QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

JOHN E. ROEMER E ROBERTO VENEZIANI

What we owe our children, they their children, …..

n. 328 - Settembre 2001

Abstract - Egalitarian theorists, since Rawls, have in the main advocated equalizing some objective measure of individual well-being, such as primary goods, functioning, or resources, rather than subjective welfare. This discussion, however, has assumed, implicitly, a static environment. By analyzing a society that survives for many generations, we demonstrate that equality of opportunity for some objective condition is incompatible with human development over time. We argue that this incompatibility can be resolved by equalizing opportunities for *welfare*. Thus, 'subjectivism' seems necessary if we are to hope for a society which can both equalize opportunities and support the development of human capacity over time.

JEL classifications: C61, D63, D9, H21.

Keywords: justice, development, dynamic programming, optimal taxation.

John E. Roemer*, Depts. of Political Science and Economics, Yale University.
Roberto Veneziani**, Dept. of Economics, L.S.E. and Political Sciences.

* I thank Alberto Bisin for useful discussions of dynamic programming. Criticism from G.A. Cohen has also been important.
** I am very grateful to Heski Bar-Isaac and the participants to seminars at the University of Milan and the London School of Economics for many helpful discussions and comments.

1. Introduction.

Egalitarians - and more specifically, socialists - have long cherished two ideals: that that society is best which promotes human development over time, and equality of condition among members of society.[1] More recently, since Rawls's rejuvenation of egalitarian studies, several qualifications have been put forth as to what the equalisandum should be. Most, although not all, participants in the discussion have advocated what we call an objectivist view, that the equalisandum should be something which is measurable independently of the views of the individuals who have it - primary goods, functionings, or resources (Rawls (1971), Sen (1980), and Dworkin (1981), respectively). The principal non-objectivist equalisandum is, of course, welfare or utility, which can only be measured knowing the utility function of the individual in question, and can only be compared interpersonally if an interpersonally comparable unit scale exists. None of the major writers advocates equality of welfare as an ethic.

Moreover, in recent years, various theories of equal opportunity have been proposed including Arneson (1989), Cohen (1989), and Roemer (1998), and we would say that Dworkin's (1981) equality-of-resources is indeed an equal-opportunity theory as well. So we might well say that egalitarians advocate, as well as human development, equality of opportunity for some condition. That condition could be something objective like functionings or primary goods, or the subjective welfare.

What we argue in this article is that the three desiderata

1. protracted human development;
2. equality of opportunity for some condition;
3. the condition be an objective characteristic of the individual;

are inconsistent. Because the first desideratum makes sense only in a dynamic context, equality of condition, or equality of opportunity for some condition, becomes equality (of opportunity) among all adults who ever live. Our claim says that if the equalisandum is

[1] Socialists have said (before consciousness about gender-neutral language) that, in the good society there will be 'self-realization of man' and 'self-realization of men.' The latter means that, over the course of a life, a person becomes self-realized, in the sense of developing her capacities. The former means that, over generations, human beings become more knowledgeable and developed. Here, we take human development to mean 'self-realization of man.'

objective - something like functioning - then achieving such equality will imply the absence of human development over time. It is only by taking the equalisandum to be welfare of a particular kind, a non-objectivist concept, that equality of opportunity is consistent with human development.

If our claimed inconsistency is correct, then egalitarians are faced with a choice: either dropping their advocacy of equality (of opportunity), or of human development, or of objectivist equalisanda. We think that the most attractive choice is to drop the objectivist view.

In other words, we claim to show that, if we move away from the static thought experiments imagined by Rawls and the objectivist writers heretofore, then objectivism ceases to be attractive (if it ever was). We must say, however, that our inquiry does not show that *justice* requires that we endorse subjectivism (the view that welfare is what must count for an egalitarian). For we advocate dropping objectivism because of its inconsistency with equality of opportunity *and human development;* and while the equality-of-opportunity part of that compound phrase refers to a state of justice, the 'human development' part does not. That is, we do not claim that justice requires human development, or even, more weakly, that justice requires human development in an environment where it is possible. Human development over time is, for us, an obvious good, but we do not know what to call the state of a society which has it, the way a society with equality of opportunity is in a state of justice.

2. The dynamic environment.

We will model the problem in a stark way. There is a society that exists for many (an infinite) number of generations. At each generation there are adults and children. Each adult has one child, and so the population size is constant. Adults, at least at the beginning date zero (0), have different wage rates - indeed, we shall seek simplicity by declaring that only two wage rates exist at date 0. Taxation of adult income is used to finance education of that generation's children, as well as to redistribute income among adults.

We suppose that an adult's wage is a measure of her family's socio-economic status (SES), where SES has an impact on the docility[2] of children. More specifically, the economic

[2] In the classical sense -- educability.

outcome of educating a child is the wage she will earn as an adult, and it takes more educational resources to bring a low SES child up to a given (adult) wage rate than it does a high SES child. We take the view that all children have identical inborn talent, and that the wage a child eventually earns as an adult is a function of her talent, the educational resources invested in her, and the SES status of her parent, our summary of the environmental factor. To be specific, we suppose there are two functions h: $\mathbf{R}_+ \rightarrow \mathbf{R}_+$ and g: $\mathbf{R}_+ \rightarrow \mathbf{R}_+$ such that a child of a parent who has a wage of w will, as an adult, earn a wage of $h(x)g(w)$, if x is the fraction of GNP per capita that is invested in her through the educational process. We assume:

Assumption 1: h and g are continuous and strictly increasing. Moreover, $h(0) = 0$ and $g(0) = 0$.

Our economic environment dispenses with two important aspects of reality - that children are differentially talented, and that children expend differential effort[3] - since we think they are unnecessary to expose the problem we want to concentrate upon. Capital and natural resources exist only implicitly in this model.

At each generation, adults must tax themselves, and the tax revenues, in the form of educational finance, must be distributed between the two types of child, those from low wage parents and those from high wage parents. The result of that education will be adults at the next date who have (perhaps) two wage levels, and the problem repeats itself. All children of a given SES receive the same educational investment, and hence have the same wage as adults.

We shall suppose that taxation takes the following form. First, all adult incomes are pooled, and each adult receives the average income. Then each adult pays the same fraction of her income as a tax. At date 0, a fraction f_L of the adults earn the low wage w_L^0, and fraction f_H earn the high wage, w_H^0. Thus, $f_L + f_H = 1$, and we define mean income at date 0 as $\mu^0 = f_L\, w_L^0 + f_H\, w_H^0$. If the tax rate is τ^0, then the after-tax income of every adult is $(1 - \tau^0)\, \mu^o$.

[3] One is, of course, free to interpret the difficulty in educating low SES children as due to their lower talent. This is formally equivalent to our model, yet might lead to different ethics. (Some would say that it's alright for low talent people to earn less than high talent people, although it's not alright for kids from disadvantaged backgrounds to earn less than equally talented kids from advantaged backgrounds.)

We wish to abstract from incentive problems; in particular, taxation does not alter labor supply, nor does anticipation of their future after-tax income affect how hard children work in school. These would be poor assumptions if we were interested in advising policy-makers, but our investigation here is of a different kind. We are interested in exposing certain logical inconsistencies in a conception of 'the good society', and it is appropriate for this inquiry to assume that individual citizens are almost perfectly cooperative. We limit their cooperative spirit only by assuming that private incentives would come into play if we redistributed adult income so that low wage earners ended up with *more* income that high wage earners. (The best we can do is to equalize all after-tax incomes.)

In the theory of equal opportunity (see Roemer [1998]), it is assumed that individuals have different circumstances and exert different efforts. Here, we abstract away from differential effort. A person's circumstances -- those characteristics beyond her control that influence her outcome -- are two in number, the SES (wage) of her parent, and the date at which she is born. We shall take children as adults-in-formation, and are concerned with equalizing opportunities among adults for some condition X, which we shall call 'welfare.' The instruments we have available are the tax rates and the distribution of educational finance among child types at each date. Since effort is nugatory, the theory of equal opportunity expounded in Roemer (1998) says that our objective is to maximize the minimal level of 'welfare' among all adults across types, where an adult's type is a pair (w, t), w being her parent's wage, and t being the date at which she is born. Informally speaking, the SES of a child's parents and the date at which she is born are circumstances beyond her control, and equality of opportunity requires that we equalize, so far as possible, the welfare of individuals with such different circumstances.

Thus, our problem is to maximize the least level of 'welfare' across all adults who ever live. To be specific, at each date we must choose a tax rate of adult income, τ, and, if there are adults with two wage levels (there are never more than two), an allocation of educational finance (r_L, r_H) among children of the two types, where $f_L r_L + f_H r_H = 1$. A child from an L family will receive educational investment in the amount $\tau\mu r_L$ and a child from an H family will receive $\tau\mu r_H$. Thus, if w_L and w_H were the parents' wages, then the children will earn, as adults, $h(\tau r_L)g(w_L)$ and $h(\tau r_H)g(w_H)$.

We next define the notion of *functioning*. We say that an adult's level of functioning is a function $F(w, y)$ of her wage, w, and her consumption (after-tax income), y - F: $\mathbf{R}_+ \times \mathbf{R}_+ \rightarrow \mathbf{R'}$, where **R'** represents the extended real line. We attempt to capture A. Sen's (1980) idea of functioning, which G. A. Cohen (1993) has characterized as 'midfare,' something midway between consumption and welfare. To wit, we imagine that a person's wage is a measure of her level of human capital and individuals derive welfare directly from their human capital. Functioning involves a degree of self-esteem and self-realization, and these, we propose, depend positively on an individual's level of human capital. Human capital, in turn, is reflected in the wage.

Let $F^* = \inf F(w, y)$. In what follows we will assume:

Assumption 2: F is continuous and monotone increasing in both arguments. Furthermore, $\lim_{w \to 0} F(w, y) = F^*$, for all y, and $\lim_{y \to 0} F(w, y) = F^*$, for all w.

In section 5, we shall assume:

Assumption 2': $F(w, y) = \gamma \log w + (1 - \gamma) \log y$, where $0 < \gamma < 1$.

We define *human development* as *an increase in functioning level of adults over time*. We believe this is consistent with the standard concept of human development, which is not an increase in welfare as such, but rather an increase in human capacity. Capacity, in our stark model, is a function of consumption and the wage, or more directly, of consumption, self-esteem, and self-realization. The wage is important as the reflection of education; in addition, it can be argued that self-esteem is a capacity enhancer, and that, too, is captured by the wage. Children embody the knowledge of past generations, through the educational process, and we have attempted to capture this in our specification of the educational technology.

This model has similarities to Arrow (1973) and Dasgupta (1974), in which the maximin criterion was examined in a dynamic framework. The main substantive difference is that we posit two types of individual, at least at the early dates, while Arrow and Dasgupta work with a representative agent. Thus, we are interested in what intergenerational equality requires with respect to *intra*-generational wage differentials, a question that neither Arrow nor Dasgupta posed.

3. Equality of opportunity for functioning: Model I.

Our first exercise is to take the 'welfare' of an adult to be her functioning level. Thus, our problem becomes to

Sup Inf $[F_L^0, F_H^0, F_L^1, F_H^1, \ldots]$, (3.1)

where F_J^t is the functioning level of adults in the 'J dynasty' at date t. The 'low dynasty' is the set of persons consisting of the low wage adults at date 0 and all their descendants; likewise for the 'high dynasty.' The instruments of the optimization are $\{\tau^t, r_L^t\}_{t=0,1\ldots} \equiv \{\tau^0, r_L^0, \tau^1, r_L^1, \ldots\}$, where we note that r_H^t is determined by r_L^t via the accounting identity $f_L\, r_L + f_H\, r_H = 1$. The level of functioning of J adults at date t is $F_J^t = F(w_J^t, (1 - \tau^t)\mu^t)$, where μ^t is mean income at date t, and the wages are given recursively by $w_J^t = h(\tau^{t-1}\, r_J^{t-1})g(w_J^{t-1})$, $\forall t > 0$, $J = H, L$. Hence,

$w_J^t = h(\tau^{t-1}r_J^{t-1})*g\{h(\tau^{t-2}r_J^{t-2})*g\{h(\tau^{t-3}r_J^{t-3})*....*g\{w_J^0\}\}....\}\}$, $\forall i$, $J = L, H$. (3.2)

It is important to note that, at some date, the wages of the two adult types may be equalized, and if that is the case, then we stipulate that, thereafter, since there is only one type of child, there is no longer any decision concerning how to allocate educational finance - all children receive the same investment. We need not consider the possibility that a child in the H dynasty has a wage lower than one in the L dynasty at a given date, for that will never be an aspect of an optimal solution. It thus follows that at any date, the functioning level of L adults will be less than or equal to the functioning level of H adults (where L and H refer to the *dynasties,* not to the wages of particular adults), because the two types have same consumption. Consequently, the equality-of-opportunity program takes the form:

Sup Inf $[F_L^0, F_L^1, \ldots]$ (3.1')

s.t. $w_H^t \geq w_L^t$, $t = 1, 2, \ldots$

We immediately observe:

Proposition 1. Let A1, A2 hold. At the solution to (3.1'), $F_L^0 = F_L^t$, $\forall t$.

Proof: 1. Clearly, $0 < \tau^t < 1$, $\forall t$, since $\lim_{\tau^t \to 1} F_L^t = F^*$, and if $\tau^{t'} \to 0$ then, by (3.2), $F_L^t \to F^*$, $\forall t \geq t'$, which by A2 are certainly not optimal.

2. Suppose $F_L^0 > F_L^{t'}$, some $t' > 0$. Then it is possible to increase τ^0 a little and leave all other variables the same, so that F_L^0 is still above the minimum, while, by (3.2) w_L^t, and thus F_L^t, increase $\forall t > 0$.

3. Suppose $F_L^{t'} > F_L^0$, $t' > 0$. By part 1, decrease $\tau^0, \ldots, \tau^{t'-1}$ so that consumption and therefore the levels of F_L^t increase $\forall t \leq t'-1$. Next, increase $\tau^{t'}$ enough so that w_J^t, $J = L, H$, $t > t' + 1$ increase. We have now increased F_L^t, $\forall t \neq t'$ and we can make all these changes in tax rates small enough so that $F_L^{t'}$ is not the smallest functioning level.▯

P1 establishes that equality of opportunity for functioning is inconsistent with human development, in the sense that a fraction f_L of adults at every date remain at the (low) level of functioning of date 0 *L* adults. If, as is reasonable, $f_L > .5$, then the majority of all adults are held to a low level of human capacity.

Do the *H* adults get reduced, over time, to this same low level of functioning? Not necessarily. Let, e.g., A2' hold: if γ is sufficiently close to 0, then consumption is very important in functioning, and it may pay to keep the wages of the *H* adults above the *L* adults' wages in order to bring about a relatively high mean income.

The maximin social welfare function is sometimes criticized for spending huge amounts of resources to raise the level of welfare of a very small group of individuals who are very poor welfare producing machines. Let us note this criticism does not apply here. Nobody is extremely handicapped in our environment - there are no terribly inefficient 'welfare'-creating individuals. It is true, however, that *L* adults at date 0 comprise an arbitrarily small fraction of the adults who have lived up to date *T*, as *T* becomes large, and all *L* adults are held to *their* level of functioning. This is surely a form of 'extremism' of maximin, although it has a different character from the form of extremism we referred to in the first sentence of this paragraph. If we contemplate sacrificing the *L* adults at date 0, we are led to ask, why do they have less than an equal right to welfare than those at later dates? The answer 'Because it is too costly to their descendents *not* to sacrifice them' invites sacrificing the *L* adults, or indeed all adults, at any finite number of dates beginning at date 0. After all, this group, too, constitutes an arbitrarily small fraction of all adults who shall ever live.

4. Equality of opportunity for welfare: Model II.

We now suppose that adults care about the functioning levels of their children, as well as their own. We define the utility of an adult of dynasty *J* at date *t* as $u_J^t = u(F_J^t, F_J^{t+1})$, where F_J^t is her own functioning level, F_J^{t+1} is the functioning level of her child, and *u* is monotone increasing in both arguments and continuous. Hence, if $u^* \equiv inf\, u(F_J^t, F_J^{t+1})$, with a slight abuse in notation, we may write $u(F^*, F^*) = u^*$. We will also rule out an extreme form of altruism by assuming

Assumption 3: $u(X, F^*) \geq u(F^*, X)$.[4]

Our equal-opportunity program now becomes

Sup Inf $[u_L^0, u_L^1, \ldots]$

$$\text{s.t.} \quad w_H^t \geq w_L^t, \ \forall t, \tag{4.1}$$

where the requirement $w_H^t \geq w_L^t$, $\forall t$, is surely superfluous. We now have:

Proposition 2. Let A1, A2, A3 hold. Let m denote the value of program (4.1). At the solution to (4.1), $u_L^0 = m$. Furthermore, there are no two consecutive dates t and $t + 1$ such that $u_L^t > u_L^0$ and $u_L^{t+1} > u_L^0$.

Proof. 1. Clearly, in the optimum $\tau^t > 0$, $\forall t$. Similarly, $\tau^t \to 1$ *and* $\tau^{t+1} \to 1$ imply that $u_L^t \to u^*$, therefore it must be either $\tau^t < 1$, or $\tau^{t+1} < 1$, or both.

2. Suppose $u_L^0 > m$. Assume that F is defined at $\tau = 1$. Case 1. If $\tau^0 < 1$, increase τ^0 a little. This raises u_L^t, $\forall t > 0$, and does not lower u_L^0 to m. Case 2. Let $\tau^0 = 1$, and thus $u_L^0 = u(F^*, X)$, $u(F^*, X) > m$. By A3 it follows that $u_L^1 > m$, $\forall \tau^1, \tau^2$. Hence, by part 1, increase τ^1 a little: both u_L^0 and u_L^1 remain above m, while u_L^t increases, $\forall t \geq 2$.

3. Suppose $u_L^1 > m$ and $u_L^2 > m$. (The same argument holds for any two consecutive u_L^t and u_L^{t+1}, $t \geq 1$.) Then decrease τ^1 a little, which increases u_L^0 above m. Assume that F is defined at $\tau = 1$. Case 1. If $\tau^2 < 1$, increase τ^2 so that both w_L^3 and w_H^3 (and thus w_J^t and u_J^t, $\forall t \geq 3$, $J = L, H$) are at least as high as before the perturbations.[5] Since u_L^1 and u_L^2 were initially greater than m, they still are. Hence, by part 2, it follows that u_L^1 and u_L^2 cannot be both greater than m. Case 2. Let $\tau^2 = 1$ and thus $u_L^2 = u(F^*, X)$, $u(F^*, X) > m$. By A3, $u_L^3 > m$, $\forall\ \tau^3, \tau^4$. Hence,

---

[4] A3 shortens the proof of P2 considerably, however none of the main results changes if A3 is dropped. Notice that if A2' holds and $F^t \to F^*$ implies $u(F^t, F^{t+1}) \to u^*$ (e.g. if u is additive), then $\tau^t < 1$, $\forall t$ (cf. L1 below), and P2 immediately follows.

[5] If $h(\tau^t r_j^t) = k(\tau^t r_J^t)^{c_1}$, and $g(w_j^t) = (w_J^t)^{c_2}$, $J = L, H$, it is not difficult to show that, given r_L^t, r_L^{t+1} constant, if $d\tau^{t+1}/d\tau^t = - c_2 \tau^{t+1}/\tau^t$, then w_L^{t+i} and w_H^{t+i}, remain unchanged, $\forall t \geq 0$, $\forall i \geq 2$.

by part 1, increase τ^3 a little, so that u_L^1, u_L^2 and u_L^3 remain above m, while u_L^t increases, $\forall t >$ 3.

If F^t is not defined in $\tau^t = 1$, let $\tau^t \to 1$ and notice that $u_L^t \to u(F^*, X)$, with $u(F^*, X) > m$. Then by A3 all arguments in parts 2 and 3 above follow.▯

If each adult cares about her child's and her grandchild's level of functioning, then the same argument shows that no *three* consecutive utilities can be greater than the value of the program, which is achieved at date 0. Thus, allowing parents to care about the functioning levels of a finite sequence of their descendents does not enable us to escape the conclusion that protracted human development fails to occur. For it is clear that if the utility level of the L dynasty returns to u_L^0 periodically, then the functioning level of one generation must return, periodically, to F_L^0 or F_L^1 or lower, by u's monotonicity. In this society, history repeats itself, condemning every n^{th} generation to the level of human development of the primeval ancestor.

It is worth noting that u can be any continuous monotonic utility function. In particular, an adult may well prefer that her child functions at a higher level than she, in the sense that, for all X and small $\delta > 0$, $u(X - \delta, X + \delta) > u(X, X)$.[6] This is perhaps somewhat surprising: even if adults *want* their children to function at a higher level than themselves, there is no protracted human development in the optimum.

5. Equality of opportunity for welfare: Model III.

We now suppose that adults care about their own level of functioning and their child's *utility.* Suppose there is a concept of utility such that

$$u_J^t = F^t + \beta u_J^{t+1} \quad \forall t \geq 0; J = L, H \qquad (5.1)$$

Hence, we can write u_J^0 recursively as

$$u_J^0 = \sum_{t=0}^{N} (\beta)^t F_J^t + (\beta)^{N+1} u_J^{N+1}, \text{ for any } N; J = L, H$$

Suppose that the discounted sum of functioning levels of this dynasty is bounded above, i.e., $\sum_{t=0}^{\infty} \beta^t F_J^t$ is bounded above for every feasible sequence $\{w^t_H, w^t_L\}_{t=0,1,2\ldots}$, given

[6] Notice that this limited form of altruism is consistent with A3.

($w^0{}_H$, $w^0{}_L$), and therefore $\beta < 1$.[7] Hence, without loss of generality, we can assume $\lim_{N\to\infty} \beta^{N+1} u^{N+1} = 0$, and the utility of any adult born in period $i \geq 0$ is:

$$u_J^i = \sum_{t=i}^{\infty} \beta^{t-i} F_J^t, \qquad \forall i \geq 0; J = L, H \tag{5.2}$$

Thus, the utility of any adult is the discounted sum of her dynasty's levels of functioning. Caring about the *welfare* of your child forces you, implicitly, to care about the *functioning* of your descendents, all the way down. It is reasonable to suppose that this formulation is psychologically accurate. Are we parents content if our children are functioning well, or does our contentment depend upon their *happiness,* where their happiness derives from the happiness of *their* children?

Our equal-opportunity-for-welfare program is stated again as (4.1), where the notation now refers to the new concept of utility. Again, the value of program (4.1) is achieved at the date 0 utility. (If it weren't, increase τ^0,[8] which will increase F_L^t, and thus u_L^t, $\forall t \geq 1$.) Consequently, program (4.1) is equivalent to the program:

$$\begin{aligned} &\sup u_L^0 \\ &\text{s.t. } w_H^t \geq w_L^t, \forall t \end{aligned} \tag{5.3}$$

Clearly, at the solution to (5.3), we have $u_L^0 \leq u_L^t$, $\forall t$. Assume:

Assumption 1': $h(x) = kx^{c_1}$, $g(w) = w^{c_2}$, where $k > 0$, $0 \leq c_1, c_2 \leq 1$.

Moreover, let A2' hold. The *sequence problem* (SP) can be written as

$$v^*(w_L^0, w_H^0) = \sup \sum_{t=0}^{\infty} \beta^t \{\gamma \log w_L^t + (1-\gamma)\log(1-\tau^t) + (1-\gamma)\log[f_L w_L^t + f_H w_H^t]\}$$
$$s.t. w_L^{t+1} = k(\tau^t)^{c_1}(r_L^t)^{c_1}(w_L^t)^{c_2}$$
$$w_H^{t+1} = k(\tau^t)^{c_1}(r_H^t)^{c_1}(w_H^t)^{c_2} = k(\tau^t)^{c_1}\frac{1}{(f_H)^{c_1}}(1-f_L r_L^t)^{c_1}(w_H^t)^{c_2}$$
$$w_H^t \geq w_L^t, \forall t$$

As we will show, the solution to SP will depend on the initial wage ratio, $\rho^0 \equiv w_H^0/w_L^0$. Therefore, let us first solve the *single-wage* SP, i.e. the SP for $\rho^0 = 1$.

---

[7] In what follows, by A1' and A4, F is bounded above so that such a condition is satisfied.

[8] In L1 below we prove that if F is bounded above, in the solution to (4.1), it will be $\tau^0 < 1$.

$$v^*(w^0) = \sup \sum_{t=0}^{\infty} \beta^t \left[\log w^t + (1-\gamma)\log(1-\tau^t)\right]$$
$$s.t. w^{t+1} = k(\tau^t)^{c_1}(w^t)^{c_2} \tag{5.4}$$

where the only instruments are tax rates - all children receive an equal per capita share of educational investment - and v^* denotes the supremum function.

In order to analyze the single-wage SP, let $W \subseteq \mathbf{R}_+$ denote the state space and let Γ: $W \rightarrow W$ denote the feasibility correspondence, where $\Gamma(w) = [0, k(w)^{c_2}]$, and thus $\Gamma(w) \neq \varnothing$, $\forall w$. Next, let $A = \{(w, y) \in W \times W | y \in \Gamma(w)\}$ be the graph of Γ. The one period return function at date t is a function Φ: $A \rightarrow \mathbf{R'}$ whose value is $F(w_L^t, (1 - \tau^t)\mu^t)$ but where τ^t is expressed as a function of (w^t, w^{t+1}). By substituting for τ^t, $\Phi(w^t, w^{t+1})$ = $\log w^t + (1-\gamma)\log\left(1 - \frac{1}{k^{1/c_1}} \frac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}\right)$ and SP can be written as

$$v^*(w^0) = \sup \sum_{t=0}^{\infty} \beta^t \left[\log w^t + (1-\gamma)\log\left(1 - \frac{1}{k^{1/c_1}} \frac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}\right)\right]$$
$$w^{t+1} \in [0, k(w^t)^{c_2}] \tag{5.5}$$

By A1', $w = k^{\frac{1}{1-c_2}}$ is the highest sustainable value of w, and therefore, without loss of generality, we restrict the analysis to a subset, $W' \subset W$, $W' = \{w \in W | w \leq w', w' > k^{\frac{1}{1-c_2}}\}$, w' finite.[9] Hence, Φ is bounded above by $\Phi(w', 0)$ and for all $w^0 \in W'$ and all feasible sequences $\{w^t\}_{t=0,1,\ldots}$, $\lim_{N\rightarrow\infty} \sum_{t=0}^{N} \beta^t \Phi(w^t, w^{t+1})$ exists in $\mathbf{R} \cup \{-\infty\}$ and (5.5) is well defined. Moreover, as shown in Appendix 1, if $c_1 + c_2 \leq 1$, Φ is strictly concave. Thus, we henceforth assume:

Assumption 4: $c_1 + c_2 \leq 1$.

Bellman's functional equation (FE) can be written as

$$v(w^0) = \sup_{w^1 \in [0, k(w^0)^{c_2}]} \left[\log w^0 + (1-\gamma)\log\left(1 - \frac{(w^1)^{1/c_1}}{k^{1/c_1}(w^0)^{c_2/c_1}}\right) + \beta v(w^1)\right] \tag{5.6}$$

[9] If $c_2 = 1$, the state space is not bounded. However, by A3, $c_2 = 1$ implies $c_1 = 0$: education plays no role, the optimal tax rate is zero and, actually, there is no genuine dynamic decision.

where $v(w)$ denotes the solution to FE. We now prove that the function $v(w) = \phi + \psi \log w$, where ϕ and ψ are unknown constants to be determined, solves (5.6):

$$v(w^0) = \sup_{w^1 \in [0, k(w^0)^{c_2}]} \left[\log w^0 + (1-\gamma)\log\left(1 - \frac{(w^1)^{1/c_1}}{k^{1/c_1}(w^0)^{c_2/c_1}}\right) + \beta\phi + \beta\psi \log w^1 \right]$$

The first order condition for this problem is

$$\frac{(1-\gamma)}{c_1} \frac{(w^1)^{1/c_1}}{k^{1/c_1}(w^0)^{c_2/c_1}} = \beta\psi\left(1 - \frac{(w^1)^{1/c_1}}{k^{1/c_1}(w^0)^{c_2/c_1}}\right)$$

and therefore

$$w^1 = k\left(\frac{\beta\psi c_1}{(1-\gamma)+\beta\psi c_1}\right)^{c_1}(w^0)^{c_2}$$

The postulated function solves FE if

$$\phi + \psi \log w^0 = \beta\phi + (1+\beta\psi c_2)\log w^0 + (1-\gamma)\log\frac{(1-\gamma)}{(1-\gamma)+\beta\psi c_1} + \beta\psi \log k + \beta\psi c_1 \log\frac{\beta\psi c_1}{(1-\gamma)+\beta\psi c_1}$$

or, by the method of undetermined coefficients,

$$\psi = \frac{1}{(1-\beta c_2)}$$

$$\phi = \frac{(1-\gamma)}{(1-\beta)}\log\frac{(1-\gamma)(1-\beta c_2)}{(1-\gamma)(1-\beta c_2)+\beta c_1} + \frac{\beta}{(1-\beta)(1-\beta c_2)}\log k + \frac{\beta c_1}{(1-\beta)(1-\beta c_2)}\log\frac{\beta c_1}{(1-\gamma)(1-\beta c_2)+\beta c_1} \quad (5.7)$$

We now have:[10]

Proposition 3. Let A1', A2', A4 hold. Let $w_L^0 = w_H^0 = w^0$, then $v^*(w^0) = \phi + [1/(1 - \beta c_2)]\log w^0$ solves (5.5), where ϕ is given by (5.7). The optimal policy is given by

$$w^{*t+1} = k\left(\frac{\beta c_1}{(1-\gamma)(1-\beta c_2)+\beta c_1}\right)^{c_1}(w^{*t})^{c_2}$$

Proof. 1. Notice that $\lim\sup_{t\to\infty} \beta^t v(w^t) = \lim\sup_{t\to\infty} \beta^t \{\phi + [1/(1 - \beta c_2)]\log w^t\}$. Clearly, $\lim_{t\to\infty} \beta^t \phi = 0$. Moreover, $[1/(1 - \beta c_2)] \lim\sup_{t\to\infty} \beta^t \log w^t \leq [1/(1 - \beta c_2)] \lim_{t\to\infty} \beta^t (\log k^t w^0) = [1/(1 - \beta c_2)] \lim_{t\to\infty} \beta^t [\log k^t + \log w^0]$. Given that $\lim_{t\to\infty} \beta^t t \log k = 0$, it follows that $\lim\sup_{t\to\infty} \beta^t v(w^t) = 0$, for all feasible sequences $\{w^t\}_{t=0,1,\ldots}$.

[10] An alternative proof of P3 is provided in App.1, based on the Euler equations.

Next, as shown in App.1, the sequence $\{w^{*t}\}_{t=0,1,\ldots}$, is such that (a) for every feasible sequence $\{w^t\}_{t=0,1,\ldots}$, $\sum_{t=0}^{\infty} \beta^t \Phi(w^{*t}, w^{*t+1}) \geq \sum_{t=0}^{\infty} \beta^t \Phi(w^t, w^{t+1})$ and (b) $\lim_{t\to\infty} \beta^t v(w^{*t}) = 0$. Hence, by the theorems on dynamic optimization (see e.g. Stokey and Lucas, 1989, pp.72-5), $v(w^0) = v^*(w^0)$.

2. The second part of the proposition is an immediate consequence of the first.▯

As concerns the relationship between equality and growth:[11]

Corollary 1. Let A1', A2', A4 hold. In an egalitarian economy with $w_L^0 = w_H^0$, the optimal wage eventually converges to $\overline{w} = (k)^{\frac{1}{1-c_2}}\left(\frac{\beta c_1}{(1-\gamma)(1-\beta c_2)+\beta c_1}\right)^{\frac{c_1}{1-c_2}}$

We now proceed to study the program (5.4) when $w_L^0 \neq w_H^0$, i.e. $\rho^0 > 1$. Let now $W \subseteq \mathbf{R}^2$ denote the state space, with generic element $w = (w_L, w_H)$ and let Γ: $W \to W$ denote the feasibility correspondence, where now

$$\Gamma(w) = \left\{\hat{w} \in W | \exists 0 \leq \tau \leq 1, \exists (r_L, r_H): f_L r_L + f_H r_H = 1, \hat{w}_L \leq k(\tau r_L)^{c_1}(w_L)^{c_2}, \hat{w}_H \leq k(\tau r_H)^{c_1}(w_H)^{c_2}\right\}$$

so that $\Gamma(w) \neq \varnothing$, $\forall w$. The one-period return function $\Phi(w_L^t, w_H^t, w_L^{t+1}, w_H^{t+1})$ is

$$\gamma \log w_L^t + (1-\gamma)\log\left(1 - \frac{1}{k^{1/c_1}}\left[f_H \frac{(w_H^{t+1})^{1/c_1}}{(w_H^t)^{c_2/c_1}} + f_L \frac{(w_L^{t+1})^{1/c_1}}{(w_L^t)^{c_2/c_1}}\right]\right) + (1-\gamma)\log[f_L w_L^t + f_H w_H^t]$$

and if $v^*(w_L^0, w_H^0)$ denotes the supremum function, we can write SP as

$$v^*(w_L^0, w_H^0) = \sup \sum_{t=0}^{\infty} \beta^t \Phi(w_L^t, w_H^t, w_L^{t+1}, w_H^{t+1})$$
$$(w_L^{t+1}, w_H^{t+1}) \in \Gamma(w_L^t, w_H^t) \qquad (5.8)$$

Again, define a vector $w' = (w_L', w_H')$, with w_L' and w_H' finite and such that $w_H' > \frac{1}{(f_H)^{c_1/(1-c_2)}}(k)^{1/(1-c_2)}$ and $w_L' > \frac{1}{(f_L)^{c_1/(1-c_2)}}(k)^{1/(1-c_2)}$. Without loss of generality, we restrict the analysis to the subset $W' = \{w \in W | w \leq w'\}$. Hence, Φ is bounded above by $\Phi(w_L', w'_H, 0, 0)$ and for all $w^0 \in W'$ and all feasible sequences $\{w^t\}_{t=0,1,\ldots}$,

[11] If $c_2 = 1$ we get unbounded growth (provided $k \geq 1$). However, as argued in fn. 9, this case can be ruled out.

$\lim_{N\to\infty}\sum_{t=0}^{N}\beta^t\Phi(w_L^t, w_H^t, w_L^{t+1}, w_H^{t+1})$ exists in **R** $\cup$ {-∞} and (5.8) is well defined. Moreover, Φ is differentiable in all arguments and, by A4, strictly concave.

We now prove:

Proposition 4. $v^*(w_L^0, w_H^0)$ is increasing in both arguments. Moreover,

$$v^*(w_L^0, w_H^0) \geq \phi + \gamma\log w_L^0 + (1-\gamma)\log[f_L w_L^0 + f_H w_H^0] + \frac{\beta c_1}{(1-\beta c_2)}\log\frac{(w_L^0)^{c_2/c_1}(w_H^0)^{c_2/c_1}}{[f_L(w_L^0)^{c_2/c_1} + f_H(w_H^0)^{c_2/c_1}]}$$

$$v^*(w_L^0, w_H^0) \leq \phi + \frac{1}{(1-\beta c_2)}\log w_H^0$$

where ϕ is given by (5.7).

Proof. 1. Let $\{w_L^t, w_H^t\}_{t=0,1,\dots}$ denote a feasible path of the states, given initial conditions (w_L^0, w_H^0). If the initial conditions are (w'^0_L, w_H^0), $w'^0_L > w_L^0$, the path $\{w'^t_L, w'^t_H\}_{t=0,1,\dots}$ such that $w'^t_L = w_L^t$, $\forall\ t \geq 1$, $w'^t_H = w_H^t$, $\forall\ t \geq 0$, is clearly feasible with $\sum_{t=0}^{\infty}\beta^t\Phi(w'^t_L, w'^{t+1}_L, w'^t_H, w'^{t+1}_H) \geq \sum_{t=0}^{\infty}\beta^t\Phi(w_L^t, w_L^{t+1}, w_H^t, w_H^{t+1})$, and since this is true for every feasible path, $v^*(w'^0_L, w_H^0) \geq v^*(w_L^0, w_H^0)$.

2. Firstly, notice that, by P3 and the monotonicity of v^*, it follows that $v^*(w_H^0, w_H^0) = \phi + 1/(1 - \beta c_2)\log w_H^0 \geq v^*(w_L^0, w_H^0)$. Secondly, notice that it is always feasible to equalize the wages in $t = 1$, i.e. to set $r_L^0 = \frac{(w_H^0)^{c_2/c_1}}{f_L(w_H^0)^{c_2/c_1} + f_H(w_L^0)^{c_2/c_1}}$. Hence, by P3,

$$v^*(w_L^0, w_H^0) \geq \sup_{\tau^0\in[0,1]}\gamma\log w_L^0 + (1-\gamma)\log(1-\tau^0) + (1-\gamma)\log[f_L w_L^0 + f_H w_H^0] + \beta\phi + \frac{\beta}{1-\beta c_2}\log k$$
$$+\frac{\beta c_1}{1-\beta c_2}\log\tau^0 + \frac{\beta c_1}{1-\beta c_2}\log\frac{(w_H^0)^{c_2/c_1}}{f_L(w_H^0)^{c_2/c_1} + f_H(w_L^0)^{c_2/c_1}} + \frac{\beta c_2}{1-\beta c_2}\log w_L^0$$

and maximizing the right hand side with respect to τ^0 the result follows. ▯

As concerns the optimal path of the controls, we now prove

Lemma 1. Let A1', A2' hold. For any finite t, in the optimum, $r_L^t > 0$ and $0 < \tau^t < 1$.

Proof. By A2', $\lim_{w_L^t\to 0} F(w_L^t, (1 - \tau^t)[f_L w_L^t + f_H w_H^t]) = -\infty$, $\forall\tau^t \in [0, 1]$, while, by A1', $r_L^t = 0$ or $\tau^t = 0$ imply $w_L^j = 0$, $\forall\ j \geq t + 1$, and hence in the optimum $r_L^t > 0$ and $\tau^t > 0$. Given the boundedness of F, a similar argument can be used to prove that $\tau^t < 1$.▯

Our strategy to solve (5.8) will be to recursively construct a function $v(w_L^0, w_H^0)$ that solves FE; then we will prove that $v(w_L^0, w_H^0) = v^*(w_L^0, w_H^0)$. As a first step, consider the Euler equations. Given the inequality constraint on wages, it is more convenient to write the one-period return as a function Ψ of w_L and the wage ratio $\rho \equiv (w_H/w_L)$.[12] Thus, $\Psi(w_L^t, \rho^t, w_L^{t+1}, \rho^{t+1})$ can be written as

$$\log w_L^t + (1-\gamma)\log\left(1 - \frac{1}{k^{1/c_1}}\frac{(w_L^{t+1})^{1/c_1}}{(w_L^t)^{c_2/c_1}}\left[f_H\frac{(\rho^{t+1})^{1/c_1}}{(\rho^t)^{c_2/c_1}} + f_L\right]\right) + (1-\gamma)\log[f_L + f_H\rho^t]$$

Thus, in an interior solution, the Euler Equations and can be written as

$$\frac{(1-\gamma)}{c_1}\frac{(w_L^{t+1})^{1/c_1}}{k^{1/c_1}(w_L^t)^{c_2/c_1}}\left[\frac{f_H(\rho^{t+1})^{1/c_1}}{(\rho^t)^{c_2/c_1}} + f_L\right]\Bigg/\left[1 - \frac{(w_L^{t+1})^{1/c_1}}{k^{1/c_1}(w_L^t)^{c_2/c_1}}\left[\frac{f_H(\rho^{t+1})^{1/c_1}}{(\rho^t)^{c_2/c_1}} + f_L\right]\right]$$

$$\beta + \frac{\beta(1-\gamma)c_2}{c_1}\frac{(w_L^{t+2})^{1/c_1}}{k^{1/c_1}(w_L^{t+1})^{c_2/c_1}}\left[\frac{f_H(\rho^{t+2})^{1/c_1}}{(\rho^{t+1})^{c_2/c_1}} + f_L\right]\Bigg/\left[1 - \frac{(w_L^{t+2})^{1/c_1}}{k^{1/c_1}(w_L^{t+1})^{c_2/c_1}}\left[\frac{f_H(\rho^{t+2})^{1/c_1}}{(\rho^{t+1})^{c_2/c_1}} + f_L\right]\right]$$

$$\frac{1}{c_1}\frac{(w_L^{t+1})^{1/c_1}}{k^{1/c_1}(w_L^t)^{c_2/c_1}}\left[\frac{f_H(\rho^{t+1})^{1/c_1}}{(\rho^t)^{c_2/c_1}}\right]\Bigg/\left[1 - \frac{(w_L^{t+1})^{1/c_1}}{k^{1/c_1}(w_L^t)^{c_2/c_1}}\left[\frac{f_H(\rho^{t+1})^{1/c_1}}{(\rho^t)^{c_2/c_1}} + f_L\right]\right]$$

$$\frac{\beta f_H}{f_L + f_H\rho^{t+1}} + \frac{\beta c_2}{c_1}\frac{(w_L^{t+2})^{1/c_1}}{k^{1/c_1}(w_L^{t+1})^{c_2/c_1}}\left[\frac{f_H(\rho^{t+2})^{1/c_1}}{(\rho^{t+1})^{c_2/c_1}}\right]\Bigg/\left[1 - \frac{(w_L^{t+2})^{1/c_1}}{k^{1/c_1}(w_L^{t+1})^{c_2/c_1}}\left[\frac{f_H(\rho^{t+2})^{1/c_1}}{(\rho^{t+1})^{c_2/c_1}} + f_L\right]\right]$$

Actually, the Euler equations can be re-written, and made more intelligible, in terms of the controls τ^t and r_L^t:

$$(1-\gamma)\frac{\tau^t}{1-\tau^t} = \beta c_1 + \beta c_2(1-\gamma)\frac{\tau^{t+1}}{1-\tau^{t+1}} \tag{5.9}$$

$$\frac{\tau^t}{1-\tau^t}(1 - f_L r_L^t) = \frac{\beta c_1 (f_H)^{1-c_1}(1 - f_L r_L^t)^{c_1}(\rho^t)^{c_2}}{f_L(r_L^t)^{c_1} + (f_H)^{1-c_1}(1 - f_L r_L^t)^{c_1}(\rho^t)^{c_2}} + \beta c_2\frac{\tau^{t+1}}{1-\tau^{t+1}}(1 - f_L r_L^{t+1}) \tag{5.10}$$

Consider $t = 0$. Given that $\rho^0 = 1 \Rightarrow \rho^t = 1$, $\forall\ t > 0$, we conjecture that there exists a number $\bar{\rho}_0 \geq 1$ such that if $\rho^0 \in [1, \bar{\rho}_0]$ then in the optimum $\rho^1 = 1$, and thus $\rho^t = 1$, $\forall\ t > 1$. From the dynamic constraints it is possible to express r_L^t as a function of ρ^t and ρ^{t+1}. Thus, substituting for r_L^0 and r_L^1 in (5.10) and setting $\rho^1 = \rho^2 = 1$, a necessary condition for, $\rho^1 = 1$ to be optimal is

[12] The FOCs deriving from the maximization of $\Psi(w_L^t, w_L^{t+1}, \rho^t, \rho^{t+1}) + \beta\,\Psi(w_L^{t+1}, w_L^{t+2}, \rho^{t+1}, \rho^{t+2})$ subject to $\rho^{t+1} \geq 1$ are the same as the FOCs deriving from the maximization of $\Phi(w_L^t, w_L^{t+1}, w_H^t, w_H^{t+1}) + \beta\,\Phi(w_L^{t+1}, w_L^{t+2}, w_H^{t+1}, w_H^{t+2})$ subject to $w_H^{t+1} \geq w_L^{t+1}$.

$$-\frac{\tau^0}{1-\tau^0}\frac{f_H}{f_H+f_L(\rho^0)^{c_2}}+\beta c_1 f_H+\beta c_2\frac{\tau^1}{1-\tau^1}f_H\le 0 \quad (5.10')$$

Next, by P3, if $\rho^1 = 1$ then in the optimum $\tau^1 = \tau^* \equiv \beta c_1/[(1 - \gamma)(1 - \beta c_2) + \beta c_1]$. Hence, by (5.9) $\tau^0 = \tau^*$ and (5.10') becomes:

$$\rho^0 \le \left[\frac{\gamma(1-\beta c_2)}{f_L(1-\beta c_2)(1-\gamma)+\beta c_2}+1\right]^{\frac{c_1}{c_2}} \equiv \bar{\rho}_0 \quad (5.10'')$$

Given the parameter restrictions, $\bar{\rho}_0 > 1$, moreover $\bar{\rho}_0$ is higher, the higher γ and the lower f_L, β, and c_2. We can now prove:[13]

Proposition 5. Let A1', A2', A4 hold. If $\rho^0 \in [1, \bar{\rho}_0]$ then in the optimum $\tau^t = \tau^* = \beta c_1/[(1 - \gamma)(1 - \beta c_2) + \beta c_1]$, $\forall\, t$, and $\rho^t = 1$, $\forall\, t \ge 1$.

Proof. Let Δ denote the difference between the objective function evaluated at $\{w^{*t}_L, w^{*t}_H\}_{t=0,1,\dots}$, the path of the two states in the proposed solution, and at $\{w_L^t, w_H^t\}_{t=0,1,\dots}$, any feasible path. Let $\Phi_{w_J^i}(w_L^t, w_H^t, w_L^{t+1}, w_H^{t+1}) \equiv \partial\Phi/\partial w_J^i$, $J = L, H$, $i = t, t + 1$. By the concavity of Φ

$$\Delta = \lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[\Phi(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)-\Phi(w^t_L, w^{t+1}_L, w^t_H, w^{t+1}_H)\right]\ge$$
$$\lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[\Phi_{w_L^t}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)(w^{*t}_L-w^t_L)+\Phi_{w_L^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)(w^{*t+1}_L-w^{t+1}_L)\right]+$$
$$\lim_{T\to\infty}\sum_{t=0}^{T}\beta^t\left[\Phi_{w_H^t}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)(w^{*t}_H-w^t_H)+\Phi_{w_H^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)(w^{*t+1}_H-w^{t+1}_H)\right]$$

In the proposed solution

$$\Phi_{w_L^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)+\beta\Phi_{w_L^{t+1}}(w^{*t+1}_L, w^{*t+2}_L, w^{*t+1}_H, w^{*t+2}_H)\ge 0$$
$$\Phi_{w_H^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)+\beta\Phi_{w_H^{t+1}}(w^{*t+1}_L, w^{*t+2}_L, w^{*t+1}_H, w^{*t+2}_H)\le 0$$
$$\Phi_{w_L^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)+\beta\Phi_{w_L^{t+1}}(w^{*t+1}_L, w^{*t+2}_L, w^{*t+1}_H, w^{*t+2}_H)=$$
$$-\Phi_{w_H^{t+1}}(w^{*t}_L, w^{*t+1}_L, w^{*t}_H, w^{*t+1}_H)-\beta\Phi_{w_H^{t+1}}(w^{*t+1}_L, w^{*t+2}_L, w^{*t+1}_H, w^{*t+2}_H)$$

and $w_L^{*t} = w_H^{*t}$, $\forall t$. Therefore since $w_L^{*0} = w_L^0$, $w_H^{*0} = w_H^0$,

[13] We adapt the proof of Thm.4.15 (Stokey and Lucas, 1989, p.98) to the case of a corner solution.

$$\Delta \geq \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \left[\Phi_{w_L^{t+1}}(w*_L^t, w*_L^{t+1}, w*_H^t, w*_H^{t+1}) + \beta\Phi_{w_L^{t+1}}(w*_L^{t+1}, w*_L^{t+2}, w*_H^{t+1}, w*_H^{t+2})\right](w_H^{t+1} - w_L^{t+1})$$
$$+\lim_{T\to\infty} \beta^T \Phi_{w_L^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})(w*_L^{T+1} - w_L^{T+1})$$
$$+\lim_{T\to\infty} \beta^T \Phi_{w_H^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})(w*_H^{T+1} - w_H^{T+1})$$

Hence, given $w_H^{t+1} \geq w_L^{t+1}$, $\forall t$, the first term on the right hand side is non-negative. Moreover, since $\Phi_{w_J^{T+1}}(w_L^T, w_L^{T+1}, w_H^T, w_H^{T+1}) \leq 0$, $J = L, H$ and

$$\Phi_{w_L^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})w*_L^{T+1} = -f_L \frac{\beta}{(1-\beta c_2)}$$
$$\Phi_{w_H^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})w*_H^{T+1} = -f_H \frac{\beta}{(1-\beta c_2)}$$

it follows that,

$$\lim_{T\to\infty} \beta^T \Phi_{w_L^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})(w*_L^{T+1} - w_L^{T+1}) \geq$$
$$\lim_{T\to\infty} \beta^T \Phi_{w_L^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})w*_L^{T+1} = 0$$
$$\lim_{T\to\infty} \beta^T \Phi_{w_H^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})(w*_H^{T+1} - w_H^{T+1}) \geq$$
$$\lim_{T\to\infty} \beta^T \Phi_{w_H^{T+1}}(w*_L^T, w*_L^{T+1}, w*_H^T, w*_H^{T+1})w*_H^{T+1} = 0$$

and therefore, $\Delta \geq 0$. ▯

Hence, if $\rho^0 \in [1, \bar{\rho}_0]$, define v_0: $W' \to \mathbf{R}$ as

$$v_0(w_L^0, w_H^0) = \phi + \gamma \log w_L^0 + (1-\gamma)\log(f_L w_L^0 + f_H w_H^0) + \frac{\beta c_1}{1-\beta c_2} \log \frac{(w_H^0)^{c_2/c_1}(w_L^0)^{c_2/c_1}}{f_L(w_H^0)^{c_2/c_1} + f_H(w_L^0)^{c_2/c_1}}$$

where ϕ is given by (5.7). Moreover, let $\tau_0 : [1, \bar{\rho}_0] \to [0,1]$ and $r_0 : [1, \bar{\rho}_0] \to \mathbf{R}_+$ denote the optimal control functions, where $\tau_0(\rho^0) = \tau*$, while $r_0(\rho^0) = \frac{(\rho^0)^{c_2/c_1}}{f_H + f_L(\rho^0)^{c_2/c_1}}$ and therefore the optimal redistributive policy depends only on the wage *ratio*, ρ^0.[14]

As regards v_0, it is continuously differentiable in both variables with

$$\frac{\partial v_0(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\gamma}{w_L^0} + \frac{(1-\gamma)f_L}{[f_L w_L^0 + f_H w_H^0]}$$
$$+\frac{(1-\gamma)c_2}{w_L^0 c_1} \frac{f_L[\varpi_{0,L}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1}(w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}}\left(\frac{f_H[\varpi_{0,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L[\varpi_{0,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}}\right)\right]$$

[14] Notice that r_0 is the optimal r_L^0. The optimal r_H^0 can be derived from the constraint $f_L r_0 + f_H r_H^0 = 1$.

$$\frac{\partial v_0(w_L^0, w_H^0)}{\partial w_H^0} = \frac{(1-\gamma) f_H}{[f_L w_L^0 + f_H w_H^0]}$$

$$+ \frac{(1-\gamma)c_2}{w_H^0 c_1} \frac{f_H[\varpi_{0,H}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1}(w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}}\left(\frac{f_H[\varpi_{0,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L[\varpi_{0,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}}\right)\right]$$

where $\varpi_{0,L}$: $W' \to \mathbf{R}_+$ and $\varpi_{0,H}$: $W' \to \mathbf{R}_+$ are the optimal wage functions, if $\rho^0 \in [1, \bar{\rho}_0]$ (they can be derived from τ_0 and r_0), and thus $\varpi_{0,L}(w_L^0, w_H^0)$ and $\varpi_{0,H}(w_L^0, w_H^0)$ denote the optimal values of w_L^1 and w_H^1, respectively. More explicitly,

$$\frac{\partial v_0(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\gamma}{w_L^0} + \frac{(1-\gamma) f_L}{[f_L w_L^0 + f_H w_H^0]} + \frac{\beta c_2}{1-\beta c_2} \frac{1}{w_L^0} \frac{f_L (w_H^0)^{c_2/c_1}}{f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}}$$

$$\frac{\partial v_0(w_L^0, w_H^0)}{\partial w_H^0} = \frac{(1-\gamma) f_H}{[f_L w_L^0 + f_H w_H^0]} + \frac{\beta c_2}{1-\beta c_2} \frac{1}{w_H^0} \frac{f_H (w_L^0)^{c_2/c_1}}{f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}}$$

Hence, v_0 is strictly increasing in both variables and, as shown in Appendix 2, strictly concave. Moreover, it is straightforward to show that if $\rho^0 \in [1, \bar{\rho}_0]$ v_0 solves[15]

$$v_0(w_L^0, w_H^0) = \max_{w^1 \in \Gamma(w^0)} [\Phi(w_L^0, w_H^0, w_L^1, w_H^1) + \beta v_0(w_L^1, w_H^1)]$$

at the corner solution $\rho^1 = 1$, and $\tau^1 = \tau^*$. Assuming v_0 to be the value function if $\rho^0 \in [1, \bar{\rho}_0]$, in Appendix 3 we prove that there exists an interval $(\bar{\rho}_0, \bar{\rho}_1]$ such that if $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, it is optimal to set $\rho^1 \in [1, \bar{\rho}_0]$, and therefore $\rho^j = 1$, $j \geq 2$. Thus, if $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, where $\varpi_{1,L}$: $W' \to \mathbf{R}_+$ and $\varpi_{1,H}$: $W' \to \mathbf{R}_+$ denote the optimal wage functions, and define $v_1 : W' \to \mathbf{R}$ as

$$v_1(w_L^0, w_H^0) = \gamma \log w_L^0 + (1-\gamma)\log(f_L w_L^0 + f_H w_H^0) + (1-\gamma)\log(1-\tau_1) + \beta v_0(\varpi_{1,L}(w_L^1, w_H^1), \varpi_{1,H}(w_L^1, w_H^1))$$

and continue the iterative procedure, assuming that v_1 is the value function on $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, and verifying that there exists a $\bar{\rho}_2$ such that, if $\rho^0 \in (\bar{\rho}_1, \bar{\rho}_2]$, it is optimal to set $\rho^1 \in (\bar{\rho}_0, \bar{\rho}_1]$, $\rho^2 \in [1, \bar{\rho}_0]$ and $\rho^j = 1$, $j \geq 3$. In general, in App. 3, we prove that in the solution to FE, there exist derive an infinite sequence of intervals $(\bar{\rho}_{k-1}, \bar{\rho}_k]$ such that, if $\rho^0 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$, then $\rho^{k+1} = 1$, $k \geq 0$ (if $k = 0$, $\bar{\rho}_{k-1} = 1$).

We can now prove:

Theorem 1. Let A1', A2', A4 hold. Consider an inegalitarian economy in which $w_L^0 \neq w_H^0$. Let $\rho^t \equiv w_H^t / w_L^t$. For any finite ρ^0, in the solution to the program (5.3), equality is reached in a

[15] We henceforth use the "max" notation because, as we shall see, the supremum is actually attained.

finite number of periods. Once equality is reached wages grow forever as described in P3 and converge in the limit to the steady state $\overline{w}$ in C1.

Proof. We need to show that the increasing function $v(w_L^0, w_H^0)$ solving Bellman's FE, obtained in Appendix 3 is the value function.

Firstly, by the monotonicity of v, $v(w_L^t, w_H^t) \leq v(w_L', w_H')$, $\forall (w_L^t, w_H^t) \in W'$. Hence, $\limsup_{t\to\infty} \beta^t v(w_L^t, w_H^t) \leq \lim_{t\to\infty} \beta^t v(w_L', w_H') = 0$. Next, by L1 and P5, the optimal sequence $\{w*_L^t, w*_H^t\}_{t=0,1\ldots}$ is bounded away from zero. Hence, by the monotonicity of v, $\lim_{t\to\infty} \beta^t v(w*_L^t, w*_H^t) = 0$, and, by the theorems on recursive dynamic optimization, $v(w_L^0, w_H^0) = v*(w_L^0, w_H^0)$ and the policies derived from v in App. 3 are indeed optimal.▯

In other words, the optimal path involves equating the wages of the contemporaneous members of the two dynasties in a finite number of periods: if $\rho^0 \in (\overline{\rho}_{k-1}, \overline{\rho}_k]$, convergence occurs in $k + 1$ periods. Once equality is reached, human development continues forever.

5. Conclusion.

Earlier, we remarked on the similarity between the present paper, Arrow (1973) and Dasgupta (1974) (A-D, henceforth). The main differences between A-D's models and ours are: (1) A-D have a representative agent each period, and so the only issue is to maximin welfare of that agent's descendents across time, whereas in the present model, there is an issue of *intragenerational* as well as *intergenerational* justice; (2) in A-D, agents care only about consumption, not about functioning (i.e., not about the wage per se); (3) in A-D, investment is modeled as capital investment, rather than educational investment. Mathematically, the main difference is that the planner has only one instrument each period in A-D, whereas in our model, she has two instruments. (This is, of course, due to difference (1) above.) Nevertheless, A-D's results are qualitatively similar to ours: an increase in consumption over time is compatible with maximin only if the equalisandum is welfare, in which case parents care about the consumption stream of their entire dynasty. Thus, the present paper may be considered an intellectual descendent of Arrow (1973) and Dasgupta (1974).

Our concern with intragenerational inequality, not expressed in the earlier literature, led us to deduce that, as long as individuals value their human capital as well as their consumption, then the maximin program will eventually equalize the levels of human capital of all individuals. We remark, however, that this result may well depend on our assumption A4, of non-increasing returns in the educational technology.

Let us recapitulate. One of the major foci of discussion in egalitarian theory of the last thirty years has been the nature of the equalisandum. The main participants in the discussion have moved away from taking welfare as that equalisandum, although it is important to note that Arneson (1989) has argued for choosing *opportunity for welfare* as the equalisandum. ('Opportunity for welfare' is, in general, quite different from 'welfare' as an equalisandum. That difference is due to differential effort, which in the present article, does not appear.) However, this debate has been carried out within the confines of a static environment, a 'model' with a single generation. Here, we have maintained that equality of opportunity, for whatever kind of condition, is an ethically viable conception in a multi-generation world, and that in such a context, it calls for equalizing opportunities across all types of adult, where an adult's type is characterized by the date at which he is born and the SES of the family in which he grew up. It is beyond this article's scope to *argue* that justice requires that a person fare no better than another simply by virtue of being born at a different date[16]. An asymmetric version of this principle is familiar in discussions of sustainable development and environmental preservation: we should leave to future generations a world as bountiful as the one left to us by our ancestors. But the other part is, we believe, just as compelling: we are under no ethical mandate to leave our descendents a world more bountiful than our own, *although we may decide to do so if that increases our welfare* by contemplating the happiness it will bring our children, and their children…

In studying the multi-generation world, we have learned that, if we choose what we call an objectivist equalisandum - we have taken 'functioning' as an appealing one - then equality of opportunity for that condition implies there will be no further human development, where human development is conceived of not as an increase in human welfare, but rather in human capacities to function. Thus, two major characteristics of what comprises the good society, as it has been conceived of by egalitarians for several hundred years, are incompatible. We showed that if we equalize opportunities for welfare, where an

[16] This is contestable. Some argue that equality of condition among living persons is all that an egalitarian ethic requires. One rationale is that self-esteem is affected by comparing one's condition to those of contemporaries, not to the dead, or to those not yet born.

adult's welfare depends upon her own level of functioning and the functioning levels of a finite stream of her descendents, the unpleasant inconsistency continues to hold. If, however, we choose a thorough-going kind of welfare as the condition for which opportunities should be equalized - one which declares that an individual's welfare depends not just on his capacities and the capacities of his children, but rather on his own capacities and his child's *welfare* - then human development and equality of opportunity are mutually consistent.

The most appealing solution to the unpleasant inconsistency is, we believe, to drop the objectivist requirement.[17] It is opportunities for welfare that we should advocate equalizing. This, incidentally, conforms to Arneson's (1989) recommendation, although the reasons brought to bear here are entirely different from those he presents. But we must add that this escape from the unpleasant inconsistency is predicated upon a psychological premise - that adults care about their own functioning, and the *welfare* of their children.

[17] Before agreeing with us, however, the reader should consult Silvestre (in press), who works with a different economic environment from ours, in which, he shows, an increase in welfare over time and egalitarianism are consistent, even when adults do not care about the welfare of their children.

REFERENCES

Arneson, R., 1989. Equality and equality of opportunity for welfare. Philosophical Studies 56, 77-93.

Arrow, K., 1973. Rawls's principle of just saving. Swedish Journal of Economics, 323-335

Cohen, G.A., 1989. On the currency of egalitarian justice. Ethics 99, 906-44.

Cohen, G.A., 1993. Equality of what? On welfare, goods, and capabilities. In: Nussbaum, M., Sen, A. (Eds.), The Quality of Life. Clarendon Press, Oxford.

Dasgupta, P.S., 1974. On Some Alternative Criteria for Justice between Generations. Journal of Public Economics 3, 415-23.

Dworkin, R., 1981. What is equality? Part 2: Equality of resources. Philosophy & Public Affairs 10, 283-345.

Rawls, J., 1971. A Theory of Justice. Harvard University Press, Cambridge (MA.).

Roemer, J.E., 1998. Equality of Opportunity. Harvard University Press, Cambridge (MA).

Sen, A., 1980. Equality of what? In: McMurrin, S. (Ed.), The Tanner Lectures on Human Values, Vol.1. University of Utah Press, Salt Lake City.

Silvestre, J. Forthcoming. Progress and conservation under Rawls's maximin principle. Social Choice and Welfare.

Stokey, N.L., Lucas, R.E., (1989). Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge (MA).

APPENDIX.

Appendix 1: Proof that $\tau^t = \beta c_1/[(1 - \gamma)(1 - \beta c_2) + \beta c_2]$, $\forall\, t$, is optimal.

First of all, let us prove:

Lemma A.1: If $c_1 + c_2 \leq 1$ then $\Phi(w^t, w^{t+1})$ is strictly concave.

Proof. Consider the function $J = \dfrac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}$. Since $\dfrac{\partial J}{\partial w^{t+1}} = \dfrac{1}{c_1}\dfrac{(w^{t+1})^{1-c_1/c_1}}{(w^t)^{c_2/c_1}}$ and $\dfrac{\partial J}{\partial w^t} = -\dfrac{1}{w^t}\dfrac{c_2}{c_1}\dfrac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}$. The entries of the Hessian are $\dfrac{\partial^2 J}{\partial (w^{t+1})^2} = \dfrac{1-c_1}{(c_1)^2}\dfrac{(w^{t+1})^{1-2c_1/c_1}}{(w^t)^{c_2/c_1}}$, $\dfrac{\partial^2 J}{\partial (w^t)^2} = +\dfrac{c_2(c_2+c_1)}{(c_1)^2}\dfrac{1}{(w^t)^2}\dfrac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}$, $\dfrac{\partial^2 J}{\partial w^t \partial w^{t+1}} = \dfrac{\partial^2 J}{\partial w^{t+1} \partial w^t} = -\dfrac{c_2}{(c_1)^2}\dfrac{1}{w^t}\dfrac{(w^{t+1})^{1-c_1/c_1}}{(w^t)^{c_2/c_1}}$.

Let D_i denote the principal minor of order i of the Hessian: $D_1 > 0$, $\forall c_1, c_2$: $0 < c_1, c_2 < 1$. On the other hand, $D_2 \geq 0 \Leftrightarrow \dfrac{(1-c_1)c_2(c_1+c_2)}{(c_1)^4} - \dfrac{(c_2)^2}{(c_1)^4} \geq 0$. Hence, if $c_1 + c_2 \leq 1$, J is convex and Φ is strictly concave.▯

We can now prove:

Proposition A.1: Let A1', A2', A4 hold. In the solution to (5.5) it will be $\tau^t = \tau^* \equiv \beta c_1/[(1 - \gamma)(1 - \beta c_2) + \beta c_2]$, $\forall\, t$.

Proof. Consider now the Euler equations (E.A1) for SP:

$$\frac{(1-\gamma)(w^{t+1})^{1/c_1}}{c_1 k^{1/c_1}(w^t)^{c_2/c_1}} \Bigg/ \left(1 - \frac{(w^{t+1})^{1/c_1}}{k^{1/c_1}(w^t)^{c_2/c_1}}\right) = \beta + \frac{\beta(1-\gamma)c_2}{k^{1/c_1}c_1}\frac{(w^{t+2})^{1/c_1}}{(w^{t+1})^{c_2/c_1}} \Bigg/ \left(1 - \frac{(w^{t+2})^{1/c_1}}{k^{1/c_1}(w^{t+1})^{c_2/c_1}}\right)$$

and the transversality condition

$$\lim_{t\to\infty} \beta^t \left[\frac{1}{w^t} + (1-\gamma)\frac{1}{k^{1/c_1}}\frac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}} \Bigg/ w^t\left(1 - \frac{1}{k^{1/c_1}}\frac{(w^{t+1})^{1/c_1}}{(w^t)^{c_2/c_1}}\right)\right] w^t = 0 \qquad \text{(T.A1)}$$

or, expressing (E.A1) and (T.A1) in terms of the controls,

$$(1-\gamma)\frac{\tau^t}{1-\tau^t} = \beta c_1 + \beta c_2(1-\gamma)\frac{\tau^{t+1}}{1-\tau^{t+1}} \qquad \text{(E.A1')}$$

$$\lim_{t\to\infty} \beta^t \left[1 + (1-\gamma)\frac{\tau^t}{1-\tau^t}\right] = 0 \qquad \text{(T.A1')}$$

and since $\tau^t = \tau^*$, $\forall\, t$, satisfies (E.A1') and (T.A1'), by L.A1, the result follows. ▯

P.A1 implies that $w^{*t+1} = k(\tau^*)^{c_1}(w^{*t})^{c_2}$ is optimal for program (5.5), as claimed in P3. Actually, this result provides another way to derive $v^*(w^0)$. In fact,

$$v^*(w^0) = \sum_{t=0}^{\infty} \beta^t \left[\log w^{*t} + (1-\gamma)\log(1-\tau^*)\right] = \frac{(1-\gamma)}{(1-\beta)} \log(1-\tau^*) + \sum_{t=0}^{\infty} \beta^t \left[\log w^{*t}\right]$$

and since by P.A1

$$w^{*t} = k(\tau^*)^{c_1} (w^{*t-1})^{c_2} = (w^0)^{(c_2)^t} \prod_{i=0}^{t-1} (k)^{(c_2)^i} (\tau^*)^{c_1 (c_2)^i} = (k)^{\frac{1-(c_2)^t}{1-c_2}} (\tau^*)^{\frac{c_1[1-(c_2)^t]}{1-c_2}} (w^0)^{(c_2)^t}$$

$\forall t \geq 1$, the supremum function is

$$v^*(w^0) = \frac{(1-\gamma)}{(1-\beta)} \log(1-\tau^*) + \sum_{t=0}^{\infty} \beta^t (c_2)^t \log w^0 + \sum_{t=1}^{\infty} \beta^t \frac{1-(c_2)^t}{1-c_2} \log k + \sum_{t=1}^{\infty} \beta^t \frac{c_1[1-(c_2)^t]}{1-c_2} \log \tau^*$$

and therefore $v^*(w^0) = \phi + 1/(1 - \beta c_2) \log w^0$, where ϕ is given by (5.7).

Appendix 2: Proof that $v_0(w_L^0, w_H^0)$ is strictly concave.

Proof. It suffices to prove that $\Lambda = \log \dfrac{(w_H^0)^{c_2/c_1} (w_L^0)^{c_2/c_1}}{f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}}$ is concave. Since

$$\frac{\partial \Lambda}{\partial w_L^0} = \frac{c_2}{c_1} \frac{f_L (w_H^0)^{c_2/c_1}}{w_L^0 [f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]} \text{ and } \frac{\partial \Lambda}{\partial w_H^0} = \frac{c_2}{c_1} \frac{f_H (w_L^0)^{c_2/c_1}}{w_H^0 [f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]}, \text{ then:}$$

$$\frac{\partial^2 \Lambda}{\partial (w_L^0)^2} = -\frac{c_2}{c_1} \frac{f_L (w_H^0)^{c_2/c_1} [f_L (w_H^0)^{c_2/c_1} + \frac{c_1 + c_2}{c_1} f_H (w_L^0)^{c_2/c_1}]}{(w_L^0)^2 [f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]}$$

$$\frac{\partial^2 \Lambda}{\partial (w_H^0)^2} = -\frac{c_2}{c_1} \frac{f_H (w_L^0)^{c_2/c_1} [\frac{c_1 + c_2}{c_1} f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]}{(w_L^0)^2 [f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]}$$

$$\frac{\partial^2 \Lambda}{\partial (w_H^0) \partial (w_L^0)} = \frac{\partial^2 \Lambda}{\partial (w_L^0) \partial (w_H^0)} = \left(\frac{c_2}{c_1}\right)^2 \frac{f_H (w_L^0)^{c_2/c_1} f_L (w_H^0)^{c_2/c_1}}{w_L^0 w_H^0 [f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}]}$$

If D_i is the principal minor of order i of the Hessian: $D_1 < 0$, while $D_2 > 0$

$$[f_L (w_H^0)^{c_2/c_1} + \frac{c_1 + c_2}{c_1} f_H (w_L^0)^{c_2/c_1}][\frac{c_1 + c_2}{c_1} f_L (w_H^0)^{c_2/c_1} + f_H (w_L^0)^{c_2/c_1}] > \left(\frac{c_2}{c_1}\right)^2 f_L (w_H^0)^{c_2/c_1} f_H (w_L^0)^{c_2/c_1} \quad \blacksquare$$

Appendix 3: Analysis of Bellman's equation.

Firstly, we show that there exists an interval $(\bar{\rho}_0, \bar{\rho}_1]$ such that if $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, it is optimal to set $\rho^1 \in [1, \bar{\rho}_0]$, and $\rho^j = 1$, $j \geq 2$.

A necessary condition for $\rho^1 \in (1, \bar{\rho}_0)$ to be optimal is

$$\frac{(1-\gamma)}{w_L^1} \frac{1}{c_1 k^{1/c}} f_L \frac{(w_L^1)^{1/c_1}}{(w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c}} \left(f_H \frac{(w_H^1)^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{(w_L^1)^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right] = \beta \frac{\partial v_0 (w_L^1, w_H^1)}{\partial w_L^1}$$

$$\frac{(1-\gamma)}{w_H^1} \frac{1}{c_1 k^{1/c}} f_H \frac{(w_H^1)^{1/c_1}}{(w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c}} \left(f_H \frac{(w_H^1)^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{(w_L^1)^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right] = \beta \frac{\partial v_0 (w_L^1, w_H^1)}{\partial w_H^1}$$

Substituting for $\partial v_0(w_L^1, w_H^1)/\partial w_L^1$ and $\partial v_0(w_L^1, w_H^1)/\partial w_H^1$, expressing the FOCs in terms of the controls and the wage ratio and summing them:

$$(1-\gamma)\frac{\tau^0}{1-\tau^0} = \beta c_1 + \beta c_2 (1-\gamma)\frac{\tau_0}{1-\tau_0} \Rightarrow \tau^0 = \tau^* \equiv \frac{\beta c_1}{(1-\gamma)(1-\beta c_2)+\beta c_1}$$

while taking the FOC corresponding to w_H^1 and substituting for the optimal τ^0:

$$\frac{1}{(1-\beta c_2)}(1-f_L r_L^0) = \frac{(1-\gamma) f_H \rho^1}{[f_L + f_H \rho^1]} + \frac{\beta c_2}{(1-\beta c_2)}(1-f_L r_0(\rho^1))$$

or, equivalently,

$$\frac{1}{(1-\beta c_2)} \frac{f_H (\rho^1)^{1/c_1}}{[f_L (\rho^0)^{c_2/c_1} + f_H (\rho^1)^{1/c_1}]} = \frac{(1-\gamma) f_H \rho^1}{[f_L + f_H \rho^1]} + \frac{\beta c_2}{(1-\beta c_2)} \frac{f_H (\pi_0(\rho^1))^{1/c_1}}{[f_L (\rho^1)^{c_2/c_1} + f_H (\pi_0(\rho^1))^{1/c_1}]}$$

where $\pi_0(\rho^1)$ is the optimal level of ρ^2, given $\rho^1 \in [1, \bar{\rho}_0]$ (it can be derived from $r_0(\rho^1)$) and we know that $\pi_0(\rho^1) = 1$. Hence, define $h_1 : (1, \bar{\rho}_0) \rightarrow \mathbf{R}_+$ as

$$\frac{1}{(1-\beta c_2)} \frac{f_H (\rho^1)^{1/c_1}}{[f_L (h_1(\rho^1))^{c_2/c_1} + f_H (\rho^1)^{1/c_1}]} = \frac{(1-\gamma) f_H \rho^1}{[f_L + f_H \rho^1]} + \frac{\beta c_2}{(1-\beta c_2)} \frac{f_H}{[f_L (\rho^1)^{c_2/c_1} + f_H]}$$

i.e. we conjecture that $h_1(\rho^1)$ is the initial wage ratio, ρ^0, that makes it optimal to choose ρ^1, $\rho^1 \in [1, \bar{\rho}_0]$. It is possible to express h_1 explicitly:

$$h_1(\rho^1) = \frac{(\rho^1)^{1/c_2}}{(f_L)^{c_1/c_2}} \left[\frac{[f_L + f_H \rho^1][f_L (\rho^1)^{c_2/c_1} + f_H]}{(1-\beta c_2)(1-\gamma)\rho^1 [f_L (\rho^1)^{c_2/c_1} + f_H] + \beta c_2 [f_L + f_H \rho^1]} - f_H \right]^{\frac{c_1}{c_2}}$$

Clearly, h_1 is continuous and $h_1(\rho^1) > 0$. Moreover, if $\rho^1 = 1$, then

$$h_1(1) = \left[\frac{\gamma(1-\beta c_2)(1-\gamma)}{f_L (1-\beta c_2)(1-\gamma) + \beta c_1} + 1 \right]^{\frac{c_1}{c_2}} = \bar{\rho}_0$$

As shown in App.5, h_1 is differentiable and $\mathrm{d}h_1(\rho^1)/\mathrm{d}\rho^1 > 1$, $\forall \rho^1 \in [1, \bar{\rho}_0]$, and thus we can define $\bar{\rho}_1 = h_1(\bar{\rho}_0)$, with $\bar{\rho}_1 > \bar{\rho}_0$ and $\bar{\rho}_1 - \bar{\rho}_0 \geq \bar{\rho}_0 - 1$. Next, let $\pi_1 = h_1^{-1} : (\bar{\rho}_0, \bar{\rho}_1] \rightarrow (1, \bar{\rho}_0]$. From the properties of h_1, it follows that π_1 is continuous, differentiable and strictly increasing, with $\pi_1(\bar{\rho}_0) = \pi_0(\bar{\rho}_0) = 1$. Hence, patching together π_0 and π_1 one obtains an increasing and continuous function. The optimal control functions are $\tau_1 : (\bar{\rho}_0, \bar{\rho}_1] \rightarrow [0,1]$, and $r_1 : (\bar{\rho}_0, \bar{\rho}_1] \rightarrow \mathbf{R}_+$, where $\tau_1(\rho^0) = \tau^*$ and $r_1(\rho^0)$ can be derived from π_1. Thus, if $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, define $v_1 : W' \rightarrow \mathbf{R}$, by

$$v_1(w_L^0, w_H^0) = [\gamma \log w_L^0 + (1-\gamma)\log[f_L w_L^0 + f_H w_H^0] + (1-\gamma)\log(1-\tau_1) + \beta v_0(\varpi_{1,L}(w_L^0, w_H^0), \varpi_{1,H}(w_L^0, w_H^0))]$$

where $\varpi_{1,L}$: $W' \rightarrow \mathbf{R}_+$ and $\varpi_{1,H}$: $W' \rightarrow \mathbf{R}_+$ denote the optimal wage functions, for $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, which can be derived from τ_1 and r_1.

Notice that v_1 is continuous and, as shown in App. 4, strictly concave. Moreover, if $\rho^0 = \bar{\rho}_0$, $v_1(w_L^0, w_H^0) = v_0(w_L^0, w_H^0)$, and the two functions can be patched together. Finally, from the differentiability of v_0 it follows that v_1 it is continuously differentiable on $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$ with

$$\frac{\partial v_1(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\gamma}{w_L^0} + \frac{(1-\gamma) f_L}{[f_L w_L^0 + f_H w_H^0]} +$$

$$\frac{(1-\gamma)c_2}{w_L^0 c_1} \frac{f_L [\varpi_{1,L}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1} (w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(f_H \frac{[\varpi_{1,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{[\varpi_{1,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

$$\frac{\partial v_1(w_L^0, w_H^0)}{\partial w_H^0} = \frac{(1-\gamma) f_H}{[f_L w_L^0 + f_H w_H^0]} +$$

$$\frac{(1-\gamma)c_2}{w_H^0 c_1} \frac{f_H [\varpi_{1,H}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1} (w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(f_H \frac{[\varpi_{1,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{[\varpi_{1,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

and if $\rho^0 = \bar{\rho}_0$, $\dfrac{\partial v_1(w_L^0, w_H^0)}{\partial w_L^0} = \dfrac{\partial v_0(w_L^0, w_H^0)}{\partial w_L^0}$ and $\dfrac{\partial v_1(w_L^0, w_H^0)}{\partial w_H^0} = \dfrac{\partial v_0(w_L^0, w_H^0)}{\partial w_H^0}$.

Since $v_1(w_L^0, w_H^0)$ solves FE on $\rho^0 \in (\bar{\rho}_0, \bar{\rho}_1]$, we can proceed as above to show that there exists a value $\bar{\rho}_2 : \rho^0 \in (\bar{\rho}_1, \bar{\rho}_2]$, in the optimum $\rho^1 \in (\bar{\rho}_0, \bar{\rho}_1]$, and $\rho^t = 1$, $t \geq 2$. In general, proceed by induction and consider the k + 1-th stage, $k \geq 1$. Let $h_k : (\bar{\rho}_{k-2}, \bar{\rho}_{k-1}] \to \mathbf{R}_+$ denote a differentiable function such that $h_k(\rho^1)$ represents the value of ρ^0 that makes it optimal to choose $\rho^1 \in (\bar{\rho}_{k-2}, \bar{\rho}_{k-1}]$ (if $k = 1$, $\bar{\rho}_{k-2} = 1$). Let $dh_k(\rho^1)/d\rho^1 > 1$, $\forall\, \rho^1 \in (\bar{\rho}_{k-2}, \bar{\rho}_{k-1}]$ and define $\bar{\rho}_k = h(\bar{\rho}_{k-1})$ and $\pi_k = h_k^{-1} : (\bar{\rho}_{k-1}, \bar{\rho}_k] \to (\bar{\rho}_{k-2}, \bar{\rho}_{k-1}]$. If $\rho^0 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$, define v_k: W→ **R** as

$$v_k(w_L^0, w_H^0) = [\gamma \log w_L^0 + (1-\gamma)\log(f_L w_L^0 + f_H w_H^0) + (1-\gamma)\log(1-\tau_k) + \beta v_{k-1}(\varpi_{k,L}(w_L^0, w_H^0), \varpi_{k,H}(w_L^0, w_H^0))]$$

where $\varpi_{k,L}$: *W'* → $\mathbf{R}_+$ and $\varpi_{k,H}$: *W'* → $\mathbf{R}_+$ denote the optimal wage functions, for $\rho^0 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$, given the controls $\tau_k : (\bar{\rho}_{k-1}, \bar{\rho}_k] \to [0,1]$, $\tau_k = \tau^*$ and $r_k : (\bar{\rho}_{k-1}, \bar{\rho}_k] \to \mathbf{R}_+$. Let v_k be strictly concave and continuously differentiable on $\rho^0 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$ with

$$\frac{\partial v_k(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\gamma}{w_L^0} + \frac{(1-\gamma) f_L}{[f_L w_L^0 + f_H w_H^0]}$$

$$+ \frac{(1-\gamma)c_2}{w_L^0 c_1} \frac{f_L [\varpi_{k,L}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1} (w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(\frac{f_H [\varpi_{k,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L [\varpi_{k,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

$$\frac{\partial v_k(w_L^0, w_H^0)}{\partial w_H^0} = \frac{(1-\gamma) f_H}{[f_L w_L^0 + f_H w_H^0]}$$

$$+ \frac{(1-\gamma)c_2}{w_H^0 c_1} \frac{f_H [\varpi_{k,H}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1} (w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(\frac{f_H [\varpi_{k,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L [\varpi_{k,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

and if $\rho^0 = \bar{\rho}_{k-1}$, $\frac{\partial v_k(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\partial v_{k-1}(w_L^0, w_H^0)}{\partial w_L^0}$ and $\frac{\partial v_k(w_L^0, w_H^0)}{\partial w_H^0} = \frac{\partial v_{k-1}(w_L^0, w_H^0)}{\partial w_H^0}$.

We conjecture that there exists a $\bar{\rho}_{k+1}$: if $\rho^0 \in (\bar{\rho}_k, \bar{\rho}_{k+1}]$, in the optimum $\rho^j \in (\bar{\rho}_{k-j}, \bar{\rho}_{k-j+1})$, $j \geq 1$. A necessary condition for $\rho^1 \in (\bar{\rho}_{k-1}, \bar{\rho}_k)$ to be optimal is

$$\frac{(1-\gamma)f_L}{w_L^1 k^{1/c_1}} \frac{(w_L^1)^{1/c_1}}{c_1 (w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}}\left(f_H \frac{(w_H^1)^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{(w_L^1)^{1/c_1}}{(w_L^0)^{c_2/c_1}}\right)\right] = \beta \frac{\partial v_k(w_L^1, w_H^1)}{\partial w_L^1}$$

$$\frac{(1-\gamma)f_H}{w_H^1 k^{1/c_1}} \frac{(w_H^1)^{1/c_1}}{c_1 (w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}}\left(f_H \frac{(w_H^1)^{1/c_1}}{(w_H^0)^{c_2/c_1}} + f_L \frac{(w_L^1)^{1/c_1}}{(w_L^0)^{c_2/c_1}}\right)\right] = \beta \frac{\partial v_k(w_L^1, w_H^1)}{\partial w_H^1}$$

Manipulating the two FOCs as above one obtains

$$\tau^0 = \frac{\beta c_1}{(1-\gamma)(1-\beta c_2) + \beta c_1}$$

$$\frac{f_H(\rho^1)^{1/c_1}}{(1-\beta c_2)[f_L(\rho^0)^{c_2/c_1} + f_H(\rho^1)^{1/c_1}]} = \frac{(1-\gamma)f_H \rho^1}{[f_L + f_H \rho^1]} + \frac{\beta c_2 f_H (\pi_k(\rho^1))^{1/c_1}}{(1-\beta c_2)[f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}]}$$

where $\pi_k(\rho^1)$ gives the optimal level of ρ^2, given $\rho^1 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$. Hence, as above, define $h_{k+1}: (\bar{\rho}_{k-1}, \bar{\rho}_k) \to \mathbf{R}_+$ as the function that gives the value of ρ^0 that makes it optimal to choose $\rho^1 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$. Therefore:

$$h_{k+1}(\rho^1) = \frac{(\rho^1)^{1/c_2}}{(f_L)^{c_1/c_2}} \left[\frac{[f_L + f_H \rho^1][f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}]}{(1-\beta c_2)(1-\gamma)\rho^1[f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}] + \beta c_2[f_L + f_H \rho^1](\pi_k(\rho^1))^{1/c_1}} - f_H\right]^{\frac{c_1}{c_2}}$$

Clearly, h_{k+1} is continuous and $h_{k+1}(\rho^1) > 0$. Moreover, if $\rho^1 = \bar{\rho}_{k-1}$, then

$$h_{k+1}(\bar{\rho}_{k-1}) = \frac{(\bar{\rho}_{k-1})^{1/c_2}}{(f_L)^{c_1/c_2}} \left[\frac{[f_L + f_H \bar{\rho}_{k-1}][f_L(\bar{\rho}_{k-1})^{c_2/c_1} + f_H(\bar{\rho}_{k-2})^{1/c_1}]}{(1-\beta c_2)(1-\gamma)\bar{\rho}_{k-1}[f_L(\bar{\rho}_{k-1})^{c_2/c_1} + f_H(\bar{\rho}_{k-2})^{1/c_1}] + \beta c_2[f_L + f_H \bar{\rho}_{k-1}](\bar{\rho}_{k-2})^{1/c_1}} - f_H\right]^{\frac{c_1}{c_2}} = \bar{\rho}_k$$

As shown in App.5, h_{k+1} is differentiable and $\mathrm{d}h_{k+1}(\rho^1)/\mathrm{d}\rho^1 \geq 1$ $\forall\, \rho_1 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$. Hence, let $\bar{\rho}_{k+1} = h_{k+1}(\bar{\rho}_k)$, with $\bar{\rho}_{k+1} > \bar{\rho}_k$, and $\bar{\rho}_{k+1} - \bar{\rho}_k \geq \bar{\rho}_k - \bar{\rho}_{k-1}$. Let $\pi_{k+1} = h_{k+1}^{-1}: (\bar{\rho}_k, \bar{\rho}_{k+1}] \to (\bar{\rho}_{k-1}, \bar{\rho}_k]$: π_{k+1} is clearly continuous, differentiable and strictly increasing, with $\pi_{k+1}(\bar{\rho}_k) = \pi_k(\bar{\rho}_k)$. Thus, patching together π_{k+1} and π_k one obtains an increasing and continuous function. From π_{k+1} it is immediate to derive the optimal r_L^0 and therefore we can write the optimal control functions $\tau_{k+1}: (\bar{\rho}_k, \bar{\rho}_{k+1}] \to [0,1]$, $\tau_{k+1} = \tau^*$, and $r_{k+1}: (\bar{\rho}_k, \bar{\rho}_{k+1}] \to \mathbf{R}_+$. Thus, if $\rho^0 \in (\bar{\rho}_k, \bar{\rho}_{k+1}]$ define $v_{k+1}: W' \to \mathbf{R}$, by

$$v_{k+1}(w_L^0, w_H^0) = [\gamma \log w_L^0 + (1-\gamma)\log(f_L w_L^0 + f_H w_H^0) + (1-\gamma)\log(1-\tau_{k+1}) + \beta v_k(\varpi_{k+1,L}(w_L^0, w_H^0), \varpi_{k+1,H}(w_L^0, w_H^0))]$$

where $\varpi_{k+1,L}$: $W' \to \mathbf{R}_+$ and $\varpi_{k+1,H}$: $W' \to \mathbf{R}_+$ denote the optimal wage functions, for $\rho^0 \in (\bar{\rho}_k, \bar{\rho}_{k+1}]$, which can be derived from τ_{k+1} and r_{k+1}. If $\rho^0 = \bar{\rho}_k$, $v_{k+1}(w_L^0, w_H^0) = v_k(w_L^0, w_H^0)$ and the two functions can be patched together. Moreover, v_{k+1} is clearly continuous and as for v_1, it is possible to show that it is strictly concave and continuously differentiable for $\rho^0 \in (\bar{\rho}_k, \bar{\rho}_{k+1})$ with

$$\frac{\partial v_{k+1}(w_L^0, w_H^0)}{\partial w_L^0} = \frac{\gamma}{w_L^0} + \frac{(1-\gamma) f_L}{[f_L w_L^0 + f_H w_H^0]}$$

$$+ \frac{(1-\gamma)c_2}{w_L^0 c_1} \frac{f_L [\varpi_{k+1,L}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1}(w_L^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(\frac{f_H [\varpi_{k+1,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L [\varpi_{k+1,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

$$\frac{\partial v_{k+1}(w_L^0, w_H^0)}{\partial w_H^0} = \frac{(1-\gamma) f_H}{[f_L w_L^0 + f_H w_H^0]}$$

$$+ \frac{(1-\gamma)c_2}{w_H^0 c_1} \frac{f_H [\varpi_{k+1,H}(w_L^0, w_H^0)]^{1/c_1}}{k^{1/c_1}(w_H^0)^{c_2/c_1}} \Bigg/ \left[1 - \frac{1}{k^{1/c_1}} \left(\frac{f_H [\varpi_{k+1,H}(w_L^0, w_H^0)]^{1/c_1}}{(w_H^0)^{c_2/c_1}} + \frac{f_L [\varpi_{k+1,L}(w_L^0, w_H^0)]^{1/c_1}}{(w_L^0)^{c_2/c_1}} \right)\right]$$

and if $\rho^0 = \bar{\rho}_k$, $\dfrac{\partial v_{k+1}(w_L^0, w_H^0)}{\partial w_L^0} = \dfrac{\partial v_k(w_L^0, w_H^0)}{\partial w_L^0}$ and $\dfrac{\partial v_{k+1}(w_L^0, w_H^0)}{\partial w_H^0} = \dfrac{\partial v_k(w_L^0, w_H^0)}{\partial w_H^0}$.

We can now define a function v: $W' \to \mathbf{R}$ such that if $\rho^0 \in (\bar{\rho}_{k-1}, \bar{\rho}_k]$ $v(w_L^0, w_H^0) = v_k(w_L^0, w_H^0)$, $k \geq 0$ (if $k = 0$, $\bar{\rho}_{k-1} = 1$). Given the properties of the v_k's, v is strictly increasing and continuously differentiable in both variables and strictly concave. Moreover, v solves Bellman's FE by construction.

Appendix 4. Proof of the concavity of $v_1(w_L^0, w_H^0)$.

Definition. Let $w^t \in W'$ and $\hat{w}^t \in W'$ and let $w^{t+1} \in \Gamma(w^t)$, $\hat{w}^{t+1} \in \Gamma(\hat{w}^t)$. The feasibility correspondence Γ is *convex* if and only if $\forall\ \theta \in (0, 1)$

$\theta w^{t+1} + (1 - \theta) \hat{w}^{t+1} \in \Gamma(\theta w^t + (1 - \theta) \hat{w}^t)$

First of all, let us prove

Lemma A.2: if $c_1 + c_2 \leq 1$, Γ is convex.

Proof. If $w^{t+1} \in \Gamma(w^t)$, then $w_L^{t+1} \leq w_H^{t+1}$ and

$$0 \leq w_L^{t+1} \leq k (r_L^t)^{c_1} (w_L^t)^{c_2}$$

$$0 \leq w_H^{t+1} \leq k \frac{1}{(f_H)^{c_1}} (1 - f_L r_L^t)^{c_1} (w_H^t)^{c_2}$$

and likewise for $\hat{w}^{t+1}$. Hence, let us prove that $\Lambda = k (r_L^t)^{c_1} (w_L^t)^{c_2}$ is concave. Since $\dfrac{\partial \Lambda}{\partial w_L^t} = k c_2 (r_L^t)^{c_1} (w_L^t)^{c_2 - 1}$ and $\dfrac{\partial \Lambda}{\partial r_L^t} = k c_1 (r_L^t)^{c_1 - 1} (w_L^t)^{c_2}$, then $\dfrac{\partial^2 \Lambda}{\partial (w_L^t)^2} = k c_2 (c_2 - 1)(r_L^t)^{c_1} (w_L^t)^{c_2 - 2}$, $\dfrac{\partial^2 \Lambda}{\partial (r_L^t)^2} = k c_1 (c_1 - 1)(r_L^t)^{c_1 - 2} (w_L^t)^{c_2}$ and $\dfrac{\partial^2 \Lambda}{\partial w_L^t \partial r_L^t} = k c_1 c_2 (r_L^t)^{c_1 - 1} (w_L^t)^{c_2 - 1}$. Let D_i denote the principal minor of order i of the Hessian.

Clearly, $D_1 < 0$, while $D_2 = (c_1 - 1)(c_2 - 1) - c_1 c_2 \geq 0 \Leftrightarrow c_1 + c_2 \leq 1$. And a similar condition holds for the other constraint, proving the convexity of Γ.▤

Proof. Consider $w^0 \in W'$ and $\hat{w}^0 \in W'$. Let $w^1 \in \Gamma(w^0)$, $\hat{w}^1 \in \Gamma(\hat{w}^0)$ be the corresponding optimal policies. Let $w^1(\theta) \equiv \theta\, w^1 + (1 - \theta)\,\hat{w}^1$ and $w^0(\theta) \equiv \theta\, w^0 + (1 - \theta)\,\hat{w}^0$. By L.A2, $w^1(\theta) \in \Gamma(w^0(\theta))$. Moreover

$$
\begin{aligned}
&v_1(w^0(\theta)) \geq \Phi(w^0(\theta), w^1(\theta)) + \beta v_0(w^1(\theta)) > \theta\Phi(w^0, w^1) + (1-\theta)\Phi(\hat{w}^0, \hat{w}^1) + \beta v_0(w^1(\theta)) > \\
&\theta\Phi(w^0, w^1) + (1-\theta)\Phi(\hat{w}^0, \hat{w}^1) + \beta\theta v_0(w^1) + \beta(1-\theta)v_0(\hat{w}^1) = \theta v_1(w^0) + (1-\theta)v_1(\hat{w}^0)
\end{aligned}
$$

where the inequalities derive from the fact that $w^1(\theta)$ is not necessarily optimal and from the strict concavity of Φ and v_0, and the last equality is true since w^1 and $\hat{w}^1$ are optimal.▤

Appendix 5. Proof that $dh_{k+1}(\rho^1)/d\rho^1 > 1$, $\forall k \geq 0$.

Proof. From the formula in App. 3

$$
\frac{dh_{k+1}(\rho^1)}{d\rho^1} = \frac{(\rho^1)^{\frac{1}{c_2}-1}}{c_2(f_L)^{c_1/c_2}} \left\{ \frac{[f_L + f_H\rho^1][f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}]}{(1-\beta c_2)(1-\gamma)\rho^1[f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}] + \beta c_2[f_L + f_H\rho^1](\pi_k(\rho^1))^{1/c_1}} - f_H \right\}^{\frac{c_1}{c_2}-1}
$$

$$
* \frac{1}{\left[(1-\beta c_2)(1-\gamma)\rho^1[f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}] + \beta c_2[f_L + f_H\rho^1](\pi_k(\rho^1))^{1/c_1}\right]^2} * \{A_{k+1} - B_{k+1} - C_{k+1}\}
$$

where all terms but $\{A_{k+1} - B_{k+1} - C_{k+1}\}$ are clearly positive and

$$
\begin{aligned}
A_{k+1} = {} & [(1+c_2)(f_L)^2(\rho^1)^{c_2/c_1} + (1+c_1)(f_H)^2\rho^1(\pi_k(\rho^1))^{1/c_1} + (1+c_1+c_2)f_L f_H(\rho^1)^{\frac{c_2}{c_1}+1} + f_L f_H \rho^1 \frac{d\pi_k(\rho^1)}{d\rho^1}\pi_k(\rho^1))^{\frac{1}{c_1}-1} \\
& + f_L f_H \pi_k(\rho^1))^{1/c_1} + (f_H)^2(\rho^1)^2 \frac{d\pi_k(\rho^1)}{d\rho^1}\pi_k(\rho^1))^{\frac{1}{c_1}-1}] * \left[(1-\beta c_2)(1-\gamma)\rho^1[f_L(\rho^1)^{c_2/c_1} + f_H(\pi_k(\rho^1))^{1/c_1}] + \beta c_2[f_L + f_H\rho^1](\pi_k(\rho^1))^{1/c_1}\right]
\end{aligned}
$$

$$
\begin{aligned}
A_{k+1} = {} & (1+c_2)(1-\beta c_2)(1-\gamma)(f_L)^3(\rho^1)^{2\frac{c_2}{c_1}+1} + (1+c_2)(1-\beta c_2)(1-\gamma)f_H(f_L)^2(\rho^1)^{\frac{c_2}{c_1}+1}(\pi_k(\rho^1))^{\frac{1}{c_1}} + (1+c_2)\beta c_2(f_L)^3(\rho^1)^{\frac{c_2}{c_1}}(\pi_k(\rho^1))^{\frac{1}{c_1}} \\
& + (1+c_2)\beta c_2 f_H(f_L)^2(\rho^1)^{\frac{c_2}{c_1}+1}(\pi_k(\rho^1))^{\frac{1}{c_1}} + (1+c_1+c_2)f_L(f_H)^2(1-\beta c_2)(1-\gamma)(\rho^1)^{\frac{c_2}{c_1}+2}(\pi_k(\rho^1))^{\frac{1}{c_1}} + \\
& + (1+c_1+c_2)(1-\beta c_2)(1-\gamma)(f_L)^2 f_H(\rho^1)^{2\frac{c_2}{c_1}+2} + \beta c_2(1+c_1+c_2)(f_L)^2 f_H(\rho^1)^{\frac{c_2}{c_1}+1}(\pi_k(\rho^1))^{\frac{1}{c_1}} + \\
& + \beta c_2(1+c_1+c_2)(f_H)^2 f_L(\rho^1)^{\frac{c_2}{c_1}+2}(\pi_k(\rho^1))^{\frac{1}{c_1}} + (f_L)^2 f_H(1-\beta c_2)(1-\gamma)(\rho^1)^{\frac{c_2}{c_1}+1}(\pi_k(\rho^1))^{\frac{1}{c_1}} + f_L(f_H)^2(1-\beta c_2)(1-\gamma)\rho^1(\pi_k(\rho^1))^{\frac{2}{c_1}} \\
& + \beta c_2(f_L)^2 f_H(\pi_k(\rho^1))^{\frac{2}{c_1}} + \beta c_2 f_L(f_H)^2\rho^1(\pi_k(\rho^1))^{\frac{2}{c_1}} + (1+c_1)(1-\beta c_2)(1-\gamma)f_L(f_H)^2(\rho^1)^{\frac{c_2}{c_1}+2}(\pi_k(\rho^1))^{\frac{1}{c_1}} + \\
& (1+c_1)(f_H)^3(1-\beta c_2)(1-\gamma)(\rho^1)^2(\pi_k(\rho^1))^{\frac{2}{c_1}} + (1+c_1)\beta c_2 f_L(f_H)^2\rho^1(\pi_k(\rho^1))^{\frac{2}{c_1}} + (1+c_1)\beta c_2(f_H)^3(\rho^1)^2(\pi_k(\rho^1))^{\frac{2}{c_1}} \\
& + (1-\beta c_2)(1-\gamma)f_H(f_L)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+2} + (1-\beta c_2)(1-\gamma)f_L(f_H)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}(\rho^1)^2 + \\
& + \beta c_2 f_H(f_L)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}\rho^1 + 2\beta c_2 f_L(f_H)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}(\rho^1)^2 + \\
& + (1-\beta c_2)(1-\gamma)f_L(f_H)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+3} + [(1-\beta c_2)(1-\gamma) + \beta c_2](f_H)^3\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}(\rho^1)^3
\end{aligned}
$$

$$B_{k+1}=[(1-\beta c_2)(1-\gamma)(c_1+c_2)f_L(\rho^1)^{\frac{c_2}{c_1}}+(1-\beta c_2)(1-\gamma)c_1 f_H(\pi_k(\rho^1))^{1/c_1}+c_1\beta c_2 f_H(\pi_k(\rho^1))^{1/c_1}+(1-\beta c_2)(1-\gamma)f_H\rho^1(\pi_k(\rho^1))^{\frac{1}{c_1}-1}\frac{d\pi_k(\rho^1)}{d\rho^1}$$
$$+\beta c_2 f_L(\pi_k(\rho^1))^{\frac{1}{c_1}-1}\frac{d\pi_k(\rho^1)}{d\rho^1}+\beta c_2 f_H\rho^1(\pi_k(\rho^1))^{\frac{1}{c_1}-1}\frac{d\pi_k(\rho^1)}{d\rho^1}]*\left[(f_L)^2(\rho^1)^{\frac{c_2}{c_1}+1}+(f_H)^2(\rho^1)^2(\pi_k(\rho^1))^{1/c_1}+f_Lf_H\rho^1(\pi_k(\rho^1))^{1/c_1}+f_Lf_H(\rho^1)^{\frac{c_2}{c_1}+2}\right]$$

$$B_{k+1}=(1-\beta c_2)(1-\gamma)(c_1+c_2)(f_L)^3(\rho^1)^{2\frac{c_2}{c_1}+1}+(1-\beta c_2)(1-\gamma)(c_1+c_2)f_H(f_L)^2(\rho^1)^{2\frac{c_2}{c_1}+2}$$
$$+(1-\beta c_2)(1-\gamma)(c_1+c_2)f_L(f_H)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+2}+(1-\beta c_2)(1-\gamma)(2c_1+c_2)f_H(f_L)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+1}$$
$$+(1-\beta c_2)(1-\gamma)c_1f_L(f_H)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+2}+[(1-\beta c_2)(1-\gamma)+\beta c_2]c_1f_L(f_H)^2(\pi_k(\rho^1))^{\frac{2}{c_1}}\rho^1+(1-\beta c_2)(1-\gamma)c_1(f_H)^3(\pi_k(\rho^1))^{\frac{2}{c_1}}(\rho^1)$$
$$+c_1\beta c_2f_H(f_L)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+1}+c_1\beta c_2(f_H)^3(\pi_k(\rho^1))^{\frac{2}{c_1}}(\rho^1)^2+\beta c_2c_1f_L(f_H)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+2}$$
$$+[(1-\beta c_2)(1-\gamma)+2\beta c_2](f_L)^2f_H\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+2}+(1-\beta c_2)(1-\gamma)(f_H)^2f_L\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+3}$$
$$+[(1-\beta c_2)(1-\gamma)+2\beta c_2](f_H)^2f_L\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}(\rho^1)^2+[(1-\beta c_2)(1-\gamma)+\beta c_2](f_H)^3\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}(\rho^1)^3$$
$$+\beta c_2(f_L)^3\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+1}+\beta c_2f_H(f_L)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{2}{c_1}-1}\rho^1+\beta c_2f_L(f_H)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+3}$$

$$C_{k+1}=f_H*\left[(1-\beta c_2)(1-\gamma)\rho^1[f_L(\rho^1)^{c_2/c_1}+f_H(\pi_k(\rho^1))^{1/c_1}]+\beta c_2[f_L+f_H\rho^1](\pi_k(\rho^1))^{1/c_1}\right]^2$$

$$C_{k+1}=(1-\beta c_2)^2(1-\gamma)^2f_H(f_L)^2(\rho^1)^{2\frac{c_2}{c_1}+2}+2(1-\beta c_2)(1-\gamma)[(1-\beta c_2)(1-\gamma)+\beta c_2]f_L(f_H)^2(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+2}$$
$$+2(1-\beta c_2)(1-\gamma)\beta c_2(f_L)^2f_H(\pi_k(\rho^1))^{\frac{1}{c_1}}(\rho^1)^{\frac{c_2}{c_1}+1}+[(1-\beta c_2)(1-\gamma)+\beta c_2]^2(f_H)^3(\pi_k(\rho^1))^{\frac{2}{c_1}}(\rho^1)^2$$
$$+2\beta c_2[(1-\beta c_2)(1-\gamma)+\beta c_2](f_H)^2f_L(\pi_k(\rho^1))^{\frac{2}{c_1}}\rho^1+(\beta c_2)^2(f_L)^2f_H(\pi_k(\rho^1))^{\frac{2}{c_1}}$$

Grouping the different terms according to the exponents of ρ^1

$$A_{k+1}-B_{k+1}-C_{k+1}=(1-\beta c_2)(1-\gamma)[1-(1-\beta c_2)(1-\gamma)](f_L)^2f_H(\rho^1)^{2\frac{c_2}{c_1}+2}+(1-c_1)(1-\beta c_2)(1-\gamma)(f_L)^3(\rho^1)^{2\frac{c_2}{c_1}+1}$$
$$+(1+c_2)\beta c_2(f_L)^3(\rho^1)^{\frac{c_2}{c_1}}(\pi_k(\rho^1))^{\frac{1}{c_1}}+2[(1-c_1)(1-\beta c_2)(1-\gamma)+(1+c_2)\beta c_2-\beta c_2(1-\beta c_2)(1-\gamma)]f_H(f_L)^2(\rho^1)^{\frac{c_2}{c_1}+1}(\pi_k(\rho^1))^{\frac{1}{c_1}}$$
$$+[2(1-\beta c_2)(1-\gamma)-2(1-\beta c_2)^2(1-\gamma)^2-2\beta c_2(1-\beta c_2)(1-\gamma)+(1+c_2)\beta c_2]f_L(f_H)^2(\rho^1)^{\frac{c_2}{c_1}+2}(\pi_k(\rho^1))^{\frac{1}{c_1}}$$
$$+[(1-c_1)(1-\beta c_2)(1-\gamma)+2\beta c_2-2\beta c_2((1-\beta c_2)(1-\gamma)+\beta c_2)]f_L(f_H)^2\rho^1(\pi_k(\rho^1))^{\frac{2}{c_1}}$$
$$+[(1-\beta c_2)(1-\gamma)+\beta c_2][1-((1-\beta c_2)(1-\gamma)+\beta c_2)](f_H)^3(\rho^1)^2(\pi_k(\rho^1))^{\frac{2}{c_1}}+\beta c_2(1-\beta c_2)(f_L)^2f_H(\pi_k(\rho^1))^{\frac{2}{c_1}}$$
$$-\beta c_2f_L(f_H)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+3}-2\beta c_2f_H(f_L)^2\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+2}-\beta c_2(f_L)^3\frac{d\pi_k(\rho^1)}{d\rho^1}(\pi_k(\rho^1))^{\frac{1}{c_1}-1}(\rho^1)^{\frac{c_2}{c_1}+1}$$

All terms in the formula of $dh_{k+1}(\rho^1)/d\rho^1$ apart from the last three are strictly positive. However, a *sufficient* condition for $dh_{k+1}(\rho^1)/d\rho^1 > 0$ is that $\frac{d\pi_k(\rho^1)}{d\rho^1}\frac{\rho^1}{\pi_k(\rho^1)}\leq 1$. Since $\pi_k = h_k^{-1}:(\bar{\rho}_{k-1},\bar{\rho}_k]\rightarrow(\bar{\rho}_{k-2},\bar{\rho}_{k-1}]$ and it represents the optimal value of ρ^2, given

$\rho^1 \in [\bar{\rho}_{k-1}, \bar{\rho}_k]$, the above condition is equivalent to $\frac{dh_k(\rho^2)}{d\rho^2}\frac{\rho^2}{h_k(\rho^2)} \geq 1$, i.e. (we omit the superscript of ρ^2 to simplify notation):

$$\frac{(\rho)^{\frac{1}{c_2}-1}}{c_2(f_L)^{c_1/c_2}} * \left\{\frac{[f_L + f_H\rho][f_L(\rho)^{c_2/c_1} + f_H(\pi_{k-1}(\rho))^{1/c_1}]}{(1-\beta c_2)(1-\gamma)\rho[f_L(\rho)^{c_2/c_1} + f_H\pi_{k-1}(\rho))^{1/c_1}] + \beta c_2[f_L + f_H\rho]\pi_{k-1}(\rho))^{1/c_1}} - f_H\right\}^{\frac{c_1}{c_2}-1}$$
$$* \frac{1}{\left[(1-\beta c_2)(1-\gamma)\rho[f_L(\rho)^{c_2/c_1} + f_H\pi_{k-1}(\rho))^{1/c_1}] + \beta c_2[f_L + f_H\rho]\pi_{k-1}(\rho))^{1/c_1}\right]^2} * \{A_k - B_k - C_k\}$$
$$* (\rho)^{-\frac{1}{c_2}+1}(f_L)^{c_1/c_2} * \left\{\frac{[f_L + f_H\rho][f_L(\rho)^{c_2/c_1} + f_H(\pi_{k-1}(\rho))^{1/c_1}]}{(1-\beta c_2)(1-\gamma)\rho[f_L(\rho)^{c_2/c_1} + f_H\pi_{k-1}(\rho))^{1/c_1}] + \beta c_2[f_L + f_H\rho]\pi_{k-1}(\rho))^{1/c_1}} - f_H\right\}^{-\frac{c_1}{c_2}} \geq 1$$

or, simplifying and rearranging terms,

$$\{A_k - B_k - C_k\} \geq$$
$$c_2 * \left[(1-\beta c_2)(1-\gamma)\rho[f_L(\rho)^{c_2/c_1} + f_H\pi_{k-1}(\rho))^{1/c_1}] + \beta c_2[f_L + f_H\rho]\pi_{k-1}(\rho))^{1/c_1}\right] *$$
$$* \left\{[f_L + f_H\rho][f_L(\rho)^{c_2/c_1} + f_H(\pi_{k-1}(\rho))^{1/c_1}] - f_H[(1-\beta c_2)(1-\gamma)\rho[f_L(\rho)^{c_2/c_1} + f_H\pi_{k-1}(\rho))^{1/c_1}] + \beta c_2[f_L + f_H\rho]\pi_{k-1}(\rho))^{1/c_1}]\right\}$$

The right hand side of this expression can be written as

$$c_2(1-\beta c_2)(1-\gamma)(f_L)^3(\rho)^{2\frac{c_2}{c_1}+1} + 2c_2(1-\beta c_2)(1-\gamma)f_H(f_L)^2(\rho)^{\frac{c_2}{c_1}+1}(\pi_{k-1}(\rho))^{\frac{1}{c_1}} + c_2\beta c_2(f_L)^3(\rho)^{\frac{c_2}{c_1}}(\pi_{k-1}(\rho))^{\frac{1}{c_1}}$$
$$+ 2c_2\beta c_2 f_H(f_L)^2(\rho)^{\frac{c_2}{c_1}+1}(\pi_{k-1}(\rho))^{\frac{1}{c_1}} + 2c_2 f_L(f_H)^2(1-\beta c_2)(1-\gamma)(\rho)^{\frac{c_2}{c_1}+2}(\pi_{k-1}(\rho))^{\frac{1}{c_1}}$$
$$+ c_2\beta c_2(f_H)^2 f_L(\rho)^{\frac{c_2}{c_1}+2}(\pi_{k-1}(\rho))^{\frac{1}{c_1}} + c_2 f_L(f_H)^2[(1-\beta c_2)(1-\gamma) + 2\beta c_2]\rho(\pi_{k-1}(\rho))^{\frac{2}{c_1}}$$
$$+ c_2\beta c_2(f_L)^2 f_H(\pi_{k-1}(\rho))^{\frac{2}{c_1}} + c_2(f_H)^3[(1-\beta c_2)(1-\gamma) + \beta c_2](\rho)^2(\pi_{k-1}(\rho))^{\frac{2}{c_1}} + c_2(1-\beta c_2)(1-\gamma)(f_L)^2 f_H(\rho)^{2\frac{c_2}{c_1}+2}$$
$$- c_2(1-\beta c_2)^2(1-\gamma)^2 f_H(f_L)^2(\rho)^{2\frac{c_2}{c_1}+2} - 2c_2(1-\beta c_2)(1-\gamma)[(1-\beta c_2)(1-\gamma) + \beta c_2]f_L(f_H)^2(\pi_{k-1}(\rho))^{\frac{1}{c_1}}(\rho)^{\frac{c_2}{c_1}+2}$$
$$- 2c_2(1-\beta c_2)(1-\gamma)\beta c_2(f_L)^2 f_H(\pi_{k-1}(\rho))^{\frac{1}{c_1}}(\rho)^{\frac{c_2}{c_1}+1} - c_2[(1-\beta c_2)(1-\gamma) + \beta c_2]^2(f_H)^3(\pi_{k-1}(\rho))^{\frac{2}{c_1}}(\rho)^2$$
$$- 2c_2\beta c_2[(1-\beta c_2)(1-\gamma) + \beta c_2](f_H)^2 f_L(\pi_{k-1}(\rho))^{\frac{2}{c_1}}\rho - c_2(\beta c_2)^2(f_L)^2 f_H(\pi_{k-1}(\rho))^{\frac{2}{c_1}}$$

Subtracting the latter expression from $\{A_k - B_k - C_k\}$:

$$(1-c_2)(1-\beta c_2)(1-\gamma)[1-(1-\beta c_2)(1-\gamma)](f_L)^2 f_H(\rho)^{2\frac{c_2}{c_1}+2} + (1-c_1-c_2)(1-\beta c_2)(1-\gamma)(f_L)^3(\rho)^{2\frac{c_2}{c_1}+1}$$
$$+ \beta c_2(f_L)^3(\rho)^{\frac{c_2}{c_1}}(\pi_{k-1}(\rho))^{\frac{1}{c_1}} + 2[(1-c_1-c_2)(1-\beta c_2)(1-\gamma) + \beta c_2 - (1-c_2)\beta c_2(1-\beta c_2)(1-\gamma)]f_H(f_L)^2(\rho)^{\frac{c_2}{c_1}+1}(\pi_{k-1}(\rho))^{\frac{1}{c_1}}$$
$$+ [2(1-c_2)(1-\beta c_2)(1-\gamma) + \beta c_2 - 2(1-c_2)(1-\beta c_2)^2(1-\gamma)^2 - 2(1-c_2)\beta c_2(1-\beta c_2)(1-\gamma)]f_L(f_H)^2(\rho)^{\frac{c_2}{c_1}+2}(\pi_{k-1}(\rho))^{\frac{1}{c_1}}$$
$$+ [(1-c_1-c_2)(1-\beta c_2)(1-\gamma) + 2(1-c_2)\beta c_2 - 2(1-c_2)\beta c_2((1-\beta c_2)(1-\gamma) + \beta c_2)]f_L(f_H)^2\rho(\pi_{k-1}(\rho))^{\frac{2}{c_1}}$$
$$+ (1-c_2)[(1-\beta c_2)(1-\gamma) + \beta c_2][1-((1-\beta c_2)(1-\gamma) + \beta c_2)](f_H)^3(\rho)^2(\pi_{k-1}(\rho))^{\frac{2}{c_1}} + (1-c_2)\beta c_2(1-\beta c_2)(f_L)^2 f_H(\pi_{k-1}(\rho))^{\frac{2}{c_1}}$$
$$- \beta c_2 f_L(f_H)^2\frac{d\pi_{k-1}(\rho)}{d\rho}(\pi_{k-1}(\rho))^{\frac{1}{c_1}-1}(\rho)^{\frac{c_2}{c_1}+3} - 2\beta c_2 f_H(f_L)^2\frac{d\pi_{k-1}(\rho)}{d\rho}(\pi_{k-1}(\rho))^{\frac{1}{c_1}-1}(\rho)^{\frac{c_2}{c_1}+2} - \beta c_2(f_L)^3\frac{d\pi_{k-1}(\rho)}{d\rho}(\pi_{k-1}(\rho))^{\frac{1}{c_1}-1}(\rho)^{\frac{c_2}{c_1}+1}$$

and again all terms in the latter formula, apart from the last three, are strictly positive. In order to prove that $dh_k(\rho)/d\rho > 1$, $\forall k \geq 1$, we prove by induction that $\frac{dh_k(\rho)}{d\rho}\frac{\rho}{h_k(\rho)} \geq 1$,

$\forall k \geq 1$, $\forall\, \rho \in [\bar{\rho}_{k-2}, \bar{\rho}_{k-1}]$. Firstly, let k = 1. Since, $\pi_0(\rho) = 1$ and $\dfrac{d\pi_0(\rho)}{d\rho} = 0$, $\forall\, \rho \in [1, \bar{\rho}_0]$, from the above formulas it follows that $dh_1(\rho)/d\rho > 0$ and $\dfrac{dh_1(\rho)}{d\rho}\dfrac{\rho}{h_1(\rho)} \geq 1$, $\forall\, \rho \in [1, \bar{\rho}_0]$. Therefore, since $h_1(1) = \bar{\rho}_0$, we can define $\bar{\rho}_1 = h_1(\bar{\rho}_0)$, $\bar{\rho}_1 > \bar{\rho}_0$, and it follows that $\rho/h_1(\rho) < 1$, and thus $dh_1(\rho)/d\rho > 1$, $\forall\, \rho \in [1, \bar{\rho}_0]$.

Consider now the induction step. Assume $\dfrac{dh_k(\rho)}{d\rho}\dfrac{\rho}{h_k(\rho)} \geq 1$, $k \geq 1$, $\forall\, \rho \in [\bar{\rho}_{k-2}, \bar{\rho}_{k-1}]$. From the above formulas $dh_{k+1}(\rho)/d\rho > 0$ and $\dfrac{dh_{k+1}(\rho)}{d\rho}\dfrac{\rho}{h_{k+1}(\rho)} \geq 1$. Therefore, since $h_{k+1}(\bar{\rho}_{k-1}) = \bar{\rho}_k$, we can define $h_{k+1}(\bar{\rho}_k) = \bar{\rho}_{k+1}$, $\bar{\rho}_{k+1} > \bar{\rho}_k$, and it follows that $\rho/h_{k+1}(\rho) < 1$ and thus $dh_{k+1}(\rho)/d\rho > 1$, $\forall\, \rho \in [\bar{\rho}_{k-1}, \bar{\rho}_k]$.