FUNDAÇÃO GETULIO VARGAS
ESCOLA DE PÓS-GRADUAÇÃO EM ECONOMIA

Cassiano Breno Machado Alves

Essays on Taxation and Regulation

Variational Approach, Couples Taxation, and Dynamic Procurement

Rio de Janeiro
2017

Cassiano Breno Machado Alves

Essays on Taxation and Regulation

Variational Approach, Couples Taxation, and Dynamic Procurement

Tese apresentada ao Curso de Doutorado em Economia da Escola de Pós Graduação em Economia como requesito parcial para a obtenção do título de Doutor em Economia.

Área de Concentração: Teoria Econômica, Finanças Públicas

Orientador: Humberto Luiz Ataíde Moreira
Co-orientador: Carlos Eugênio E. L. da Costa

Rio de Janeiro
2017

Ficha catalográfica elaborada pela Biblioteca Mario Henrique Simonsen/FGV.

Alves, Cassiano Breno Machado

Essays on taxation and regulation: variational approach, couples taxation, and dynamic procurement / Cassiano Breno Machado Alves. - 2017.

106f.

Tese (Doutorado) - Fundação Getulio Vargas, Escola de Pós-Graduação em Economia.

Orientador: Humberto Luiz de Ataíde Moreira.

Coorientador: Carlos Eugênio Ellery Lustosa da Costa.

Inclui Bibliografia.

1. Impostos. 2. Política Tributária. 3. Orçamento Familiar. I. Moreira, Humberto Luiz Ataíde. II. Costa, Carlos Eugênio da. III. Fundação Getulio Vargas. Escola de Pós-Graduação em Economia. IV. Título.

CDD - 336.2

CASSIANO BRENO MACHADO ALVES

"ESSAYS ON TAXATION AND REGULATION: VARIATIONAL APPROACH, COUPLES TAXATION, AND DYNAMIC PROCUREMENT"

Tese apresentada ao Curso de Doutorado em Economia da Escola de Pós-Graduação em Economia para obtenção do grau de Doutora em Economia.

Data da defesa: 20/09/2017

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Humberto Moreira

Humberto Luiz Ataíde Moreira
Orientador (a)

Carlos Eugenio Ellery Lustosa da Costa

Felipe S. Iach

Felipe Saraiva Iachan

Braz Ministerio de Camargo

Sergei Vieira

Sergei Vieira Silva

Resumo

Esta tese coleciona 3 dos projetos por mim desenvolvidos durante meu período de doutoramento na Escola de Pós Graduação em Economia da Fundação Getulio Vargas. A exposição será feita em ordem contrária a confecção dos mesmos, posto que com a maior maturidade e domínio dos temas abordados as contribuições são mais substanciais, na leitura deste autor.

No primeiro capítulo apresento um artigo que nasce da percepção que métodos variacionais levam à formulas para a estrutura tributária ótima que podem não ser válidas em ambientes mais complexos. Este trabalho é co-autorado com meus orientadores Carlos E. da Costa e Humberto Moreira. Para ilustrar tal ponto desenvolvemos um modelo no qual a falta de ordenação da característica não observada pelo governo gera uma não validade das condições inerentes à aplicação de métodos variacionais.

No segundo capítulo apresento o estudo do desenho ótimo da estrutura tributária em uma economia formada por casais levando em conta o processo decisório dentro do domicílio. Este artigo também co-autorado com meus orientadores e com Felipe Lobel aluno deste mesmo departamento. Neste artigo mostramos o impacto na estrutura tributária ótima quando o planejador social toma a utilidade de cada indivíduo, ao invés da utilidade aggregada do domicílio, como unidade básica na formulação do seu critério de bem-estar.

A terceira parte aborda o problema de regulação em um ambiente dinâmico e discute como o fenômeno conhecido como o *ratchet effet* é afetado ao se permitir tipos randômicos em uma relação entre um regulador (principal) e firma licitante (agent) na qual o primeiro não pode se comprometer a contratos de longo prazo. Neste caso toda informação revelada influencia os novos termos desta relação.

KEYWORDS: Optimal Taxation, Couples Taxation, Variational Approach and Dynamic Regulation.

Abstract of this Thesis

This thesis contains 3 articles developed as a partial requirement for the degree of Doctor in Economics at Escola de Pós Graduação em Economia from Getulio Vargas Foundation.
In the first chapter, I present a paper disscuss situations where the variational method fails to identify the Optmal tax system. This paper is co-authored with my advisors Carlos E. da Costa e Humberto Moreira. In the second chapter, we study the feature of an optimal tax system when we take the family structure in account. This article is also co-authored with my advisors and Felipe Lobel a student in this same department. In the last chapter, we study the problem of regulating a firm in a dinamic environment and we study how the ratchet effect changes when the type of agent is a random variable.

KEYWORDS: Optimal Taxation, Couples Taxation, Variational Approach and Dynamic Regulation.

List of Figures

List of Tables

Contents

Chapter 1

On the Variational Approach to the Analysis of Tax Systems: A Cautionary Note

Abstract

We study optimal income taxation in a class of static, one-dimensional Mirleesian economies where preferences do not satisfy the Spence-Mirrlees condition (SMC). We characterize necessary conditions for the optimal taxation using a structural mechanism design approach based on type assignment functions. Because the SMC is violated, local incentive constraints no longer suffice for implementability, and an additional set of global incentive constraints must be explicitly taken into account. When these global constraints bind, they create a tension between infra-marginal types whose ICs are not handled by any local approach. Local perturbations (or small reforms) of the optimal tax schedule may have global (first-order) impacts on welfare, thus invalidating some of the assumptions underlying local variational methods.

1.1 Introduction

Following Piketty's (1997) and Saez's (2001) contributions, optimal income tax theory has geared towards the use of variational methods as its main approach in the analysis of tax systems, as an alternative to Mirrlees' (1971b) structural mechanism design approach.[1] These techniques are based on studying local perturbations (or small reforms)

[1]In a similar vain, the so-called sufficient statistics approach proposed by Chetty (2009) uses variational techniques to study the welfare impact of policy changes in applications other than the design of income tax systems. The sufficient statistics approach applied to the design of optimal income taxes is, in fact, the main contribution of Piketty (1997) and Saez (2001).

of a baseline nonlinear tax system under some regularity assumptions regarding the optimum. They have the significant advantage of producing formulae based on empirically relevant objects, usually elasticities, that are very transparent about the economic forces driving the size and shape of the optimal tax schedule. Since these elasticities are often observable and/or may be credibly estimated with available real-world data using quasi-experimental designs, the variational approach brought optimal tax theory closer to actual policy-making.

Another perceived advantage is that it does not require the strong restrictions needed to characterize optimal taxes using a fully structural mechanism design approach as in Mirrlees' (1971b) original work. In particular, the Spence-Mirrless condition (SMC) for agents' utility need not be assumed.[2] Hence, the method would be applicable in complex environments where imposing these restrictions seems too strong. All of these gains come at the expense of making technical regularity assumptions on endogenous objects. Golosov, Tsyvinski, and Werquin (2014) provide a systematic analysis of the underlying assumptions necessary for applying the method in the design of optimal tax systems. Intuitively, local perturbations around the optimal tax system should not have any impact on welfare. The required sufficient assumption is the Lipschitz continuity of the agents' decisions with respect to the tax policy, which is an endogenous object. Hence it is not possible to check its validity in any given application.

The goal of this paper is twofold. First, we present a class of simple Mirrleesian economies that dispense with the SMC and remain tractable enough for the Mirrlees' approach (1971b), and we investigate whether the regularity assumptions in Golosov, Tsyvinski, and Werquin (2014) are verified at the optimum tax scheme. Second, we propose a different class of solutions for the optimal design of income tax systems featuring segregation of income groups. These solutions emerge naturally when tackling the mechanism design problem using type assignment functions.[3]

For the class of economies we study, the SMC is replaced by the assumption that the marginal disutility of producing taxable income is increasing in one region and decreasing in another of the parameter space, where the boundary between the regions is described by a monotonic function of income earned. This assumption, first adopted in a monopoly pricing context by Araújo and Moreira (2010), implies that local incentive compatibility constraints are not sufficient to ensure full incentive compatibility. Typically, efficient allocations are not monotonous generating discrete pooling, i.e. the sets of agents for which the same bundles are assigned need not be connected. Building on this work, Araujo,

[2]The SMC is often referred to as the single-crossing property. This condition guarantees that agents' marginal utility is ordered with respect to their idiosyncratic characteristic or type.

[3]The Mechanism Design Approach typically works in the space of direct mechanism that for each type associates an allocation. Our method uses a type assignment functions which inverts the logic: for each possible bundle a set of types is assigned.

Moreira, and Vieira (2015) exhibit a welfare improvement by relaxing the requirement of convex-valued mechanisms.

Our proposed tax schedule promotes segregation of income groups, which allows the government to more efficiently balance the distortions within each group in an environment where the tax system has to take into account global incentive compatibility constraints. Intuitively, the government handles these constraints by making sure that a pivotal agent is indifferent between migrating from the high to the low-income group.

We assess the properties of schedules that promote segregation of income groups and show that, in particular, they do not satisfy the Lipschitz continuity assumption described in Golosov, Tsyvinski, and Werquin (2014). Hence, we cannot make use of the Gateaux differential to identify the optimum. In fact, when global incentive constraints are binding, local perturbations of the tax system at the optimum have infra-marginal welfare impacts that are not negligible. By distorting the taxation on the pivotal type, global incentive compatibility requires distorting the allocation of a positive measure set of individuals. Thus, the impact on welfare is proportional to the shadow cost of this global incentive constraint measured by the associated Lagrangian multiplier. As a consequence of these complexities, the use of variational methods may fail in providing an optimal income tax system. Moreover, elasticities emerging from the use of variational methods may not even be sufficient statistics to the welfare analysis.[4]

It is important to stress that the violation of SMC is a necessary but not sufficient condition to the failure of variational methods in characterizing optimal tax systems. It is crucial to have a strong tension between intra-marginal benefits and infra-marginal incentives generated by binding global incentive compatibility constraints.

One of the features of the Variational Approach is to be agnostic about the underlying information frictions in the economic environment. However, unless we fully specify the nature of the information friction, we have no hope of assuring the validity of the Variational Approach in identifying the optimum. Apart from its technical contributions, this paper should be thought of as a cautionary tale regarding the usage of variational methods in complex environments where the failure of SMC and the associated impossibility of ordering marginal utilities according to types naturally emerges.[5]

The rest of the paper is organized as follows. After this introduction, we briefly

[4] Piketty, Saez, and Stantcheva (2014), Hendren (2013) and Scheuer and Werning (2015) have examples of complex environments where the elasticities that should be considered are not the usual elasticity of taxable income.

[5] The possibility of binding global incentive constraints was recognized a long time ago by **?** (**?**) in the Moral Hazard context. Unlike the screening enviroment of Mirrlees (1971b), in the Moral Hazard context he couldn't find a natural assumption in terms of fundamentals that guarantees the sufficiency of the First Order Approach.

discuss the related literature. In Section 1.2, we present a class of economies where the agent's utility does not satisfy the Spence-Mirrlees condition. In Section 3, we discuss the complications emerging in this environment and characterize the relevant incentive constraints. In Section 4, we propose a novel approach to the optimal income taxation based on type assignment function and compare with the one emerging from the variational approach. The last section is reserved for the conclusion and discussion of future steps.

Literature review

This paper lies in the intersection of two different branches of the literature: (1) adverse selection models without the SMC and (2) optimal taxation using variational techniques.

Although perturbation methods have been in use since at least Sheshinski (1972), Piketty (1997); Dahlby (1998); Saez (2001) have shown how to extend them from the parametric restrictions on schedules that were imposed in the early literature. These methods proved to be a significant generalization since they allowed us to assess optimal allocations, thus placing them on the same footing as what was accomplished by Mirrlees (1971b).[6]

These methods have expanded the scope of optimal tax theory to address migration (as in Lehmann, Simula, and Trannoy (2014)), dynamics (as in Golosov, Tsyvinski, and Werquin (2014)), general equilibrium effects (as in Sachs, Tsyvinski, and Werquin (2016)), etc. among others. They have also brought theory closer to applications since tax formulae are usually expressed in elasticities which can be recovered from the data.

On a related literature, the sufficient statistic approach uses variational methods to identify formulas for the welfare impact of policy reforms based on high-level elasticities. For instance, Feldstein (1999) first pointed out that the elasticity of taxable income is a sufficient statistic to welfare analysis. This approach has since been extended by Chetty (2009), Piketty, Saez, and Stantcheva (2014), among others as a middle-ground between structural and reduced form methods. For instance, Saez's (2001) uses variational methods to re-write the Mirrlees' (1971b) formula for the optimal income tax rate in terms of labor supply elasticities.

Methodologically, our paper is built upon Araujo, Moreira, and Vieira (2015) and closer to the literature of screening problems without the SMC. Notable examples are Araújo and Moreira (2010) and SchottmÃČÅŠller (2015). The literature of models that do

[6]Recall that under the Taxation Principle – Hammond (1979, 1987) – any constrained efficient allocation can be implemented through a suitable design of budget sets.

not satisfy SMC is tightly related to the multidimensional screening literature (see Rochet and Choné (1998), Rochet and Stole (2003) and Armstrong and Rochet (1999)). For instance, the failure of the SMC could also emerge naturally due to multidimensionality; for example, in the problem of taxing couples (see Alves, da Costa, and Moreira (2017a)).

The use of type assignment functions as an alternative approach to mechanism design dates back to Goldman, Leland, and Sibley (1984) and Noldeke and Samuelson (2007). To the best of our knowledge, we are the first to use this method in the optimal taxation context.

1.2 Environment

The economy is inhabited by a continuum of agents with measure equal to one. Each agent is parametrized by a type $\theta \in \Theta \subset \mathbb{R}_+$ and earns a taxable income $z \in Z \subset \mathbb{R}_+$. Agents spend their income in a consumption good $c \in \mathbb{R}_+$ incurring in utility given by the quasi-linear function

$$V(c, z, \theta) = c - v(z, \theta), \tag{1.1}$$

common to all agents.[7] The assumption that all individuals share the same utility function differing from each other only due to the realization of the type θ implies that all redistribution should be founded in horizontal equity motives where all individuals with the same type are treated equally.

Function $v(z, \theta)$ is an important piece of our analysis. It represents the utility cost incurred by a type-θ agent to earn z units of income. We assume $v(z, \theta)$ to be increasing on earnings, convex and three times continuously differentiable on $Z \times \Theta$. The utility function considered here has the advantage of being simple to handle and flexible enough to accommodate several utility functions used in the public economics literature papers. The interpretation of parameter θ depends on the application under consideration. For instance, it can be the labor market productivity as in the original Mirrlees's (1971b) economy, a labor taste parameter, a discount factor in dynamic models or the amalgam of several variables from a multidimensional model. In Appendix A we present some common environments in the public economics literature and the implied function $v(z, \theta)$ as well as the interpretation for θ.

In autarky, individuals consume all their income and optimally choose the taxable

[7]Since this the only good in the economy we can normalize its price to one, or equivalently, that income is measure in units of the consumption good.

income in order to satisfy the following first-order necessary condition:[8]

$$1 - v_z(z^A(\theta), \theta) = 0. \tag{1.2}$$

Two properties of the autarky allocation are worth noting. First, all individuals get the same marginal disutility of taxable income supply. They generate income until the marginal cost in terms of the utility of making the extra unit is equal to 1, which is the marginal benefit. Second, when $v_{z\theta} > 0$, higher types get lower incomes in autarky, and when $v_{z\theta} < 0$, higher types get higher incomes.

The government sets a non-linear tax policy $T : Z \to \mathbb{R}$ assigning a tax liability/subsidy for each possible income level to achieve its redistributive goals. To make our model tractable, we assume that the type space is a connected interval, $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$. The cross-section distribution of types θ is common-knowledge and denoted by F with absolutely continuous and strictly positive density $f : \Theta \to \mathbb{R}_{++}$. Type θ is a private information of agents and the government has to design incentive schemes to reveal this private information.

The revelation principle guarantees that the set of allocations that can be reached given the information structure and the resource constraint are payoff equivalent to the set of allocations implemented by direct revelation mechanisms. In this taxation context a direct mechanism is a pair of functions $z, c : \Theta \to Z \times \mathbb{R}_+$ specifying an allocation consisting in a income-consumption pair for each type θ revealed. The design of such mechanisms has to induce truthful revelation of types. This is considered by imposing incentive compatibility constraints formally defined below. We use interim Bayesian Nash equilibrium as our implementability concept.

Definition 1 *An allocation $z, c : \Theta \to Z \times \mathbb{R}_+$ is said to be* ***incentive-feasible*** *when: (i) the incentive compatibility constraints:*

$$c(\theta) - v(z(\theta), \theta) \geq c(\hat{\theta}) - v(z(\hat{\theta}), \theta), \tag{1.3}$$

for all $\hat{\theta}, \theta \in \Theta$, and; (ii) the government's budget constraint:

$$\int_{\Theta} [z(\theta) - c(\theta)] f(\theta) d\theta \geq 0; \tag{1.4}$$

are satisfied.

[8] Note that the second-order sufficient condition is satisfied since we assumed function $v(z, \theta)$ to be convex in z.

1.2.1 The tax assignment function and the taxation principle

Under the mechanism design approach one chooses in the space of direct mechanisms the one which implements the allocation leading to maximum welfare. The variational approach, on the other hand, works directly on the space of tax schedules, often called indirect mechanisms. One of the goals of this paper is to clarify how important phenomena in the space of direct mechanisms translate to tax schedules and vice-versa.

Incentive compatibility in the mechanism design approach translates into behavioral responses, usually measured by elasticities, in the space of indirect mechanisms. More specifically, when calculating the welfare impacts of a tax reform, the tax designer anticipates the changes in taxable income supply and internalizes it to get the best tax schedule. The non-linear income tax schedule $T : Z \to \mathbb{R}$ comprises all taxation that depends on the primary income, not only personal income tax. The net-of-tax income $c = z - T(z)$ is consumed and so the taxation implied by the direct mechanism is $T(z) = z - c$. The taxation principle provides the tool to connect direct and indirect mechanisms. It states that any incentive-feasible allocation can also be implemented via an income tax system and conversely any allocation implied by a tax system is incentive-feasible. This result is a powerful tool in the challenging task of formally characterizing the connection between incentives and behavioral responses.

Proposition 1 (Taxation Principle) *An allocation $z, c : \Theta \to Z \times \mathbb{R}_+$ is incentive-feasible if and only if it is tax-implementable. In other words, an allocation is incentive-feasible if and only if there exists a tax schedule $T : Z \to \mathbb{R}$ such that for all $\theta \in \Theta$:*

(i) $T(z(\theta)) = z(\theta) - c(\theta)$;

(ii) $z(\theta) \in \arg\max_{z \in Z(\Theta)} V(z - T(z), z, \theta)$, *and;*

(iii) the government's budget constraint holds: $\int_{\theta \in \Theta} T(z(\theta)) f(\theta) d\theta \geq 0$.

Proof: See the Appendix. ■

Now, instead of prescribing an allocation for the type announcements, the government designs a menu of choices, or budget sets, and lets individuals self-select their income. As we saw in Proposition 4, the allocation generated by this tax system is incentive-feasible.

When facing a differentiable tax schedule $T : Z \to \mathbb{R}$,[9] the type-θ agent's problem is to choose an income level $z \in Z$ to maximize after-tax utility $V(z - T(z), z, \theta)$. Explicitly,

[9]Non-differentiability creates kinks on the tax schedule where bunching may occur. Bunching is an important feature of the optimal tax that will explore later.

the problem is

$$\max_{z\in Z}\ z - T(z) - v(z,\theta). \tag{1.5}$$

The optimal choice of production z satisfies the first-order condition:

$$T'(z) = 1 - v_z(z,\theta). \tag{1.6}$$

This first-order condition implicitly defines a Marshallian earnings (taxable income) supply function,[10] which shall be denoted by $z_\theta(T)$.[11] Comparing this equation with the autarky solution we see that the tax liability creates a wedge between the marginal cost of an extra unit of income and its benefit, which is drecrease by the marginal tax rate. This wedge creates a distortion on the choice of z that is related to the marginal rate at the individual optimal level.

The elasticity of taxable income (ETI) with respect to the marginal tax rate , which is an important measure of the behavioral responses, is given by[12]

$$\zeta_\theta(z) \equiv \frac{dz_\theta}{d(1-T')}\frac{(1-T')}{z_\theta} = \frac{v_z(z,\theta)}{v_{zz}(z,\theta)z}. \tag{1.7}$$

One important technical point needs to be made here. The sufficiency of this first-order necessary condition as well as this formula for the elasticity heavily relies on the tax schedule being well-behaved. In fact, a sufficient condition widely used in the literature is the convexity of the tax system. In this case the objective function is concave and the second-order condition is satisfied because

$$-T''(z) - v_{zz}(z,\theta) \leq 0. \tag{1.8}$$

This assumption is too strong, however: restricting the tax schedule to be convex will lead to welfare losses, in general.[13]

The space of indirect mechanisms is the space of functionals defined over Z, and we have to use the appropriate objects for this space. In particular, in the variational

[10] In fact, the Marshallian supply function coincides with the Hicksian since we assume quasi-linear utility. Therefore, in this case, the compensated and uncompensated elasticities are the same.

[11] The subscript on the taxable income supply is important to make explicit the dependence on the type generating heterogeneity.

[12] Note that

$$\frac{\partial c}{\partial z} = v_z \rightarrow \frac{\partial c}{\partial z}\frac{1}{zv_{zz}} = \frac{v_z}{zv_{zz}} \rightarrow \frac{\partial c}{\partial z} = \zeta z v_{zz}.$$

[13] This is for example the case in the Utilitarian problem with a bounded set of types.

approach – see Golosov, Tsyvinski, and Werquin (2014) – it is assumed that the functionals $z_\theta : T(\cdot) \to Z$ which assign for each possible tax schedule a choice of income $z \in Z$ are Lipschitz continuous. The Gateaux derivative with respect to the tax system in the direction of the reform $h : Z \to \mathbb{R}$ is given by

$$dz_\theta(T, h) = \lim_{\alpha \to 0} \frac{z_\theta(T + \alpha h) - z_\theta(T)}{\alpha}. \tag{1.9}$$

This gives the behavioral response of a type-θ agent to a small reform in direction h.

1.2.2 The failure of SMC and its implications

Typically taxation models include assumptions which discipline the relationship between marginal rate of substitution and types. In more broad contexts of mechanism design and signaling games, they are known as the Spence-Mirrlees (or single-crossing) condition. They state that the marginal rate of substitution between consumption and taxable income is decreasing in the type. In our model the marginal rate of substitution between consumption and income is given by

$$\left.\frac{dc}{dz}\right|_{V(c,z,\theta)=cte} = v_z(z, \theta). \tag{1.10}$$

Therefore, the utility function $V(c, z, \theta)$ satisfies the usual Spence-Mirrlees condition if $v_{z\theta} < 0$, for all $z \in Z$ and $\theta \in \Theta$. Graphically, in a diagram $Z \times C$, the indifference curves of individuals with lower types should be stepper (when $v_{z\theta} < 0$) and cross each other at most once. The assumption below, first used in Araújo and Moreira (2010), relaxes the requirement that the utility function satisfies the SMC.[14]

Assumption 1 *The condition $v_{z\theta}(z, \theta) = 0$ defines implicitly a monotonic function $z_0 : \Theta \to Z$ such that $v_{z\theta}(z, \theta) > 0$ for $z < z_0(\theta)$ and $v_{z\theta}(z, \theta) < 0$ for $z > z_0(\theta)$. For simplicity, assume additionally that $v_{zz\theta} > 0$.*[15]

The function $z_0(\theta)$, henceforth referred as a separating curve, splits the space $\Theta \times Z$ into two regions where the signal of the cross-derivative remains constant. Let us define $CS_- = \{(\theta, z) \in \Theta \times Z : v_{z\theta}(z, \theta) < 0\}$ and $CS_+ = \{(\theta, z) \in \Theta \times Z : v_{z\theta}(z, \theta) > 0\}$ these regions. The function $z_0(\theta)$ can be increasing or decreasing depending on the application. See Appendix A for examples of economies generating this particular type of failure of the SMC. Figure 1 plots the separating curve generated in the economy from Example 1.[16]

[14] This assumption was also used in SchottmÃČÅŠller (2015), Choné and Gauthier (2017) and Araujo, Moreira, and Vieira (2015).

[15] The assumption $v_{zz\theta} > 0$ is not crucial but will be convenient to guarantee concavity of government objective function.

[16] In Appendix A this example is identified by Example A5.

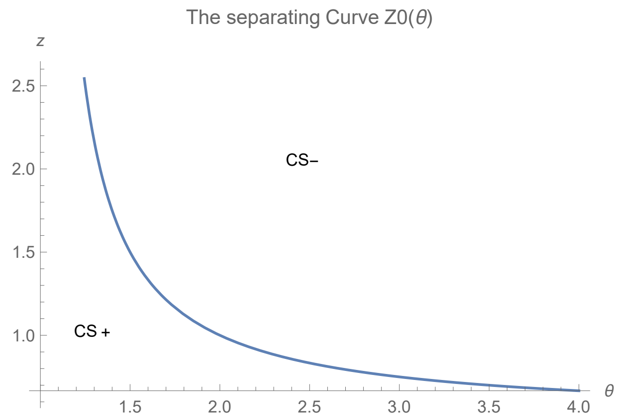


Figure 1.1: The separating curve

Example 1 (Guiding Example) *We assume that every individual in the economy has labor market productivity γ to generate taxable income (measured in units of the output good) $z = \gamma l$ at a cost in terms of utility given by $\delta h(l)$, where δ represents a taste for labor parameter. Individuals with high γ generate more income for a given level of labor supply. Individuals with low δ have lower disutility when supplying a given level l of labor. Assuming γ and δ depends on a common parameter θ in the following way: $\gamma = \theta$ and $\delta = \psi(\theta)$ the function $v(z, \theta)$ takes the form:*

$$v(z, \theta) = \psi(\theta) h\left(\frac{z}{\theta}\right) \tag{1.11}$$

assuming that ψ and h are such that there is a negative relation between δ and γ we will have the failure of SMC as in Assumption 1. The parameter θ creates a tension between productivity and tastes for labor. See Example A4 in Appendix A for more details.

In the region to the left of the separating curve the "laziness" effect dominates the "productivity" effect (high δ relative to γ) and in the region to the opposite is true. Therefore, there is no clear ordering of the marginal utility (or MRS) according to type θ. In particular, as one can see in Figure 2, indifference curves for different types may cross twice.

The natural ordering created by the SMC is an important tool to reduce the complexity of the IC constraints that had to be considered. Indeed, when the SMC is satisfied the set of implementable allocations is fully characterized by local conditions. These local incentive compatibility constraints are the usual ones found in the literature, summarized in the following lemma (the necessity of these conditions does not depend on the validity of the SMC).

Let $V : \Theta \to \mathbb{R}_+$ be the informational rent function of the agent in an incentive-feasible mechanism $c, z : \Theta \to \mathbb{R}_+ \times Z$. Hence, $V(\theta) = c(\theta) - v(z(\theta), \theta)$.

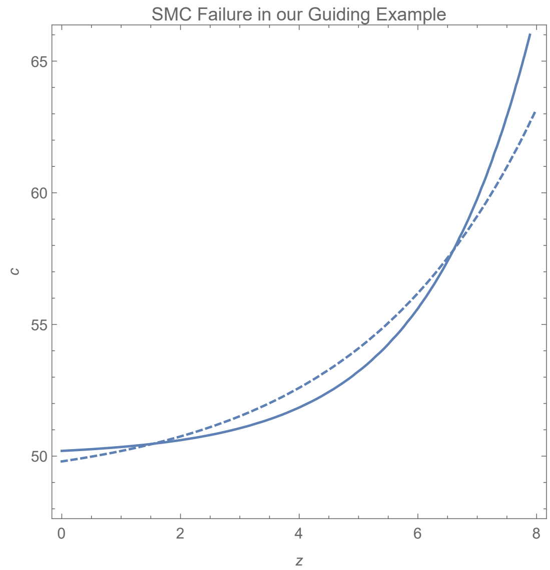


Figure 1.2: A pair of indifference curves for the economy in Example 1

Lemma 1 *Suppose that $z : \Theta \to Z$ is part of a bounded incentive-feasible allocation $c, z : \Theta \to \mathbb{R}_+ \times Z$. Then, the following are necessary local conditions for implementability.*

(i) ***(Envelope condition)*** *The agent's informational rent function, is given by*

$$V(\theta) = V(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} v_\theta(z(s), s)ds, \tag{1.12}$$

for all $\theta \in \Theta$;

(ii) ***(Monotonicity condition)*** *$z : \Theta \to Z$ is non-decreasing in the region CS_- and non-increasing in the region CS_+.*

Proof: See the Appendix. ■

Condition (i) of Lemma 5 guarantees payoff equivalence among any incentive-feasible mechanisms, and it is equivalent to the first-order condition of agents' revelation problems. Condition (ii) is related to the local second-order necessary conditions of these problems. We can pin-down the consumption from the definition of informational rent: $c(\theta) = V(\theta) + v(z(\theta), \theta)$.

When the SMC is imposed, at any allocation, the MRS for different types are ordered in the same way. Under Assumption 1, in contrast, the ordering of MRS between any two types will in general depend on the specific allocation one is considering – see Figure **??**. It is, then, natural to ask whether other non-local necessary conditions for implementability must be taken into account since, without a natural ordering, it is possible to have different and "distant" agents choosing the exact same allocation. In this case, a whole new set

of global incentive compatibility constraints must be imposed. The necessary condition described in the following proposition is derived from these global incentive compatibility constraints.

Lemma 2 (Discrete pooling condition) *Let $z : \Theta \to Z$ be an implementable allocation and $T(z)$ be the associated taxation schedule that implements it (as in Proposition 4). Suppose that $z : \Theta \to Z$ pools two different types at $z \in Z$ (i.e. $z = z(\theta) = z(\hat{\theta})$) a income level where $T : Z \to \mathbb{R}$ is differentiable. Then, we must have*

$$v_z(z, \theta) = v_z(z, \hat{\theta}). \tag{1.13}$$

Proof: Let $z : \Theta \to Z$ be an allocation. We know from Proposition 4 (taxation principle) that there exists a taxation schedule $T : Z \to \mathbb{R}$ such that the problem of the individual is

$$z(\theta) \in \arg \max_{z \in Z(\Theta)} V(z - T(z), z, \theta). \tag{1.14}$$

Given the differentiability of the tax schedule, the optimum can be characterized by the first-order condition at $z(\theta)$ that is given by

$$1 - T'(z(\theta)) = v_z(z(\theta), \theta). \tag{1.15}$$

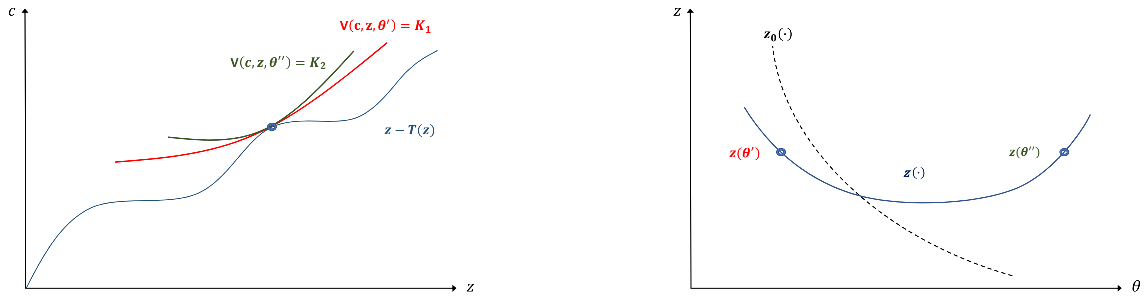
and analogously for $z(\hat{\theta})$ Therefore, since $z(\theta) = z(\hat{\theta}) = z$, we must have

$$v_z(z, \theta) = 1 - T'(z(\theta)) = 1 - T'(z(\hat{\theta})) = v_z(z, \hat{\theta}). \tag{1.16}$$

Proving the necessary condition for implementability. ■ The condition in Lemma 6, henceforth denoted by discrete pooling condition, is a direct consequence of the taxation principle and requires that discretely pooled types obtain the same marginal tax rate despite being "very different". It is important to distinguish discrete pooling from continuous pooling where all types in a neighborhood receives the same allocation. Continuous pooling usually happens when the monotonicity constraint (as in Lemma 5 (ii)) is binding and an ironing procedure is necessary to guarantee incentive-compatibility.

Figure 1.3a and 1.3b show how discrete pooling occurs in the space of indirect and direct mechanisms, respectively. Here we should note that under SMC it is impossible to have this kind of discrete pooling behavior since "very different" agents will have very different slopes for their indifference curves at a given point. That is not the case in our example.

Figure 3.4a and 3.4b show how continuous pooling occurs in the space of indirect

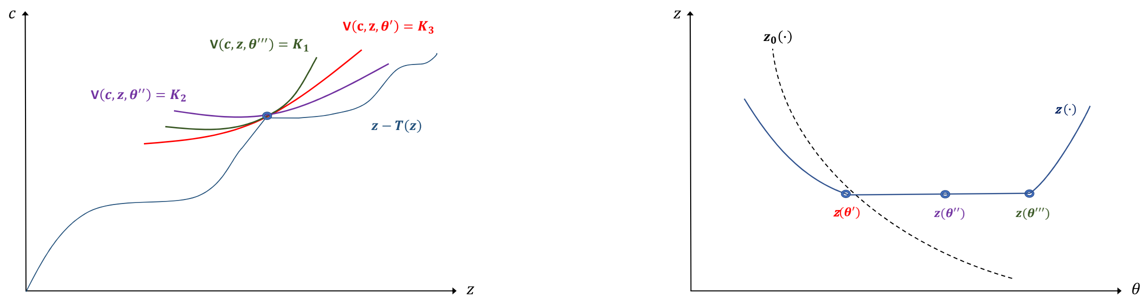


(a) Indirect Mechanism

(b) Direct Mechanism

Figure 1.3: Discrete Pooling

and direct mechanisms, respectively. Continuous pooling shows up as a flat region in the direct mechanism and a kink in the tax schedule in the indirect mechanism. The kink induces different but close agents to make the same choice even though they do not have the same MRS at their choice. At the kink, the marginal tax rate is not defined, but only the right and left derivatives are defined. Therefore, if type θ is choosing at the kink (i.e., his/her MRS is in between the right and the left derivatives) all types in a neighborhood will do as well. Note that types θ' and θ'' are continuously and discretely pooled.



(a) Indirect Mechanism

(b) Direct Mechanism

Figure 1.4: Continuous Pooling

Now let us use the GIF and pseudo-inverses to characterize global incentive constraints in a convenient way. Let $z : \Theta \to Z$ be an arbitrary allocation (direct mechanism)

and define $\psi_b, \psi_s : Z \to \Theta$ the pseudo-inverse of z as $\psi_b(z) = \inf\{\theta \in \Theta; z(\theta) \leq z\}$ and $\psi_s(z) = \sup\{\theta \in \Theta; z(\theta) \leq z\}$. Additionally, whenever $\{\theta \in \Theta; z(\theta) \leq z\} = \emptyset$ define $\psi_b(z) = \underline{\theta}$ and $\psi_s(z) = \overline{\theta}$.

Proposition 2 (Global IC) *Let $z : \Theta \to Z$ be a bounded allocation and $\psi_b, \psi_s : Z \to \Theta$ the pseudo-inverse functions of z. Therefore, for all $\theta, \hat{\theta} \in \Theta$, $z : \Theta \to Z$ is incentive compatible if and only if*

$$\int_{z(\hat{\theta})}^{z(\theta)} [v_z(z, \psi_s(z)) - v_z(z, \psi_b(z))]\, dz \geq 0. \tag{1.17}$$

Proof: Using the GIF the allocation $z : \Theta \to Z$ is incentive compatible if and only if for all $\theta, \hat{\theta} \in \Theta$, $\Phi(\theta, \hat{\theta}; z(\cdot)) \geq 0$. Note that

$$\begin{aligned} 0 \leq \Phi(\theta, \hat{\theta}; z(\cdot)) &= \\ = \int_{\hat{\theta}}^{\theta} \left[\int_{z(\hat{\theta})}^{z(s)} v_{z\theta}(t, s) dt \right] ds &= \int_{z(\hat{\theta})}^{z(\theta)} \left[\int_{\psi_b(z)}^{\psi_s(z)} v_{z\theta}(t, s) ds \right] dt \\ &= \int_{z(\hat{\theta})}^{z(\theta)} [v_z(t, \psi_s(t)) - v_z(t, \psi_b(t))]\, dt, \end{aligned} \tag{1.18}$$

where the second equality comes from Fubini's theorem and the last from the fundamental theorem of calculus. Change of variables gets the result. ■

The next corollary describes a particular global incentive-compatibility constraint that will be shown to be very relevant. It gives the conditions for which a pivotal type θ_d does not envy the type in the upper bound of type space when the allocation features a U-shape format.

Corollary 1 *Let $z : \Theta \to Z$ be a bounded, U-shaped incentive-compatible allocation. Therefore, an arbitrary type $\theta_d \in \Theta$ does not envy $\overline{\theta}$ if*

$$\int_{z_l}^{z_h} [v_z(z, \psi_s(z)) - v_z(z, \theta_d)]\, dz = 0, \tag{1.19}$$

where $\psi_b, \psi_s : Z \to \Theta$ the pseudo-inverse functions of z, $z_l = z(\theta_d)$ and $z_h = \inf\{z \in Z : \psi_s(z) = \overline{\theta}\}$.

Proof: Note that for any $z \in Z$ such that $\psi_b(z) \neq \psi_s(z)$ ($z \in [\underline{z}, z_l)$) we have discrete pooling and consequently $v_z(z, \psi_s(z)) - v_z(z, \psi_b(z)) = 0$ by Proposition 6 since $z : \Theta \to Z$ is incentive-compatible. ■

1.2.3 Government

In our quasi-linear environment, the government needs a strong reason to redistribute income across households.[17] We assume that the government follows a weighted utilitarian social welfare criterion. More specifically, the government wants to maximize a weighted average of individuals' utilities where the weights of types are given by the density function $g : \Theta \to \mathbb{R}_+$. If the distribution G first-order stochastically dominates the F,[18] the government puts higher weight on individuals' utility with lower θ. The government's problem in the structural mechanism design formulation is to choose an incentive-feasible allocation $c, z : \Theta \to \mathbb{R}_+ \times Z$ to maximize the welfare criterion.[19] We can incorporate the local incentive-compatibility and budget constraints in the objective function and rewrite the mechanism design problem of the government in the following way (see Appendix C for details on the derivation):

$$\max_{z:\Theta\to Z} \quad \int_{\underline{\theta}}^{\overline{\theta}} W(z(\theta), \theta) d\theta \tag{1.20}$$

subject to the global incentive constraints defined in equations (1.13) and (1.17) from Propositions 5 and 6. The social welfare function augmented of incentive-feasibility constraints is given by the function $W : Z \times \Theta \to \mathbb{R}$ defined as

$$W(z, \theta) = \left[z - v(z, \theta) + v_\theta(z, \theta) \frac{G(\theta) - F(\theta)}{f(\theta)} \right] f(\theta). \tag{1.21}$$

1.3 Optimal Taxation

Let us define some relaxed sub-problems that will help to build our proposed mechanism. All these solutions emerge naturally in our approach based in type assignment functions.[20]

Relaxed Solution Let us denote $z^R : \Theta \to Z$ the solution of the relaxed problem where all global-incentive compatibility constraints are ignored. Note that this relaxed problem

[17] The majority of the papers in taxation assumes the government to follow a utilitarian social welfare criterion. Under this criterion, only the difference between marginal utility of consumption across households is enough to fulfill the government's desire to redistribute income. However, with quasi-linear utility the marginal utility of consumption is constant and equal to 1 for all individuals.

[18] The distribution G first-order stochastically dominates the F iff $G(\theta) \geq F(\theta)$, for all $\theta \in (-\infty, \infty)$.

[19] Under this welfare criterion in a economy where agents have quasi-linear utility the first-best allocation is trivial. Since quasi-linearity implies transferable utility the first-best requires making all agents supply income at the efficient level and transfer everything to the agent with higher weight.

[20] We refer the reader who may be interested in the details of these relaxed problems to Araújo and Moreira (2010) and Araujo, Moreira, and Vieira (2015).

can be solved pointwise with the following Euler equation given by

$$W_z(z^R(\theta), \theta) = 0. \tag{1.22}$$

For future reference define $T_R : Z \to \mathbb{R}$ the tax schedule that would be implied by this allocation, i.e., $T_R(z(\theta)) = z_R(\theta) - c_R(\theta)$. It is important to note that this tax schedule not necessarily implements this allocation. Indeed, it will not when the allocation is not incentive compatible.

Discrete Pooling Solution Let $z^{DP} : \Theta \to Z$ denote the solution of the government problem where the discrete pooling condition as defined in equation (1.13) and monotonicity constraints as defined in Lemma 5 (ii) are considered. z^{DP} describes the solution when two "very different" types are discretely pooled. The Euler equation for this problem when types $\theta, \hat{\theta}$ are pooled at income level z (i.e., $z = z^{DP}(\hat{\theta}) = z^{DP}(\theta)$) is given by:

$$\frac{W_z(z, \theta)}{v_{z\theta}(z, \theta)} f(\theta) = \frac{W_z(z, \hat{\theta})}{v_{z\theta}(z, \hat{\theta})} f(\hat{\theta}). \tag{1.23}$$

The mechanism proposed by Araújo and Moreira (2010) has the solutions of these two subproblems as its elements connected through a vertical ironing procedure.[21]

For future reference define $T_{DP} : Z \to \mathbb{R}$ the tax schedule that would be implied by this allocation, i.e., $T_{DP}(z(\theta)) = z_{DP}(\theta) - c_{DP}(\theta)$.

Isoperimetric Solution The isoperimetric problem, as proposed by Araujo, Moreira, and Vieira (2015), has these same elements but instead of connecting the solutions z^R and z^{DP} through vertical ironing, it allows for discontinuity of the mechanism at a pivotal type θ_d. The introduction of the discontinuity requires us to take into account an additional relevant global IC - that the pivotal type does not envy the allocation prescribed to the highest type as characterized in Corollary 2.[22]

The Euler equation for this iso-perimetric problem when the GIC is binding is given by

$$W_z(z, \psi_s(z)) + \delta v_{z\theta}(z, \psi_s(z)) = 0, \tag{1.24}$$

where δ is the Lagrange multiplier associated with the binding GIC.

To illustrate these objects, Figure 1.5 plots the numerical solutions of these relaxed

[21] The vertical ironing procedure creates additional distortions on the mechanism to guarantee incentive-compatibility for the vertically pooled types.

[22] For $z \in [z_l, z_h]$ as defined in Corollary 2.

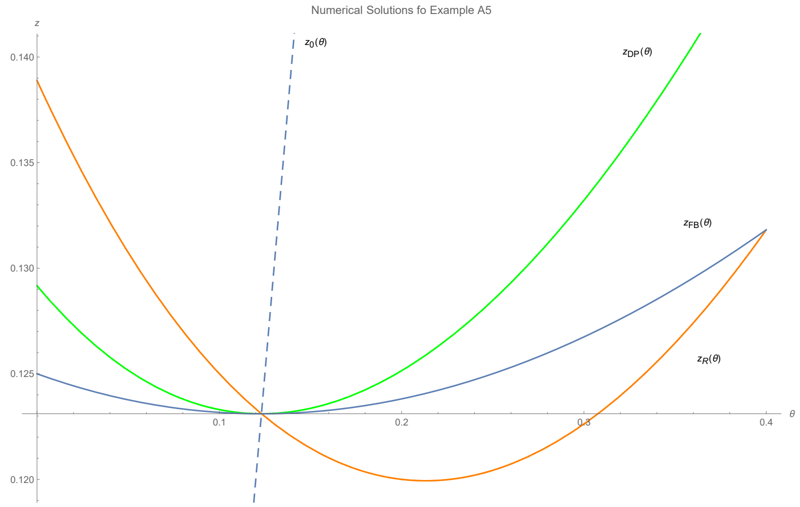


Figure 1.5: Numerical Solutions for Example A5

problems for the economy from Example A5 (see Appendix A).

1.3.1 An approach to the optimal taxation problem based on type assignment functions

The Variational Approach works in the space of indirect mechanisms, where the government proposes a tax schedule and lets individuals self-select. Although it is easier to get first-order conditions using this technique, it is hard to have a clear sense of all objects involved. In particular, it is hard to guarantee sufficiency of these first-order constraints to derive formulae for the optimal taxation.

On the other hand, the traditional mechanism design approach is very clear about the role of each structure but it is also extremely difficult to apply when the characteristics of the problem deviate from the standard model. Our suggestion is to work with the type assignment functions as proposed by Noldeke and Samuelson (2007). We can think of type assignment functions as the pseudo-inverses of a direct mechanism.[23]

The main advantage of using this approach relies on the fact that global incentive constraints are naturally described in terms of income levels instead of types. For the local IC constraints, we have the opposite. They are naturally expressed in terms of types (recall Lemma 5); by using this approach, we take the best characteristics of both.

The first step to implement this approach is to re-write the objective function of the

[23]Without SMC, it is common to have non-monotone solutions and the inverse function of any such mechanism is not well defined. However, the pseudo-inverses (to the right and to the left) are well defined.

government in equation (1.20) as follows:

$$\int_{\underline{\theta}}^{\overline{\theta}} W(z(\theta),\theta)d\theta = \int_{\underline{\theta}}^{\overline{\theta}} \left(\int_{\underline{z}}^{z(\theta)} W_z(z,\theta)dz \right) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} W(\underline{z},\theta)d\theta$$
$$= \int_{\underline{z}}^{\overline{z}} \left[\mathcal{W}(z,\psi_s(z)) - \mathcal{W}(z,\psi_b(z)) \right] dz + \int_{\underline{\theta}}^{\overline{\theta}} W(\underline{z},\theta)d\theta, \quad (1.25)$$

where $\mathcal{W} : Z \times \Theta \to \mathbb{R}$ is given by

$$\mathcal{W}(z,\theta) = \int_{\underline{\theta}}^{\theta} W_z(z,s)ds \quad (1.26)$$

and $\psi_b(\cdot)$ and $\psi_s(\cdot)$ are the type assignment functions. The first equality follows from the Fundamental Theorem of Calculus and the second from Fubini's Theorem.

Therefore we can rewrite government's optimal taxation problem in terms of type-assignment functions as choosing $\psi_b, \psi_s : Z \to \Theta$ to maximize the functional in Equation 1.25 subject to all global incentive constraints. Formally, the problem is

$$\max_{\psi_b(\cdot),\psi_s(\cdot)} \int_{\underline{z}}^{\overline{z}} \left[\mathcal{W}(z,\psi_s(z)) - \mathcal{W}(z,\psi_b(z)) \right] dz \quad (1.27)$$

subject to all global incentive constraints: for all $z_l, z_h \in Z$

$$\int_{z_l}^{z_h} \left[v_z(z,\psi_s(z)) - v_z(z,\psi_b(z)) \right] dz \geq 0. \quad (1.28)$$

Note that the set of all global incentive constraints includes in particular: (i) discrete pooling constraints, $v_z(z,\psi_b(z)) - v_z(z,\psi_s(z)) = 0$ whenever $\underline{\theta} < \psi_b(\theta) < \psi_s(\theta) < \overline{\theta}$; (ii) monotonicity constraints: ψ_b non-increasing and ψ_s non-decreasing; and (iii) and the local incentive constraints in Lemma 5 (i).

This problem has several convenient properties. First, notice that this problem is entirely described in terms of income levels. The discrete pooling condition is a point-wise constraint and the global incentive constraint is a collection of isoperimetric constraints, which is very common in Calculus of Variations. As a consequence, this formulation disentangles the local IC of global ICs, by allowing us to treat them separately.

Numerical Simulations of Examples A4 and A5 display U-shaped solutions (i.e. $z : \Theta \to Z$ such that there exists $\theta_0 \in int(\Theta)$ with $z(\cdot)$ decreasing for all $\theta < \theta_0$ and $z(\cdot)$ increasing for all $\theta > \theta_0$). Inspired by this let's restrict the class of problems to be

considered.[24]

Assumption 2 *Assume that the relaxed solution is U-shaped and cross the separating curve in its decreasing portion. Assume additionally that $z_R(\underline{\theta}) > z_R(\overline{\theta})$. Lastly, assume that the solutions are uniquely determined.*

Figure 5 illustrates the geometry of our class of problems. The typical geometry has $z_{DP}(\theta) \in [z_0(\theta), z_R(\theta)]$,[25] $z_R(\overline{\theta})$ is equal to the first best value. Te same thing happens in the point where the solutions meet, i.e., $z_R(\theta) = z_{DP}(\theta) = z_0(\theta)$. Therefore, potentially we may have the usual non-distortion at the top result as well as non-distortion at the middle. We will discuss this in greater depth later.

We believe that this class of solutions comprises many interesting applications in public economics and provides a well-behaved solution that has very intuitive and interesting characteristics. In particular, it reduces the set of global ICs to be considered.

1.3.2 A solution featuring segregation

We propose a taxation schedule where the government segments individuals in two groups: low and high income.[26] The intuition for doing such a policy is as follows. By dividing into two groups, the government can further explore the trade-off between equity and efficiency within each group. The only additional difficulty is that the government should guarantee that individuals do not envy the allocations of the other group. These are the global incentive conditions that may be binding, thus helping to broaden the design of tax systems. In particular, when the conditions are binding, the underlying assumptions for the use of variational methods are violated.

Let $\theta_d \in [\underline{\theta}, \overline{\theta}]$ be a pivotal type whose behavior will determine the segregation of groups. Before going into details, let's motivate the solution in some steps. First consider the problem

$$\max_{\psi_b(\cdot),\psi_s(\cdot)} \int_{\underline{z}}^{\overline{z}} \left[\mathcal{W}(z, \psi_s(z)) - \mathcal{W}(z, \psi_b(z))\right] dz \tag{1.29}$$

subject to:

(i) $\left[\overline{\theta} - \psi_s(z)\right] \left[v_z(z, \psi_b(z)) - v_z(z, \psi_s(z))\right] \leq 0;$[27]

[24] SchottmÃČÅŠller (2015) study a class of monotone solutions where global incentive constraints are binding even though no discrete pooling happens.

[25] With some abuse of notation, we can change the limits of the interval whenever $z_0(\theta) > z_R(\theta)$.

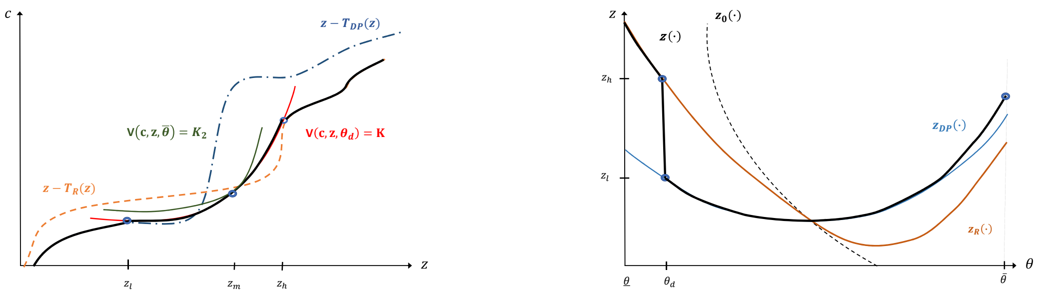
[26] The segregation in groups was inspired by Araujo, Moreira, and Vieira (2015).

[27] This is a compact way of writing the condition $v_z(z, \psi_b(z)) - v_z(z, \psi_s(z)) \leq 0$ being valid whenever $\psi_s(z) < \overline{\theta}$.

(ii) $[\psi_b(z) - \underline{\theta}]\,[v_z(z, \psi_b(z)) - v_z(z, \psi_s(z))] \geq 0$;

(iii) ψ_b and ψ_s monotonous.

In this exposition we will assume that the only relevant constraint is the condition (i). This problem is the type assignment formulation of the problem considered in Araújo and Moreira (2010). It involves parts of the relaxed solution $z_R(\cdot)$ and the discrete pooling solution $z_{DP}(\cdot)$. The transition between groups is done through a vertical ironing procedure at an endogenously determined pivotal type θ_d. This ironing procedure means that the government commits to offer any income level in the interval $[z(\theta_d^-), z(\theta_d^+)]$. Figures 1.6a and 1.6b illustrate the vertical ironing procedure.



(a) Indirect Mechanism

(b) Direct Mechanism

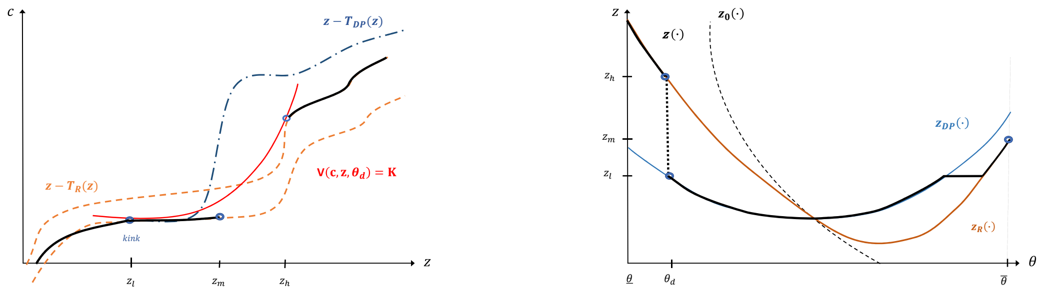
Figure 1.6: Vertical Ironing as in Araújo and Moreira (2010)

Remark 1 *The correspondence solution in this problem is convex valued and, therefore, consistent with the variational method.*

It turns out that this requirement is very restrictive since several types in the upper part of the type space are discretely pooled with the connected segment. Hence, an additional distortion should be implemented to guarantee incentive compatibility. In other words, the discrete pooling IC becomes binding more often than necessary.

In fact, we can do better. Let us construct our solution in three intuitive (and entertaining) steps. Let θ_d be the pivotal type where the vertical ironing process occurs. At this point, we can think of this pivotal type as endogenously determined in the problem of (1.27). Later, this type will parametrize a class of solutions. Define $z_h = z(\theta^-)$ and $z_l = z(\theta^+)$ the right and left limits of the allocation at this pivotal type.

Step 1 *Jump allows us to return to the relaxed solution benefiting from a lower level of distortion.*



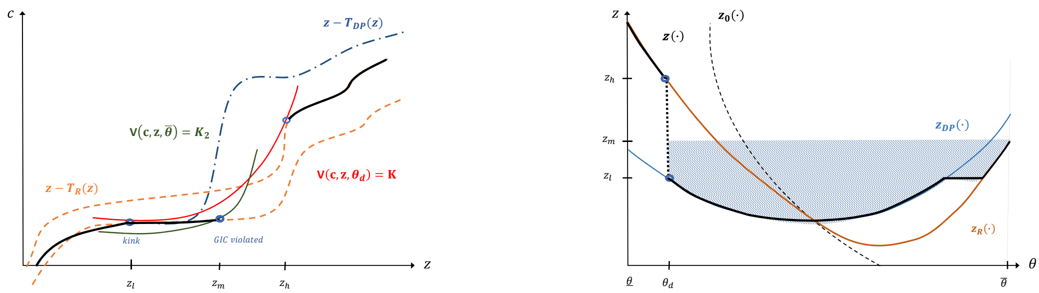
(a) Indirect Mechanism

(b) Direct Mechanism

Figure 1.7: Step 1

Instead of offering all level of income $z \in [z(\theta_d^-), z(\theta_d^+)]$, the government could offer a discontinuous allocation at θ_d. In this case, the set of agents on the right part of the type space would not be discretely pooled anymore. Hence, the allocation could return to the less distorted relaxed solution. The transition between these solutions is done through continuous pooling to guarantee monotonicity, creating a kink in the tax schedule. Define $z_m = z_R(\overline{\theta})$. This is clearly a welfare improvement since it returns to the less distorted relaxed solution (recall that the relaxed problem is a more constrained problem).

Step 2 *A global incentive compatibility constraint may be violated.*



(a) Indirect Mechanism

(b) Direct Mechanism

Figure 1.8: Step 2

In the following Lemma 7 we present the set of potential global incentive compatibility constraints that may be binding. Indeed, the only global constraints that matter are the ones where the pivotal type θ_d is indifferent between the lowest allocation on the

high-income group z_h and the highest allocation in the low-income group z_l; and that the highest type $\overline{\theta}$ does not envy the highest allocation at the low-income group z_l.

Lemma 3 *Under the conditions of Assumption 2 the set of potentially binding global incentive-compatibility constraints reduce only to the following two:*

$$\int_{z_l}^{z_h} \left[v_z(z,\psi_s(z)) - v_z(z,\theta_d)\right] dz = 0 \tag{1.30}$$

and

$$\int_{z_l}^{z_m} \left[v_z(z,\psi_b(z)) - v_z(z,\overline{\theta})\right] dz = 0. \tag{1.31}$$

These constraints represent the marginal types that are more willing to get the allocation in the other group.

These constraints are conflicting i.e., if one is binding the other one will necessary be slack. Therefore, depending on the problem, only one may be binding. For the sake of simplicity, assume that the global IC, given by equation (1.31), is the one that can potentially be binding.

The shaded area in Figure 8 marks region where this GIC is integrated over. If the area weighted by $v_{z\theta}$ is bigger to the right side of the separating curve, the allocation will not be incentive-compatible.

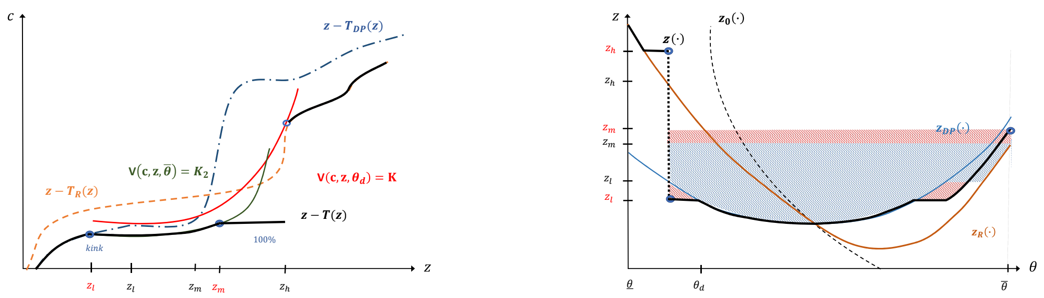
Remark 2 *If both GICs are slack, the solution in the previous step cannot be improved, i.e., it is the optimal in this class.*

Step 3 *Distort allocations to restore incentive compatibility.*

The last step requires distorting the allocation using $z_m, z_h, and z_l$ as margins to restore the GIC. By construction, we have a class of solutions parametrized by the pivotal type. The proposed solution in this paper is to maximize over the parameter θ_d, henceforth referred to as the segregation mechanism. In Appendix C we present the mathematical formulation of the optimization program.

Whenever $\theta_d \in int(\Theta)$ the solution features segregation in low and high income groups. The extreme cases $\theta_d = \underline{\theta}$ and $\theta_d = \overline{\theta}$ are the solutions proposed in Theorem 3 of Araújo and Moreira (2010) (see Figure 4 in the page 1125 for the intuition), which are consistent with the variational method.

As we can see in Figure 1.9a and 1.9b the segregation mechanism creates an extra distortion on high types to balance the impact of the binding GIC. Another unusual characteristic is the 100% marginal tax rate at the optimum. This is a empirical phenomenon that do not have any rationalization based on optimality in the taxation literature. At the best of our knowledge, we provided the first theory of why should a government implement 100% marginal tax rate in a optimal tax schedule.



(a) Indirect Mechanism

(b) Direct Mechanism

Figure 1.9: Step 3

Remark 3 *By construction, the monotonicity constraint of ψ_b and ψ_s being increasing and decreasing, respectively, is satisfied.*

Remark 4 *If there exists $z \in [\underline{z}, \overline{z}]$ such that $\psi_b^+(z) = \psi_b^-(z) \equiv \theta_d$, the optimal allocation does not present jump. Therefore, the solution in Araújo and Moreira (2010) is a degenerate case of our proposed solution.*

1.3.3 Necessary conditions

The next propositions characterize the optimal marginal tax rate with respect to the type assignment functions.

Proposition 3 *Let $\psi_b, \psi_s : Z \to \Theta$ the type assignment functions (correspondences) that solves the segregation mechanism for a given θ_d. The first-order necessary condition for the optimal marginal tax rate is given by:*

(i) for $z \in (z_h, \overline{z}]$, $\psi_s(z) = \theta_d$ and $\psi_b(z)$ satisfies

$$T'(z) = 1 - v_z(z, \psi_b(z)) = \left[\frac{G(\psi_b(z)) - F(\psi_b(z))}{f(\psi_b(z))}\right] v_{z\theta}(z, \psi_b(z)); \qquad (1.32)$$

(ii) for $z \in (z_m, z_h)$, $\psi_s(z) = \psi_b(z) = \emptyset$ and we have 100% marginal tax rate. This is indeed a sufficient but not necessary condition;

$$T'(z) = 1; \tag{1.33}$$

(iii) for $z \in [\underline{z}, z_l)$, we have discrete pooling and, $\psi_s(z)$ and $\psi_b(z)$ are defined by

$$\left[1 - v_z(z, \psi_b(z)) - \left(\frac{G(\psi_b(z)) - F(\psi_b(z))}{f(\psi_b(z))}\right) v_{z\theta}(z, \psi_b(z))\right] \frac{f(\psi_b(z))}{v_{z\theta}(z, \psi_b(z))} = \left[1 - v_z(z, \psi_s(z)) - \left(\frac{G(\psi_s(z)) - F(\psi_s(z))}{f(\psi_s(z))}\right) v_{z\theta}(z, \psi_s(z))\right] \frac{f(\psi_s(z))}{v_{z\theta}(z, \psi_s(z))}; \tag{1.34}$$

(iv) for $z \in (z_l, z_m]$, $\psi_b(z) = \theta_d$ and $\psi_s(z)$ satisfies

$$T'(z) = 1 - v_z(z, \psi_s(z)) = \left[\frac{G(\psi_s(z)) - F(\psi_s(z)) - \delta}{f(\psi_s(z))}\right] v_{z\theta}(z, \psi_s(z)), \tag{1.35}$$

where δ is the Lagrange multiplier associated with the binding GIC (either Equation 1.30 or 1.31 from Lemma 7).

Remark 5 *For $z \in \{z_l, z_h\}$ we have continuous pooling and the type assignment function is actually a correspondence.*

Theorem 1 *Whenever the segregation mechanism is not degenerate, that is, $\theta_d \in (\underline{\theta}, \overline{\theta})$, the conditions underlying the Variational Method are violated. Moreover, in this case, the segregation mechanism dominates the solution given by the variational method in terms of welfare.*

Proof: Given the formulation of our problem, this proof is straightforward. First notice that whenever the segregation mechanism is not degenerate the resulting tax schedule is discontinuous. Moreover, since the GIC constraint is binding in this case, local perturbations of the tax schedules will break the GIC of all types close to the pivotal type (or close to $\overline{\theta}$ if the binding IC is the second equation of Lemma 7). Therefore, local reforms will have discontinuous impact on agents' decisions. Therefore, $z_\theta(T)$ is not Lipschitz-continuous. ■

The intuition for the result in Theorem 1 goes as follows. According to the Variational Method, the Gateaux differential of the welfare function, taking the optimal mechanism as the baseline, should be zero in all directions. However, we can find directions where the Gateaux derivative of the welfare function using the segregation mechanism as

the baseline will not be zero. As mentioned in the introduction, local perturbations may have first-order impact on the welfare.

Conjecture 1 *The taxation scheme proposed in Proposition 7 is the optimal income tax mechanism in the class of cÃădlag mechanisms.*

Example 2 *Suppose that the common prior for the parameter θ is a uniform distribution over the interval $[0,1]$. The first-order necessary conditions for the marginal tax rate in Proposition 7 can be written as:*

(i) for $z \in (z_h, \overline{z}]$, $\psi_s(z) = \theta_d$ and $\psi_b(z)$ satisfies

$$T'(z) = [G(\psi_b(z)) - \psi_b(z)]\, v_{z\theta}(z, \psi_b(z)) \tag{1.36}$$

(iii) for $z \in [\underline{z}, z_l)$, we have discrete pooling and $\psi_s(z)$ and $\psi_b(z)$ are defined by

$$T'(z)\left[\frac{1}{v_{z\theta}(z,\psi_b(z))} - \frac{1}{v_{z\theta}(z,\psi_s(z))}\right] = [G(\psi_b(z)) - G(\psi_s(z))] - [\psi_s(z) - \psi_b(z)] \tag{1.37}$$

(iv) for $z \in (z_l, z_m]$, $\psi_b(z) = \theta_d$ and $\psi_s(z)$ satisfies

$$T'(z) = [G(\psi_s(z)) - \psi_s(z) - \delta]\, v_{z\theta}(z, \psi_s(z)), \tag{1.38}$$

where δ is the Lagrange multiplier associated with the active GIC.

1.4 Conclusion

In this paper, we studied the design of income tax schedules in a class of economies where the Spence-Mirrlees condition is violated. We showed that the lack of ordination created by the absence of this condition typically generates allocations featuring non-monotonicities and discontinuities. These properties are novelties in the taxation literature and may invalidate the usage of variational methods. We also proposed a new methodology to tackle the design of income taxes based on type assignment functions.

The next steps in the literature could involve a better understanding of the "correct" elasticities implied by our analysis, as well as an investigation of its behavior, and the possible usage of quasi-experimental designs to estimate them. Our main conclusion is that additional caution should be taken when using variational methods in complex environments, given the potential non-regularity problems.

Appendix

Appendix A - examples of the function $v(z, \theta)$

The function $v(z, \theta)$ is flexible enough to summarize a big class of economies used in several applications in the public economics literature.

Example A1 Diamond (1998) was the first to use quasi-linear utility function in a typical Mirleessian economy. Types represent individual idiosyncratic productivity θ in the labor market. Individuals generate taxable income $z = \theta l$ By supplying l "units" of labor at a cost in utility terms given by a function $h(l)$ assumed to be sufficiently well behaved. In this model, the function $v(\cdot)$ takes the form

$$v(z, \theta) = h\left(\frac{z}{\theta}\right). \tag{1.39}$$

Typically $h(\cdot)$ is increasing, convex and twice continuously differentiable function.

Example A2 Another very common used utility function in taxation literature is

$$v(z, \theta) = \theta h(z), \tag{1.40}$$

where type θ represents an idiosyncratic labor taste parameter and $h(\cdot)$ represents the disutility incurred in supplying $l = z$ unities of labor.

Example A3 In Alves, da Costa, and Moreira (2017a) the functional form

$$v(z, \theta) = \min_{z_a, z_b \geq 0} \left\{ h\left(\frac{z_a}{\theta_a}\right) + h\left(\frac{z_b}{\theta_b}\right) \text{ s.t. } z_a + z_b = z \right\} \tag{1.41}$$

is used to denote couples efficient labor supply decision in a unitary model of household. In this application type $\theta = (\theta_a, \theta_b)$ is a two-dimensional vector representing the idiosyncratic labor productivity parameter of each spouse in a couple and $h(l)$ represents the disutility incurred in supplying $l = z/\theta_i$ unities of labor.

The next two examples do not satisfy the Spence-Mirrlees condition property and are helpful in guiding our discussion.

Example A4 This example mixes the structure of examples A1 and A2 to create a very reasonable economy where agents' utilities do not satisfy SMC. Assume that every individual in the economy has labor market productivity γ to generate taxable income (measured in units of the output good) $z = \gamma l$ at a cost in terms of utility given by $\delta h(l)$, where δ represents a taste labor parameter. Individuals with high γ generates more

income for a given level of labor supply. Individuals with low δ have lower disutility when supplying a level l of labor. Assuming γ and δ depend on a common parameter θ in the following way: $\gamma = \theta$ and $\delta = \psi(\theta)$. In this economy the function $v(z, \theta)$ takes the form[28]

$$v(z, \theta) = \psi(\theta) h\left(\frac{z}{\theta}\right). \tag{1.42}$$

For these functional forms,

$$v_z(z, \theta) = \frac{\psi(\theta)}{\theta} h'\left(\frac{z}{\theta}\right), \quad v_{zz}(z, \theta) = \frac{\psi(\theta)}{\theta^2} h''\left(\frac{z}{\theta}\right), \tag{1.43}$$

and

$$v_{z\theta}(z, \theta) = \frac{\psi(\theta)}{\theta^2} h''(l) l \left[\frac{\frac{d}{d\theta}(\psi(\theta)/\theta)}{\psi(\theta)/\theta} \frac{h'(l)}{h''(l) l} \theta - 1 \right]. \tag{1.44}$$

Therefore, $v_{z\theta}(z, \theta) \geq 0$ (≤ 0) if and only if

$$\frac{\frac{d}{d\theta}(\psi(\theta)/\theta)}{\psi(\theta)/\theta} \frac{h'(l)}{h''(l) l} \theta - 1 \leq 0 \tag{1.45}$$

or equivalently

$$\left[\frac{\psi'(\theta)}{\psi(\theta)} \theta - 1 \right] \frac{h'(l)}{h''(l) l} \leq 1 \quad (\geq 1). \tag{1.46}$$

Let $\epsilon(l)$ be the elasticity of labor supply,

$$\epsilon(l) = \frac{h'(l)}{h''(l) l}. \tag{1.47}$$

It is constant equal to ϵ for the iso-elastic case,

$$h(l) = \frac{1}{1 + 1/\epsilon} l^{1+1/\epsilon}. \tag{1.48}$$

It is equals to ϵ/l for the exponential case,

$$h(l) = \exp(l/\epsilon). \tag{1.49}$$

Now assume that $\psi(\theta) = \theta - b$ with $b < \underline{\theta}$, the condition defining the function $z_0(\theta)$ simplifies to

$$\left[\frac{\theta}{\theta - b} - 1 \right] \leq \frac{1}{\epsilon(l)}. \tag{1.50}$$

[28] In the numerical exercises we make $h(l) = \exp(\epsilon l)$ and $\psi(\theta) = \theta - b$.

In the iso-elastic case we have a vertical line at

$$\theta = b(\epsilon - 1). \tag{1.51}$$

In the more interesting case of exponential function we have a decreasing function separating the regions as assumed in Assumption 1. Indeed,

$$z_0(\theta) = \epsilon \frac{b}{1 - b/\theta}. \tag{1.52}$$

Example A5 This example is adapted from Araujo, Moreira, and Vieira (2015) to our taxation context. The utility function is

$$v(z, \theta) = z\theta^2 - b^2(\theta + 4)\frac{z^2}{2} + 1. \tag{1.53}$$

In this example type θ does not have a clear interpretation but it can be thought as a combination of different forces affecting the agent's utility in a non-trivial way. Despite the artificiality, the simplicity of this example is very convenient to allow us to quickly assess the potential pitfalls of using a variational approach in an environment where the SMC condition fails. Indeed, the separating curve $z_0(\theta)$ (Assumption 1) in this particular example is a increasing line given by

$$z_0(\theta) = \frac{\theta}{b^2}. \tag{1.54}$$

To simplify things even further, we assume that the government follows a Rawlsian social welfare criterion.[29] Assume that types are uniformly distributed on the interval $\Theta = [0, 1]$. We simulated this example using Wolfram's Mathematica. We can use the solution in Figure **??** to illustrate the relaxed subproblems in a concrete example.

Appendix B - Proofs

Proof of Proposition 4

Proof: Let $z, c : \Theta \to Z \times \mathbb{R}_+$ be an incentive-feasible allocation. Let $T : Z \to \mathbb{R}$ be defined as follows. For all $z \in z(\Theta)$, define $T(z) = z(\theta) - c(\theta)$. For $z \in z(\Theta)^c$ make $T(z) = z$ and $z \in z(\Theta)^c$ make $T(z) = z$. Let F^* to denote the income distribution in the

[29] As in Alves, da Costa, and Moreira (2017a), the solution for this problem is equivalent to a solution of a dual program where the government maximizes tax revenue subject to the incentive constraints and a minimum utility requirement for the least well-off individual.

economy. Note that,

$$\int_Z T(z)dF^*(z) = \int_{Z(\Theta)} T(z)dF^*(z) + \int_{Z(\Theta)^c} T(z)dF^*(z)$$
$$= \int_{Z(\Theta)} T(z(\theta))dF^*(z(\theta)) + \int_{Z(\Theta)^c} zdF^*(z) \geq \int_\Theta [z(\theta) - c(\theta)]f(\theta)d\theta \geq 0. \quad (1.55)$$

where in the first inequality we use that $Z(\Theta) \subset Z \subset \mathbb{R}_+$ and the last inequality comes from the feasibility of the allocation, proving (i).

For (ii) take $\theta \in \Theta$ arbitrary, note that for any $z \in z(\Theta)^c$

$$V(z - T(z), z, \theta) = z - T(z) - v(z, \theta) = z - z - v(z, \theta) \leq -v(0, \theta) \quad (1.56)$$

Therefore, $z \in z(\Theta)^c$ cannot be optimal [30]

Now take any $z \in z(\Theta)$, therefore $z = z(\hat{\theta})$, for some $\hat{\theta} \in \Theta$. By incentive compatibility, we have

$$\begin{aligned}
&V(z(\theta) - T(z(\theta)), z(\theta), \theta) = z(\theta) - T(z(\theta)) - v(z(\theta), \theta) \\
&= z(\theta) - [z(\theta) - c(\theta)] - v(z(\theta), \theta) = c(\theta) - v(z(\theta), \theta) \\
&\geq c(\hat{\theta}) - v(z(\hat{\theta}), \theta) = z(\hat{\theta}) - [z(\hat{\theta}) - c(\hat{\theta})] - v(z(\hat{\theta}), \theta) \\
&\qquad = z(\hat{\theta}) - T(z(\hat{\theta})) - v(z(\hat{\theta}), \theta) = V(z - T(z), z, \theta),
\end{aligned}$$

proving (ii). Therefore, any incentive-feasible allocation can be implemented by a non-linear tax schedule.

Now take a tax schedule $T : Z \to \mathbb{R}$ and let $z, c : \Theta \to Z \times \mathbb{R}_+$ the allocation implemented by this tax schedule where $z(\theta)$ represents the income that type-θ agent chooses to get when facing the proposed tax schedule. Let us show that it is incentive-feasible. The agent consumes all his income net-of taxes, then $c(\theta) = z(\theta) - T(z(\theta))$ and the government budget constraint implies

$$\int_\Theta [z(\theta) - c(\theta)]f(\theta)d\theta = \int_\Theta z(\theta) - [z(\theta) - T(z(\theta))]f(\theta)d\theta$$
$$= \int_\Theta T(z(\theta))f(\theta)d\theta = \int_Z T(z)dF^*(z) \geq 0, \quad (1.57)$$

where the last inequality follows from the budget balance condition. Therefore, the allocation is feasible. To prove the incentive compatibility note that since $z(\theta)$ is the choice

[30] Assume for simplicity that $0 \in z(\Theta)$.

facing the tax schedule, for any other $z \in Z$ and in particular $z(\hat{\theta})$. We have

$$
\begin{aligned}
c(\theta) - v(z(\theta), \theta) &= z(\theta) - T(z(\theta)) - v(z(\theta), \theta) \\
&\geq z(\hat{\theta}) - T(z(\hat{\theta})) - v(z(\hat{\theta}), \theta) = c(\hat{\theta}) - v(z(\hat{\theta}), \theta),
\end{aligned}
$$

where the inequality follows from the choice. Therefore, the allocation reached with an income tax is incentive compatible. ■

Proof of Lemma 5

These are usual results in the mechanism design literature. We will present a proof in the less general case when the mechanism is differentiable because it is more simple. For more general versions of this result we refer to the classical Milgrom and Segal (2002) and Rochet (1987). **Proof:** Assume that $z : \Theta \to \mathbb{R}_+$ is differentiable in the interior of CS_+ (and CS_- respectively). Let $V(\theta', \theta)$ be the utility that θ-agent gets by announcing to be type θ' in the mechanism. Incentive compatibility requires $V(\theta, \theta) \geq V(\theta', \theta)$, for all $\theta' \in \Theta$. By differentiability we can characterize the optimum using first and second-order conditions. The first-order condition for incentive compatibility is

$$
\left. \frac{\partial}{\partial \theta'} V(\theta', \theta) \right|_{\theta' = \theta} = 0; \tag{1.58}
$$

and the second-order condition for incentive compatibility is

$$
\left. \frac{\partial^2}{\partial \theta'^2} V(\theta', \theta) \right|_{\theta' = \theta} \leq 0. \tag{1.59}
$$

The FOC boils down to

$$
\frac{d}{d\theta} c(\theta) - v_z(z(\theta), \theta) \frac{d}{d\theta} z(\theta) = 0. \tag{1.60}
$$

On the other hand, using it, we have

$$
\begin{aligned}
\dot{V}(\theta) &= \frac{\partial}{\partial \theta} V(\theta, \theta) \\
&= \frac{d}{d\theta} c(\theta) - v_z(z(\theta), \theta) \frac{d}{d\theta} z(\theta) - v_\theta(z(\theta), \theta) \\
&= -v_\theta(z(\theta), \theta)
\end{aligned} \tag{1.61}
$$

and applying the fundamental theorem of calculus we have (i). Taking the total derivative of the FOC we have

$$
\frac{d^2}{d\theta'^2} V(\theta, \theta) = -\frac{d^2}{d\theta' d\theta} V(\theta, \theta). \tag{1.62}
$$

Then, from the SOC we have

$$\frac{d^2}{d\theta' d\theta} V(\theta, \theta) = v_{z\theta}(z(\theta), \theta) \frac{d}{d\theta} z(\theta) \geq 0. \tag{1.63}$$

Therefore, $\frac{dz(\theta)}{d\theta}$ and $v_{z\theta}(z, (\theta)\theta)$ should have opposite signs, proving (ii). ■

Proof of Lemma 7

Proof: This is a very intuitive result that follows from the function $v(z, \theta)$ being continuously differentiable and the monotonicity constraints of the type assignment functions $\psi_b(\cdot), \psi_s(\cdot)$. ■

Derivation of the global incentive function

Using Lemma 5 (i), it is convenient to write the incentive compatibility constraint using the following global incentive function (GIF):

$$\Phi(\theta, \hat{\theta}; z(\cdot)) \equiv \int_{\hat{\theta}}^{\theta} \left[\int_{z(\hat{\theta})}^{z(s)} v_{z\theta}(t, s) dt \right] ds. \tag{1.64}$$

Fix a mechanism $z : \Theta \to \mathbb{R}_+$ and take $\theta \in \Theta$. Incentive compatibility requires $V(\theta, \theta) \geq V(\theta', \theta)$, for all $\theta' \in \Theta$. This is equivalent to

$$\begin{aligned} & V(\theta) - V(\theta', \theta) = V(\theta) - c(\theta') + v(z(\theta'), \theta) \geq 0 \iff \\ & V(\theta) - V(\theta') + v(z(\theta'), \theta) - v(z(\theta'), \theta') \geq 0 \iff \\ & \qquad \int_{\theta'}^{\theta} [v_\theta(z(s), s) - v(z(\theta'), s)] \, ds = \int_{\theta'}^{\theta} \left[\int_{z(\theta')}^{z(s)} v_{z\theta}(t, s) dt \right] ds \geq 0. \end{aligned}$$

Therefore, defining

$$\Phi(\theta, \theta', z(\cdot)) = \int_{\theta'}^{\theta} \left[\int_{z(\theta')}^{z(s)} v_{z\theta}(t, s) dt \right] ds, \tag{1.65}$$

we have $\Phi(\theta, \theta', z(\cdot)) \geq 0$ for all $\theta, \theta' \in \Theta$ if and only if $z : \Theta \to \mathbb{R}_+$ is incentive compatible.

By construction, an allocation $z : \Theta \to \mathbb{R}_+$ is incentive compatible iff, for all $\theta, \hat{\theta} \in \Theta$, $\Phi(\theta, \hat{\theta}; z(\cdot)) \geq 0$. As we can see above, under the SMC, $v_{z\theta}(t, s)$ would have a constant sign and the necessary monotonicity condition would also be sufficient for incentive compatibility.[31] Without the SMC, we do not have a natural ordering of agents and the impact of $v_{z\theta}(t, s)$ to the right side of the separating curve has to be compensated with the impact

[31]Notice that for $\theta, \hat{\theta}$ such that $z(\theta) = z(\hat{\theta})$, we have $\Phi(\hat{\theta}, \theta) = -\Phi(\theta, \hat{\theta})$.

to the left side on the GIF to get incentive compatibility.

Appendix C

Incorporating local incentive-compatibility and budget constraints in the objective function.

The mechanism design problem of the government is to choose the best allocation rule $c, z : \Theta \to \mathbb{R}_+ \times Z$ assigning a pair of income and consumption to each type inducing truthful revelation. Formally,

$$\max_{c(\cdot),z(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} V(c(\theta), z(\theta), \theta)g(\theta)d\theta \tag{1.66}$$

subject to the budget constraint,

$$\int_{\theta\in\Theta} [z(\theta) - c(\theta)]f(\theta)d\theta \geq 0; \tag{1.67}$$

and the incentive-compatibility constraints: for all $\theta \in \Theta$,

$$\theta \in \arg\max_{\theta'\in\Theta} V(c(\theta'), z(\theta'), \theta). \tag{1.68}$$

This formulation is very convenient because it makes easy to incorporate the local incentive-compatibility constraints and the government budget constraints into the objective function.

Recall the definition of the informational rent get by an agent with type θ in an incentive-compatible mechanism is given by $V(\theta) = c(\theta) - v(z(\theta), \theta)$.[32] Using the envelope condition (Lemma 5 (i)) we can eliminate consumption from the problem.

$$c(\theta) = V(\theta) + v(z(\theta), \theta) = V(\underline{\theta}) + v(z(\theta), \theta) - \int_{\underline{\theta}}^{\theta} v_\theta(z(s), s)ds. \tag{1.69}$$

[32]Recall that Lemma 5 (i) implies payoff equivalence, up to a constant, for all incentive-compatible mechanism.

Plugging it into the budget constraint we have

$$
\begin{aligned}
0 = \int_{\underline{\theta}}^{\overline{\theta}} [z(\theta) - c(\theta)] f(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - V(\underline{\theta}) - v(z(\theta), \theta) + \int_{\underline{\theta}}^{\theta} v_\theta(z(s), s) ds \right] f(\theta) d\theta \\
= -V(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - v(z(\theta), \theta) + \int_{\underline{\theta}}^{\theta} v_\theta(z(s), s) ds \right] f(\theta) d\theta \\
= -V(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - v(z(\theta), \theta) + v_\theta(z(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} \right] f(\theta) d\theta \quad (1.70)
\end{aligned}
$$

Where we used the fact that the budget constraint must be satisfied with equality since agents utility are strictly increasing in consumption. Re-arranging to have

$$
V(\underline{\theta}) = \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - v(z(\theta), \theta) - v_\theta(z(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} \right] f(\theta) d\theta. \quad (1.71)
$$

Equation (1.75) shows the guaranteed utility level that the lower bound on type space gets in any incentive-feasible allocation. Together with Lemma 5, we see that the government guarantees an utility level that is taxed/subsidized away as type increases increases.

Substituting Equation 1.69 in the objective function of the government we have

$$
\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} g(\theta) \left[c(\theta) - v(z(\theta), \theta) \right] d\theta = \int_{\underline{\theta}}^{\overline{\theta}} g(\theta) \left[V(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} v_\theta(z(s), s) ds \right] d\theta \\
= V(\underline{\theta}) - \int_{\underline{\theta}}^{\overline{\theta}} g(\theta) \left[\int_{\underline{\theta}}^{\theta} v_\theta(z(s), s) ds \right] d\theta = V(\underline{\theta}) - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(z(\theta), \theta) \left[1 - G(\theta) \right] d\theta \\
= \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - v(z(\theta), \theta) + v_\theta(z(\theta), \theta) \frac{G(\theta) - F(\theta)}{f(\theta)} \right] f(\theta) d\theta \quad (1.72)
\end{aligned}
$$

Therefore, defining the function $W : Z \times \Theta \to \mathbb{R}$ given by

$$
W(z, \theta) = \left[z - v(z, \theta) + v_\theta(z, \theta) \frac{G(\theta) - F(\theta)}{f(\theta)} \right] f(\theta). \quad (1.73)
$$

and the problem of the government can be written as

$$
\max_{z(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} W(z(\theta), \theta) d\theta \quad (1.74)
$$

subject to all additional global incentive compatibility constraints.

Lemma 4 *Let $z : \Theta \to Z$ be an incentive-feasible allocation. Then the informational rent*

of the lowest type is given by

$$V(\underline{\theta}) = \int_{\underline{\theta}}^{\overline{\theta}} \left[z(\theta) - v(z(\theta),\theta) + v_\theta(z(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} \right] f(\theta) d\theta. \tag{1.75}$$

This lemma provides the guaranteed utility level that the lower bound on type space gets in any incentive-feasible allocation. Together with Lemma 5, we see that the government guarantees an utility level that is taxed/subsidized away as type increases. It is given by the average "virtual welfare" of all agents in the economy.

The Segregation Problem

Let $\theta_d \in \Theta$ be a pivotal type that parametrizes the program and let $Z = [\underline{z}, \overline{z}]$ to be the set of possible income values. The Segregation Problem is defined as follows.[33]

$$\max_{\psi_b,\psi_s,z_l,z_m,z_h} \int_{\underline{z}}^{z_l} \left[\mathcal{W}(z,\theta_d) - \mathcal{W}(z,\psi_b(z))\right] dz + \int_{z_l}^{z_m} \left[\mathcal{W}(z,\psi_s(z)) - \mathcal{W}(z,\theta_d)\right] dz + \int_{z_h}^{\overline{z}} \left[\mathcal{W}(z,\psi_s(z)) - \mathcal{W}(z,\psi_b(z))\right] dz \tag{1.76}$$

subject to:

(i) Monotonicity, ψ_b non-increasing and ψ_s non-decreasing.

(ii) For all $z \in [z_h, \overline{z}]$,

$$\left[\psi_b(z) - \underline{\theta}\right]\left[v_z(z,\psi_b(z)) - v_z(z,\psi_s(z))\right] \geq 0, \tag{1.77}$$

and

$$\underline{\theta} \leq \psi_b(z) \leq \psi_s(z) \leq \theta_d; \tag{1.78}$$

(iii) For all $z \in [\underline{z}, z_l]$,

$$\left[\overline{\theta} - \psi_s(z)\right]\left[v_z(z,\psi_b(z)) - v_z(z,\psi_s(z))\right] \leq 0; \tag{1.79}$$

[33]Since $\psi_b, \psi_s : Z \to \Theta$ should assign a type for every possible income level, for $z \in (z_m, z_l)$ assign $\psi_b(z) = \theta_d$ and $\psi_s(z) = \overline{\theta}$ completing the definition of the type assignment functions.

and

$$\theta_d \leq \psi_b(z) \leq \psi_s(z) \leq \overline{\theta}; \tag{1.80}$$

(iv) and the GIC from Lemma 7,

$$\int_{z_l}^{z_h} v_z(z, \psi_s(z)) - v_z(z, \theta_d) dz \geq 0. \tag{1.81}$$

$$\int_{z_l}^{z_m} v_z(z, \psi_b(z)) - v_z(z, \overline{\theta}) dz \geq 0. \tag{1.82}$$

The solution for this problem is parametrized by the pivotal type. We can choose it in order to maximize the welfare criterion. Let us denote it by the optimal segregation mechanism.

Call δ_{θ_d} and $\delta_{\overline{\theta}}$ the Lagrange multiplier associated with GIC in equations (1.30) and (1.31). Since by Lemma 7 only one of the GIC will be binding, i.e. $\delta_{\theta_d} > 0$ if and only if $\delta_{\overline{\theta}} = 0$ and vice-versa, we can define the Lagrange multiplier of the active constraint as $\delta = \delta_{\theta_d} + \delta_{\overline{\theta}}$.

There are some very interesting aspects of this formulation. First of all, This program can be solved point-wisely which makes the problem much more tractable than the direct mechanism design and variational approaches.

The usual solutions proposed in the literature are degenerate cases of the segregating mechanism. In particular, whenever $\theta_d \in \{\underline{\theta}, \overline{\theta}\}$ the resulting solution will be consistent with the Variational method as defined in Golosov, Tsyvinski, and Werquin (2014).

Chapter 2

Should the government take sides in couples quarrels?

Abstract

We derive the optimal joint tax schedule for household earnings. By restricting the tax schedules to depend only on total household income and the number of earners, and by restricting preferences to be of the isoelastic type we are able to reduce the dimensionality of the screening problem and to deal with different degrees of assortative mating, heterogeneity in the spouses' bargaining power and participation decisions. We assess the role of dissonance, a divergence between the planner's and the household's objectives, in shaping the optimal schedule. We perform numerical illustrations to show in which extent those heterogeneities affect optimal taxation.

2.1 Introduction

Households taxation theory has always been limited by the multidimensional nature of the screening problem to which it gives rise. As a consequence, most contributions to the literature restricts its scope, by focusing on pre-defined functional forms for taxes, Boskin and Sheshinski (1983a), relevant decision margins, Kleven, Kreiner, and Saez (2009), etc. In this paper, we contribute to this literature by exploring the design of optimal tax schedules restricted to those which are based on household total income.

Despite a growing tendency towards the use of individual taxes, the use of joint schedules is still pervasive. The usual rationale is that one should not treat differently couples with identical earnings. Although we do not side strictly with this view, we recognize the compelling argument by Gordon and Kopczuk (2014), who shows that from

an equity perspective, joint taxes are a better departure point for tax design than separate taxes.[1] This motivates us to take on the problem of optimal joint tax design.

From a technical perspective, this restriction allows us to maintain full generality on skill distributions, the relative bargaining power of spouses and correlation between spouses' skills. Indeed, the inherent multi-dimensionality that plagues optimal household taxation, here endogenously collapses into a single-dimensional problem where single-crossing is preserved. This allows us to handle a rich enough distribution of household types that can be successfully taken to the data. Moreover, we are able to explore the consequences of dissonance, as defined by Apps and Rees (2011), to the design of tax schedules.

Households in our model are comprised of a primary and a secondary earner, each with his or her own utility function, and are characterized by a pair of individual labor market productivities and relative Pareto weights. We assume Pareto weights to be exogenous to policy, but we do allow for families to be arbitrarily heterogeneous with respect to this Parameter. Most analyses assume that the social planner aims to redistribute across families, under the view that households' welfare is a well defined concept from both a positive and a normative perspective. This kind of assumption is, however, devoid of any meaning if one is committed to methodological individualism, a central tenet in the economic analysis. We, instead, focus on each individual's welfare, which often leads to differences between the relative importance that the planner attaches to each spouse's well-being, and the relative importance spouses attach themselves. We show that this type of dissonance, using the terminology of Apps and Rees (2011), leads to a deviation from the traditional Mirrlees' (1971b) formula that takes into account the bargaining power of spouses both to assess the marginal social value of income in the hands of each spouse and to assess the impact of marginal distortions on the amount of leisure consumed by spouses.

We perform counterfactual exercises comparing the standard and modified tax formulas, when we vary the spouses' bargaining power and the degree of assortative mating.

Thus far we have not discussed the possibility that one spouse might opt not to work. In many households, however, secondary earners enter and leave the labor market relatively frequently. To account for this possibility we allow labor supply choices on the extensive margin in which case tax schedules are allowed to vary with the number of earners. The fact that participation depends on the tax schedules, raises some new theoretical issues. We benefit from the technical innovations in Rothschild and Scheuer

[1] Gordon and Kopczuk (2014), starts with a schedule that depends only on total household earnings (or on individual earning in the case of individual taxation) and estimates the benefits of conditioning taxes on other observables. They show that, departing from separate taxes requires more conditioning and accomplishes less in terms of promoting equity.

(2013) and Gomes, Lozachmeur, and Pavan (2017) to characterize these schedules while respecting the endogeneity of the participation decision.

The rest of the paper is organized as follows. After this introduction, there is a brief literature review. Section 2 presents the model. Sections 3 develops the optimal taxation results. Section 4 perform numerical simulations. Section 5 discusses the implications of the tax system on the extensive margin decisions and Section 6 concludes.

Related literature

Until the 1970's, families were modeled on a common preference base. In particular, all family members' resources were assumed to be pooled in order to maximize a single objective function. However, seminal contributions of Manser and Brown (1980) and McElroy and Horney (1981) caused the family economics literature to progress towards the understanding of joint family decisions driven by divergent interests within the household.[2] There are many policy implications when families are no longer seen as a common preference institution. Tax policies should be revised after this huge progress on the family economics literature and, in particular, on family taxation. Although we do maintain the assumption that power is not affected by policy, we allow policy to respond to differences in power across couples.

Kleven, Kreiner, and Saez (2009) analyze the general nonlinear optimal income tax for couples. In their model, the primary earner chooses labor supply in a continuum, but the secondary earner makes decisions on participating or not in the labor market. If the secondary earner opts to participate in the job market, the labor supply is given, whereas in our environment both spouses choose in the intensive margin, or both margins. They show under which conditions the optimal tax scheme displays a positive tax on secondary earnings and that taxes on secondary earnings decrease with primary earnings.

Jacquet and Lehmann (2016) obtain an optimal tax formula when individuals face multidimensional heterogeneity. According to the authors, optimal tax formula should average, with specific weights, the sufficient statistics of individuals who earn the same income. One could think that the household is a unitary bidimensional tax agent where each dimension refers to a spousal productivity. However, their problem is not deeply concerned with bargaining power affecting the mechanism, and neither with labor market participation.

Guner, Kaygusuz, and Ventura (2012) quantify the effects of tax reforms taking carefully into account the labor supply of married females as well as the current demographic

[2] Lundberg and Pollak (1996) for example, the adoption of bargaining models to allow for independent agency of men and women in marriage.

structure. Married females have a heterogeneous cost of participating. They analyze how the structure of taxation can affect the participation decision, which is empirically responsive to tax perturbations. In our setting, both men and women have a heterogeneous participation cost and they do not have the option to file separately if both spouses are working.

Boskin and Sheshinski (1983b) explore the non-optimality of tax schedules, like the one adopted in the US, once husbands and wives face equal marginal tax rates. They discuss the possibilities of taxing spouses equally or differently, and even to subsidize one of the family's member. They show that taxing earnings of husbands and wives at the same rate is inefficient because it reduces disproportionately the female labor market participation.

Rothschild and Scheuer (2013) and Gomes, Lozachmeur, and Pavan (2017) study the design of optimal tax systems in a Roy model, where productivity is sector-specific. Their approach provides interesting insights about taxation and endogenous sector choice. In our economy, however, there is only one sector and two tax systems (depending on whether one or two spouses are working). Hence, the couple's decision of which tax schedule to choose is similar to the choice of an individual electing a sector-specific tax schedule in a two-sector economy.

2.2 Model

The economy is inhabited by a continuum of households (or families) with measure normalized to one. Each household is comprised by two spouses indexed by $i \in \{a; b\}$, where a stands for women and b for men.[3] Households are parametrized by a two-dimensional vector denoting each spouses labor market productivity (or wage) $w_i \in [\underline{w}; \bar{w}]$ in a linear production function of income given by $z_i = w_i l_i$. Individuals derive utility from the consumption of a private consumption good $x_i \in \mathbb{R}_+$ and disutility from marketable labor supply l_i resulting in a total utility given by $U(x_i, l_i) = u(x_i) - h\left(l_i\right)$, common to all individuals.

Let's assume for convenience the functional forms $u(x) = \ln x$ and $h(l) = \frac{l^{1+\gamma}}{1+\gamma}$. The parameter $\gamma > 0$ in the disutility of labor supply is the inverse of elasticity of labor supply, constant by assumption. Using $l_i = \frac{z_i}{w_i}$, we can re-parameterize agent's utility as

$$U(x_i, z_i, \theta_i) = \ln x_i - \theta_i \frac{z_i^{1+\gamma}}{1+\gamma}, \tag{2.1}$$

[3] We restrict our analysis to traditional couple since there is no clear labeling of spouses with the same gender when the model is taken to data.

where the parameter $\theta_i \equiv w_i^{-(1+\gamma)}$ is referred as type. Define $\underline{\theta} = \overline{w}^{-(1+\gamma)}$ and $\overline{\theta} = \underline{w}^{-(1+\gamma)}$ to write the type space as $\Theta = [\underline{\theta}, \overline{\theta}] \times [\underline{\theta}, \overline{\theta}]$.

Types are private information, but the joint distribution function of the random vector (θ_a, θ_b), denoted $F(\cdot)$, is common knowledge. We assume that $F(\cdot)$ is absolutely continuous with strictly positive density, $f(\cdot)$.

Tax system We assume that the tax basis for the government is total household income, denoted by $z \equiv z_a + z_b$. We are agnostic about the reasons for such choice, although we do acknowledge its existence to a rationale based on fairness between spouses in a couple as well as its empirical relevance.[4][5]

This restriction on the space of feasible mechanism will be convenient to reduce the numbers of margins that a government has to take into account in order to align incentives in a multidimensional environment.

The government with redistributive purposes taxes households using a non-linear joint income tax schedule, $T : Z \to \mathbb{R}$, assigning a tax liability for each possible level of family income. Interim Bayesian Nash equilibrium is our implementability concept.

2.2.1 Family's problem

In our simple environment, the problem of a family facing a tax schedule $T : Z \to \mathbb{R}$ is

$$\max_{z_a, z_b} \ \alpha \left[\ln x_a - \theta_a \frac{z_a^{1+\gamma}}{1+\gamma} \right] + (1-\alpha) \left[\ln x_b - \theta_b \frac{z_b^{1+\gamma}}{1+\gamma} \right], \tag{2.2}$$

subject to[6]

$$x = z - T(z),$$

$$x_a + x_b = x,$$

and

$$z = z_a + z_b.$$

[4]Indeed, this tax basis was chosen in many countries such as the US, Ireland and Germany.

[5]Cremer, Lozachmeur, and Pestieau (2012) provides conditions under which government should tax couples jointly.

[6]These three constraints can be easily summarized in just one:

$$x_a + x_b \leq z_a + z_b - T(z_a + z_b)$$

Families' disposable income $x = z - T(z)$ is used in the consumption of a private good allocated between spouses $x = x_a + x_b$ in an efficient way. This efficiency is captured by assigning female Pareto weights, $\alpha \in (0, 1)$, common to all households, constant and exogenously given. In Appendix A we show that we can separate the household consumption and labor supply collective decision. The following lemma summarizes the consequences of our assumptions about the household behavior.

Lemma 5 *The efficient solution to the household consumption allocation problem is* $x_a = \alpha x$ *and* $x_b = (1 - \alpha)x$.

Lemma 6 *Each spouses' earned income,* z_i *for* $i = a, b$*, can be expressed in terms of household's income* z*, productivity gap* θ_a/θ_b *and the Pareto coefficient* α*:*

$$z_a = z\left[\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}} + 1\right]^{-1},$$

and

$$z_b = z\left[\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}} + 1\right]^{-1}.$$

As the previous lemma shows, individual's income varies negatively with individual bargain power. One can interpret this result by thinking that joint household production depends on individuals' effort. This means greater bargain power will lead to an individual benefiting from joint household production, but the other spousal being relatively more responsible for labor supplying. Individual income also varies positively with spousals' productivity gap, which means that more productive spousals work relatively more, fixing all other primitives.

Using the previous results, we can rewrite the household welfare function in a simpler way.

Proposition 4 *The household welfare function for a type* (θ_a, θ_b) *family can be expressed as a function of total earned income,* z*, and disposable income* x*, as*

$$V(x, z, \theta_a, \theta_b, \alpha) = \kappa(\alpha) + \ln x - \omega(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma}$$

where $\kappa(\alpha) \equiv \alpha \ln(\alpha) + (1-\alpha)\ln(1-\alpha)$ *and*

$$\omega(\theta_a, \theta_b, \alpha) := \alpha\theta_a\left[1 + \left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)} + (1-\alpha)\theta_b\left[1 + \left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}.$$

Proof: See the appendix. ■

According to Proposition 4, the household objective can be restated in terms of the household aggregate consumption (disposable income) and aggregate income only.

Two features are worth noting in this household objective. First, family types only enter the objective through $\omega(\theta_a, \theta_b, \alpha)$. Therefore, families whose combination of $(\theta_a, \theta_b, \alpha)$ generates the same ω, share the same objective.[7] We can think of $\omega(\theta_a, \theta_b, \alpha)$ as a sufficient statistic for the family utility of a given allocation. Second, the part of the utility that depends on α not through ω is separable from the allocation and, hence, not affected by the tax policy.

Therefore, given our assumptions, the family's problem facing a tax schedule $T : Z \to \mathbb{R}$ is given by the maximization of

$$\kappa(\alpha) + \ln(x) - \omega(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma}, \tag{2.3}$$

subject to

$$x = z - T(z).$$

Proposition 4 allows us to think about the family's problem as if it were an individual's problem. The first-order necessary condition for the type ω family's problem is

$$\frac{1 - T'(z)}{z - T(z)} = \omega z^{\gamma}. \tag{2.4}$$

or equivalently, $1 - T'(z) = x\omega z^{\gamma}$.

2.3 Optimal taxation

The government taxes households using a non-linear joint income tax schedule $T : Z \to \mathbb{R}$ to maximize a utilitarian social welfare function.

2.3.1 Families utility

Assume that the planner adopts as welfare criterion the maximization of households' average utility. As Proposition 4 suggests, the planner will take household's utility into account, ignoring intra-household decision. As a consequence, families with the same type

[7] Total utility would vary if we allow α's to vary across households.

ω should get the same treatment and the mechanism design problem is analogous to the standard one solved in Mirrlees (1971b). Therefore, if the planner smoothes consumption across families under incentive compatibility constraints, the planner's problem is equivalent to smooth individual's utilities. Not surprisingly, the optimal non-linear tax schedule is the standard one, as shown in the next proposition.

Let $\psi(\cdot)$ be the probability density function of the random variable $\omega(\theta_a, \theta_b, \alpha)$ implied by the joint distribution of the random vector (θ_a, θ_b), denoted by $F(\cdot)$. Let $\Omega \equiv \omega(\Theta, \alpha)$ denote the set of possible values for ω. By the revelation principle, the planner's problem based solely on the family's utility can be restated in the space of direct mechanism as[8]

$$\max_{x,z:\Omega \to Z} \int_{\Omega} \left[\ln x(\omega) - \omega \frac{z(\omega)^{1+\gamma}}{1+\gamma} \right] \psi(\omega) d\omega, \tag{2.5}$$

subject to incentive compatibility constraints: for every $\omega \in \Omega$

$$\omega \in \arg\max_{\tilde{\omega} \in \Omega} \left[\ln x(\tilde{\omega}) - \omega \frac{z(\tilde{\omega})^{1+\gamma}}{1+\gamma} \right], \tag{2.6}$$

and a budget constraint:

$$\int_{\Omega} \left[z(\omega) - x(\omega) \right] \psi(\omega) d\omega \geq 0. \tag{2.7}$$

Proposition 5 *Suppose that the social welfare function is based on the household's utility. In an interval where $z(\omega)$ is strictly increasing, the marginal tax rate for the optimal non-linear income joint taxation will follows exactly as in Mirrlees' model:*

$$T'(z(\omega)) = \frac{z(\omega)^{\gamma}}{\psi(\omega)} \int_{\omega}^{\overline{\omega}} \left[x_v(v(s), z(s), s) - \frac{1}{\lambda} \right] \psi(s) ds \tag{2.8}$$

where $\lambda \in \mathbb{R}_+$ is the Lagrange multiplier associated with the budget constraint.

Proof: See the appendix. ■

Corollary 2 *Suppose that the social welfare function is based on the household's utility. The social value of income measured by the shadow price (Lagrange multiplier) of government budget constraint is equal to the inverse of mean consumption (disposable income):*

$$\lambda = \mathbb{E}^{\psi}[x]^{-1}. \tag{2.9}$$

[8]Since the term $\kappa(\alpha)$ is not affected by the allocation we can ignore it.

Using the structure of our model, we can rewrite the formula for the marginal taxation in a clearer way.

Corollary 3 *Suppose that the social welfare function is based on the household's utility. The optimal non-linear income joint taxation is given by:*

$$\frac{T'(z(\omega))}{1-T'(z(\omega))} = \frac{1}{x(\omega)\psi(\omega)\omega} \int_{\omega}^{\overline{\omega}} \left[x(s) - \mathbb{E}^{\psi}[x]\right] \psi(s) ds. \tag{2.10}$$

This is the usual Mirrlees' formula, which in the case of log preferences implies that the size of distortion increases with the mass of families more productive than the given family. In particular there is no distortion at the top because the mass of families above the highest type is null.

2.3.2 Individual's utility

Equation (2.3) represents type ω households' objective. Whether any welfare meaning should be attached to it is a different matter. As Chiappori and Meghir (2014) pointed out, any utility measure that does not take into account intra-household inequality creates an inconsistency with any utilitarian welfare criteria. In this sense, even restricted to tax schedules based on the total household income, the planner may still have an incentive to take into account individuals', instead of households', utilities for the tax design. In this section, we want to shed light on the implications of these considerations to optimal tax theory.

Let $V_a(x, z, \theta_a, \theta_b)$ be the indirect utility derived by spouse a in a couple (θ_a, θ_b) from the allocation (x, z) and $V_b(x, z, \theta_a, \theta_b)$ be analogously defined.

Proposition 6 *The indirect utility of a spouse in a family with total income z, disposable income x, types θ_a, θ_b and bargaining power α can be written as*

$$V_a(x, z, \theta_a, \theta_b) = \ln \alpha - \kappa(\alpha) + V(x, z, \theta_a, \theta_b) + [\omega - \omega_a] \frac{z^{1+\gamma}}{1+\gamma},$$

and

$$V_b(x, z, \theta_a, \theta_b) = \ln(1-\alpha) - \kappa(\alpha) + V(x, z, \theta_a, \theta_b) + [\omega - \omega_b] \frac{z^{1+\gamma}}{1+\gamma},$$

where $V(x, z, \theta_a, \theta_b)$, $\kappa(\alpha)$ and $\omega(\theta_a, \theta_b, \alpha)$ are the same as in Proposition 4, and

$$\omega_a(\theta_a, \theta_b, \alpha) \equiv \theta_a \left[1 + \left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)},$$

and

$$\omega_b(\theta_a, \theta_b, \alpha) \equiv \theta_b \left[1 + \left(\frac{1-\alpha}{\alpha} \frac{\theta_b}{\theta_a} \right)^{\frac{1}{\gamma}} \right]^{-(1+\gamma)}.$$

are the "adjusted types" of spouses a and b.

Proof: See the appendix. ■

Adjusted types are synthetic parameters that captures individuals' productivities, bargain power coefficients and wage gap. Adjusted types vary negatively with individuals' productivities and bargain power coefficients. It varies positively with the productivity gap between spousals.

Corollary 4 *The family type is a mean of the adjusted types with weights given by the bargaining power:*

$$\omega := \alpha\omega_a + (1-\alpha)\omega_b.$$

These adjusted types represent the effective type after considering the household dynamics in a reduced form utility derived from the consumption of the bundle (z, x). From Proposition 6, we can see that the higher is the bargaining power α and the productivity gap $\frac{\theta_a}{\theta_b}$ the higher will be the adjusted type of spouse a. For spouses b we have the opposite effect: the higher is the bargaining power α and the productivity gap $\frac{\theta_a}{\theta_b}$ the lower will be the adjusted type of spouse a.

Recall that $f(\theta_a, \theta_b)$ is probability density function of the random vector (θ_a, θ_b) and $\psi(\omega)$ is probability density function of the random variable $\omega(\theta_a, \theta_b, \alpha)$. We can equivalently work with the random vector (ω, ω_a) whose distribution is implied by $F(\theta_a, \theta_b)$, denoted by $\Psi(\omega, \omega_a)$. We can factor $\Psi(\omega, \omega_a)$ as

$$\Psi(\omega, \omega_a) = \phi(\omega_a|\omega)\psi(\omega). \tag{2.11}$$

Let us denote $\Omega \equiv \omega(\Theta, \alpha)$ the set of possible values for ω and $\Omega_a \equiv \omega_a(\Theta, \alpha)$ the set of possible values for ω_a and $\underline{\omega}$,$\bar{\omega}$,$\underline{\omega}_a$,$\bar{\omega}_a$ the inf and sup values of these sets, respectively.

In Appendix B, we show how to derive the social welfare criteria of a utilitarian social planner that takes into account individuals' utilities. We present the result in the following lemma.

Lemma 7 *The social welfare criteria of a utilitarian social planner that takes in account*

individuals' utilities and weights equally spouses within the family is

$$\int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega}_a}^{\overline{\omega}_a} \left[W(x,z,\omega,\omega_a) \right] \phi(\omega_a|\omega)\psi(\omega) d\omega_a d\omega, \tag{2.12}$$

where

$$W(x,z,\omega,\omega_a) = \left[\frac{1}{2} - \alpha\right] \ln \frac{\alpha}{1-\alpha} + V(x,z,\omega) + \left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a - \omega}{1-\alpha}\right)\frac{z^{1+\gamma}}{1+\gamma} \tag{2.13}$$

written in terms of the adjusted type of the family ω and the adjusted type of spouse a, ω_a.

The function $W : \mathbb{R}_+ \times \mathbb{R} \times \Omega \times \Omega_a \to \mathbb{R}$ represents the utility derived by the planner from offering the bundle (x,z) in an allocation to family with adjusted types ω and ω_a. The dissonance between family's utility and planner's utility is exacerbated with higher adjusted types discrepancy and bargaining power. As expected, if $\alpha = \frac{1}{2}$, there is no tension between spouses in the household's collective decision, and consequently the planner's utility becomes the family's utility function. Note also that for couples whose spouses have the same adjusted type the household decision process only affects the planner welfare in level.

The last term internalizes how the collective decision process affects the decision margins of the individuals vis-a-vis the family's decision and it is the only term affected by the taxation policy.

An interesting parallel can be made between our model and the theory of taxation of behavioral agents in Farhi and Gabaix (2015). In the government's perspective, the utility derived by the family is distorted, thus generating a discrepancy between the decision utility and the experienced utility (in the planner's viewpoint). Therefore, we can interpret a couple as a behavioral agent in our model.

Given our formulation it is convenient to write the problem as choosing an utility level for the family $v(\omega)$, an income level $z(\omega,\omega_a)$ depending on both adjusted types and letting the consumption be implicitly defined by $x(v(\omega), z(\omega,\omega_a), \omega, \omega_a)$. The planner's problem determines $v : \Omega \to \mathbb{R}$ and $z : \Omega \times \Omega_a \to \mathbb{R}$ to solve:

$$\max_{v(\cdot),z(\cdot)} \int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega}_a}^{\overline{\omega}_a} \left[v(\omega) + \left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a - \omega}{1-\alpha}\right)\frac{z(\omega)^{1+\gamma}}{1+\gamma} \right] \phi(\omega_a|\omega)\psi(\omega) d\omega_a d\omega. \tag{2.14}$$

subject to incentive compatibility constraints:

$$(\omega,\omega_a) \in \arg\max_{\tilde{\omega}\in\Omega, \tilde{\omega}_a \in \Omega_a} \left[\ln x(\tilde{\omega},\tilde{\omega}_a) - \omega \frac{z(\tilde{\omega},\tilde{\omega}_a)^{1+\gamma}}{1+\gamma} \right], \tag{2.15}$$

for every $\omega \in \Omega$ and $\omega_a \in \Omega_a$, and budget constraint:

$$\int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega}_a}^{\overline{\omega}_a} \left[z(\omega, \omega_a) - x(v(\omega), z(\omega, \omega_a), \omega, \omega_a)\right] \phi(\omega_a|\omega)\psi(\omega) d\omega_a d\omega \geq 0. \tag{2.16}$$

The assumption of iso-elastic disutility of labor allows us to disentangle the dissonance between the collective behavior in the family and the individualistic preferences of the government. In fact, by Proposition 4 the family type ω is a sufficient statistics for the welfare of the family. By the revelation principle, all families with the same type choose the same allocation when facing a given tax schedule. Hence, despite the dissonance in the planner's and families' welfare, any incentive compatible allocation should give the same allocation for the families with the same type. Hence, without loss, we can focus on direct mechanism $v : \Omega \to \mathbb{R}$ and $z : \Omega \to \mathbb{R}$. The following lemma states the necessary and sufficient conditions for implementability.

Lemma 8 *Suppose that the social welfare function is based on the individual's utility. The set of implementable allocations is characterized by the local incentive conditions, i.e., $v : \Omega \to \mathbb{R}$ and $z : \Omega \to \mathbb{R}$ piecewise differentiable is an implementable allocation if and only if*

i) For each $\omega \in \Omega$, $\dot{v}(\omega) = z(\omega)^{1+\gamma}/(1+\gamma)$;

ii) $z(\omega)$ is increasing in ω.

The next proposition characterizes the optimal taxation.

Proposition 7 *Suppose that the social welfare function is based on individual's utility. In all intervals where $z(\omega)$ is strictly increasing, the marginal tax rate for the optimal (non-linear) income joint taxation is*

$$T'\left(z(\omega)\right) = \frac{z(\omega)^\gamma}{\psi(\omega)} \int_{\omega}^{\overline{\omega}} \left[x_v(v(s), z(s), s) - \frac{1}{\lambda}\right] \psi(s) ds - \\ \frac{z(\omega)^\gamma}{\lambda} \left(\alpha - \frac{1}{2}\right) \left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right), \tag{2.17}$$

where $\lambda \in \mathbb{R}_+$ is the Lagrange multiplier associated with the budget constraint.

Proof: See the appendix. ■

Corollary 5 *Suppose that the social welfare function is based on the individuals' utility. The social value of income measured by the shadow price (the Lagrange multiplier) of*

government budget constraint is equal to the inverse of mean consumption (disposable income):

$$\lambda = \mathbb{E}^{\psi}[x]^{-1}. \tag{2.18}$$

Using the structure of our model, we can rewrite the formula for marginal taxation in a clearer way.

Corollary 6 *Suppose that the social welfare function is based on the household's utility. If $z(\cdot)$ is strictly increasing at ω, the marginal tax rate faced by type ω households is*

$$\frac{T'\left(z(\omega)\right)}{1-T'\left(z(\omega)\right)} = \frac{1}{x(\omega)\psi(\omega)\omega}\int_{\omega}^{\overline{\omega}}\left[x(s)-\mathbb{E}^{\psi}[x]\right]\psi(s)ds - \frac{\mathbb{E}^{\psi}[x]}{x(\omega)\omega}\left(\alpha-\frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega]-\omega}{1-\alpha}\right). \tag{2.19}$$

The last term in the formula for the optimal joint taxation represents the additional distortion that the government promotes, due to the dissonance of its evaluation of the social welfare and the families' utility.

The analysis of the limit cases leads to some interesting results. First, note that when spouses have the same power within the household (i.e., $\alpha = \frac{1}{2}$), we are back to the standard case where the planner's welfare is identical to the family's one. In this case, there is no disagreement between the individual and family interests. In particular, we can treat the family as a representative agent.

Additionally, when spouse a has the highest bargaining power (i.e., $\alpha = 1$), $\omega = \omega_a$ and, consequently, $\mathbb{E}[\omega_a|\omega] = \omega$. We are then back to the standard case where the planner's welfare is identical to the family's. The intuition leans on the fact that the family's utility is equal to the spouse a's utility and the government cannot do any better than follow the dominant spouse with an unbounded welfare cost.

When the most productive spouse has lower bargaining power (i.e. $\omega_a < \omega$ and $\alpha < \frac{1}{2}$; or $\omega_a > \omega$ and $\alpha > \frac{1}{2}$), the optimal marginal tax is smaller than the one in the standard case. The highest skilled spouse's disutility of labor supply is not very relevant on the family's problem since it has a minor impact on the family's welfare. Moreover, it is smaller than the low-skilled spouse's disutility of labor supply.

A smaller marginal tax provides a higher incentive for households to increase their income and, by consequence, incentives productive types to work more. There is also a welfare redistribution from high-skilled to low-skilled workers within the household. In other words, the household bargaining plays the redistributive role previously played by

the taxation mechanism.

On the other hand, when the most productive spouse has higher bargaining power, the family's income depends more on higher skilled private interest. Therefore, more progressive tax schedules are redistributed from families with more productive chiefs to families with less productive chiefs. Thus, the optimal marginal tax increases when compared to the one derived in the standard case.

2.3.3 Combining joint and individual instruments

In this section, we analyze the welfare impact of a tax reform that introduces extra tax instruments besides the joint taxation. The goal is to access how restrictive is the use of joint income as the tax basis.

Let individual taxes (or subsidies) allowed to be enforced to household participants. For instance, let us say that spouses a are the ones subjected to this individual instrument imposing an income tax with constant rate $t \in \mathbb{R}$ that can be either positive or negative. In our simple collective environment with this additional tax instrument, the problem of the family facing tax schedule $T : Z \to \mathbb{R}$ and income tax $t \in \mathbb{R}$ is given by:

$$\max_{z_a, z_b} \ \alpha \left[\ln x_a - \theta_a \frac{z_a^{1+\gamma}}{1+\gamma} \right] + (1-\alpha) \left[\ln x_b - \theta_b \frac{z_b^{1+\gamma}}{1+\gamma} \right] \tag{2.20}$$

subject to[9]

$$x = z - T(z),$$

$$x_a + x_b = x,$$

and

$$z = z_a(1-t) + z_b.$$

Keeping the assumptions of the collective approach of household behavior and analogously to Proposition 4, we can write the family's welfare in a simple way:

Proposition 8 *The family social welfare function of a family with total income z, dis-*

[9]These three constraints can be easily summarized in just one as

$$x_a + x_b \leq z_a(1-t) + z_b - T(z_a(1-t) + z_b)$$

posable income x, types θ_a, θ_b and bargaining power α can be written as

$$V(x,z,\theta_a,\theta_b,\alpha,t)=\kappa(\alpha)+\ln x-\omega(\theta_a,\theta_b,\alpha,t)\frac{z^{1+\gamma}}{1+\gamma} \tag{2.21}$$

where $\kappa(\alpha)\equiv\alpha\ln(\alpha)+(1-\alpha)\ln(1-\alpha)$ is fully determined by α hence exogenous and

$$\omega(\theta_a,\theta_b,\alpha,t):=\alpha\theta_a\left[1+\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\frac{1}{1-t}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}+ \\ (1-\alpha)\theta_b\left[1+\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}(1-t)\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}$$

is a "family type" that is function of spouses' types and the bargaining power and the tax rate.

Proof: See the appendix. ■

Analogously to the previous case, we have the same expression for the family utility except that the family type now depends also on t $\omega(\theta_a,\theta_b,\alpha,t)$. Therefore, the planner's problem can be rewritten in the same way.

Tax authorities can promote welfare improvement combining individual and joint taxes. The following proposition gives conditions under which these improvements are possible.

Proposition 9 *If $\alpha\theta_a<(1-\alpha)\theta_b$, subsidizing spouses a leads to a welfare improvement. On the other hand, if $\alpha\theta_a>(1-\alpha)\theta_b$, taxing spouses a leads to a welfare improvement.*

Proof: See the appendix. ■

2.4 Simulations

In this section we simulate the optimal marginal tax rate derived in Propositions 5 and 7 using real data from the 2016 March CPS to have a clearer assessment of the optimal taxation and the impacts of dissonance. We use a sample of married couples which are the primary family in the household, and whose income is strictly positive, using wages as a proxy for types.[10] In Appendix C, we have a comprehensive exposition of the implementation strategy based in a discrete grid of the income distribution adapted from Mankiw, Weinzierl, and Yagan (2009) to our model.

[10] We will consider families where one of the spouses earns no income when implementing the extension in Section 5 allowing for participation decisions.

Note that the marginal tax rate in Proposition 7 depends on the distribution of $\omega, \psi(\omega)$ and on the distribution of ω_a, conditional on $\omega, \phi(\omega_a \mid \omega)$. However, we can calculate the implied distribution of $\omega, \psi(\omega)$ from the knowledge of the empirical wage distribution.

Figure 1 shows information about the relation of family and female income in our sample while Figure 2 repeats the same for adjusted types for different levels of α. Both present the scatter plot of the families in our sample along with the estimate of the conditional expectation implemented non-parametrically using a kernel regression for three levels of α.

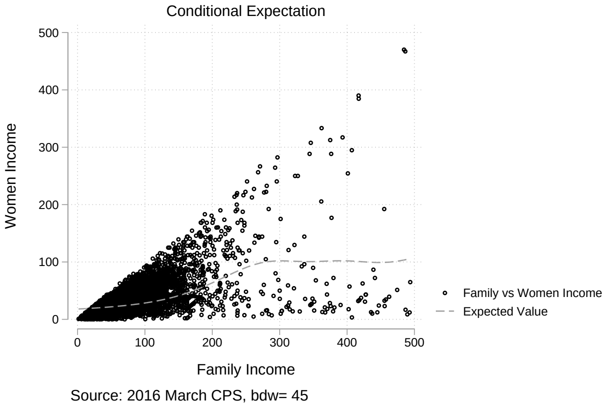
Recall the formula for the optimal marginal tax rate in Proposition 7. The last term on the right-hand side of this formula is the additional distortion that the planner imposes to influence the intra-household decision process due to the existence of dissonance. It depends heavily on the relation of the parameter α and the conditional expectation $\mathbb{E}[\omega_a|\omega]$.

First, notice that the three panels in Figure 1 are exactly the same, since this is the raw income data used as proxy for types and does not depend on α. We keep all three to facilitate visual comparisons. In Figure 2, we can see how the adjusted types change with α. The intuition for that is that types are adjusted to compensate for the bargaining power within the family. For instance, the adjusted types work as if we boost types for spouses with less bargaining power in the reduced form utility. Comparing the three panels in Figure 2, we see that the impact of collective bargaining can be empirically relevant.

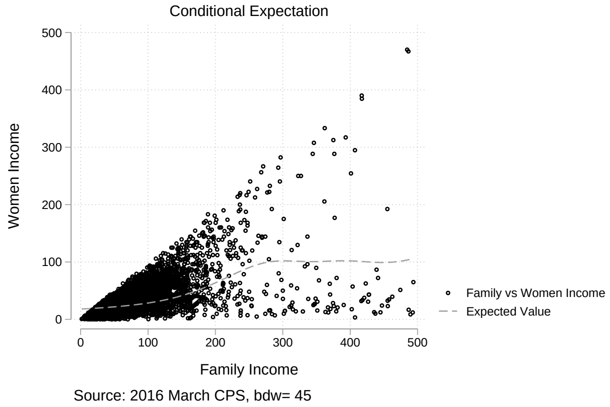
In Figure 3, we see how the dissonance affects the shape of optimal marginal tax rate. In fact, when women have low bargaining power $\alpha = .3$, the tax system becomes very regressive comparing to the benchmark case of equal weights and no dissonance. When women have high bargaining power $\alpha = .7$, the changing goes in the other direction.

It is important to note that these effects are heavily dependent on the degree of assortative matching. In these simulations we are implicitly using the degree inherent to our sample. In a close future, we intend to make comparative statics of the optimal taxation with respect to the assortative matching. [11]

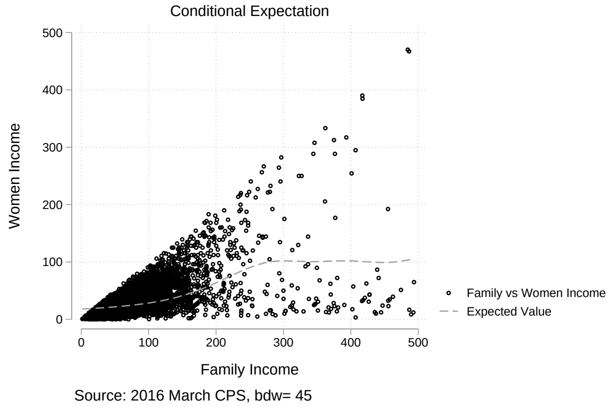
[11] In Figure 3 panels (b) and (c) the tax rate is not smooth; this is due to a numeric incompatibility of our methodology when some bin in our grid for the income range does not have any observations in our sample. In the simulations for the case $\alpha = .3$, we manually interpolate the data in this bins. This was a very time-consuming process and we only did for the empirically most relevant case, according to our opinion.)



(a) $\alpha = .3$

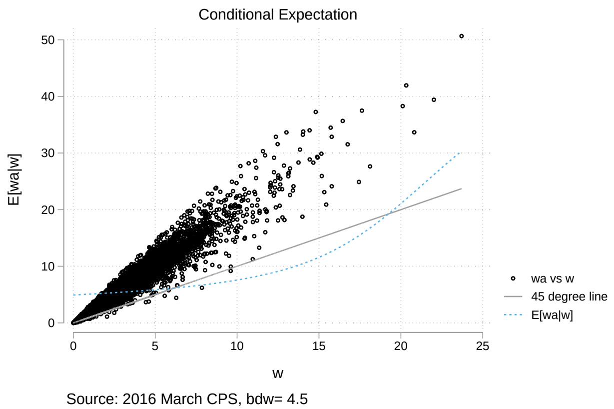


(b) $\alpha = .5$

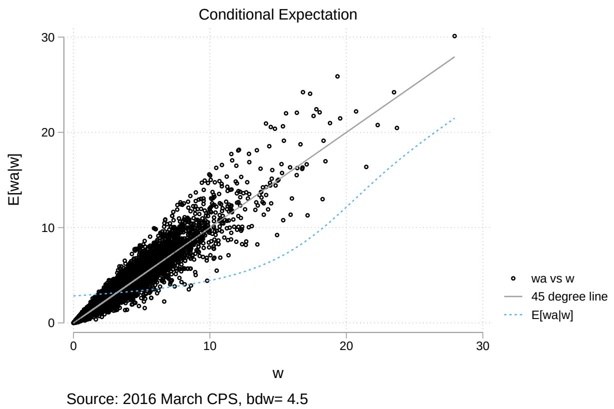


(c) $\alpha = .7$

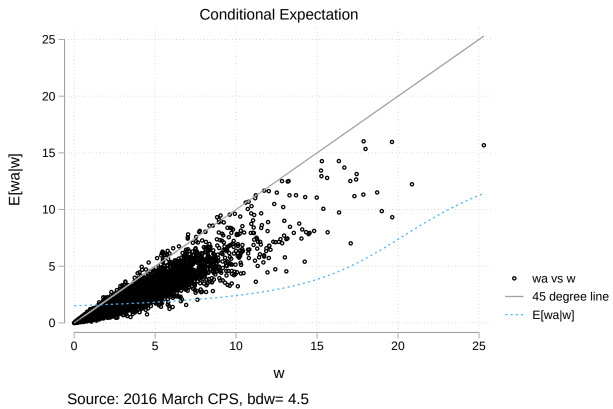
Figure 2.1: Expected value of $\mathbb{E}[z_a|z]$



(a) $\alpha = .3$

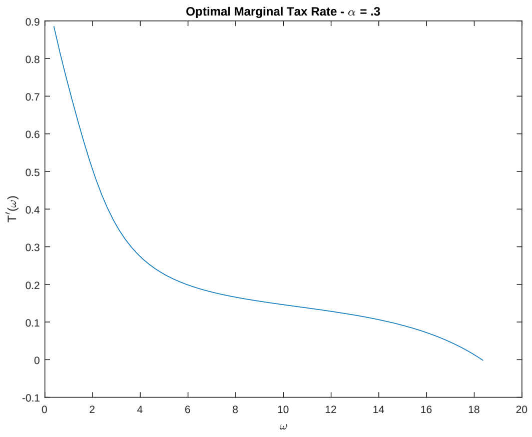


(b) $\alpha = .5$

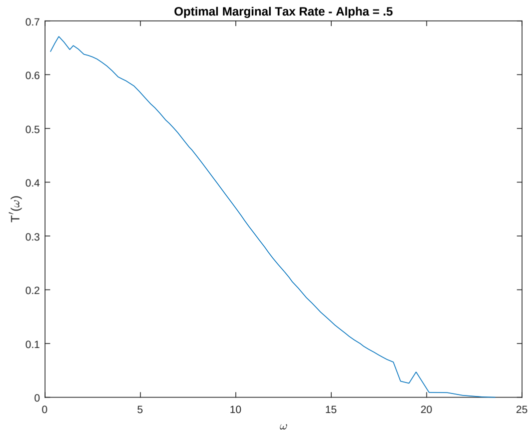


(c) $\alpha = .7$

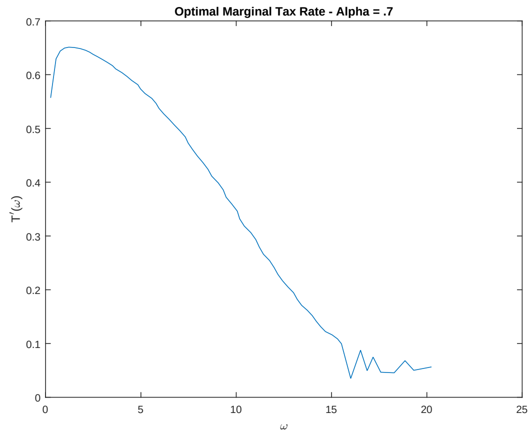
Figure 2.2: Expected value of $\mathbb{E}[\omega_a|\omega]$



(a) $\alpha = .3$



(b) $\alpha = .5$



(c) $\alpha = .7$

Figure 2.3: Optimal Marginal Tax Rate

Note that figure 2.1 is invariant with α because at this first stage we are just regressing household and wives income. As in Mankiw, Weinzierl, and Yagan (2009) we consider wages as a proxy for productivities. So it is important to understand the relation between spousals' and household's productivities.

In figure 2.2 we can see that as α increases, the $\mathbb{E}[\omega_a|\omega]$ decreases. One interpretation is that for a given family type, the greater is woman's bargain power, her welfare relies more on that power and less on her original productivity. So we expect decreasing woman's adjusted types.

As in lemma 6 individuals' income decrease with their bargain coefficient. Whenever an individual becomes more powerful within the household, she is more suitable to work less and rely more on the spousal's effort.

2.5 Labor market participation

When considering the labor supply of married women, the participation choice is an important margin. In this section we allow individuals to make both extensive and intensive margin labor supply decisions, by deciding whether or not to participate in the labor market, and how much effort to make in case of participation. To incorporate extensive margin decisions, we assume that spouses have different costs of participating in the market. Given that the majority of individuals who leave the labor force are women, we assume that there is an opportunity cost A for women to participate in the labor market. Moreover, in this section, we hold α constant across households.

We allow the government to design different tax schedules: $T_1(\cdot)$, for households with one earner, and $T_2(\cdot)$, for households with two earners. Equivalently, we consider two different allocations $(v_2(\omega), z_2(\omega))$, for households with two earners, and $(v_1(\theta_b), z(\theta_b))$, for household in which only the husband works.

It will be therefore convenient to redefine the problem in terms of (θ_b, ω) types. That is, we define a couple by the husband's productivity θ_b and the household's productivity ω. Letting the support of the joint distribution of (θ_a, θ_b) to be $\Theta \times \Theta$, with $\Theta = (0, \infty)$, the same is the support for the joint distribution of (θ_b, ω) is $\Theta \times \tilde{\Theta}$ where $\tilde{\Theta}(\theta_b) := \{\omega \in \mathbb{R}_+ | \omega < (1-\alpha)\theta_b\}$ and $\tilde{\Theta} = \cup_{\theta_b \in \Theta} \tilde{\Theta}(\theta_b)$.

We use $F_{\omega|b}(\omega|\theta_b)$ to denote the conditional distribution of ω, given θ_b.[12] Using our

[12] If $F_{a|b}(\cdot|\theta_b)$ is the conditional distribution of θ_a given θ_b, then

$$F_{\omega|b}(\omega|\theta_b) = F_{a|b}(\{\omega^{-1/\gamma} - ((1-\alpha)\theta_b)^{-1/\gamma}\}^{-\gamma}\alpha^{-1}|\theta_b).$$

definition of ω, a (θ_a, θ_b)-couple with earned income z attains utility

$$V(x, z, \omega) = \kappa(\alpha) + \ln x - \omega \frac{z^{1+\gamma}}{1+\gamma} - \alpha A,$$

where $\kappa(\alpha) \equiv \alpha \ln(\alpha) + (1-\alpha)\ln(1-\alpha)$, if both spouses work, and

$$V_b(x, z, \theta_b) = \kappa(\alpha) + \ln x - (1-\alpha)\theta_b \frac{z^{1+\gamma}}{1+\gamma},$$

if only the husband works.

Given the tax schedules $T_1(\cdot)$ and $T_2(\cdot)$, the family's optimization problem can be written as

$$\max \left\{ \max_{z_2} \left\{ \ln(z_2 - T_2(z_2)) - \omega \frac{z_2^{1+\gamma}}{1+\gamma} \right\} - \alpha A, \right. \\ \left. \max_{z_1} \left\{ \ln(z_1 - T_1(z_1)) - (1-\alpha)\theta_b \frac{z_1^{1+\gamma}}{1+\gamma} \right\} \right\}$$

Define the extensive margin choice rule $\mathcal{C} : \tilde{\Theta} \times \Theta_b \mapsto \{1, 2\}$ which specifies to every couple (ω, θ_b) whether there should be one or two spouses working. The first thing we show is that for any pair of schedules $T_1(\cdot)$ and $T_2(\cdot)$ this choice rule takes the form of an increasing function $\omega^\circ = c(\theta_b)$ which establishes for each θ_b a value ω° such that every θ_b couple with $\omega < \omega^\circ$ will have both spouses working and for every couple with $\omega \geq \omega^\circ$ only the husband will work.

Lemma 9 *Given $T_1(\cdot)$ and $T_2(\cdot)$ there exists an increasing threshold function $c(\theta_b)$, $\mathcal{C}(\omega, \theta_b) = 1$ if $\omega \geq c(\theta_b)$ and $\mathcal{C}(\omega, \theta_b) = 2$, otherwise.*

Proof: Let ω and θ_b be such that $v_2(\omega) - \alpha A = v_1(\theta_b)$. For any $\omega' < \omega$ that shares the same θ_b we have

$$v_2(\omega') \geq \ln(x_2(\omega)) - \omega' \frac{z_2(\omega)^{1+\gamma}}{1+\gamma} > v_2(\omega) = v_1(\theta_b) + \alpha A.$$

Hence, ω'-couple decides that both spouses should work. If we consider, instead, a couple with the same θ_b but $\omega < \omega''$ then we can show that $v_2(\omega) > v_2(\omega'')$, which implies that couple ω'' decides that only the husband must work.

Next, for an arbitrary θ_b let ω be such that $v_2(\omega) - v_1(\theta_b) = \alpha A$. Now consider slightly increasing θ_b, thus reducing v_1 while preserving the difference constant at αA. Because $v_2(\cdot)$ is decreasing in ω, this requires also increase in ω. ∎

For every $\theta_b \in \Theta$ define the set $\tilde{\Theta}(\theta_b)$ through

$$\tilde{\Theta}(\theta_b) := \left\{ \omega \in \mathbb{R}_+ \middle| \exists \theta_a \in \Theta \text{ such that } \omega = \{(\alpha\theta_a)^{-1/\gamma} + ((1-\alpha)\theta_a)^{-1/\gamma}\}^{-\gamma} \right\}$$

Although different possibilities regarding $c(\cdot)$ may arise at the optimum, we shall focus on the case illustrated by Figure 2.4 in which, for all ω, there is a value θ_a' such that $\mathcal{C}(\theta_a, \theta_b) = 1$. When this is the case, we establish the solution for the second step of the program.[13]

For any $\theta_b \in \Theta$, we define

$$\mathcal{G}_1(\theta_b|c) = \int_0^{\theta_b} \int_{c(\tilde{\theta}_b)}^{(1-\alpha)\tilde{\theta}_b} f(\omega, \tilde{\theta}_b) d\omega d\tilde{\theta}_b = \int_0^{\theta_b} \left[1 - F_{\omega|b}(c(\tilde{\theta}_b)|\tilde{\theta}_b)\right] f_b(\tilde{\theta}_b) d\tilde{\theta}_b$$

Lemma 10 in the Appendix shows that the way we define the intervals of integration is without loss.[14] We may also define

$$\mathcal{G}_2(\omega|c) := \int_0^{\omega} \int_{c^{-1}(\tilde{\omega})}^{\infty} f(\omega, \tilde{\theta}_b) d\tilde{\theta}_b d\omega =$$

$$\int_0^{\omega} \left[1 - F_{\theta_b|\omega}(c^{-1}(\tilde{\omega})|\tilde{\omega})\right] f_\omega(\tilde{\omega}) d\tilde{\omega},$$

with associated density $\left[1 - F_{\theta_b|\omega}(c^{-1}(\tilde{\omega})|\tilde{\omega})\right] f_\omega(\tilde{\omega})$.

For our purposes it will be important to further note that from $\mathcal{G}_2(c(\theta_b)|c)$, the density $\}_2(\theta_b|c) := c'(\theta_b)\left[1 - F_{\theta_b|\omega}(\theta_b|c(\theta_b))\right] f_\omega(c(\theta_b))$ is defined.

In order to characterize $c(\cdot)$, write

$$v_2(\omega) := \max_{z_1} \left\{ \ln(z_2 - T_2(z_2)) - \omega \frac{z_2^{1+\gamma}}{1+\gamma} \right\} - \alpha A,$$

and

$$v_1(\theta_b) := \max_{z_1} \left\{ \ln(z_1 - T_1(z_1)) - (1-\alpha)\theta_b \frac{z_1^{1+\gamma}}{1+\gamma} \right\}.$$

In this case, $v_2(c(\theta_b)) = v_1(c(\theta_b))$, for all θ_b, implicitly defines function $c(\cdot)$. Now,

[13] Another possibility is that for some subset(s) of Θ, all couples will have both spouses working at the optimum. In this case, it is more convenient to work with the program defined as a function of ω instead of θ_b.

[14] Modulo the alternative configuration for which it is the set of couples for which both agents work that contains all possible types θ_b.

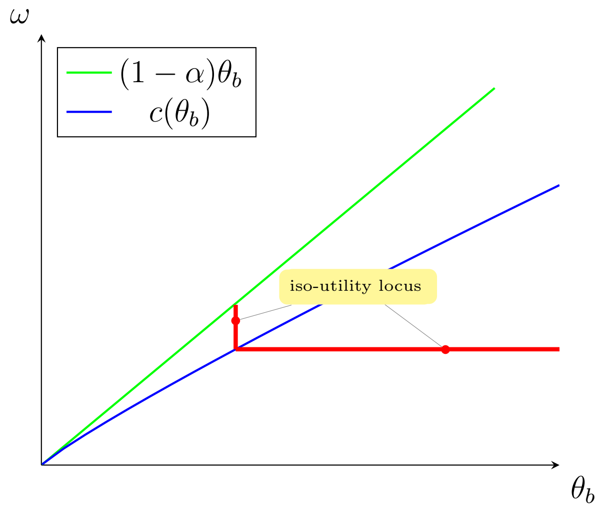


Figure 2.4: **Participation decision** The figure shows the support for the joint distribution of ω and θ_b along with the threshold function, $c(\cdot)$. In red we see the set o agents which share the same flow utility, $v_1(\theta_b) = v_2(c(\theta_b)) - \alpha A$.

differentiating this equation we obtain

$$\dot{v}_1(\theta_b) = \dot{v}_2(c(\theta_b))c'(\theta_b),$$

which, given the envelope condition

$$\dot{v}_1(\theta_b) = -(1-\alpha)\frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma}, \tag{2.22}$$

becomes

$$(1-\alpha)^{\frac{1}{1+\gamma}} z_1(\theta_b) = z_2(c(\theta_b))c'(\theta_b)^{\frac{1}{1+\gamma}}. \tag{2.23}$$

Following Gomes, Lozachmeur, and Pavan (2014) it is now possible to solve the planner's program in two steps. First, take $c(\cdot)$ as given and solve the constrained planner's program. This will give the best allocation among those that induce $c(\cdot)$. In the second step, we choose the optimal $c(\cdot)$.

Note that the restriction $v_1(\theta_b) = v_2(c(\theta_b))$, implies that, for a given $c(\cdot)$, the planner's objective can be simply written as

$$\max \int_{\underline{\theta}_b}^{\bar{\theta}_b} v_1(\theta_b) \}(\theta_b|c) d\theta_b, \tag{2.24}$$

where

$$\}(\theta_b|c) = F_{\theta_b|\omega}(c(\theta_b)|\theta_b)\phi_b(\theta_b) + c'(\theta_b)F_{\theta_b|\omega}(\theta_b|c(\theta_b))\phi_\omega(c(\theta_b)).$$

As for the resource constraint, note that

$$x_1(\theta_b) = \exp\left\{v_1(\theta_b) + (1-\alpha)\theta_b \frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma}\right\}.$$

Next,

$$v_2(c(\theta_b)) = \ln x_2(c(\theta_b)) - c(\theta_b)\frac{z_2(c(\theta_b))^{1+\gamma}}{1+\gamma}$$

implies

$$v_2(c(\theta_b)) = \ln x_2(c(\theta_b)) - (1-\alpha)\frac{c(\theta_b)}{c'(\theta_b)}\frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma},$$

which gives

$$x_2(c(\theta_b)) = \exp\left\{v_1(\theta_b) + \frac{c(\theta_b)}{c'(\theta_b)}(1-\alpha)\frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma}\right\}.$$

Hence, the resource constraint reduces to

$$\begin{aligned}&\int\left\{z_1(\theta_b) - \exp\left\{v_1(\theta_b) + (1-\alpha)\theta_b\frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma}\right\}\right\}\}_1(\theta_b|c)d\theta_b + \\ &\int\left\{z_1(\theta_b)\left(\frac{1-\alpha}{c'(\theta_b)}\right)^{\frac{1}{1+\gamma}} - \exp\left\{v_1(\theta_b) + (1-\alpha)\frac{c(\theta_b)}{c'(\theta_b)}\frac{z_1(\theta_b)^{1+\gamma}}{1+\gamma}\right\}\right\}\}_2(\theta_b|c)d\theta_b \geq B,\end{aligned} \tag{2.25}$$

where

$$\}_1(\theta_b|c) := \left[1 - F_{\omega|b}(c(\theta_b)|\theta_b)\right] f_b(\theta_b),$$

and

$$\}_2(\theta_b|c) := c'(\theta_b)\left[1 - F_{\theta_b|\omega}(\theta_b|c(\theta_b))\right] f_\omega(c(\theta_b)).$$

Thus, the planner's program is to maximize (2.24) subject to (2.25) and the intensive margin incentive constraint, comprised of the envelope condition (2.22), and the monotonicity condition, $z_1(\theta_b)$ decreasing.

As usual, we ignore the monotonicity constraint and verify ex-post its validity. Solving the relaxed versions of the problem above is as simple as solving a usual Mirrlees'

program. The first-order condition with respect to $z_1(\theta_b)$ yields

$$\frac{-\mu(\theta_b)}{\lambda}(1-\alpha)z_1(\theta_b)^{\gamma+1} = T_1'(z_1(\theta_b))z_1(\theta_b)\}_1(\theta_b|c)+ \\ T_2'(z_2(c_2(\theta_b)))z_2(c(\theta_b))\}_2(\theta_b|c), \quad (2.26)$$

where $\mu(\cdot)$ is the Lagrange multiplier associated with the intensive margin IC constraint (2.22) and λ is the Lagrange multiplier associated with the resource constraint (2.25).

The first-order condition with respect to $v_1(\theta_b)$ allows us to calculate

$$\lambda = \left\{\int_0^{\infty}\left\{x_1(\theta_b)\}_1(\theta_b|c) + x_2(c(\theta_b))\}_2(\theta_b|c)\right\}d\theta_b\right\}^{-1} \quad (2.27)$$

and

$$\mu(\theta_b) = \int_0^{\theta_b}\left\{[1-\lambda x_1(a)]\}_1(a|c) + [1-\lambda x_2(c(a))]\}_2(a|c)\right\}da. \quad (2.28)$$

When compared to a typical Mirrlees' program, the only difference is that it is the average (between the two schedules) of marginal tax rates that matter for incentive provision in the intensive margin at each level of θ_b.

As for the second step in the characterization procedure, assume that we know the optimal $(z_1(\theta_b))_{\theta_b}$ allocation, $(z_1^\star(\theta_b))_{\theta_b}$. Given, $(z_1^\star(\cdot))$, the problem is now to find a continuous, strictly increasing (over $(\theta_b', \bar{\theta}_b)$) function such that $c(\theta_b) = 0$, for all $\theta_b < \theta_b'$, that maximizes the program above.

Assume that we have already chosen θ_b'. Given $v_1(\theta_b')$, then $z_1^\star(\cdot)$ pins down the whole path for $v_1(\theta_b)$ through (2.22) and the boundary condition.

Define $J(c(\theta_b), c'(\theta_b), \theta_b)$ through

$$J(c(\theta_b), c'(\theta_b), \theta_b) := v_1^\star(\theta_b)\}(\theta_b|c)+ \\ \lambda\left\{z_1^\star(\theta_b) - \exp\left\{v_1^\star(\theta_b) + \theta_b\frac{z_1^\star(\theta_b)^{1+\gamma}}{1+\gamma}\right\}\right\}\}_1(\theta_b|c)+ \\ \lambda\left\{z_1^\star(\theta_b)\left(\frac{1-\alpha}{c'(\theta_b)}\right)^{\frac{1}{1+\gamma}} - \exp\left\{v_1^\star(\theta_b) + (1-\alpha)\frac{c(\theta_b)}{c'(\theta_b)}\frac{z_1^\star(\theta_b)^{1+\gamma}}{1+\gamma}\right\}\right\}\}_2(\theta_b|c),$$

This is a calculus of variation problem which solution is characterized by the Euler equation

$$\frac{d}{d\theta_b}\frac{\partial J(c(\theta_b), c'(\theta_b), \theta_b)}{\partial c'(\theta_b)} = \frac{\partial J(c(\theta_b), c'(\theta_b), \theta_b)}{\partial c(\theta_b)}.$$

2.6 Conclusion

In this paper, we analyze taxation of the joint income of couples considering a collective model for household decision. One of the main novelties of our paper is the possibility of tackling incentive compatibility constraints without a further restriction on agents decision space. More specifically by using isoelastic utilities we maintaining full generality of spouses intensive margin labor decisions.

The other novelty of our paper is the analysis of the impact of dissonance between the planner and household welfare when the planner considers individuals utility in his social welfare criteria. We perform simulation using data from the American economy finding a strong influence of the bargaining power within the household in the shape of the optimal joint income tax.

Following up, we intend to take this calibrated model to perform counterfactual comparative statics exercises varying the degree of assortative matching. Parallelly we plan to expand the analysis of the participation decisions which still is in an embryonary stage.

Appendix

Appendix A - Collective decision

The consumption allocation problem within the household

The disposable income $x = z - T(z)$ is allocated to the consumption of private goods of the two spouses according to efficient Nash bargaining protocol. Under this protocol the within household consumption allocation problem is

$$\max_{x_a, x_b} [x_a]^{\alpha}[x_b]^{1-\alpha} \tag{2.29}$$

subject to

$$x_a + x_b = x, \tag{2.30}$$

where α is the bargaining power of spouses a. The first-order condition (FOC) for this problem is

$$\frac{\alpha}{1-\alpha} = \frac{x_a}{x_b} \tag{2.31}$$

substituting in the constraint $x_a + x_b = x$ we have the solution

$$x_a = \alpha x, \tag{2.32}$$
$$x_b = (1-\alpha)x. \tag{2.33}$$

This gives the proof of Lemma 5.

The collective choice of labor supply

In our collective approach to household choice, spouses' labor supply maximizes the family welfare. Given our assumptions of efficient allocation within the household, the division of labor supply is made in order to minimize the disutility costs of getting a given income level $z \in \mathbb{R}$. Therefore, the collective choice of labor supply is the solution to

$$\min_{z_a, z_b} \alpha \left(\theta_a \frac{z_a^{1+\gamma}}{1+\gamma}\right) + (1-\alpha)\left(\theta_b \frac{z_b^{1+\gamma}}{1+\gamma}\right) \tag{2.34}$$

subject to

$$z_a + z_b = z. \tag{2.35}$$

The FOC for this problem is

$$\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b} = \left(\frac{z_a}{z_b}\right)^{\gamma} \tag{2.36}$$

or equivalently,

$$z_b = \left[\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right]^{\frac{1}{\gamma}} z_a. \tag{2.37}$$

Using $z = z_a + z_b$, we have

$$z_a = z\left[\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}} + 1\right]^{-1}, \tag{2.38}$$

$$z_b = z\left[\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}} + 1\right]^{-1}. \tag{2.39}$$

This gives the proof of Lemma 6.

Appendix B - Proofs

Proof of Proposition 4

Proof: Using Lemma 5 and 6 we can restate the family utility as

$$V(x, z, \theta_a, \theta_b, \alpha) = \alpha\ln\alpha + (1-\alpha)\ln(1-\alpha) + \ln x - \tag{2.40}$$

$$\left\{\alpha\theta_a\left[1+\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)} + (1-\alpha)\theta_b\left[1+\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}\right\}\frac{z^{1+\gamma}}{1+\gamma} \tag{2.41}$$

defining

$$\omega(\theta_a, \theta_b, \alpha) = \alpha\theta_a\left[1+\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)} + (1-\alpha)\theta_b\left[1+\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}, \tag{2.42}$$

and

$$\kappa(\alpha) = \alpha\ln\alpha + (1-\alpha)\ln(1-\alpha). \tag{2.43}$$

Thus, we have

$$V(x, z, \theta_a, \theta_b, \alpha) = \kappa(\alpha) + \ln x - \omega\frac{z^{1+\gamma}}{1+\gamma}. \tag{2.44}$$

$\blacksquare$

Proof of Proposition 5

This is a standard result since Lemmas 5, 6 and Proposition 4 allow us to write the planner's problem exactly as in Mirrlees' case, when we interpret individuals as a household. We present the proof for completeness.

Proof: Let $v(\omega)$ denote the utility of family with type ω in the mechanism. Taking $v(\omega)$ and $z(\omega)$ as the choice variables, $x(\omega)$ is implicitly defined by

$$v(\omega) = \ln x(v(\omega), z(\omega), \omega) - \omega \frac{z(\omega)}{1+\gamma}^{1+\gamma}. \tag{2.45}$$

The family utility satisfies the single-crossing property on the parameter θ since

$$\frac{dx}{dz}|_{v(\omega)=v} = \frac{h'(z)}{u'(x)} = \omega z x. \tag{2.46}$$

The slope of the indifference curve is increasing in the type. Hence, incentive compatibility can be fully characterized by the local constraints. In Alves, da Costa, and Moreira (2017b) we discuss the characterization of optimal mechanisms when the single-crossing property does not hold.

The planner solves the following program:

$$\max_{v(\cdot), z(\cdot)} \int_{\Omega} v(\omega)\psi(\omega) d\omega$$

subject to the local first-order incentive constraint:

$$\dot{v}(\omega) = \frac{z(\omega)^{1+\gamma}}{1+\gamma}; \tag{2.47}$$

the monotonicity constraint: $z(\omega)$ is increasing; and budget constraint:

$$\int_{\Omega} [z(\omega) - x(v(\omega), z(\omega), \omega)]\psi(\omega) d\omega = 0. \tag{2.48}$$

By the implicit function theorem,

$$x_v = \frac{1}{u'(x(v(\omega), z(\omega), \omega))} = x(v(\omega), z(\omega), \omega), \tag{2.49}$$

$$x_z = \frac{h'(z(\omega))}{u'(x(v(\omega), z(\omega), \omega))} = \omega z(\omega) x(v(\omega), z(\omega), \omega). \tag{2.50}$$

Under this approach the income $z(\omega)$ is the control variable and utility of the family $v(\omega)$

is the state variable. Hence, the Hamiltonian associated to this problem is

$$\mathcal{H}(v(\omega), z(\omega), \mu(\omega), \lambda, \omega) = v(\omega)\psi(\omega) + \lambda[z(\omega) - x(v(\omega), z(\omega), \omega)]\psi(\omega) - \mu(\omega)\frac{z(\omega)^{1+\gamma}}{1+\gamma},$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier associated to the government budget constraint and $\mu(\omega)$ is the co-state variable. At the optimum, the control variable $z(\omega)$ maximizes the Hamiltonian function:

$$\lambda\psi(\omega)[1 - x_z(\cdot)] = \mu(\omega)z(\omega)^{\gamma}. \tag{2.51}$$

The FOC of the Hamiltonian is

$$\dot{\mu}(\omega) = -\frac{\partial}{\partial v}\mathcal{H} = -[\psi(\omega) - \lambda x_v(\cdot)\psi(\omega)] \tag{2.52}$$

or

$$\dot{\mu}(\omega) = \psi(\omega)[\lambda x_v - 1]. \tag{2.53}$$

and the transversality condition is: $\mu(\underline{\omega}) = \mu(\overline{\omega}) = 0$. Integrating on both sides from θ to $\overline{\theta}$, we have

$$\mu(\overline{\omega}) - \mu(\omega) = \int_{\omega}^{\overline{\omega}} \dot{\mu}(s)ds = \int_{\omega}^{\overline{\omega}} \psi(s)[\lambda x_v(\cdot) - 1]ds. \tag{2.54}$$

Using the transversality condition we have

$$\mu(\omega) = \lambda \int_{\omega}^{\overline{\omega}} \psi(s)[\frac{1}{\lambda} - x_v(v(s), z(s), s)]ds. \tag{2.55}$$

Note that from the transversality condition $\mu(\underline{\omega}) = 0$ we can pin down the social value of an extra unit of income:

$$0 = \mu(\underline{\omega}) = \lambda \int_{\underline{\omega}}^{\overline{\omega}} \psi(s)[\frac{1}{\lambda} - x_v(v(s), z(s), s)]ds. \tag{2.56}$$

Therefore

$$\lambda = \frac{1}{\int_{\underline{\omega}}^{\overline{\omega}}[x_v(v(s), z(s), s)]\psi(s)ds} = \frac{1}{\int_{\underline{\omega}}^{\overline{\omega}}[x(v(s), z(s), s)]\psi(s)ds}$$

or

$$\lambda = \mathbb{E}^{\psi}[x]^{-1}.$$

Substituting $\mu(\omega)$ defined in (2.55) into equation (2.51) we have

$$[1 - x_z(v(\omega), z(\omega), \omega)] = \frac{z(\omega)^\gamma}{\psi(\omega)} \int_\omega^{\varpi} [\frac{1}{\lambda} - x_v(v(s), z(s), s)]\psi(s)ds.$$

This gives us an expression for the optimal joint income tax rate that follows exactly the one of the standard Mirrlees' model:

$$T'(z(\omega)) = \frac{z(\omega)^\gamma}{\psi(\omega)} \int_\omega^{\varpi} \left[x_v(v(s), z(s), s) - \frac{1}{\lambda}\right] \psi(s)ds, \tag{2.57}$$

whenever this differential equation gives an increasing function $z(\omega)$. Using $\lambda = \mathbb{E}^\psi[x]^{-1}$ and $x_v = x(v(\omega), z(\omega), \omega)$ we can re-write, with some abuse of notation ($x_v(v(s), z(s), s) = x(s)$), the formula for the optimal marginal taxation as

$$T'(z(\omega)) = \frac{z(\omega)^\gamma}{\psi(\omega)} \int_\omega^{\varpi} \left[x(s) - \mathbb{E}^\psi[x]\right] \psi(s)ds. \tag{2.58}$$

■

Proof of Proposition 6

Proof: Let $V_a(x, z, \theta_a, \theta_b)$ and $V_b(x, z, \theta_a, \theta_b)$ be indirect utility derived by spouses a and b in a couple (θ_a, θ_b) from the allocation (x, z).

Using Lemmas 5 and 6, we have

$$V_a(x, z, \theta_a, \theta_b) = \ln \alpha x - \omega_a(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma},$$

where

$$\omega_a(\theta_a, \theta_b, \alpha) \equiv \theta_a \left[1 + \left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}.$$

Analogously,

$$V_b(x, z, \theta_a, \theta_b) = \ln(1-\alpha)x - \omega_b(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma},$$

where

$$\omega_b(\theta_a, \theta_b, \alpha) \equiv \theta_b \left[1 + \left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}.$$

Note that given our definitions

$$\omega(\theta_a, \theta_b, \alpha) = \alpha \left[\omega_a(\theta_a, \theta_b, \alpha)\right] + (1-\alpha)\left[\omega_b(\theta_a, \theta_b, \alpha)\right].$$

From Proposition 4,

$$\begin{aligned}
&V(x, z, \theta_a, \theta_b) = \kappa(\alpha) + \ln x - \omega(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma} = \\
&\kappa(\alpha) + \ln(\alpha x \frac{1}{\alpha}) - \left[\alpha\omega_a(\theta_a, \theta_b, \alpha) + (1-\alpha)\omega_b(\theta_a, \theta_b, \alpha)\right]\frac{z^{1+\gamma}}{1+\gamma} = \\
&\kappa(\alpha) + \ln(\alpha x) - \ln(\alpha) - (1-\alpha)\left[\omega_b(\theta_a, \theta_b, \alpha) - \omega_a(\theta_a, \theta_b, \alpha)\right]\frac{z^{1+\gamma}}{1+\gamma} - \omega_a(\theta_a, \theta_b, \alpha)\frac{z^{1+\gamma}}{1+\gamma} = \\
&\qquad \kappa(\alpha) - \ln(\alpha) - (1-\alpha)\left[\omega_b(\theta_a, \theta_b, \alpha) - \omega_a(\theta_a, \theta_b, \alpha)\right]\frac{z^{1+\gamma}}{1+\gamma} + V_a(x, z, \theta_a, \theta_b).
\end{aligned}$$

Using

$$\begin{aligned}
&(1-\alpha)\left[\omega_b(\theta_a, \theta_b, \alpha) - \omega_a(\theta_a, \theta_b, \alpha)\right] = \\
&\left[(1-\alpha)\omega_b(\theta_a, \theta_b, \alpha) + \alpha\omega_a(\theta_a, \theta_b, \alpha)\right] - \omega_a(\theta_a, \theta_b, \alpha) = \\
&\qquad \omega(\theta_a, \theta_b, \alpha) - \omega_a(\theta_a, \theta_b, \alpha),
\end{aligned}$$

and rearranging terms we have

$$V_a(x, z, \theta_a, \theta_b) = \ln(\alpha) - \kappa(\alpha) + V(x, z, \theta_a, \theta_b) + \left[\omega(\theta_a, \theta_b, \alpha) - \omega_a(\theta_a, \theta_b, \alpha)\right]\frac{z^{1+\gamma}}{1+\gamma}.$$

Analogously,

$$V_b(x, z, \theta_a, \theta_b) = \ln(1-\alpha) - \kappa(\alpha) + V(x, z, \theta_a, \theta_b) + \left[\omega(\theta_a, \theta_b, \alpha) - \omega_b(\theta_a, \theta_b, \alpha)\right]\frac{z^{1+\gamma}}{1+\gamma}.$$

■

Utilitarian welfare criteria based on individuals

Let us assume that the planner is utilitarian across and within households. Recall that $V_a(x, z, \theta_a, \theta_b)$ and $V_b(x, z, \theta_a, \theta_b)$ are the indirect utility derived by spouses a and b in a couple (θ_a, θ_b) from the allocation (x, z).

The planner's welfare of a given allocation $(x, z) : \Theta \to \mathbb{R}_+ \times \mathbb{R}$ is (neglecting the explicit dependence of (x, z) on types to simplify the formula):

$$\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{1}{2}V_a(x(\cdot), z(\cdot), \theta_a, \theta_b) + \frac{1}{2}V_b(x(\cdot), z(\cdot), \theta_a, \theta_b)\right] f(\theta_a, \theta_b) d\theta_a d\theta_b.$$

Given the characterization in Proposition 6 of $V_a(x, z, \theta_a, \theta_b)$ and $V_b(x, z, \theta_a, \theta_b)$, we can re-write the welfare criteria as a function of (ω, ω_a). In fact, substituting V_a, V_b and $\kappa(\alpha)$ we have

$$
\begin{aligned}
&\frac{1}{2}V_a(x, z, \theta_a, \theta_b) + \frac{1}{2}V_b(x, z, \theta_a, \theta_b) = \\
&\frac{1}{2}\ln(\alpha) + \frac{1}{2}\ln(1-\alpha) - \alpha\ln(\alpha) - (1-\alpha)\ln(1-\alpha) + \\
&V(x, z, \theta_a, \theta_b) + \left[\omega - \frac{\omega_a + \omega_b}{2}\right]\frac{z^{1+\gamma}}{1+\gamma} \\
&= \left[\frac{1}{2} - \alpha\right]\ln\frac{\alpha}{1-\alpha} + V(x, z, \theta_a, \theta_b) + \left[\omega - \frac{\omega_a + \omega_b}{2}\right]\frac{z^{1+\gamma}}{1+\gamma} \\
&= \left[\frac{1}{2} - \alpha\right]\ln\frac{\alpha}{1-\alpha} + V(x, z, \theta_a, \theta_b) + \left(\alpha - \frac{1}{2}\right)[\omega_a - \omega_b]\frac{z^{1+\gamma}}{1+\gamma} \\
&= \left[\frac{1}{2} - \alpha\right]\ln\frac{\alpha}{1-\alpha} + V(x, z, \theta_a, \theta_b) + \left(\alpha - \frac{1}{2}\right)\left[\omega_a - \left(\frac{\omega - \alpha\omega_a}{1-\alpha}\right)\right]\frac{z^{1+\gamma}}{1+\gamma} \\
&\qquad = \left[\frac{1}{2} - \alpha\right]\ln\frac{\alpha}{1-\alpha} + V(x, z, \theta_a, \theta_b) + \left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a - \omega}{1-\alpha}\right)\frac{z^{1+\gamma}}{1+\gamma},
\end{aligned}
$$

where in the second equality we substitute V_a, V_b and $\kappa(\alpha)$. In the third and fifth equalities we used $\omega = \alpha\omega_a + (1-\alpha)\omega_b$.

Therefore, we can define the planner's social welfare function based on individuals in terms of family type and spouse a type as

$$
W(x, z, \omega, \omega_a) = \left[\frac{1}{2} - \alpha\right]\ln\frac{\alpha}{1-\alpha} + V(x, z, \omega, \omega_a) + \left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a - \omega}{1-\alpha}\right)\frac{z^{1+\gamma}}{1+\gamma}. \quad (2.59)
$$

Proof of Proposition 7

Using Lemmas 5, 6 and Proposition 6 we can rewrite the planner's problem as a function of the family's type ω and ω_a in a simpler way, and then apply the usual tools of optimal control problems to find the optimal allocation in the space of piecewise continuous functions.

Proof: Let $v(\omega)$ denote the utility get by type θ in the mechanism. Taking $v(\omega)$ and $z(\omega)$ as choice variables, $x(\omega)$ is implicitly defined by

$$
v(\omega) = \ln x(v(\omega), z(\omega), \omega) - \omega\frac{z(\omega)}{1+\gamma}^{1+\gamma}. \quad (2.60)
$$

The utility function of the family satisfies the single-crossing property with respect to

parameter ω since

$$\frac{dx}{dz}|_{v(\omega)=v} = \frac{h'(z)}{u'(x)} = \omega z x. \tag{2.61}$$

The slope of the indifference curve is increasing in the type. Hence, incentive compatibility can be fully characterized by the local and monotonicity constraints.

First, notice that we can re-write the objective function as

$$\int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega_a}}^{\overline{\omega_a}} \left[v(\omega) + \left(\alpha - \frac{1}{2}\right) \left(\frac{\omega_a - \omega}{1-\alpha}\right) \frac{z(\omega)^{1+\gamma}}{1+\gamma} \right] \phi(\omega_a|\omega)\psi(\omega) d\omega_a d\omega =$$
$$\int_{\underline{\omega}}^{\overline{\omega}} v(\omega) \left[\int_{\underline{\omega_a}}^{\overline{\omega_a}} \phi(\omega_a|\omega) d\omega_a \right] \psi(\omega) d\omega +$$
$$\int_{\underline{\omega}}^{\overline{\omega}} \left(\alpha - \frac{1}{2}\right) \frac{z(\omega)^{1+\gamma}}{1+\gamma} \int_{\underline{\omega_a}}^{\overline{\omega_a}} \left[\left(\frac{\omega_a - \omega}{1-\alpha}\right) \phi(\omega_a|\omega) d\omega_a \right] \psi(\omega) d\omega.$$

Using

$$\left[\int_{\underline{\omega_a}}^{\overline{\omega_a}} \phi(\omega_a|\omega) d\omega_a \right] = 1,$$

and

$$\int_{\underline{\omega_a}}^{\overline{\omega_a}} \omega_a \phi(\omega_a|\omega) d\omega_a \equiv \mathbb{E}[\omega_a|\omega],$$

we can re-write the objective function as

$$\int_{\underline{\omega}}^{\overline{\omega}} \left[v(\omega) + \left(\alpha - \frac{1}{2}\right) \left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right) \frac{z(\omega)^{1+\gamma}}{1+\gamma} \right] \psi(\omega) d\omega.$$

Thus, the planner solves the following program:

$$\max_{v(\cdot), z(\cdot)} \int_{\underline{\omega}}^{\overline{\omega}} \left[v(\omega) + \left(\alpha - \frac{1}{2}\right) \left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right) \frac{z(\omega)^{1+\gamma}}{1+\gamma} \right] \psi(\omega) d\omega. \tag{2.62}$$

subject to local first-order local incentive constraint:

$$\dot{v}(\omega) = \frac{z(\omega)^{1+\gamma}}{1+\gamma}; \tag{2.63}$$

and monotonicity constraints: that $z(\omega)$ is increasing; and budget constraint:

$$\int_{\underline{\omega}}^{\overline{\omega}} [z(\omega) - x(v(\omega), z(\omega), \omega)] \psi(\omega) d\omega \geq 0. \tag{2.64}$$

By the implicit function theorem,

$$x_v = \frac{1}{u'\left(x(v(\omega), z(\omega), \omega)\right)} = x(v(\omega), z(\omega), \omega),$$
$$x_z = \frac{h'\left(z(\omega)\right)}{u'\left(x(v(\omega), z(\omega), \omega)\right)} = \omega z(\omega) x(v(\omega), z(\omega), \omega). \quad (2.65)$$

In this approach the income $z(\omega)$ is the control variable and utility of the family $v(\omega)$ is the state variable.

The Hamiltonian associated to this problem is

$$\mathcal{H}(v(\omega), z(\omega), \mu(\omega), \lambda, \omega) = \left[v(\omega) + \left(\alpha - \frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right)\frac{z(\omega)^{1+\gamma}}{1+\gamma}\right]\psi(\omega)$$
$$+ \lambda\left[z(\omega) - x(v(\omega), z(\omega), \omega)\right]\psi(\omega) - \mu(\omega)\frac{z(\omega)^{1+\gamma}}{1+\gamma},$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier associated to the government budget constraint, and $\mu(\omega)$ is the co-state variable. At the optimum, the control $z(\omega)$ maximizes the Hamiltonian function giving the FOC:

$$\left[\left(\alpha - \frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right) z(\omega)^{\gamma} + \lambda\left(1 - x_z(\cdot)\right)\right]\psi(\omega) = \mu(\omega) z(\omega)^{\gamma} \quad (2.66)$$

and the FOC of the Hamiltonian is

$$\dot{\mu}(\omega) = -\frac{\partial}{\partial v}\mathcal{H}(\cdot) = -[1 - \lambda x_v(\cdot)]\psi(\omega) \quad (2.67)$$

or

$$\dot{\mu}(\omega) = \psi(\omega)[\lambda x_v(\cdot) - 1] \quad (2.68)$$

and the transversality conditions $\mu\left(\underline{\omega}\right) = \mu\left(\overline{\omega}\right) = 0$. Integrating on both sides from θ to $\overline{\omega}$ and applying the fundamental theorem of calculus we have

$$\mu\left(\overline{\omega}\right) - \mu(\theta) = \int_{\theta}^{\overline{\omega}} \dot{\mu}(s) ds = \int_{\theta}^{\overline{\omega}} \psi(s)[\lambda x_v(\cdot) - 1] ds. \quad (2.69)$$

Using the transversality condition we have

$$\mu(\theta) = \lambda \int_{\theta}^{\overline{\theta}} \psi(s)[\frac{1}{\lambda} - x_v(v(s), z(s), s)] ds. \quad (2.70)$$

Note that from the transversality condition $\mu(\underline{\theta}) = 0$ we can pin down the social value of

an extra unit of income:

$$0 = \mu(\underline{\theta}) = \lambda \int_{\underline{\theta}}^{\overline{\theta}} \psi(s)[\frac{1}{\lambda} - x_v(v(s), z(s), s)]ds. \tag{2.71}$$

Therefore,

$$\lambda = \frac{1}{\int_{\underline{\theta}}^{\overline{\theta}} [x_v(v(s), z(s), s)]\psi(s)ds} = \frac{1}{\int_{\underline{\theta}}^{\overline{\theta}} [x(v(s), z(s), s)]\psi(s)ds} \tag{2.72}$$

or

$$\lambda = \mathbb{E}^{\psi}[x]^{-1}. \tag{2.73}$$

Substituting $\mu(\theta)$ in (2.70) into equation (2.66) we have

$$\left[\left(\alpha - \frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right) z(\omega)^{\gamma} + \lambda\left(1 - x_z(\,\cdot\,)\right)\right] \psi(\omega) = $$
$$z(\omega)^{\gamma}\lambda \int_{\theta}^{\overline{\theta}} \psi(s)[\frac{1}{\lambda} - x_v(v(s), z(s), s)]ds.$$

This gives us the following expression for the marginal tax rate

$$T'\left(z(\omega)\right) = \frac{z(\omega)^{\gamma}}{\psi(\omega)} \int_{\theta}^{\overline{\theta}} \left[x_v(v(s), z(s), s) - \frac{1}{\lambda}\right] \psi(s)ds - \frac{z(\omega)^{\gamma}}{\lambda}\left(\alpha - \frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right).$$

Using $\lambda = \mathbb{E}^{\psi}[x]^{-1}$ and $x_v = x(v(\theta), z(\theta), \theta)$, we can re-write with some abuse of notation $(x_v(v(s), z(s), s) = x(s))$ the formula for the optimal marginal taxation as

$$T'\left(z(\omega)\right) = \frac{z(\omega)^{\gamma}}{\psi(\omega)} \int_{\theta}^{\overline{\theta}} \left[x(s) - \mathbb{E}^{\psi}[x]\right] \psi(s)ds - z(\omega)^{\gamma}\mathbb{E}^{\psi}[x]\left(\alpha - \frac{1}{2}\right)\left(\frac{\mathbb{E}[\omega_a|\omega] - \omega}{1-\alpha}\right). \tag{2.74}$$

■

Proof of Proposition 8

The proof is analogous to Proposition 4, we present here for completeness.

Proof: In our collective approach to household choice, the spouses' labor supply maximizes the family welfare. Given our assumptions of efficient allocation within the household, the division of labor supply is made in order to minimize the disutility costs of getting a given income level $z \in \mathbb{R}$. Therefore, the collective choice of labor supply is the

solution to

$$\min_{z_a,z_b}\ \alpha\left[\theta_a\frac{z_a^{1+\gamma}}{1+\gamma}\right]+(1-\alpha)\left[\theta_b\frac{z_b^{1+\gamma}}{1+\gamma}\right] \tag{2.75}$$

subject to

$$z_a(1-t)+z_b=z. \tag{2.76}$$

The FOC for this problem is

$$\frac{\alpha\theta_a z_a^{\gamma}}{1-t}=(1-\alpha)\theta_b z_b^{\gamma}. \tag{2.77}$$

Using $z=z_a(1-t)+z_b$ we have

$$z_a=z\left[\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\frac{1}{1-t}\right)^{\frac{1}{\gamma}}+1\right]^{-1} \tag{2.78}$$

$$z_b=z\left[\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}(1-t)\right)^{\frac{1}{\gamma}}+1\right]^{-1}. \tag{2.79}$$

Substituting in the family utility function we have

$$V(x,z,\theta_a,\theta_b,\alpha,t)=\kappa(\alpha)+\ln x-\omega(\theta_a,\theta_b,\alpha,t)\frac{z^{1+\gamma}}{1+\gamma}, \tag{2.80}$$

where

$$\omega(\theta_a,\theta_b,\alpha,t)\equiv\alpha\theta_a\left[1+\left(\frac{\alpha}{1-\alpha}\frac{\theta_a}{\theta_b}\frac{1}{(1-t)}\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)}+$$
$$(1-\alpha)\theta_b\left[1+\left(\frac{1-\alpha}{\alpha}\frac{\theta_b}{\theta_a}(1-t)\right)^{\frac{1}{\gamma}}\right]^{-(1+\gamma)} \tag{2.81}$$

and

$$\kappa(\alpha)=\alpha\ln\alpha+(1-\alpha)\ln(1-\alpha). \tag{2.82}$$

■

Proof of Proposition 9

Proof: Note that if $t>0$ and $\alpha\theta_a<(1-\alpha)\theta_b$, then from Proposition 6 $\omega(\theta_a,\theta_b,\alpha)<\omega(\theta_a,\theta_b,\alpha,t)$, and, therefore, $v(\omega)>v(\omega^t)$. On the other hand, if $t<0$ and $\alpha\theta_a<$

$(1-\alpha)\theta_b$, then $\omega(\theta_a,\theta_b,\alpha) > \omega(\theta_a,\theta_b,\alpha,t)$, and, therefore, $v(\omega) < v(\omega^t)$.

It is straightforward that if $t < 0$ and $\alpha\theta_a < (1-\alpha)\theta_b$, we have $\omega_a(\theta_a,\theta_b,\alpha,t) > \omega_a(\theta_a,\theta_b,\alpha)$ and $\omega(\theta_a,\theta_b,\alpha,t) < \omega(\theta_a,\theta_b,\alpha)$, which lead to $[\omega_a^t - \omega^t] > [\omega_a - \omega]$. In this case,

$$v(\omega^t) + \left[\left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a^t - \omega^t}{1-\alpha}\right)\right] > v(\omega) + \left[\left(\alpha - \frac{1}{2}\right)\left(\frac{\omega_a - \omega}{1-\alpha}\right)\right].$$

Therefore, it is welfare improving to subsidize women when $\alpha\theta_a < (1-\alpha)\theta_b$. ■

Results used in the Participation Section

Lemma 10 *At the optimum there is a θ_b' such that for $\theta_b > \theta_b'$ there is $\theta_a' \in \Theta$ such that $\mathcal{C}(\theta_a,\theta_b) = 1$ for all $\theta_a > \theta_a'$. Moreover, if $\theta_b' > 0$ then, for all $\theta_b < \theta_b'$, $\mathcal{C}(\theta_a,\theta_b) = 2$ for all $\theta_a \in \Theta$.*

Proof: Assume that this is not the case. I.e., assume that there is $\theta_b'' > \theta_b'$ such that $\mathcal{C}(\theta_a,\theta_b) = 2$ for all $\theta_a \in \Theta$, $\theta_b \in [\theta_b'', a)$, where $a > \theta_b''$ is possibly infinite. Without loss we define $c(\theta_b) = (1-\alpha)\theta_b$ for $\theta_b \in [\theta_b'', a)$. Now consider replacing $c(\cdot)$ with $\bar{c}(\cdot)$ which we define as follows. Choose δ_1 and δ_2 satisfying

$$\max\left\{\Big|v_2(c(\theta_b'') - \delta_1) - c(\theta_b''))\Big|; \Big|v_2(c(a+\delta_2) - c(a))\Big|\right\} < \kappa,$$

$$c(\theta_b'') - \delta_1 = a + \delta_2 = \epsilon$$

for some small positive κ and ϵ. Now, $\bar{c}(\cdot)$ is equal to $c(\cdot)$ in $\Theta \setminus \big[\theta_b'' - \delta_1, a+\delta_2\big)$, and is defined through $\bar{c}(\theta_b) = (1-\alpha)\theta_b - \epsilon$ in $\big[\theta_b'' - \delta_1, a+\delta_2\big)$. Note that $\dot{\bar{c}}(\theta_b) = 1-\alpha$ in $\big[\theta_b'' - \delta_1, a+\delta_2\big)$. Using (2.23) one easily verifies that the allocation $v_1(\cdot)$ in this interval is defined through $z_1(\theta_b)^{1+\gamma} = z_2(c(\theta_b))^{1+\gamma}$. For small enough ϵ, $c(\theta_b) \approx (1-\alpha)\theta_b$, which implies $\ln x_1(\theta_b) = \ln x_2(c(\theta_b)) - \alpha A$. The utility change with the new allocation is infinitesimal for a finte reduction in cost. ■

Appendix C - numerical analysis

In this section, we compute numerically the optimal tax schedules derived in Propositions 5 and 7, using real data from the 2017 March CPS to have a clear idea of the shape of the optimal taxation in both cases. We use the sample of couples without dependents, using wages as a proxy for types.

Note that the marginal tax rate in Proposition 7 depends on the distribution of $\omega, \psi(\omega)$ and on the distribution of ω_a, conditional on $\omega, \phi(\omega_a \mid \omega)$. However, we can calculate the implied distribution of $\omega, \psi(\omega)$ from the knowledge of the wage distribution.

It is important to notice that we only need to estimate the conditional expectation $\mathbb{E}[\omega_a|\omega]$ and we do that non-parametrically using a simple kernel regression. This simplifies our analysis by avoiding the calculation of the conditional distribution of $\omega, \phi(\omega_a \mid \omega)$ without relying on any additional restrictive assumptions.

After calculating these objects, we simulate the marginal tax rate by adapting the algorithm proposed by Mankiw, Weinzierl, and Yagan (2009), based on a discrete grid of the empirical distributions. The algorithm goes as in the following steps:

Step 1: Given $\psi(\omega)$, discretize in a probability mass function $\pi(\omega)$ in the following way: divide the interval considered $[\underline{\omega}, \overline{\omega}]$ in N bins of equal width Δ, creating a grid where $\{\omega_i\}_{i=1}^N$ with $\underline{\omega} \leq \omega_1 \leq \cdots \leq \omega_n \leq \cdots \leq \omega_N \leq \overline{\omega}$ are the midpoints in the grid. The probability mass function at each ω is given by,

$$\pi(\omega) = \Psi(\omega + \Delta/2) - \Psi(\omega - \Delta/2), \tag{2.83}$$

where $\Psi(\cdot)$ is the cumulative distribution function of ω.

Step 2: For each $\{\omega_i\}_{i=1}^N$ estimate the conditional expectation using the Nadaraya and Watson kernel regression as

$$\hat{\mathbb{E}}[\omega_a|\omega_i] = \frac{1/(Nh)\sum_{j=1}^N K(\frac{\omega_j - \omega_i}{h})\omega_{aj}}{1/(Nh)\sum_{j=1}^N K(\frac{\omega_j - \omega_i}{h})}. \tag{2.84}$$

where $K(\cdot)$ is the Gaussian density and h is the bandwidth.

Now we start the loop for the calculation of the optimal marginal tax rate $T_k(\omega_i)$.

Step 3: Start with a simple guess for the marginal tax rate. For instance, a flat tax schedule $T_0(\omega_i) = .35$ for all ω_i in the grid and a lump-sum transfer $t_0 = .0001$.

Step 4: Since $T_{k-1}(\cdot)$ at each iteration is only calculated at points of the grid, extrapolate $T_{k-1}(\cdot)$ to be defined as well at points outside the grid in the following way:

$$T'_{k-1}(\omega) = \begin{cases} T'_{k-1}(\omega_1), \text{ for } \underline{\omega} \leq \omega \leq \omega_1; \\ T'_{k-1}(\omega_2), \text{ for } \omega_1 \leq \omega \leq \omega_2; \\ \dots \\ T'_{k-1}(\omega_n), \text{ for } \omega_{n-1} \leq \omega \leq \omega_n; \\ \dots \\ T'_{k-1}(\omega_N), \text{ for } \omega_{N-1} \leq \omega \leq \overline{\omega}; \end{cases} \tag{2.85}$$

generating the tax schedule

$$T_{k-1}(\omega_i) = \begin{cases} T'_{k-1}(\omega_1)z(\omega_i) - t_{k-1}, \text{ for } \underline{\omega} \leq \omega_i \leq \omega_1; \\ T'_{k-1}(\omega_1)z(\omega_1) + T'_{k-1}(\omega_2)[z(\omega_i) - z(\omega_1)] - t_{k-1}, \text{ for } \omega_1 \leq \omega_i \leq \omega_2; \\ \dots \\ T'_{k-1}(\omega_1)z(\omega_1) + \sum_{j=2}^{n-1} T'_{k-1}(\omega_j)[z(\omega_j) - z(\omega_{j-1})] + \\ \qquad T'_{k-1}(\omega_n)[z(\omega_i) - z(\omega_{n-1})] - t_{k-1}, \text{ for } \quad \omega_{n-1} \leq \omega_i \leq \omega_n; \end{cases} \tag{2.86}$$

Step 5: Given $T_{k-1}(\cdot)$, calculate the optimal labor supply at each ω_i by solving

$$max_z \ln(z - T_{k-1}(z)) - \omega_i \frac{z^{1+\gamma}}{1+\gamma}. \tag{2.87}$$

The necessary first-order condition for the type ω_i problem is

$$1 - T'_{k-1}(z) = \omega z^{\gamma} (z - T_{k-1}(z)) . \tag{2.88}$$

and the optimal $z^{\star}(\omega)$ is implicitly defined in this FOC.

Step 6: Given the optimal choice of $z(\omega)$, we can define the consumption as

$$x(\omega) = z^{\star}(\omega) - T_{k-1}(z^{\star}(\omega)) \tag{2.89}$$

and use all the information to update the marginal tax rate in each point of the grid using 6 and the first-order condition of the planner's problem as follows

$$\begin{aligned} T'_k(\omega_i) = \frac{z^{\star}(\omega_i)^{\gamma}}{\psi(\omega_i)} \int_{\omega_i}^{\overline{\omega}} \left[x(s) - \mathbb{E}^{\psi}[x]\right] \psi(s) ds \\ - z^{\star}(\omega_i)^{\gamma} \mathbb{E}^{\psi}[x] \left(\alpha - \frac{1}{2}\right) \left(\frac{\mathbb{E}[\omega_a|\omega_i] - \omega_i}{1-\alpha}\right). \end{aligned} \tag{2.90}$$

Since the family utility satisfies the single crossing property w.r.t. ω, this FOC and the monotonicity of $z^{\star}(\omega)$ are sufficient for optimality.

Step 7: The formula in *Step 6* is unfeasible to estimate given our approximation of the wage distribution in a grid. Therefore, we must use a discrete approximation. One

possibility is as follows:

$$
\begin{aligned}
&T_k'(\omega_i) \approx \frac{1}{x(\omega_i)\omega_i}\frac{1}{\pi(\omega_i)/\Delta}\left[\sum_{j=i+1}^{N}\left[x(\omega_j) - \hat{\mathbb{E}}[x]\right]\pi(\omega_j)\right] \\
&- \frac{1}{x(\omega_i)\omega_i}\hat{\mathbb{E}}[x]\left(\alpha - \frac{1}{2}\right)\left(\frac{\hat{\mathbb{E}}[\omega_a|\omega_i] - \omega_i}{1-\alpha}\right) = \\
&\frac{1}{x(\omega_i)\omega_i}\frac{1}{\pi(\omega_i)/\Delta}\left[\sum_{j=i+1}^{N} x(\omega_j)\pi(\omega_j) - \hat{\mathbb{E}}[x](1-\Pi(\omega_i))\right] \\
&\qquad\qquad - \frac{1}{x(\omega_i)\omega_i}\hat{\mathbb{E}}[x]\left(\alpha - \frac{1}{2}\right)\left(\frac{\hat{\mathbb{E}}[\omega_a|\omega_i] - \omega_i}{1-\alpha}\right), \qquad (2.91)
\end{aligned}
$$

where $\hat{\mathbb{E}}[\omega_a|\omega_i]$ is calculated at *Step 2* and

$$
\hat{\mathbb{E}}[x] = \sum_{j=1}^{N} x(\omega_j)\pi(\omega_j). \qquad (2.92)
$$

Step 8: Update t to ensure the budget constraint in the following way:

$$
t_k = \sum_{j=1}^{N} T_k(\omega_j)\pi(\omega_j). \qquad (2.93)
$$

Step 9: Return to *Step 4* until

$$
\max_i\{|T_k'(\omega_i) - T_{k-1}'(\omega_i)|\} < 10^{-3}. \qquad (2.94)
$$

Step 10: After getting a fixed-point of this operator, the only thing to be checked is the monotonicity to guarantee that the found tax schedule is implementable.

Note that this approximation is increasingly precise as $\Delta \to 0$ and $\omega_N \to +\infty$. From this numerical exercise, we can make some counterfactuals. For example, evaluating how the marginal tax rates vary as the spouses' bargaining power changes within the household.

Chapter 3

A bridge between commitment and non-commitment: Stochastic types in Laffont and Tirole's dynamic regulation model

Abstract

This paper investigates the introduction of type dynamic in the Laffont and Tirole's regulation model. The regulator and the firm are engaged in a two period relationship governed by short-term contracts, where, the regulator observes cost but cannot distinguish how much of the cost is due to effort on cost reduction or efficiency of firm's technology, named type. There is asymmetric information about the firm's type.

Our model is developed in a framework in which the regulator learns with firm's choice in the first period and uses that information to design the best second period incentive scheme. The regulator is aware of the possibility of changes in types and takes that into account. We show how type dynamic builds a bridge between commitment and non-commitment situations. In particular, the possibility of changing types mitigates the ratchet effect. We show that when degree of type dynamics is small the equilibrium shows separation and the welfare achieved is close to his upper bound (given by the commitment allocation).

3.1 Introduction

Regulating a firm is one of the most important issues in modern economies. The task of inducing firms to act efficiently is a challenge; especially when the economic environment is complex. The branch of economic theory known as contract theory has shown how a commonplace activity of procurement is affected by the asymmetry of information. In fact, it is well known that different information between two parts, in a relationship, generates the necessity of designing incentive schemes to mitigate the effects of opportunistic behavior taken by the most informed part.

The theory of regulation took into account these lessons and a new look to this questions arose. Laffont (1994), in a very interesting survey, describes how the use of contract theory tools has changed regulation agenda. In this work, we focus on two aspects of regulation: lack of commitment and type dynamic.

Our framework was first presented in Laffont and Tirole (1986), under which observability of costs is used by the regulator in the design of an incentive contract when adverse selection and moral hazard takes place. Their main finding is that there is a trade-off involving informational rent extraction and ex-post efficiency. Laffont and Tirole (1988) extend the analysis to a two-period dynamic setup and discuss the role of ratchet effect in the design of incentive contract when the regulator cannot commit himself to a second-period scheme and with a continuum of types.

Lack of commitment emerges when the economic environment (for example, legal prohibitions or the incapability of forecasting the relevant future) restricts the regulator to run a long-term relationship by a series of short-term contracts. [1]

According to Laffont and Tirole (1993): "If the regulated firm produces at a low-cost today, the regulator might infer that low costs are not that hard to achieve and tomorrow offer a demanding incentive scheme. That is, the firm jeopardizes future rents by being efficient". The literature calls this phenomenon the "ratchet effect". Laffont and Tirole (1988) show that the ratchet effect, in a model with continuous of types, leads to much pooling at the first-period allocation. The existence of pooling allocations and ratchet effect are the big issues of this setup.

Laffont and Tirole (1987) and Laffont and Tirole (1993) developed a two-type version of the model in Laffont and Tirole (1988). The impossibility of separation equilibrium is no longer an issue. Equilibria usually show some degree of pooling/separation, but full separation is possible as well. The main result found here is that separation will occur with a small discount factor, pooling with a large discount factor or when the uncertainty

[1] Laffont and Tirole (1993) (see in section 9.1) list a number of situations generating lack of commitment.

about agent's ability is negligible.

The literature also has focused on another feature of dynamic contracting. Another issue is type dynamic, which we define as the possibility of change on agent's type.

Our model is an extension of the two type version Laffont and Tirole's model. The primary focus of our paper is the interaction between type dynamic and the ratchet effect. In particular, as we will further see, type dynamic builds a bridge between commitment and non-commitment situations. In fact, the possibility of changing types mitigates the "ratchet effect".

Our main finding is that the lack of commitment plays a role only in economic environments where the asymmetric information is persistent. In other words, if the asymmetric information is volatile the regulator doesn't have any advantage in using the learned information. Furthermore, based on our simulation, we conjecture two properties for the optimal contract.

We show in the independent case that the optimal contract without commitment differs from the one with commitment. Precisely, the regulator has power enough to extract the full rent get by an efficient firm in the second period (without distorting the allocation) but, due to the lack of commitment, he is not able to induce the inefficient firm to do the first-best effort.

In our numerical simulations, if the efficient type is as likely as the inefficient type in both periods, the equilibrium welfare is very close to his upper bound. The pooling allocations occur for a high enough persistence, and the expected rent is increasing is the persistence level.

In section 2, we present the model, which is an extension of Laffont and Tirole's dynamic regulation model. In fact, we allow that firm's efficiency parameter changes between periods. To control the degree of persistence, we make assumptions about how the type dynamic occurs. We present the full information setup and the dynamic setup under commitment for reference.

Then, in section 3, we present the dynamic setup without commitment. As in Laffont and Tirole (1993) and Laffont and Tirole (1988) we have three possible kinds of semi-separation equilibria. We also completely characterize them.

In section 4 we present our results. We calculate the optimal contract in the independent case under non-commitment. We investigate how the parameter of persistence has influenced the distortion from a separation allocation about the commitment. In fact, with type dynamic, even with highly correlated types, the effect is local, and it is very difficult to have global implications about the behavior of the solution under imperfect

correlation. As an alternative to deal with this difficult, we develop numerical simulation in section 4.4. We finally conclude in section 5. All proofs are found in the Appendix.

Related Literature

Baron and Besanko (1984) study long-term relationship under full commitment and type dynamic. In particular, in their equation (21), they impose a structure for the type efficiency that resembles ours. Their main findings are that, in the independent case, the optimal contract yields to the first best allocation in the second period and, in the perfect correlation case, the optimal contract is the repetition of the static contract.

As we will further see, our findings differ from theirs. In our model, the optimal contract in the independent case is not the repetition of the static contract. Moreover, in the perfect correlation case our model is isomorphic to the Laffont and Tirole's model and the optimal first-period contract, in general, involves stochastic mechanisms. The source of this difference is the lack of commitment in our model. The firm now is very concerned about how the regulator will use any information learned and acts strategically.

Our work shares similarities with Ramos (2011). Both works extend the original Laffont and Tirole's model into a type dynamic framework. But here we assume only two types and we control for type persistence and Ramos (2011) has types with i.i.d. Continuous distribution.

Others papers (e.g. Roberts (1983) and Townsend (1982)) have the same result found in Baron and Besanko (1984) for the i.i.d. case. The intuition is that, once in the commitment situation with i.i.d. types, the contracts foresee the action taken by the agents in the future. At the moment of designing the contract, there is no asymmetric information, except for the first period. So, the results must coincide with the full information case. However, in our model, the asymmetric information is renewed each period.

Kusuda (2006) takes the Laffont and Tirole's model and repeats it successively in a matching framework with different regulators and agents. Kusuda (2006) allows the agent's type to vary in the relationships. Nevertheless, in our model, we let the firm's type to vary within a relationship.

Battaglini (2005), in an infinitely repeated setting by a monopolist to a consumer whose preferences follow a Markov process, analyzes the effects of type dynamic in the design of incentive contracts under commitment. Others like Courty and Li (2000), Kennan (2001) and Hendel and Lizzeri (1997) analyze relationship governed by short term contracts in contexts different than ours. For example, Courty and Li (2000) analyze how sequential mechanisms help producers to price discriminate when consumers learn

private information about their demand over time. In contrast to our model, the private information is revealed sequentially. On the other hand, in ours, the private information reborn each period.

Pavan, Segal, and Toikka (2011) analyze the design of dynamic contracts when the agents' types follow a stochastic process, decisions are made sequentially, and the payoff does not need to be time separable. Their article focus in a general characterization of these contracts and, in particular, an envelope formula (which generalizes Mirrlees's formula[2]) for the derivative of the agent's payoff, in equilibrium, with respect to his current type summarizes all the first-order constraints in an incentive-compatible scheme.

3.2 The Model

The model is an extension of Laffont and Tirole's regulation setup in order to allow type dynamic. Consider a two-period relationship between a firm (the agent) and a regulator (principal) in which the firm implements, each period, an indivisible project with cost structure given by:

$$C_\tau = \theta_\tau - e_\tau, \tag{3.1}$$

where e_τ is the private action taken by the firm and represents the effort performed in reducing costs and θ_τ is an efficiency parameter known only by the firm.[3]

We assume that the cost is observable by the regulator, but he cannot distinguish how much of the cost is due to effort or efficiency. The inability to distinguish the composition of cost is the source of asymmetric information and permits the firm to act opportunistically.

Given a transfer schedule and a choice of effort, the firm's utility in each period is

$$U(e_\tau, t_\tau) = t_\tau - \psi(e_\tau), \tag{3.2}$$

where t_τ is the net monetary transfer received from the regulator and $\psi(e_\tau)$ is his disutility incurred to achieve effort level e_τ. We make some regularity assumptions.

Assumption 3 *$\psi(\cdot)$ is three times continuously differentiable with $\psi' > 0$, $\psi'' > 0$, $\psi''' > 0$ and there exists e^* such that $\psi'(e^*) = 1$.*

[2]This formula give us a necessary condition for a mechanism to be incentive compatible. (see Mirrlees (1971a))

[3]We can interpret the project as a process of procurement for a public good not sold in the market, so we can ignore any price influence.

In this economy we assume that public spending is financed by distortionary taxation. Let $\lambda > 0$ be the social cost of raising public funds and S be the social benefit of the project. The consumer's welfare is given by

$$S - (1+\lambda)[C_\tau + t_\tau]. \tag{3.3}$$

The regulator is utilitarian in the following sense: his objective is to choose $(t_\tau, C_\tau)_{\tau=1,2} \in (\mathbb{R} \times \mathbb{R})^2$ to maximize intertemporal social surplus. The regulator's static utility is the sum between consumers' and firm's utility. Therefore, the regulator utility is

$$W(C_\tau, t_\tau, e_\tau) = S - (1+\lambda)[C_\tau + t_\tau] + U(e_\tau, t_\tau). \tag{3.4}$$

Finally, we assume that both the regulator and firm discount future by the same discount rate $\delta > 0$.

In the asymmetric information framework the parameter θ_τ is a private information of the firm and is the source of the whole informational asymmetry. Differently from the original model, we allow a change in the type from period one to period two. With the purpose of controlling the degree of type persistence, we make assumptions about how the stochastic process governing the firm's type evolves.

Assumption 4 *Firm's types (θ_1, θ_2) evolves over time according to the following rule:*

- *θ_1 is a discrete random variable with support $\{\theta_L, \theta_H\}$, $\theta_L < \theta_H$. Let ν_1 be the prior, shared by the firm and the regulator, of firm being efficient in period one:*

$$\nu_1 = \Pr(\theta_1 = \theta_L); \tag{3.5}$$

- *θ_2 is defined by a lottery. With probability μ, type remains the same of the previous period, and with probability $1-\mu$, it will be drawn again from a new discrete random variable with the same support. Let ν_2 be the probability, known by both the firm and the regulator, of the firm being efficient in period two, in case of a new drawing:*

$$\nu_2 = \Pr(\theta_2 = \theta_L | ND); \tag{3.6}$$

where ND refers to new draw.

One possible interpretation for the changing in firm's type is that the firm is engaged in a restructuring process which can be successful or not in making her efficient.

We assume that the regulator is aware of the possibility of type changes. Our approach clearly generalizes Laffont and Tirole (1993) once with $\mu = 1$ our model is

exactly the original model. On the other hand, with $\mu = 0$ we have the type independent model, where types are independently drawn in each period. Compared to the original model, it is more difficult for the regulator to learn about the firm's type. For example, even with full revelation in the first period, the second period has asymmetric information. Define $\Delta\theta = \theta_H - \theta_L$.

The parameter $\mu \in [0, 1]$ is a global measure of informativeness, it determines the weight of θ_1 in determining θ_2.[4]

3.2.1 Full Information Setup

In the full information setup, the regulator knows the realization of firm's type and can directly contract on effort. His objective is to maximize intertemporal welfare subject to participation constraints.

Definition 2 ***(Full information problem)*** *If regulator observes the type and can contract on effort directly, he would solve*

$$\max_{e_\tau, t_\tau} \sum_{\tau=1,2} \delta^{\tau-1}\{S_\tau - (1+\lambda)(\theta_\tau - e_\tau + \psi(e_\tau)) - \lambda[t_\tau - \psi(e_\tau)]\} \tag{3.7}$$

subject to $t_\tau - \psi(e_\tau) \geq 0$ *for* $\tau = 1, 2$.

In the following proposition, we solve this problem.

Proposition 10 *The solution of full information problem (First Best) is given by* $\psi'(e^*_\tau) = 1$ *and* $t^*_\tau = \psi(e^*_\tau)$.

Proof: See Laffont and Tirole (1993). ■

In the full information situation, the regulator chooses the level of effort in order to equalize the marginal benefits of reducing the cost $(1 + \lambda)$ and the social loss in terms of firms disutility $(1 + \lambda)\psi'(e)$. Since the regulator dislikes leaving any rent, he will transfer the minimum necessary to the firm not to quit the relationship or, equivalently, makes the constraint binding at the solution. As we easily see this problem is stationary, and the solution is the same for the static case.

[4] Baron and Besanko (1984) use a similar structure in one of their cases. In their work they have two continuous random variables: the first period type θ_1 and a shock ε. The second period type is a convex combination of θ_1 and ε, i.e., $\theta_2 = \gamma\theta_1 + (1 - \gamma)\varepsilon$. The parameter γ has analogous interpretation of our μ.

For future reference define

$$W^{FI}(\nu) = S - (1+\lambda)[\nu\theta_L + (1-\nu)\theta_H - e^* + \psi(e^*)] \tag{3.8}$$

as the one period expected welfare achieved in the first best contract when the regulator's belief that $\theta_\tau = \theta_L$ is ν.

3.2.2 Dynamics with Commitment

In this section, we assume that the regulator commits himself to a long-term scheme. The regulator will not use any information learned after the first period to modify the contract, even in the case that the contract in course is not optimal.

Baron and Besanko (1984), Laffont and Tirole (1986) and Roberts (1983) showed that, under commitment, the optimal contract in the type independent case, is to follow the optimal static contract in the first period and to set the first-best allocation in the second period. Therefore, the welfare in the commitment situation (with $\mu = 0$) is

$$W^{AI}(\nu_1) + \delta W^{FI}(\nu_2) \tag{3.9}$$

And, in the full persistence case, the optimal contract is the repetition of the static contract taking into account the beliefs in each period. Therefore, the welfare in the commitment situation (with $\mu = 1$) is

$$W^{AI}(\nu_1) + \delta W^{AI}(\nu_1) \tag{3.10}$$

where $W^{AI}(\nu)$ is the welfare in a one-period assimetric information setup when the regulator has a prior ν and $W^{FI}(\nu)$ is the welfare in a one-period full information setup. In the next section we derive it.

The static setup

We now calculate the optimal contract in the static setup. In this setup, a contract is a vector (t_L, C_L, t_H, C_H) conditional on types revealed.

Formally, because only the transfer t and the level of cost C are observable, the contract must be a function $t(C)$ that specifies a monetary transfer contingent on a cost level observed. But we can, with no loss of generality by the revelation principle, design a contract based on the announced type. So the contract would be $(t(\theta), C(\theta))_{\theta=\theta_L,\theta_H}$. Defining $t_i \equiv t(\theta_i)$ and $C_i \equiv C(\theta_i)$ for $i = L, H$ we have the usual notation.

The efforts associated with this contract are $e_L = \theta_L - C_L$ and $e_H = \theta_H - C_H$. For

expositional convenience, we can rewrite the problem in function of efforts.[5]

In order to guarantee the participation of firm, the regulator must transfer enough to overcome the firm's outside opportunity, which we normalize to 0. These are the individual rationality constraints, referred by (IR_L) and (IR_H).

Additionally, the regulator must design the contract so that the firm chooses exactly the one expected to him. In other words, the firm with type θ_i should prefer the contract (t_i, C_i) instead of (t_j, C_j) for $i, j = L, H$; $i \neq j$. These are incentive compatibility constraints, refered by (IC_L) and (IC_H).

The set of constraints faced by the regulator is:

$$\begin{aligned} U_L = t_L - \psi(e_L) &\geq t_H - \psi(e_H - \Delta\theta) \ (IC_L) \\ U_H = t_H - \psi(e_H) &\geq t_L - \psi(e_L + \Delta\theta) \ (IC_H) \\ U_L = t_L - \psi(e_L) &\geq 0 \ (IR_L) \\ U_H = t_H - \psi(e_H) &\geq 0 \ (IR_H) \end{aligned}$$

Note that (IC_L) and $(IR_H) \Rightarrow (IR_L)$, since $U_L = t_L - \psi(e_L) \geq t_H - \psi(e_H - \Delta\theta) \geq t_H - \psi(e_H) = U_H \geq 0$. The first inequality is the (IC_L), the second follows from the fact that ψ is increasing and the last is (IR_L), so we can ignore this last constraint. We can rewrite (IC_L) as $U_L \geq U_H + \Phi(e_H)$, where $\Phi(e) \equiv \psi(e) - \psi(e - \Delta\theta)$.[6] This function plays an important role in the analysis. As we will further see, it defines the rent gotten by the efficient firm and represents the saving in disutility achieved by the efficient firm when the regulator thinks that he is inefficient.

Definition 3 ***(One period asymmetric information problem)*** *Given a prior ν of the firm being efficient, the regulator under static asymmetric information (second best) solves*

$$\max_{(t_L, e_L, t_H, e_H)} S - \nu\{(1+\lambda)[\theta_L - e_L + \psi(e_L)] + \lambda U_L\} - (1-\nu)\{(1+\lambda)[\theta_H - e_H + \psi(e_H)] + \lambda U_H$$

subject to (IC_L), (IC_H), (IR_L) and (IR_H).

The convexity of $\Phi(\cdot)$ guarantees that the regulator problem is concave. The solution to this problem involves neglecting the (IC_H) constraint and checking if it is satisfied at the solution. Since the regulator dislikes leaving any rent to the firm, the remaining constraints are binding at the optimum.

[5] Laffont and Tirole (1993) show that the optimal contract is not stochastic (see page 119 for this proof).

[6] $\Phi(\cdot)$ is increasing and convex since $\psi'' > 0$ and $\psi''' \geq 0$.

Proposition 11 *The solution of one period asymmetric information (Second Best) is given by*

$$\psi'(e_L) = 1 \qquad (e_L = e^*) \tag{3.11}$$

$$\psi'(e_H) = 1 - \frac{\lambda}{1+\lambda}\frac{\nu}{1-\nu}\Phi(e_H) \qquad (e_H < e^*) \tag{3.12}$$

and transfers $t_L = \psi(e^*) + \Phi(e_H)$ *leaving a positive rent of* $U_L = \Phi(e_H)$ *to the efficient type and* $t_H = \psi(e_H)$ *leaving a positive no rent* $U_H = 0$ *to the inefficient type.*

Proof: See Laffont and Tirole (1993). ■

The regulator faces a trade-off between rent extraction and power of the incentive scheme. To induce the efficient type to reveal itself, the regulator raises the rent left, but he decreases the effort required by the inefficient type.

These following value functions will be used extensively in the dynamic setup:

$$W^{AI}(\nu) = S - \nu[(1+\lambda)(\theta_L - e^* + \psi(e^*)) + \lambda\Phi(e_H)] - (1-\nu)(1+\lambda)(\theta_H - e_H + \psi(e_H)), \tag{3.13}$$

where e_H is defined by (14). Let $\mathcal{U}(\nu)$ be the reduced form of firm's rent, i.e., $\mathcal{U}(\nu) \equiv \Phi(e_H(\nu))$.

Once the posterior is endogenous, in some cases with high posterior the probability of the firm being efficient, it is worthwhile for the regulator to shut down the inefficient firm. For example, in the first period separation with complete type persistence, if firm reveals itself being efficient he will remain efficient with high probability. Hence, the regulator may find it worth keeping only the efficient firm in the relationship. In the proof of Proposition 4, we calculate a cut-off value (Γ) for the posterior such that for all $\nu \geq \Gamma$ it is always worthwhile for the regulator to shut down the inefficient firm. In these cases we have $\mathcal{U}(1) \equiv 0$ and $W^{AI}(1) \equiv W^{FI}(1)$.

3.3 Dynamics Without Commitment

We now focus our attention on a long run (two periods) relationship. From now on, we assume that the regulator is not able, in the first period, to commit himself to a second-period scheme. In other words, the regulator must offer the second-period contract after the entire resolution of the first period, so he can learn about firm's type and uses that information in the design of the second-period contract.

Formally, the regulator designs the second-period contract depending on the contract

chosen by the firm in the first period $t_1(C_1)$.[7] He sees the realization of cost C_1 and updates his beliefs about the firm's type. Then, in the second period, he offers a contract $t_2(C_2)$. The firm, on the other hand, knows that her choice has an influence on the regulator's belief. In equilibrium, both firm and regulator take into account the dynamic effect of their choices and make their decisions optimally.

Definition 4 ***(Strategies)*** *The regulator's strategy in this game is a pair of transfer functions (incentive scheme)*$\{t_1(C_1); t_2(C_2; t_1(\cdot), C_1)\}$ *and the firm's strategy in this game is a set of effort functions* $\{e_1(t_1(\cdot), \theta), e_2(t_1(\cdot), t_2(\cdot), \theta, C_1)\}$ *and acceptance function* $\{\chi_1(t_1(\cdot), \theta), \chi_2(t_1(\cdot), t_2(\cdot), \theta, C_1)\}$ *where* $\chi_\tau = 1$ *if the firm accepts the scheme at date* τ*, and* $\chi_\tau = 0$ *otherwise.*

The ratchet effect is a consequence of the lack of commitment. At the end of period one the regulator learns about firm's type and wants to use that information to expropriate rents. If he were able to commit not to use that information, the ratchet effect would no longer be a problem.

We model the equilibrium of this game as a Perfect Bayesian Equilibrium.

Definition 5 ***(Perfect Bayesian Equilibrium - PBE)*** *A profile* (t_1, t_2) *consists in a PBE if the following conditions are hold:*

- *Given a second period transfer schedule* $t_2(\cdot)$*, the firm's choice* $[e_2, \chi_2]$ *are optimal.*
- *the transfer schedule* $t_2(\cdot)$ *is optimal for the regulator given his updated beliefs* $(\overline{\nu}, \underline{\nu})$*.*
- $[e_1, \chi_1]$ *are optimal for the firm given first period transfer* $t_1(\cdot)$ *and taking in account its effect on regulator's posterior, through* C_1*.*
- t_1 *is optimal for the regulator taking into account its effect on all others strategies.*
- *the posteriors* $(\overline{\nu}, \underline{\nu})$ *are derived from the prior, the firm's strategies in first period and the observed cost, using Bayes rule and taking in account the possible change of type.*

Different from the static case, in this setup, we cannot rule out stochastic contracts. Additionally, we assume that the regulator offers a two-contract first period menu $(t_L^1, C_L^1, t_H^1, C_H^1)$ and in the second period $(t_L^2, C_L^2, t_H^2, C_H^2)$. Let $(x, 1-x)$ be the randomization, over first period contracts, induced for efficient type and $(y, 1-y)$ for the inefficient

[7] We cannot directly use mechanism because the revelation principle is no longer valid.

one. x is the probability, induce by the regulator, of the efficient firm chooses the first period contract (t_L^1, C_L^1) and, analogously, y is the probability of the inefficient firm chooses the same contract. Clearly, $x, y \in [0,1]$.[8]

We solve for equilibria using backward induction. The problem in the second period is analogous to the static setup developed in section 2.2.1. So, the second-period optimal contract is the optimal static contract, characterized by Proposition 2, for the updated probability.[9]

3.3.1 Updating Beliefs

After knowing the choice of the first period, the regulator will update his belief according to Bayes rule and taking into account the possible change in types. Define $\underline{\nu}$ being the posterior probability that the firm is efficient associated with the choice of C_L^1, i.e., $\underline{\nu} \equiv \Pr[\theta_2 = \theta_L | C_1 = C_L^1]$, or

$$\underline{\nu} = \mu \frac{x\nu_1}{x\nu_1 + y(1-\nu_1)} + (1-\mu)\nu_2. \tag{3.14}$$

Recall form that is the updated posterior for the original Laffont and Tirole's model. Similarly, define $\overline{\nu}$ being the posterior probability that the firm is efficient associated with the choice of C_H^1, i.e., $\overline{\nu} \equiv \Pr[\theta_2 = \theta_L | C_1 = C_H^1]$, or

$$\underline{\nu} = \mu \frac{x\nu_1}{x\nu_1 + y(1-\nu_1)} + (1-\mu)\nu_2. \tag{3.15}$$

Analogously, is the updated posterior for the original Laffont and Tirole's model.

3.3.2 Long Run Constraints

Recall from the previous section that in the second period the firm gets rent of $\mathcal{U}(\nu)$ if the regulator believes he is efficient with probability ν and zero otherwise. [10] In our

[8] Offer as many contracts as the number of types does not lead to any loss of generality, see Bester and Strausz (2004).

[9] We assume, in this section, that is never optimal shut down the inefficient type. In section 4, we relax this hypothesis.

[10] By Proposition 9.6 of Laffont and Tirole (1993) we can assume that type θ_L chooses the contract (t_L^1, C_L^1) and type θ_H chooses (t_H^1, C_H^1) with strict positive probability.

context of type dynamic, the constraints become

$$t_L^1 - \psi(\theta_L - C_L^1) + \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\underline{\nu}) \geq t_H^1 - \psi(\theta_L - C_H^1) + \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\overline{\nu}) \; (IC_L)$$
$$t_H^1 - \psi(\theta_H - C_H^1) + \delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu}) \geq t_L^1 - \psi(\theta_H - C_L^1) + \delta[(1-\mu)\nu_2]\mathcal{U}(\underline{\nu}) \; (IC_H)$$
$$t_L^1 - \psi(\theta_L - C_L^1) + \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\underline{\nu}) \geq 0 \; (IR_L)$$
$$t_H^1 - \psi(\theta_H - C_H^1) + \delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu}) \geq 0 \; (IR_H).$$

It is important to note there is a dispersion of the information in the second period. The firm after the first period knows his type and calculates the expected future rent taking in account only his actual type and the distribution of a possible second period draw. The regulator, on the other hand, instead of knowing the actual type uses a posterior inferred by the first-period choices.

Once the regulator is aware of the change in types, he always offers two pairs of contracts: one that extracts all rents of the inefficient type and another that leaves positive rents, with size depending on first period choice, for the efficient type.

Differently from the static setup, the incentive constraints are, in general, "upwards" and "downwards" binding. These phenomena lead the regulator to offer stochastic mechanisms. As in the static setup we can show that (IC_L) and $(IR_H) \Rightarrow (IR_L)$, so we can ignore this last constraint. Note also that for $\mu = 1$ we have the original constraints of Laffont and Tirole's model.

3.3.3 The Dynamic Problem

As we saw previously, the regulator chooses the first-period contract, specifying a transfer conditional on costs observed in the first period. Facing this contract, the firm (which at this moment knows his type) chooses the first-period level of effort. Then the first-period cost is realized, and the regulator transfers the amount prescripted by the incentive contract.

Observing the cost realization the regulator updates his beliefs about the firm's efficiency and in the second period he chooses the optimal regulation mechanism given his updated beliefs. Then the firm (which at this moment also knows his new second-period type) chooses again the level of effort and a new realization of cost takes place.

The regulator wants to maximize the expected welfare subject to long run constraints. For expositional convenience we rewrite the problem in function of efforts $e_L^1 = \theta_L - C_L^1$, $e_H^1 = \theta_H - C_H^1$.

Definition 6 ***(Regulator's problem)*** *The first period regulator's problem in the dy-*

namic setup with type dynamic is to choose $(t_L^1, e_L^1, t_H^1, e_H^1)$ *in order to*

$$\begin{aligned}
\max\ & S + \nu_1\{x[-(1+\lambda)(\theta_L - e_L^1 + t_L^1) + t_L^1 - \psi(e_L^1)] \\
& +(1-x)[-(1+\lambda)(\theta_H - e_H^1 + t_H^1) + t_H^1 - \psi(e_H^1 - \Delta\theta)]\} \\
& +(1-\nu_1)\{y[-(1+\lambda)(\theta_L - e_L^1 + t_L^1) + t_L^1 - \psi(e_L^1 + \Delta\theta)] \\
& +(1-y)[-(1+\lambda)(\theta_H - e_H^1 + t_H^1) + t_H^1 - \psi(e_H^1)]\} \\
& +\delta[\nu_1 x + (1-\nu_1)y]W^{AI}(\underline{\nu}) + \delta[\nu_1(1-x) + (1-\nu_1)(1-y)]W^{AI}(\overline{\nu})
\end{aligned}$$

subject to (IC_L), (IC_H), (IR_L) *and* (IR_H)

3.3.4 Equilibrium Characterization

As we saw previously, unlike the standard models the constraint for inefficient type (IC_H) can also be binding. There are several kinds of equilibria based on which constraints are binding. The pooling equilibrium is an extreme case of any of these equilibria. In Appendix A, we write the welfare function for all kinds of equilibria and we maximize with respect to efforts.

Type I Equilibrium

In a type I equilibrium, (IC_L) and (IR_H) contraint are binding. The type θ_H has no incentive to lie and choose the contract (t_H^1, C_H^1) and the type θ_L randomizes between the two contracts. In other words, $x > 0$ and $y = 0$. In equilibrium transfers are $t_L^1 = \psi(e_L^1) + \Phi(e_H^1) - \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\underline{\nu}) + \delta\mu\mathcal{U}(\overline{\nu})$ and $t_H^1 = \psi(e_H) - \delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu})$. The regulator updates his beliefs to

$$\underline{\nu} = \mu + (1-\mu)\nu_2 \tag{3.16}$$

$$\overline{\nu} = \mu\{\frac{(1-x)\nu_1}{1-x\nu_1}\} + (1-\mu)\nu_2 \tag{3.17}$$

We interpret the welfare function (see Appendix A) in the following way. With probability $\nu_1 x$ type θ_L reveals itself to being efficient in the first period (by choosing the contract (t_L^1, C_L^1)) and then get rent of $\Phi(e_H^1)$ plus the amount $\delta\mu\mathcal{U}(\overline{\nu})$ that compensates the firm in period 1 for not lying (choosing the contract (t_H^1, C_H^1)).

To induce the efficient firm to separate, the regulator must increase his transfer in the first period and maintain the (IC_L) constraint satisfied. In a type I equilibrium this transfer is not high enough to make the (IC_H) constraint violated. To do so he increases the static rent by the value $\delta[\mu + (1-\mu)\nu_2][\mathcal{U}(\overline{\nu}) - \mathcal{U}(\underline{\nu})]$, which is the extra rent (positive or negative) get in the second period in case of lying. However, the regulator can save

resources (comparing the static case) decreasing all transfers by the extra second period rent get by the inefficient firm $\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu})$ keeping all incentive constraints satisfied without violating any participation constraints.

With probability $\nu_1(1-x)+(1-\nu_1)$ the firm chooses (t_H^1, C_H^1) and the regulator doesn't know for sure if the firm is efficient or not. With probability $\nu_1(1-x)$ the firm is lying, implements the cost C_H with a economy in disutility of effort and gets rent of $\Phi(e_H^1)+\delta\mu\mathcal{U}(\overline{\nu})$. With probability $(1-\nu_1)$ the firm with type θ_H reveals itself to be inefficient and gets no rent.

In order to guarantee the inefficient firm participation, the regulator must compensate the disutility incurred with the contract. Then he must transfer more than $\psi(e_H)$. But he can use the future expected rent $\delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu})$ for this. As we saw, since, in this equilibrium the inefficient firm do not want to lie about his type, the regulator can decrease all rents by this amount.

In the original Laffont and Tirole's model the amount of compensation is $\delta\mathcal{U}(\nu^1)$ different from the one of our model. The efficient firm must be compensated by the extra rent that he would have if he had lied about being efficient in both periods. This rent is a function of the regulator's belief of firm being efficient, i.e. $\overline{\nu}=\mu\frac{(1-x)\nu_1}{1-x\nu_1}+(1-\mu)\nu_2$. On the other hand, the regulator knows that even if the firm reveals to be efficient with a probability of $(1-\mu)(1-\nu_2)$ the firm will be inefficient in the second period. So the regulator may not transfer the amount of $\mathcal{U}(\underline{\nu})$.

The second period welfare is analogous to the static welfare for the updated probability (see equation (16)). The optimal first period effort in a type I equilibrium is characterized by

$$\psi'(e_L^1)=1, \tag{3.18}$$

$$\psi'(e_H^1)=1-\frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\Phi'(e_H^1)+\frac{\nu_1}{1-\nu_1}(1-x)[1-\psi'(e_H-\Delta\theta)]. \tag{3.19}$$

The efficient firm achieves the first best level of effort and the effort level of the inefficient firm is distorted above the static second best level (see equation 13).

Note that the lack of commitment and the possibility of changing types allows the regulator to expropriate future rents by decreasing first period transfers. In the extreme case of independent types, the second period rent is known a priori, and the regulator can expropriate it without violate any constraint.

Type II Equilibrium

In a type II equilibrium, only the (IC_H) and (IR_H) contraint are binding. The type θ_L has no incentive to lye and choose the contract (t_L^1, C_L^1) and type θ_H randomizes between the two contracts. In other words, $x = 1$ and $y \geq 0$. Transfers are $t_L^1 = \psi(e_L^1 + \Delta\theta) - \delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu})$ and $t_H^1 = \psi(e_H^1) - \delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu})$. The regulator updates his beliefs to

$$\underline{\nu} = \mu\frac{\nu_1}{\nu_1 + y(1-\nu_1)} + (1-\mu)\nu_2 \tag{3.20}$$

$$\overline{\nu} = (1-\mu)\nu_2 \tag{3.21}$$

We interpret the welfare function (see appendix A) in the following way: With probability ν_1 the type θ_L reveals itself to being efficient in the first period (by choosing the contract (t_L^1, C_L^1)) and then get rent of $\Phi(e_L^1 + \Delta\theta) - \delta\mu\mathcal{U}(\underline{\nu})$. This amount induces the inefficient firm to separate since he does an effort of $e_L^1 + \Delta\theta$ to achieve cost C_L^1. This transfer maintains the (IC_H) constraint satisfied.

To induce the efficient firm to separate, the regulator increase his transfer in the first period to make the (IC_L) constraint slack. In a type II equilibrium, this transfer is high enough to make the (IC_H) constraint binding. To do so he increases the static rent by increasing the required effort for the inefficient type until the level required to achieve cost C_L^1. However, the regulator can save resources (comparing the static case) decreasing the efficient transfers by the extra second-period rent get by the inefficient firm if he had lie $\delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu})$ keeping all incentive constraints satisfied without violating any participation constraints.

The second period welfare is analogous to the static welfare for the updated probability (see equation (16)) and the optimal first-period effort in type II equilibrium is

$$\psi'(e_L^1) = 1 - \frac{\lambda}{1+\lambda}\Phi'(e_L^1 + \Delta\theta) + \frac{1-\nu_1}{\nu_1}y[1 - \psi'(e_L + \Delta\theta)], \tag{3.22}$$

$$\psi'(e_H^1) = 1. \tag{3.23}$$

The inefficient firm achieves the first best level of effort and the effort level of the efficient firm is distorted below first best level (see equation (13)).

Type III Equilibrium

In a type III equilibrium, (IC_L), (IC_H) and (IR_H) are binding. Both types randomize between the two contracts. In other words, $x > 0$ and $1 > y \geq 0$ and we have an

implied constraint (see Appendix A)

$$\Phi(e_L^1 + \Delta\theta) + \delta\mu\mathcal{U}(\underline{\nu}) = \Phi(e_H^1) + \delta\mu\mathcal{U}(\overline{\nu}) \tag{3.24}$$

The regulator updates his beliefs to

$$\underline{\nu} = \mu\frac{x\nu_1}{x\nu_1 + y(1-\nu_1)} + (1-\mu)\nu_2 \tag{3.25}$$

$$\overline{\nu} = \mu\frac{(1-x)\nu_1}{(1-x)\nu_1 + (1-y)(1-\nu_1)} + (1-\mu)\nu_2 \tag{3.26}$$

In a type III equilibrium all the effects elucidated previously occurs. The changes in transfers that compensates the firm in period 1 for not lying (choosing the contract (t_H^1, C_H^1)) is high enough to make the inefficient firm indifferent between telling the truth (choosing the contract (t_H^1, C_H^1)) and lying about being efficient (choosing the contract (t_L^1, C_L^1)). Therefore, both (IC_L) and (IC_L) are binding.

The optimal effort in type III equilibrium is

$$\psi'(e_L^1) = 1 + \frac{1-\nu_1}{\nu_1}\frac{y}{x}[1 - \psi'(e_L + \Delta\theta)] - [\frac{\gamma}{\nu_1 x(1+\lambda)} + \frac{\lambda}{1+\lambda}]\Phi'(e_L^1 + \Delta\theta)$$

$$\psi'(e_H^1) = 1 + \frac{\nu_1}{1-\nu_1}\frac{1-x}{1-y}[1 - \psi'(e_H - \Delta\theta)] + [\frac{\gamma}{(1-\nu_1)(1-y)(1+\lambda)} - \frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\frac{1-x}{1-y}]\Phi'(e_H^1)$$

where γ is the Lagrange multiplier associated to the restriction (28). The direction of distortion depends on the signal the Lagrange multiplier, which we were not able to determine.

3.4 Properties of the Optimal Contract

In this section, we explore some properties of the equilibrium in our model. We are, in particular, interested in understanding how the degree of persistence influences the optimal contract.

In our simulations, see Section 4.4, we take into account the possibility of shutting down the inefficient firm. As we saw in Section 2.1.1, in some circumstances it may be worthwhile for the regulator to shut down the inefficient firm. However, the cut-off value for the beliefs that defines if the regulator will or won't keep both firms is dependent on the degree of persistence μ. As we see in the next proposition, for μ small, this is not an issue for the regulator.

Proposition 12 *There exists $\tilde{\mu} \in (0,1)$ such that the regulator never shutdown the inefficient firm in the second period for all $\mu \leq \tilde{\mu}$.*

The intuition for this result is the following. If there is enough uncertainty in the second period the regulator wants to keep both firms in the relationship, because shutdown the probability of firm being inefficient is never too small.

3.4.1 Optimal Contract with Type Independence ($\mu = 0$)

As one could see in Proposition 3 if we restrict our analysis only to the extreme case of type independence we can fully characterize the optimal contract.

We call type independence when firm's type is drawn in each period from a new independent, in the statistical sense, random variable. In our model, this happens with $\mu = 0$. In this section, we calculate the optimal contract in this situation.

Now, the regulator's beliefs are always updated to

$$\underline{\nu} = \overline{\nu} = \nu_2. \tag{3.27}$$

The long run constraints (IC_L) and (IC_H) are identical as the ones from the static setup. It suggests the ratchet effect is no longer an issue for the regulator.

The next two lemmata help us to understand the specificities of the type independent model.

Lemma 11 *In the type independent model, type I equilibrium always shows separation. Moreover, the first-period efforts are the same of the optimal static contract.*

Lemma 12 *In the type independent model, type II equilibrium is never optimal.*

Lemma 13 *In an independent type model, the type III equilibrium always shows full pooling, and the efficient type first-period effort is the same of the type II equilibrium. Lastly, type III equilibrium cannot be optimal.*

It is intuitive to see why it must hold. Since the regulator doesn't get any advantage inducing firms to randomize, the regulator must pool firms in the first period to maintain both types indifferent between the two contracts.

With the previous lemma we have proved the following:

Proposition 13 *The optimal contract in the independent type model ($\mu = 0$) achieves a higher welfare than the a repetition of the optimal static contract and lower than the commitment allocation. Precisely its welfare is given by*

$$W^{AI}(\nu_1) + \lambda\delta\nu_2\mathcal{U}(\nu_2) + \delta W^{AI}(\nu_2). \tag{3.28}$$

This result is different from the one found in Baron and Besanko (1984), Townsend (1982) and Roberts (1983). In all these papers, the second-period optimal contract achieves first best welfare, but it is not the case for our model. In fact, since the relationship has to be governed by short-term contracts, the asymmetry of the information in the moment of the design of the second-period contract moves the optimal contract away from the first-best allocation. It happens in our model because the source of asymmetric information is renewed each period.

However, the regulator can improve the welfare, comparing to the repetition of static allocation, by decreasing the first-period rent in order to appropriate the expected rent generated in the second period. With $\mu = 0$ this rent is known for sure at the moment of the design of the first-period contract. The regulator can decrease all transfer by this amount without any impact on incentives. The second-period allocation only differs from the first best allocation in the effort realized by the inefficient firm, which has to be distorted away from the efficient to separate firms.

Corollary 7 *For μ small enough, the equilibrium always shows separation.*

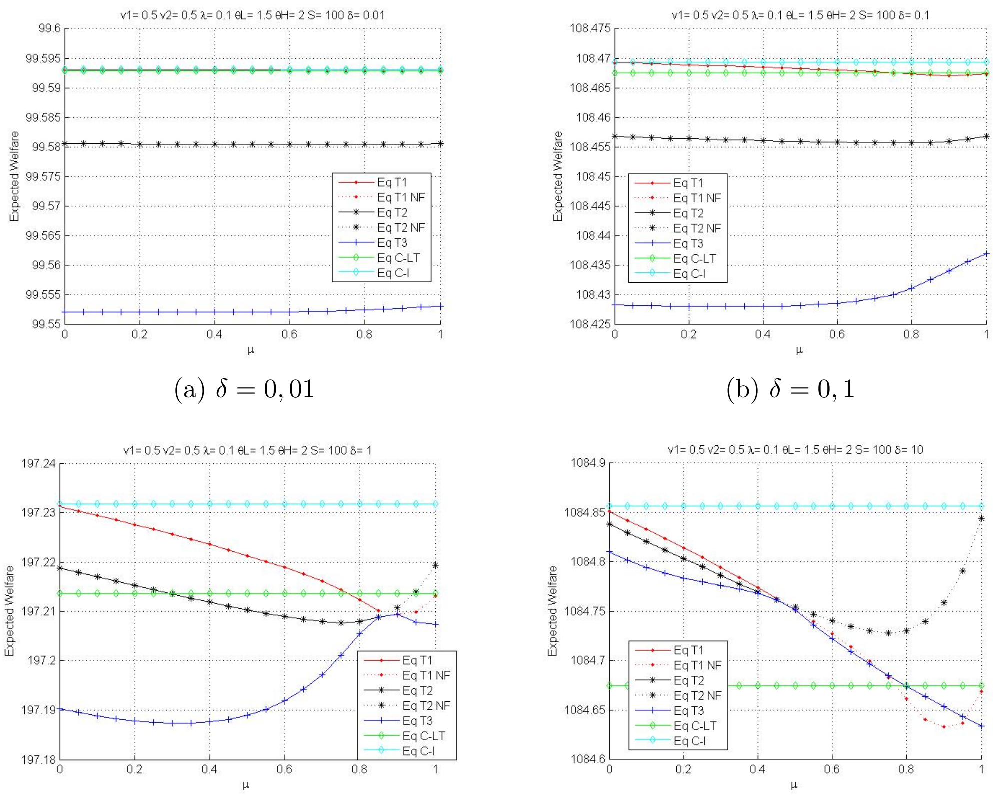
3.4.2 Numerical Simulations

Making inference about the equilibrium properties, when we have an intermediate degree of type persistence, has showed to be hard. In order to avoid these difficulties we consider a special case of quadratic effort function and we calculate the optimal contract through numerical computation. Then we make inferences about its properties. The disutility of effort takes the form $\psi(e) = \max\{0, \frac{e^2}{2}\}$.

Figure 1 shows the dependence of equilibrium expected welfare on the parameter μ for several values of discount factor δ. In Figure 2, we simulate the welfare function for several combinations of v_1 and v_2 and, in Figure 3, we make the same exercise varying $\Delta\theta$.

The green line is the welfare achived in the commitment alocation of the full persistence model (i.e. the contract that implements the static assimetric information welfare in both periods). The light-blue line is the welfare achieved in the commitment allocation of the type independent model (i.e. the contract that implements the static assimetric information welfare in first period and first best in the second). As we can see in all figures the welfare of the commitment allocation (Idependent model) is an upper bound to any non-commitment allocation for the entery range of μ. The red line is the welfare in a type I equilibrium; the line becomes dashed when this equilibrium is no longer feasible, i.e., when the incentive constraint for the inefficient firm is not satisfied. The black line

is the welafre in a type II equilibrium. Finally, the blue line is the welfare in a type III equilibrium allocation.



(a) $\delta = 0,01$

(b) $\delta = 0,1$

(c) $\delta = 1$

(d) $\delta = 10$

Figure 3.1: Comparative statics on δ.

Whenever $\nu_1 \geq \nu_2$ the welfare achieved in equilibrium is greater as smaller is the value of μ. For $\nu_1 < \nu_2$ the opposite occurs. In fact, for small persistence the equilibrium is always a type I equilibrium and the optimal welfare is very close to the upper bound (the commitment allocation). Our simulations allow us to make some inferences about the properties of equilibrium. In particular, we conjecture the following statements:

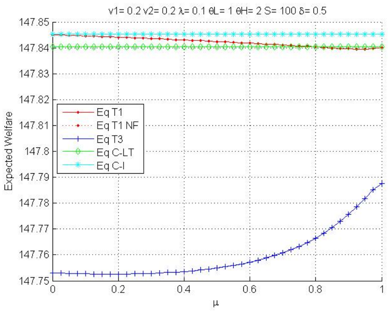
Conjecture 2 *There exists $\overline{\mu} \in [0,1]$, such that, for all $\mu > \overline{\mu}$, type I equilibrium is not feasible.*

It is important to note that if $\overline{\mu} = 1$, type I equilibrium is always feasible.

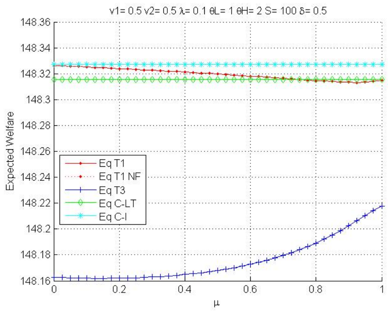
Conjecture 3 *Type II equilibrium is never optimal.*

Conjecture 4 *Type I equilibrium dominates type III equilibrium whenever it is feasible.*

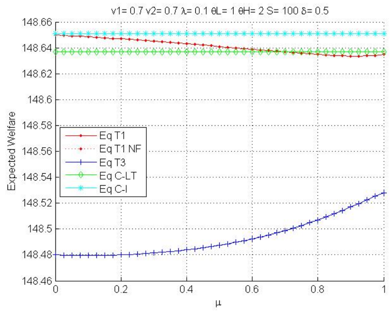
Now with table 1 and table 2 we can see how the parameter μ affects the equilibrium allocation and each variable for several values of δ (table 1) and $\Delta\theta$ (table 2).



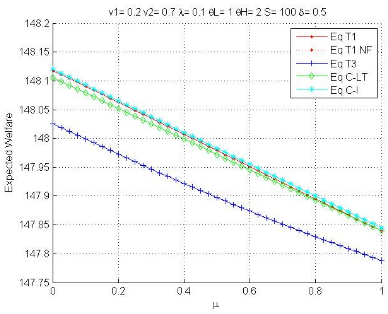
(a) $\nu_1 = 0,2$; $\nu_2 = 0,2$



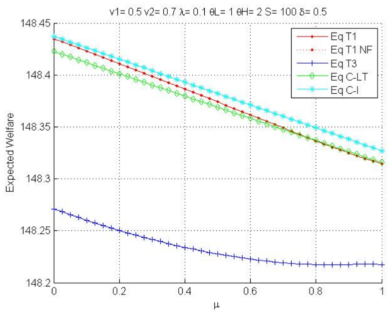
(b) $\nu_1 = 0,5$ $\nu_2 = 0,5$



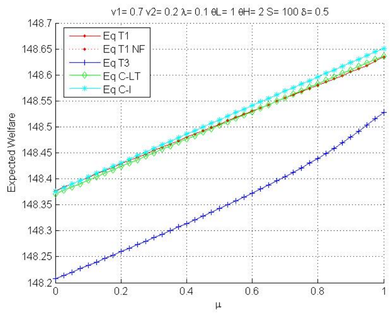
(c) $\nu_1 = 0,7$ $\nu_2 = 0,7$



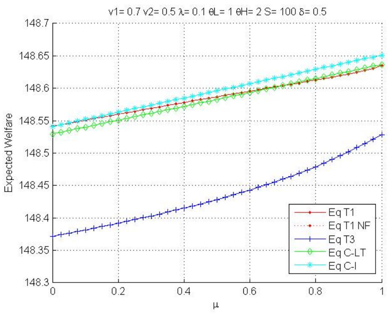
(d) $\nu_1 = 0,2$ $\nu_2 = 0,7$



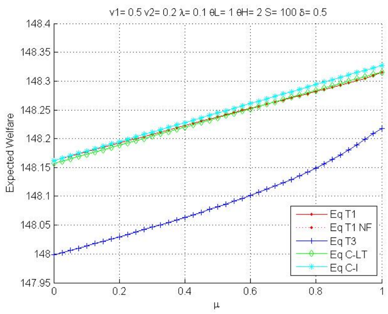
(e) $\nu_1 = 0,5$ $\nu_2 = 0,7$



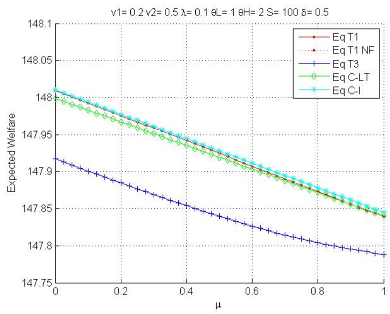
(f) $\nu_1 = 0,7$ $\nu_2 = 0,2$



(g) $\nu_1 = 0,7$ $\nu_2 = 0,5$

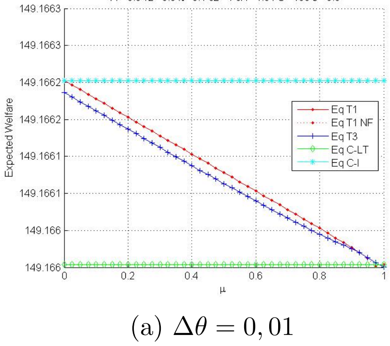


(h) $\nu_1 = 0,5$ $\nu_2 = 0,2$

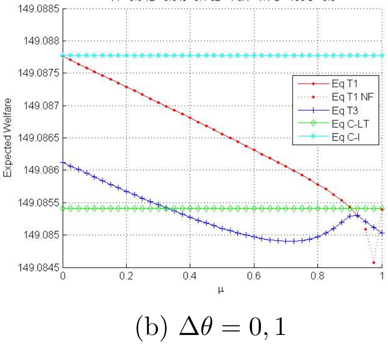


(i) $\nu_1 = 0,2$ $\nu_2 = 0,5$

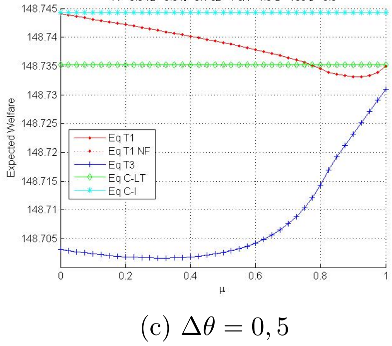
Figure 3.2: Comparative statics on ν_1 and ν_2.



(a) $\Delta\theta = 0,01$



(b) $\Delta\theta = 0,1$



(c) $\Delta\theta = 0,5$

Figure 3.3: Comparative statics on $\Delta\theta$.

δ	0.01	0.1	1	10	0.01	0.1	1	10	0.01	0.1	1	10
	$\mu = 0$				$\mu = 0.5$				$\mu = 1$			
Eq. type	I	I	I	I	I	I	I	III	I	I	III	III
x	1	1	1	1	1	1	1	0	1	1	0	0.49
y	-	-	-	-	-	-	-	1	-	-	1	0.49
e_L^1	1	1	1	1	1	1	1	0.42	1	1	0.35	0.73
e_H^1	0.95	0.95	0.95	0.95	0.95	0.95	0.95	1.53	0.95	0.95	1.60	1.23
t_L^1	0.85	0.83	0.68	-0.91	0.85	0.85	0.81	-0.49	0.86	0.89	0.36	0.75
t_H^1	0.45	0.43	0.28	-1.31	0.45	0.45	0.36	0.4	0.46	0.46	1.28	0.75
Exp Welfare	99.6	108.5	197.2	1084.9	99.6	108.5	197.2	1084.8	99.6	108.5	197.2	1084.7

$\theta_L = 1.5$, $\theta_H = 2$, $S = 100$, $\nu_1 = 0.5$, $\nu_2 = 0.5$, $\lambda = 0.1$.

Table 3.1: Simulations for δ with μ fixed.

In table 1, we learned that if δ is small (e.g. $\delta = 0.01$ or $\delta = 0.1$) the equilibrium (no matter the value of μ) is always full revelation. For a greater value of δ (e.g. $\delta = 1$ or $\delta = 10$) the equilibrium becomes pooling, with double randomization, if we have a high enough persistence level. As we saw in Section 3.4 if the persistence is small the extra rent left by the regulator is not high enough to intefere in the choice of the inefficient type. This fact is consistent with Proposition 5. From Proposition 5 we know that for a very low μ equilibrium always shows separation.

The greater the discount factor δ the lower will be the degree of persistence μ necessary to permit pooling. For example, for $\delta = 1$, with $\mu = 1$ we have pooling but with $\mu = 0.5$ we don't. However, for $\delta = 10$ we have type III equilibrium with pooling with $\mu = 0.5$ and $\mu = 1$.

A very interesting phenomena happens with a high discount rate and independent types ($\delta = 10$ and $\mu = 0$). The regulator leaves negative transfer in the first period for both firms. It happens because, in this case, the second period rent is so high that allows to the regulator leaves negative transfer in the first period without violating the participation.

Note that, when the equilibrium type III is optimal, the effort level of the inefficient firm e_H is greater than the first best effort.

In table 2, we see that for a very low level of $\Delta\theta$ (e.g. $\Delta\theta = 0.01$ and $\Delta\theta = 0.1$) we have separation for intermediate persistence ($\mu = 0.5$) and complete pooling for full persistence ($\mu = 1$). The effort level of inefficient type exceeds the first best level only for $\Delta\theta = 0.1$ and $\mu = 1$.

The expected rent, calculated in table 1, are increasing in μ. For $\delta = 10$ the expected rend left by the regulator are $0.75; -0.045; -1.11$ for the values of μ equals to $1; 0.5; 0$ respectively. For $\delta = 1$ the expected rend left by the regulator are $0.82; 0.585; 0.48$ for the values of μ equals to $1; 0.5; 0$ respectively. For $\delta = 0.1$ and $\delta = 0.01$ the expected rent remains constant at the level $0.675; 0.65; 0.63$ for the values of μ equals to $1; 0.5; 0$

$\Delta\theta$	0.01	0.1	0.5	0.01	0.1	0.5	0.01	0.1	0.5
	$\mu = 0$			$\mu = 0,5$			$\mu = 1$		
Eq. Type	I	I	I	I	I	I	III	III	I
x	1	1	1	1	1	1	0.51	0.1	1
y	-	-	-	-	-	-	0.42	0.9	-
e_L^1	1	1	1	1	1	1	0.99	0.89	1
e_H^1	0.99	0.99	0.95	0.99	0.99	0.95	1	1.09	0.95
t_L^1	0.51	0.57	0.76	0.51	0.58	0.83	0.50	0.49	1.04
t_H^1	0.50	0.47	0.36	0.50	0.48	0.41	0.50	0.60	0.46
Exp Welfare	149.2	149.1	148.7	149.2	149.1	148.7	149.2	149.1	148.7

$\theta_L = 1$, $S = 100$, $\nu_1 = 0.5$, $\nu_2 = 0.5$, $\lambda = 0.1$, $\delta = 0.5$.

Table 3.2: Simulations for $\Delta\theta$ with μ fixed.

respectively.

The expected rent, calculated in table 2, are increasing in μ as well. For $\Delta\theta = 0.5$ the expected rend left by the regulator are $0.75; 0.62; 0.56$ for the values of μ equals to $1; 0.5; 0$ respectively. For $\Delta\theta = 0.1$ the expected rend left by the regulator are $0.545; 0.53; 0.52$ for the values of μ equals to $1; 0.5; 0$ respectively. For $\Delta\theta = 0.01$ the expected rent remains constant at the level 0.50 for all values of μ.

In short, as expected from Proposition 5, the equilibrium with $\mu = 0$ always shows separation. But even with an intermediate value of μ (e.g. $\mu = 1/2$) the pooling allocation occurs only with a very high level of discount value δ (e.g. $\delta > 10$) and we don't have pooling for any value of $\Delta\theta$. So, we need some persistence to guarantee pooling with small $\Delta\theta$. Concluding, the numerical results found in Laffont and Tirole (1993) are not robust to a low degree of type persistence.

3.5 Conclusion

In this paper, we aim to expand our understanding of dynamic regulation with the addition of one more feature to the seminal paper of Laffont and Tirole (1993). We learn that little type persistence works as if we were in commitment environment. We also explore some differences from our model in comparison with the original one. However, a precise characterization of the optimal contract had not shown to be possible and numerical simulations were necessary to more understanding of the optimal contract.

Appendix

Appendix A - Welfare functions

Type I Equilibrium

The welfare function for a type I equilibrium is

$$
\begin{aligned}
&S - \nu_1 x[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1))] - \nu_1 x\lambda\Phi(e_H^1) - \nu_1 x\lambda\delta[\mu\mathcal{U}(\overline{\nu}) - (\mu + (1-\mu)\nu_2)\mathcal{U}(\underline{\nu})] \\
&-\nu_1(1-x)[(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1 - \Delta\theta))] - \nu_1(1-x)\lambda\Phi(e_H^1) + \nu_1(1-x)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu}) \\
&-(1-\nu_1)(1+\lambda)[\theta_H - e_H^1 + \psi(e_H^1)] + (1-\nu_1)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu}) \\
&+\delta[\nu_1 x]W^{AI}(\underline{\nu}) + \delta[\nu_1(1-x) + (1-\nu_1)]W^{AI}(\overline{\nu}).
\end{aligned}
$$

Maximizing the welfare with respect to first period efforts yields

$$
\begin{aligned}
&\psi'(e_L^1) = 1 && (3.29) \\
&\psi'(e_H^1) = 1 - \frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\Phi'(e_H^1) + \frac{\nu_1}{1-\nu_1}(1-x)[1-\psi'(e_H - \Delta\theta)] && (3.30)
\end{aligned}
$$

Type II Equilibrium

The welfare function for type II equilibrium is

$$
\begin{aligned}
&S - \nu_1[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1))] - \nu_1\lambda\Phi(e_L^1 + \Delta\theta) + \nu_1\lambda\delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu}) \\
&-(1-\nu_1)y[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1 + \Delta\theta))] + (1-\nu_1)y\lambda\delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu}) \\
&-(1-\nu_1)(1-y)[(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1))] + (1-\nu_1)(1-y)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu}) \\
&+\delta[\nu_1 + (1-\nu_1)y]W^{AI}(\underline{\nu}) + \delta[(1-\nu_1)(1-y)]W^{AI}(\overline{\nu}).
\end{aligned}
$$

Maximizing the welfare with respect to first period efforts yields

$$
\begin{aligned}
&\psi'(e_L^1) = 1 - \frac{\lambda}{1+\lambda}\Phi'(e_L^1 + \Delta\theta) + \frac{1-\nu_1}{\nu_1}y[1-\psi'(e_L + \Delta\theta)] && (3.31) \\
&\psi'(e_H^1) = 1 && (3.32)
\end{aligned}
$$

Type III Equilibrium

Since both (IC_L) and (IC_H) are binding, we have

$$t_L^1 - \psi(e_L^1) + \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\underline{\nu}) = t_H^1 - \psi(e_H^1 - \Delta\theta) + \delta[\mu + (1-\mu)\nu_2]\mathcal{U}(\overline{\nu})$$
$$t_H^1 - \psi(e_H^1) + \delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu}) = t_L^1 - \psi(e_L^1 + \Delta\theta) + \delta[(1-\mu)\nu_2]\mathcal{U}(\underline{\nu})$$
$$t_H^1 = \psi(e_H^1) - \delta[(1-\mu)\nu_2]\mathcal{U}(\overline{\nu})$$

this implies

$$\Phi(e_L^1 + \Delta\theta) - \delta\mu\mathcal{U}(\underline{\nu}) = \Phi(e_H^1) - \delta\mu\mathcal{U}(\overline{\nu}). \tag{3.33}$$

The welfare function for type III equilibrium is

$$\begin{aligned}
&S - \nu_1 x(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1)) - \nu_1 x\lambda\Phi(e_L^1 + \Delta\theta) + \nu_1 x\lambda\delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu}) \\
&-\nu_1(1-x)(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1 - \Delta\theta)) - \nu_1(1-x)\lambda\Phi(e_H^1) + \nu_1(1-x)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu}) \\
&-(1-\nu_1)y(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1 + \Delta\theta)) + (1-\nu_1)y\lambda\delta(1-\mu)\nu_2\mathcal{U}(\underline{\nu}) \\
&-(1-\nu_1)(1-y)(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1)) + (1-\nu_1)(1-y)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\overline{\nu}) \\
&+\delta[\nu_1 x + (1-\nu_1)y]W^{AI}(\underline{\nu}) + \delta[\nu_1(1-x) + (1-\nu_1)(1-y)]W^{AI}(\overline{\nu})
\end{aligned}$$

s.t.

$$\Phi(e_L^1 + \Delta\theta) - \delta\mu\mathcal{U}(\underline{\nu}) = \Phi(e_H^1) - \delta\mu\mathcal{U}(\overline{\nu}). \tag{3.34}$$

Maximizing the welfare with respect to first period efforts yields

$$\psi'(e_L^1) = 1 + \frac{1-\nu_1}{\nu_1}\frac{y}{x}[1 - \psi'(e_L + \Delta\theta)] - [\frac{\gamma}{\nu_1 x(1+\lambda)} + \frac{\lambda}{1+\lambda}]\Phi'(e_L^1 + \Delta\theta)$$
$$\psi'(e_H^1) = 1 + \frac{\nu_1}{1-\nu_1}\frac{1-x}{1-y}[1 - \psi'(e_H - \Delta\theta)] + [\frac{\gamma}{(1-\nu_1)(1-y)(1+\lambda)} - \frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\frac{1-x}{1-y}]\Phi'(e_H^1),$$

where γ is the Lagrange multiplier associated with restriction.

Appendix B - Proofs

Proof of Proposition 3

Proof: Let $\tilde{\nu}$ be the regulator's belief in second period. Depends on firm's the first period choice, $\tilde{\nu} = \mu\overline{\nu} + (1-\mu)\nu_2$ or $\tilde{\nu} = \mu\underline{\nu} + (1-\mu)\nu_2$. In the second period the regulator

prefers to produce with the two types if and only if

$$\begin{aligned} &\tilde{\nu}[S-(1+\lambda)(\theta_L - e^* + \psi(e^*)) + \lambda\Phi(e_H)] \\ &+(1-\tilde{\nu})[S-(1+\lambda)(\theta_H - e_H + \psi(e_H))] \\ &\geq \tilde{\nu}[S-(1+\lambda)(\theta_L - e^* + \psi(e^*))] \end{aligned}$$

we can rewrite this condition to

$$\tilde{\nu} \leq \frac{[S-(1+\lambda)(\theta_L - e^* + \psi(e^*))]}{[S-(1+\lambda)(\theta_L - e^* + \psi(e^*))] + \lambda\Phi(e_H)}.$$

Call $\Gamma = \frac{[S-(1+\lambda)(\theta_L - e^* + \psi(e^*))]}{[S-(1+\lambda)(\theta_L - e^* + \psi(e^*))] + \lambda\Phi(e_H)}$. Note that $\tilde{\nu}$ will be maximum if $\overline{\nu} = 1$ or $\underline{\nu} = 1$. In both of this case we have the maximum value $\tilde{\nu} = \mu + (1-\mu)\nu_2$. Replacing in previous the condition, we have

$$\mu \leq \frac{\Gamma - \nu_2}{1 - \nu_2}.$$

To complete our argument, we just need to see that the left hand side of the inequality converges to zero with μ and the right hand side does not (because $\tilde{\nu}$ converges to ν_2, when μ converges to zero). ■

Proof of Lemma 1

Proof: In type I equilibrium with $\mu = 0$ the welfare function is

$$\begin{aligned} &S - \nu_1 x[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1))] - \nu_1\lambda\Phi(e_H^1) + \nu_1\lambda\delta\nu_2\mathcal{U}(\nu_2) \\ &-\nu_1(1-x)[(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1 - \Delta\theta))] \\ &-(1-\nu_1)(1+\lambda)[\theta_H - e_H^1 + \psi(e_H^1)] + (1-\nu_1)\lambda\delta\nu_2\mathcal{U}(\nu_2) + \delta W^{AI}(\nu_2) \end{aligned}$$

Maximizing the welfare with respect to first period efforts yields

$$\psi'(e_L^1) = 1 \tag{3.35}$$

$$\psi'(e_H^1) = 1 - \frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\Phi'(e_H^1) + \frac{\nu_1}{1-\nu_1}(1-x)[1-\psi'(e_H - \Delta\theta)] \tag{3.36}$$

Note that $\frac{\partial W}{\partial x} > 0$ because

$$\begin{aligned} \frac{\partial W}{\partial x} &= -\nu_1[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1))] + \nu_1[(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1 - \Delta\theta))] + \delta\frac{dW^{AI}}{d\nu}\frac{d\nu_2}{dx} \\ &= \nu_1(1+\lambda)[\theta_H - \theta_L + e_L - e_H + \psi(e_H^1 - \Delta\theta) - \psi(e_L)]. \end{aligned}$$

By the sub gradient inequality this term is positive. Therefore, the regulator will induce $x = 1$ in the optimal contract. Substituing $x = 1$ we have the same efforts of the static setup

$$\psi'(e_L^1) = 1 \tag{3.37}$$

$$\psi'(e_H^1) = 1 - \frac{\lambda}{1+\lambda}\frac{\nu_1}{1-\nu_1}\Phi'(e_H^1) \tag{3.38}$$

■

Proof of Lemma 2

Proof: The welfare function for type II equilibrium with $\mu = 0$ is

$$\begin{aligned}
&S - \nu_1[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1))] - \nu_1\lambda\Phi(e_L^1 + \Delta\theta) + \nu_1\lambda\delta\nu_2\mathcal{U}(\nu_2)\\
&-(1-\nu_1)y[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1 + \Delta\theta))] + (1-\nu_1)y\lambda\delta\nu_2\mathcal{U}(\nu_2)\\
&-(1-\nu_1)(1-y)[(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1))] + (1-\nu_1)(1-y)\lambda\delta\nu_2\mathcal{U}(\nu_2)\\
&+\delta W^{AI}(\nu_2).
\end{aligned}$$

Note that $\frac{\partial W}{\partial y} < 0$ because

$$\frac{\partial W}{\partial y} = (1-\nu_1)(1+\lambda)[\theta_H - e_H + \psi(e_H^1) - (\theta_H - (e_L + \Delta\theta) + \psi(e_L + \Delta\theta))].$$

This term is negative since $e_H = e^*$ in an optimal type II equilibrium. Therefore, the regulator will induce $y = 0$ in the optimal contract. From now the argument follow Lemma 9.1 of Laffont and Tirole (1993). ■

Proof of Lemma 3

Proof: In the type III equilibrium with $\mu = 0$ we have $\Phi(e_L^1 + \Delta\theta) = \Phi(e_H^1)$. Since $\Phi' > 0$, it implies

$$e_L^1 + \Delta\theta = e_H^1 \tag{3.39}$$

or alternatively $C_L^1 = C_H^1$.

The welfare function now is

$$
\begin{aligned}
& S - \nu_1 x(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1)) \\
& -\nu_1(1-x)(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1 - \Delta\theta)) - \nu_1\lambda\Phi(e_L^1 + \Delta\theta) + \nu_1\lambda\delta(1-\mu)\nu_2\mathcal{U}(\nu_2) \\
& -(1-\nu_1)y(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1 + \Delta\theta)) \\
& -(1-\nu_1)(1-y)(1+\lambda)(\theta_H - e_H^1 + \psi(e_H^1)) + (1-\nu_1)\lambda\delta(1-\mu)\nu_2\mathcal{U}(\nu_2) \\
& +\delta W^{AI}(\nu_2)
\end{aligned}
$$

Maximizing with respect to e_L^1

$$
\psi'(e_L^1) = 1 - \frac{\lambda}{1+\lambda}\Phi'(e_L^1 + \Delta\theta) + \frac{1-\nu_1}{\nu_1} y[1 - \psi'(e_L + \Delta\theta)]. \tag{3.40}
$$

Note that this effort is the same found in type II equilibrium.

Note that $\frac{\partial W}{\partial y} = 0$ because

$$
\begin{aligned}
\frac{\partial W}{\partial y} = & -(1-\nu_1)[(1+\lambda)(\theta_L - e_L^1 + \psi(e_L^1 + \Delta\theta))] \\
& +(1-\nu_1)[(1+\lambda)(\theta_H - (e_L^1 + \Delta\theta) + \psi(e_L^1 + \Delta\theta))] \\
& +\frac{dW^{AI}}{d\nu}\frac{d\nu_2}{dy} = 0
\end{aligned}
$$

and also $\frac{\partial W}{\partial x} = 0$.

Since randomization has no effect in the welfare, and type III equilibrium includes type I equilibrium as a special case, then the III equilibrium is always dominated by type I equilibrium. ■

Proof of Corollary 1

Proof: As we saw in Proposition 5, equilibrium with $\mu = 0$ the regulator separate two types. In other other, the optimum welfare, with $\mu = 0$ in a type I equilibrium dominates the optimum welfare in type III equilibrium. By the maximum theorem, the value function is continuous on parameter μ and also the effort function and the optimal induced randomization. Additionally as we saw in proof of Lemma 2, the derivative of the welfare function, in a type I equilibrium with $\mu = 0$, with respect to x is positive. Then, in a neighborhood of $\mu = 0$ we still have type I equilibrium dominating type III and a positive derivative on x. Proving the stated result. ■

Bibliography

Alves, C. B., C. E. da Costa, and H. Moreira (2017a): "Couples Taxation in a multidimensional Mirrleesian economy," .

——— (2017b): "On the Variational Approach to the Analysis of Tax System: A Cautionary Note," .

Apps, P., and R. Rees (2011): "Optimal taxation and tax reform for two-earner households," *CESifo Economic Studies*, 57(2), 283–304.

Araújo, A., and H. Moreira (2010): "Adverse selection problems without the Spence-Mirrlees condition," *Journal Economic Theory*, 145(3), 1113–1141.

Araujo, A., H. Moreira, and S. Vieira (2015): "The marginal tariff approach without single-crossing," *Journal of Mathematical Economics*, 61, 166 – 184.

Armstrong, M., and J.-C. Rochet (1999): "Multi-dimensional screening:: A user's guide," *European Economic Review*, 43(4), 959–979.

Baron, D., and D. Besanko (1984): "Regulation and information in a continuing relationship," *Information Economics and Policy*, 1(3), 267–302.

Battaglini, M. (2005): "Long-term contracting with markovian consumers," *American Economic Review*, pp. 637–658.

Bester, H., and R. Strausz (2004): "Contracting with imperfect commitment and the revelation principle: the single agent case," *Econometrica*, 69(4), 1077–1098.

Boskin, M. J., and E. Sheshinski (1983a): "Optimal tax treatment of the family: Married couples," *Journal of Public Economics*, 20(3), 281–297.

——— (1983b): "Optimal tax treatment of the family: Married couples," *Journal of Public Economics*, 20(3), 281–297.

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1(1), 451–488.

Chiappori, P.-A., and C. Meghir (2014): "Intrahousehold Inequality," Working Paper 20191, National Bureau of Economic Research.

Choné, P., and S. Gauthier (2017): "Optimal rationing within a heterogeneous population," *Journal of Public Economic Theory*, 19(3), 732–738.

Courty, P., and H. Li (2000): "Sequential screening," *Review of Economic Studies*, 67(4), 697–717.

Cremer, H., J.-M. Lozachmeur, and P. Pestieau (2012): "Income taxation of couples and the tax unit choice," *Journal of Population Economics*, 25(2), 763–778.

Dahlby, B. (1998): "Progressive taxation and the social marginal cost of public funds," *Journa of Public Economics*, 67, 105 — 122.

Diamond, P. A. (1998): "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates," *American Economic Review*, 88(1), 83–95.

Farhi, E., and X. Gabaix (2015): "Optimal taxation with behavioral agents," Discussion paper, National Bureau of Economic Research.

Feldstein, M. (1999): "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics*, 81(4), 674–680.

Goldman, M. B., H. E. Leland, and D. S. Sibley (1984): "Optimal Nonuniform Prices," *The Review of Economic Studies*, 51(2), 305–319.

Golosov, M., A. Tsyvinski, and N. Werquin (2014): "A Variational Approach to the Analysis of Tax Systems.," Mimeo. Princeton University.

Gomes, R., J.-M. Lozachmeur, and A. Pavan (2017): "Differential taxation and occupational choice," *The Review of Economic Studies*, p. rdx022.

Gomes, R. D., J.-M. Lozachmeur, and A. Pavan (2014): "Differential taxation and occupational choice," .

Gordon, R. H., and W. Kopczuk (2014): "The choice of the personal income tax base," *Journal of Public Economics*, 118, 97–110.

Guner, N., R. Kaygusuz, and G. Ventura (2012): "Taxation and household labour supply," *The Review of economic studies*, 79(3), 1113–1149.

Hammond, P. J. (1979): "Straightforward Individual Incentive Compatibility in Large Economies," *Review of Economic Studies*, 46(0), 263–282.

Hammond, P. J. (1987): “Markets as Constraints: Multilateral Incentive Compatibility in Continuum Economies,” *Review of Economic Studies*, 54(3), 399–412.

Hendel, I., and A. Lizzeri (1997): “Adverse selection in durable goods markets,” Discussion paper, National Bureau of Economic Research.

Hendren, N. (2013): “The Policy Elasticity,” Working Paper 19177, National Bureau of Economic Research.

Jacquet, L., and E. Lehmann (2016): “Optimal taxation with heterogeneous skills and elasticities: Structural and sufficient statistics approaches,” in *Working Paper*.

Kennan, J. (2001): “Repeated bargaining with persistent private information,” *Review of Economic Studies*, 68(4), 719–755.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009): “The optimal income taxation of couples,” *Econometrica*, 77(2), 537–560.

Kusuda, Y. (2006): “Dynamic Contract Model under Asymmetric Information,” Ph.D. thesis, CITY UNIVERSITY OF NEW YORK.

Laffont, J. (1994): “The new economics of regulation ten years after,” *Econometrica: Journal of the Econometric Society*, pp. 507–537.

Laffont, J., and J. Tirole (1986): “Using cost observation to regulate firms,” *The Journal of Political Economy*, pp. 614–641.

——— (1987): “Comparative statics of the optimal dynamic incentive contract,” *European Economic Review*, 31(4), 901–926.

——— (1988): “The dynamics of incentive contracts,” *Econometrica: Journal of the Econometric Society*, pp. 1153–1175.

——— (1993): *A theory of incentives in procurement and regulation.* the MIT Press.

Lehmann, E., L. Simula, and A. Trannoy (2014): “Tax Me if You Can! Optimal Nonlinear Income Tax between Competing Governments,” *Quarterly Journal of Economics*, 129(4), 1995 – 2030.

Lundberg, S., and R. A. Pollak (1996): “Bargaining and distribution in marriage,” *The journal of economic perspectives*, 10(4), 139–158.

Mankiw, N. G., M. Weinzierl, and D. Yagan (2009): “Optimal Taxation in Theory and Practice,” *The Journal of Economic Perspectives*, 23(4), 147–174.

MANSER, M., AND M. BROWN (1980): "Marriage and household decision-making: A bargaining analysis," *International economic review*, pp. 31–44.

MCELROY, M. B., AND M. J. HORNEY (1981): "Nash-bargained household decisions: Toward a generalization of the theory of demand," *International economic review*, pp. 333–349.

MILGROM, P., AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

MIRRLEES, J. (1971a): "An exploration in the theory of optimum income taxation," *The review of economic studies*, 38(2), 175–208.

MIRRLEES, J. A. (1971b): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38(2), pp. 175–208.

NOLDEKE, G., AND L. SAMUELSON (2007): "Optimal bunching without optimal control," *Journal of Economic Theory*, 134(1), 405 – 420.

PAVAN, A., I. SEGAL, AND J. TOIKKA (2011): "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization and Information Disclosureo," Discussion paper, mimeo Northwestern University and Stanford University.

PIKETTY, T. (1997): "La redistribuition fiscale face au chômage," *Revue Française d' Èconomie*, 12, 157 – 201.

PIKETTY, T., E. SAEZ, AND S. STANTCHEVA (2014): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 6(1), 230–71.

RAMOS, P. S. (2011): "Regulacao Dinamica de Monopolios," Discussion paper, Dissertacao de Mestrado Puc-Rio.

ROBERTS, K. (1983): "Long term contracts," *Mimeo.*

ROCHET, J.-C. (1987): "A necessary and sufficient condition for rationalizability in a quasi-linear context," *Journal of Mathematical Economics*, 16(2), 191 – 200.

ROCHET, J.-C., AND P. CHONÉ (1998): "Ironing, sweeping, and multidimensional screening," *Econometrica*, pp. 783–826.

ROCHET, J.-C., AND L. A. STOLE (2003): "The economics of multidimensional screening," *Econometric Society Monographs*, 35, 150–197.

ROTHSCHILD, C., AND F. SCHEUER (2013): "Redistributive taxation in the roy model," *The Quarterly Journal of Economics*, 128(2), 623–668.

Sachs, D., A. Tsyvinski, and N. Werquin (2016): "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," Working Paper 22646, National Bureau of Economic Research.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–229.

Scheuer, F., and I. Werning (2015): "The Taxation of Superstars," Working Paper 21323, National Bureau of Economic Research.

SchottmÃČÅŠller, C. (2015): "Adverse selection without single crossing: Monotone solutions," *Journal of Economic Theory*, 158, 127 – 164.

Sheshinski, E. (1972): "The Optimal Linear Income Tax," *Review of Economic Studies*, 39, 297 – 302.

Townsend, R. (1982): "Optimal multiperiod contracts and the gain from enduring relationships under private information," *The Journal of Political Economy*, pp. 1166–1186.