Optimal taxation in occupational choice models: an application to the work decisions of couples

Guy Laroque
Nicola Pavoni

Optimal Taxation in Occupational Choice Models:

An Application to the Work Decisions of Couples[*]

Guy Laroque[†] and Nicola Pavoni[‡]

May 3, 2017

Abstract

We study a general model of occupational choice and optimal income taxation where agents have private cost of work that differ across occupations and have both deterministic and random components. We apply our framework to study the work decisions of couples in an extensive set up and give necessary and sufficient conditions under which joint-working households should be subsidized compared to single-worker households.

1 Introduction

We lay down a discrete choice model in a continuum economy. Each agent must choose one of a finite number of available alternatives, i.e. one of a finite number of occupations. The agents differ in their private costs of fulfilling their tasks, so that the economy features many dimensions of heterogeneity. The government knows the distribution of costs, but

[†]UCL, IFS and Sciences Po.

[‡]Bocconi and IFS.

1

cannot observe the individual characteristics. However it sees the occupational choices of the agents, and sets up taxes or subsidies conditional on these choices. We analyze the second best government program. The first order conditions have interesting structural features that generalizes the properties derived in the extensive model of labor supply, for instance in Choné and Laroque (2011) and Choné and Laroque (2005). The vector of pairwise differences in tax/subsidy across occupations satisfies a linear system featuring the social weights attached to the various alternatives and the pairwise elasticities of migration from one occupation to another.

Providing a general approach to optimal taxation and delivering general formulas for taxes may have a few important advantages. Many recent works within the optimal taxation literature (e.g., taxation of couples) address questions that require frameworks with multiple dimensions of heterogeneity. Due to the notorious technical complications of the multi-dimensional screening problem, most analysis are performed imposing several additional assumptions with unclear consequences on the final results.

We indeed apply the general setup to study the (extensive) labor supply decisions of couples. There are four possibilities: nobody works, only the man works, only the woman works, or both work. We focus on the comparison of the tax schedules faced by a member of a household, depending on the activity of her/his partner. We call *positive reinforcement* the situation when the financial incentives to work of an agent are larger when the partner is at work compared to the case where he/she stays at home: secondary earnings are *subsidized*. Conversely, *negative reinforcement* denotes the case where the incentives to work are smaller when the partner works than when s/he is unemployed. In the symmetric case, where the two partners have equal productivities while working, we derive a formula for reinforcement at the optimum, involving the social weights and the migration elasticities mentioned above. We find that there is a stronger case for subsidizing working couples than taxing them compared to single earners.

Three strands of literature are linked to our results. First of all, our approach encompasses recent models used to study optimal taxation in occupational choice models. These include: Rothschild and Scheuer (2013) and Rothschild and Scheuer (2015), Ales, Kurnaz, and Sleet

(2015) and Lockwood, Nathanson, and Weyl (forthcoming). The first two works focus on general equilibrium effects and externalities across sectors while Ales et al. focus on the effect of technical change on optimal taxation. They all assume that taxation is uniform across sectors (as the income tax schedule is the same across sectors and sales taxes are not considered). Gomes, Lozachmeur, and Pavan (2016) allow for occupation specific taxation and show the optimality of having production inefficiencies in a context where skills are imperfectly transferable across occupations. We use our model to address very different questions (optimal taxation of couples).

Recently there has been a number of studies of couple taxation. These include: Frankel (2014), Kleven, Kreiner, and Saez (2009), and Immervoll, Kleven, Kreiner, and Verdelin (2011). Frankel (2014) studies a simple model where the individuals can be of one of two types, low or high productivity. Their utility functions are linear in consumption. The government wants to favour the couples who are both of the low type. He finds that negative reinforcement is attenuated when the level of assortative mating increases. Our paper is complementary to Kleven, Kreiner, and Saez (2009) along two main dimensions. First, they do not allow for assortative mating (at least in the derivation of the analytical results). Second, they work with an exogenous definition of primary and secondary earners in the household. This, in particular rules out symmetric roles between partners within the couple. Our main contribution with respect to this literature is that we study all participation margins allowing for joint deviations. We allow for a 'non-precise' order among states for couples. In addition, we remove the assumption of independence allowing for any sort of assortative mating. We study in detail the symmetric case that has been ignored by the literature. Immervoll, Kleven, Kreiner, and Verdelin (2011) has two main limitations compared to our analysis. They first of all assume uni-dimensional heterogeneity by having a deterministic relationship between the productivity of the two partners. Moreover, they leave out an important margin: the possibility of moving from unemployment directly into employment for both partners. As we will explain below, when positive reinforcement is an optimal policy the margin between unemployment and employment for both partners is operative and crucial in shaping taxes. The generality of our approach allows this crucial margin to be operative.

Finally, our results can be related to the industrial organization (IO) literature, where positive reinforcement can be seen as bundling discounts. It includes: Armstrong (2011), McAfee, McMillan, and Whinston (1989), Spence (1980), Salinger (1995). There is an important analogy between our analysis and that done in the IO literature, however there are more relevant differences than what might appear at first sight. First of all, we do not only study the Rawlsian case (corresponding to the monopolist problem considered in IO).[1] Second, the level of consumption and welfare when inactive in our model is an endogenous variable while it is exogenously set (to zero) in IO. Another assumption typically made in IO is the nonnegativity of profits in each product which - in our framework - would imply positive taxation to single partners. This amounts to an implicit condition on Pareto weights in our framework which we may not want to impose.

The paper is organized as follows. Section 2 lays down a general occupation model with a finite number of choices. It derives the properties of the agents' behavior and the first order conditions associated with the program of the benevolent planner. Section 3 specializes the model to the labour supply of couples. In this section we mainly study the situations where the two partners in the couple are taxed equally. We provide a statistics whose sign is equal to that of reinforcement at the optimum. We discuss the scope for positive reinforcement, depending on the social weights given to couples where both work, on the elasticity of unemployment with respect to the unemployment benefit and on the elasticity of the number of two-earners couples with respect to their net incomes. Section 4 considers two extensions to this case. We first keep the non-discriminatory condition and allow for multiple productivities. Then we study the general case where the non-discriminatory condition is removed.

---

[1]A further difference here is that in our set up, the agent who is the worse off is endogenously determined while it is exogenous in the IO literature.

2 General Setup

There is a finite[2] number of occupations $i = 0, \ldots, I$, and a continuum of agents indexed by their type $\alpha \in \mathcal{A} \subset \mathbb{R}^N$. The distribution of types is described by the continuous cumulate F and the measure of the set $\mathcal{A}$ is 1. Occupation i yields a before tax production or income ω^i, $i = 0, 1, 2, \ldots I$, which is observed by the government. Let $c = (c^i, i = 0, \ldots I)$ be the vector of net income levels. Since agent's type is private information, after tax income c^i cannot depend on α but is a function of the observed decision i. The utility of agent α in occupation i when facing the net income schedule c is $u^i(c^i, \alpha)$.

Assumption 1 $\forall \, (i, \alpha)$, $u^i(\cdot, \alpha)$ *is increasing and continuously differentiable.*

Let $u_1^i(c^i, \alpha)$ be the derivative of u^i with respect to the first argument evaluated at (c^i, α).

Assumption 2 $\forall \, (i, c^i)$, *both $u^i(c^i, \cdot)$ and $u_1^i(c^i, \cdot)$ are continuous.*

These assumptions are meant to provide regularity conditions to the problem, but they will not be used in any specific proof. The key property that will be used throughout the paper is the possibility of taking derivatives of the measure of the sets associated to each occupation, $\mu(A^i(c))$, with respect to each element of c.

2.1 Occupation Choice Functions (Labor Supply)

Consider a group of agents facing the net income schedule $c := (c^i, i = 0, \ldots, I)$. The subset of agents that chooses i is $A^i(c)$:

$$A^i(c) := \{\alpha | u^i(c^i, \alpha) > u^j(c^j, \alpha) \text{ for all } j \neq i\}. \tag{1}$$

We assume enough regularities in utilities and the distribution F so that we can put a strict inequality sign, the measures of the sets are well defined and sum up to 1 over all the occu-

---

[2]It should be clear from what follows that considering countable many (or a continuum of) occupations is mainly a matter of 'techniques'. We believe that virtually nothing is lost in terms of economic intuition by considering finitely many occupations.

pations and are differentiable with respect to the after tax incomes.[3] The number of agents who choose i is the measure of the set A^i:

$$\mu(A^i(c)) := \int_{A^i(c)} \mathrm{d}F(\alpha).$$

Since everyone chooses one of the alternatives, and by assumption the measure of the points at the borders of the sets is zero,

$$\sum_{i=0}^{I} \mu(A^i(c)) = 1.$$

Since for each i the utility function u^i is increasing in c^i - and it does not depend on c^j, $j \neq i$ - it is easy to see that *the set $A^i(c)$ is non-decreasing in c^i and non-increasing in c^j for $j \neq i$*. When increasing c^i, all the agents that join A^i come from the other sets:

$$\frac{\partial \mu(A^i(c))}{\partial c^i} = - \sum_{j=0, j \neq i}^{I} \frac{\partial \mu(A^j(c))}{\partial c^i}. \tag{2}$$

2.2   The Planner Program and First Order Conditions

The benevolent planner maximizes a social objective subject to a budget constraint. The program of the utilitarian planner has the following form

$$\begin{cases} \max_{c} \quad \sum_{i=0}^{I} \int_{A^i(c)} \beta(\alpha) \psi(u^i(c^i, \alpha)) \mathrm{d}F(\alpha) \\ \sum_{i=0}^{I} \left[\omega^i - c^i \right] \mu(A^i(c)) = G, \end{cases} \tag{3}$$

where ψ is a weakly increasing function, and $\beta(\alpha) \geq 0$ with $\int \beta(\alpha) dF(\alpha) = 1$. Distributional motives are typically captured by imposing concavity on $\psi(\cdot)$ or some monotonicity on $\beta(\cdot)$.

Recall that we assumed that the measures of the choice sets are differentiable with respect to the after tax incomes. The second best program is typically a non convex program, where the first order conditions are necessary, but may lead to local minima. The first order conditions are nevertheless worth having and help to understand the trade-offs between equity and efficiency.

---

[3] A simple case which we shall use later is one where $\alpha = (\alpha^i, i = 0, 1, \dots, I)$ and $u^i(c^i, \alpha) = c^i - \alpha^i$ for all i. The distribution of α has support on a product of intervals $\Pi_{i=0}^{i=I}[\underline{\alpha}_i, \overline{\alpha}^i]$, $\underline{\alpha}_i < \overline{\alpha}^i$, and is absolutely continuous with respect to the Lebesgue measure.

Suppose that some coordinate c^i of the government instruments is free to move locally. With λ the Lagrange multiplier of the budget constraint, the first order condition is

$$\frac{\partial \mathcal{L}}{\partial c^i} = \int_{A^i(c)} [\beta(\alpha)\psi'(u^i(c^i,\alpha))u_1^i(c^i,\alpha) - \lambda]\mathrm{d}F(\alpha) + \lambda \sum_{j=0}^{I}(\omega^j - c^j)\frac{\partial\mu(A^j(c))}{\partial c^i} = 0, \quad (4)$$

where u_1^i indicates the derivative of u^i with respect to the first argument (i.e., the net income in state i) and all entries of the consumption vector c are at their optimal levels. In order to lighten notation, we often mute the dependence of the sets A^i on the consumption vector c.

It is useful to divide through by λ and to introduce the *average social weight of the agents* that choose state i (again, all values are computed at the optimal levels of consumption):

$$P(A^i) := \frac{1}{\mu(A^i)} \int_{A^i} \frac{\beta(\alpha)\psi'(u^i(c^i,\alpha))u_1^i(c^i,\alpha)}{\lambda}\mathrm{d}F(\alpha).$$

We now want to discuss some general properties of the system of first order conditions for this class of problems. We first consider the analysis that describes the tax levels, then the difference in taxes.

Levels. The first order conditions (4), for $i = 0, \ldots, I$, are

$$\mu(A^i)[P(A^i) - 1] = -\sum_{j=0}^{I}\frac{\partial\mu(A^j(c))}{\partial c^i}[\omega^j - c^j] = -\sum_{j=0}^{I}\frac{\partial\mu(A^j(c))}{\partial c^i}t^j. \quad (5)$$

where we denote by $t^i := \omega^i - c^i$ the *total tax* paid in state i. The full problem looks for the simultaneous solution of $I+2$ equations (the first order conditions plus the budget constraint) with $I+2$ unknowns (the t^i or c^i and the multiplier λ). Acknowledging the endogeneity of the Pareto weights, the measures of the sets and their derivatives, we will proceed as follows. We take the Pareto weights, the measure of the sets, and the derivatives of the measures of the sets as given, we solve for the system (5) of $I+1$ equations and $I+1$ unknowns. Then, if such values are obtained at the optimum they must solve the budget constraint, which delivers the value of the multiplier λ. The budget constraint of (3), written with the tax levels $t^i = \omega^i - c^i$, becomes

$$\sum_{i=0}^{I} t^i \mu(A^i(c)) = G.$$

Note that $t^i = G$ for all i and $P(A^i) = 1$ for all i is a solution to the system of equations made of the budget constraint and of the above first order conditions. This requires however that the parameters, in particular the $\beta(\cdot)$s, are compatible with the Pareto weights and that the after tax incomes $c^i = \omega^i - G$ belong to the domain of u^i.

Differences. In order to study the properties of the system when the aim is difference in taxes, we can use (2) to eliminate some terms in the system (5). One possibility is to eliminate the (own) derivative of $\mu(A^i)$ with respect to c^i. We obtain:

$$\mu(A^i)[P(A^i) - 1] = \sum_{j \neq i}^{I} \frac{\partial \mu(A^j(c))}{\partial c^i}[\omega^i - \omega^j - (c^i - c^j)]. \tag{6}$$

or, again using the tax definition

$$\mu(A^i)[P(A^i) - 1] = \sum_{j \neq i}^{I} \frac{\partial \mu(A^j(c))}{\partial c^i}[t^i - t^j]. \tag{7}$$

We have again $I + 2$ equations and the same number of unknowns. Notice moreover, that if we take as parameters the Pareto weights, the measures of the sets together with their derivatives, system (6) is linear in the differences of taxes.

The left hand side of (6) represents the *net value of increasing consumption to agents in set A^i in the economy absent production considerations and incentives*: at the margin, this is given by the difference between the average Pareto weight for this group and *the cost of funds* required to increase consumption to all agents in the set, multiplied by the size of the population in the set. As it is clear from the definition of the Pareto weights, we have normalized all values so that the cost of funds equals one, so the cost of funds is our 'numeraire'.

The right hand side represents the *budget effects induced by the changes of the sets A^j shaped by the incentives*. For example, agents moving from A^j to A^i because of the change in c^i will generate a gross (production) return/cost of $\omega^i - \omega^j$ and a differential cost of $c^i - c^j$ for each agent moving. So the total budget effect is given by the difference between the two effects (i.e., $t^i - t^j$) multiplied by the number of the migrants from A^j to A^i (i.e., $\frac{\partial \mu(A^j(c))}{\partial c^i}$).

8

A second possibility is to use (2) to eliminate the derivative of one set with respect to each consumption level, for example $\partial\mu(A^0(c))/\partial c^i$ through

$$\frac{\partial\mu(A^0(c))}{\partial c^i} = -\sum_{j=1}^{I}\frac{\partial\mu(A^j(c))}{\partial c^i}.$$

Substituting into (5) yields

$$\mu(A^i)[P(A^i) - 1] = -\sum_{j=1}^{I}\frac{\partial\mu(A^j(c))}{\partial c^i}[t^j - t^0].$$

Ignoring row 0, we have the following system of dimension $I \times I$

$$\mu(P - 1) = H\Delta t \tag{8}$$

where H is a square matrix with generic element $h_{ij} = -\frac{\partial\mu(A^j)}{\partial c^i}$ for $j = 1, \ldots, I$ and $i = 1, \ldots, I$, and Δt is a I vector with generic element $t^i - t^0$ for $i = 1, 2, \ldots I$. Note that from (8) we have a sort of 'anything goes' result since for any vector of tax differences, measures and Pareto weights, we can always find a matrix H that solves condition (8).[4] Assuming H invertible, the vector of tax differences must solve:

$$\Delta t = H^{-1}\mu(P - 1), \tag{9}$$

where $\mu(P-1)$ stands for the I vector of generic term $\mu(A^i)(P(A^i)-1)$. The budget constraint, together with the first line, indexed 0, of the difference system, would serve to determine the remaining unknowns, the tax t_0 and the multiplier λ. The inverse matrix H^{-1} has an interesting variational interpretation. When positive, each entry $h_{ij}^{(-1)}$ of row i of H^{-1} represents the element of proportionality that we should consider in the perturbation strategy for c^j. The perturbation implies an increase in the number of agents in occupation i and a decrease in the number of agents in occupation 0 with factor of proportionality equal to 1. By adopting such a perturbation strategy, all movements across occupations other than those in i or 0 fully offset each other and hence equal zero.

---

[4]Of course, the second order conditions must be verified. They impose conditions on the changes of Pareto weights and measures.

The system (9) provides a general formula for taxes relative to unemployment. By following the same steps, replacing A^0 and c^0 for A^i and c^i respectively, we obtain a similar $I \times I$ system delivering taxes with respect to state i.

A typical class of theoretical questions the literature on optimal taxation addresses is the determination of the set of Pareto weights for which, given certain conditions on H and μ, a specific property of the vector Δt holds.

2.3   Special Cases and Interpretation

To interpret our general formulae, we apply them to some special cases that are known in the literature. From now onwards let $\alpha(\cdot)$ be a vector function $\alpha : \mathcal{A} \to \mathbb{R}^{I+1}$ that maps the random type α into a vector of costs of work. A special case of this function is the following. Let $N = I + 1$ and $\alpha := (\alpha^i, i = 0, 1, 2, \ldots, I)$. Then $\alpha^i(\alpha)$ may represent the projection function on the i-th coordinate of the vector α.

Additive Separable Utility   Suppose $u^i(c^i, \alpha) = u(c) - \alpha^i(\alpha)$ and denote $u_1(\cdot) = u'(\cdot)$. Since an increase of $dc^i = \varepsilon/u'(c^i)$ in the payment c^i in occupation i changes utility by the same amount ε in each occupation, it leaves the distribution of agents fixed. Hence from (3) and the definition of $P(A^i)$

$$\frac{d\mathcal{L}}{\lambda} = \sum_{i=0}^{I} \left[\mu(A^i) P(A^i) dc^i - \mu(A^i) dc^i \right].$$

Therefore optimality requires:

$$\sum_{i=0}^{I} \frac{\mu(A^i)}{u'(c^i)} [P(A^i) - 1] = 0. \tag{10}$$

Using the fact that the agents at the border between sets A^i and A^j satisfy the relation $u(c^i) - \alpha^i(\alpha) = u(c^j) - \alpha^j(\alpha)$, which is linear in utility of net income, if we denote $u^i := u(c^i)$, we have[5]

$$\frac{\partial \mu(A^i)}{\partial u^j} = \frac{\partial \mu(A^j)}{\partial u^i}, \tag{11}$$

---

[5]The proof follows from making explicit the definition of the sets and their variations when there is either a change in u^i or in u^j. By definition $A^i(u) = \{\alpha | u^i - u^j > \alpha^i - \alpha^j, \text{ for all } j \neq i\}$. Consider a small change

In our case, this condition can also be written as[6]

$$\frac{1}{u'(c^j)}\frac{\partial \mu(A^i(c))}{\partial c^j} = \frac{1}{u'(c^i)}\frac{\partial \mu(A^j(c))}{\partial c^i}. \tag{12}$$

Pecuniary Cost Model The pecuniary cost model is a special case of the additive model where $u(c^i) = c^i$. In this model, alternative i entails a pecuniary cost $\alpha^i(\alpha)$, hence a net value of $\omega^i - \alpha^i(\alpha)$ in the absence of government intervention. The assumption of pecuniary costs rules out income effects so that aggregate supply satisfies a gross substitute property, which follows from the behavior of the sets A^i defined in (1) specialized to this case. Conditions (12) and (11) specialize to

$$\frac{\partial \mu(A^i)}{\partial c^j} = \frac{\partial \mu(A^j)}{\partial c^i}, \tag{13}$$

while condition (10) specializes to:[7]

$$\sum_{i=0}^{I} \mu(A^i)[P(A^i) - 1] = 0. \tag{14}$$

Note that in both these models the specialization of the utility functions u^i deliver both consistency restrictions on the Pareto weights (namely, conditions (10) and (14)) and conditions on cross derivatives (conditions (12) and (13)). This is intuitive. Recall that the system of equations (8) defined by the first order conditions is made of I independent equations that deliver the differences $t^i - t^0$ for $i = 1, 2, \ldots, I$. It must be that the neglected first order condition can be obtained as an identity by appropriately summing the I first order conditions.

$du^j = du$. The corresponding change in A^i is

$$\frac{dA^i(u)}{du^j} = \left\{ \alpha | u^i - u^j - du^j > \alpha^i - \alpha^j > u^i - u^j, u^i - u^h > \alpha^i - \alpha^h, \text{ for all } h \neq i, j \right\}.$$

Equivalently, taking $du^i = -du^j$, and replacing $u^i - \alpha^i$ with $u^j - \alpha^j$ with $\pm du$:

$$\frac{dA^i(u)}{du^j} = \left\{ \alpha | u^j - u^i > \alpha^j - \alpha^i > u^j - u^i - du^i, u^j - u^h > \alpha^j - \alpha^h, \text{ for all } h \neq i, j \right\} = \frac{dA^j(u)}{du^i}.$$

---

[6]The proof uses the fact that c^j does not enter u^i for $j \neq i$. Note indeed that the inequalities defining the different sets are not only linear in the difference in utilities $u(c^i) - u(c^j)$, an increase in c^i induces changes to the set A^j only because agents move into set A^i from set A^j, $i \neq j$.

[7]This condition can also be obtained directly by summing up the first order conditions using the symmetry of the cross derivatives in (13).

11

For this to happen, the elements in the matrix H must be somewhat related to each other, and this relationship should be consistent with the relationship linking the elements of the vector $\mu(P-1)$. For the additive separable model (and its special case, the pecuniary model) indeed, it can be directly verified that by summing up the elements on the left hand side and on the right hand side of the system $\mu(P-1) = H\Delta t$ one obtains the first order condition for c^0.

Extensive Margin (Choné & Laroque)   The simplest case is when $I = 1$, i.e., the most standard extensive model, in which case, we have only one tax rate $t^1 - t^0 = \omega - (c^1 - c^0)$, and this rate solves (see condition (7) for $i = 0, 1$):

$$P(A^1) - 1 = \frac{\frac{\partial \mu(A^0)}{\partial c^1}}{\mu(A^1)}[t^1 - t^0] = -\frac{\frac{\partial \mu(A^1)}{\partial c^1}}{\mu(A^1)}[t^1 - t^0];$$

$$P(A^0) - 1 = \frac{\frac{\partial \mu(A^1)}{\partial c^0}}{\mu(A^0)}[t^0 - t^1] = \frac{\frac{\partial \mu(A^0)}{\partial c^0}}{\mu(A^0)}[t^1 - t^0],$$

where we used the specialization to the two states case of condition (2): $\frac{\partial \mu(A^0)}{\partial c^1} = -\frac{\partial \mu(A^1)}{\partial c^1}$ and $\frac{\partial \mu(A^1)}{\partial c^0} = -\frac{\partial \mu(A^0)}{\partial c^0}$. Since own derivatives are positive ($\frac{\partial \mu(A^i)}{\partial c^i} > 0$), the sign of $t^1 - t^0$ is fully determined by the sign of $P(A^1) - 1$ or, equivalently, the sign of $1 - P(A^0)$, with $t^1 - t^0 > 0$ if and only if $P(A^1) < 1$. In the additive separable model, condition (10) together with the first order conditions, implies

$$\frac{\mu(A^1)[P(A^1) - 1]}{u'(c^1)} = -\frac{\mu(A^0)[P(A^0) - 1]}{u'(c^0)} = -\frac{\partial \mu(A^1)}{\partial u^1}[t^1 - t^0] = -\frac{\partial \mu(A^0)}{\partial u^0}[t^1 - t^0]. \quad (15)$$

Three states   Immervoll, Kleven, Kreiner, and Verdelin (2011) study a model of couple taxation with restrictions on the set of borders that make it effectively a simplified version of our model with $I = 2$.[8] In this case, each occupation has at most two interlinked borders.

---

[8]Immervoll, Kleven, Kreiner, and Verdelin (2011) also postulate that there is no agent in the border between the set A^2 and A^0, that is $\frac{\partial \mu(A^2)}{\partial c^0} = \frac{\partial \mu(A^0)}{\partial c^2} = 0$. Guaranteeing this 'no-border condition' unconditional on the payments is effectively requiring only one dimension of heterogeneity. This is hence a simplified version of the Mirrlees model we discuss below.

The matrix H takes the form:

$$H = \begin{bmatrix} -\frac{\partial \mu(A^1)}{\partial c^1} & -\frac{\partial \mu(A^2)}{\partial c^1} \\ -\frac{\partial \mu(A^1)}{\partial c^2} & -\frac{\partial \mu(A^2)}{\partial c^2} \end{bmatrix}.$$

Assuming H invertible - namely, $|H| := \frac{\partial \mu(A^1)}{\partial c^1} \frac{\partial \mu(A^2)}{\partial c^2} - \frac{\partial \mu(A^2)}{\partial c^1} \frac{\partial \mu(A^1)}{\partial c^2} \neq 0$ - (9) gives the following necessary condition for tax rates:

$$t^1 - t^0 = \frac{1}{|H|} \left\{ \frac{\partial \mu(A^2)}{\partial c^1} \mu(A^2)[P(A^2) - 1] - \frac{\partial \mu(A^2)}{\partial c^2} \mu(A^1)[P(A^1) - 1] \right\}$$

$$t^2 - t^0 = \frac{1}{|H|} \left\{ \frac{\partial \mu(A^1)}{\partial c^2} \mu(A^1)[P(A^1) - 1] - \frac{\partial \mu(A^1)}{\partial c^1} \mu(A^2)[P(A^2) - 1] \right\}$$

The tax rate between occupations 1 and 2 can be obtained from $t^2 - t^1 = t^2 - t^0 - (t^1 - t^0)$. Immervoll, Kleven, Kreiner, and Verdelin (2011) consider the case where $\frac{\partial \mu(A^2)}{\partial c^0} = \frac{\partial \mu(A^0)}{\partial c^2} = 0$ and hence $\frac{\partial \mu(A^1)}{\partial c^2} = -\frac{\partial \mu(A^2)}{\partial c^2}$. We can hence recover the expression for taxes in their paper:

$$t^2 - t^1 = \frac{1}{|H|} \left[\frac{\partial \mu(A^2)}{\partial c^1} + \frac{\partial \mu(A^1)}{\partial c^1} \right] \mu(A^2)[1 - P(A^2)] = \frac{\mu(A^2)}{\frac{\partial \mu(A^2)}{\partial c^2}}[1 - P(A^2)],$$

where we used the fact that, in this case, the determinant equals $|H| = \left[\frac{\partial \mu(A^2)}{\partial c^1} + \frac{\partial \mu(A^1)}{\partial c^1} \right] \frac{\partial \mu(A^2)}{\partial c^2}$.

Intensive Margin (Mirrlees)   Consider the following specialization of our model. We let a gross income ω^i represent the production of the worker in occupation $i = 0, 1, \ldots I$ (we are considering a finite number of possible income levels). To simplify the discussion we assume additive separability:

$$u^i(c^i, \alpha) = u(c^i) - \alpha^i(\alpha),$$

with u increasing and concave. Assume that ω^i is strictly increasing in i and, for each α, $\alpha^i(\alpha)$ is strictly increasing in i. The index i hence represents a measure of 'intensity' in the space of income.[9]

A crucial assumption in the optimal taxation literature is *increasing difference*. This requires the definition of a (complete) order $\succsim$ on the space of types. The increasing difference

---

[9]We impose no conditions on ω^i, but of course in order to have positive support for each occupation we would need to make some regularity assumptions. We disregard occupations that are not chosen by any worker. Appropriate re-labeling allows us to use our indexes without loss of generality.

assumption (or single crossing of indifference curves) amounts to requiring that for all $j > k$ we have[10]

$$\alpha^j(\alpha') - \alpha^k(\alpha') \leq \alpha^j(\alpha) - \alpha^k(\alpha) \text{ whenever } \alpha' \gtrsim \alpha, \text{ with strict inequality whenever } \alpha' \succ \alpha.$$

Under the assumption of monotonicity of $\alpha^i(\alpha)$ in i and increasing difference it is easy to show the following three facts. (a) net incomes c^i increase in i; (b) occupation i has at most two borders: $i - 1$ and $i + 1$; (c) if type α chooses occupation i then type $\alpha' \gtrsim \alpha$ chooses occupation $j \geq i$.

The fact that the single crossing condition is based on a complete order on $\mathcal{A}$ suggests that this model has only one true dimension of heterogeneity. Using the analogy to the utility representation of preferences indeed, under some regularity assumptions in $\gtrsim$, there exists a function $\theta : \mathcal{A} \to \mathbb{R}$ with the following property: $\theta(\alpha') \geq \theta(\alpha) \iff \alpha' \gtrsim \alpha$. The function θ is uniquely defined up to a monotone transformation. We can then redefine types in the unidimensional codomain of θ.

Since only borders $i - 1$ and $i + 1$ are possibly active, the general first order condition (7) - i.e., when we eliminate the own derivative - for this model specializes to:

$$\mu(A^i)[P(A^i) - 1] = \frac{\partial \mu(A^{i-1})}{\partial c^i}[t^i - t^{i-1}] + \frac{\partial \mu(A^{i+1})}{\partial c^i}[t^i - t^{i+1}].$$

We can now recover the optimal taxation formulas. The typical perturbation considered is an increase in net income c^s by $\frac{u'(c^i)}{u'(c^s)}\varepsilon$ for all $s \geq i$. This perturbation implies that no agent would change occupation for $s > i$.[11] The condition requiring that such a perturbation not improve on the optimal allocation reads:

$$t^i - t^{i-1} = \frac{u'(c^i)}{\frac{\partial \mu(A^{i-1})}{\partial c^i}} \sum_{s=i}^{I} \frac{\mu(A^s)[P(A^s) - 1]}{u'(c^s)} = \frac{\mu(A^i)}{\frac{\partial \mu(A^i)}{\partial u^i}} \sum_{s=i}^{I} \frac{\mu(A^s)[1 - P(A^s)]}{u'(c^s)\mu(A^i)}, \qquad (16)$$

where, again, u^i is a shortcut for $u(c^i)$ and we use the identity $\frac{\partial \mu(A^i)}{\partial c^i}\frac{1}{u'(c^i)} \equiv \frac{\partial \mu(A^i)}{\partial u^i}$. In the above, we also used the fact that state i has at most two borders, so that condition (2) specialises to

---

[10]We could have considered a more general model. Single crossing would again require to define a complete order over the type space such that $u^j(c, \alpha) - u^k(c, \alpha) > u^j(c, \alpha') - u^k(c, \alpha')$ for all c, $j > k$, and $\alpha \gtrsim \alpha'$.

[11]Mechanically, this perturbation allows to isolate the component $\frac{\partial \mu(A^{i-1})}{\partial c^i}[t^i - t^{i-1}]$ in the first order conditions by taking summation over $s \geq i$.

$\frac{\partial \mu(A^i)}{\partial c^i} = -\frac{\partial \mu(A^{i-1})}{\partial c^i} - \frac{\partial \mu(A^{i+1})}{\partial c^i}$. This formula is analogous to the one obtained in the literature for the marginal tax rates; in our framework with finite income levels, we obtain expressions of 'average taxes' by dividing both sides by $\omega^i - \omega^{i-1}$. For each i, let $\Pi(i) := \sum_{s=i}^{I} \frac{\mu(A^s)[1-P(A^s)]}{u'(c^s)}$. Note that $\Pi(0) = 0$ by construction. Moreover, the usual distributional motive suggests that $\Pi(i)$ increases with i and hence $\Pi(i) > 0$ implying that *the average tax is always positive.*

Similarly, if we increase consumption payment c^s by $\frac{u'(c^i)}{u'(c^s)}$ for all $s < i$, so that no agent would change state for $s < i - 1$, we obtain as optimality condition:

$$t^{i-1} - t^i = \frac{\mu(A^i)}{\frac{\partial \mu(A^i)}{\partial u^i}} \sum_{s=0}^{i-1} \frac{\mu(A^s)[1 - P(A^s)]}{u'(c^s)\mu(A^i)}.$$

The tight relationship between the two tax expressions generalizes the argument we made in the $I = 1$ case (see equations (15)) and in this case, they are a corollary to condition (10) rewritten as:

$$\sum_{s=i}^{I} \frac{\mu(A^s)[P(A^s) - 1]}{u'(c^s)} = -\sum_{s=0}^{i-1} \frac{\mu(A^s)[P(A^s) - 1]}{u'(c^s)}.$$

Intensive and Extensive (Saez (2002)) Consider now the combination of the intensive and extensive model, and maintain the separable utility assumption: $u(c^i) - \alpha^i(\alpha)$. Suppose that agents also have the border with the set A^0.[12] Since only borders $i - 1$, $i + 1$, and 0 are possibly active, the general first order condition (7) for this model specializes to:

$$\mu(A^i)[P(A^i) - 1] = \frac{\partial \mu(A^{i-1})}{\partial c^i}[t^i - t^{i-1}] + \frac{\partial \mu(A^{i+1})}{\partial c^i}[t^i - t^{i+1}] + \frac{\partial \mu(A^0)}{\partial c^i}[t^i - t^0].$$

[12]In terms of assumptions on the cost function $\alpha^i(\alpha)$ we amend the Mirrlees assumption as follows. We maintain the *increasing difference* assumption as above for all $j, k > 0$. We impose no restriction on $\alpha^0(\cdot)$ in terms of the order $\succsim$. The single crossing condition implies that each agent α faces at most three borders. Each type that chooses occupation $i > 0$ might be at the border with $i - 1$, or with $i + 1$ or at the border with occupation 0. Note indeed that the arguments excluding any other border for $k > 0$, $k \neq i - 1, i + 1$ apply directly. This model admits two genuine dimensions of heterogeneity across agents.

We can again recover the optimal taxation formula by considering a change in c^s by the amount $\frac{1}{u'(c^s)}$ for $s \geq i$. This perturbation modifies formula (16) as follows:

$$
t^i - t^{i-1} = \frac{u'(c^i)}{\frac{\partial \mu(A^i)}{\partial c^i}} \sum_{s=i}^{I} \frac{\mu(A^s)}{u'(c^s)} \left[1 - P(A^s) + \frac{\frac{\partial \mu(A^0)}{\partial c^s}}{\mu(A^s)} (t^s - t^0) \right]
$$

$$
= \frac{u'(c^i)}{\frac{\partial \mu(A^i)}{\partial c^i}} \sum_{s=i}^{I} \frac{\mu(A^s)}{u'(c^s)} \left[1 - P(A^s) + \frac{\frac{\partial \mu(A^s)}{\partial c^0}}{\mu(A^s)} \frac{u'(c^s)}{u'(c^0)} (t^s - t^0) \right]
$$

where in the last term in the square bracket we used (12). The average tax between $i-1$ and i gets modified by the additional term $\frac{\frac{\partial \mu(A^0)}{\partial c^s}}{\mu(A^s)}(t^s - t^0)$ which is negative whenever $t^s - t^0 > 0$. Note that this expression is analogous to that in Saez (2002), page 1055. It expresses the tax rate $t^i - t^{i-1}$ implicitly, as a function of other tax rates: hence the difficulty to resolve without understanding the determinants of the sign of $t^s - t^0$. Recall that when the matrix H is invertible, from (9) we obtain the expression for the vector $(t^s - t^0, s = 1, \ldots, I)$ which can be plugged directly into the above equation and express the tax rate $t^i - t^{i-1}$ as a function of only elasticities and Pareto weights. In particular, the matrix H in this case takes the form:

$$
H = \begin{bmatrix}
h_{11} & h_{12} & 0 & 0 & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & 0 & h_{ii-1} & h_{ii} & h_{ii+1} & 0 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & 0 & 0 & h_{II-1} & h_{II}
\end{bmatrix},
$$

where, recall, $h_{ij} = -\frac{\partial \mu(A^j)}{\partial c^i}$. Any symbolic algebra program (such as Matlab, Maple, or Mathematica) delivers the inverse as a relatively ugly expression, and the inverse times $|H|$ as a less ugly expression thanks to the many zeros in the matrix.

As we will see, the model allows for the couple labor supply decision interpretation. In the next section, we will focus on the extensive margins for each spouse. In Section 4.3 we will briefly discuss the case with intensive margins. The key element in the structure of the problem, the one ultimately determining the expressions for optimal taxes, is the structure of 'borders' across occupations. Namely, formulas for optimal taxes are crucially shaped by the set of occupations that have (at least) one common border with each occupation i.

16

2.4 Elasticities, Average Taxes, and Income Effects

It is common in optimal taxation literature to express tax rates in terms of elasticities. The 'consumption elasticity' of the set A^j when payment $i \neq j$ changes is defined as follows[13]

$$\eta_i^j := -\frac{\partial \mu(A^j)}{\partial(c^i - c^j)} \frac{c^i - c^j}{\mu(A^j)}. \tag{17}$$

Since for $j \neq i$, $\frac{\partial \mu(A^j)}{\partial c^i} = \frac{\partial \mu(A^j)}{\partial(c^i-c^j)} \leq 0$,[14] the sign of the elasticity depends on the sign of $c^i - c^j$.

If $\omega^i \neq \omega^j$ we can define the *average tax rate* between state i and j as

$$\tau^{i,j} = \tau^{j,i} := \frac{t^i - t^j}{\omega^i - \omega^j},$$

so that $(c^i - c^j)/(\omega^i - \omega^j) = 1 - \tau^{i,j}$, and we can rewrite the system of first order conditions as:

$$P(A^i) - 1 = -\sum_{j \neq i}^I \eta_i^j \frac{\mu(A^j)}{\mu(A^i)} \frac{\tau^{i,j}}{1 - \tau^{i,j}}. \tag{18}$$

In general, while we always have (by definition) $\tau^{i,j} = \tau^{j,i}$, the elasticity η_i^j is not directly related to η_j^i. The main discrepancy might be imputed to 'income effects'. In the pecuniary cost model indeed, it is easy to show that, thanks to (13), $\eta_j^i = \eta_i^j$. In the additive case a similar condition (i.e., $\hat{\eta}_j^i = \hat{\eta}_i^j$) holds true for 'utility elasticities' defined as:

$$\hat{\eta}_i^j := -\frac{\partial \mu(A^j)}{\partial u^i} \frac{u(c^i) - u(c^j)}{\mu(A^j)} = -\frac{\partial \mu(A^j)}{\partial c^i} \frac{u(c^i) - u(c^j)}{u'(c^i)\mu(A^j)}. \tag{19}$$

In this paper we will work with tax levels (t^i or Δt) rather than with tax rates τ since incomes ω^i and ω^j may be equal, making the tax *rate* undefined. In the model of Saez (2002)) η_i^0 can be interpreted as *participation elasticities* while η_{i+1}^i can be interpreted as *intensive elasticities*.

---

[13]Note that although $c^i - c^j$ is linear in the two consumption levels, in general the elasticity depends on which of the two payments (c^i or c^j) changes. Indeed the changes in the measure of the sets are determined by changes in utility (not in consumption) of the agents in the border between the two sets A^i and A^j. However, note that as long as utilities are monotone in consumption, we cannot distinguish consumption across agents with different α within a given state. This observation, together with the linearity of the budget constraint in consumption justifies the choice of these elasticities. Note that the estimation of these elasticities is complicated as they are not even symmetric. They would be symmetric in absence of income effects.

[14]The equality in the derivative is immediate since the derivatives are computes keeping c^j constant.

These numbers are not easy to obtain in the data. The estimated values of η_i^0 depend both on the sex of the participant and on the income. For males, η_i^0 is around .2; for females, it ranges between .6 and 1 (single mothers with low income have a participation elasticity of .6 while for wives the elasticity may be as high as .9). These numbers decrease with income of the household (perhaps they are even zero for households with earnings above the median). The intensive elasticity for men is again approximately .2 while for married woman it is around .25 for low income earners and up to .5 for middle and high income earners.

3 Taxation of couples

The model we presented in the previous section is well adapted to represent the labor supply decisions of men and women in couples in as far as they involve multidimensional random shocks. We specialize the setup into an extensive model. The economy is populated by couples made of a man and of a woman. Every agent can either work full time or not work. The couple therefore chooses one of the four states: $i = 0$ (no one works), 1 (man works, woman does not), 2 (man does not work, woman does) or 3 (both work). The couple's production is ω^0 when nobody works, respectively ω^1 and ω^2 when only one member of the couple works and ω^3 when both are working. We study a unitary model where the choices of the couple are derived from the maximization of a utility function u^i of total after-tax income c^i.

We assume that the productivity of any member of the couple is independent of the activity of her/his partner:

Assumption 3 (Additivity of production)

$$\omega^0 = 0, \qquad \omega^3 = \omega^1 + \omega^2.$$

The assumption $\omega^0 = 0$ is just a normalization (as all results would be identical if we increased all $\omega's$ and public expenditure G by the same quantity).

3.1 Reinforcement

We want to discuss conditions under which the optimal program gives larger (resp. smaller) incentives to work to the members of the couple when their partner works than when (s)he does not work. This is the property of *positive* (resp. *negative*) *reinforcement*, which in our setup comes down formally to the inequality

$$c^3 + c^0 \geq c^1 + c^2 \quad \text{(positive reinforcement)}, \tag{20}$$

$$c^3 + c^0 \leq c^1 + c^2 \quad \text{(negative reinforcement)}. \tag{21}$$

Consider a woman that decides to work in a family where the husband does not work. The family financial incentive to work is given by the difference between c^2 and c^0. Consider now the financial incentive to work for a woman when her husband already works. It is equal to the difference between c^3 and c^1. So, if $c^2 - c^0 > c^3 - c^1$ the woman tax on income is larger (i.e. the financial incentive to work is lower) when she belongs to a household where the husband already works. Using the definition of taxes, $t^i = \omega^i - c^i$, we have

$$t^1 + t^2 - t^3 - t^0 = \omega^1 + \omega^2 - \omega^3 - \omega^0 - c^1 - c^2 + c^3 + c^0.$$

Now under Assumption 3, $\omega^1 + \omega^2 - \omega^3 - \omega^0$ is equal to zero. Hence, if we define as *reinforcement term* the net sum $c^3 + c^0 - c^1 - c^2$, this term equals $t^1 + t^2 - (t^3 + t^0)$, and positive reinforcement refers to the inequality $t^1 - t^0 > t^3 - t^2$ or, equivalently if we look at agent 2: $t^2 - t^0 > t^3 - t^1$.

Remark 1 *Note that when the planner has no redistribution motives, that is when utilities are linear and the welfare function does not have redistributional motives, then the problem is to only maximize production efficiency. In this case, we have $P(A^i) = 1$ for $i = 0, 1, 2, 3$; $\omega^i - c^i + c^0 = 0$ and $c^3 = \omega^3 + c^0$. When $\omega^3 = \omega^1 + \omega^2$ we get neither positive nor negative reinforcement. Moreover, if in addition $\omega^0 = 0$ and we do not have any extra resources to pay the unemployed agents $(G = 0)$, we also have $c^0 = 0$.*

3.2 Graphical representation in the linear case

Consider Figure 1 which provides a graphical representation of the sets $A^i(c)$ in the model where there are no income effects and $N = I + 1$ (i.e., $u^i(c^i, \alpha) = c^i - \alpha^i(\alpha) = c^i - \alpha^i$) and $\alpha^0(\alpha) + \alpha^3(\alpha) = \alpha^1(\alpha) + \alpha^2(\alpha)$ for all α. The plan describes the values taken by the parameters α. The horizontal axis bears $x = \alpha^1 - \alpha^0$, equal to $\alpha^3 - \alpha^2$, while $y = \alpha^2 - \alpha^0$, equal to $\alpha^3 - \alpha^1$, is on the vertical axis. For any consumption vector, each set $A^i(c)$ is defined by three linear inequalities in the αs.[15] The two panels of Figure 1 represent the two typical cases of positive and negative reinforcements, positive reinforcement on the left, negative on the right. Recall indeed that negative reinforcement corresponds to the case $c^3 + c^0 < c^1 + c^2 \iff c^3 - c^1 < c^2 - c^0 \iff c^3 - c^2 < c^1 - c^0$. A useful feature of this case is that *negative reinforcement implies that there is no border between the sets A^3 and A^0.*[16]

---

[15]For instance, $A^0(c)$ is the set of αs such that

$$c^0 - \alpha^0 > c^1 - \alpha^1, \quad c^0 - \alpha^0 > c^2 - \alpha^2, \quad c^0 - \alpha^0 > c^3 - \alpha^3.$$

Using the equality $\alpha^0 + \alpha^3 = \alpha^1 + \alpha^2$, this can be rewritten as

$$c^0 - c^1 > \alpha^0 - \alpha^1 = -x, \quad c^0 - c^2 > \alpha^0 - \alpha^2 = -y, \quad c^0 - c^3 > \alpha^0 - \alpha^3 = -x - y,$$

which yields

$$x > c^1 - c^0, \quad y > c^2 - c^0,$$

the final inequality being always satisfied if $x + y > c^1 + c^2 - 2c^0 > c^3 - c^0$, or $c^1 + c^2 > c^0 + c^3$, the case of negative reinforcement, while it binds when $c^1 + c^2 < c^0 + c^3$, the positive reinforcement case, where we need to keep

$$x + y > c^3 - c^0 = c^3 - c^2 + c^2 - c^0 = c^3 - c^1 + c^1 - c^0,$$

in the definition of the set $A^0(c)$.

[16]In Appendix B we describe a class of preferences - inclusive of the additive separable utility - that excludes the possibility of any border between A^3 and A^0 when negative reinforcement holds, under similar restrictions across the vectors of parameters α.

20

Positive reinforcement   **Negative reinforcement**

$\alpha^2 - \alpha^0 = \alpha^3 - \alpha^1$ $\alpha^2 - \alpha^0 = \alpha^3 - \alpha^1$

A^1 A^0 A^1 A^0

$c^3 - c^1$ $c^2 - c^0$

$c^2 - c^0$ $c^3 - c^1$

A^3 A^2 A^3 A^2

$c^1 - c^0$ $c^3 - c^2$ $c^3 - c^2$ $c^1 - c^0$ $\alpha^1 - \alpha^0 = \alpha^3 - \alpha^2$

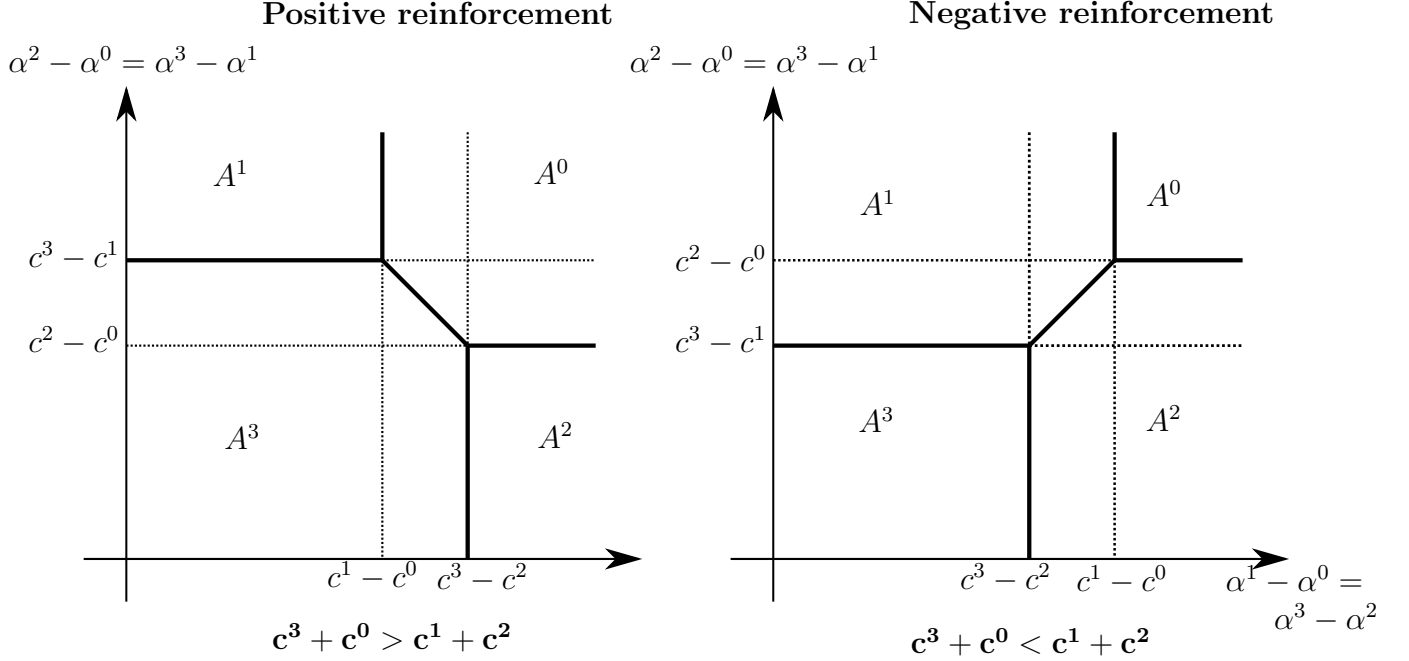$\mathbf{c^3 + c^0 > c^1 + c^2}$ $\mathbf{c^3 + c^0 < c^1 + c^2}$

Figure 1: *The right figure represents the case with positive reinforcement (subsidizing working couples) while the left one represents the case with negative reinforcement (taxing working couples).*

3.3 Optimality of subsidy to jointly working couples when single working couples are taxed

As it is transparent from the first order optimality conditions, the Pareto weights play a crucial role in delivering positive or negative reinforcement. For example, we expect that positive reinforcement will be favoured by a government that attaches large weights to joint working couples. This simple conjecture is confirmed by the next proposition.

Proposition 1 *Assume the optimal payments are such that $\omega^j - c^j + c^0 = t^j - t^0 \geq 0$ for $j = 1, 2$ and that under negative enforcement $\frac{\partial \mu(A^0(c))}{\partial c^3} = 0$. If at the optimum we have $P(A^3) \geq 1$, then the program must display positive reinforcement.*

Almost all proofs are reported in Appendix A.

Rawlsian Planner with no income effects   To show positive reinforcement, the previous proposition assumes that working couples are taxed. This assumption is implicit in the
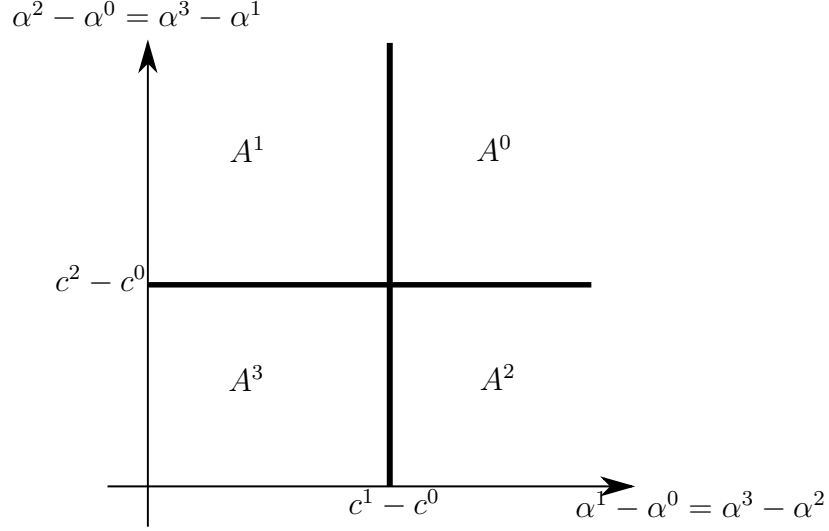
Figure 2: The symmetric case with independent taxation

nonnegative profit condition for each product in the industrial organization literature (e.g., Armstrong (2011)). More precisely, in the IO literature, it is assumed that the firm is making positive profits on the product, namely $\omega^i - c^i \geq 0$, and that $c^0 = 0$. These two assumptions imply, in particular, $\omega^i - c^i + c^0 \geq 0$. The IO literature also assumes Rawlsian planner and no income effects.[17] The proposition that follows builds on the analogy with the IO literature and shows a first important result regarding the *desirability of subsidizing joint working couples*:

Proposition 2 *Assume Rawlsian planner and no income effects. Then*

$$\omega^i - c^i + c^0 > 0 \quad for \quad i = 1, 2, 3. \tag{22}$$

If, in addition for $i = 1, 2$, the marginals F_i defined on $\alpha^i - \alpha^0$ for $i = 1, 2$, are independent and log-concave, positive reinforcement is optimal.

The proof of this proposition is based on a generalized version of the independence assumption we will make in Section 3.7 and in particular Lemma 11 in Appendix A. The intuition

---

[17]For details on the Rawlsian planner case see Appendix D.

regarding the optimality of the positive reinforcement is given in Figure 2. There we have drawn the optimal $c^1 - c^0$ and $c^2 - c^0$ lines in case the planner were aiming to maximize revenues under independent taxation, that is under the restrictions $c^1 - c^0 = c^3 - c^2$ and $c^2 - c^0 = c^3 - c^1$. The optimal level of taxes balances the cost of returns of any perturbation. In particular, consider a decrease in both c^1 and c^3. At the optimum, in order to have zero effect on total revenues, it must be that the reduction in revenues induced on the infra-marginals (those in areas A^1 and A^3) is fully balanced by the increase in revenues generated by the additional agents that from A^0 will move into A^1 (at the top of the figure) and from A^2 will move into A^3. That is, the vertical line will marginally move to the right. The fact that the distribution of $\alpha^1 - \alpha^0$ - indicated as $F_1(\cdot)$ in the proposition - is independent of the level of $\alpha^2 - \alpha^0$ implies that this aggregate argument is also valid line by line. Consider now the budget effects of a marginal decrease in c^3 alone. The area A^3 will expand. The horizontal expansion will be such that the reduction in budget is fully compensated by the new taxes paid by movers from A^2 into A^3. This change however, also expands A^3 vertically. There is hence a net increase in tax revenues generated by the couples moving from A^1 into A^3. This argument implies that a marginal decrease in c^3 for the level that is optional under independent taxation, increases revenues. The log concavity assumption implies that the argument is, in fact, global and at the optimum we have $c^3 - c^2 < c^1 - c^0$.

3.4   Non Discriminatory Condition

Another case where we can analyze the occurrence of positive or negative reinforcement in a relatively easy fashion is the situation where the two members of the couple pay the same taxes when they are single earners: $t^1 = t^2$. This is likely to occur at the optimum when the men and women have identical economic characteristics and therefore are treated equally in any gender neutral tax scheme. The restriction $t^1 = t^2$ to the optimal contract simplifies the mathematics and allows to concentrate on the fundamental aspects of reinforcement. In Section 4.2 we will consider the general case.

Assumption 4 (Non discriminatory condition) *The policy treats equally the single worker*

couples (gender neutral tax code) $t^2 = t^1 := t$.

When the distributions of costs $(\alpha_0, \alpha_1, \alpha_3)$ and $(\alpha_0, \alpha_2, \alpha_3)$ are identical and the two members of the couple have symmetric productivities (i.e., $\omega^1 = \omega^2 = \omega, \omega^3 = 2\omega$), the assumption $t^1 = t^2$ is without loss of generality. It turns out that we never formally use the symmetry of the distribution and of productivities either in our derivations. We may still have it in mind as a reference to give a simple economic interpretation to the non discriminatory condition.

Proposition 3 *Suppose we restrict the tax scheme to have $t^1 = t^2 = t$. If (t^3, t, t^0) is an optimal scheme, then the sign of the reinforcement term $2t - t^3 - t^0$ is the same as that of the following expression (evaluated at the optimum):*

$$R := \frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3} - \frac{\partial \mu(A^0)}{\partial c^3}}[P(A^3) - 1] + \frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0} - \frac{\partial \mu(A^3)}{\partial c^0}}[P(A^0) - 1]. \tag{23}$$

Remark 2 *Note that when both $P(A^0) = 1$ and $P(A^3) = 1$, the optimal program has zero reinforcement. Indeed, when the tax code is gender neutral, the two by two system of equations of the proof implies $X^0 = X^3 = 0$ and hence $t^0 - t^1 = t^3 - t^1 = 0$, so that all couples face the same taxes $t = G$.*

The expression (23) for the reinforcement term R is a weighted sum of 'corrected' elasticities for the measure of the sets $\mu(A^3)$ and $\mu(A^0)$. A nice intuition for the result - and alternative line of proof - is the following. Suppose $R > 0$. Consider now an increase in both c^3 and c^0 by $\frac{1}{\frac{\partial \mu(A^3)}{\partial c^3} - \frac{\partial \mu(A^0)}{\partial c^3}}$ and $\frac{1}{\frac{\partial \mu(A^0)}{\partial c^0} - \frac{\partial \mu(A^3)}{\partial c^0}}$, respectively. If this change had no effects on the budget, $R > 0$ implies that such perturbation would increase welfare. It can be verified directly that this particular perturbation is such that the number of people moving into A^3 from $A^1 \cup A^2$ is the same as those moving into A^0 (and all movements between A^1 and A^2 have no effects on the budget since $t^1 = t^2$). If negative reinforcement were true, by definition, such perturbation would relax the BC violating optimality. As a consequence we cannot have negative reinforcement when $R > 0$.

We now move to an interesting special case where the reinforcement term takes a simpler form and it is even easier to describe intuitively.

3.5 Exploring further the non discriminatory case

Recall the discussion around Figure 1. In the additive case, negative reinforcement implies that both $\frac{\partial \mu(A^3(c))}{\partial c^0} = 0$ and $\frac{\partial \mu(A^0(c))}{\partial c^3} = 0$. This observation leads to the following sufficient condition for positive reinforcement that comes as a corollary to Proposition 3.

Corollary 4 *Assume that whenever the scheme has negative reinforcement we have $\frac{\partial \mu(A^3(c))}{\partial c^0} = \frac{\partial \mu(A^0(c))}{\partial c^3} = 0$. If at the optimum we have:*[18]

$$\hat{R} := \frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3}}[P(A^3) - 1] + \frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}}[P(A^0) - 1] > 0, \tag{24}$$

then the tax scheme displays positive reinforcement.

Proof. Start from the expression (23). If negative reinforcement implies $\frac{\partial \mu(A^3(c))}{\partial c^0} = \frac{\partial \mu(A^0(c))}{\partial c^3} = 0$, negative reinforcement implies a negative $\hat{R}$. But it cannot be. Since reinforcement is either negative or positive, $\hat{R} > 0$ is a sufficient condition for positive reinforcement. **Q.E.D.**

The above condition bears obviously on endogenous variables. In Section 3.7 we provide assumptions on primitives that imply condition (24). In any case, condition (24) has a simple intuitive meaning we are going to describe. Have a look at the right panel of Figure 1. A nice feature of not having a direct border between A^3 and A^0 is that we can consider separately two key margins. We first have the margin where the household has to decide whether one partner enters into the labor force or remains with the other member out of the labor force. The second margin only regards couples that already have one partner working, and it regards the decision on whether the second partner should work as well or not. The key financial incentive related to the first margin is $c - c^0$ while the financial incentive governing the second margin is $c^3 - c$. Positive or negative reinforcement follows from the relationship between these two financial incentives.

---

[18]From condition (2) together with $\frac{\partial \mu(A^3(c))}{\partial c^0} = \frac{\partial \mu(A^0(c))}{\partial c^3} = 0$, under negative reinforcement, we have:

$$2\frac{\partial \mu(A)}{\partial c^0} = -\frac{\partial \mu(A^0)}{\partial c^0}, \quad \text{and} \quad 2\frac{\partial \mu(A)}{\partial c^3} = -\frac{\partial \mu(A^3)}{\partial c^3}.$$

We hence could get an equivalent expression by replacing these terms in (24).

In the standard setup, where one margin at a time is considered, the tax on workers would solve:

$$[P(A^0) - 1]\frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}} = \omega - (c - c^0) = (t - t^0).$$

Analogously, the 'sequential margins model' applied to the second decision, would deliver:

$$[P(A^3) - 1]\frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3}} = -[\omega^3 - (c^3 - c)] = -(t^3 - t),$$

where, recall, $t^3 = \omega^3 - c^3$, $t = \omega - c$, and $t^0 = -c^0$. Note that negative reinforcement is defined as $t^3 - t > t - t^0$, that is, the second margin is taxed more than the first margin. This arrangement is optimal when, as in Proposition 4, we have

$$[P(A^3) - 1]\frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3}} < -[P(A^0) - 1]\frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}}.$$

In other terms, negative reinforcement together with the lack of borders between A^0 and A^3 (hence right panel of Figure 1) allows us to derive a necessary condition for negative reinforcement (the negation of condition (24)) as a comparison between two conditions that look identical to a model where margins can be analyzed in sequence, as in the standard Mirrlees model. Negative reinforcement is hence analogous to 'regressivity', where high income earners (earning $\omega^3 = 2\omega$) pay a lower marginal tax on the last unit ω than the tax paid by low income earners (earning only ω). We know that in the single margin case the ratio $[P(A^j) - 1]\frac{\mu(A^j)}{\frac{\partial \mu(A^j)}{\partial c^j}}$ fully determines the *sign* of the tax, it is just natural that the relative size of the tax is given by the ratio of these two conditions. That is, it is natural to have expressions for the reinforcement term that are described as ratios between Pareto weights and elasticities. That is precisely the way we restate Corollary 4 below:

Corollary 5 *Define*

$$r := \frac{\frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}}}{\frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3}}} = \frac{\frac{\partial \log \mu(A^3)}{\partial c^3}}{\frac{\partial \log \mu(A^0)}{\partial c^0}} \geq 0,$$

computed at the optimal program. If the Pareto weights satisfy

$$P(A^3) - 1 \geq -r[P(A^0) - 1], \tag{25}$$

the optimal program exhibits positive reinforcement.

Remark 3 *Under the conditions of Corollary 4, since r is positive, condition (25) implies that when both $P(A^0) \geq 1$ and $P(A^3) \geq 1$, the optimal program has positive reinforcement.*

Remark 3 indicates that we have positive reinforcement when the government cares a lot for the unemployed (type 0) and fully working couples (type 3). It then probably tax single worker families. This is confirmed by the optimality conditions. If we set $P(A^0) \geq 1$ in the first order condition with respect to c^0 and we assume $t^2 = t^1$, since $\frac{\partial \mu(A(c))}{\partial c^0} < 0$, implies $\omega - (c - c^0) = t - t^0 \geq 0$, that is, single working households are taxed compared to non-working households. Since $P(A^3) \geq 1$ implies $t^3 - t \leq 0$, we obtain positive reinforcement.

3.6 Pareto weights and reinforcement: a graphical representation

It is instructive to represent graphically the set of Pareto weights that lead to positive reinforcement, according to Remark 3, in the simple linear pecuniary setup *excluding income effects*. This requires

$$\mu(A^0)[P(A^0) - 1] + \mu(A)[P(A) - 1] + \mu(A^3)[P(A^3) - 1] = 0.$$

We provide two representations. The first one focuses on the weights of the unemployed A^0 and of the full working A^3 households, using the identity to eliminate the single worker households. The second uses a 'Machina' triangle and treats symmetrically A^0, A and A^3.

Figure 3 represents condition (25) in the $(P(A^0) - 1, P(A^3) - 1)$ plane. The social weights being non negative, the domain is contained in the orthant $[P(A^0) - 1 \geq -1, P(A^3) - 1 \geq -1]$. It is bounded above by the line

$$\mu(A^0)[P(A^0) - 1] - \mu(A) + \mu(A^3)[P(A^3) - 1] \leq 0,$$

or

$$\mu(A^0)P(A^0) + \mu(A^3)P(A^3) \leq 1.$$

This is the dotted line in Figure 3. It joins the representative points $\mathcal{R}^0$ of Rawls where $P(A^0) - 1 = \frac{1 - \mu(A^0)}{\mu(A^0)}$, and $P(A) = P(A^3) = 0$ (hence $P(A^3) - 1 = -1$); and the other extreme

27

case where the social planner is only interested in the working couples, denoted as $\mathcal{R}^3$, where $P(A^3) - 1 = \frac{1-\mu(A^3)}{\mu(A^3)}$, and $P(A) = P(A^0) = 0$ (hence $P(A^0) - 1 = -1$). The feasible set of social weights is the triangle comprised between the axes and the dotted line. Positive reinforcement takes place in the portion of the domain delimited by the solid line defined by (25). The slope of the solid line is $-r$ while that of the dotted line is $-\mu(A^0)/\mu(A^3)$, i.e. the product of $-r$ by the positive quantity

$$\frac{\frac{\partial \mu(A^0)}{\partial c^0}}{\frac{\partial \mu(A^3)}{\partial c^3}}.$$

Interestingly, for $P(A^3)$ smaller than 1, even for $P(A^3)$ equal to zero, this domain is non-empty. Moreover, it is easy to see that when $P(A^3) < 1$ the set of $P(A^0)$ for which there is positive reinforcement increases with r, all other things equal.

In Figure 4, we give an alternative presentation. The social weights are embedded into a three dimensional positive orthant. The three dimensional space is intersected with the plan

$$\mu(A^0)P(A^0) + \mu(A)P(A) + \mu(A^3)P(A^3) = 1.$$

Figure 4 represents this intersection. The three summits of the triangle are respectively $(0, 0, 1/\mu(A^3))$, $(1/\mu(A^0), 0, 0)$ and $0, 1/\mu(A), 0)$. Equation (25) is a plan in the three dimensional space: its trace on the triangle is the solid line. It intersects the side $P(A^0) = 0$ of the triangle at a point where $P(A^3) = 1 + r$, according to (25), and the side $P(A^3) = 0$ at a point where $P(A^0) = 1 + 1/r$. The region of positive reinforcement can be identified by computing the direction of the inequality sign of (25) at the summits of the triangle.

3.7 The role of the distribution of costs: rawlsian planner, symmetry and linear utilities

Corollary 5 and the remark that follows show that positive reinforcement is compatible with a very small $P(A^3)$, even $P(A^3) = 0$. A Rawlsian planner would take $P(A) = P(A^3) = 0$. If in addition, we have no income effects then it must be that $P(A^0) = 1/\mu(A^0)$ and hence $P(A^0) - 1 = [1 - \mu(A^0)]/\mu(A^0)$.[19] In this case we have the following specialization of
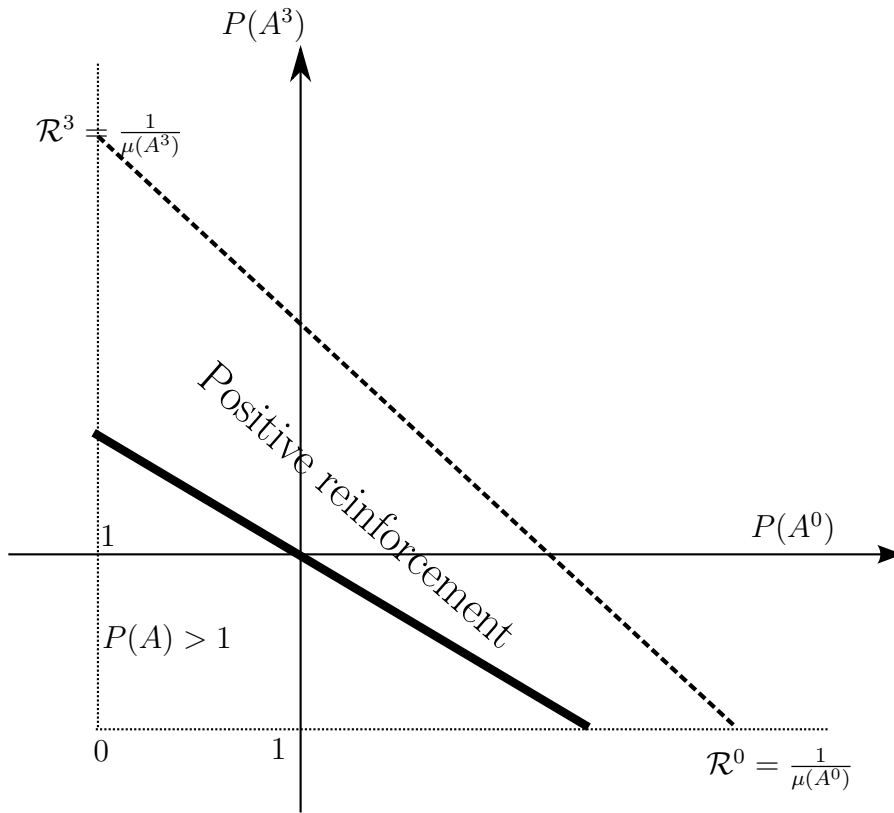
---

[19]For details on the Rawlsian planner case see Appendix D.

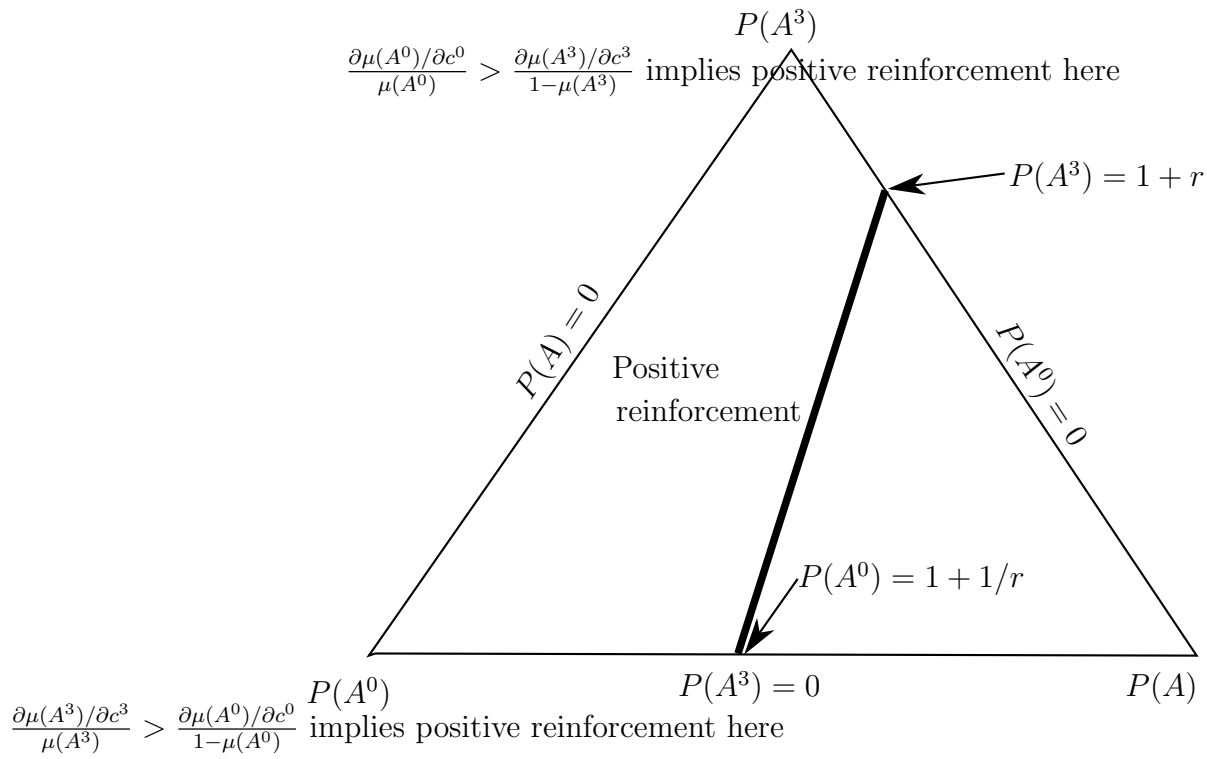Figure 3: Two dimensional representation of sufficient conditions for positive reinforcement

$$\frac{\partial \mu(A^0)/\partial c^0}{\mu(A^0)} > \frac{\partial \mu(A^3)/\partial c^3}{1-\mu(A^3)} \text{ implies positive reinforcement here}$$

$P(A^3)$

$P(A^3) = 1 + r$

$P(A) = 0$

$P(A^0) = 0$

Positive reinforcement

$P(A^0) = 1 + 1/r$

$P(A^0)$

$P(A^3) = 0$

$P(A)$

$$\frac{\partial \mu(A^3)/\partial c^3}{\mu(A^3)} > \frac{\partial \mu(A^0)/\partial c^0}{1-\mu(A^0)} \text{ implies positive reinforcement here}$$

Figure 4: A three dimensional representation of the social weights

Proposition 3:

Remark 4 *Consider the model with no income effects, and a Rawlsian planner. Positive reinforcement is optimal whenever*

$$r = \frac{\dfrac{\partial \log \mu(A^3)}{\partial c^3}}{\dfrac{\partial \log \mu(A^0)}{\partial c^0}} \geq \frac{\mu(A^0)}{1 - \mu(A^0)}.$$

Later in this section we give conditions on the distribution of α's for the inequality to be satisfied. A simple manipulation shows that the inequality can be rewritten as

$$\frac{\dfrac{\partial \log \mu(A^3)}{\partial c^3}}{-\dfrac{\partial \log[1 - \mu(A^0)]}{\partial c^0}} \geq 1 \tag{26}$$

Positive reinforcement is therefore optimal when the semi-elasticity of the measure of the set A^3 with respect to c^3 is larger than the opposite of the semi-elasticity of the measure of the set $A \cup A^3$ with respect to c^0. The interpretation for this inequality is better seen from considering Figure 1. The numerator is equal to the semi-elasticity of the measure of the households in A^3 with respect to their after tax income c^3. The denominator is the semi-elasticity of the measure of the households in $A \cup A^3$ with respect to a decrease in c^0, the after tax income they would get if they decided to join the unemployed. The condition says that positive reinforcement prevails when the numerator is larger than the denominator, the elasticity of the smaller set is larger than the one of the bigger (inclusive) one.

Distributional Assumptions: Further Analysis   We now investigate the meaning and plausibility of restrictions on the semi elasticities in (26). Define the function

$$\Phi(x, y) = \Pr\left\{\alpha^1 - \alpha^0 \leq x \ \& \ \alpha^2 - \alpha^0 \leq x \mid \alpha^1 - \alpha^0 \leq y \ \text{ OR } \ \alpha^2 - \alpha^0 \leq y\right\}.$$

Note that $\Phi(x, y) \leq 1$.

Assumption 5 *Assume that $\Phi(x, x + \varepsilon)$ is (i) increasing [(ii) decreasing] in x for all $\varepsilon \geq 0$.*

To give an interpretation of Assumption 5, note that

$$\Phi(x, x+\varepsilon) = \frac{\mu(A^3)(x)}{\mu(A \cup A^3)(x+\varepsilon)},$$

which refers to the right panel of Figure 1 in the symmetric case, assuming that $c^1 = c^2$ hence $c^3 - c^1 = c^3 - c^2 = x$ and $c^1 - c^0 = c^2 - c^0 = x + \varepsilon$. Assumption 5(i) hence says that as we increase c^3 and decrease c^0 by the same amount, the measure $\mu(A^3)(x)$ increases with x faster than the measure $\mu(A \cup A^3)(x+\varepsilon)$. This condition can be interpreted as asserting that assortative mating in couples increases with skills (we can interpret high skilled people those with low working cost). Formally, we have

Lemma 6 *Assume negative reinforcement, symmetry and Assumption 5 (i) [resp. (ii)], then*

$$\frac{\frac{\partial \mu(A^3(c))}{\partial c^3}}{\mu(A^3(c))} - \frac{\frac{\partial \mu(A^0(c))}{\partial c^0}}{1 - \mu(A^0(c))} \geq 0 \text{ [resp.} \leq 0] \qquad \text{for all } c.$$

To repeat the interpretation, the meaning of condition:

$$\frac{\frac{\partial \mu(A^3)}{\partial c^3}}{\mu(A^3)} - \frac{\frac{\partial \mu(A^0)}{\partial c^0}}{[1 - \mu(A^0)]} \geq 0,$$

is as follows. The first term $\frac{\frac{\partial \mu(A^3)}{\partial c^3}}{\mu(A^3)}$ represents the (positive) hazard rate of joint working couples with respect to the level of their after tax income. The second term $-\frac{\frac{\partial \mu(A^0)}{\partial c^0}}{[1-\mu(A^0)]} = \frac{\sum_{j\neq 0} \frac{\partial \mu(A^j)}{\partial c^0}}{[1-\mu(A^0)]} = \frac{\frac{\partial[1-\mu(A^0)]}{\partial c^0}}{[1-\mu(A^0)]}$ represents the (negative) hazard rate of the number of all working couples (with either one member or both members at work) with respect to the level of unemployment benefits. This condition speaks to perturbations that increase both c^3 and c^0 by the same amount: it says that the first hazard rate dominates the second one.

Suppose now that $\Phi(x, x)$ is increasing in x. The question we ask is: what extra assumptions are required to get Assumption 5 *(i)*? We use the log separability of Φ.

Lemma 7 *Assume $\Phi(\cdot, \cdot)$ is increasing along the ray $x = y$. In addition, assume $[1 - \mu(A^0)](\cdot)$ is log-concave, or $\mu(A^3)(\cdot)$ is log-convex, or both. Then Assumption 5 (i) is satisfied.*

As a corollary, we also have the case considered in Proposition 2.

Corollary 8 *Assume for $i = 1, 2$, that $\alpha^i - \alpha^0$ are distributed independently and symmetrically and that the cumulate F is log concave. Then Assumption 5 (i) is satisfied.*

We now get the same result with an alternative assumption.

Assumption 6 *Define the following function:*

$$\Psi(x, y) = \Pr\left\{\alpha^1 - \alpha^0 \geq x \ \& \ \alpha^2 - \alpha^0 \geq x \mid \alpha^1 - \alpha^0 \geq y \ OR \ \alpha^2 - \alpha^0 \geq y\right\} = \frac{\mu(A^0)(x)}{1 - \mu(A^3)(y)}.$$

Assume $\Psi(y + \varepsilon, y)$ (i) decreases [(ii) increases] with y for all $\varepsilon \geq 0$.

Note that under independence Assumption 6*(i)* is satisfied for $\varepsilon = 0$.

Proposition 9 *Suppose that Assumptions 7 (in Appendix B) and 3 hold and that $P(A^0) \geq 1$ [resp. $P(A^0) \leq 1$]. If in addition, Assumption 6 (i) [resp. Assumption 6 (ii)] holds and $\mu(A^3))P(A^3) + (1 - \mu(A^3))P(A^0) \geq 1$ then only $c^3 + c^0 > c^1 + c^2$ (positive reinforcement) can be optimal.*

Remark 5 *It is easy to see that we might be able to find some weakening of the condition $\mu(A^3))P(A^3) + (1 - \mu(A^3))P(A^0) \geq 1$ using some appropriate weights γ^0 and γ^3. The condition we just stated on the Pareto weights implies $P(A^0) \geq \frac{\mu(A^1)P(A^1) + \mu(A^2)P(A^2)}{\mu(A^1) + \mu(A^2)}$ that under symmetry boils down to $P(A^0) \geq P(A)$. See Lemma 11 in Appendix A for details.*

4   Extensions

In this section we consider two generalizations to our results. The first one maintains the assumption of non-discriminatory taxation and generalizes formula to the case with many productivities and Assumption 4, which in this context can be interpreted as perfect assortative matching. In Section 4.2 we discuss the general solution to our model of couple taxation allowing for asymmetry within the couple.

4.1 The Case with Many Productivities

Consider now the possibility of households with heterogenous gross productivities. We assume perfect assortative mating. This, together with Assumption 3, implies:[20] $\omega^0 = 0, \omega^1 = \omega^2 = \omega, \omega^3 = 2\omega$, where we abused in notation and use ω to indicate the family productivity. Because of perfect assortative mating in the economy, whenever one partner works, all about the family is known. On the other hand, when both partners are unemployed productivity is not observable.

Proposition 10 *Fix ω and assume $t^1(\omega) = t^2(\omega)$. The sign of the reinforcement term for agents ω is the same as that of the following expression:*

$$\sum_{i=1,2} \left[\frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^0)}{\partial c^i(\omega)} \right] \times RR(\omega),$$

where

$$RR(\omega) := \frac{\mu(A^3(\omega))[P(A^3(\omega)) - 1]}{\frac{\partial \mu(A^3(\omega))}{\partial c^3(\omega)} - \frac{\partial \mu(A^0)}{\partial c^3(\omega)}} + \frac{\mu(A(\omega))[P(A(\omega)) - 1]}{\sum_{i=1,2} \left[\frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^0)}{\partial c^i(\omega)} \right]}.$$

The intuition for the $RR(\omega)$ term is as in the single ω case. We just wrote it using the set $A(\omega) = A^1(\omega) \cup A^2(\omega)$ because the set A^0 is common to all ω's. The extra complication is that now the sign of $\sum_{i=1,2} \left[\frac{\partial \mu(A^0(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)} \right] = \sum_{i=1,2} \frac{\partial \mu(A(\omega))}{\partial c^i(\omega)} + 2 \frac{\partial \mu(A^0(\omega))}{\partial c^i(\omega)}$ is, in general, indeterminate. Note that the difference $\frac{\partial \mu(A^0(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)}$ is related to the reinforcement term since agents moving from A^0 into A^1 pay $t - t^0$ while agents going from A^1 into A^3 cost $t^3 - t$. So this measures how many people have a net return of $t - t^0 - (t^3 - t)$ - that is, the negative of the reinforcement term - for the planner (in terms of budget consequences due to behavioral reactions alone).

We can ask questions such as: *is the reinforcement term increasing or decreasing with the skill of the couple?* Simulations would be valuable in this case and are left to the future.

4.2 General Solution

After gaining intuition on a few special cases, we turn to the general formula. The following delivers the solution for the taxes and reinforcement. We now specialize the system (9) to the

---

[20] A similar analysis holds in a 'gender gap' model, where $\omega^1 = \kappa \omega^2$, with $\kappa \neq 1$ constant across families.

case of optimal taxation of couples:

$$\Delta t = H^{-1}(\mu(P-1)),$$

where $\Delta t := \left(t^1 - t^0, t^2 - t^0, t^3 - t^0\right), \mu(P-1) := \left(\mu(A^1)[P(A^1) - 1], \mu(A^2)[P(A^2) - 1], \mu(A^3)[P(A^3) - 1]\right),$

and

$$H = \begin{bmatrix} -\frac{\partial \mu(A^1)}{\partial c^1} & -\frac{\partial \mu(A^2)}{\partial c^1} & -\frac{\partial \mu(A^3)}{\partial c^1} \\ -\frac{\partial \mu(A^1)}{\partial c^2} & -\frac{\partial \mu(A^2)}{\partial c^2} & -\frac{\partial \mu(A^3)}{\partial c^2} \\ -\frac{\partial \mu(A^1)}{\partial c^3} & -\frac{\partial \mu(A^2)}{\partial c^3} & -\frac{\partial \mu(A^3)}{\partial c^3} \end{bmatrix}.$$

Recall that the value of $\mu(A^0)[P(A^0) - 1]$ and multiplier must be co-determined using the budget constraints and the condition for the weighted sum of the Pareto weights in case of income effects. The reinforcement term can be obtained by the following linear combination of the three rows of the solution: $(1, 1, -1)$. Note indeed that all these values can be seen as differences with respect to t^2, so $t^1 - t^0 + t^2 - t^0 - (t^3 - t^0) = t^1 - t^0 - (t^3 - t^2)$, which is our reinforcement term.

If we *assume away income effects* we gain some symmetries and the Hessian becomes

$$H = \begin{bmatrix} a_1 & -b & -e \\ -b & a_2 & -d \\ -e & -d & a_3 \end{bmatrix},$$

and hence

$$H^{-1} = \frac{1}{\Delta} \begin{bmatrix} a_2 a_3 - d^2 & ed + a_3 b & a_2 e + bd \\ ed + a_3 b & a_1 a_3 - e^2 & a_1 d + eb \\ a_2 e + bd & a_1 d + eb & a_2 a_1 - b^2 \end{bmatrix}.$$

Note that - of course- $\mu(A^0)[P(A^0) - 1]$ will not appear since it is residual as the weighted sum of Pareto weight equals one when there are no income effects. As we discussed above, in the pecuniary model we have: $\sum_{i=0,1,2,3} \mu(A^i)[P(A^i) - 1] = 0$. Suppose now that property P is satisfied. *Under negative reinforcement* we have $\frac{\partial \mu(A^3)}{\partial c^0} = 0$ and hence $\frac{\partial \mu(A^3)}{\partial c^3} = \frac{\partial \mu(A^1)}{\partial c^3} + \frac{\partial \mu(A^2)}{\partial c^3}$. In terms of our matrix H we have $a_3 = -(e + d)$ and hence the (reduced) matrix of the first

order conditions becomes:

$$H = \begin{bmatrix} a_1 & -b & -e \\ -b & a_2 & -d \\ -e & -d & -(e+d) \end{bmatrix}$$

and hence the expression for the inverse gets simplified accordingly.[21]

4.3 Couple Taxation with Intensive Margin (Kleven, Kreiner, and Saez (2009))

In terms of notation, it will be better to consider both subscripts and superscripts. Suppose the primary earner can choose among N occupations of increasing intensity of work, while the secondary earner can only choose wether to work or not work. The primary earner's level of intensity will be indexed by $i = 1, 2, 3, \ldots, N$ while the occupation of the secondary earner will be indexed by $\ell = 0, 1$. Overall, we have $2 \times N$ occupations. We will denote by c_ℓ^i the consumption associated to occupations (i, ℓ). It will be natural to assume that the production is additive, namely, the income generated in occupation (i, ℓ) will be $\omega_\ell^i = \omega^i + \overline{\omega}\ell$. We have $\omega^i - c_\ell^i - (\omega^{i-1} - c_\ell^{i-1}) := t_\ell^i - t_\ell^{i-1}$. The concept of jointness in Kleven, Kreiner, and Saez (2009)) regards determining a relationship between $t_1^i - t_1^{i-1}$ and $t_0^i - t_0^{i-1}$ which holds for all i. In particular, they define positive jointness the situation where $t_1^i - t_1^{i-1} > t_0^i - t_0^{i-1}$ for all i and negative jointness when the inequality is reversed for all i. It is easy to see that the first order condition with respect to $c^i\ell$ solves:

$$\mu(A_\ell^i)[P(A_\ell^i) - 1] = \frac{\partial\mu(A_\ell^{i-1})}{\partial c_\ell^i}[t_\ell^i - t_\ell^{i-1}] + \frac{\partial\mu(A_\ell^{i+1})}{\partial c_\ell^i}[t_\ell^i - t_\ell^{i+1}] + \sum_{k \in K(i,\ell)} \frac{\partial\mu(A_{-\ell}^k)}{\partial c_\ell^i}[t_\ell^i - t_{-\ell}^k].$$

where $-\ell$ indicates the complement of ℓ, and $K(i, \ell)$ is the set of occupations with a border with (i, ℓ). We can again recover the optimal tax formula proposed by the authors by considering a

---

[21]We have

$$H^{-1} = \frac{1}{\Delta} \begin{bmatrix} -d^2 - a_2 d - a_2 e & (d-b)e - bd & a_2 e + bd \\ (d-b)e - b*d & -e^2 - a_1 e - a_1 d & be + a_1 d \\ a_2 e + bd & be + a_1 d & -b^2 + a_1 a_2 \end{bmatrix}.$$

change in c^s by the same amount for $s \geq i$. Assuming linear utility, this perturbation delivers our version of the formula in the paper:

$$t_\ell^i - t_\ell^{i-1} = \sum_{s=i}^I \frac{\mu(A_\ell^s)}{\frac{\partial \mu(A_\ell^i)}{\partial c_\ell^i}} \left[1 - P(A_\ell^s) + \sum_{k \in K(s,\ell)} \frac{\frac{\partial \mu(A_{-\ell}^k)}{\partial c^s}}{\mu(A_\ell^s)}(t_\ell^s - t_{-\ell}^k) \right].$$

Again, our general formula will deliver expressions for taxes as functions of measures and Pareto weights and not implicit as it is here. It would be just a matter of matrix inversion. It is immediate to extend this to the case where the secondary earner is allowed to choose more than two levels of intensity.

References

ALES, L., M. KURNAZ, AND C. SLEET (2015): "Technical Change, Wage Inequality, and Taxes," *American Economic Review*, 105(10), 3061–3101.

ARMSTRONG, M. (2011): "Bundling Revisited: Substitute Products and Inter-Firm Discounts," Discussion Paper 574, University of Oxford.

CHONÉ, P., AND G. LAROQUE (2005): "Optimal Incentives for Labor Force Participation," *Journal of Public Economics*, 89, 395–425.

——— (2011): "Optimal taxation in the extensive model," *Journal of Economic Theory*, 146(2), 425–453.

FRANKEL, A. (2014): "Taxation of Couples under Assortative Mating," *American Economic Journal: Economic Policy*, 6, 155–177.

GOMES, R., J.-M. LOZACHMEUR, AND A. PAVAN (2016): "Differential Taxation and Occupational Choice," Discussion paper, Toulouse School of Economics.

IMMERVOLL, H., H. KLEVEN, C. KREINER, AND N. VERDELIN (2011): "Optimal tax and transfer programmes for couples with extensive labor supply responses," *Journal of Public Economics*, 95, 1485–1500.

Kleven, H., C. Kreiner, and E. Saez (2009): "The optimal income taxation of couples," *Econometrica*, 77, 537–560.

Lockwood, B., C. Nathanson, and G. Weyl (forthcoming): "Taxation and the Allocation of Talent," *Journal of Political Economy.*

McAfee, R. P., J. McMillan, and M. Whinston (1989): "Multiproduct Monopoly, Commodity Bundling and Correlation of Values," *Quarterly Journal of Economics*, 104, 371–384.

Rothschild, C., and F. Scheuer (2013): "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 69, 277–311.

——— (2015): "A Theory of Income Taxation under Multidimensional Skill Heterogeneity," Discussion Paper 5165, CESifo working paper.

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073.

Salinger, M. (1995): "A Graphical Analysis of Bundling," *The Journal of Business*, 68, 85–98.

Spence, A. M. (1980): "Multi-product Quantity-dependent Prices and Profitably Constraints," *Review of Economic Studies*, 47, 821–842.

A Proofs

Proof of Proposition 1.   We will show it by contradiction. So assume negative reinforcement, namely $t^3 - t^1 - (t^2 - t_0) > 0$. Recall the first order condition for c^3 in (7) for $I = 2$. In the case of negative reinforcement and the assumption in the proposition reads:

$$\mu(A^3)[P(A^3) - 1] = \sum_{j=0,1,2} \frac{\partial \mu(A^j(c))}{\partial c^3}[t^3 - t^j] \tag{27}$$

Now, since $\frac{\partial \mu(A^j(c))}{\partial c^3} \leq 0$, if $P(A^3) \geq 1$ it must be that for some $j \in \{1, 2\}$

$$t^3 - t^j \leq 0. \tag{28}$$

Suppose without loss of generality that $j = 1$. Since $t^2 - t^0 \geq 0$, from (28), we have

$$t^3 - t^1 - (t^2 - t^0) \leq 0,$$

which generates a contradiction to the assumed negative reinforcement. This hence implies that the optimal contract must have positive reinforcement. ∎

Proof of Proposition 2. See right after Lemma 11.

Proof of Proposition 3. The first order conditions (7) become

$$
\begin{aligned}
\mu(A^0)[P(A^0) - 1] &= & &\frac{\partial \mu(A^1)}{\partial c^0}(t^0 - t^1) &+\frac{\partial \mu(A^2)}{\partial c^0}(t^0 - t^2) &+\frac{\partial \mu(A^3)}{\partial c^0}(t^0 - t^3) \\
\mu(A^1)[P(A^1) - 1] &= \frac{\partial \mu(A^0)}{\partial c^1}(t^1 - t^0) & &&+\frac{\partial \mu(A^2)}{\partial c^1}(t^1 - t^2) &+\frac{\partial \mu(A^3)}{\partial c^1}(t^1 - t^3) \\
\mu(A^2)[P(A^2) - 1] &= \frac{\partial \mu(A^0)}{\partial c^2}(t^2 - t^0) &+\frac{\partial \mu(A^1)}{\partial c^2}(t^2 - t^1) & &&+\frac{\partial \mu(A^3)}{\partial c^2}(t^2 - t^3) \\
\mu(A^3)[P(A^3) - 1] &= \frac{\partial \mu(A^0)}{\partial c^3}(t^3 - t^0) &+\frac{\partial \mu(A^1)}{\partial c^3}(t^3 - t^1) &+\frac{\partial \mu(A^2)}{\partial c^3}(t^3 - t^2). & &
\end{aligned}
\tag{29}
$$

We rearrange the first and last rows in order to get rid of t^2 (in the first row, we add to both sides $\partial \mu(A^2)/\partial c^0 (t^2 - t^1)$, in the last row we add $\partial \mu(A^2)/\partial c^3 (t^2 - t^1)$) we get:

$$
\begin{aligned}
\mu(A^0)[P(A^0) - 1] + \frac{\partial \mu(A^2)}{\partial c^0}(t^2 - t^1) &= \left[\frac{\partial \mu(A^1)}{\partial c^0} + \frac{\partial \mu(A^2)}{\partial c^0}\right](t^0 - t^1) + \frac{\partial \mu(A^3)}{\partial c^0}(t^0 - t^3) \\
\mu(A^3)[P(A^3) - 1] + \frac{\partial \mu(A^2)}{\partial c^3}(t^2 - t^1) &= \frac{\partial \mu(A^0)}{\partial c^3}(t^3 - t^0) + \left[\frac{\partial \mu(A^1)}{\partial c^3} + \frac{\partial \mu(A^2)}{\partial c^3}\right](t^3 - t^1).
\end{aligned}
$$

Let $X^0 = \mu(A^0)(P(A^0) - 1) + \frac{\partial \mu(A^2)}{\partial c^0}(t^2 - t^1)$ and $X^3 = \mu(A^3)(P(A^3) - 1) + \frac{\partial \mu(A^2)}{\partial c^3}(t^2 - t^1)$. The system can be considered as a system of two equations in $t^0 - t^1$ and $t^3 - t^1$. A simple manipulation yields

$$
\begin{aligned}
X^0 &= \left[\frac{\partial \mu(A^1)}{\partial c^0} + \frac{\partial \mu(A^2)}{\partial c^0} + \frac{\partial \mu(A^3)}{\partial c^0}\right](t^0 - t^1) - \frac{\partial \mu(A^3)}{\partial c^0}(t^3 - t^1) \\
X^3 &= -\frac{\partial \mu(A^0)}{\partial c^3}(t^0 - t^1) + \left[\frac{\partial \mu(A^0)}{\partial c^3} + \frac{\partial \mu(A^1)}{\partial c^3} + \frac{\partial \mu(A^2)}{\partial c^3}\right](t^3 - t^1),
\end{aligned}
$$

or

$$\frac{\partial\mu(A^0)}{\partial c^0}(t^0 - t^1) + \frac{\partial\mu(A^3)}{\partial c^0}(t^3 - t^1) = -X^0$$

$$\frac{\partial\mu(A^0)}{\partial c^3}(t^0 - t^1) + \frac{\partial\mu(A^3)}{\partial c^3}(t^3 - t^1) = -X^3.$$

The solution of this two by two system is

$$t^0 - t^1 = \frac{1}{\Delta}\left[-\frac{\partial\mu(A^3)}{\partial c^3}X^0 + \frac{\partial\mu(A^3)}{\partial c^0}X^3\right]$$

$$t^3 - t^1 = \frac{1}{\Delta}\left[\frac{\partial\mu(A^0)}{\partial c^3}X^0 - \frac{\partial\mu(A^0)}{\partial c^0}X^3\right],$$

where Δ is the determinant $\partial\mu(A^0)/\partial c^0\partial\mu(A^3)/\partial c^3 - \partial\mu(A^3)/\partial c^0\partial\mu(A^0)/\partial c^3 > 0.$[22] If we force the no-discriminatory condition $t^2 = t^1 := t,$[23] the sum of the two equations yields

$$t^3 - 2t + t^0 = \frac{\left[\left(\frac{\partial\mu(A^3)}{\partial c^0} - \frac{\partial\mu(A^0)}{\partial c^0}\right)\mu(A^3)[P(A^3) - 1]\right] + \left[\left(\frac{\partial\mu(A^0)}{\partial c^3} - \frac{\partial\mu(A^3)}{\partial c^3}\right)\mu(A^0)[P(A^0) - 1]\right]}{\Delta},$$

If we divide both sides by the positive number

$$\left(\partial\mu(A^3)/\partial c^0 - \partial\mu(A^0)/\partial c^0\right) \times \left(\partial\mu(A^0)/\partial c^3 - \partial\mu(A^3)\partial c^3\right)$$

we have shown our result. ∎

Proof of Lemma 6.   We show it for case (i). Since we are in the negative reinforcement case, and symmetric, the number $\mu(A^3)$ can be written as a sole function of one parameter.

$$\mu(A^3)(x) = \Pr\{\alpha^1 - \alpha^0 \leq x \ \& \ \alpha^2 - \alpha^0 \leq x\},$$

similarly, under negative reinforcement we can define the measure $1 - \mu(A^0)$ as a function of one parameter,

$$[1 - \mu(A^0)](y) = \Pr\{\alpha^1 - \alpha^0 \leq y \ OR \ \alpha^2 - \alpha^0 \leq y\}.$$

---

[22]All cross-elements have negative sign while the own derivatives have positive sign. From equation (2) we know that the cross elements are dominated by the own elements in absolute value.

[23]Note that if we impose the constraint $t^1 = t^2$ to the system, we would only have a multiplier in the second and third rows of the first order conditions (of course entering with a different sign in the two lines), but the first order conditions for c^0 and c^3 will not be affected.

Now, as mentioned above, our function Φ solves

$$\Phi(x, y) = \frac{\mu(A^3(x))}{\mu(A \cup A^3(y))}.$$

In addition, Φ is log separable. According to this new notation, we have

$$\frac{\frac{\partial \mu(A^3)(c)}{\partial c^3}}{\mu(A^3)(c)} = \frac{d}{dc^3} \ln[\mu(A^3(c^3 - c))] = \frac{d}{dx} \ln\left(\Phi(c^3 - c, c - c^0)\right).$$

and

$$\frac{-\frac{\partial \mu(A^0)(c)}{\partial c^0}}{[1 - \mu(A^0)(c)]} = -\frac{d}{dc^0} \ln[1 - \mu(A^0(c - c^0))] = \frac{d}{dy} \ln\left(\Phi(c^3 - c, c - c^0)\right)$$

Then we can write the required inequality as:

$$\frac{d}{dx} \ln\left(\Phi(c^3 - c, c - c^0)\right) + \frac{d}{dy} \ln\left(\Phi(c^3 - c, c - c^0)\right) \geq 0.$$

Finally, if we define $\varepsilon := c - c^0 - (c^3 - c) \geq 0$ we have that the above condition is guaranteed by $\ln\left(\Phi(x, x + \varepsilon)\right)$ being increasing in x for all $\varepsilon \geq 0$. $\blacksquare$

Proof of Lemma 7. For $\varepsilon \geq 0$ we have

$$\Phi(x, x + \varepsilon) = \frac{\mu(A^3)(x)}{[1 - \mu(A^0)](x + \varepsilon)}.$$

Hence for $\varepsilon \geq 0$

$$\frac{d}{dx} \ln\left(\Phi(x, x + \varepsilon)\right) \geq \frac{d}{dx} \ln\left(\Phi(x, x)\right)$$

whenever $\frac{d}{dx} \ln[1 - \mu(A^0)](x + \varepsilon)$ decreases with ε.

$$\Phi(y - \varepsilon, y) = \frac{\mu(A^3)(y - \varepsilon)}{[1 - \mu(A^0)](y)}.$$

Hence for $\varepsilon \geq 0$

$$\frac{d}{dy} \ln\left(\Phi(y - \varepsilon, y)\right) \geq \frac{d}{dy} \ln\left(\Phi(y, y)\right)$$

whenever $\frac{d}{dx} \ln \mu(A^3)(y - \varepsilon)$ increases with ε. $\blacksquare$

Proof of Lemma 8. Note that under independence (and symmetry) $\Phi(\cdot, \cdot)$ is increasing along the ray $x = y$. This is so since

$$\Phi(x, x) = \frac{[F(x)]^2}{1 - [1 - F(x)]^2} = \frac{[F(x)]^2}{2F(x) - [F(x)]^2} = \frac{F(x)}{2 - F(x)}.$$

Then

$$\frac{d}{dx} \Phi(x, x) = \frac{f(x)}{2 - F(x)} + \frac{F(x)f(x)}{[2 - F(x)]^2} = \frac{2f(x)}{[2 - F(x)]^2} > 0.$$

$\blacksquare$

Proof of Proposition 9. The proof is exactly the same as the proof below for Proposition 10. Before summing over the first order conditions w.r.t. c^3 and c^0 (30), just divide by $\mu(A^0)$ and $1 - \mu(A^3)$ respectively. ∎

Lemma 11 *Suppose independence, Assumptions 3 and Property P. In addition, assume $\omega^i - c^i + c^0 \geq 0$ for $i = 1, 2$, and that F_i is log concave for $i = 1, 2$, and there exist two non negative numbers $\gamma^3 \geq 1$ and $\gamma^0 \leq \frac{\mu(A^0)}{1-\mu(A^0)}$ such that $\gamma^3[P(A^3)-1] + \gamma^0[P(A^0)-1] \geq 0$. Then a program where $c^3 + c^0 < c^1 + c^2$ cannot be optimal.*

Proof of Lemma 11 Under Property P the derivative of the Lagrangian with respect to c^0 and c^3 at contracts satisfying condition (21) *(i.e., negative reinforcement)* are as follows (the other two first order conditions do not change):

$$c^3: \qquad \mu(A^3)[P(A^3) - 1] - \sum_{j=1,2} \frac{\partial \mu(A^j(c))}{\partial c^3}[\omega^3 - \omega^j - (c^3 - c^j)];$$

$$\tag{30}$$

$$c^0: \qquad \mu(A^0)[P(A^0) - 1] - \sum_{j=1,2} \frac{\partial \mu(A^j(c))}{\partial c^0}[-\omega^j - (c^0 - c^j)].$$

Divide now the first line of (30) by $\gamma^3 \mu(A^3)$ and the second line of it by $\gamma^0[1 - \mu(A^0)]$ and sum them over, where $\gamma^i = \frac{1}{p^1}$ for $i = 0, 3$. Then we get the expressions

$$p^3[P(A^3)-1] + p^0 \frac{\mu(A^0)}{1-\mu(A^0)}(P(A^0)-1) - \frac{\sum_{j=1,2} \frac{\partial \mu(A^j(c))}{\partial c^3}[\omega^3 - \omega^j - (c^3 - c^j)]}{\gamma^3 \mu(A^3)} - \frac{\sum_{j=1,2} \frac{\partial \mu(A^j(c))}{\partial c^0}[-\omega^j - (c^0 - c^j)]}{\gamma^0[1 - \mu(A^0)]}.$$

By assumption the first term is non-negative, we can now concentrate on the second and third terms. If we write them explicitly and collect for the same ω's, we get

$$(\omega^2 - c^2 + c^0) \left\{ \frac{\frac{\partial \mu(A^2)}{\partial c^0}}{\gamma^0[1 - \mu(A^0)]} - \frac{\frac{\partial \mu(A^1)}{\partial c^3}}{\gamma^3 \mu(A^3)} \right\} \tag{31}$$

$$+(\omega^1 - c^1 + c^0) \left\{ \frac{\frac{\partial \mu(A^1)}{\partial c^0}}{\gamma^0[1 - \mu(A^0)]} - \frac{\frac{\partial \mu(A^2)}{\partial c^3}}{\gamma^3 \mu(A^3)} \right\} \tag{32}$$

$$+ \frac{c^3 + c^0 - c^1 - c^2 - \omega^3 + \omega^1 + \omega^2}{\gamma^3 \mu(A^3)} \left[\frac{\partial \mu(A^1)}{\partial c^3} + \frac{\partial \mu(A^2)}{\partial c^3} \right]. \tag{33}$$

The terms (31) and (32) have explicit developments given by (37) in Appendix C. Now, note that the expression to the right hand side of the first line decreases with F_1 and for $F_1 = 0$ it takes the value

42

$\frac{f_2(c^2-c^0)}{F_2(c^2-c^0)}$. Since by negative reinforcement we have $c^2 - c^0 > c^3 - c^1$ log concavity implies positivity of (31). A similar derivation implies positivity of (32). To see it note that

$$\frac{-\frac{\partial \mu(A^1)}{\partial c^0}}{1 - \mu(A^0)} = \frac{f_1(c^1 - c^0)[1 - F_2(c^2 - c^0)]}{1 - (1 - F_1(c^1 - c^0))(1 - F_2(c^2 - c^0))}$$

$$\frac{-\frac{\partial \mu(A^2)}{\partial c^3}}{\mu(A^3)} = \frac{f_1(c^3 - c^2)}{F_1(c^3 - c^2)}.$$

Again the top expression decreases with F_2 and for $F_2 = 0$ it takes the value $\frac{f_1(c^1-c^0)}{F_1(c^1-c^0)}$. Since by negative reinforcement we have $c^1 - c^0 > c^3 - c^2$ log concavity implies positivity of (32). All this implies that it cannot be that the derivative of the Lagrangian is zero at this point. We hence have shown our result for $p^0 = p^3 = 1$. It is clear that the condition on p's reinforce all our inequalities, hence the result. ∎

Proof of Proposition 2.   Consider the case where the worse off household has both its members unemployed, i.e. is in occupation 0. Under the assumption that $\omega^0 = 0$, using the identity linking the measures of sets to eliminate $\mu(A^0)$, the first order conditions reduce to

$$(\omega^1 - c^1 + c^0)\frac{\partial \mu(A^1)}{\partial c^1} + (\omega^2 - c^2 + c^0)\frac{\partial \mu(A^2)}{\partial c^1} + (\omega^3 - c^3 + c^0)\frac{\partial \mu(A^3)}{\partial c^1} = \mu(A^1),$$

$$(\omega^1 - c^1 + c^0)\frac{\partial \mu(A^1)}{\partial c^2} + (\omega^2 - c^2 + c^0)\frac{\partial \mu(A^2)}{\partial c^2} + (\omega^3 - c^3 + c^0)\frac{\partial \mu(A^3)}{\partial c^2} = \mu(A^2),$$

$$(\omega^1 - c^1 + c^0)\frac{\partial \mu(A^1)}{\partial c^3} + (\omega^2 - c^2 + c^0)\frac{\partial \mu(A^2)}{\partial c^3} + (\omega^3 - c^3 + c^0)\frac{\partial \mu(A^3)}{\partial c^3} = \mu(A^3).$$

The (3×3) matrix $\partial \mu(A^i)/\partial c^j$ is *symmetric* dominant diagonal with positive diagonal terms, negative off diagonal terms, and positive row sums. The symmetry is guaranteed by the no income effects. The above equalities imply (22). The conclusion is hence an immediate consequence of Lemma 11. ∎

Proof of Proposition 10.   Recall that the first order conditions with heterogenous ω is the same as the single ω for $i \neq 0$. We are not indicating the ω but note that if in the population there is a fraction $f(\omega)$ of people with a given ω, we could think of a set $A^0(\omega)$ with measure

$$\mu(A^0(\omega)) = f(\omega) - \sum_{i=1,2,3} \mu(A^i(\omega)),$$

where $\mu(A^i(\omega))$ are observed by us as someone works here. Of course, note importantly that with this definition of $A^0(\omega)$ we can keep the convenient equality for the the sum of the derivatives:

$$\sum_{i=0,1,2,3} \frac{\partial\mu(A^i(\omega))}{\partial c^k} = 0 \quad \forall k.$$

On the other hand, $\sum_{i=0,1,2,3} \mu(A^i(\omega))[P(A^i(\omega)) - 1]$ is typically different from one even for the case with no income effects. In addition, recall that in order to characterize the reinforcement term we can neglect one first order condition. If we impose the no-discrimination condition $t^1 = t^2$, we denote $A := A^2 \cup A^1$ and for all ω, $t^3 = \omega^1 + \omega^2 - c^3$ and $t^i = \omega^i - c^i$, and since for all i $dc^i = dt^i$, we consider a perturbation such that $dc^1 = dc^2$, the first order conditions (6) become:[24]

$$
\begin{aligned}
\mu(A)[P(A) - 1] &= \sum_{i=1,2} \frac{\partial\mu(A^0)}{\partial c^i}(t - t^0) + \sum_{i=1,2} \frac{\partial\mu(A^3)}{\partial c^i}(t - t^3) \\
\mu(A^3)[P(A^3) - 1] &= \frac{\partial\mu(A^0)}{\partial c^3}(t^3 - t^0) + \frac{\partial\mu(A)}{\partial c^3}(t^3 - t).
\end{aligned}
\tag{34}
$$

where we abused notation and used $\mu(A^1)[P(A^1) - 1] + \mu(A^2)[P(A^2) - 1] := \mu(A)[P(A) - 1]$. Recall that $t^3 - t^0 = t^3 - t + t - t^0$, hence the system can be written as a 2X2 as follows

$$
\begin{aligned}
a(t - t^0) + b(t^3 - t) &= X \\
c(t - t^0) + d(t^3 - t) &= Y.
\end{aligned}
$$

where $a = \sum_{i=1,2} \frac{\partial\mu(A^0)}{\partial c^i}$ $b = -\sum_{i=1,2} \frac{\partial\mu(A^3)}{\partial c^i}$, $c = \frac{\partial\mu(A^0)}{\partial c^3}$, $d = \frac{\partial\mu(A^0)}{\partial c^3} + \frac{\partial\mu(A)}{\partial c^3}$. The solution of this two by two system is given by the inverse - which is the matrix

$$
\begin{bmatrix}
d & -b \\
-c & a
\end{bmatrix}
$$

divided by the determinat $\Delta = ad - bc$ - multiplied by the vector $(X, Y) = (\mu(A)[P(A) - 1], \mu(A^3)[P(A^3) - 1])$; and it delivers:

$$
\begin{aligned}
t - t^0 &= \frac{1}{\Delta}[dX - bY] \\
t^3 - t &= \frac{1}{\Delta}[aY - cX],
\end{aligned}
$$

[24]Note importantly, that when there are wealth effects, the perturbation $dc^1 = dc^2$ might generate changes of people across sets A^1 and A^2. The assumption $t^1 = t^2$ makes these changes irrelevant. They will be multiplied by $t^1 - t^2 = 0$.

with $\Delta > 0$.[25] As a consequence, for each ω the reinforcement term is given by:

$$
\begin{aligned}
t - t^0 - (t^3 - t) &= \frac{(d+c)X - (b+a)Y}{\Delta} \\
&= \frac{\mu(A)[P(A)-1]\left[\frac{\partial\mu(A)}{\partial c^3} + 2\frac{\partial\mu(A^0)}{\partial c^3}\right] + \mu(A^3)[P(A^3)-1]\sum_{i=1,2}\left[\frac{\partial\mu(A^3)}{\partial c^i} - \frac{\partial\mu(A^0)}{\partial c^i}\right]}{\Delta} \\
&= \frac{\mu(A)[P(A)-1]\left[\frac{\partial\mu(A^0)}{\partial c^3} - \frac{\partial\mu(A^3)}{\partial c^3}\right] + \mu(A^3)[P(A^3)-1]\sum_{i=1,2}\left[\frac{\partial\mu(A^3)}{\partial c^i} - \frac{\partial\mu(A^0)}{\partial c^i}\right]}{\Delta}
\end{aligned}
$$

where we used the fact that from $\frac{\partial\mu(A^0)}{\partial c^3} = -\frac{\partial\mu(A^3)}{\partial c^3} - \frac{\partial\mu(A)}{\partial c^3}$ we have $\frac{\partial\mu(A)}{\partial c^3} + 2\frac{\partial\mu(A^0)}{\partial c^3} = \frac{\partial\mu(A^0)}{\partial c^3} - \frac{\partial\mu(A^3)}{\partial c^3} < 0$. If we ignore the determinant and divide the whole expression by $-\left[\frac{\partial\mu(A^0)}{\partial c^3} - \frac{\partial\mu(A^3)}{\partial c^3}\right] > 0$, we obtain the expression in the main statement. ∎

B   Special Cases: Property P and No Income Effects

We now show conditions that impose some structure to our problem. *It is a joint assumption on preferences and set of contracts.*

Property 1 (P) *Given the preferences of the individuals $(u^i, i = 1, 2, \ldots, I)$, we say that the tuple (c^0, c^1, c^2, c^3) satisfies Property P if for all α the following conditions hold:*

- $u^3(c^3, \alpha) \geq \max\{u^1(c^1, \alpha), u^2(c^2, \alpha)\} \Rightarrow u^3(c^3, \alpha) \geq u^0(c^1 + c^2 - c^3, \alpha)$

- $u^0(c^0, \alpha) \geq \max\{u^1(c^1, \alpha), u^2(c^2, \alpha)\} \Rightarrow u^0(c^0, \alpha) \geq u^3(c^1 + c^2 - c^0, \alpha)$.

Lemma 12 *Whenever Property P holds, negative reinforcement (i.e, condition (21)) implies*

$$
\frac{\partial\mu(A^3(c))}{\partial c^0} = \frac{\partial\mu(A^0(c))}{\partial c^3} = 0.
$$

Immediate from Property P. Just note that the statement is true for $c^3 + c^0 = c^1 + c^2$ so - from monotonicity - it holds *a fortiori* when condition (21) holds. ∎

---

[25]Indeed we have: $\Delta = ad - bc = -\frac{\partial\mu(A^3)}{\partial c^3} \times \frac{\partial\mu(A^0)}{\partial c} + \frac{\partial\mu(A^3)}{\partial c} \times \frac{\partial\mu(A^0)}{\partial c^3} > 0$. This is so since the only positive number among the four is $\frac{\partial\mu(A^3)}{\partial c^3}$.

B.1 Separable cost of work.

We now present an environment where Property P holds *for all tuples* (c^0, c^1, c^2, c^3) *that satisfy condition (21), i.e., negative reinforcement.* Suppose, there are two increasing functions v and u, with u weakly concave, such that for all i

$$u^i(c, \alpha) = v(u(c) - \alpha^i), \tag{35}$$

where α^i is the i-th entry of the vector α. This model includes in particular the pecuniary cost model (for u linear) we discussed above but allows for wealth effects. Suppose moreover that for all households in the economy:

Assumption 7

$$\alpha^3 + \alpha^0 \geq \alpha^1 + \alpha^2, \quad with \quad \alpha^3 \geq \alpha^1, \alpha^2 \geq \alpha^0. \tag{36}$$

This assumption recalls the increasing difference assumption and covers, for example, the case where $\alpha^0 = 0$ and the cost of going to work is independent of the activity of the other member, so that $\alpha^3 = \alpha^1 + \alpha^2$. More generally, it states that the cost for both members of the couple to work on the market α^3 is at least as large as the sum of the cost for one to go to work, say α^1, plus the net cost supported by the other to go on the market α^2 after deducting the home production α^0 gathered when both are unemployed. Note that for each realization of (α^1, α^2) both α^3 and α^0 can still be random.

Proposition 13 *Assume the utilities u^i can be represented by (35) and that Assumption 7 holds. Then when payments satisfy negative reinforcement they satisfy Property P and*

$$\frac{\partial \mu(A^3(c))}{\partial c^0} = \frac{\partial \mu(A^0(c))}{\partial c^3} = 0.$$

Notice that since v is strictly monotone and the same function for all states, we can only work with its argument $u(c^i) - \alpha^i$. We will show that Property P holds whenever payments satisfy negative reinforcement, hence the second part of the result will follow from Lemma 12.

We start with the first implication of Property P, i.e.

$$u(c^3) - \alpha^3 \geq \max\{u(c^1) - \alpha^1, u(c^2) - \alpha^2\} \Rightarrow u(c^3) - \alpha^3 \geq u(c^1 + c^2 - c^3) - \alpha^0.$$

We first show that the left hand inequality implies $c^3 \geq \max\{c^1, c^2\}$. Indeed it implies that

$$u(c^3) \geq u(c^1) + \alpha^3 - \alpha^1 \geq u(c^1),$$

46

where we used $\alpha^3 - \alpha^1 \geq 0$ from Assumption 7. By monotonicity of u we have $c^3 \geq c^1$. Similarly, one gets $c^3 \geq c^2$.

We now prove the first implication defining Property P. The first inequality implies that $2u(c^3) - 2\alpha^3 \geq u(c^1) + u(c^2) - (\alpha^1 + \alpha^2)$. If we use $\alpha^3 + \alpha^0 \geq \alpha^1 + \alpha^2$ we get $u(c^3) - \alpha^3 \geq u(c^1) + u(c^2) - u(c^3) - \alpha^0$. To show the final inequality, it hence suffices to show that $u(c^1) + u(c^2) - u(c^3) \geq u(c^1 + c^2 - c^3)$ or bringing $u(c^3)$ on the other side and dividing by $\frac{1}{2}$ both sides we have: $\frac{1}{2}u(c^1) + \frac{1}{2}u(c^2) \geq \frac{1}{2}u(c^1 + c^2 - c^3) + \frac{1}{2}u(c^3)$.

Now notice that since $c^3 \geq \max\{c^1, c^2\}$ the payments in the right hand side constitute a mean preserving spread of those of the left hand side, where the lotteries have equal probability. Hence the inequality is true by concavity of u.

The proof of the other implication of Property P must be divided in two steps. Consider first the case where for all α we have $u(c^3) - \alpha^3 < \max\{u(c^1) - \alpha^1, u(c^2) - \alpha^2)\}$. In this case, the desired implication follows trivially: $u(c^0) - \alpha^0 \geq \max\{u(c^1) - \alpha^1), u(c^2) - \alpha^2)\} > u(c^3) - \alpha^3$.

Finally, consider the case where for some $\underline{\alpha}$ we have $u^3(c^3, \underline{\alpha}) \geq \max\{u^1(c^1, \underline{\alpha}), u^2(c^2, \underline{\alpha})\}$. This is the top inequality of Property (P) holding for $\underline{\alpha}$. As we saw above, this implies that $c^3 \geq \max\{c^1, c^2\}$. As a consequence, under negative reinforcement, we must have $c^0 \leq c^1 + c^2 - c^3 \leq \min\{c^1, c^2\}$. The proof is now essentially identical to that done above. Just note that in this case, the lottery: $c^1 + c^2 - c^0$ with probability $\frac{1}{2}$ and c^0 with probability, $\frac{1}{2}$ is a mean preserving spread of the lottery c^1 or c^2 with equal probability. ■

Remark 6 *It is easy to show that when u is linear, i.e., when there are no wealth effects, we can dispense of the second part of Assumption 7 and only require $\alpha^3 + \alpha^0 \geq \alpha^1 + \alpha^2$.*

Figure 5 provides a graphical representation of the sets A^i in the pecuniary cost case where $\alpha^0 + \alpha^3 = \alpha^1 + \alpha^2$. The plan describes the values taken by the parameters α. The horizontal axis bears $\alpha^1 - \alpha^0$, equal to $\alpha^3 - \alpha^2$, while $\alpha^2 - \alpha^0$, equal to $\alpha^3 - \alpha^1$, is on the vertical axis.

The right panel of Figure 5 is a graphical representation of Lemma 12: Since the sets A^3 and A^0 have no borders in common, changing marginally c^0 has no effect on A^3 and, symmetrically, small changes in c^3 have no effect on A^0.
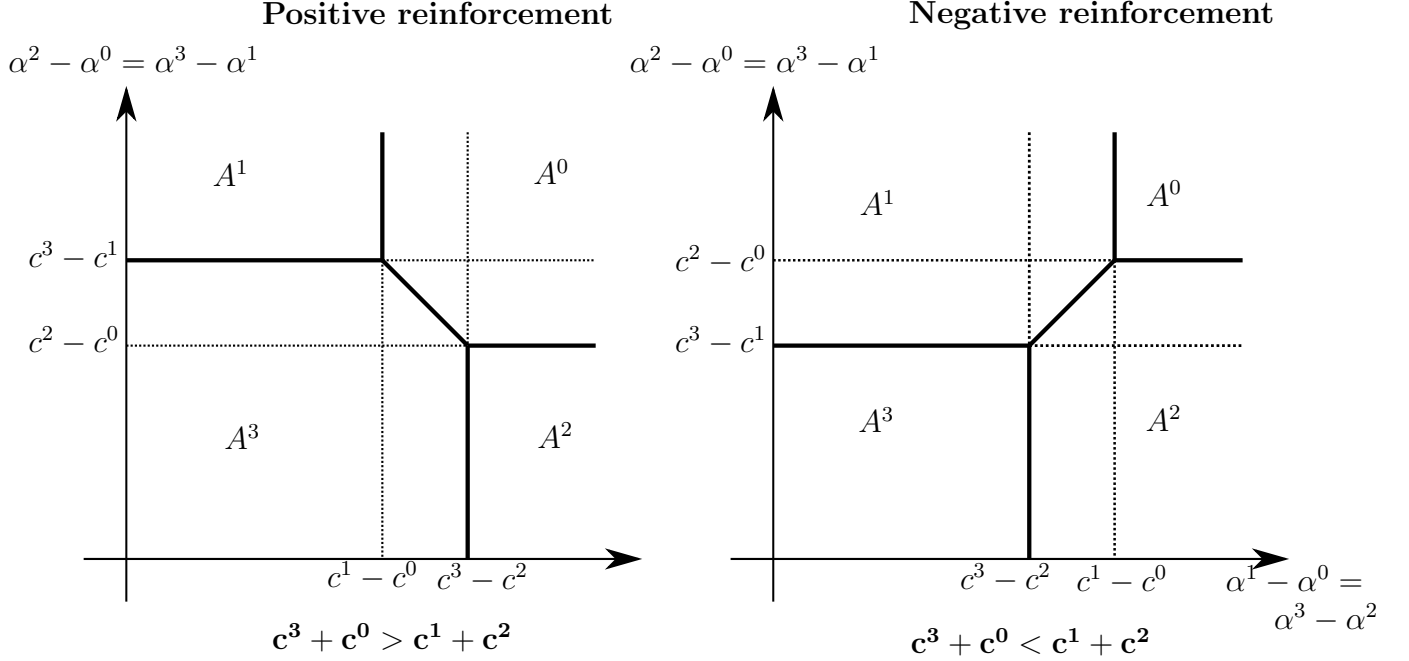
Figure 5: *The right figure represents the case with positive reinforcement (subsidizing working couples) while the left one represents the case with negative reinforcement (taxing working couples).*

C   The Pecuniary case with Independent Distribution: Computations

Consider the original system where to get the reduced matrix H instead of eliminating the row corresponding to the first order conditions with respect to c^0 we eliminate the row corresponding to the first order conditions with respect to c^2. When the α's are independently distributed we have:

$$\frac{-\frac{\partial \mu(A^2)}{\partial c^0}}{1 - \mu(A^0)} = \frac{f_2(c^2 - c^0)[1 - F_1(c^1 - c^0)]}{1 - (1 - F_1(c^1 - c^0))(1 - F_2(c^2 - c^0))} \tag{37}$$

$$\frac{-\frac{\partial \mu(A^1)}{\partial c^3}}{\mu(A^3)} = \frac{f_2(c^3 - c^1)}{F_2(c^3 - c^1)}. \tag{38}$$

This follows directly from the proof of Lemma 12 where we saw that Assumption 3 and negative reinforcement allow to simplify the expression describing the sets A^0 and A^3 into

$$A^0 = \{\alpha | c^0 - \alpha^0 > c^2 - \alpha^2, c^0 - \alpha^0 > c^1 - \alpha^1\},$$

$$A^3 = \{\alpha | c^3 - \alpha^3 > c^2 - \alpha^2, c^3 - \alpha^3 > c^1 - \alpha^1\}.$$

Such a simplification does not occur in the expressions of the sets A^1 and A^2, which involve three inequalities. However their 'differential' take a simple form. Indeed for A^2:

$$\frac{\partial A^2}{\partial c_0} := \left\{ \alpha | c^2 - \alpha^2 \geq c^1 - \alpha^1; c^2 - \alpha^2 \geq c^3 - \alpha^3; c^2 - c^0 \leq \alpha^2 - \alpha^0 \leq c^2 - c^0 + dc^0 \right\}.$$

We can ignore the middle inequality, since by negative reinforcement, $c^3 \leq c^1 + c^2 - c^0$, so that

$$c^3 - \alpha^3 \leq c^1 + c^2 - c^0 - \alpha^1 - \alpha^2 + \alpha^0 \leq c^1 - \alpha^1 + dc^0.$$

Also getting $\alpha^2 - c^2$ from the last term on the right hand side and substituting in the first

$$\frac{\partial A^2}{\partial c_0} = \left\{ \alpha | c^1 - c^0 \leq \alpha^1 - \alpha^0; c^2 - c^0 \leq \alpha^2 - \alpha^0 \leq c^2 - c^0 + dc^0 \right\}.$$

A similar derivation can be made for the differential of A^1 with respect to c^3.

Suppose again that property P is satisfied. Recall that if we eliminate t^2 and its related row, the system of first order conditions become simpler. *Under negative reinforcement* the (reduced) matrix of the first order conditions becomes:

$$H = \begin{bmatrix} a_0 & -b & 0 \\ -b & a_1 & -d \\ 0 & -d & a_3 \end{bmatrix},$$

where

$$\begin{aligned}
a_0 &= f_1(c^1 - c^0)(1 - F_2(c^2 - c^0)) + f_2(c^2 - c^0)(1 - F_1(c^1 - c^0)); \\
b &= f_1(c^1 - c^0)(1 - F_2(c^2 - c^0)); \\
a_1 &= \frac{\partial \mu(A^1)}{\partial c^1} = f_2(c^3 - c^2)F_1(c^3 - c^1) + \int \cdots + f_1(c^2 - c^0)[1 - F_2(c^1 - c^0), \\
d &= F_1(c^3 - c^2)f_2(c^3 - c^1), \\
a_3 &= f_1(c^3 - c^2)F_2(c^3 - c^1) + F_1(c^3 - c^2)f_2(c^3 - c^1).
\end{aligned}$$

and hence

$$H^{-1} = \frac{1}{\Delta} \begin{bmatrix} a_1 a_3 - d^2 & a_3 b & bd \\ a_3 b & a_0 a_3 & a_0 d \\ bd & a_0 d & a_1 a_0 - b^2 \end{bmatrix}.$$

So, the reinforcement term equals

$$R = \frac{-1}{\Delta} \left\{ \left[d^2 - a_1 a_3 + a_3 b - bd \right] X_0 + \left[-a_3 b + a_0 a_3 - a_0 d \right] X_1 + \left[-bd + a_0 d + b^2 - a_0 a_1 \right] X_3 \right\}$$

where of course X_2 does not appear since it is residual as the sum of Pareto weights equals one in the pecuniary cost model. We also have:

$$\Delta = a_0 a_1 a_3 - a_0 d^2 - a_3 b^2;$$

$$d^2 - a_1 a_3 + a_3 b - bd = ...$$

$$\frac{-a_3 b + a_0 a_3 - a_0 d}{A} = \frac{f_1(c^3 - c^2) f_2(c^2 - c^0)}{F_1(c^3 - c^2)(1 - F_2(c^2 - c^0))} - \frac{f_1(c^1 - c^0) f_2(c^3 - c^1)}{F_2(c^3 - c^1)(1 - F_1(c^1 - c^0))};$$

$$-bd + a_0 d + b^2 - a_0 a_1 = ...$$

where $A = F_2(c^3 - c^1)(1 - F_1(c^1 - c^0)) F_1(c^3 - c^2)(1 - F_2(c^2 - c^0)) > 0$.

D   Rawlsian Planner

We shall also study the behaviour of a Rawlsian government that aims at maximizing the welfare of the worst off person in the economy. For concreteness, suppose that the distribution of α contains a point $\overline{\alpha}$ such that for all agents α, $\alpha^i \leq \overline{\alpha}^i$ for all i. We claim that the agent $\overline{\alpha}$ is the worst off person in the economy. Indeed suppose that s/he chooses decision j. Then s/he is worse off that all those that picked up j since with a cost $\overline{\alpha}^j$ larger than their α^j. But s/he is also worse off than the others, since by revealed preference, the choice i that they make brings them a net income at least as good as selecting j.

The Rawlsian government selects the tax schedule c that maximizes this quantity subject to the budget constraint.[26]

We use the transformation of the Rawlsian criterion suggested in footnote 26.

---

[26]In the situation where the worse off agents are indifferent between several decisions, the Rawlsian criterion defined above may be difficult to manipulate. A possible transformation of the program consists in introducing an element ρ in the simplex of $\mathcal{R}^{I+1}$, to give implicit weights to the various choices. The government then chooses both c and ρ so as to maximize

$$\sum_{i=1}^{I} \rho^i (c^i - \overline{\alpha}^i),$$

subject to the budget constraint with ρ in the simplex. The two programs are equivalent. Indeed the welfare obtained in the initial program can be achieved in the second one, by concentrating the weights on the optimal choices of the household. Conversely a solution of the second program is such that $\rho^i > 0$ implies $c^i - \overline{\alpha}^i \geq c^j - \overline{\alpha}^j$ for all j and therefore gives the same value as the first program.

$$\begin{cases} \max \sum_{i=0}^{I} \rho^i (c^i - \overline{\alpha}^i) \mathrm{d}F(\alpha) \\ \sum_{i=0}^{I} \rho^i = 1, \\ \rho^i \geq 0 \text{ for all } i, \\ \sum_{i=0}^{I} \left[\omega^i - c^i \right] \mu(A^i(c)) = G. \end{cases} \qquad (39)$$

A marginal equal change in all the c^i's shows that the Lagrange multiplier λ of the budget constraint is equal to 1. The first order condition with respect to c^i is

$$\frac{\partial \mathcal{L}}{\partial c^i} = \rho^i - \mu(A^i(c)) + \sum_{j=0}^{I} (\omega^j - c^j) \frac{\partial \mu(A^j(c))}{\partial c^i} = 0.$$

This implies that for all i

$$1 - \mu(A^i(c)) + \sum_{j=0}^{I} (\omega^j - c^j) \frac{\partial \mu(A^j(c))}{\partial c^i} \geq 0.$$

For the set of indices K such that the Rawlsian agent attains her optimum in occupation k, ρ^k is positive so that

$$-\mu(A^k(c)) + \sum_{j=0}^{I} (\omega^j - c^j) \frac{\partial \mu(A^j(c))}{\partial c^k} < 0.$$

On the complement of K, for all i in $I - K$, the first order conditions reduce to

$$-\mu(A^i(c)) + \sum_{j=0}^{I} (\omega^j - c^j) \frac{\partial \mu(A^j(c))}{\partial c^i} = 0.$$

Counting unknowns and equations, we have $I + 1$ unknowns, the consumptions c^i, $i = 0, 1, \ldots, I$, and $I + 1$ equations: the above $I + 1 - K$ first order conditions, the $K - 1$ equalities of the $c^k - \overline{\alpha}^k$ for k in K and the overall budget constraint.