity.

(iii) The policy space

Let C be the space of continuous functions on the domain \mathbf{R}_+ . A *policy* is a pair of functions $(\psi, r) \in C^2$ such that

$$\int_0^\infty (\psi(h) + r(h)) d\mathbf{F}(h) \leq \mu \quad (3.3)$$

and two social norms hold, as described below. The interpretation is that $\psi(h)$ is the after-tax income of an adult of type h , and $r(h)$ is the public educational investment in a child from an h -family.

All educational investment is public. Specification of a policy (ψ, r) solves the four political problems described in Section 1.

Thus the indirect utility function $v: T \times H \rightarrow \mathbf{R}$ is given by

$$\begin{aligned} v(\psi, r, h) &= \log(\psi(h) - C_0) + \gamma \log \alpha h^b r(h)^c \\ &= \log(\psi(h) - C_0) + \gamma \log \alpha h^b + \gamma c \log r(h) \\ &\doteq \log(\psi(h) - C_0) + \gamma c \log r(h) \end{aligned} \quad (3.4)$$

where, in the last line of (3.4), we have dropped a gratuitous constant term.

The two social norms are:

$$\text{for all } h, \psi'(h) + r'(h) \geq 0, \quad (3.5a)$$

$$\psi'(h) + r'(h) \leq m, \quad (3.5b)$$

where ‘prime’ indicates derivative, and the inequalities are meant to hold where the derivatives exist. m is a positive constant, a parameter of the problem. We call $\psi(h) + r(h)$ the total resource bundle allocated to an h household, so (3.5ab) restrict the rates at which the total resource bundle changes with h . We call (3.5ab) social norms, as they are not motivated by political competition or incentive compatibility considerations.

There is a natural incentive compatibility condition, that adult utility be non-decreasing in h , so that no adult would have an incentive to work at a lower income-earning capacity than her true capacity. The local version of this condition is:

$$\frac{\psi'(h)}{\psi(h)} + \gamma c \frac{r'(h)}{r(h)} \geq 0. \quad (3.5aa)$$

Some might prefer to substitute (3.5aa) for (3.5a) in the model, but doing so renders the analysis below much more difficult. (It converts what will be a convex optimization problem on an infinite-dimensional space to a non-convex problem.) In the interests of simplicity, and not diffusing attention from our main concern, we use (3.5a) in lieu of (3.5aa). We conjecture, however, that the results we report would remain identical if (3.5a) were replaced with (3.5aa).

Thus, our policy space is

$$T = \{(\psi, r) \in C^2 \mid (3.3), (3.5a), \text{ and } (3.5b) \text{ hold}\}.$$

B. Quasi-PUNEs

For a given point $(h^*, y) \in \mathbf{R}_+ \times \mathbf{R}$, consider the following two programs:

$$\left. \begin{aligned}
& \max_{(\psi, r) \in C^2} \int_0^{h^*} \log(\psi(h) - C_0) + \gamma c \log r(h) d\mathbf{F}(h) \\
& \text{s.t.} \\
& 0 \leq \psi'(h) + r'(h) \leq m \quad (3.61) \\
& \int_0^\infty (\psi(h) + r(h)) d\mathbf{F}(h) \leq \mu \quad (3.62) \\
& \log(\psi(h^*) - C_0) + \gamma c \log r(h^*) \geq y \quad (3.63)
\end{aligned} \right\} \quad (3.6)$$

$$\left. \begin{aligned}
& \max_{(\psi, r) \in C^2} \int_{h^*}^\infty \log(\psi(h) - C_0) + \gamma c \log r(h) d\mathbf{F}(h) \\
& \text{s.t.} \\
& 0 \leq \psi'(h) + r'(h) \leq m \quad (3.71) \\
& \int_0^\infty (\psi(h) + r(h)) d\mathbf{F}(h) \leq \mu \quad (3.72) \\
& \log(\psi(h^*) - C_0) + \gamma c \log r(h^*) \geq y \quad (3.73)
\end{aligned} \right\} \quad (3.7)$$

Let (h^*, y) be such that solutions (ψ^L, r^L) and (ψ^R, r^R) exist to (3.6) and (3.7), respectively, and such that inequalities (3.63) and (3.73) bind at the solutions. We will show that the following hold:

1. $0 \leq h \leq h^* \Rightarrow v(\psi^L, r^L, h) \geq v(\psi^R, r^R, h)$, and
2. $h^* \leq h \leq \infty \Rightarrow v(\psi^R, r^R, h) \geq v(\psi^L, r^L, h)$.

It will follow that (ψ^L, r^L) and (ψ^R, r^R) constitute a quasi-PUNE at (h^*, y) , and that solutions of programs (3.6) and (3.7) at which (3.63) and (3.73) bind comprise precisely the quasi-PUNEs for our problem.

Our first task is to characterize the set

$$\Gamma = \{(h^*, y) \in \mathbf{R}_+ \times \mathbf{R} \mid \text{solutions to (3.6) and (3.7) exist at which (3.63) and (3.73) bind}\}.$$

Note that $T = \{(\psi, r) \in C^2 \mid (3.61) \text{ and } (3.62) \text{ hold}\}$. Consider the following three programs:

$$\left. \max_{(\psi, r) \in T} \int_0^{h^*} v(\psi, r; h) d\mathbf{F}(h) \right\} \quad (3.8)$$

$$\left. \max_{(\psi, r) \in T} \int_{h^*}^{\infty} v(\psi, r; h) d\mathbf{F}(h) \right\} \quad (3.9)$$

$$\left. \max_{(\psi, r) \in T} v(\psi, r; h^*) \right\} \quad (3.10)$$

Let their solutions be denoted τ^L, τ^R , and τ^* , respectively. Let $y^*(h^*)$ be the value of program (3.10), i.e.

$$y^*(h^*) = v(\tau^*, h^*),$$

and define

$$y^L(h^*) = v(\tau^L, h^*), \quad y^R(h^*) = v(\tau^R, h^*).$$

We have:

Proposition 2. For h^* given, $(h^*, y) \in \Gamma$ iff

$$\max[y^L(h^*), y^R(h^*)] \leq y \leq y^*(h^*). \quad (3.11)$$

Proof:

1. Suppose $y > y^*(h^*)$. Then there is no feasible solution to (3.6) or (3.7), for (3.63) will never hold on T . Thus we must have $y \leq y^*(h^*)$ if $(h^*, y) \in \Gamma$.
2. Suppose $y < y^L(h^*)$. Then constraint (3.63) is not binding at the solution to (3.6), since the solution to (3.8) is indeed the solution to (3.6). Similarly, if $y < y^R(h^*)$, then constraint (3.73) is not binding at the solution to (3.7). Thus $(h^*, y) \in \Gamma$ implies $y \geq \max[y^L(h^*), y^R(h^*)]$.

3. Conversely, if (3.11) holds, then the opportunity sets of (3.6) and (3.7) are non-empty, and at the optimal solutions, (3.63) and (3.73) must bind, because
- $$y \geq \max[y^L(h^*), y^R(h^*)]. \quad \blacksquare$$

Actually, the proof of Proposition 2 has ignored the compactness issue - whether non-emptiness of the opportunity sets for programs (3.6) and (3.7) implies the attainment of (optimal) solutions. We shall show below that if (3.11) holds, solutions are indeed attained.

Thus we have argued that

$$\Gamma = \{(h^*, y) \mid \max[y^L(h^*), y^R(h^*)] \leq y \leq y^*(h^*)\}.$$

The virtue of the quasi-PUNE notion is now evident: we can characterize the set Γ , and thus the set of quasi-PUNEs, merely by solving the three programs (3.8), (3.9), and (3.10). No fixed-point machinery is needed to do this.

We next solve these three programs.

Proposition 3. Let (Ψ, r) be a solution to (3.6), (3.7), (3.8), (3.9), or (3.10).

Then

$$r(h) = \begin{cases} \gamma c(\Psi(h) - C_0), & \text{if } \Psi(h) \geq C_0 & (3.12a) \\ 0, & \text{if } 0 \leq \Psi(h) < C_0 & (3.12b) \end{cases}$$

Lemma 1. Let $X \geq C_0$ be the total resource dedicated to household h . Then the household's optimal distribution of X between consumption ψ and educational investment r is

$$\psi = C_0 + \frac{1}{1 + \gamma c} (X - C_0)$$

$$r = \frac{\gamma c}{1 + \gamma c} (X - C_0).$$

If $X < C_0$ then the optimal allocation is

$$\psi = X, \quad r = 0.$$

Proof: The household would choose consumption ψ to maximize

its utility, which leads immediately to the claim. ■

Proof of Proposition 3:

Let $(\hat{\psi}, \hat{r})$ be a solution to program (3.6), and suppose that the claim were false. Define

$\hat{z}(h) = \hat{\psi}(h) + \hat{r}(h)$, and define

$$\left. \begin{aligned} \psi(h) &= C_0 + \frac{1}{1 + \gamma c} (\hat{z}(h) - C_0) \\ r(h) &= \frac{\gamma c}{1 + \gamma c} (\hat{z}(h) - C_0) \end{aligned} \right\} \text{if } \hat{z}(h) \geq C_0$$

$$\left. \begin{aligned} \psi(h) &= \hat{z}(h) \\ r(h) &= 0 \end{aligned} \right\} \text{if } \hat{z}(h) < C_0.$$

It is straightforward to check that $(\psi, r) \in T$. Furthermore, each household receives the same total resource at $(\hat{\psi}, \hat{r})$ and at (ψ, r) . But according to Lemma 1, for every h , $(\psi(h), r(h))$ is the optimal way for household h to allocate the total resource assigned to it between consumption and education. Therefore the objective function of (3.6) increases if we substitute (ψ, r) for $(\hat{\psi}, \hat{r})$, a contradiction. (To be precise, the argument shows that $(\hat{\psi}, \hat{r})$ must equal (ψ, r) except possibly on a set of \mathbf{F} -measure zero. Continuity then completes the argument.) ■

Remark 4. If we replaced social norm (3.5a) with incentive compatibility (3.5aa), Proposition 3, although probably true, is much more difficult to prove. It is for this reason that we employ (3.5a).

Now suppose that the solutions to (3.8), (3.9), and (3.10) all give at least C_0 in total resource to every household. Then we know, for each of those solutions, that (3.12a) holds. By substituting from (3.12a), we can reduce (3.8), (3.9), and (3.10) to the following three programs:

$$\left. \begin{array}{l} \max_{\psi \in C} \int_0^{h^*} \log(\psi(h) - C_0) d\mathbf{F}(h) \\ \text{s.t.} \\ \int_0^\infty \psi(h) d\mathbf{F}(h) \leq k_1 \\ 0 \leq \psi'(h) \leq k_2 \end{array} \right\} \quad (3.8a)$$

$$\left. \begin{array}{l} \max_{\psi \in C} \int_{h^*}^{\infty} \log(\psi(h) - C_0) d\mathbf{F}(h) \\ \text{s.t.} \\ \int_0^{\infty} \psi(h) d\mathbf{F}(h) \leq k_1 \\ 0 \leq \psi'(h) \leq k_2, \end{array} \right\} \quad (3.9a)$$

and

$$\left. \begin{array}{l} \max_{\psi \in C} \log(\psi(h^*) - C_0) \\ \text{s.t.} \\ \int_0^{\infty} \psi(h) d\mathbf{F}(h) \leq k_1 \\ 0 \leq \psi'(h) \leq k_2, \end{array} \right\} \quad (3.10a)$$

where

$$k_1 = \frac{\mu + \gamma C_0}{1 + \gamma c} \quad \text{and} \quad k_2 = \frac{m}{1 + \gamma c}.$$

Conversely, if the solutions to (3.8a), (3.9a), and (3.10a) all satisfy $\psi(h) > C_0$ for $h > 0$, then by consideration of (3.12a), we have also solved (3.8), (3.9), and (3.10).

We assume:

$$\text{A2. } \mu > C_0 \text{ and } m \leq \frac{\mu - C_0}{\mu}.$$

We have:

Proposition 4. If A2 holds, then:

a. *The solution to (3.8a) is*

$$\Psi_{h^*}^L(h) \equiv k_1;$$

in particular, we have $\Psi_{h^*}^L(h) > C_0$.

b. The solution to (3.9a) is

$$\begin{aligned} \Psi_{h^*}^R(h) &= \Psi_0^R + k_2 h, \\ \text{where } \Psi_0^R &= k_1 - k_2 \mu; \\ \text{in particular } \Psi_0^R &\geq C_0. \end{aligned}$$

c. The solution to (3.10a) is illustrated in Figure 1. It is given by

$$\Psi_{h^*}(h) = \begin{cases} \Psi_0^* + k_2 h, & \text{if } 0 \leq h \leq h^* \\ \Psi_0^* + k_2 h^* & \text{if } h > h^* \end{cases},$$

where Ψ_0^* is the solution of the equation

$$\Psi_0^* + k_2 \int_0^{h^*} h d\mathbf{F}(h) + k_2 h^* (1 - \mathbf{F}(h^*)) = k_1. \quad (3.13)$$

Indeed, $\Psi_0^* \geq C_0$, as pictured in figure 1.

Note that the three functions $\Psi_{h^*}^L$, $\Psi_{h^*}^R$, and Ψ_{h^*} are all greater than C_0 for $h > 0$.

Proof: See appendix.

Although we offer a formal proof of Proposition 4 in the appendix, the following intuitive argument will probably convince the reader of the proposition's validity.

First, consider program (3.8a). The *benefit* to household h is $\log(\psi(h) - C_0)$; the *cost* (to the optimizer) of supplying household h is $\psi(h)$; hence the benefit-cost ratio, $\frac{\log(\psi(h) - C_0)}{\psi(h)}$, is non-increasing in h , because $\psi'(h) \geq 0$ is required. So the optimizer should give as much of the resource as possible to low h : front-loading, so to speak. The binding constraint is $\psi'(h) \geq 0$: so the planner allocates $\psi(h) \equiv k_1$.

Second, consider program (3.9a). The planner wants to minimize the resource going to $[0, h^*]$. So, whatever the optimal value of ψ at h^* , ψ should descend to the left of h^* as fast as possible – that is, at rate k_2 , to $h=0$. to the right of h^* , ψ should ascend as fast as possible to (high) value and then maintain that value forever. But $m \leq \frac{\mu - C_0}{\mu}$ means that the function $\psi_{h^*}^R$ has exactly these properties – indeed, $\psi_{h^*}^R$ never reaches the ‘high plateau.’ The budget constraint is

$$\psi_0^R + \int_0^{\infty} k_2 h d\mathbf{F}(h) = k_1;$$

if $m = \frac{\mu - C_0}{\mu}$, this reduces to

$$\psi_0^R + \frac{\mu - C_0}{1 + \gamma c} = \frac{\mu + \gamma c C_0}{1 + \gamma c},$$

and hence $\psi_0^R = C_0$. If $m < \frac{\mu - C_0}{\mu}$, then $\psi_0^R > C_0$.

Third, consider (3.10a). Clearly it is a waste to give any resource to $h > h^*$, so we must have

$$\psi(h) = \psi(h^*) \text{ for } h > h^*.$$

Now the optimizer wants to minimize what goes to $[0, h^*)$, conditional upon reaching a high value at h^* , so ψ should descend rapidly (at rate k_2) to the left of h^* . Thus the stated function ψ_{h^*} will be the solution as long as $\psi_0^* \geq C_0$. Equation (3.13) is the budget constraint, and the constraint on m implies that $\psi_0^* \geq C_0$. This is obvious because it is obvious that $\psi_0^* > \psi_R^*$, because $\psi_{h^*}^R$ gives more resource to $[h^*, \infty)$ than ψ_{h^*} does.

In like manner, if A2 holds, then, using Proposition 3, we can reduce programs (3.6) and (3.7) to:

$$\left. \begin{array}{l} \max_{\psi \in C} \int_0^{h^*} \log(\psi(h) - C_0) d\mathbf{F}(h) \\ \text{s.t.} \\ 0 \leq \psi'(h) \leq k_2 \\ \int_0^\infty \psi(h) d\mathbf{F}(h) \leq k_1 \\ \log(\psi(h^*) - C_0) \geq \hat{y}, \end{array} \right\} \quad (3.6a)$$

and

$$\left. \begin{aligned}
 & \max_{\psi \in C} \int_{h^*}^{\infty} \log(\psi(h) - C_0) d\mathbf{F}(h) \\
 & \text{s.t.} \\
 & 0 \leq \psi'(h) \leq k_2 \\
 & \int_0^{\infty} \psi(h) d\mathbf{F}(h) \leq k_1 \\
 & \log(\psi(h^*) - C_0) \geq \hat{y}, \quad (3.73a)
 \end{aligned} \right\} \quad (3.7a)$$

where $\hat{y} = \frac{1}{1 + \gamma c} (y - \gamma c \log \gamma c)$.

Of course, the analogous result to Proposition 2 holds, that is:

Proposition 2a. Let $\hat{\Gamma} = \{(h^*, y) \in \mathbf{R}_+ \times \mathbf{R} \mid (3.6a) \text{ and } (3.7a) \text{ have solutions at which } (3.63a) \text{ and } (3.73a) \text{ bind}\}$. Define

$$\begin{aligned}
 \hat{y}^L(h^*) &= \log(\psi_{h^*}^L(h^*) - C_0) \\
 \hat{y}^R(h^*) &= \log(\psi_{h^*}^R(h^*) - C_0) \\
 \hat{y}^*(h^*) &= \log(\psi_{h^*}(h^*) - C_0)
 \end{aligned}$$

Then

$$\max[\hat{y}^L(h^*), \hat{y}^R(h^*)] \leq \hat{y} \leq \hat{y}^*(h^*) \quad (3.14a)$$

Conversely, if (3.14a) holds, then $(h^*, \hat{y}) \in \hat{\Gamma}$.

Proof: As in Proposition 2.

$\hat{\Gamma}$ is our parameterization of the set of quasi-PUNEs associated with the ‘reduced’ problem, where we have substituted out for r . From consideration of the three

programs (3.8), (3.9), and (3.10), it is clear that the interval of admissible values \hat{y} is non-empty for every h^* .

We next derive what the quasi-PUNE looks like at $(h^*, \hat{y}) \in \hat{\Gamma}$.

Proposition 5. Suppose A2 holds. Let $(h^, \hat{y}) \in \hat{\Gamma}$. Then:*

a. *The solution to (3.6a) is illustrated in Figure 2. It is defined by:*

$$\psi^L(h) = \begin{cases} \hat{\psi}_0^L & , \quad 0 \leq h \leq h_L \\ \hat{\psi}_0^L + k_2(h - h_L) & , \quad h_L \leq h \leq h^* \\ C_0 + e^{\hat{y}} & , \quad h > h^* \end{cases}$$

where $(\hat{\psi}_0^L, h_L)$ is the simultaneous solution of the two equations:

$$\log(\hat{\psi}_0^L + k_2(h^* - h_L) - C_0) = \hat{y}, \quad (3.15a)$$

$$\hat{\psi}_0^L + k_2 \int_{h_L}^{h^*} (h - h_L) d\mathbf{F}(h) + (1 - F(h^*))k_2(h^* - h_L) = k_1. \quad (3.15b)$$

In particular, $\hat{\psi}_0^L \geq C_0$.

b. *The solution of (3.7a) is illustrated in Figure 3. It is defined by:*

$$\psi^R(h) = \begin{cases} \hat{\psi}_0^R + k_2 h, & 0 \leq h \leq h_R \\ \hat{\psi}_0^R + k_2 h_R, & h > h_R \end{cases}$$

where $(\hat{\psi}_0^R, h_R)$ is the simultaneous solution of:

$$\log(\hat{\psi}_0^R + k_2 h^* - C_0) = \hat{y}, \quad (3.15c)$$

$$\hat{\psi}_0^R + k_2 \int_0^{h_R} h d\mathbf{F}(h) + (1 - F(h_R))k_2 h_R = k_1. \quad (3.15d)$$

In particular, $\psi_0^R \geq C_0$.

Proof. See appendix.

Again, we provide an intuitive argument for Proposition 5. First, consider program (3.6a); we argue that ψ^L is the solution. We know that $\psi^L(h^*) - C_0 = e^{\hat{y}}$, because constraint (3.63a) binds at a quasi-PUNE. Obviously, we must have $\psi^L(h) = C_0 + e^{\hat{y}}$ for $h > h^*$ because it is a waste to give resources to $[h^*, \infty)$. Now on $[0, h^*)$, the optimizer wants to pile on the resource as early as possible. She does so up to some value h_L , at which ψ scoots up as rapidly as possible to its value at h^* . Equation (3.15a) defines the value $\psi^L(h^*)$ and equation (3.15b) is the budget constraint.

The reader can provide the reasoning for part b.

Finally, we must verify that $\hat{\psi}_0^L$ and $\hat{\psi}_0^R$ are both at least C_0 . From figures 2 and 3, it suffices to show that $\hat{\psi}_0^R \geq C_0$. But it is clear, from consideration of programs (3.7a) and (3.9), that $\hat{\psi}_0^R \geq \psi_0^R$, since the additional constraint (3.73a) in (3.7a) only forces the constant term to be larger. Therefore the claim follows.

Proposition 4 and Proposition 5 completely characterize the manifold of quasi-PUNEs.

A quasi-PUNE at $(h^*, y) \in \Gamma$ is derived from a point $(h^*, \hat{y}) \in \hat{\Gamma}$ by defining $y = (1 + \gamma c)\hat{y} + \gamma c \log \gamma c$; the associated quasi-PUNE is given by $(\psi^L, \gamma c(\psi^L - C_0))$ and $(\psi^R, \gamma c(\psi^R - C_0))$ where ψ^L and ψ^R are the solutions of (3.6a) and (3.7a). We have one thing left to check: that every member of each party at least weakly favors her party's policy to the other party's policy. This claim is easy to verify. Indeed, by superimposing figure 3 upon figure 2 we have figure 4; we see that the after-tax income functions of the two parties coincide on the interval $[h_L, h^*]$ of types, and indeed, each member of a party weakly favors her party's policy to the other's. Of course, the two educational investment functions are just scaled down multiples of translations of the functions graphed in Figure 4. We illustrate them in Figure 5 for future reference.

The key fact needed in the analysis to follow is that if A2 holds, then at any quasi-PUNE, and therefore at any PUNEED (ψ, r) :

$$h > 0 \Rightarrow \psi(h) > C_0 \Rightarrow r(h) > 0.$$

C. PUNEED

We know every PUNEED is a quasi-PUNE. We now show that the set of PUNEEDs is non-empty. To do so, we compute the PUNEED where the militants of each party get to play their ideal policies.

Let $h^* = \mu$, and let $L = [0, h^*)$ and $R = [h^*, \infty)$. Then the ideal policies of the militants are the solutions of (3.8a) and (3.9a); thus they are:

$$\begin{aligned}\psi^L(h) &= k_1 \\ \psi^R(h) &= \psi_0^R + k_2 h = k_1 + k_2(h - \mu).\end{aligned}$$

The unique indifferent voter is

$$h = \mu = h^*.$$

In particular, $h < h^*$ implies h prefers ψ^L to ψ^R and $h > h^*$ prefers ψ^R to ψ^L . The militants in each party will assent to no deviation in policy. Thus (ψ^L, ψ^R, h^*) is a PUNEED.

Hence the set of PUNEED is non-empty. Indeed, other considerations lead us to believe that there is a 2-manifold of PUNEED, but that need not occupy us here.

§4 Democratic dynamics

We now imagine a sequence of overlapping generations, at dates $t = 0, 1, \dots$. The probability distribution of adult wages at date 0 is \mathbf{F}^0 . Political competition is organized over the questions of taxation and educational investment, and a PUNEED is realized, inducing a policy lottery. One party wins the election, and its educational investment policy is implemented, giving rise to a distribution of wages at date 1, \mathbf{F}^1 . This process

continues forever, inducing a sequence $\{\mathbf{F}^t\}$ of wage distributions. We are interested in the asymptotic distribution of wages.

Over time, it is not reasonable to suppose that C_0, m , and α remain constant. We therefore denote their values at date t by C_0^t, m^t , and α^t . Let μ^t be the average human capital at date t . We now assume the time-dated version of A2:

$$\text{A2}^t \quad \mu^t > C_0^t \text{ and } m^t < \frac{\mu^t - C_0^t}{\mu^t}.$$

The consequence of A2^t is that, at every PUNEEP (ψ^t, r^t) at date t , $r^t(h) > 0$ for $h > 0$.

The coefficient of variation of \mathbf{F}^t is

$$\eta^t = \int \left(\frac{h}{\mu^t} - 1\right)^2 d\mathbf{F}^t(h). \quad (4.1)$$

We are, in particular, interested in the limit of $\{\eta^t\}$. Does it exist, and if so, is it positive or zero?

We will work with an altered sequence of distributions, normalized to maintain the mean constantly at μ^0 . Define the distribution function

$$\hat{F}^t(h) = F^t\left(\frac{\mu^0}{\mu^t} h\right), \quad (4.2)$$

and let $\hat{\mathbf{F}}^t$ be the associated probability measure. Then the mean of $\hat{\mathbf{F}}^t$ is μ^0 . Since $\hat{\mathbf{F}}^t$ has the same coefficient of variation as \mathbf{F}^t , we will study the coefficients of variation of the sequence $\{\hat{\mathbf{F}}^t\}$.

Proposition 6. *If A1 and A2 hold, then the distribution function \hat{F}^{t+1} cuts the distribution function \hat{F}^t once from below. That is,*

$$(\exists h')(0 < h < h' \Rightarrow \hat{F}^{t+1}(h) < \hat{F}^t(h) \text{ and } h > h' \Rightarrow \hat{F}^{t+1}(h) > \hat{F}^t(h)).$$

Proof.

Let (Ψ, r) be the PUNEEP at date t . Since the mapping $h \rightarrow \alpha^t h^b r(h)^c$ is strictly monotone increasing, mothers and sons occupy the same ranks in their respective wage distributions, that is:

$$\forall h \quad F^{t+1}(\alpha^t h^b r(h)^c) = F^t(h) \quad (4.3)$$

Hence, from (4.2):

$$\forall h \quad \hat{F}^{t+1}\left(\frac{\mu^{t+1}}{\mu^0} \alpha^t h^b r(h)^c\right) = \hat{F}^t\left(\frac{\mu^t}{\mu^0} h\right) \quad (4.4)$$

Let $\theta: \mathbf{R}_+ \rightarrow \mathbf{R}_+$ be defined by:

$$\forall h \in \mathbf{R}_+ \quad \frac{\mu^t}{\mu^0} h \rightarrow \frac{\mu^{t+1}}{\mu^0} \alpha^t h^b r(h)^c.$$

Then we may rewrite (4.4) as $\hat{F}^{t+1}(h) = \hat{F}^t(\theta^{-1}(h))$, and so $\hat{F}^{t+1}(h) \underset{<}{\geq} \hat{F}^t(h)$ as

$$\hat{F}^t(\theta^{-1}(h)) \underset{<}{\geq} \hat{F}^t(h) \text{ as } \theta^{-1}(h) \underset{<}{\geq} h \text{ as } h \underset{<}{\geq} \theta(h) \text{ as } h \underset{<}{\geq} \frac{\mu^{t+1}}{\mu^t} \alpha^t h^b r(h)^c \text{ as } \frac{h^{1-b}}{r(h)^c} \underset{<}{\geq} \alpha^* \equiv \frac{\mu^{t+1}}{\mu^t} \alpha^t.$$

We next argue that the function $\zeta(h) = \frac{h^{1-b}}{r(h)^c}$ is strictly increasing on \mathbf{R}_+ , taking

on values from zero to “infinity,” which means that

$$(\exists h') | 0 \leq h < h' \Rightarrow \zeta(h) < \alpha^* \text{ and } h > h' \Rightarrow \zeta(h) > \alpha^*.$$

This will prove the proposition.

Suppose Left won the election at date t . The graph of $r(h)$ is pictured in Figure 5 (recall $r(h) = \gamma c(\psi(h) - C_0)$). Obviously $\zeta(h)$ is strictly increasing on the intervals $[0, h_L)$ and $[h^*, \infty)$, where r is constant. On the interval $[h_L, h^*]$, we have

$r(h) = \beta_0 + \gamma ck_2 h$, where $\beta_0 \geq 0$. (See figure 4: $\beta_0 \geq 0$ because $\psi_0^R(0) \geq C_0$.) Therefore on this interval

$$\zeta(h) = \frac{h^{1-b}}{(\beta_0 + \gamma ck_2 h)^c}.$$

Therefore we have

$$\frac{d}{dh} \log \zeta(h) = \frac{1-b}{h} - \frac{\gamma c^2 k_2}{(\beta_0 + \gamma ck_2 h)}$$

and so

$$\frac{d}{dh} \log \zeta(h) > 0 \Leftrightarrow \frac{1-b}{c} > \frac{\gamma ck_2 h}{\beta_0 + \gamma ck_2 h}. \quad (4.5)$$

Since $\beta_0 \geq 0$, the r.h.s. of the last inequality is no larger than unity, and hence $\zeta(h)$ is strictly increasing on $[h_L, h^*]$ if $\frac{1-b}{c} > 1$. But this last inequality is guaranteed by A1.

Now suppose that Right won the election at date t . Again consult figure 5.

Exactly, the same kind of argument shows that ζ is strictly increasing. ■

Since the sequence $\{\hat{\mathbf{F}}^t\}$ is mean-preserving and $\hat{\mathbf{F}}^{t+1}$ cuts $\hat{\mathbf{F}}^t$ once from below, we have that $\hat{\mathbf{F}}^{t+1}$ second-order stochastically dominates $\hat{\mathbf{F}}^t$. In particular, we know

$$\forall x > 0 \quad \int_0^x \hat{\mathbf{F}}^{t+1}(h) dh < \int_0^x \hat{\mathbf{F}}^t(h) dh. \quad (4.6)$$

Proposition 7. The sequence $\{\hat{\mathbf{F}}^t\}$ converges weakly to a distribution function, denoted

$\hat{\mathbf{F}}$.

Proof:

Define $\Phi^t(x) = \int_0^x \hat{F}^t(h) dh$. From (4.6), $\Phi^t(x)$ is strictly decreasing in t , for any $x > 0$. Therefore, $\lim_{t \rightarrow \infty} \Phi^t(x)$ exists: call the limit $\Phi(x)$. We claim \hat{F}^t converges weakly to $\Phi'(x)$ (denoted $\hat{F}^t \Rightarrow \Phi'(x)$). We know $\Phi(x)$ is an increasing function, so its derivative exists almost everywhere.

Because $\Phi^{t'}(x) = \hat{F}^t(x)$, it suffices to show that

$$\forall x > 0 \quad \lim_t \Phi^{t'}(x) = \Phi'(x) \quad , \quad (4.7)$$

or

$$\lim_{t \rightarrow \infty} \lim_{\delta \rightarrow 0} \frac{\Phi^t(x + \delta) - \Phi^t(x)}{\delta} = \lim_{\delta \rightarrow 0} \frac{\Phi(x + \delta) - \Phi(x)}{\delta}.$$

But we know

$$\lim_{\delta \rightarrow 0} \lim_{t \rightarrow \infty} \frac{\Phi^t(x + \delta) - \Phi^t(x)}{\delta} = \lim_{\delta \rightarrow 0} \frac{\Phi(x + \delta) - \Phi(x)}{\delta}, \quad (4.8)$$

and the limits on the l.h.s. of (4.8) commute; we are done. ■

Lemma 2. Let A2 hold. Let $r^L(r^R)$ be the investment function of the Left (Right) at any quasi-PUNE. Let $y = sh + \beta$ be the line containing any chord of the graph of $r^L(r^R)$.

Then s and β are non-negative.

Proof. Inspect Figure 5. ■

Let E be the distribution function with all its mass at the point μ^0 ; that is

$$E(h) = \begin{cases} 0 & \text{if } h < \mu^0 \\ 1 & \text{if } h \geq \mu^0 \end{cases}.$$

We have:

Theorem 1. If A1 and A2, then $\hat{F} = E$.

Proof:

We have observed that each ‘dynasty’ occupies a fixed rank in the mean-normalized distributions \hat{F}^t , across time, regardless of which parties win the various elections.

To prove the claim, it suffices to show that for any $\varepsilon, \delta \in (0,1)$,

$(\exists t^*)(t > t^* \Rightarrow 1 - \varepsilon$ of the mass of $\hat{\mathbf{F}}^t$ lies on an interval of length at most δ).

Given small, positive ε and δ , define h_1 and h_2 by

$$F^0(h_1) = \frac{\varepsilon}{2}, \quad F^0(h_2) = 1 - \frac{\varepsilon}{2}.$$

Let h_1^t, h_2^t be the descendants of h_1 and h_2 at date t . Whether Left or Right wins at date t , we have

$$\frac{h_2^{t+1}}{h_1^{t+1}} \leq \left(\frac{s^t h_2^t + \beta^t}{s^t h_1^t + \beta^t} \right)^c \left(\frac{h_2^t}{h_1^t} \right)^b \quad (4.9)$$

where $y = s^t h + \beta^t$ is the equation of the chord containing $r^J(h_1^t)$ and $r^J(h_2^t)$ (for $J = L$ or R) in the realized PUNEED at date t .

But since $\beta^t \geq 0$ and $s^t \geq 0$ always (Lemma 2), we have:

$$\frac{s^t h_2^t + \beta^t}{s^t h_1^t + \beta^t} \leq \frac{h_2^t}{h_1^t},$$

and hence, from (4.9),

$$\frac{h_2^{t+1}}{h_1^{t+1}} \leq \left(\frac{h_2^t}{h_1^t} \right)^{b+c},$$

and so by recursion

$$\frac{h_2^{t+1}}{h_1^{t+1}} \leq \left(\frac{h_2}{h_1} \right)^{(b+c)^t};$$

since $b + c < 1$, it follows that $\frac{h_2^{t+1}}{h_1^{t+1}} \rightarrow 1$. In particular, for any small positive number

η there exists t^* such that $t > t^* \Rightarrow \frac{h_2^t}{h_1^t} < 1 + \eta$; since ranks of dynasties are preserved

across time, this means that fraction $1 - \varepsilon$ of the mass of the distributions $\hat{\mathbf{F}}^t$ lies between the mean-normalized variants of h_1^t and h_2^t .

Denote these mean-normalized variants by \hat{h}_1^t and \hat{h}_2^t . We claim $\{\hat{h}_1^t\}$ are bounded above by some number M . For suppose this were false and that for any $M > 0$, we eventually have $\hat{h}_1^t > M$ and therefore $\hat{h}_2^t > M$.

Then $\mu^0 > \int_{\hat{h}_1^t}^{\hat{h}_2^t} h d\hat{\mathbf{F}}^t(h) = (1 - \varepsilon)M' > (1 - \varepsilon)M$, where M' is some number in the interval $[\hat{h}_1^t, \hat{h}_2^t]$, a contradiction.

Now let $M = \sup_t h_1^t$. Then the length of the interval $[\hat{h}_1^t, (1 + \eta)\hat{h}_1^t]$, which is $\eta\hat{h}_1^t$ is less than ηM . By choosing $\eta < \frac{\delta}{M}$, we therefore show that $1 - \varepsilon$ of the mass of $\hat{\mathbf{F}}^t$ lies on an interval of length at most δ . ■

Finally let us contrast the dynamics of democracy with the dynamics of laissez-faire, a system in which there is no taxation. In this case, each household partitions its income, h , optimally between consumption and educational investment.

If $C_0 = 0$, then $r(h) = \frac{\gamma c}{1 + \gamma c} h$, and so the human capital of the son is

$$h' = \alpha \left(\frac{\gamma c}{1 + \gamma c} \right)^c h^{b+c}.$$

If A1 holds, then the coefficient of variation of the distribution of human capital approaches zero. Democracy will speed up the convergence, but convergence to equality will occur, absent democracy.

If $C_0 > 0$, then $r(h) = \text{Max}[\frac{\gamma c}{1 + \gamma c}(h - C_0), 0]$. We examine the special case in

which, for all t , $C'_0 = C_0$ and $\alpha' = \alpha$. We compute that the son's human capital is greater than the mother's if and only if

$$\alpha \left(\frac{\gamma c}{1 + \gamma c} \right)^c > \frac{h^{1-b}}{(h - C_0)^c}. \quad (4.10)$$

The function on the r.h.s. of (4.10) approaches infinity asymptotically near C_0 and as h approaches infinity, as we illustrate in Figure 6. As long as α is sufficiently large, there will be two values, denoted h_1 and h_2 , where (4.10) is an equality. The dynamics are illustrated in Figure 6. Fraction $1 - \mathbf{F}^0(h_1)$ of the population end up eventually with zero human capital, and the rest end up with human capital h_2 . Thus, without democracy, we have, asymptotically, a highly polarized society.

§5 Topping Off

We have assumed until now that educational funding is purely public. But winning publicly financed education has been itself a significant victory of democracy. So it would have been more realistic to begin with the supposition that education could be privately or publicly financed.

First, note that at any PUNEEP, under our assumptions, no household will desire to top off public education with additional private education, because every PUNEEP partitions the household's total resource bundle just as the optimizing household would. So there will be no demand for further private education at these equilibria.

Now suppose that it is not assumed, initially, that education will be publicly financed. Thus, a party may propose a policy (ψ, r) assuming that citizens will top off privately, if $r < \gamma c(\psi - C_0)$. Thus, the h -household facing the policy $(\psi(h), r(h))$ solves for its private educational investment, which we denote $r^P(h)$:

$$\text{Max}_{r^P} \log(\psi(h) - r^P - C_0) + \gamma c \log(r(h) + r^P).$$

The solution is

$$r^P(h) = \frac{\gamma c(\psi(h) - C_0) - r(h)}{1 + \gamma c},$$

assuming that $r(h) < \gamma c(\psi(h) - C_0)$. Let $z(h) = \psi(h) + r(h)$. Then

$$\begin{aligned} \psi(h) - r^P(h) - C_0 &= \frac{z(h) - C_0}{1 + \gamma c}, \text{ and} \\ r^P(h) + r(h) &= \frac{\gamma c(z(h) - C_0)}{1 + \gamma c}. \end{aligned}$$

Without loss of generality, we may therefore write the household's indirect utility function as

$$v(\psi, r; h) = \log(z(h) - C_0).$$

So we may write the program of the Left party (for instance) at h^* as:

$$\begin{aligned}
& \underset{z(\cdot)}{\text{Max}} \int_0^{h^*} \log(z(h) - C_0) d\mathbf{F}(h) \\
& \text{s.t. } 0 \leq z'(h) \leq m \\
& \int_0^{\infty} z(h) d\mathbf{F}(h) \leq \mu \\
& \log(z(h^*) - C_0) \geq y^*.
\end{aligned}$$

It is thus clear that the set of PUNEEDs where private financing of education is not precluded is isomorphic to the set of PUNEEDs where only public financing is possible. It is a matter of indifference whether education is publicly funded or whether households finance some or all education privately: the children receive identical educational investments in both cases. The key to convergence to wage equality is that A1 and A2 hold.

§6 *Random talent*

In this section, we alter the determination of the child's future wage to admit a random factor: we assume that, if h is the wage of the mother and h' the wage of the son, then

$$h' = \varepsilon \alpha h^b r^c \tag{6.1}$$

where ε is a random variable. We shall assume that ε is lognormally distributed with a mean of unity; thus $\log \varepsilon$ is normally distributed with mean $\frac{-\sigma}{2}$ and variance σ^2 .

If we assume that a parent knows her child's 'talent,' ε , before the election occurs, then the politics are identical to what we have analyzed already: ε enters simply as a positive constant on the child's wage, which does not affect any parent's preferences

over policies. Thus our analysis of PUNEEDs stands. The limit distribution of wages, however, will change.

Let $h_2 > h_1$ be two wages at date 0. It is of interest to calculate the probability that $h_2^t < h_1^t$, that is, the probability that the ranks of the t^{th} descendants will have reversed.

We know, according to the proof of Theorem 1, that

$$\frac{h_2^1}{h_1^1} \leq \frac{\varepsilon_2^1}{\varepsilon_1^1} \left(\frac{h_2}{h_1} \right)^{b+c}$$

where ε_i^1 is the talent realization of the child of h_i , for $i = 1, 2$. Recursively, we have:

$$\frac{h_2^t}{h_1^t} \leq \frac{\prod_{\tau=1}^t (\varepsilon_2^\tau)^{(b+c)^{t-\tau}}}{\prod_{\tau=1}^t (\varepsilon_1^\tau)^{(b+c)^{t-\tau}}} \left(\frac{h_2}{h_1} \right)^{(b+c)^t},$$

where ε_i^τ is the talent realization of the τ^{th} descendant of h_i , for $i = 1, 2$. Taking

logarithms:

$$\log \frac{h_2^t}{h_1^t} \leq \sum_{\tau=1}^t (b+c)^{t-\tau} \log \varepsilon_2^\tau - \sum_{\tau=1}^t (b+c)^{t-\tau} \log \varepsilon_1^\tau + (b+c)^t \log \frac{h_2}{h_1}. \quad (6.2)$$

Define $\log \frac{h_2^t}{h_1^t}$ as the random variable R^t . Define Z^t to be the random variable

$$\sum_{\tau=1}^t (b+c)^{t-\tau} X_2^\tau - \sum_{\tau=1}^t (b+c)^{t-\tau} X_1^\tau$$

where X_1^τ and X_2^τ are i.i.d. random variables which are $N(-\frac{\sigma}{2}, \sigma)$. Let $\rho = \frac{h_2}{h_1}$. Then (6.2)

implies that

$$\Pr[R^t \leq m] \geq \Pr[Z^t \leq m - (b+c)^t \log \rho] \quad (6.3)$$

and so, in particular,

$$\Pr[R^t \leq 0] \geq \Pr[Z^t \leq -(b+c)^t \log \rho].$$

Now Z^t is a normally distributed random variable with mean zero and

variance $2\sigma^2 \sum_{\tau=1}^t (b+c)^{2(t-\tau)} = 2\sigma^2 \left(\frac{1-(b+c)^{2t}}{1-(b+c)^2} \right)$. So for large t , Z^t is approximately

normally distributed with mean zero and variance $2\sigma^2/(1-(b+c)^2) \equiv \sigma_\infty^2$.

If, following Bénabou (in press), we take $\sigma = 1$ and $b+c = 0.7$, then $\sigma_\infty^* = 1.980$.

Then from (6.3), we see that the probability that $\frac{h_2^t}{h_1^t} \leq q$, for large t , is at least

(approximately) the probability that $N(0, \sigma_\infty)$ is less than $\log q$. For the values of

σ and $(b+c)$ given above, we compute these probabilities for various values of q :

q	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Pr	.519	.537	.553	.567	.581	.594	.606	.617	.627	.637

§7 Discussion

We have shown, that under assumptions A1 and A2, democracy leads asymptotically to equalization of wages, in the precise sense that the coefficient of variation of the distribution of human capital approaches zero. A1 is an assumption about the educational technology, which we assumed to hold timelessly. A2 comprises two assumptions; first, that $\mu^t > C_0^t$, which says that the society is sufficiently developed that a family enjoying the mean income would rather send its child to school than to work; and second, that $m^t \leq \frac{\mu^t - C_0^t}{\mu^t}$, which says that it is politically unacceptable for a political party to propose a policy where the total resource bundle

(that is, after-tax income plus educational investment in the child) rises at too rapid a rate. We might suppose that μ' increases considerably more rapidly than C'_0 with time, and so eventually the constraint simply says that $m^t < 1$. Now a society with $m^t > 1$ would seem to be quite regressive, and so the 'social norm' embodied in A2 will eventually be quite mild. So we can loosely summarize our analysis as saying that, once a democracy reaches a sufficiently high level of economic development, convergence to equality of human capital is assured.

We reiterate that we have demonstrated this result with essentially no other restrictions on the policy space than the two social norms. In this sense, the political competition is quite unrestricted.

Why have we not observed more rapid convergence to equality of wages in advanced democracies, then? Three answers can be suggested. First, there is a random talent or effort element, as discussed in section 6, in wage determination. Our analysis says that, with such a stochastic element, we will not reach equality of wages, but rather a situation in which the wages of distant descendants of the first ancestors are independent of the wages of those ancestors. Second, there are technological shocks that change the distribution of human capital. The worsening inequality of incomes in the US and UK in the past twenty years is to be understood as a change in the wage-earning capacities of individuals due to non-neutral technological change or increased competition. Our model has assumed these kinds of shock away. We assumed only neutral technical change in allowing the technological parameter α to vary with time. Third, democracy is not very

old, and assumption A2 perhaps only holds in a small number of countries, and perhaps only recently. Without income taxation, which is to say under *laissez-faire*, we have that the rate of increase of the total resource bundle with respect to h is unity, and A2 tells us we can only expect convergence to equality, in this case, if $C_0 = 0$. Surely, A2 has held, if indeed it does hold now, in most advanced democracies, for less than a century.

In the United States, funding for public education of h -households does increase with h : this is accomplished through the linking of educational finance with the local property tax base. In the political equilibria of our model, this is the case—that is, $r(h)$ is non-decreasing in h , and increasing in h in a region. In many European countries, equal public educational investment in children of all backgrounds is closer to the truth. Of course, section 5 tells us that, in these equilibria, parents will top off the public investment in their children. I conjecture that, at least in the Nordic countries, this does not occur. We may understand this as the consequence of the operation of another social norm – not one we have modeled here. There is, however, an alternative explanation, that the educational of other people’s children is a public good. Whatever it is that forces the investment of education in children not to increase with the human capital of the parents will, of course, lead more rapidly to convergence to equality, given A1.

Theorem 1 has an implication for contemporary debates in democratic theory. Democratic theorists are divided into two groups, according to whether they define democracy in a minimalist or maximalist fashion. The minimalist view (see, for example, Przeworski et al (2000)) is that democracy is best conceived as a system with political competition between parties, *tout court*. The maximalist version frequently goes by the

name of deliberative democracy (see, for instance, Elster (1998)); here democracy requires as well as political competition, a thorough-going discussion among citizens – a forum – at which citizens convince each other to take account of their mutual needs. Maximalists tend to think that political competition alone will not suffice to bring about a decent society (read: equality or justice).

Our analysis tends to *support* this conclusion, in the sense that we have no political explanation for the constraint that

$$z'(h) \leq \frac{\mu - C_0}{\mu}.$$

Of course, we have not shown that if this constraint fails, then democracy will *not* bring convergence to equality. All our analysis permits us to say is that, if it fails, then there are quasi-PUNEs at which a positive mass of children will receive zero educational investment. (Perhaps, that is to say, a more delicate analysis could lead us to eliminate any PUNEES that invest zero in a positive mass of children.) Thus, at this point, it is attractive to say that the deliberative aspect of democracy induces a social norm that assures that the required constraint on the rate of increase of total resources holds.

Appendix: Proof of Propositions 3 and 5

The proof of all the parts of these propositions is the same. I will prove only Proposition 5 part a, as an illustration of the technique.

Let ψ^L be the function defined in Prop.5, part a. We know that the optimal solution must coincide with ψ^L on $[h^*, \infty)$; the only issue is to prove that ψ^L is indeed the solution on $[0, h^*)$.

Suppose ψ^L were not optimal. Let $\psi^L + g$ be the true solution, where g is a variation defined on $[0, h^*]$. We must have $g(h^*) = 0$ because we know ψ^L takes the right value at h^* .

Proposition B1. *Suppose $\psi^L + g$ is in T and $g(h^*) = 0$. Suppose there exists a non-negative function $\lambda: [h_L, h^*] \rightarrow \mathbf{R}_+$ and a number $\delta > 0$ such that the function*

$$\Delta(\varepsilon) := \int_0^{h^*} \log(\psi^L(h) + \varepsilon g(h) - C_0) d\mathbf{F}(h) + \int_{h_L}^{h^*} \lambda(h) (k_2 - (\psi^L(h) + \varepsilon g(h))') dh + \delta \left(K - \int_0^{h^*} (\psi^L(h) + \varepsilon g(h)) d\mathbf{F}(h) \right)$$

is maximized at $\varepsilon = 0$, where $K = \int_0^{h^} \psi^L(h) d\mathbf{F}(h)$. Then the value of program (3.6a) at ψ^L*

is at least as great as its value at $\psi^L + g$.

Proof:

1. Note that at $\varepsilon=0$ the second and third terms in the definition of $\Delta(\varepsilon)$ disappear, because

$$\psi^{L'} = k_2 \text{ on } [h_L, h^*] \text{ and } K = \int_0^{h^*} \psi^L d\mathbf{F}. \text{ Therefore } \Delta(0) = \int_0^{h^*} \log(\psi^L(h) - C_0) d\mathbf{F}(h), \text{ which}$$

is the value of program (3.6a) at ψ^L .

2. Feasibility of $\psi^L + g$ implies that

$$(\psi^L(h) + g(h))' \leq k_2 \text{ on } [h_L, h^*]$$

and
$$\int_0^{h^*} (\psi^L(h) + g(h)) d\mathbf{F}(h) \leq K.$$

Therefore at $\varepsilon=1$, the second and third terms in expression $\Delta(1)$ are non-negative, because of the non-negativity of λ and δ .

3. But $\Delta(0) \geq \Delta(1)$. It follows that

$$\int_0^{h^*} \log(\psi^L - C_0) d\mathbf{F} \geq \int_0^{h^*} \log(\psi^L + g - C_0) d\mathbf{F}. \quad \blacksquare$$

Proposition B2. *Define:*

$$\delta = \frac{1}{\hat{\psi}_0^L - C_0}, \text{ and}$$

$$\lambda(h) = \int_{h_L}^h \left(\frac{f(s)}{\hat{\psi}_0^L - C_0} - \frac{f(s)}{\psi - C_0} \right) ds$$

on $[h_L, h^*]$, where f is the density function of \mathbf{F} . Then $\delta \geq 0, \lambda(h) \geq 0$ on $[h_L, h^*]$, and Λ is maximized at $\varepsilon=0$.

Of course, Propositions B1 and B2 together prove that ψ^L is the solution of program (3.6a).

Proof of Prop. B2:

1. We have shown in the text that $\delta \geq 0$.

2. Because $\psi^L(h) > \hat{\psi}_0^L$ for $h \in (h_L, h^*]$, it follows that

$$\lambda'(h) = \frac{f(h)}{\hat{\psi}_0^L - C_0} - \frac{f(h)}{\psi^L(h) - C_0} > 0$$

on the interval. Since $\lambda(h_L) = 0$ by definition, it follows that $\lambda \geq 0$ on $[h_L, h^*]$.

3. Note that Δ is a concave function. Thus to show that Δ is maximized at zero, it

suffices to show that zero is a local maximum of Δ . We proceed to demonstrate this

by showing that $\frac{d\Delta}{d\varepsilon}(0) = 0$.

$$\begin{aligned} \frac{d\Delta}{d\varepsilon}(0) &= \int_0^{h^*} \frac{g(h)}{\psi^L(h) - C_0} d\mathbf{F}(h) - \int_{h_L}^{h^*} \lambda(h) g'(h) dh - \delta \int_0^{h^*} g(h) d\mathbf{F}(h) \\ 4. &= \delta \int_0^{h_L} g(h) d\mathbf{F}(h) + \int_{h_L}^{h^*} \frac{g(h)}{\psi^L(h) - C_0} d\mathbf{F}(h) - \lambda(h) g(h) \Big|_{h_L}^{h^*} + \int_{h_L}^{h^*} \lambda'(h) g(h) dh - \\ &\quad \delta \int_0^{h_L} g(h) d\mathbf{F}(h) - \delta \int_{h_L}^{h^*} g(h) d\mathbf{F}(h). \end{aligned}$$

The key step in the above expansion is the use of integration by parts to integrate the

$\lambda(h)g'(h)$ term. Now, noting that $\lambda(h_L) = 0$ and $g(h^*) = 0$, and substituting for $\lambda'(h)$,

the latest expression reduces to

$$\begin{aligned} &\int_{h_L}^{h^*} \frac{g(h)}{\psi^L(h) - C_0} d\mathbf{F}(h) + \int_{h_L}^{h^*} \left(\frac{1}{\hat{\psi}_0^L - C_0} - \frac{1}{\psi^L(h) - C_0} \right) g(h) d\mathbf{F}(h) \\ &- \delta \int_{h_L}^{h^*} g(h) d\mathbf{F}(h) = 0, \end{aligned}$$

as was to be shown. ■

Remark. The reader who wishes to employ this technique to prove the other parts of

Propositions 3 and 5, should he or she not be convinced by the intuitive arguments

presented in the text, should write down the Lagrangian function $\Delta(\varepsilon)$ with the unknown

parameters $\lambda(\cdot)$ and δ and carry out the integration by parts. It will be immediately clear how λ and δ are to be defined. Of course, the technique only works when one 'knows' the solution to the program.

References

Bénabou, R. (in press) "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?",

Econometrica

Elster, J. (ed.) 1998. *Deliberative Democracy*, New York: Cambridge University Press

Przeworski, A., M.E. Alvarez, J.A. Cheibub, and F. Limongi, 2000. *Democracy and Development*, New York: Cambridge University Press

Roemer, J.E. 1998. "Why the poor do not expropriate the rich: An old argument in new garb," *Journal of Public Economics* 70 (3) 399-424

-- 1999. "The democratic political economy of progressive taxation," *Econometrica*, 67, 1-19.

-- 2001. *Political Competition: Theory and Applications*, Cambridge, Mass.: Harvard University Press

Figure 1 The solution to program (3.10a)

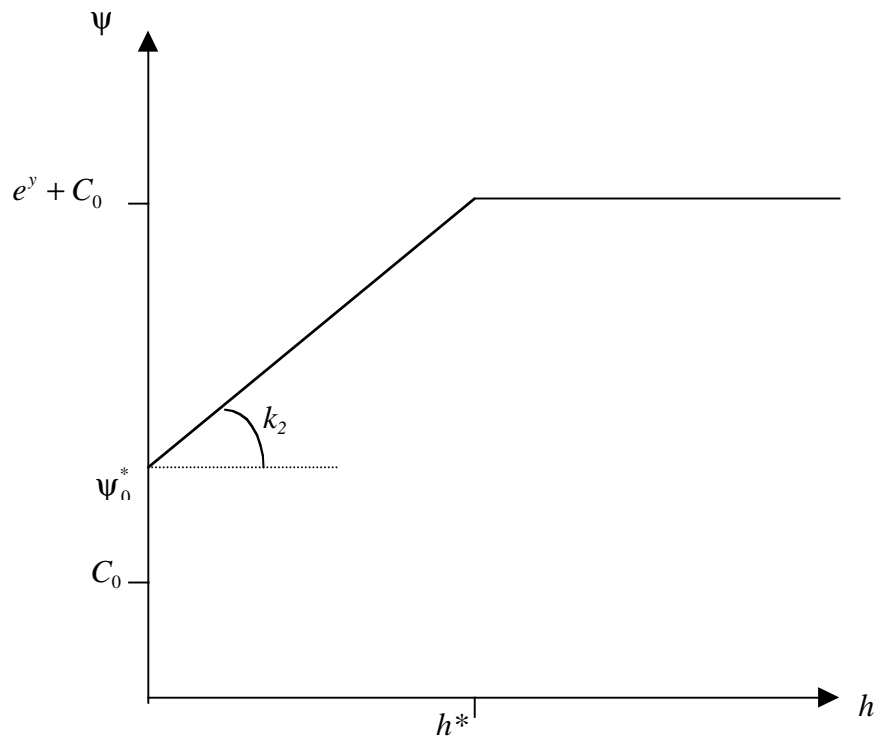


Figure 2 The solution to (3.6a)

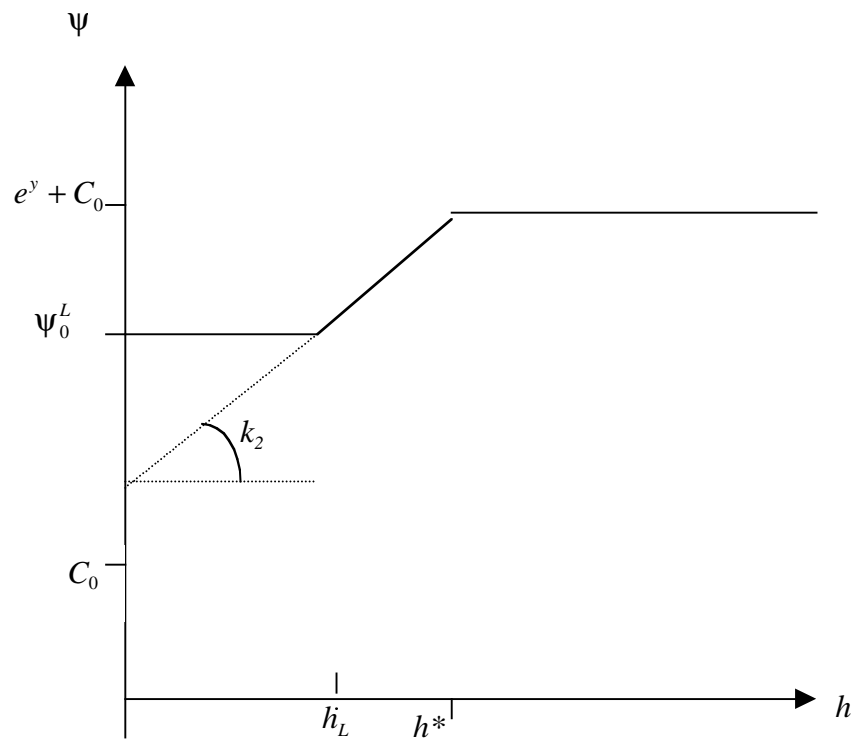


Figure 3 The solution of program (3.7a)

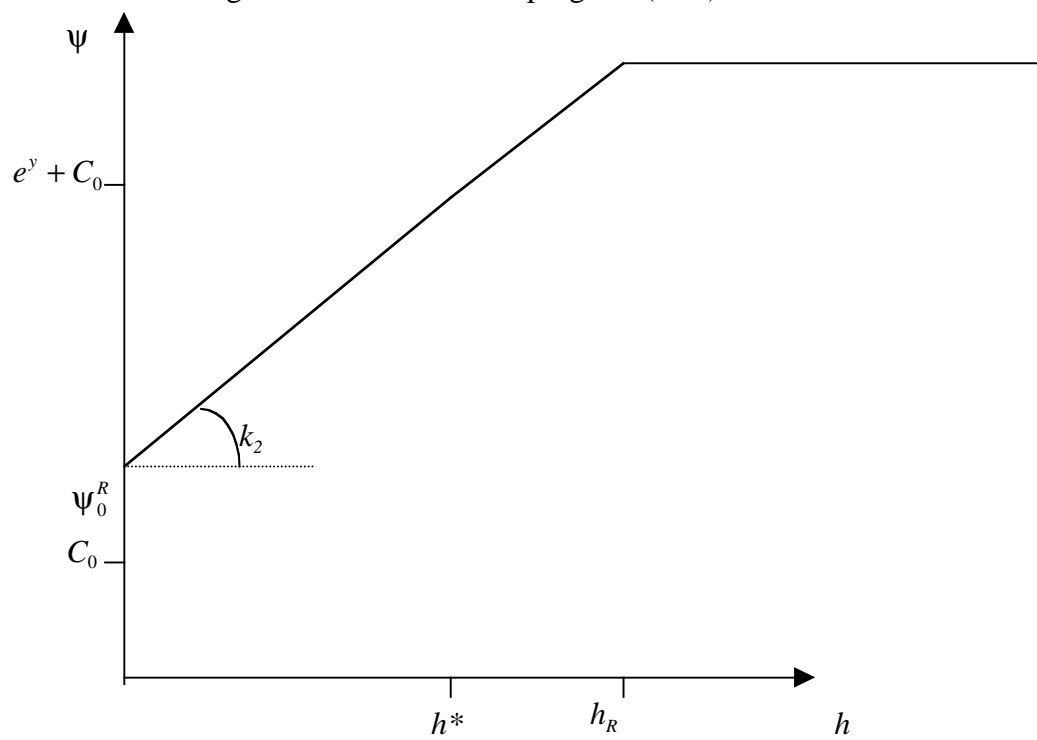


Figure 4 Left (bold) and Right policies in a PUNEEP

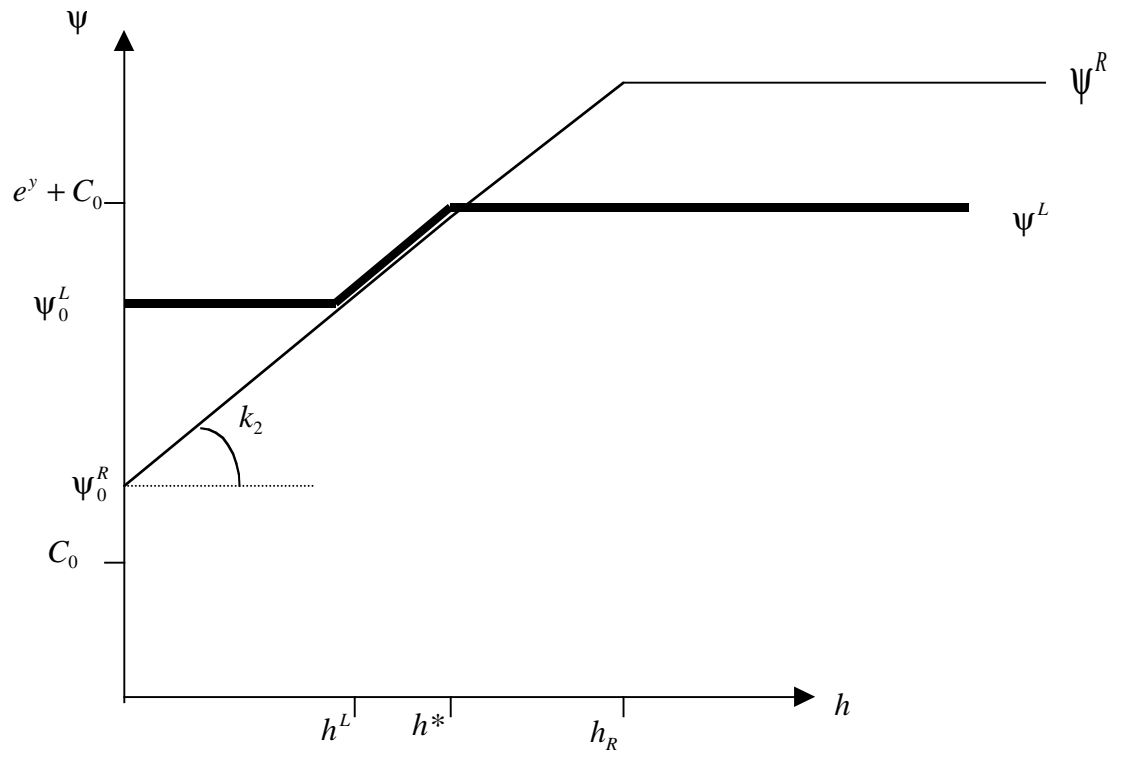


Figure 5 Left (bold) and Right educational investments in a PUNEEP

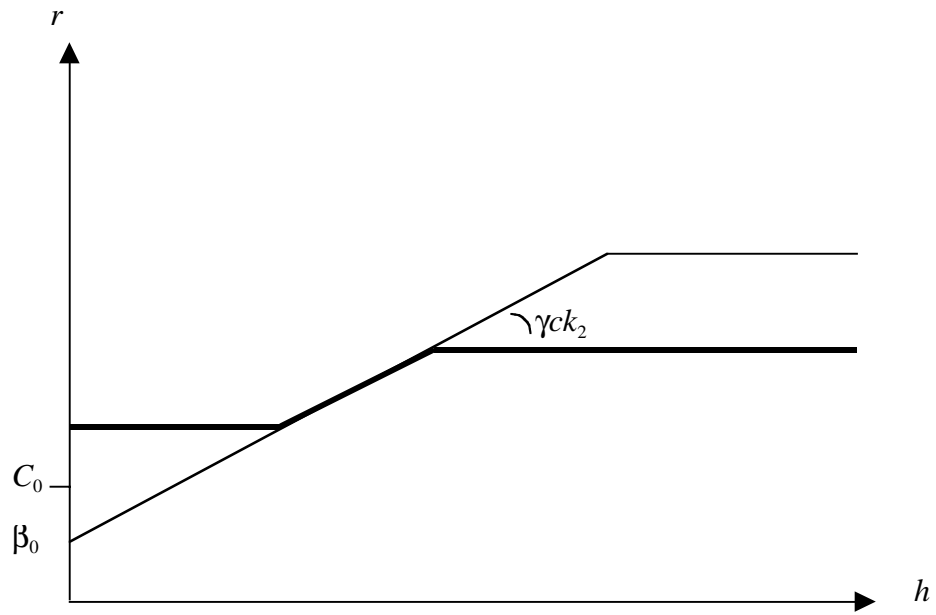


Figure 6 Dynamics under laissez-faire when $C_0 > 0$

