INCOME INEQUALITY AND EDUCATION POLICY WHEN APTITUDE IS INNATE: PICKING A POINT ON A KUZNETS CURVE

Peter Kennedy
Department of Economics
University of Victoria

6 January 2017

ABSTRACT

I characterize the educational choices made by workers in terms of their innate aptitudes and how closely those aptitudes match available job types. A worker can choose to specialize in a narrow field or pursue a more general education based on the closeness of that match. Workers fortunate enough to have an innate aptitude closely matched to a job type will become highly specialized, and will reap high market rewards. Less fortunate workers will be less well-rewarded. Thus, there arises a distribution of income in the economy. The mean, variance and skewness of that distribution is a function of education policy, relating in particular to the educational opportunities made available, and to how they are funded, whether through income tax or tuition charges. I have cast that policy choice in terms of picking a point on a Kuznets curve, which dictates the relationship between average income and income-equality in this economy.

1. INTRODUCTION

It seems increasingly common for parents and policy-makers to want all young citizens to receive a university education. This is often seen as a gateway to higher incomes, and to lower inequality because it levels the playing field in terms of career opportunities. This argument clearly has some merit. Historically in most countries, a university education was available only to the children of the wealthy. Not only did this deny the economy of the productive potential of gifted children who happened to be born poor, it also perpetuated the inequality of income that bestowed better educational opportunities on the children of the rich.

But have the benefits of university-education-for-all been oversold? While an advanced education provides a significant boost to the incomes of those who study in a field for which there is high market demand, the same cannot be said for those who study in fields that are less well-rewarded. One might argue that there must be some unmeasured compensating differential for the latter group, or else they would have chosen to study in the more lucrative field. Accordingly, if welfare is properly measured, then all students benefit from more education, no matter what their field of study. However, this argument does not account for differences in aptitude. Many students might wish to study computer engineering for the income reward it offers but they simply do not have the aptitude for it, so they study English literature instead.

Arguments that rest on differences in innate aptitude are not fashionable in some circles. There are many who believe that any person can achieve any goal if given the opportunity. The evidence suggests otherwise.

Where these arguments falter is when they conflate aptitude with notions of worthiness or merit. The student with a gift for computer engineering received a lucky draw from the aptitude lottery. That does not make her more worthy or deserving of a university education. It nonetheless might make her a more valuable recipient of educational investment for society.

Questions about optimal investments in educational might matter less if the costs of a university education were purely private, but they are not. Tuition is often heavily subsidized, with much of the cost paid for out of taxes. This can sometimes be justified on the basis of the positive externalities associated with education, but subsidized tuition is often motivated as much by redistributional goals as it is by efficiency goals. Again, there is merit to that policy. If the redistribution of wealth is a goal for policy-makers then pursuing it through the provision of subsidized educational opportunities can be a highly effective mechanism. However, if that subsidy is extended equally to all students, regardless of their innate aptitude, then the redistributional impact can be heavily diluted. Conversely, if more and cheaper education is provided to students whose market returns from that education will be modest, the tax revenues available to finance it will be similarly modest. The cost of providing highly subsidized but low-reward education may simply be too high.

The policy problem can be usefully couched in terms of a tradeoff between creating wealth, by providing educational opportunities to those with aptitudes that generate high returns, and spreading that wealth by providing subsidized opportunities for all. In this paper I cast that problem in terms of picking a point on a Kuznets curve.

I construct a model in which workers, who start out as students, have heterogeneous innate aptitudes. Some of those workers are fortunate enough to have an aptitude closely matched to an available job type, and so enjoy a high market reward for that aptitude. Others are less fortunate, and have aptitudes that are less well-rewarded in the market. As students, these workers face a decision about how specialized to become in terms of their human capital. Specialized workers are more productive in their jobs, but there is a productivity penalty for specializing in a field of study for which they do not have an aptitude. As students, the workers face a choice between studying what they are good at, and studying what will pay well after graduation. That choice leads to a highly specialized education for some, and a very general education for others.

In the context of these choices made by workers, the policy-maker decides how many educational hours should be provided, and how this should be funded, as a mix of tuition and a tax-funded subsidy.

The model is simple enough to yield an analytical solution for the Kuznets curve that arises in this economy. It dictates the relationship between the Gini coefficient (calculated from after-tax incomes) and average income in the economy. It has an inverted U-shape. I cast the education policy problem in terms of picking a point on this curve. I also identify the point on the Kuznets curve that voting would select in this economy, with full voter-turnout, and with turnout only by relatively high-income individuals. These outcomes can then be matched to the implicit weights that a policy-maker would place on average income versus income equality in choosing an educational policy.

This paper relates to a very wide span of existing literature, ranging from job-skill matching, to returns to education, to university funding, to income equality cross households. It is not my intention here to try to cover all of these connections to the literature. Instead, I direct readers to a paper that covers most of these issues, and their connections to key papers in the literature, with a level of knowledge and authority beyond anything I could usefully provide. That paper is Altonji, Arcidiacono and Maurel (2016).

The rest of this paper is organized as follows. Section 2 presents the model of job types and heterogeneous innate aptitudes. Section 3 characterizes worker choices over training, in terms of field of study and the degree of specialization in that field. Section 4 then derives the distribution of income that arises from those choices and the underlying distribution of aptitudes. Section 5 relates educational policy to the moments of that income distribution. Section 6 then casts the education policy in terms of picking a point on the Kuznets curve. Section 7 derives the median-voter outcome for education policy. Section 8 concludes.

2. THE MODEL

Consider an economy in which workers are distinguished by their innate aptitude, *a*. These workers are distributed uniformly around an *aptitude circle* of circumference *C*. There are *n* skilled-job types distributed along this circle, each one located equidistant from its closest neighbors. Thus, the distance between job-types is $c = C/n$. To avoid potential confusion over the location of a worker relative to a job-type, I specify an arbitrary but fixed direction of measurement. In particular, job-types are ordered sequentially, $\tau_1 < \tau_2, ..., < \tau_n$, in a clockwise direction. This means that $\tau_i = \tau_{i-1} + c$. See Figure 1 for an example with $n = 4$. (Ignore the triangle in that figure for the moment).

Workers must obtain an education prior to entering the workforce. As students, they make two key choices: their field of study, *x*; and the extent to which they specialize within that field, z. A field of study corresponds to a point on the aptitude circle, around which training is centered. Specialization relates to how narrowly the education is focused on that centre point. For example, an arts education could focus on European history but with varying degrees of depth versus breadth. A highly specialized program might involve multiple courses on the minutiae of 16th century Europe, and little else. A much less specialized program might involve three broad courses in European history supplemented with courses in comparative history, political science, economics and statistics.

Each student has a budget of h instructional hours, determined by state policy. Those hours must be split between depth and breadth according to the triangular relationship depicted in Figure 1. The breadth of the education, centered on field x, is denoted b. The depth of the education is the extent of specialization, z. The area of this education triangle is bz, and this area must equal h. Thus, $b = h / z$.

The cost of the education is shared between the student and the state. The student pays tuition πh, where $\pi \geq 0$ is henceforth called the price of instruction, and this payment is deferred until she is working, when it is tax-deductible. The residual aggregate cost to the state is financed by a tax on income at rate t. Both π and t are the same for all individuals, regardless of their educational choices or their eventual job-types.

I assume that labour is supplied inelastically. While an elastic labor supply would allow us to introduce distortions associated with an income tax, it complicates the interpretation of income inequality because the consumption and value of leisure varies across workers. Moreover, most of the empirical work on inequality, and the popular debate on income inequality, focuses on income, not on the more comprehensive notion of “full income” which incorporates the value of leisure. Accordingly, my model abstracts from this distinction but in consequence it also abstracts from the distorting effect of the income tax on the labor-leisure choice.

A worker receives an income proportional to her productivity. The productivity of a worker whose aptitude happens to be perfectly matched to a job type, and who chooses her field to match that job type, is increasing in z, at a diminishing rate. I specify a logarithmic form for this relationship. In particular, if $a = \tau = x$ then income is simply proportional to $\log(z)$. This worker will choose to specialize as narrowly on τ as is practically possible.

A less fortunate worker, whose aptitude is not perfectly matched to a job type, faces two key tradeoffs in her choice over depth versus breadth. First, a high degree of specialization makes a worker highly skilled in a very narrow set of tasks but leaves her less skilled at tasks that are more distant from the focal point of her education. Thus, a lack of breadth imposes a *skill-mismatch penalty* (*SMP*) if the worker does not work exactly in her field of specialization. I specify the SMP for a worker who has trained in field *x* with breadth *b*, and works in job-type τ, as

(1) $$SMP = \frac{\delta(x-\tau)^2}{b}$$

where $\delta \geq 0$ captures the strength of the *SMP*. Thus, an income penalty occurs if $x \neq \tau$ but that penalty can be reduced if the worker has lots of breadth in her training.

The second tradeoff relates to a potential mismatch between field of study and aptitude. Consider a worker whose aptitude does not match perfectly with an available job type. For example, a history student may have great talent for the study of 16^{th} century Europe but jobs in that field are few and far between. So rather than pursue his true interest single-mindedly, he broadens his program to include courses in economics and statistics. However, these courses do not come naturally; he finds them difficult to master. In consequence, he suffers an *aptitude-mismatch penalty* (*AMP*). This penalty rises with the extent to which he specializes in the mismatched field; more breadth in his training helps to bridge the gap between his field and his aptitude by introducing connecting threads and common concepts. Accordingly, I specify the AMP for a worker with aptitude *a* who has trained in field *x* with breadth *b* as

(2) $$AMP = \frac{\theta(x-a)^2}{b}$$

where $\theta \geq 0$ captures the strength of the AMP.

I assume that the SMP and the AMP are additive. Thus, setting $b = h/z$ gives us the following expression for after-tax income

(3) $$y_\tau(a,x,z) = (1-t)\left(\beta\left(\log(z) - \frac{z}{h}(\delta(x-\tau)^2 + \theta(x-a)^2)\right) - \pi h\right)$$

where $\beta > 0$ is an exogenous technological parameter that determines the productivity of labor in this economy.

3. EDUCATION CHOICES

I assume that a student chooses her education to maximize her future after-tax income; there is no field-specific "consumption value" from study that might otherwise incline her to study a field close to her aptitude. Thus, if a student plans to work in job-type *i* then her field and extent of specialization are chosen to maximize (3). These choices are

(4) $$\hat{x}_i(a) = \frac{\delta\tau_i + \theta a}{\delta + \theta}$$

and

(5) $$\hat{z}_i(a) = \frac{h(\delta + \theta)}{\delta\theta(a - \tau_i)^2}$$

respectively. Expression (4) tells us that a worker's chosen field is a weighted average of her aptitude and the job-type in which she plans to work. This choice reflects the fact that matching skill to job-type is not the only consideration when it comes to educational choices; aptitude – and its implications for the capacity to acquire skills in a given area – are also important. The relative weight given to these two considerations is determined by the relative strengths of the skill-mismatch penalty and the aptitude-mismatch penalty respectively, as reflected in the relative size of δ and θ. Only in the limiting case where the worker happens to have an aptitude perfectly suited to an available job-type ($a = \tau_i$) does the chosen field match perfectly with the job type.

Expression (5) tells us that the chosen extent of specialization is decreasing in the distance between aptitude and job-type. This choice also reflects the joint effects of the aptitude-mismatch and skill-mismatch penalties. The worker chooses a field that is not perfectly matched to either her aptitude or her planned job type, and this leads her to add generality to her education rather than specialize narrowly in a field with no jobs or in a field to which she is not well-suited.

Substituting $\hat{x}_i(a)$ and $\hat{z}_i(a)$ into $y_\tau(a, x, z)$ yields income as a function of aptitude and instructional hours:

(6) $$\hat{y}_i(a,h) = (1-t)\left(\beta\left(\log(\hat{z}_i(a)) - 1\right) - \pi h\right)$$

This tells us that workers who specialize earn more than those with more general skills. This does *not* mean that specialization is a good strategy for all workers. Specialists earn more than generalists because of their good fortune in possessing aptitudes that are closely-matched to available job types; it is this close matching that makes specialization lucrative. Generalists rationally choose not to specialize despite the fact that empirical evidence from this economy would show that specialists do better.

I can now use (6) to partition workers on the aptitude circle into those who train towards job-type *i* rather than job-type $i-1$ or $i+1$. In particular, a worker with aptitude *a* chooses job type *i* if and only if $\hat{y}_i(a,h) \geq \max\{\hat{y}_{i-1}(a,h), \hat{y}_{i+1}(a,h)\}$. This means that a worker with aptitude *a* chooses job-type *i* if and only if

(7) $$a \in \left[\tau_i - \frac{c}{2}, \tau_i + \frac{c}{2}\right]$$

That is, each worker chooses the job-type that is closest to her aptitude. I henceforth refer to (7) as the i^{th} arc of the aptitude circle; workers located in that arc, by definition, choose job-type *i*.

The lowest-income workers in arc *i* are those for whom $a = \tau_i - c/2$ or $a = \tau_i + c/2$. These are the workers on the lower and upper boundaries of the arc respectively. Income for these workers is

(8) $$y_L = (1-t)\left(\beta\left(\log(\varphi(h)) - 1\right) - \pi h\right) \quad \forall i$$

where

(9) $$\varphi(h) = \frac{4h(\delta+\theta)}{\delta\theta c^2}$$

I henceforth assume that the state is constitutionally bound to ensure that $y_L > 0$. We will see later that this places a restriction on feasible policies.

The symmetry of job-type locations around the aptitude circle means that the distributions of educational choices and associated incomes within any arc are the same for all arcs. This means that the analysis can be couched in terms of a representative arc, and drop the *i* subscript from all variables. I denote the aptitudes of workers at the lower and upper boundaries of this representative arc as a_{BL} and a_{BH} respectively.

4. THE DISTRIBUTION OF INCOME AND MEASURES OF INEQUALITY

In the Appendix I show that the distribution of income within a representative arc – and hence, for the economy as whole – is an exponential distribution, with cumulative density given by

$$(10) \qquad F(y) = 1 - \left(\frac{\varphi(h)}{\exp\left(1 + \frac{y + \pi h(1-t)}{\beta(1-t)}\right)} \right)^{\frac{1}{2}}$$

on support $[y_L, \infty)$. The mean of this distribution is

$$(11) \qquad \mu = \frac{2\beta(1-t)}{\psi(h)}$$

where

$$(12) \qquad \psi(h) = \frac{2\beta}{\beta\left(1 + \log(\varphi(h))\right) - \pi h}$$

This term $\psi(h)$ will appear repeatedly throughout the analysis, and we can think of policy-changes to *h* and π in terms of changes to $\psi(h)$. Note that $\psi(h) > 0$ when $y_L > 0$. The variance of the distribution is $\sigma^2 = 4\beta^2(1-t)^2$, and the median income is

$$(13) \qquad y_M = (1-t)\left(\beta\left(\log(4\varphi(h)) - 1\right) - \pi h\right)$$

It is straightforward to show that $\mu > y_M$; thus, the income distribution is positively-skewed.

We can distinguish those workers who earn above-average income from those who earn below-average income by setting $\hat{y}_i(a,h) = \mu$ and solving for *a*. This yields

$$(14) \qquad \bar{a}_L = \tau - \frac{c}{2\mathbf{e}}$$

and

(15) $$\bar{a}_H = \tau + \frac{c}{2\mathbf{e}}$$

Thus, workers within the representative arc for whom $a \notin (\bar{a}_L, \bar{a}_H)$ have below-average incomes, while workers for whom $a \in (\bar{a}_L, \bar{a}_H)$ have above-average incomes. Note that $\bar{a}_L$ and $\bar{a}_H$ are both independent of the policy parameters.

We can identify similar thresholds with respect to the median income. In particular, setting $\hat{y}_i(a,h) = y_M$ and solving for *a* yields

(16) $$a_{ML} = \tau - \frac{c}{4}$$

and

(17) $$a_{MH} = \tau + \frac{c}{4}$$

Workers within the representative arc for whom $a \notin (a_{ML}, a_{MH})$ have incomes below the median income, while workers for whom $a \in (a_{ML}, a_{MH})$ have incomes above the median income. Note that this partitioning of the representative arc around a_{ML} and a_{MH} simply divides the arc into two equal parts (because the distribution of *a* is uniform on that arc).

From (10) we can derive the Lorenz curve for this economy (see the Appendix). In particular, let $L(p)$ denote the fraction of total income earned by the lowest-paid fraction *p* of workers. Then

(18) $$L(p) = p + (1-p)\log(1-p)\psi(h)$$

Perfect income equality yields a Lorenz curve where $L(p) = p$. Thus, $p - L(p)$ captures the extent to which this economy exhibits income inequality. The maximum deviation of $L(p)$ from *p* occurs at

(19) $$\bar{p} = 1 - \mathbf{e}^{-1} > \tfrac{1}{2}$$

This critical value measures the fraction of workers who earn less than the average income; that is, $F^{-1}(\bar{p}) = \mu$.

From the Lorenz curve we can calculate the income share going to the k^{th} "*q*-tile" of workers as $s_q(k) = L(k) - L(k - 1/q)$, where $k \in [0,1]$. For example, the income share of the top 1% of workers is found by setting $q = 100$ and $k = 0.99$. Similarly, the income share of the second quintile of workers is found by setting $q = 5$ and $k = 0.40$. In general,

(20) $$s_q(k) = \frac{1}{q} - \psi(h)\left((1-k+1/q)\log(1-k+1/q) - (1-k)\log(1-k) \right)$$

This income-share function exhibits a special relationship with the parameters of the model. In particular, the solution to

(21) $$\lim_{q\to\infty} \frac{\partial s_q(k)}{\partial \psi(h)} = 0$$

is $k = \overline{p}$ for any non-zero value of $\psi(h)$. This tells us that $s_q(k)$ pivots around the point $\{\overline{p}, s_q(\overline{p})\}$ in response to a change in $\psi(h)$ as $q \to \infty$. This effect is illustrated in Figure 2, which plots $s_q(k)$ against *k* for two different values of $\psi(h)$, labeled ψ_1 and $\psi_2 > \psi_1$. Recall that $\overline{p}$ measures the fraction of workers who earn less than the average income. Thus, an increase (decrease) in $\psi(h)$ causes all workers with below-average income to lose (gain) income share, and all workers with above-average income to gain (lose) income share. Figure 2 also illustrates that these gains and losses are most pronounced in the tails of the income distribution. In particular,

(22) $$\frac{\partial^2 s_q(k)}{\partial \psi \partial k} = \frac{\psi(h)}{(1-k)(1+q-qk)} > 0$$

This tells us that the largest share-losses (gains) from an increase (decrease) in $\psi(h)$ are incurred by the lowest income earners, while the highest share-gains (losses) are captured by the highest income earners. Moreover, the derivative in (22) is itself increasing in *k*; the very highest income earners gain (lose) disproportionately from an increase (decrease) in $\psi(h)$ relative to the other high-income earners below them on the income scale.

While $s_q(k)$ is perhaps the most informative indicator of income inequality, the Gini coefficient is a more convenient summary measure, and it is the measure on which I focus in the remaining sections. The Gini coefficient is calculated as twice the area under $p - L(p)$, and in this economy it takes a very simple form:

$$G(h) = 1 - 2\int_0^1 L(p)dp = \frac{\psi(h)}{2} \tag{23}$$

Thus, any increase (decrease) in $\psi(h)$ causes income inequality to rise (fall).

5. EDUCATION POLICY

I assume that instructional hours can be provided at constant marginal cost, ρ. Thus, the *per capita* cost of providing *h* hours to every student is ρh. On the other side of the state's ledger, *per capita* revenue from income tax and tuition is

$$r = t\left(\frac{\mu}{1-t}\right) + \pi h \tag{24}$$

where μ is given by (11). A balanced budget requires $r = \rho h$, and the tax rate needed to ensure this is

$$t(h) = \frac{(\rho - \pi)\psi(h)h}{2\beta} \tag{25}$$

I assume that the policy-maker is bound by its governing constitution to meet this balanced-budget requirement.

I impose four additional constitutional constraints on the policy-maker. First, the income of the poorest worker cannot be negative. This restricts π to be no greater than

$$\pi_H(h) = \frac{\beta\left(\log(\varphi(h) - 1\right)}{h} \tag{26}$$

Second, π cannot exceed ρ (or equivalently, the income tax rate cannot be negative). Thus, we require $\pi \le \min\{\rho, \pi_H(h)\}$. Third, π cannot be negative. In combination with $\pi \le \pi_H(h)$, this means that *h* cannot be less than

$$h_L = \frac{\beta \mathrm{e}}{\rho\zeta} \tag{27}$$

where

$$\zeta = \frac{4\beta(\delta + \theta)}{\rho\delta\theta c^2} \tag{28}$$

Fourth, the policy-maker cannot set $t(h) > 1$. This constraint never binds if $\pi = \rho$ but for any $\pi < \rho$ it means that *h* cannot exceed

(29) $$h_H = -\frac{\beta}{\rho} W_{-1}\left(\frac{-1}{\zeta \mathrm{e}}\right)$$

where W_{-1} is the lower branch of the Lambert W function.[1] Note that h_H is independent of π, because tuition is tax deductible.

Which of the two upper bounds on π is binding? This depends on ζ. In particular, if $\zeta < \mathbf{e}^2$, then $\min\{\rho, \pi_H(h)\} = \pi_H(h) \;\; \forall h \in [h_L, h_H]$. This tells us that if ζ is small (because ρ is large relative to β, for example) then π must always be set lower than ρ; tuition *must* be subsidized to ensure non-negative incomes for all workers. Conversely, if $\zeta > \mathrm{e}^2$ (because ρ is small relative to β) then unsubsidized tuition can be charged over some range of *h*. These restrictions on the set of $\{h, \pi\}$ pairs that are available to the policy-maker will be important for the inequality-implications of its policy choices, and they are summarized in Figure 3, for two different values of ζ. The lightly-shaded region represents the feasible set when $\zeta > \mathrm{e}^2$ ($\rho = \rho'$ small); the entire shaded region represents the feasible set when $\zeta < \mathrm{e}^2$ ($\rho = \rho''$ large).

I now derive *per capita* income and the Gini coefficient under a balanced-budget policy. Setting $t = t(h)$ in (11) and (23) yields

(30) $$\mu(h) = \beta\Big(1 + \log(\varphi(h))\Big) - \rho h$$

and

(31) $$G(h, \pi) = \frac{\beta}{\beta\Big(1 + \log(\varphi(h))\Big) - \pi h}$$

respectively.

Consider first the relationship between *G* and *h*. At any given value of $\pi > 0$, $G(h, \pi)$ reaches a minimum at $h = \beta / \pi$. At $h < \beta / \pi$, an increase in *h* causes inequality

[1] The Lambert W function (sometimes called the omega function or the product logarithm), denoted $W(u)$, is the inverse of $u = W\exp(W)$. It has two real branches defined on $u \geq -1/\mathbf{e}$, and is double-valued on $(-1/\mathbf{e}, 0)$. The lower branch has $W \leq -1$ and the upper branch has $W \geq -1$.

to fall. Why? The increase in h allows all workers to specialize more – fields and breadth of study are unchanged – and this raises incomes for all workers. However, the relative impact on income is smallest for the highest-income workers because these workers are the most specialized, and there are diminishing returns to specialization. Thus, income inequality falls. As h rises further, this equalizing effect of more education is eventually outweighed by the offsetting effect of rising tuition (when $\pi > 0$), which is independent of income and therefore constitutes a larger relative burden on lower-income workers. The smaller is π, the larger is the value of h at which inequality starts to rise with h because the tax-financed component of education spending is proportional to income. If $\pi = 0$ then inequality never rises with h; more instructional hours continues to reduce inequality, albeit at a diminishing rate.

Next consider the relationship between μ and h. It is straightforward to show that $\mu(h)$ reaches a maximum at

$$h^* \equiv \frac{\beta}{\rho} \tag{32}$$

This income-maximizing education level equates the marginal cost of instruction hours, measured by ρ, with the value-marginal product of instruction hours, as determined by the productivity parameter β. This value-marginal product declines with h because there are diminishing returns to specialization, so *per capita* income eventually reaches a turning point beyond which more instructional hours causes income to fall.

Note that h^* is independent of π. The price of tuition has no incentive effects on either field choices or specialization levels because it is not contingent on income. It therefore has no impact on gross incomes, and hence, no impact on *per capita* income for the economy as a whole, and no impact on the education level that maximizes that *per capita* income.[2]

Suppose for the moment that the policy-maker is concerned first and foremost with *per capita* income, and that h is set equal to h^*. Then *per capita* income is

[2] I should stress the importance of two assumptions at this point. Recall that the state is obligated to ensure that h is affordable for all workers (income for the poorest worker cannot be negative), and that tuition is paid only once a worker has an income. If instead there are some workers who cannot afford to purchase h instructional hours then the price of tuition must affect the income distribution. We will later see that $\pi = 0$ in the median-voter outcome, so affordability constraints would never actually bind anyway.

(33) $$\mu^* \equiv \mu(h^*) = \beta \log(\zeta)$$

and the Gini coefficient is

(34) $$G(h^*, \pi) = \frac{1}{\log(\zeta) + 1 - \frac{\pi}{\rho}}$$

Note that $G(h^*, \pi)$ is increasing in π. Thus, the policy-maker can then choose π to achieve its income-inequality goals, but *only within certain limits*. The first limitation stems from the $\pi < \pi_H(h)$ requirement. If $\zeta < e^2$ (because ρ is large relative to β) then $\pi_H(h^*) < \rho$. In this case the policy-maker cannot implement the income-maximizing policy unless tuition is subsidized. Thus, the policy-maker may be forced to use a redistributive policy that reduces inequality incidentally even if income-maximization is its only goal. To put this another way, income-maximization and inequality-reduction are to some degree complementary polices if the marginal cost of education is sufficiently high relative to overall productivity in the economy.

The second limitation on the policy-maker stems from the $\pi \geq 0$ requirement. This puts a lower bound on $G(h^*, \pi)$, denoted

(35) $$G_0^* \equiv G(h^*)\Big|_{\pi = 0} = \frac{1}{\log(\zeta) + 1}$$

If the policy-maker wishes to achieve a level of inequality lower than G_0^* then it cannot do so using tuition pricing alone. To reduce inequality below G_0^*, the policy-maker must set $h > h^*$, and this requires giving up some *per capita* income. In the next section I cast this policy choice as picking a point on a Kuznets curve.

6. PICKING A POINT ON A KUZNETS CURVE

A Kuznets curve for this economy is a relationship between *per capita* income and the Gini coefficient, constructed from equations (30) and (31). There is a different Kuznets curve for each feasible value of π but at this point I confine consideration to the curve associated with $\pi = 0$; a policy-maker concerned about inequality and *per capita* income will only resort to setting $h > h^*$ (and hence, $\mu < \mu^*$) if there is no room left for reducing

π. Setting $\pi = 0$ in (31) and taking its inverse yields the level of education required to achieve a given level of inequality *G* when $\pi = 0$:

(36) $$h_0^{-1}(G) = \frac{\beta e^{\left(\frac{1}{G}-1\right)}}{\rho\zeta}$$

where the "0" subscript indicates that $\pi = 0$. Substituting $h = h_0^{-1}(G)$ and $\pi = 0$ in (6) then tells us the relationship between a policy-induced level of inequality and the income for an individual worker with aptitude *a*:

(37) $$y_0(a,G) = \beta(1-\lambda(G))\left(\log\left(\frac{c^2}{4(\tau-a)^2}\right)+\frac{1}{G}-2\right)$$

where

(38) $$\lambda(G) = \frac{\mathbf{e}^{\left(\frac{1}{G}-1\right)}G}{\zeta}$$

Figure 4 plots this relationship for two different workers: a poorest worker (with $a = a_{BL}$) and a median-income worker (with $a = a_{ML}$). The arrows indicate the direction in which the underlying value of *h* is increasing. The heavy curve labeled $y_0(G)$ is the locus of the maxima of $y_0(a,G)$ with respect to *G* for each value of *a*; I will return to this curve in section 7.

We can now derive the Kuznets curve when $\pi = 0$ by substituting $h_0^{-1}(G)$ for h in $\mu(h)$ from (30) to yield

(39) $$\mu_0(G) = \mu(h_0^{-1}(G)) = \frac{\beta(1-\lambda(G))}{G}$$

This relationship between *per capita* income and the Gini coefficient is illustrated in the upper quadrant of Figure 5. (Ignore the lower quadrant for the moment). The arrows along the curve indicate the direction in which the underlying value of *h* is increasing. When $h < h^*$, an increase in *h* causes *per capita* income to rise and income-inequality to fall; there is no trade-off between the two metrics. At $h = h^*$, *per capita* income is maximized and any further increase in *h* involves an increasingly costly trade-off: income-inequality continues to fall but per *capita income* falls rapidly until $\mu = 0$ at

$h = h_L$. Note that there is a discontinuity at $\mu = 0$ because $G(h_H - \varepsilon) > 0$ for $\varepsilon > 0$ arbitrarily small (because incomes are close to zero but not identical) but $G_0(h_H) = 0$ and $\mu_0(h_H) = 0$ because *all* incomes are zero at $h = h_L$ when $\pi = 0$.

Now suppose that the goal of the policy-maker is to maximize a combination of *per capita* income and income equality, given by

(40) $$U(h) = \mu(h) - \omega G(h)$$

where $\omega \geq 0$. Let $h(\omega)$ denote the solution to $\partial U(h)/\partial h = 0$. A closed-form solution exists for the inverse of $h(\omega)$, and it is given by

(41) $$\omega^{-1}(h) = (\rho h - \beta)\left(\log(\varphi(h)) + 1\right)^2$$

If $\omega = 0$ then the policy-maker puts exclusive weight on *per capita* income, and $\omega^{-1}(h) = 0$ yields the income-maximizing solution: $h^* = \beta / \rho$. Conversely, if $\omega > 0$ then $h(\omega) > \beta / \rho$, and $h(\omega)$ is increasing and strictly concave in ω.

Now let us position $h(\omega)$ on the Kuznets curve. Setting $h = h_0^{-1}(G)$ in (41) yields the optimal policy in terms of *G*:

(42) $$\omega^{-1}(G) = \frac{\beta(\lambda(G) - G)}{G^3}$$

This relationship between ω and the optimal level of inequality is depicted in the lower quadrant of Figure 5. At any value of ω, such as ω' in the figure, $\omega^{-1}(G)$ identifies the associated optimal level of inequality, $G(\omega')$, and from the Kuznets curve we can then identify the associated optimal level of *per capita* income, $\mu(\omega')$. Note that $G = G(h^*)$ and $\mu = \mu(h^*)$ at $\omega^{-1}(G) = 0$ because when $\omega = 0$ the optimal policy maximizes *per capita* income. Note too that for any $\omega > \omega_H$ the optimal policy is corner-constrained at $h = h_H$.

7. VOTING OVER POLICY OPTIONS

I now ask whether an optimal policy like that in (41) can be derived as the solution to a voting problem in which worker-voters care only about their own individual incomes. I assume that the policy options offered to voters are limited by the same constraints I imposed upon the policy-maker in section 5, as summarized in Figure 3. Specifically, voters are offered a choice over $h \in [h_L, h_H]$ and $\pi \in [0, \min\{\rho, \pi(h)\}]$, as described in (26) – (29).

Consider voter preferences over these policy options. Substituting $t = t(h)$ from (25) into (6) yields income for a worker with aptitude a on a representative arc as a function of h and π in a balanced-budget setting:

(43) $$y(a,h,\pi) = \beta\left(1 - \frac{(\rho-\pi)\psi(h)h}{2\beta}\right)\left(\log\left(\frac{(\delta+\theta)h}{\delta\theta(\tau-a)^2 e}\right) - \frac{\pi h}{\beta}\right)$$

Differentiation of $y(a,h,\pi)$ with respect to π then yields a partitioning of the representative arc in terms of how π affects income. In particular, $\forall \pi \in [0,\rho]$ and $\forall h \in [h_L, h_H]$,

(44) $$\frac{\partial y(a,h,\pi)}{\partial \pi} > 0 \text{ for } a \in (\overline{a}_L, \overline{a}_H)$$

and

(45) $$\frac{\partial y(a,h,\pi)}{\partial \pi} < 0 \text{ for } a \in (a_0, \overline{a}_L) \text{ and for } a \in (\overline{a}_H, a_1)$$

This tells us that among the policy choices on offer, a worker on a representative arc with an aptitude inside the interval $(\overline{a}_L, \overline{a}_H)$ prefers the largest value of π on offer; that is, their preferred policy is $\pi = \min\{\rho, \pi(h)\}$. Conversely, a worker with an aptitude outside the interval $(\overline{a}_L, \overline{a}_H)$ prefers the lowest value of π on offer; their preferred policy is $\pi = 0$.[3] These preferences are independent of *h*.

If all workers vote over π then the preferences of the median voter will determine the outcome. The median voter here is the worker with the median income, and we know that $y_M < \mu$ in this economy. Thus, the median-voter outcome is $\pi = 0$.

Now consider preferences over *h* when $\pi = 0$. Let $h(a)$ denote the solution to

[3] Note the link between this result and our earlier discussion of income shares in Section 4. In particular, recall that from Figure 2 that $s_q(k)$ pivots around the point $\{\overline{p}, s_q(\overline{p})\}$ in response to a change in $\psi(h)$ as $q \to \infty$, where $F^{-1}(\overline{p}) = \mu$. Since $\psi(h)$ increases monotonically with π, it follows that any reduction in π will always benefit those for whom $y < \mu$ and hurt those for whom $y > \mu$.

(46) $$\left(\frac{\partial y(a,h,\pi)}{\partial h}\bigg|_{\pi=0}\right)=0$$

There is a closed-form solution for the inverse of $h(a)$, and it has two branches, given by

(47) $$a_L^{-1}(h)=\tau-\left(\frac{(\delta+\theta)h}{\delta\theta e^{\gamma(h)}}\right)^{\frac{1}{2}}$$

and

(48) $$a_H^{-1}(h)=\tau+\left(\frac{(\delta+\theta)h}{\delta\theta e^{\gamma(h)}}\right)^{\frac{1}{2}}$$

where

(49) $$\gamma(h)=\frac{\beta(\log(\varphi(h)+1)^2-\rho h}{\rho h\log(\varphi(h))}$$

These relationships between aptitude and the preferred value of *h* are depicted in Figure 6. The figure highlights three key features of these relationships. First, workers with aptitudes increasingly distant from their chosen job type prefer increasingly higher *h*. This reflects the fact that the returns to *h* are diminishing, and that *h* is financed here exclusively through the income tax (since $\pi=0$). This means that the benefits of raising *h* accrue increasingly to those with more distant aptitudes (and hence, low incomes) while the costs are borne primarily by high-income earners.

Second, workers with $a\in(\bar{a}_L,\bar{a}_H)$ prefer $h<h^*$, while workers within a representative arc with $a\notin(\bar{a}_L,\bar{a}_H)$ prefer $h>h^*$. Thus, only those workers who earn exactly the average income prefer a policy that maximizes that average income; all other workers prefer a policy under which the average income is not maximized.

The third key feature of $h(a)$ follows from the second: the median voter prefers a value of *h* less than h^*. In particular, let h_M be the solution to $a_L^{-1}(h_M)=a_{ML}$, or equivalently, $a_H^{-1}(h_M)=a_{MH}$. Then $h_M<h^*$. This tells us that the median-voter outcome is one in which *per capita* income is not maximized.

Voting over Points on a Kuznets Curve

Voting over *h* when $\pi = 0$ can be thought of as voting over points on the Kuznets curve in (39). For this purpose, it is helpful to cast preferences over *h* in terms of preferences over *G*. Recall from (37) that $y_0(a,G)$ tells us the income for a worker with aptitude *a* when $\pi = 0$ and $h = h_0^{-1}(G)$. Setting $h = h_0^{-1}(G)$ into (47) and (48), and then making these substitutions for *a* in $y_0(a,G)$ yields a relationship between income and the preferred value of *G* for a worker with that income:

(50) $$y_0(G) = \frac{\beta(1-\lambda(G))^2}{\lambda(G)(1-G)}$$

This is the heavy curve in Figure 4. It is the locus of the maxima of $y_0(a,G)$ with respect to *G* for each value of *a*, and its inverse, denoted $G_0^{-1}(y)$, tells us the level of inequality preferred by a worker with income *y*. Thus, the level of inequality preferred by the median voter is $G_M \equiv G_0^{-1}(y_M)$.

Figure 7 overlays $y_0(G)$ on the Kuznets curve. Recall that the Kuznets curve plots *per capita* income against *G*, while $y_0(G)$ plots individual income against *G*. The vertical axis nonetheless has the same unit of measure for both curves, so the two curves are directly comparable in the figure. Note that $y_0(G)$ crosses $\mu_0(G)$ at $\{G^*, \mu^*\}$; the level of inequality preferred by a worker with the *average* income is necessarily the level of inequality that maximizes average income.

Figure 7 also highlights the level of inequality preferred by the median-income worker, G_M. This value of *G* corresponds to the intersection of $y_0(G)$ with $y_0(a_{LM}, G)$, reproduced from Figure 4. The level of *per capita* income corresponding to G_M on the Kuznets curve is μ_M, at the point labeled M. This point is preferred by the median voter to any other point on the Kuznets curve.

I can now tie these results back to the question I posed at the beginning of this section: can (41) reflect a voting outcome when individuals care only about their own absolute incomes? Clearly it can. The lower quadrant of Figure 7 reproduces the lower quadrant from Figure 5, and identifies a value of ω, denoted ω_M, that would implement the solution to (41) as the median-voter outcome in this economy. There does not exist a closed-form analytical solution for ω_M, but it is intuitively clear that $\omega_M > 0$.

Figure 7 also identifies one additional point of interest on the Kuznets curve, labeled *L*. This is the point preferred by the lowest-income worker, characterized by the intersection of $y_0(G)$ with $y_0(a_{BL}, G)$. Any point on the Kuznets curve to the left of point *L* is Pareto-dominated by *L*. This in turn places an upper bound on the weight, labeled ω_L in the lower frame of Figure 7, that a planner concerned with Pareto-efficiency would ever assign to inequality in the planning problem from (40). Points on the Kuznets curve to the right of point *L* cannot be Pareto-ranked.

Implications for a Cross-Country Kuznets Curve

Now suppose that countries differ according to their productivity, as measured by β, but are identical in all other respects. What is the relationship between *per capita income* and inequality across countries if policy in each country is determined by the median voter in that country?

Setting $a = a_{ML}$ in (46) yields the relationship between *h* and β when *h* is chosen by the median voter, denoted $h_M(\beta)$. Its inverse has a closed-form solution given by

(51) $$\beta_M^{-1}(h) = \rho h\left(1 - \frac{2\log(\varphi(h))\log(\mathrm{e}/2)}{(\log(\varphi(h))+1)^2}\right)$$

This tells us that there is a positive relationship between β and instructional hours: more productive economies provide more education in the median-voter outcome. Moreover, $h_M(\beta)$ is strictly convex in β. In contrast, recall that the relationship between β and the income-maximizing level of education is linear; recall (32) above. Thus, the ratio $h_M(\beta)/h^*(\beta)$ is increasing in β; richer countries deviate most from the income-maximizing education level.

I next construct the relationship between β and the preferred level of inequality for the median voter. Recall from (32) the relationship between *h* and the induced level of inequality when $\pi = 0$: $h_0^{-1}(G)$. This is independent of β. Thus, we can substitute $h_0^{-1}(G)$ for *h* in (51) to yield a relationship between β and *G* at the median-voter outcome. Let $\beta_M^{-1}(G)$ denote this relationship. We can then set $\beta = \beta_M^{-1}(G)$ in $\mu_0(G)$ from (40) to generate a cross-country relationship between *per capita* income and inequality when education policy is decided by the median voter:

(52) $$\mu_M^C(G) = \frac{(1-G)(1-2G\log(e/2))\delta\theta\rho c^2 e^{\left(\frac{1}{G}-1\right)}}{4(\delta+\theta)G}$$

This relationship is depicted as the heavy curve in Figure 8, together with individual-country Kuznets curves for two countries with productivity parameters β_1 and $\beta_2 > \beta_1$. Note that $\mu_M^C(G)$ passes through points M_1 and M_2; it is the locus of all such median-voter outcomes for every possible value of β. The arrows along $\mu_M^C(G)$ in the figure indicate the direction in which β is increasing, as well as the direction in which $h_M(\beta)$ is increasing.

The cross-country relationship in Figure 8 tells us that rich countries have less inequality than poorer countries when the only difference between countries is their exogenously given productivity parameter, β. In reality of course, countries differ according to a vast array of factors, including the extent to which policy is influenced by voting, so it is important to view $\mu_M^C(G)$ in the context of the model.

The more important message here relates to what a relationship like $\mu_M^C(G)$ does *not* tell us. In particular, one cannot infer from an observed relationship like $\mu_M^C(G)$ that greater investment in education is a magic bullet for hitting two targets at once. It might appear from $\mu_M^C(G)$ that *per capita* income and greater equality are complements – both a consequence of higher educations levels – but for every country along $\mu_M^C(G)$ the reality is quite the opposite. Each country is at a point on its own Kuznets curve at which *per capita* income and greater equality are *substitutes*: more of one must be accompanied by less of the other. For each of these countries, additional spending on education would reduce inequality, but that would come at the expense of *per capita* income.

Partial Voter Turnout

Voter turnout is typically less than complete in real elections, and low-income individuals are typically less likely to vote than those with higher incomes. It is therefore interesting to investigate briefly the implications of partial voter turnout.

Suppose that only those workers with income greater than y_V vote. In the Appendix I show that the median-voter outcome still involves $\pi = 0$ if and only if $y_V < \bar{y}_V$, where

(53) $$\bar{y}_V = \frac{\beta \log(\zeta)(\log(\zeta/4)+1)}{\log(\zeta)+1}$$

This is the threshold level for y_V above which the median voter has an income greater than μ^* when $\pi = 0$. Thus, if $y_V > \bar{y}_V$ then $\pi = 0$ no longer emerges as the median-voter outcome. In that case, the median-voter prefers the largest value of π on offer at any given *h*.[4]

It is straightforward to show that if $y_V > \bar{y}_V$ and $\zeta > \mathrm{e}^2$ (β is large relative to ρ) then the median-voter outcome is $\pi = \rho$ and $h = h^*$; tuition is fully priced, and *per capita* income is maximized. The Gini coefficient at this outcome is

(54) $$G_P = \frac{1}{\log(\zeta)}$$

and this can be expressed in terms of G_M, the outcome when all workers vote:

(55) $$G_P = \frac{G_M}{1 - G_M\left(1 - \log\left(1 - 2\log(e/2)G_M(1 - G_M)\right)\right)}$$

Thus, $G_P > G_M$ (since $G_M < 1$); inequality is higher than when all workers vote. This outcome is depicted in Figure 9 as point *P*. Note that *P* does not lie on the Kuznets curve that applies when all workers vote, labeled $\mu_0(G)$ in the figure, which is based on $\pi = 0$. Compared to point *M* in the figure – the median-voter outcome for the same economy when all workers vote – *per capita* income is higher at point *P* (because *h* is lower) but income-inequality is also higher (because *h* is lower and π is higher).

---

[4] The simplicity of this model means that the income threshold $\bar{y}_V$ translates into a numerical voter-turnout percentage that is independent of all model parameters. That percentage is just below 74% .

The curve labeled $\mu_P^C(G)$ in Figure 9 depicts the cross-country Kuznets curve when there is partial voter turnout in all countries, and $y_V > \bar{y}_V$ in all countries (so the voting outcome is characterized by fully-priced tuition and the income-maximizing level of *h*). That is, $\mu_P^C(G)$ is the partial-voter-turnout analogue of $\mu_M^C(G)$.

Note from Figure 9 that $\mu_P^C(G)$ is kinked at $\mu = \bar{\mu}$, where

(56) $$\bar{\mu} = \frac{\delta\theta\rho c^2 \mathrm{e}^2}{2(\delta+\theta)}$$

Countries with $\mu < \bar{\mu}$ have $\zeta < \mathrm{e}^2$ (β is small relative to ρ) and this means that fully-priced tuition is not an available option because it would violate the non-negativity constraint on the lowest income. For these countries, the median-voter outcome is $h = h^*$ and $\pi = \pi_H(h^*) < \rho$. The Gini coefficient in these countries is at its upper bound of $\frac{1}{2}$ (given the non-negative lowest-income constraint), and inequality does not start to fall as income rises until income reaches the $\bar{\mu}$ threshold.

While the partial-turnout voting outcome exhibits more inequality than when all workers vote, inequality still cannot be reduced without a reduction in *per capita* income, and in that regard it shares the same key property of the full-voting outcome. Under both voting scenarios, the cross-country Kuznets curve is negatively sloped (above $\bar{\mu}$ in the partial-voter-turnout case), and lower inequality is correlated with more education, but we cannot interpret that to mean that more education is a good recipe for any individual country wishing to reduce inequality.

I do not wish to overstate the message here. There can be little doubt that long-term productivity growth – an advance in β over time – is tied to investment in education, and this link is entirely absent from my model. Nonetheless, I believe a central point here is robust: one must be very careful about drawing inferences from cross-country relationships between income, inequality and education spending, when looking for a magic bullet that can reduce inequality and boost income at the same time.

8. CONCLUSION

I have characterized the educational choices made by workers in terms of their innate aptitudes and how closely those aptitudes match available job types. A worker can choose to specialize in a narrow field or pursue a more general education based on the closeness of that match. Workers fortunate enough to have an innate aptitude closely matched to a job type will become highly specialized, and will reap high market rewards. Less fortunate workers will be less well-rewarded. Thus, there arises a distribution of income in the economy. The mean, variance and skewness of that distribution is a function of education policy, relating in particular to the educational opportunities made available, and to how they are funded, whether through income tax or tuition charges. I have cast that policy choice in terms of picking a point on a Kuznets curve, which dictates the relationship between average income and income-equality in this economy.

REFERENCES

Altonji, Joseph G., Peter Arcidiacono and Arnaud Maurel (2016), "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects", Chapter 7 in *Handbook of the Economics of Education*, Volume 5, 305 – 396.

APPENDIX

Derivation of the Income Distribution

The cumulative density of income within a representative arc is

(A1) $$F(Y) = P\{y \le Y\} = 1 - P\{a_L^{-1}(Y) \le a \le a_H^{-1}(Y)\}$$

where

(A2) $$a_L^{-1}(Y) = \tau - \frac{c}{2}\left(\frac{\varphi(h)}{\mathbf{e}^{\left(1+\frac{Y+\pi h(1-t)}{\beta(1-t)}\right)}}\right)^{\frac{1}{2}}$$

and

(A3) $$a_H^{-1}(Y) = \tau + \frac{c}{2}\left(\frac{\varphi(h)}{\mathbf{e}^{\left(1+\frac{Y+\pi h(1-t)}{\beta(1-t)}\right)}}\right)^{\frac{1}{2}}$$

are the solutions to $y_i^*(a) = Y$.

These solutions are illustrated in Figure A1. Note that the domain of the density function denoted $f(Y)$, has two distinct regions. The first of these regions is where $Y < y_L$. In this region, $f(Y) = 0$. The second region is where $Y \ge y_L$. In this region,

(A4) $$F(Y) = 1 - P\{a_L^{-1}(Y) \le a \le a_H^{-1}(Y)\} = 1 - [D(a_H^{-1}(Y)) - D(a_L^{-1}(Y))]$$

where $D(a)$ is the cumulative density of the distribution of *a* within the representative arc. Thus, in region 2

(A5) $$f(Y) = \frac{\partial D(a_L^{-1}(Y))}{\partial Y} - \frac{\partial D(a_H^{-1}(Y))}{\partial Y}$$

Since a is distributed uniformly on $[\tau - c/2, \tau + c/2]$, it follows that

(A6) $$f(Y) = \frac{1}{2\beta(1-t)}\left(\frac{\varphi(h)}{\mathbf{e}^{\left(1+\frac{Y+\pi h(1-t)}{\beta(1-t)}\right)}}\right)^{\frac{1}{2}} \quad \text{for} \quad Y \ge y_L$$

and $f(Y) = 0$ otherwise. Thus, *Y* has an exponential distribution. The associated cumulative density is

(A7) $$F(y) = \int_{y_L}^{y} f(Y)dY = 1 - \left(\frac{\varphi(h)}{\mathbf{e}^{\left(1+\frac{y+\pi h(1-t)}{\beta(1-t)}\right)}} \right)^{\frac{1}{2}}$$

Derivation of the Lorenz Curve

Let $y(p)$ denote the minimum income level below which a fraction *p* of workers earn. Thus, $F(y(p)) = p$. Solving for $y(p)$ yields

(A8) $$y(p) = (1-t)\left(\beta\left(\log\left(\frac{\varphi(h)}{(1-p)^2} \right) - 1 \right) - \pi h \right)$$

We can now find the average income of workers who earn less than $y(p)$, denoted $x(p)$:

(A9) $$x(p) = \int_{y_L}^{y(p)} yf(y)dy = (1-t)\left(\beta\left(p\log(\varphi(h)+1) + 2(1-p)\log(1-p) \right) - p\pi h \right)$$

The Lorenz curve is then constructed as

(A10) $$L(p) = \frac{x(p)}{\mu} = p + \frac{2\beta(1-p)\log(1-p)}{\beta(\log(\varphi(h))+1) - \pi h}$$

Derivation of Condition (53)

The starting point is to identify lower and upper bounds on *voter* aptitudes, denoted a_{VL} and a_{VH} respectively, that ensure that the median voter has the average income among workers. This requires that $\bar{a}_L = (a_{VL} + \tau)/2$ and $\bar{a}_H = (a_{VH} + \tau)/2$; that is, $\bar{a}_L$ and $\bar{a}_H$ must respectively constitute the midpoints of voting workers below and above τ. These conditions imply that $a_{VL} = 2\bar{a}_L - \tau$ and $a_{HL} = 2\bar{a}_H - \tau$. Making either of these substitutions for *a* in $\hat{y}(a)$ from (6), and setting $t = t(h)$, $h = h^*$ and $\pi = 0$, yields

(A11) $$\bar{y}_V = \frac{\beta\log(\zeta)(\log(\zeta/4)+1)}{\log(\zeta)+1}$$

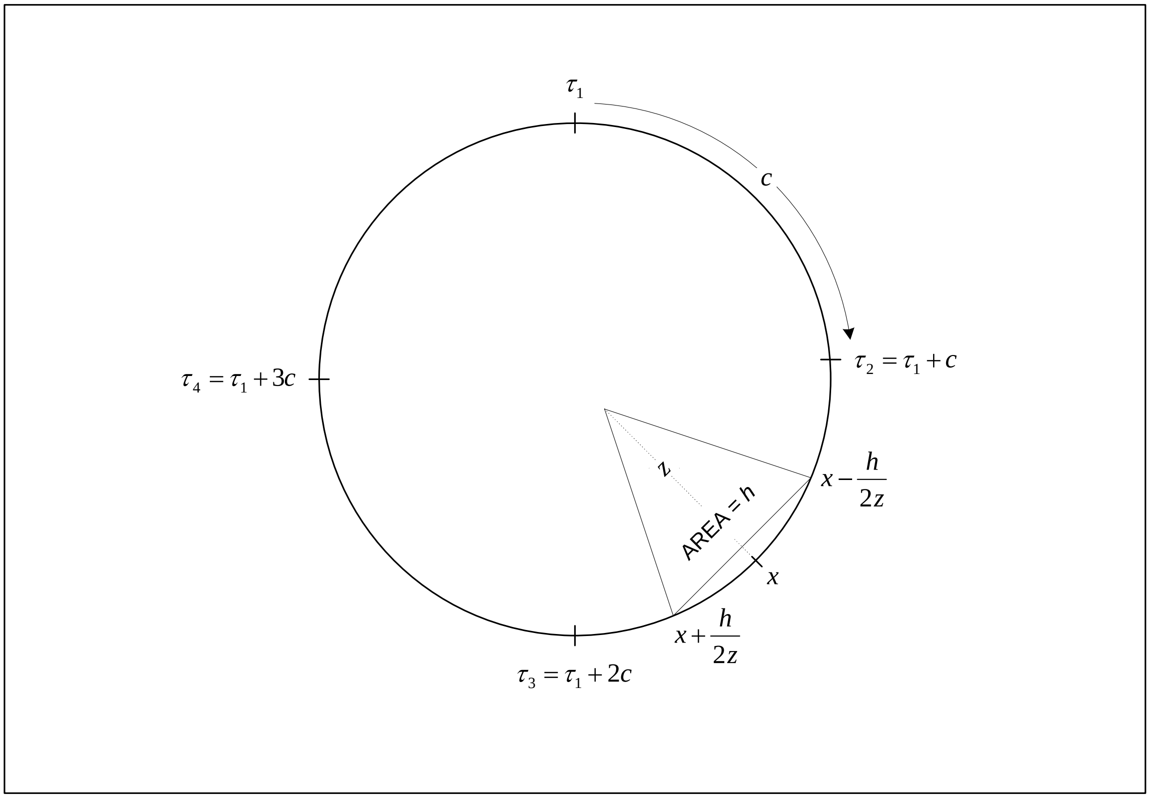

FIGURE 1

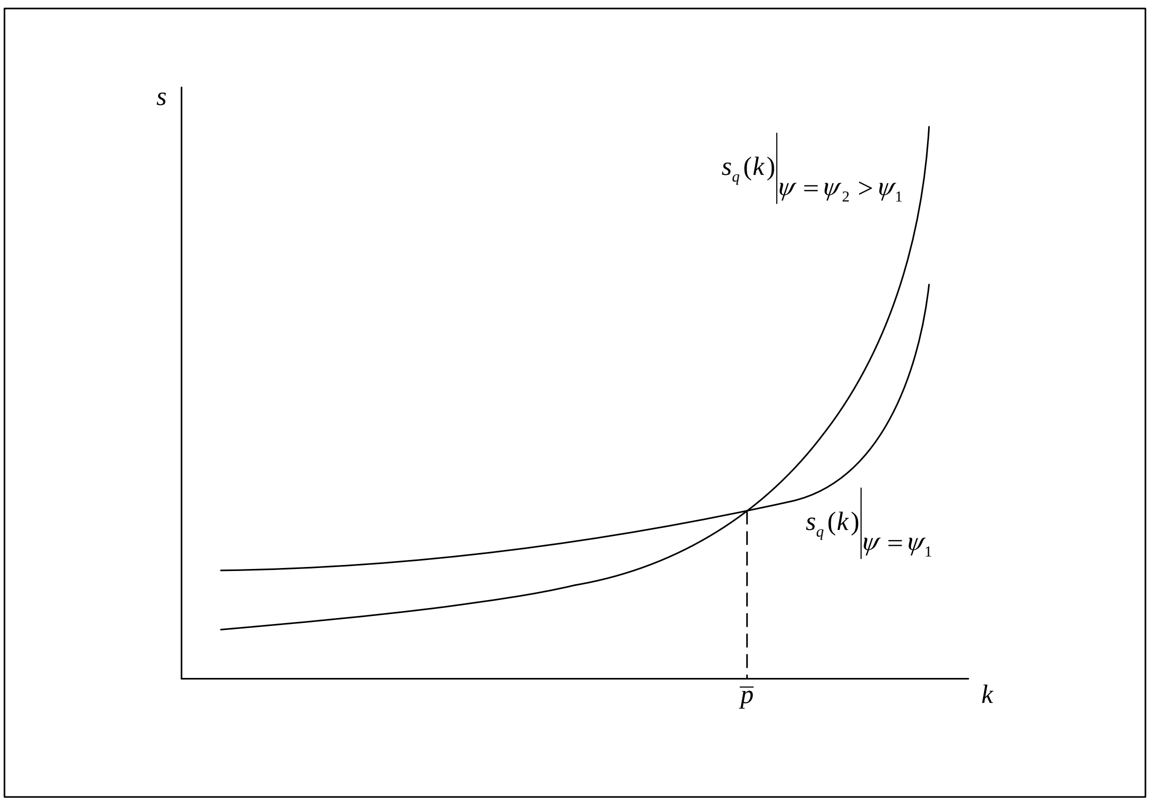

FIGURE 2

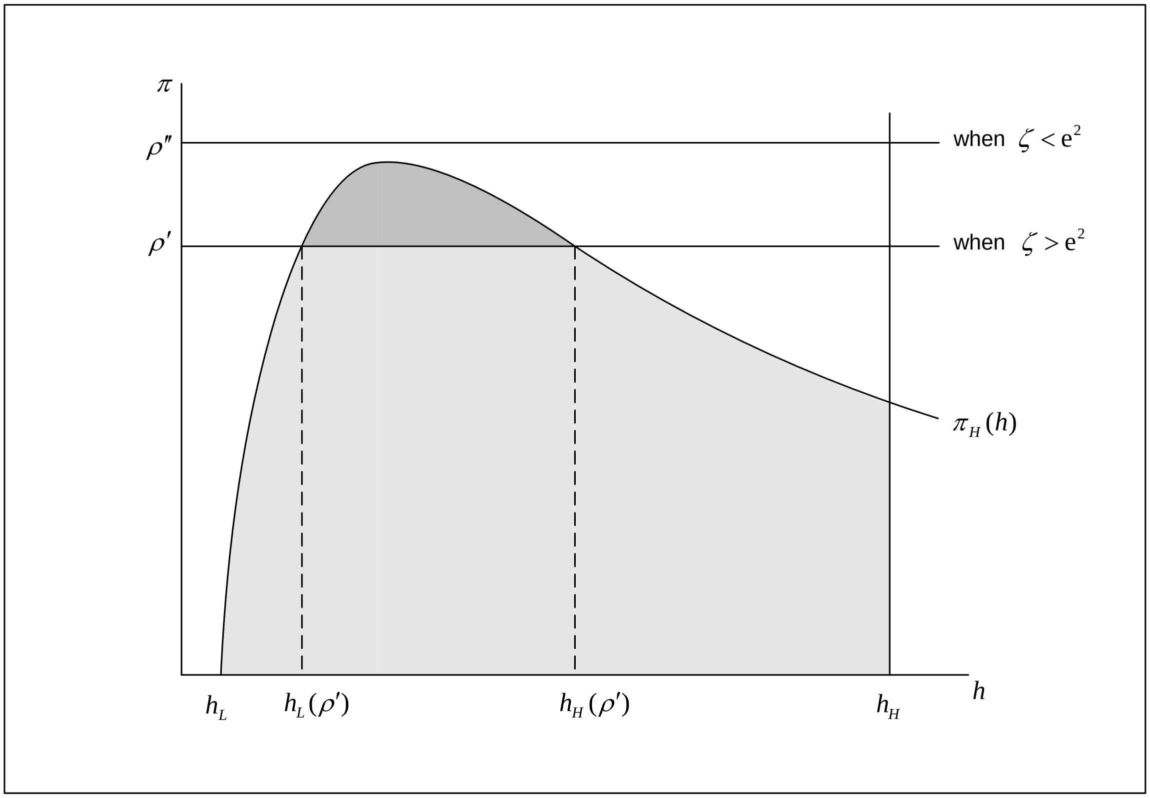

FIGURE 3

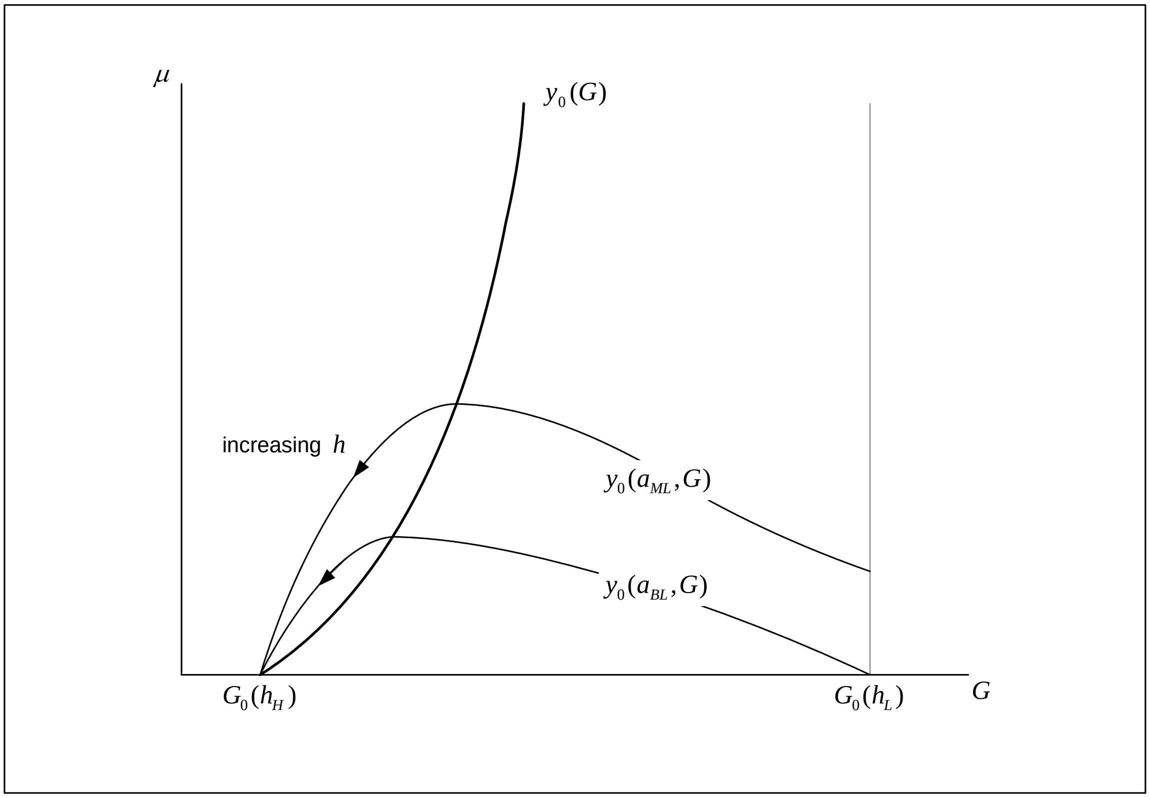

FIGURE 4

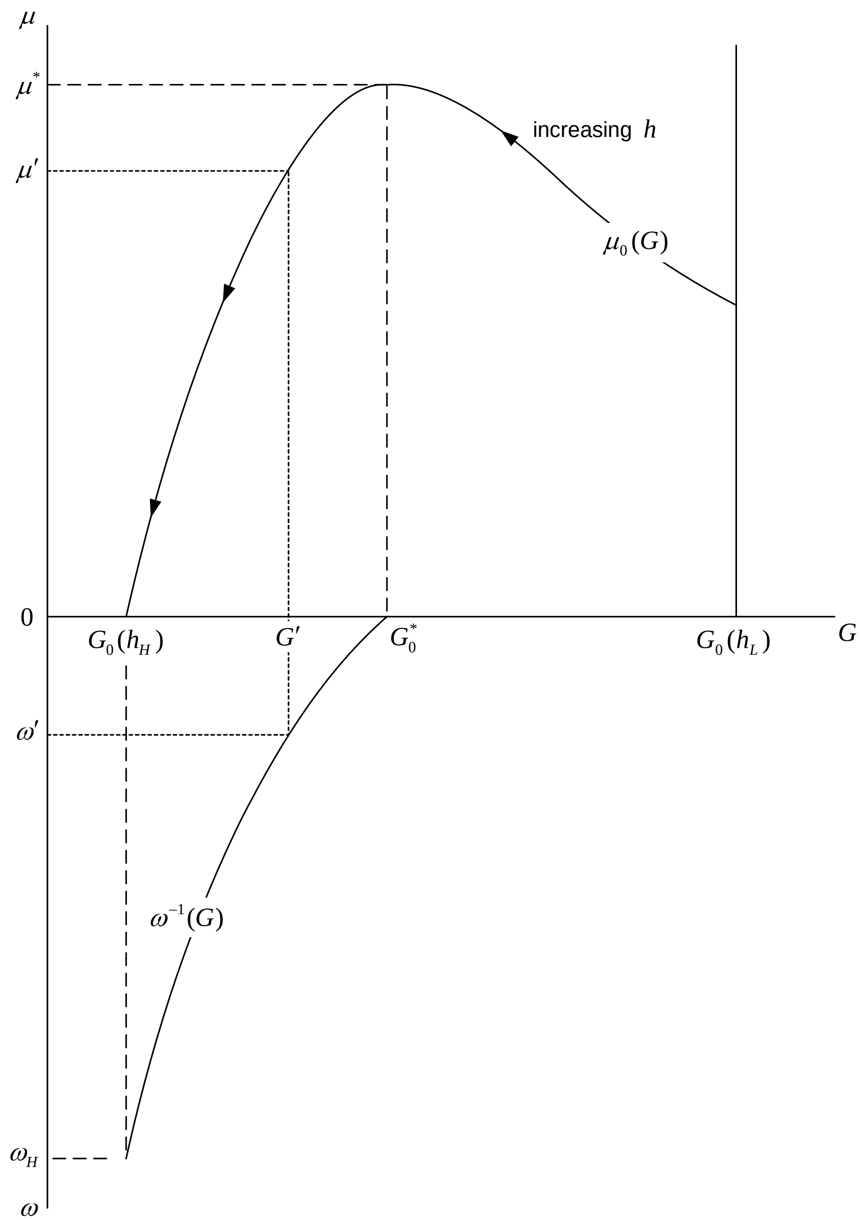

FIGURE 5

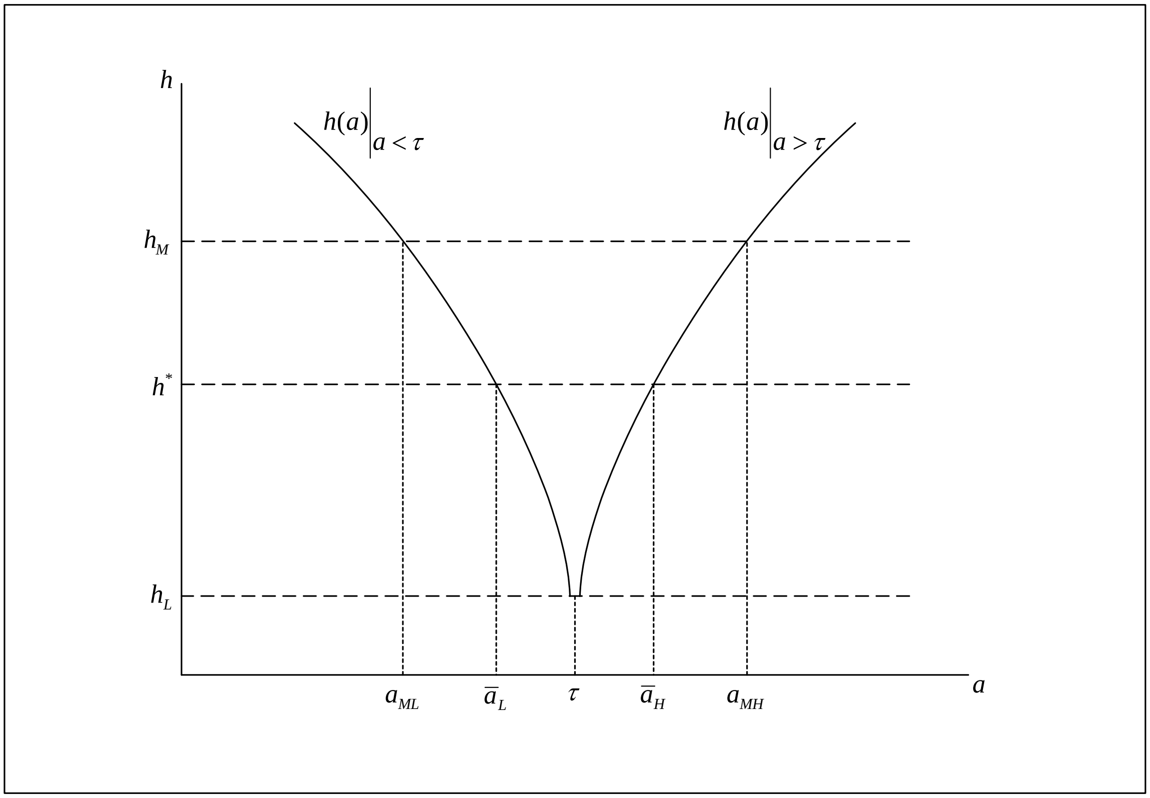

FIGURE 6

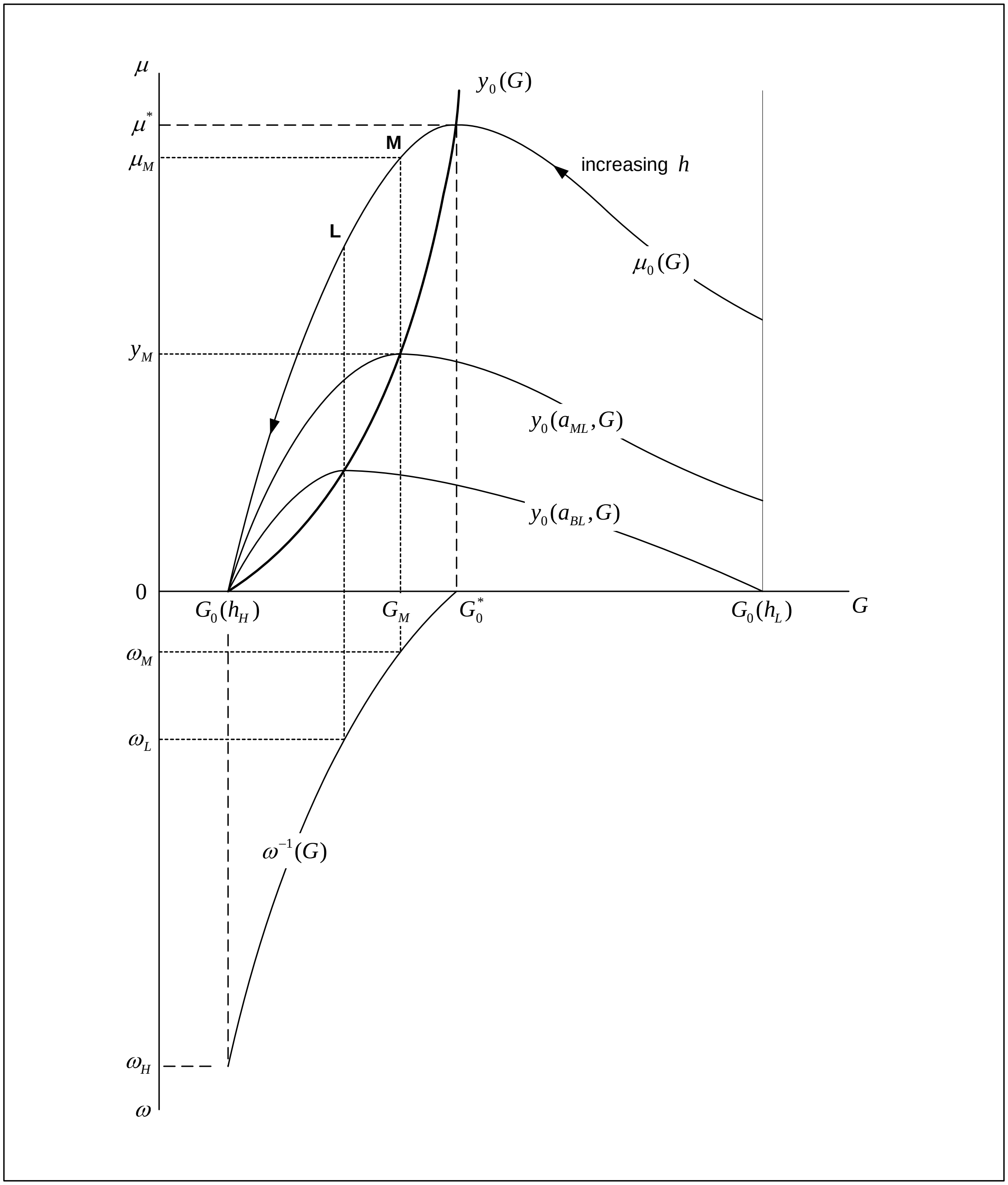
μ
μ*
μ_M
y_M
0
ω_M
ω_L
ω_H
ω
y_0(G)
M
L
increasing h
μ_0(G)
y_0(a_ML, G)
y_0(a_BL, G)
G_0(h_H)
G_M
G_0*
G_0(h_L)
G
ω^-1(G)

FIGURE 7

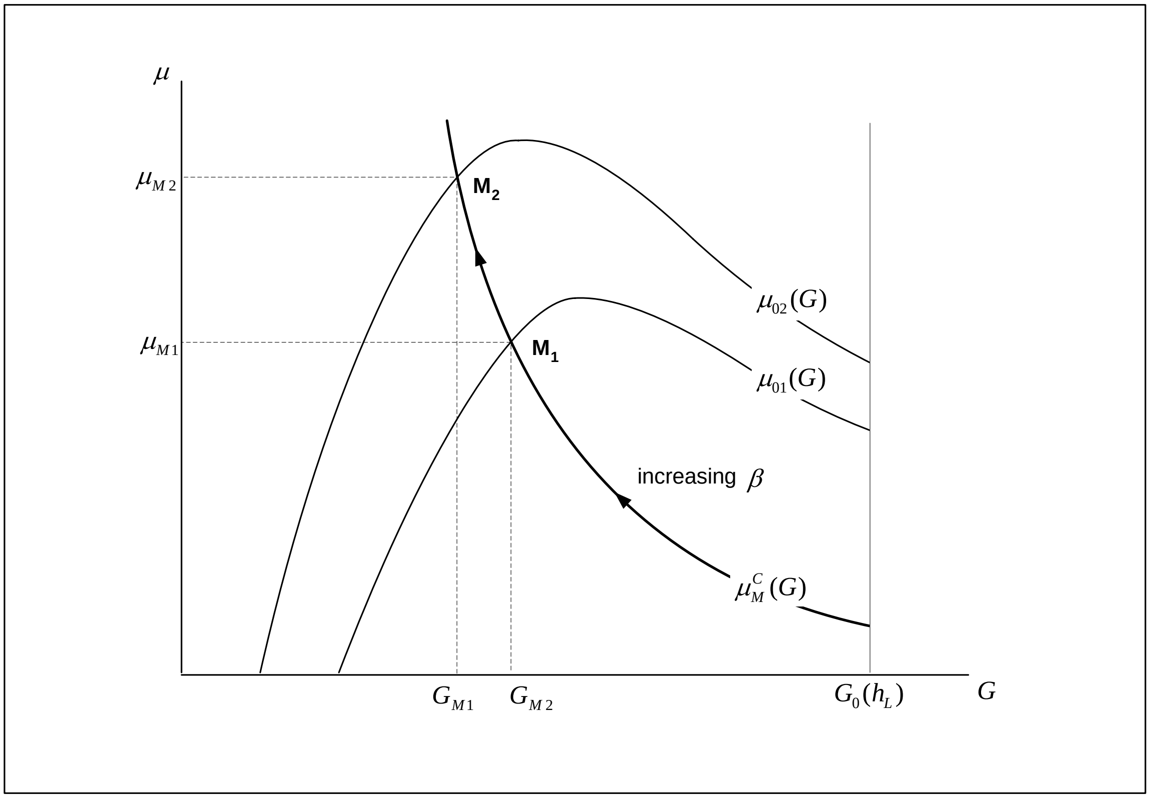

FIGURE 8

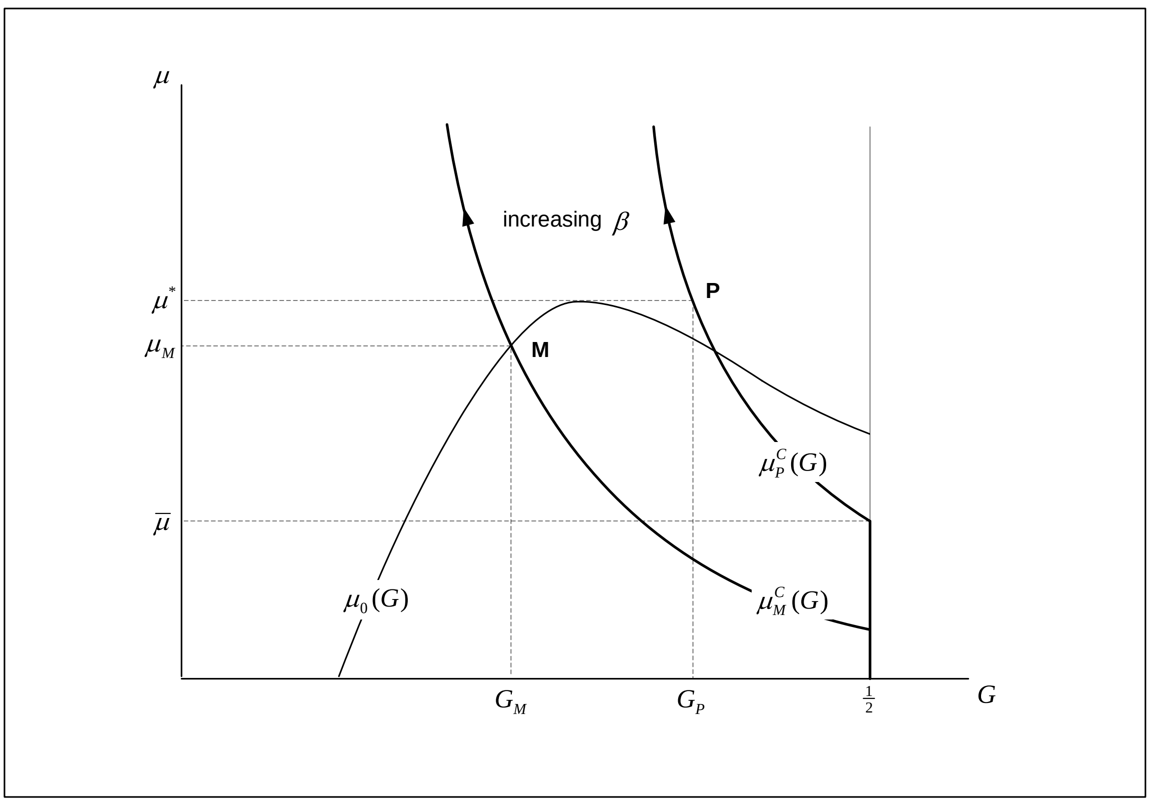

FIGURE 9

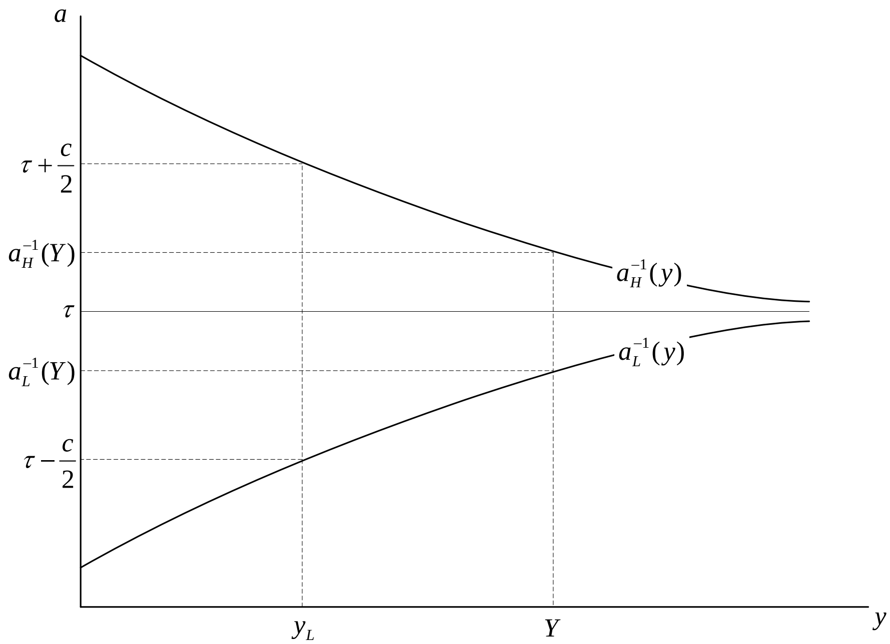

FIGURE A1