UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENERÍA INDUSTRIAL

A GENERAL APPROACH FOR SCREENING PROBLEMS WITHOUT THE SINGLE-CROSSING PROPERTY

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

TIBOR ALEJANDRO HEUMANN EPSTEIN

SANTIAGO DE CHILE
NOVIEMBRE 2011

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MÁTEMATICAS
DEPARTAMENTO DE INGENERÍA INDUSTRIAL

A GENERAL APPROACH FOR SCREENING PROBLEMS WITHOUT THE SINGLE-CROSSING PROPERTY

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

TIBOR ALEJANDRO HEUMANN EPSTEIN

PROFESOR GUIA:
NICOLÁS FIGUEROA GONZÁLEZ

MIEMBROS DE LA COMISION:
JUAN ESCOBAR CASTRO
ROBERTO COMINETTI COTTI-COMETTI
JUAN PABLO MONTERO AYALA

SANTIAGO DE CHILE
NOVIEMBRE 2011

Resumen

En este trabajo consideramos el modelo clásico de diseño de mecanismos, con un principal, que debe tomar una decisión o determinar la asignación de un bien, y agentes, que poseen información privada que es relevante para el principal. Para utilizar de manera óptima la información de los agentes, el principal diseña un menú de contratos, donde cada uno especifica la decisión que tomará el principal y las transferencias que se le darán al agente. Dado este menú, cada agente elige el contrato que más le favorece.

El objetivo del principal es diseñar un menú de contratos que maximice su bienestar, que puede coincidir o no con el bienestar social. Existe una amplia literatura que considera este problema, sin embargo la mayor parte de ésta toma como suposición fundamental que las preferencias de los agentes satisfacen la propiedad de corte único (S.C.P. por sus siglas en ingles). Esta propiedad nos garantiza que la valoración marginal de los agentes por el bien en cuestión cambia monótonamente con su información privada. Para el principal esto simplifica significativamente el diseño del menú óptimo, ya que garantiza que el problema de maximización que enfrentan los agentes, al elegir el contrato que más les favorece, es un problema de maximización cóncavo. Como los agentes enfrentan un problema cóncavo, el principal, al diseñar el menú de contratos, solo debe preocuparse localmente de la condición de primer y segundo orden de los agentes.

En esta tesis consideramos el caso en que las preferencias de los agentes no satisfacen S.C.P.. Desde un punto de vista técnico, al relajar este supuesto, se pierde la monotonicidad en las preferencias de los agentes. Esto hace que para el principal no sea suficiente analizar las condiciones de primer y segundo orden de los agentes, y deba analizar la decisión de cada agente globalmente. Por esto, el problema de maximización para el principal es mucho mas complejo de analizar ya que no basta con maximizar localmente los contratos para cada agente, si no que se debe considerar los efectos globales de cada contrato.

En esta tesis, se introduce la condición de “doble cruce”, que es un supuesto mas débil que S.C.P.. Asi, se encuentran condiciones necesarias para que un mecanismo sea implementable y también condición necesarias para la optimalidad de éste. Estas condiciones son interpretadas desde un punto de vista económico, lo que permite extender las intuiciones a una generalidad de problemas en que no se cumple S.C.P. y entender las limitaciones que impone este supuesto.

Por otro lado, ocupando las condiciones necesarias, encontramos un nuevo método para solucionar y encontrar contratos óptimos en el modelo que estudiamos. Este método permite transformar un problema de dimensión infinita en un problema bidimensional, que se puede solucionar. Ejemplificamos el método propuesto resolviendo dos ejemplos. La primera situación consiste en encontrar la forma optima de arrendar una tecnología que queda obsoleta en el tiempo a una tasa desconocida para el principal. La segunda es como regular la tecnología que usa un monopolio que produce externalidades negativas, pero cuya eficiencia para implementar distintas tecnologías es información privada del monopolio.

Summary

In this thesis we consider the classic mechanism design model with a principal, that must make a decision or determine the allocation of some good, and agents, which possess private information that is relevant for the principal. To make the best use of the agents private information the principal designs a menu of contracts, each of them specifying the decision the principal will make and transfers that will be made to the agent. Given this menu of contracts the agents chooses the one that benefits him the most.

The objective of the principal is to design a menu of contracts that maximizes his welfare, which might or might not agree with the social welfare. The literature on the subject is extensive, nevertheless the great majority is under the assumption that the agents preferences satisfy the Single Crossing Property (S.C.P.). This property guarantee that the marginal utility of the agents for the good that is being contracted upon changes monotonically with the private information. This simplifies the design of the optimal menu of contracts for the principal since it guarantees that the agents will face a concave problem when choosing the contract that benefits them the most, this in turn allows the principal to only consider the first and second order condition of the agents maximization problem when designing the menu.

In this thesis we consider the problem in which agents preferences do not satisfy the S.C.P. From a technical point of view, relaxing this assumption implies that the monotonicity in the agents preferences is lost. Thus, for the principal it is no longer sufficient to analyze only the first and second order condition that the agents will face, but the decision of each agent must be analyzed in a global way. This makes the principal's maximization problem much harder since it is not enough to analyze the design of the menu locally, but global effects must be taken in consideration for each contract.

In this thesis the Double Crossing Property is introduced, which is a weaker assumption than the S.C.P. Necessary condition are found for the implementability and optimality of a menu of contracts. This condition are interpreted from a economic standpoint, which allows to extend the intuitions to other problem in which the S.C.P. is not fulfilled and understand the limitations that this assumption entail.

On the other hand, using the necessary conditions, a method is found that allows to find the optimal menu in the model we study. This method allows to transform an infinite dimension problem in a two dimensional problem, which can be solved. Finally the method is exemplified by solving two examples. The first examples consists in finding the optimal way to lease a technology that becomes obsolete at a time rate that is unknown to the principal. The second example is how to regulate a monopoly that produces negative externalities, but the cost to implement different technologies, which produce different amount of externalities, is unknown to the principal.

Agradecimientos

Le agradezco a mi familia por haberme apoyado, guiado y ayudado en todo lo que he hecho durante mi vida, en especial a mis padres, que siempre me motivaron a buscar algo que me apasione, lo cual fue un factor fundamental para que decidiera estudiar economía.

Le agradezco a Nicolás Figueroa por haber hecho de esta tesis una gran experiencia, por todo lo que me ha enseñado y los innumerables consejos que me ha dado. Finalmente, sin su ayuda esta tesis no habría sido posible.

Le agradezco a todos los profesores del CEA por ser parte fundamental de mi formación como economista, en particular a Felipe Balmaceda por haberme ayudado y aconsejado en innumerables ocasiones durante el magíster.

Le agradezco a Olga Barrera y Julie Lagos por ayudarme, siempre con una gran disposición, con todos los problemas administrativos que surgieron durante el programa.

Le agradezco a todos mis compañeros del MAGCEA, no solo por ayudarme en los distintos cursos cuando lo necesite, si no que por convertir el programa de magíster en una gran oportunidad para hacer excelentes amigos.

A las instituciones que me apoyaron en la realización de la tesis: CONICYT, por otorgarme la Beca de Magíster; Vicerrectoría de Asuntos Académicos por otorgarme la Beca de Estadías Cortas.

Contents

List of Figures

1 Introduction

Most of the mechanism design literature, and bayesian games in general, assumes that the agents' preferences satisfy the single crossing property (S.C.P.). Under this assumption optimal contracts can be easily characterized, and are in general always monotonic (the same thing happens with comparative statics in bayesian games). Although this assumption is sometimes natural, this assumption involves an important loss of generality, and the lack of tools to tackle more difficult problems is an important bottleneck in the development of the field.

On top of the technical challenge, here are several reasons why understanding problems without the S.C.P. is important. There are natural mechanism design problems that do not satisfy the S.C.P., in section 3 we provide two. Quah and Strulovici [4] study and provide properties under which the sum of functions that satisfy the S.C.P. inherit the S.C.P., and show that these conditions are quite restrictive. Bernheim [2] and Bagwell and Bernheim [3] are two examples of signaling games in which, only by breaking the S.C.P., interesting phenomena arises. Finally, it is worth mentioning that several of the problems and complexities that arise in a single dimensional model without S.C.P. also appear in multidimensional screening, since both types of problems need to deal with non-local incentive compatibility constraints . Thus, we believe that understanding mechanism design problems that do not satisfy the S.C.P. is a fundamental step-stone to understand multidimensional screening.

If the S.C.P. is satisfied, we can order types according to their marginal valuation of the allocation (probability of winning in a auction, quantity produced in monopoly regulation, etc.), which in turn allows to consider only local incentive compatibility conditions. Basically, the iso-profit curves, in the space of the contracted good and transfers, can be ordered between types according on how "steep" they are, therefore they always cross at most once. We study a model in which we can order types according to the concavity of the utility of the contracted good, therefore in this model the iso-profit curves always cross at most twice. Although these problem might seem similar without the S.C.P. the local incentive compatibility conditions are no longer sufficient to guarantee global incentive compatibility. This simple extension leads to contracts that are qualitatively different to the ones that arise with S.C.P., allowing to extract economic intuitions that could not be found otherwise. By giving a better understanding of this model, in which the S.C.P. is not fulfilled, we can reach a better understanding of the real implications that the assumption of the S.C.P. has on a model and how it restricts the richness of the studied phenomena.

One of the main properties that is lost when the S.C.P. is not fulfilled is the monotonicity of the optimal contract. As Araujo and Moreira [1] have shown, we can have optimal policies with a U-Shaped form. Moreover, we show that without the S.C.P. there are two types of distortions that can be identified. The first one, which is standard, is when some type θ is distorted because it is part of some active I.C.C. (either θ is indifferent between his assignment and the assignment of some type $\tilde{\theta}$ or vice versa), we call this type of distortion a direct distortion. The second one, which is only found in problems in which the S.C.P. is not satisfied, is given by types that are being distorted from the optimal assignment even though there is no active I.C.C. involving that given type, which we call an indirect distortion. The latter case is a very interesting situation in which a small variation in the assignment of some type could break the I.C.C. between other two types, and highlights one of the main difficulties tackled in this paper.

Our closest reference on the literature is Araujo Moreira [1], in which they study almost the same model, so naturally we use and extend several of their results (we also try to keep the notation as similar as possible). They find necessary conditions for optimality when there is discrete pooling between types (which is obviously a case of direct distortion) and find a specific setup in which their conditions are sufficient for optimality. In this setup the only relevant global I.C.C. that need to be taken into account to find the optimal policy arise from discrete pooling types. We are able to extended their results to the general case of direct distortions, and also derive necessary conditions for optimality when indirect distortions are optimal, which is one of our main theoretical results. Another important contribution is the proposed method to find the optimal policy in a wide range of problems, almost completely independent of the principal's objective function, which greatly generalize Araujo and Moreira's results. Our proposed method breaks down the general problem into simpler problems, and the final result gives us policies with discrete pooling, indirect distortion and direct distortion by types that are not being pooled. Both examples in section 3 are solved using this method.

2 Model and Standard Results

The model is a basic principal-agent relationship. Agent's utility and the principal's utility depend on a decision to be made by the principal, indexed by $x \in [\underline{x}, \bar{x}] \subseteq R$, and some information privately know by the agent, indexed by $\theta \in [\underline{\theta}, \bar{\theta}] \subset R$. Transfers can be made between the principal and the agent, indexed by $t \in R$, which are linear in the utility of the agent and principal. The principal has a prior distribution on the agents private information, given by $P(\theta)$ with $P(\underline{\theta}) = 1 - P(\bar{\theta}) = 0$ and density $p(\theta) > 0 \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]$. The principal must design a menu of contracts that maximizes its expected utility.

Agents utility is quasi-linear in transfers, and is given by $v(x,\theta) + t$, agent's have a reserve utility of 0 and the principal's utility has the form $u(x,\theta) + \omega v(x,\theta) - (1-\omega)t$, with $\omega \in [0,1]$.

Assumption 1. *We will assume that $v_{xx\theta}(x,\theta) \geq 0$ and $v_{x\theta\theta}(x,\theta) \geq 0$ for all x, θ.*[1]

Remark 2. *From the previous assumption, and using the implicit function theorem we can define a unique decreasing function $x_0(\theta)$, such that $v_{x\theta}(x_0(\theta),\theta) = 0$. So we know that $v_{x\theta}(x,\theta) > 0 \iff x > x_0(\theta)$*

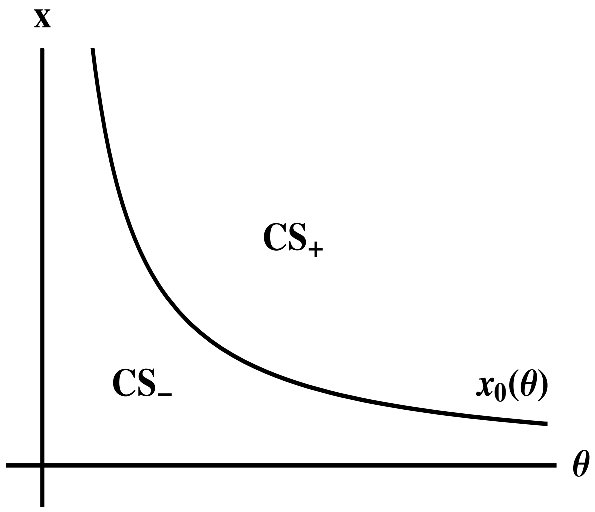

Figure 1: Agents Preferences

We define a mechanism as functions $t(\cdot) : \Theta \to R$ and $x(\cdot) : \Theta \to [\underline{x}, \bar{x}]$, and we say the mechanism $\{x(\theta), t(\theta)\}$ is incentive compatible if and only if:

$$\theta \in argmax_{\theta'} v(x(\theta'),\theta) + t(\theta')$$

By the revelation principle we know that the principal can restrict to I.C. mechanisms, and thus the principal's problem is given by:

$$\max_{\{x(\theta'),T(\theta')\}} \quad \int_{\Theta} u(x(\theta),\theta) + \omega v(x(\theta),\theta) - t(\theta) dP(\theta)$$

$$s.t. \begin{cases} v(x(\theta),\theta) + t(\theta) \geq 0 & \forall \theta \in \Theta \\ \theta \in argmax_{\theta' \in \Theta} \quad v(x(\theta'),\theta) + t(\theta') & \forall \theta \in \Theta \end{cases}$$

Using standard techniques, we can show any I.C. mechanism $x(\cdot)$ must fulfill the following conditions .

Lemma 3. *(Local I.C.C.)*

Let $\{x(\cdot), t(\cdot)\}$ be an I.C. mechanism then the following condition must be fulfilled:

1. *(F.O.C.) $V^x(\theta) := v(x(\theta),\theta) + t(\theta) = V^x(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(z),z)dz \quad \forall \theta \in \Theta$*

[1] This is a slightly stronger assumption than the one done my Araujo Moreira[1], they consider the case in which $v_x(x,\theta)$ is cuasi-convex in θ instead of strictly convex. The reason why we need this slightly stronger assumption is so we can get that the function $\delta(\theta) = v(x_1,\theta) - v(x_2,\theta)$ is quasi-convex for all $x_1, x_2 \in [\underline{x}, \bar{x}]$

2. *(S.O.C.)* $x(\cdot)$ *is non-decreasing (non-increasing) in* $CS_+(CS_-)$

Using lemma 3 we can rewrite the principal's problem as follows :

$$\max_{\{x(\cdot)\}} \quad \int_\Theta \Big[\underbrace{u(x(\theta),\theta) + v(x(\theta),\theta) - (1-\omega)z(\theta)v_\theta(x(\theta),\theta)}_{f(x,\theta)=\text{virtual surplus}} \Big] p(\theta)d\theta$$

$$s.t. \quad \begin{cases} v(x(\theta,\theta) + t(\theta) \geq 0 & \forall\theta \in \Theta \\ \theta \in argmax_{\theta'} \quad v(x(\theta'),\theta) + t(\theta') & \forall\theta \in \Theta \\ x(\cdot)\text{is non-decreasing (non-increasing) in } CS_+(CS_-) \end{cases}$$

where $z(\theta)$ is defined by:

$$z(\theta)^2 = \begin{cases} \frac{P(\theta)}{p(\theta)} & \text{if } v_\theta \leq 0 \\ -\frac{1-P(\theta)}{p(\theta)} & \text{if } v_\theta \geq 0 \end{cases}$$

Proof. The previous results are standard in the mechanism design literature, and require no further comment, the proofs can be found in Araujo and Moreira [1]. □

We will assume the virtual surplus $f(x,\theta)$ is quasi-convex in x and it is maximized at $x_1(\theta)$. Following Araujo and Moreira [1] we can define the Global Incentive Function (GIF) as follows:

$$\Phi^x(\theta,\hat{\theta}) := \int_{\hat{\theta}}^{\theta} \int_{x(\hat{\theta})}^{x(\tilde{\theta})} v_{x\theta}(\tilde{x},\tilde{\theta})d\tilde{x}d\tilde{\theta} \tag{1}$$

Araujo and Moreira [1] prove the following result:

Lemma 4. *A mechanism* $x(\cdot)$ *is I.C. if and only if* $\Phi(\theta,\theta') \geq 0$ *for all* $\theta,\theta' \in \Theta$[1].

Proof. The proof can be found in Araujo and Moreira [1]. □

Remark 5. *Lemma 4 allow us to turn a two stage maximization problem into a iso-perimetric problem.*

To explain the problems presented by the incentive compatibility constraints in this model we consider continuous mechanisms. If the policy is non-decreasing in CS_+ and non-increasing in CS_-, there will be a θ_0 such that the policy is non-decreasing for types higher than θ_0 and non-increasing for types lower than θ_0. Thus we can see that we may have two separate types $\theta_1,\theta_2 \in \Theta$ such that $x(\theta_1) = x(\theta_2) = \xi$, and obviously there must be a unique transfer associated to allocation ξ. The first challenge is to find a policy that satisfies the local I.C.C. and keeps the same transfers for pooling types.

[2]For simplicity we consider the case where $v_\theta(\cdot)$ does not change sign, or $\omega = 1$

[1]Note that the integrals of $\Phi^x(\theta,\hat{\theta})$ might be integrating in a negative direction, making the zones CS_- integrate positive value. This will be the case when analyzing decreasing policies.

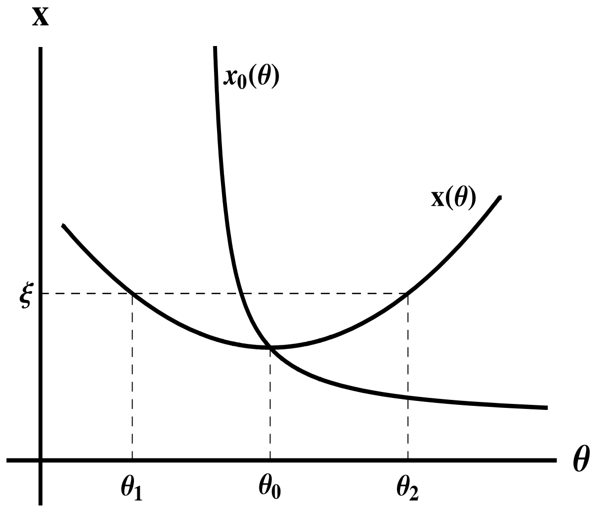

Figure 2: U-Shaped I

The second challenge is presented by the I.C.C. for decreasing policies In the following figure we have a decreasing policy $x(\cdot)$, that is locally I.C. and "close" to $x_0(\theta)$. For policy $x(\cdot)$ to be I.C., condition 1 must be satisfied. However, $\Phi^x(\theta_2, \theta_1)$ (represented by the shaded area) might be negative.

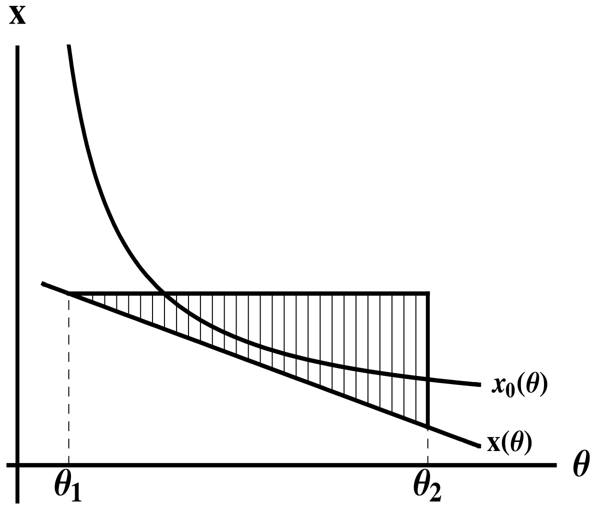

Figure 3: Decreasing Policy I

Let's consider the previous figure and take the case in which $x(\cdot)$ is I.C. and $\Phi^x(\theta_2, \theta_1) = 0$. Since the value of $\Phi^x(\theta_2, \theta_1)$ depends on the assignments given to all types in between (θ_1, θ_2) we might have that there are types in (θ_1, θ_2) that are being distorted even though all I.C.C. involving these types are not binding. So, even for types that don't have their I.C.C. binding with any other type, it might be necessary to keep them " low" so that the shaded area is positive. This is the case we call indirect distortion.

3 Examples

3.1 Leasing

Consider a principal who leases a technology from an agent. The expected lifespan of the technology is a private information of the agent, as different technologies become obsolete at different rates. We assume that at time t a technology θ becomes obsolete for the agent with probability $e^{-\theta t}$, but becomes obsolete for the principal at a higher rate with probability $e^{-(\theta+r)t}$ (r can also be a an extra discount rate).[1]

A technology θ, if not obsolete, gives an instantaneous income of θ to the agent and has a cost $c(\theta)$ with $c'(\theta) < 0$. Therefore, the utility for agent θ of having access the technology θ from time t on is:

$$v(t,\theta) = \int_{s=t}^{\infty} \theta e^{-\theta s} ds - c(\theta) = e^{-\theta t} - c(\theta)$$

The principal can extract a higher instantaneous income from the technology than the agent, $\beta\theta$. Thus, the utility for the principal to keep the technology up to time t is:

$$u(t,\theta) = \int_{s=0}^{t} \beta\theta e^{-(\theta+r)s} ds = \underbrace{\beta\frac{\theta}{\theta+r}}_{=\beta(\theta)} \left(1 - e^{-(\theta+r)t}\right)$$

The principal offers a menu of contracts $\{t(\theta), T(\theta)\}_{\theta\in\Theta}$, specifying the time at which the lease ends and the transfer made to the agent. The total income of agent θ if he chooses contract θ' is given by $V(\theta',\theta) = e^{-\theta t(\theta')} - c(\theta) + T(\theta')$.

Since the principal maximizes expected utility, the principal's problem is given by:

$$U = \max_{\{t(\theta'),T(\theta')\}} \quad \int_{\Theta} \left\{\beta(\theta)(1 - e^{-(r+\theta)t(\theta')}) - T(\theta')\right\} dP(\theta)$$

s.t. I.C.C. & P.C.

Looking at the agent's preferences it is easy to see that the single crossing property is not fulfilled, since $v_{t\theta}$ changes signs. This comes from the fact that it is unclear which agent is willing to receive less for delaying the end of the lease in a amount of time Δt. On the one hand, a technology of a higher θ receives a higher income stream and thus is willing to forfeit more transfers in order to receive the technology earlier. On the other hand, agents with with a higher θ have a higher discount value (the income stream is more front-loaded) and thus the present value of the income stream decreases with θ, making him less willing to forfeit transfers to receive the technology back earlier. In fact we can have that the first effect dominates, in which case t and θ are strategic substitutes ($v_{t\theta} < 0$), or the second, in which case t and θ are strategic complements ($v_{t\theta} > 0$).

Another way to look at the problem is also interesting. The income that agent θ gets from receiving the technology back at time t is $e^{-\theta t}$, which is also the probability of technology θ not being obsolete by time t. Thus we can see that higher θ's are technologies that become obsolete at a higher rate and thus, conditional on not being obsolete at time t, have a higher probability of becoming obsolete in the extra time Δt, requiring a higher compensation to delay the end of the lease. However higher θ's are already obsolete by time t with a higher probability and, if that is the case, require no extra compensation to delay the end of the lease. Then, they require a lower compensation to accept a delay in the expiration date of the lease.

Noting that $v_{t\theta} = e^{-\theta t}(\theta t - 1)$ we have the following figure:

[1]The same model also applies to several other situations like exclusivity, copyright, venture capital etc.

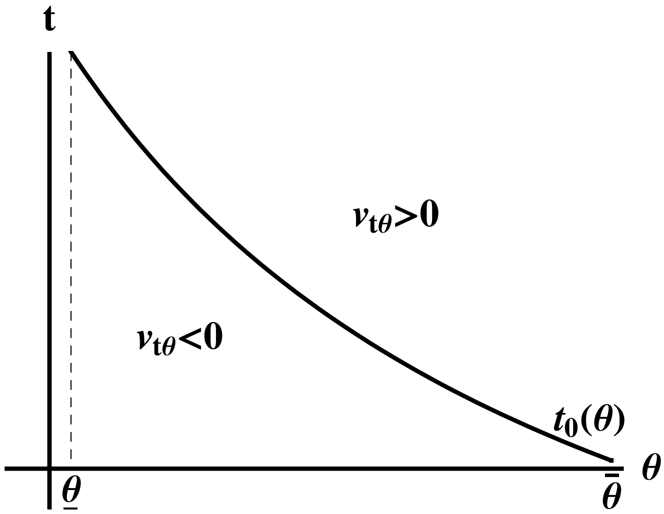

Figure 4: Agents Preferences Lease

To simplify the analysis, we make the following assumptions:

1. If the principal did not exist, it would not be profitable for the agent to develop any technology. [3] That is: $v(0,\theta) = 1 - c(\theta) \leq 0$.
2. All technologies would be implemented if there were full information. That is $\beta\frac{\theta}{\theta+r} - c(\theta) \geq 0$
3. Higher θ's are more profitable for the agent. That is $v_\theta(t,\theta) = -te^{-\theta t} - c'(\theta) \geq 0$

In particular, to get an exact solution, we assume that $c(\theta) = 1 - 2\theta$, $P(\theta)$ is uniform in the interval $[\frac{1}{2}, 1]$, $\beta = 4.5$ and $r = 1$. With this, (1) to (3) are fulfilled.

3.1.1 Maximization Problem

Lt's consider the case in which there is full information (FI), so the principal just maximizes the surplus, since he can extract all rents. We have:

$$t^{FI}(\theta) = argmax_t \quad \beta\frac{\theta}{\theta+r}(1 - e^{-(\theta+r)t}) + e^{-\theta t} - c(\theta)$$

$$\Rightarrow t^{FI}(\theta) = \frac{-1}{r}Log(\frac{1}{\beta})$$

The result is quite intuitive. From an efficiency point of view the optimal time at which the lease must end is independent of θ, and it is given by the time at which the extra stream that the principal receives offsets the extra discount rate it confronts. Thus, t^{FI} is the time in which the present value of the stream that the principal and the agent get from the technology are the same.

Going back to the problem with private information and using standard techniques we can rewrite the principal's problem as:

$$U = \max_{t(\theta)} \quad \int_\Theta \left\{ \underbrace{e^{-\theta t(\theta)} - \beta(\theta)e^{-(r+\theta)t(\theta)} + z(\theta)te^{-\theta t(\theta)}}_{=f(t,\theta)} \right\} p(\theta)d\theta + E_p[H(\theta)]$$

$$s.t. I.C.C.$$

where $z(\theta) = \frac{1-P(\theta)}{p(\theta)}$, $f(t,\theta)$ is the virtual surplus and $H(\cdot)$ does not depend on $t(\cdot)$. Pointwise maximization of the virtual surplus leads to:

$$-\theta e^{-\theta t_1(\theta)}\left(1 - \beta e^{-rt_1(\theta)}\right) - z(\theta)v_{t\theta}(t_1(\theta),\theta) = 0$$

[3] With this, the critical type is not endogenous

This has a straightforward interpretation. The principal's income is given by the surplus minus the informational rents. Maximizing the first term is equivalent to maximizing total surplus (it is maximized at t^{FI}) while the second part corresponds to minimizing the informational rents, which are minimized at $t_0(\theta)$, where $v_{t\theta} = 0$.

We know proceed to show the optimal solution. As we can see in the next figure, we can divide the type space Θ in two zones. Long lived technologies in $[\underline{\theta}, \theta_1]$ and short lived technologies in $[\theta_1, \bar{\theta}]$. Pointwise maximization leads to a lease time t higher than t^{FI} for long lived technologies, while the opposite happens for short lived technologies.

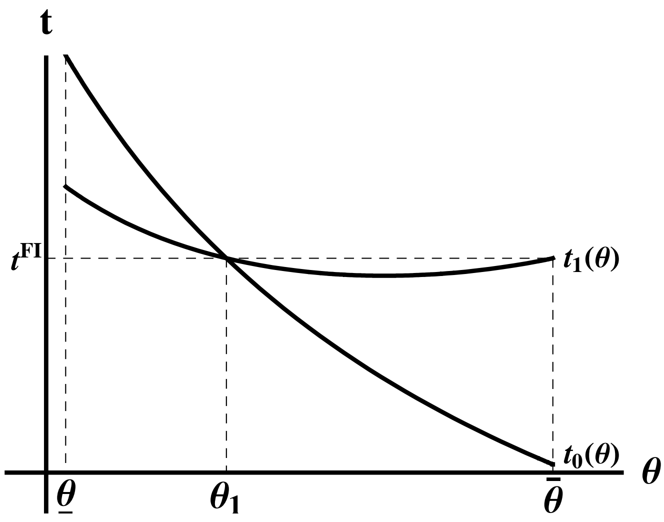

Figure 5: Pointwise Maximization Lease

However, $t_1(\cdot)$ does not satisfy even the local I.C. constraints (it is decreasing in a region where $v_{t\theta}$ is positive). Moreover, a naive "fixing" of this does not solve the problem. Consider, for example, the policy $x(\cdot)$ given by

$$x(\cdot) = \begin{cases} t_1(\theta) & \theta < \theta_1 \\ t_1(\theta_1) & \theta > \theta_1 \end{cases}$$

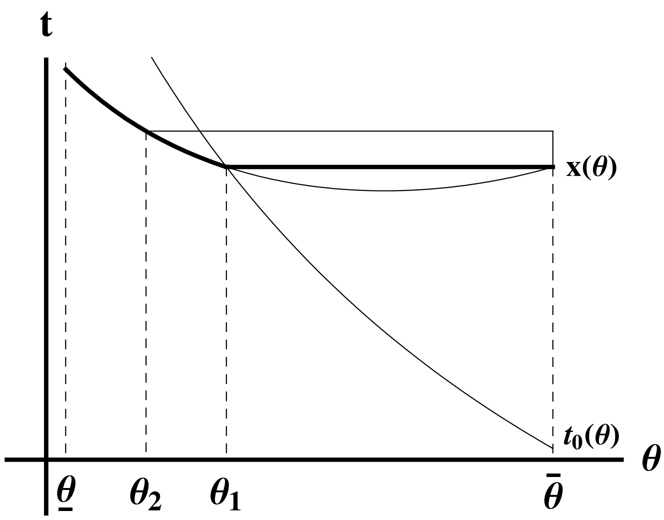

Figure 6:

Comparing the utility obtained by an agent of type $\bar{\theta}$ that tells the truth or declares some other type $\theta_2 \in [\underline{\theta}, \theta_1)$ one can verify the latter gives a greater utility. Using the local I.C. constraints the difference between both utilities can be written as:

$$\begin{aligned}
\Phi^x(\bar\theta,\theta_2) &= v(x(\bar\theta),\bar\theta)+T(\bar\theta)-(v(x(\theta_2),\bar\theta)+T(\theta_2))\\
&= \underbrace{v(x(\bar\theta),\bar\theta)+T(\bar\theta)}_{V(\bar\theta,\bar\theta)} - v(x(\theta_2),\bar\theta) - \underbrace{T(\theta_2)}_{V(\theta_2,\theta_2)-v(x(\theta_2),\theta_2)}\\
&= \underbrace{v(x(\theta_2),\theta_2)-v(x(\theta_2),\bar\theta)}_{-\int_{\theta_2}^{\bar\theta} v_\theta(x(\theta_2),s)ds} + \underbrace{V(\bar\theta,\bar\theta)-V(\theta_2,\theta_2)}_{\int_{\theta_2}^{\bar\theta} v_\theta(s,s)ds}\\
&= -\int_{\theta_2}^{\bar\theta}\int_{x(z)}^{x(\theta_2)} v_{x\theta}(z,s)dzds
\end{aligned}$$

If $\Phi^x(\bar\theta,\theta_2)<0$ for some θ_2, then the an agent of type $\bar\theta$ would prefer to declare θ_2. In this case the area that needs to be calculated is shown in the previous figure, and it is easy to note that for θ_2 close enough to θ_1 the area in which $v_{x\theta}>0$ will dominate the area in which $v_{x\theta}<0$, and thus the policy is not incentive compatible.

Intuitively, what is happening is that policy $x(\cdot)$ assigns a higher t to types lower than θ_1, which is compensated with higher transfers. Locally, lower technologies need a smaller compensation for delaying the end of the lease because the dominating effect is the bigger income stream. But a technology $\bar\theta$ is very likely obsolete by time t^{FI}, and thus his owner is willing to get an extra transfer in exchange for some extra leasing time.

The optimal policy, computed with methods developed in this paper, is shown in the next figure:

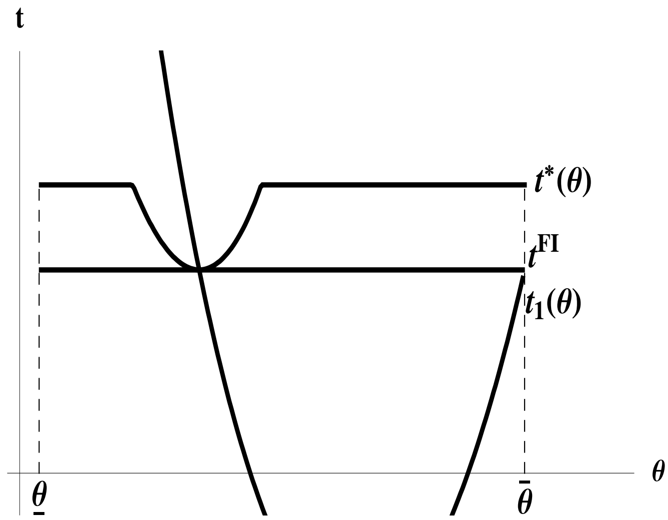

Figure 7: Optimal Policy Lease

We can see that long lived technologies are distorted from t^{FI} in the direction of the point-wise maximization, although not as much as pointwise maximization would dictate. Short-lived technologies, on the other hand, are distorted from t^{FI} away from the pointwise solution. The intuition behind these distortions lies in the global I.C. constraints. They force short lived technologies to be distorted in the same direction as the long lived technologies, to avoid non-local deviations. For this reason, since the distortion of the long lived technologies affects the choices that can be made about the short lived ones, long lived technologies are distorted less. The resulting policy consists in two zones in which the policy is bunching and a middle zone that has a U-shaped form, in which there is a discrete pooling between long and short lived technologies.

Through out the paper we will show how to find the optimal U-Shaped form in a general context. We show that the effect that the global I.C.C. have in the policy, which leads to this kind of U-shaped form in which there is discrete pooling, can be easily weighted.

3.2 Externalities

Let's consider a monopoly that has constant marginal costs equal to $x\theta$, where x is the technology adopted by the firm and θ is the firm's efficiency. The firm can decide its production level, but the principal (e.g. a governmental agency) may force it to adopt a given technology. For simplicity we assume that the firm faces a linear demand. Thus, the operational profits made by the firm are given by:

$$v(x,\theta) = \max_q \quad (A - bq)q - \theta x q = \frac{(A-\theta x)^2}{4b}$$

As a side effect, the firm produces a negative externality. More expensive technologies (higher x) are cleaner.

We assume that externalities are proportional to the quantity produced, and are reduced proportionally to the technology adopted. That is,

$$\Rightarrow \text{Externalities} = q(E - \beta x)$$

The principal offers a menu $\{x(\theta), T(\theta)\}_{\theta\in\Theta}$ and maximizes the expected value of

$$u(x,\theta) = \text{Firm's Profits} + \text{Consumer's Surplus} - \text{Externalities}$$

It is easy to see that: Firm's Profits $= \frac{(A-\theta x)^2}{4b} + T$ and Consumer's Surplus $= \frac{(A-\theta x)^2}{8b} - T$, therefore the principal's problem is given by:

$$U = \max_{\{t(\theta'),T(\theta')\}} \quad \int_\Theta \left\{ \frac{(A-\theta x)^2}{4b} + \frac{(A-\theta x)^2}{8b} - q(E-\beta x) \right\} dP(\theta)$$

$$\text{s.t. I.C.C. \& P.C.}$$

Unlike the previous example, since the principal doesn't care about transfers, the virtual surplus is the same as the total surplus.

Again, it is unclear which firm is willing to pay more to decrease x, so the S.C.P. is not fulfilled. On the one hand, more efficient firms are less affected by x (since x and θ are complements), and thus more efficient firms are less willing to pay for a marginal decrease in x. On the other hand, more efficient firms want to produce more, and thus more efficient firms might be willing to pay more to reduce x. Once again we can see that it is unclear what effect dominates. It could be that x and θ are strategic substitutes ($v_{x\theta} < 0$), or strategic complements ($v_{x\theta} > 0$).

Finally it is worth mentioning that in case the marginal costs would have the form $\theta + x$, one of the effects would disappear and the S.C.P. would be fulfilled. If this would be the case, then more efficient firms would always be more willing to pay for a decrease in x.

Going back to the maximization problem

$$U = \max_{\{t(\theta'),T(\theta')\}} \quad \int_\Theta \underbrace{\left\{ \frac{(A-\theta x)^2}{4b} + \frac{(A-\theta x)^2}{8b} - q(E-\beta x) \right\}}_{=f(x,\theta)} dP(\theta)$$

$$\text{s.t. I.C.C. \& P.C.}$$

Pointwise maximization (which is the same as the first best in this case) leads to

$$-\frac{3\theta}{4b}\left(A - \theta x^{FI}(\theta)\right) + \beta\frac{A - \theta x^{FI}(\theta)}{2b} + \frac{\theta}{2b}(E - \beta x^{FI}(\theta)) = 0$$

Looking at the next figure[4], we can see that $x^{FI}(\cdot)$ satisfies all the local I.C. constraints.

[4] All graphs are done for the case $A = 5$, $b = 1$, $E = \frac{11}{4}$, $\beta = 5$, $\underline{\theta} = 2$, $\bar{\theta} = 4$ and a uniform distribution for $p(\theta)$

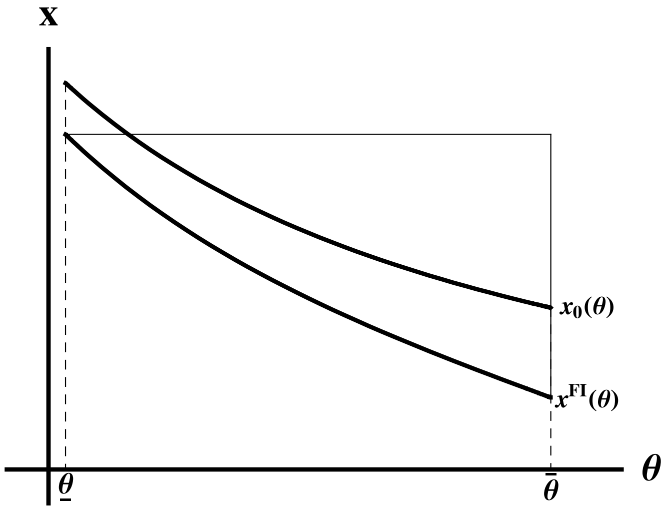

Figure 8: Pointwise Maximization Externalities

Nevertheless, x^{FI} is not I.C. We compare the income that type $\bar{\theta}$ receives by declaring $\bar{\theta}$ or $\underline{\theta}$ under this policy. Using the local I.C.C. the difference between both utilities can be written as

$$
\begin{aligned}
\Phi^{x^{FI}}(\bar{\theta},\underline{\theta}) &= v(x^{FI}(\bar{\theta}),\bar{\theta}) + T(\bar{\theta}) - (v(x^{FI}(\underline{\theta}),\bar{\theta}) + T(\underline{\theta})) \\
&= \underbrace{v(x^{FI}(\bar{\theta}),\bar{\theta}) + T(\bar{\theta})}_{V(\bar{\theta},\bar{\theta})} - v(x^{FI}(\underline{\theta}),\bar{\theta}) - \underbrace{T(\underline{\theta})}_{V(\underline{\theta},\underline{\theta})-v(x(\underline{\theta}),\underline{\theta})} \\
&= \underbrace{v(x^{FI}(\underline{\theta}),\underline{\theta}) - v(x(\underline{\theta}),\bar{\theta})}_{-\int_{\underline{\theta}}^{\bar{\theta}} v_\theta(x^{FI}(\underline{\theta}),s)ds} + \underbrace{V(\bar{\theta},\bar{\theta}) - V(\underline{\theta},\underline{\theta})}_{\int_{\underline{\theta}}^{\bar{\theta}} v_\theta(s,s)ds} \\
&= -\int_{\underline{\theta}}^{\bar{\theta}} \int_{x^{FI}(z)}^{x^{FI}(\underline{\theta})} v_{x\theta}(z,s)dzds
\end{aligned}
$$

It is unclear to the naked eye if this area is positive or negative in this case, but for these parameters the corresponding value is -0.19393. This means that the difference between the transfers paid to $\underline{\theta}$ and $\bar{\theta}$ is bigger than the cost to $\bar{\theta}$ of implementing $x^{FI}(\underline{\theta})$ instead of $x^{FI}(\bar{\theta})$, which means $x^{FI}(\cdot)$ is not I.C..

Note that the marginal rate of substitution between the transfer t and the technology x at some level $x^{FI}(\tilde{\theta})$ is given by the marginal cost of the technology to the respective type $\tilde{\theta}$ (to keep him indifferent between telling the truth and making a local deviation). Therefore, the difference between the transfers paid to $\underline{\theta}$ and $\bar{\theta}$ does not only depend on the difference between the technologies assigned to them, but also on the rate of substitution between the technology and the transfers of all technologies in between $x^{FI}(\bar{\theta})$ and $x^{FI}(\underline{\theta})$. Therefore, the difference in the transfers paid to $\underline{\theta}$ and $\bar{\theta}$ depends on the *whole* path of $x^{FI}(\cdot)$ between $\underline{\theta}$ and $\bar{\theta}$.

Since the optimal policy is not I.C., it is clear that the optimal policy must lie below $x^{FI}(\cdot)$ at least in some interval so that the computed area has bigger zone in which $v_{x\theta} < 0$, which in turn would increase the value of the computed area. Intuitively, this implies that more efficient firms are getting a given technology x, which lowers the rate of substitution between technology and transfer, therefore lowering the difference between the transfers paid to $\underline{\theta}$ and $\bar{\theta}$. The following figure presents the optimal policy $x^{*}(\cdot)$

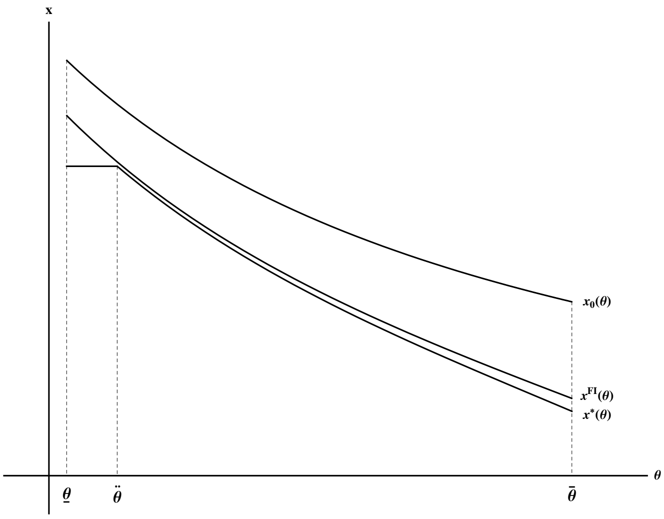

Figure 9: Optimal Policy Externalities

We can see that the optimal policy has a bunching zone in $[\underline{\theta}, \ddot{\theta}]$ and a strictly decreasing zone $[\ddot{\theta}, \bar{\theta}]$. Since the difference in transfers between $\bar{\theta}$ and $\underline{\theta}$ in $x^*(\cdot)$ depends on the whole path of $x^*(\cdot)$ in $[\ddot{\theta}, \bar{\theta}]$ this zone is distorted from $x^{FI}(\cdot)$ to lower this difference. On other hand, $x^*(\cdot)$ can be seen as truncated at $\ddot{\theta}^*$ because the higher the assignment for $\underline{\theta}$ the more difficult it is to keep the difference in transfers low enough, and thus truncating the optimal policy allows to lower the distortion of interior types.

Intuitively in zones of the policy in which $v_{x\theta}$ is bigger in absolute value we have that the difference in preferences between types are more pronounced, and thus a bigger decrease in transfers can be achieved with lower distortions. Geometrically, adding area where $v_{x\theta}$ is bigger in absolute value is more efficient. As a result in the optimal policy all the strictly decreasing part has the same ratio $\frac{f_x(x^*(\theta),\theta)p(\theta)}{v_{x\theta}(x^*(\theta),\theta)}$, which is the reason why the optimal policy in the precious figure has zones in which the distortion is bigger than others.

So far we have only given the intuition on how the global I.C.C. between $\bar{\theta}$ and $\underline{\theta}$ is managed, which is sufficient for all the global I.C.C. to be fulfilled in this particular example. This is not the general case, and through out this paper we show how the global I.C.C. are managed in a general way, and we characterize the optimal solution for these cases, but the previous intuitions are kept.

4 Global Incentive Compatibility Constraints

Before giving any results concerning the I.C.C. in the case of a continuum of types we will analyze what kind of policy are implementable in this model but concidering just two types, $\theta_H > \theta_L$. To take a case in which the S.C.P. is not fulfilled we will consider the case in which exists $\xi_t \in (\underline{x}, \bar{x})$ such that $v_x(\xi_t, \theta_H) = v_x(\xi_t, \theta_L)$. Since in this model $v_{xx\theta} > 0$, we know that $v_x(\xi', \theta_H) > v_x(\xi', \theta_L)$ for $\xi' > \xi_t$ and vice-versa. Looking at the I.C.C. between θ_H and θ_L :

$$v(x(\theta_H), \theta_H) + T(\theta_H) \geq v(x(\theta_L), \theta_H) + T(\theta_L) \quad \wedge \quad v(x(\theta_L), \theta_L) + T(\theta_L) \geq v(x(\theta_H), \theta_L) + T(\theta_H)$$

$$\Rightarrow v(x(\theta_H), \theta_H) - v(x(\theta_L), \theta_H) \geq v(x(\theta_H), \theta_L) - v(x(\theta_L), \theta_L) \iff \int_{x(\theta_L)}^{x(\theta_H)} v_x(z, \theta_H) \geq \int_{x(\theta_L)}^{x(\theta_H)} v_x(z, \theta_L) dz$$

Thus, any assignments for θ_H and θ_L that are implementable through transfers must satisfy the previous inequality. Just by looking at the previous inequality, we can see that for any given $x(\theta_H)$ there is a bounded space where $x(\theta_L)$ can be implementable. Moreover, we know that $x(\theta_H) \geq (\leq) \xi_t \Rightarrow x(\theta_L) \leq (\geq) \theta_H$, so we can know if $x(\theta_L)$ is bigger or smaller than $x(\theta_H)$ based only in $x(\theta_H)$. The previous can easily be seen graphically, the following two plots show the curves of iso-profit of θ_H and θ_L, and show the implementable space left for $x(\theta_L)$ and $x(\theta_H)$ when $x(\theta_H)$ and $x(\theta_L)$ are held fixed respectively.

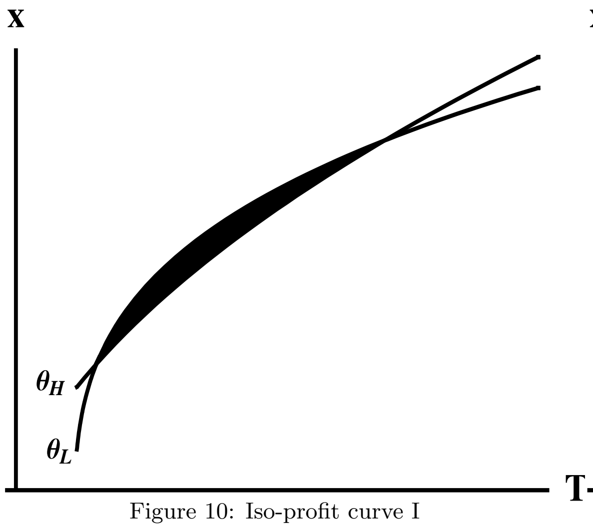

Figure 10: Iso-profit curve I

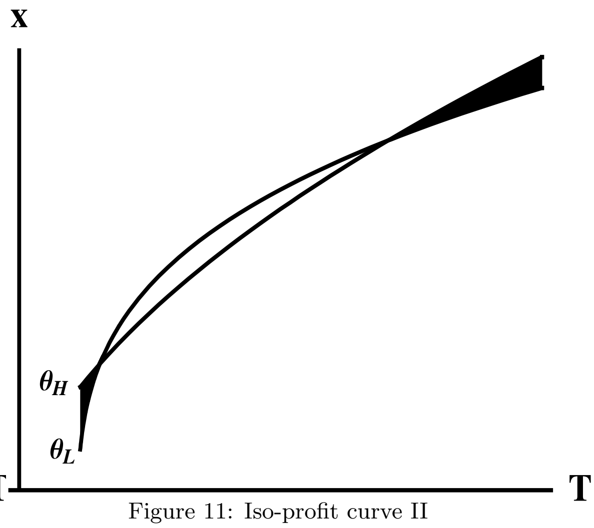

Figure 11: Iso-profit curve II

From the previous analysis it is easy to see that the point at which the isocurve between two types are tangent is important.

Definition 6. *Let $\theta \in [\underline{\theta}, \bar{\theta}]$, for a given $\xi \in [\underline{x}, \bar{x}]$ if there exists $\theta' \neq \theta$ such that $v_x(\xi, \theta) = v_x(\xi, \theta')$, we will define $\hat{\theta}(\xi, \theta) = \theta'$. Just by the assumptions made of the agent's preferences we know that if $\hat{\theta}(\xi, \theta)$ exists, then it is unique. Moreover if $\hat{\theta}(\xi, \theta) < \theta$ then we know that $v_{x\theta}(\xi, \theta) > 0 > v_{x\theta}(\xi, \hat{\theta}(\xi, \theta))$, and viceversa.*

just using the previous intuitions we can make the following observation on the I.C.C. for a continuum of types:

Proposition 7. *Let $x(\cdot)$ be an I.C. mechanism, and let θ' be such that $\hat{\theta}(x(\theta'), \theta')$ exists and $\hat{\theta}(x(\theta'), \theta') < \theta'$. Then the following must hold true:*

$$x(\theta) \begin{cases} \leq x(\theta') & \theta \in [\hat{\theta}(x(\theta'), \theta'), \theta'] \\ \geq x(\theta) & \theta \in [\underline{\theta}, \hat{\theta}(x(\theta'), \theta')] \\ = x(\theta') & \theta = \hat{\theta}(x(\theta'), \theta') \end{cases}$$

Proof. The previous result is direct from the I.C.C. between two types, described previously. It can be deducted from the fact that for a given θ and $x(\theta)$ in which exists $\hat{\theta}(\xi, \theta) < \theta$, then the tangency point is below $x(\theta)$ for $\theta' > \hat{\theta}(\xi, \theta)$ and above $x(\theta)$ for $\theta' < \hat{\theta}(\xi, \theta)$, and the rest is trivial from the the I.C.C. between two type (just rename $\theta = \theta_H$, and any $\theta_L < \theta$, and see where the tangency point is with respect to $x(\theta)$). $\square$

The previous proposition shows that the assignment of any given type induces a shape on all lower types, which consists in separating the lower types in a zone in which the assignments are lower and a second zone in which the assignments are higher. The following theorem will be very useful in characterizing the optimal policy since it helps disconnect the optimization problem by allowing us to find the optimal policy for the types that have lower and higher assignment independently.

Theorem 8. *Let $\hat{x}(\cdot)$ be an I.C. policy, then for any given interval $[\theta_1, \theta_2]$such that $x_\alpha = \hat{x}(\theta_2) = \hat{x}(\theta_1)$, we have that the policy $\hat{x}'$ defined by:*

$$\hat{x}'(\theta) = \begin{cases} \hat{x}(\theta) & \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}] \\ x_\alpha & \theta \in [\theta_1, \theta_2] \end{cases}$$

is also I.C.

The opposite is also true. That is to say, let $\hat{x}(\cdot)$ be I.C. and such that the policy is bunching in an interval $[\theta_1, \theta_2]$. For any policy $\tilde{x}(\cdot)$ defined in $[\theta_1, \theta_2]$, such that $\tilde{x}(\cdot)$ is I.C. and such that $x_\alpha = \tilde{x}(\theta_2) = \tilde{x}(\theta_1) = \hat{x}(\theta_1)$, then policy $\hat{x}'(\cdot)$ defined by:

$$\hat{x}'(\theta) = \begin{cases} \hat{x}(\theta) & \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}] \\ \tilde{x}(\theta) & \theta \in [\theta_1, \theta_2] \end{cases}$$

is also I.C.

Proof. The proof can be found in the Appendix. □

As we explained in the the description of the model the I.C.C. in this model presents two challenges, one concerning with discrete pooling and the second one concerning decreasing policies. The previous result will allow us to disentangle the complete optimization in subproblems, which will allow us to separate the challenge concerning discrete pooling with the one concerning decreasing policies. The next theorem will be useful to characterize decreasing policies, it will ensures that binding global I.C.C. are always between pairs of types, moreover it will allowing us to find a order in the active I.C.C. This will allow us to describe the global I.C.C. as "nested", which significantly limits the possible combinations of active constraints.

Theorem 9. *Let $\hat{x}(\theta)$ be an I.C. mechanism, such that $\hat{x}(\theta)$ is decreasing in some interval $[\theta_\alpha, \theta_\beta]$. For any $(\theta_1 < \theta_2 \leq \theta_3 < \theta_4) \in [\theta_\alpha, \theta_\beta]$ it can never hold true that $\Phi(\theta_4, \theta_2) = 0$ and $\Phi(\theta_3, \theta_1) = 0$.*

Proof. The proof of theorem 9 can be explained easily geometrically. Consider the following figure with a decreasing policy:

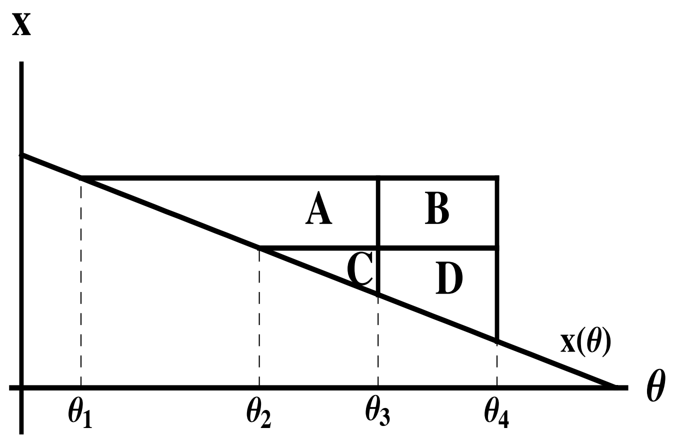

Figure 12: Decreasing policy II

The GIF function between θ_1 and θ_3 is given by area A+C weighted by $v_{x\theta}$ and the GIF function between θ_2 and θ_4 is given by area C+D weighted by $v_{x\theta}$. So if the I.C.C. between θ_1 and θ_3 is active, and the I.C.C. between θ_2 and θ_4 is active it means that the areas A+C weighted by $v_{x\theta}$ are 0 and the areas C+D weighted by $v_{x\theta}$ are 0. We also know that for the policy to be I.C.C. the area C weighted by $v_{x\theta}$ must be positive, so we have that areas A+C+D weighted by $v_{x\theta}$ is positive. Finally, it can be shown that area D weighted by $v_{x\theta}$, must be negative, and thus A+B+C+D weighted by $v_{x\theta}$ is negative, which means that θ_4 would want to jump to θ_1, and thus the mechanism is not I.C.

The details can be found in the Appendix. □

So far we have used the characteristics of the functional form of the agents preferences to find global properties of implementable policies. The next lemma is result found by Araujo and Moreira [1], which we will present here since we will use it. This lemma ensure that the local necessary conditions of types that are binding by global I.C.C. are compatible.

Lemma 10. *Let $x(\cdot)$ be an implementable decision and $\theta, \hat{\theta} \in \Theta$ be such that $\Phi(\theta, \hat{\theta}) = 0$*

(i) If $x(\cdot)$ is strictly monotonic and continuous at $\hat{\theta}$ then

$$v_x(x(\hat{\theta}), \theta) = v_x(x(\hat{\theta}), \hat{\theta})$$

. *(ii) If $x(\cdot)$ is continuous at θ , then*

$$v_\theta(x(\hat{\theta}), \theta) = v_\theta(x(\theta), \theta)$$

To interpret lemma 10 let's consider $x(\cdot)$ to be a policy that fulfills all hypothesis of 10. Equation $\Phi(\theta, \hat{\theta}) = 0$ means that θ is indifferent between the contract offered to him and the contract offered to $\hat{\theta}$.

We have that $x(\cdot)$ is strictly monotonic and continuous at $\hat{\theta}$, which means that the policy is screening locally at $\hat{\theta}$. Since θ is indifferent between the contract offered to him and the contract offered to $\hat{\theta}$, if the contracts offered locally around $\hat{\theta}$ were to be offered to θ we must have that θ chooses $\hat{\theta}$'s contract (else $x(\cdot)$ would not be I.C.). Hence preferences of θ and $\hat{\theta}$ must be indistinguishable around $x(\hat{\theta})$, which is condition (i). Mathematically, the marginal change in transfers for a marginal change of allocation at $x(\hat{\theta})$ is exactly $v_x(x(\hat{\theta}), \hat{\theta})$, and $\Phi(\theta, \hat{\theta}) = 0$ implies that θ is indifferent between his contract and the one given to $\hat{\theta}$, so condition (i) guarantees that θ will not want to change to a marginally lower allocation or a marginally higher allocation than $x(\hat{\theta})$. In other words condition(i) guarantees that θ doesn't want to change to a neighborhood near $\hat{\theta}$

Since $x(\cdot)$ is I.C.C., locally around θ all types receive the same utility as if they were choosing $x(\theta)$ (at least in a first order approximation), and thus choosing $x(\hat{\theta})$ cannot yield a higher utility than $x(\theta)$ for all types in a neighborhood around θ. Hence, condition (i) states that all types in some neighborhood around θ all types must receive (almost) the same utility from choosing $x(\theta)$ and $x(\hat{\theta})$. Mathematically to interpret the second part I will rewrite the equations as follows:

$$v_\theta(x(\hat{\theta}), \theta) = v_\theta(x(\theta), \theta) \iff \frac{\partial}{\partial \varphi} \int_{x(\theta)}^{x(\hat{\theta})} v_x(z, \varphi) dz|_{\varphi=\theta} = 0$$

Notice that the term $\int_{x(\theta)}^{x(\hat{\theta})} v_x(z, \theta) dz$ gives exactly how much θ values the difference between the allocation given to $\hat{\theta}$ and to him. Since θ is indifferent between his contract and the one given to $\hat{\theta}$, condition (ii) guarantees that no type in a neighborhood near θ values this difference more than θ, because in this case he would rather $\hat{\theta}$ contract than his. In other words condition(ii) guarantees that no type in a neighborhood of θ wants to change to $\hat{\theta}$

5 Necessary Conditions for Optimality

The next theorem is one of our main results. It gives an optimality condition for all types whose I.C.C. have some slack in the optimal policy. This is the case as we mention in the introduction of indirect distortion since non of the global I.C.C. involving the given type is active, nevertheless they need to be distorted to ensure the fulfillment of the I.C.C. among other types.

Theorem 11. *Let $x^*(\cdot)$ be an optimal policy, which is continuous in (θ_1, θ_2).*

1. *Suppose that $\forall \theta \in (\theta_1, \theta_2) \forall \theta' \in \Theta \quad \Phi^{x^*}(\theta, \theta') > 0$, then $\frac{f_x(\hat{x}(\theta), \theta)p(\theta)}{v_{x\theta}}$ is non-increasing in (θ_1, θ_2)*
2. *Suppose that $\forall \theta \in (\theta_1, \theta_2) \forall \theta' \in \Theta \quad \Phi^{x^*}(\theta', \theta) > 0$, then $\frac{f_x(\hat{x}(\theta), \theta)p(\theta)}{v_{x\theta}}$ is non-decreasing in (θ_1, θ_2)*

Proof. See Appendix. □

Remark 12. *It is worth mentioning that theorem* **??** *holds for any functional form of the agents preferences $v(x, \theta)$ or the prinicpal's objective function $f(x, \theta)$.*

Corollary 13. *Suppose that $x^*(\cdot)$ is continuous in some interval (θ_1, θ_2), and $\Phi(\theta, \theta') > 0$ and $\Phi(\theta', \theta) > 0$ for all $\theta \in (\theta_1, \theta_2)$ and $\theta' \in \Theta$. Then,*

$$\frac{f_x(x^*(\theta), \theta)p(\theta)}{v_{x\theta}(x^*(\theta), \theta)} \textit{is constant in } (\theta_1, \theta_2)$$

Note that corollary 13 implies that the optimal policy must keep the value of $\frac{f(\cdot,\cdot)p(\cdot)}{v_{x\theta}(\cdot,\cdot)}$ constant unless there is an incentive compatibility constraint that is binding, that is:

Corollary 14. *Let x^* be the optimal policy. Then for all $\theta' \in [\underline{\theta}, \bar{\theta}]$ such that x^* is continuous at θ' one of the following two conditions must hold:*

1. $\frac{d}{d\theta}\frac{f_x(x^*(\theta),\theta)p(\theta)}{v_{x\theta}(x^*(\theta),\theta)}|_{\theta=\theta'} = 0$

2. $\exists\ \theta'' \in \Theta$ *such that* $\Phi^{x^*}(\theta', \theta'') = 0 \quad \vee \quad \Phi^{x^*}(\theta'', \theta') = 0$

Corollary 13 has an interesting interpretation. First note that $v_{x\theta}(x^*(\theta),\theta)$ can be interpreted as the "price" of distorting type θ. Consider a policy such that there are no binding I.C.C. for all types in (θ_1, θ_2), and a principal who wants to increase $x(\theta_2)$ in dx' without affecting the I.C.C. between types that are outside (θ_1, θ_2) (we are disregarding the local I.C.C.). This change would would increase $\Phi^{x^*}(\theta_2, \theta_1)$ by an amount of $v_{x\theta}(x^*(\theta_2), \theta_2)dx'd\theta$ (which in turns increases $\Phi^{x^*}(\theta_A, \theta_B)$ by an amount of $v_{x\theta}(x^*(\theta_2), \theta_2)dx'd\theta$ for all types $\theta_A < \theta_1$ and $\theta_B > \theta_2$). In order to keep $\Phi^{x^*}(\theta_A, \theta_B)$ constant, it is necessary to decrease $x^*(\cdot)$ at some other type θ_1 by an amount dx such that:

$$v_{x\theta}(x^*(\theta_1), \theta_1)dxd\theta + v_{x\theta}(x^*(\theta_2), \theta_2)dx'd\theta = 0$$

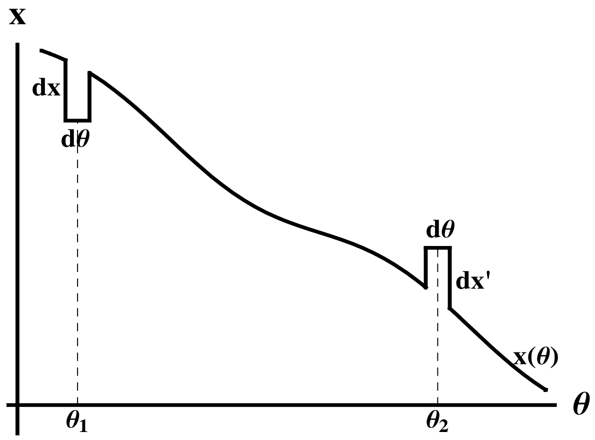

Figure 13: Variational calculus

We make the analogy to an agent maximizing utility under a fixed budget. An increase dx' in the consumption of one good $(x^*(\theta_2))$ times the price of the good must be the same as a decrease of dx in the consumption of some other good $(x^*(\theta_1))$ times its price. In this case $v_{x\theta}(x^*(\theta_2), \theta_2)$ and $v_{x\theta}(x^*(\theta_1), \theta_1)$ can be interpreted as these relevant prices. Moreover, $f_x(x^*(\theta), \theta)p(\theta)$ is the marginal utility obtained by the principal by a marginal increase of the "consumption of the good $x^*(\theta)$". Therefore, keeping the ratio $\frac{f_x(x^*(\theta),\theta)p(\theta)}{v_{x\theta}(x^*(\theta),\theta)}$ constant is analogous to keeping the ratio between marginal utility and price constant in classical consumer theory.

It is important to highlight that in the case in which the S.C.P. holds the feasibility of a policy is independent of the term $v_{x\theta}$ (it is only necessary to keep the monotonicity of the policy). Thus, the optimal policy is affected by the term $v_{x\theta}$ only through the informational rents.

There is an implication of corollary 13. Under those conditions and $\frac{\partial}{\partial x}\frac{f_x(x,\theta)p(\theta)}{v_{x\theta}(x,\theta)} < 0$ it turns out that the curve that keeps $\frac{f_x(x,\theta)p(\theta)}{v_{x\theta}(x,\theta)}$ constant is a minimum, and thus it is always better to have discontinuous jumps. But note that:

$$\frac{\partial}{\partial x}\frac{f_x(x,\theta)p(\theta)}{v_{x\theta}(x,\theta)} < 0 \iff -\frac{v_{xx\theta}}{v_{x\theta}} > -\frac{f_{xx}}{f_x}$$

Noting that the informational rents of type θ' are given by the rents given to some other type $\theta' - \Delta\theta$ plus $v_\theta(x^*(\theta'), \theta') \cdot \Delta\theta$, we can interpret $-\frac{v_{xx\theta}}{v_{x\theta}}$ as the agents risk aversion, while $-\frac{f_{xx}}{f_x}$ is the principal's risk aversion. Therefore, the principal only decides to "smooth" it's risk if he is more risk averse than the agents.

The next theorem gives us an optimality condition for types with an active I.C.C.. This theorem is an extension of a result from Araujo Moreira [1], the difference is that we use both conditions of lemma 10 , and we ensure no other global I.C.C. is broken.

Theorem 15. *Let $x^*(\cdot)$ be the optimal implementable convex-valued correspondence, and $\theta'', \theta' \in \Theta$ be such that $\Phi^{x^*}(\theta'', \theta') = 0$. If $x^*(\cdot)$ is strictly monotonic and continuous at θ' and θ'' , then:*

$$\frac{f_x(x^*(\theta''), \theta'')p(\theta'')}{v_{x\theta}(x^*(\theta''), \theta'')} = \frac{f_x(x^*(\theta'), \theta')p(\theta')}{v_{x\theta}(x^*(\theta'), \theta')}$$

6 Characterization of the Optimal Mechanism

6.1 Proof Strategy

In this section we will use the implementability conditions from section 4 and the optimality conditions from section 5 to characterize the optimal policy. It is important to note that the optimality conditions from section 5 are derived using variational calculus, and thus it is natural to reduce the space of policies to continuous policies. However, the space of continuous feasible policies endowed with the sup-norm is not closed. Noting that discontinuous feasible policies that are the limit of feasible continuous policies can be extended to a convex-valued correspondence by adding the values between the left and right limits without breaking any I.C.C., as seen in figure 14.

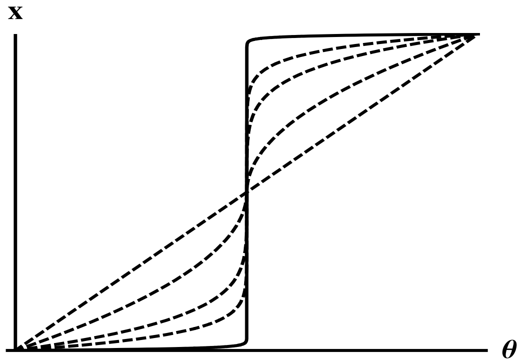

Figure 14: Continuous policies

Following Araujo and Moreira [1] we define the extended version of $x(\cdot)$ as the correspondence that contains all the values between the right and left limit of $x(\cdot)$.

Definition 16. *We define χ as the space of extended policies, or in other words the space of convex-valued correspondence.*

We will look for the optimal policy in χ [5].

First, note that for any feasible policy $x(\cdot) \in \chi$ there exists θ_0 such that $\forall \theta > \theta_0$ $x(\cdot)$ lies above $x_0(\cdot)$ and $\forall \theta < \theta_0$ $x(\cdot)$ lies below $x_0(\cdot)$, therefore $x(\cdot) \in \chi$ is quasi-convex. Thus we can identify 4 types of geometries:

- non-increasing
- non-decreasing
- U-Shaped in which $x(\bar{\theta}) \geq x(\underline{\theta})$
- U-Shaped in which $x(\bar{\theta}) < x(\underline{\theta})$

In subsection 6.2 we show that the problem of finding the optimal policy can be reduced to finding the optimal policy in the following three types of policies :

- Increasing policies, which are given by the policies $x(\cdot) \in \chi$ such that $x_\theta(\theta) \geq 0$ for all $\theta \in \Theta$
- U-Shaped policies, which are given by policies $x(\cdot) \in \chi$ in which for all ξ in the codomain of $x(\cdot)$ (except for $x(\theta_0)$) there exists θ, θ', such that $x(\theta) = x(\theta') = \xi$ and $v_{x\theta}(\xi, \theta) < 0 < v_{x\theta}(\xi, \theta')$.
- Decreasing policies, which are given by the policies $x(\cdot) \in \chi$ such that $x_\theta(\theta) \leq 0$ for all $\theta \in \Theta$

[5] Araujo and Moreira [1] show this restriction may actually reduce the principal's utility

The I.C.C. of the three types of policies previously explained present different challenges, and thus we develop different techniques for each of them. We proceed to give an overview of the challenges each type of policy presents and the respective characterization we give of the optimum.

Increasing Policies: For the increasing policies the local I.C.C. are sufficient to satisfy the global I.C.C., and thus they can be solved as if the S.C.P. was fulfilled. For this reason we do not speak at length of this kind of policies.

Decreasing Policies: The decreasing policies present the problem explained section 1, which we will now revise. It can be seen geometrically as trying to find the curve as close as possible to x_0, but keeping the shaded area positive for all θ, θ'.

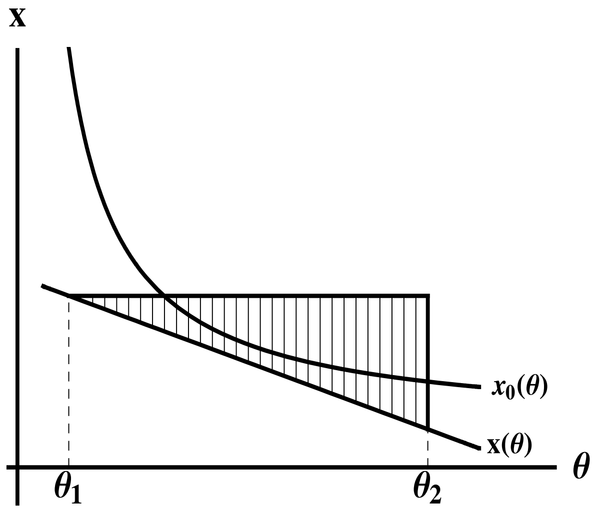

Figure 15: Decreasing policy III

It is easy to see that in this case it may be possible to modify a policy locally around some θ, and breaking the I.C.C. between some $\theta' << \theta$ and $\theta'' >> \theta$, therefore in this case the global I.C.C. are the hardest to manage since changes affect in a "global" way.

The characterization of the optimal solution in this case is done with the intuition behind corollary 13, all types θ that are kept "low" to keep the I.C. between other pair of types θ', θ'' should keep a constant ratio $\frac{f_x p}{v_{x\theta}}$. Therefore the method consists in finding a family of policies that keep the ratio $\frac{f_x p}{v_{x\theta}}$ constant (we call them isocurves), and it is obvious that among these policies the one with higher ratio are better (as long as the 1^{st} best is not achieved). For each of these policies that is not I.C. we identify critical parts of the policy that need to be "fixed" to keep the I.C.C. Through this method, for each isocurve we know that the optimal policy must lay below the isocurve in the critical parts and above it in the parts that leaves slack for improvement. This allow us to parametrize the optimal policy as a function of these isocurves.

To interpret the solution it is necessary to recall the intuition behind corollary 13, in which each type is a "good" being consumed, assignment $x(\theta)$ is the level of consumption, $f_x p$ is the marginal utility of consumption and $v_{x\theta}$ is the price paid. Therefore, the previous method is simply stating that the optimal consumption is achieved whenever all the budget is spent and the ratio between the marginal utility of a good and it's price is constant across goods. Whenever this is not possible we want to keep this ratio as high as possible.

U-Shaped Policies: In this case there exists a θ_0 such that $x(\cdot)$ is increasing for $\theta > \theta_0$ and decreasing for $\theta < \theta_0$, and therefore there are pooling types between types greater and smaller than θ_0. Thus, the main challenge is to keep the local I.C.C. for types bigger and smaller than θ_0, but having both parts of the policy agree on the transfers assigned to pooling types.

The key feature for solving this type of policy is noticing that to keep the global I.C.C. the policy for types bigger than θ_0 pin down completely and in a unique way all types lower than θ_0. Therefore, we show that this is equivalent to solving only for types bigger than θ_0, in which case only local I.C.C. need to be taken into consideration, and modifying the objective function to weight for all types smaller than θ_0 that are being pinned down. Summing up, we show that solving this type of policy is equivalent to solving a problem in which the S.C.P. is fulfilled, but with a modified objective function.

6.2 Generalized Procedure

Now we explain the general procedure, and how the optimization problem is broke down into simpler subproblems. Using lemma 7 for any feasible policy $x(\cdot) \in \chi$ and $\theta > \theta'$ the following must hold:

$$v_x(x(\theta), \theta) = v_x(x(\theta), \theta')^6 \Rightarrow x(\theta') = x(\theta)$$

That is to say, for any given type θ such that the point $(x(\theta), \theta)$ is in CS_+ and the point $(x(\theta), \theta')$ is its reflection, type θ' must be pooled with θ. For this reason, it is important to define the sub-region of CS_+ in which the points have a reflection. For this reason we introduce the function $\Sigma(\cdot)$, which gives the biggest assignment such that types have a reflection.

Definition 17. *Let's define the function $\Sigma(\cdot)$ implicitly by $v_x(\Sigma(\theta), \underline{\theta}) = v_x(\Sigma(\theta), \theta)$*
It is easy to notice the following properties:

- *If $x_0(\underline{\theta})$ exists, then and $\Sigma(\underline{\theta}) = x_0(\underline{\theta})$*
- $\forall \theta \in \Theta \backslash \{\underline{\theta}\} \quad \Sigma(\theta) > x_0(\theta)$
- $\Sigma(\cdot)$ *is strictly decreasing*

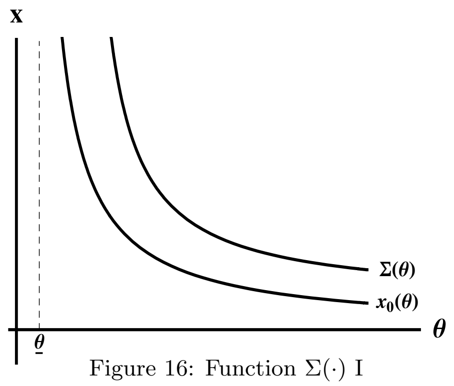

Figure 16: Function $\Sigma(\cdot)$ I

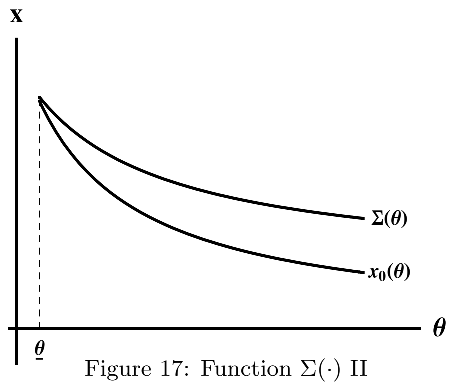

Figure 17: Function $\Sigma(\cdot)$ II

Note that for any type θ and for all assignments $x \in (x_0(\theta), \Sigma(\theta)]$ the point (x, θ) is in CS_+ and has a reflection.

We can identify 4 different types of policies:

(A) Let $x(\cdot) \in \chi$ be a feasible policy such that $x(\bar{\theta}) \in [\underline{x}, x_0(\bar{\theta})]$. In this case it is clear that $x(\cdot)$ must be non-increasing.

[6]By definition $\hat{\theta}(x(\theta), \theta) = \theta'$

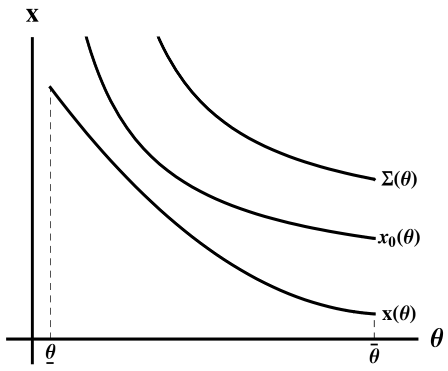

Figure 18: Subset A

We will define the subset $A \subseteq \chi$ as follows:

$$x(\cdot) \in A \iff x(\cdot) \text{ is feasible } \quad \wedge \quad x(\bar{\theta}) \leq x_0(\bar{\theta})$$

(B) Let $x(\cdot) \in \chi$ be a feasible policy such that $x(\bar{\theta}) \in (x_0(\bar{\theta}), \Sigma(\bar{\theta})]$. In this case we have that the point $(x(\bar{\theta}), \bar{\theta})$ has a reflection, given by $(x(\bar{\theta}), \hat{\theta}(x(\bar{\theta}), \bar{\theta})$[7]) and therefore $x(\hat{\theta}) = x(\bar{\theta})$, moreover using the local I.C.C. we know that:

- $x(\cdot)$ lies below $x(\bar{\theta})$ for all $\theta \in [\hat{\theta}, \bar{\theta}]$
- $x(\cdot)$ is non-increasing for all $\theta \in [\underline{\theta}, \hat{\theta}]$

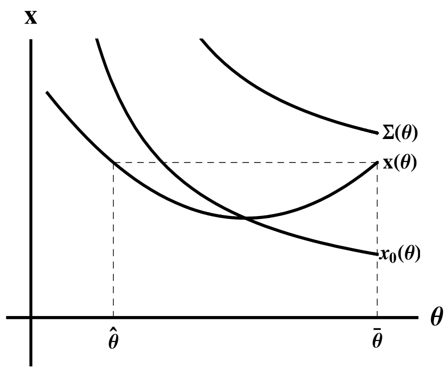

Figure 19: Subset B

We will define the subset $B \subseteq \chi$ as follows:

$$x(\cdot) \in B \iff x(\cdot) \text{ is feasible } \quad \wedge \quad x_0(\bar{\theta}) < x(\bar{\theta}) \leq \Sigma(\bar{\theta})$$

So far we have identified all policies $x(\cdot)$ such that $x(\bar{\theta}) \leq \Sigma(\bar{\theta})$, so now we proceed to describe the policies $x(\cdot)$ such that $x(\bar{\theta}) > \Sigma(\bar{\theta})$. We will consider the sub-cases in which $x(\cdot)$ and $\Sigma(\cdot)$ intersects and the case they do not (note that $x(\cdot)$ can intersect $\Sigma(\cdot)$ only if $x(\bar{\theta}) > \Sigma(\bar{\theta})$)[8].

(C) Let $x(\cdot) \in \chi$ be a feasible policy such that $x(\cdot)$ intersects with $\Sigma(\cdot)$, that is to say $\exists \theta' \in \Theta \backslash \{\underline{\theta}, \bar{\theta}\}$ such that $x(\theta') = \Sigma(\theta')$

By definition of $\Sigma(\cdot)$, the reflection of $(x(\theta'), \theta')$ is $(x(\theta'), \underline{\theta})$, therefore $x(\underline{\theta}) = x(\theta')$, moreover using the local I.C.C. we know that:

[7] To avoid excess notation we will reefer to $\hat{\theta}(x(\bar{\theta}), \bar{\theta})$ simply as $\hat{\theta}$

[8] We consider that $x(\cdot)$ and $\Sigma(\cdot)$ intersects if and only if there is an interior type θ such that $\Sigma(\theta) = x(\theta)$

- $x(\cdot)$ lies below $x(\theta')$ for all $\theta \in [\underline{\theta}, \theta']$
- $x(\cdot)$ is non-decreasing for all $\theta \in [\theta', \bar{\theta}]$

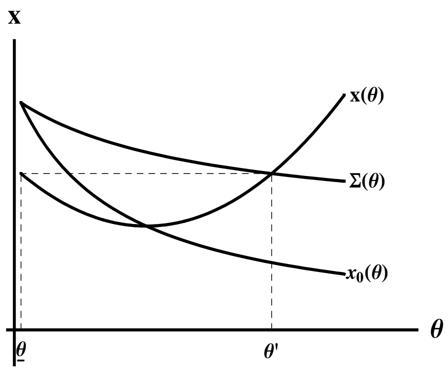

Figure 20: Subset C

We will define the subset $C \subseteq \chi$ as follows:

$$x(\cdot) \in C \iff x(\cdot) \text{ is feasible } \quad \wedge \quad \exists \theta' \in \Theta \backslash \{\underline{\theta}, \bar{\theta}\} \text{ such that } x(\theta') = \Sigma(\theta')$$

(D) Let $x(\cdot) \in \chi$ be a feasible policy such that $x(\bar{\theta}) > \Sigma(\bar{\theta})$ and $x(\cdot)$ does not intersect with $\Sigma(\cdot)$. Then it is easy to note that in this case $x(\underline{\theta}) \geq x_0(\underline{\theta})$ and $x(\cdot)$ must be non-decreasing.

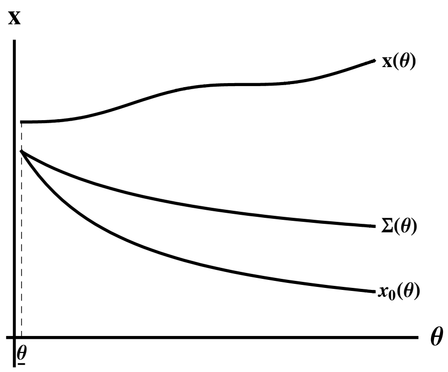

Figure 21: Subset D

We will define the subset $D \subseteq \chi$ as follows:

$$x(\cdot) \in D \iff x(\cdot) \text{ is feasible } \quad \wedge \quad x(\underline{\theta}) \geq x_0(\underline{\theta})$$

A direct proposition from the previous description is the following:

Lemma 18. *Let $x(\cdot) \in \chi$ be a feasible policy, then one and only one of the following conditions must hold:*

- *$x(\cdot) \in A$, or equivalently $x(\bar{\theta}) \leq x_0(\bar{\theta})$.*
  In this case $x(\cdot)$ must be non-increasing.
- *$x(\cdot) \in B$, or equivalently $x_0(\bar{\theta}) < x(\bar{\theta}) \leq \Sigma(\bar{\theta})$.*
  In this case $x(\cdot)$ must be U-Shaped in $[\hat{\theta}, \bar{\theta}]$ and non-increasing in $[\underline{\theta}, \hat{\theta}]$
- *$x(\cdot) \in C$, or equivalently $\exists \theta' \in \Theta \backslash \{\underline{\theta}, \bar{\theta}\} \quad x(\theta') = \Sigma(\theta')$.*
  In this case $x(\cdot)$ must be U-Shaped in $[\underline{\theta}, \theta']$and non-decreasing in $[\theta', \bar{\theta}]$
- *$x(\cdot) \in D$, or equivalently $x(\underline{\theta}) > x_0(\underline{\theta})$.*
  In this case $x(\cdot)$ must be non-decreasing.

Proof. Direct from the previous description. □

Now we will proceed to explain how to find the optimal policy in χ. Consider the following procedure:

- Sweep over all values of $\xi \in [\underline{x}, \bar{x}]$ and proceed as follows:
  - If ξ is in $[\underline{x}, x_0(\bar{\theta})]$ denote by $\Lambda^\xi(\cdot)$ the optimal policy $x(\cdot) \in A$ such that $x(\bar{\theta}) = \xi$.
  - If ξ is in $(x_0(\bar{\theta}), \Sigma(\bar{\theta})]$ denote by $\Lambda^\xi(\cdot)$ the optimal policy $x(\cdot) \in B$ such that $x(\bar{\theta}) = \xi$.

    Using theorem 8 we can see that the problem is separable and we can find, independently, the optimal non-increasing part $x_\alpha(\cdot)$ for $\theta \in [\underline{\theta}, \hat{\theta}]$ and the optimal U-Shaped part $x_\beta(\cdot)$ for $\theta \in [\hat{\theta}, \bar{\theta}]$.
    To find the non-increasing part, we consider that $x_\alpha(\theta) = \xi$ for all $\theta \in [\hat{\theta}, \bar{\theta}]$ and proceed to find the optimal policy in $[\underline{\theta}, \hat{\theta}]$. To find the optimal U-Shaped part, we find the optimal policy in $[\hat{\theta}, \bar{\theta}]$ disregarding all I.C.C. with $\theta \in [\underline{\theta}, \hat{\theta}]$.
    Then we define $\Lambda^\xi(\cdot)$ as follows:

    $$\Lambda^\xi(\theta) = \begin{cases} x_\alpha(\theta) & \theta \in [\hat{\theta}, \bar{\theta}] \\ x_\beta(\theta) & \theta \in [\underline{\theta}, \hat{\theta}] \end{cases}$$

    Being able to separate both problems allow us to disregard some of the global I.C.C. when finding the optimal policy. This is very useful since the challenges that present the I.C.C. in the non-increasing and the U-Shaped part are quite different, therefore separating the problem allow us to use different techniques for the different parts of the policies

  - If ξ is in $(\Sigma(\bar{\theta}), \Sigma(\underline{\theta}))$ denote by $\Lambda^\xi(\cdot)$ the optimal policy $x(\cdot) \in C$ such that $x(\cdot)$ intersects with $\Sigma(\cdot)$ at $\theta' = \Sigma^{-1}(\xi)$, that is to say $x(\Sigma^{-1}(\xi)) = \xi$.

    Just like before, using theorem 8 we can see that the problem is separable and we can find, independently, the optimal non-decreasing part $x_\alpha(\cdot)$ for $\theta \in [\theta', \bar{\theta}]$ and the optimal U-Shaped part $x_\beta(\cdot)$ for $\theta \in [\underline{\theta}, \theta']$.
    To find the non-decreasing part, we consider that $x_\alpha(\theta) = \xi$ for all $\theta \in [\underline{\theta}, \theta']$ and proceed to find the optimal policy in $[\theta', \bar{\theta}]$. To find the optimal U-Shaped part, we find the optimal policy in $[\underline{\theta}, \theta']$ disregarding all I.C.C. with $\theta \in [\theta', \bar{\theta}]$.
    Then we define $\Lambda^\xi(\cdot)$ as follows:

    $$\Lambda^\xi(\theta) = \begin{cases} x_\alpha(\theta) & \theta \in [\theta', \bar{\theta}] \\ x_\beta(\theta) & \theta \in [\underline{\theta}, \theta'] \end{cases}$$

  - If ξ is in $[x_0(\underline{\theta}), \bar{x}]$ denote by $\Lambda^\xi(\cdot)$ the optimal policy $x(\cdot) \in D$ such that $x(\underline{\theta}) = \xi$.
- Find the optimal policy $x^*(\cdot)$ by maximizing over all ξ in $[\underline{x}, \bar{x}]$, that is to say:

$$\xi^* = argmax_\xi \int_\Theta f(\Lambda^\xi(\theta), \theta) p(\theta) d\theta$$

$$x^*(\cdot) = \Lambda^{\xi^*}(\cdot)$$

It is easy to see that the previous procedure reduces the problem of finding the optimal policy $x^*(\cdot)$ into finding three types of policies:

Decreasing Policies: For all ξ in $[\underline{x}, \Sigma(\bar{\theta})]$ finding the optimal non-increasing policy $x(\cdot) \in \chi$ such that $x(\bar{\theta}) = \xi$.

Increasing Policies: For all ξ in $[\Sigma(\bar{\theta}), \bar{x}]$ find the optimal non-decreasing policy $x(\cdot) \in \chi$ such that $x(\underline{\theta}) = \xi$.

U-Shaped Policies: For all ξ in $[x_0(\bar{\theta}), x_0(\underline{\theta})]$ find the optimal U-Shaped policy $x(\cdot) \in \chi$ such that $\forall \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}] \quad x(\theta) = \xi$. Where $\{\theta_1, \theta_2\}$ are given by:

$$\{\theta_1, \theta_2\} = \begin{cases} \{\hat{\theta}, \bar{\theta}\} & \xi \in [x_0(\bar{\theta}), \Sigma(\bar{\theta})] \\ \{\underline{\theta}, \Sigma^{-1}(\xi)\} & \xi \in [\Sigma(\bar{\theta}), x_0(\underline{\theta})] \end{cases}$$

We now proceed to explain how to find each type of policy.

6.3 Increasing Policies

Lemma 19. *Let $x(\cdot) \in \chi$ be a non-decreasing policy, then the local I.C.C. are necessary and sufficient to guarantee the global I.C.C.*

Proof. Using lemma 19 it is clear that finding the increasing policies can be done as if the S.C.P. was fulfilled, and thus require no extra explanaition. □

6.4 Optimal U-Shaped policy

In this section we will find the optimal U-Shaped policy for the zone B of the ironing (the U-shaped form of zone C can be solved the same way). Therefore, we will find the optimal policy $\Lambda^{\xi}(\cdot)$ in $[\hat{\theta}(\bar{\theta}, \xi), \bar{\theta}]$, such that $\Lambda^{\xi}(\bar{\theta}) = \xi$. In what follows we will denote $\Lambda^{\xi}(\cdot)$ by $\Lambda(\cdot)$ and $\hat{\theta}(\bar{\theta}, \xi)$ simply by $\hat{\theta}$. We will use a different approach than the one found in Araujo and Moreira [1] which will allow us to cover a broader range of problems, although the basic intuitions are similar.

Intuitively, in any incentive compatible policy in χ the part that lies in CS_- is the reflection of the part that lies in CS_+, thus any incentive compatible policy in χ is uniquely determined by the section that lies in CS_+. Therefore we just need to find the intersection θ_0 between the policy and $x_0(\cdot)$, and the optimal shape of the policy on CS_+, $\tilde{x}(\cdot)$. With this, and considering a modified objective function that takes into account the effect of $\tilde{x}(\cdot)$ on the policy on CS_- we can solve the problem.

Lemma 20. *The optimal U-shaped part of $\Lambda(\cdot)$, such that $\Lambda(\bar{\theta}) = \xi$, can be defined by a type θ_0 and a non-decreasing policy $\tilde{x}(\cdot)$ such that $\tilde{x}(\theta_0) = x_0(\theta_0)$ and $\tilde{x}(\bar{\theta}) = \xi$, where $\Lambda(\cdot)$ is given by:*

$$\Lambda(\theta) = \begin{cases} \tilde{x}(\theta) & \forall \theta \in [\theta_0, \bar{\theta}] \\ \tilde{x}(\hat{\theta}(\Lambda(\theta), \theta)) & \forall \theta \in [\hat{\theta}, \theta_0) \end{cases}$$

θ_0 and $\tilde{x}(\cdot)$ are the solution to the following problem[9]:

$$\Psi := \min_{\theta_0, \tilde{x}(\cdot)} \int_{x_0(\theta_0)}^{\xi} \int_{x_0^{-1}(z)}^{\tilde{x}^{-1}(z)} f_x(z, y)p(y) - f_x(z, \hat{\theta}(z, y))p(\hat{\theta}(z, y)) \frac{v_{x\theta}(z, y)}{v_{x\theta}(z, \hat{\theta}(z, y))} dz dy$$

$$s.t. \begin{cases} \tilde{x}(\cdot) \text{ is non-decreasing} \\ \tilde{x}(\theta_0) = x_0(\theta_0) \\ \tilde{x}(\bar{\theta}) = \xi \end{cases}$$

Problem Ψ consists in minimizing the area enclosed by the curves $x_0(\cdot)$, $\tilde{x}(\cdot)$ and ξ (shaded area in figure 22), weighted by $f_x(x, \theta)p(\theta) - f_x(x, \hat{\theta}(x, \theta))p(\hat{\theta}(x, \theta)) \frac{v_{x\theta}(x, \theta)}{v_{x\theta}(x, \hat{\theta}(x, \theta))}$.

[9] If $\tilde{x}(\cdot)$ has a bunching zone, the inverse $\tilde{x}^{-1}(\cdot)$ is not uniquely defined. In case $\tilde{x}^{-1}(\cdot)$ is not uniquely defined we consider the smallest element in $\tilde{x}^{-1}(\cdot)$

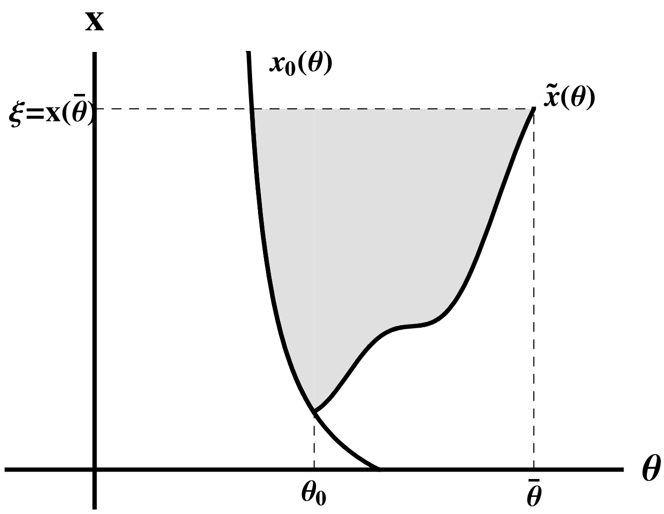

Figure 22: Optimal U-Shape I

Proof. The proof of lemma 20 consists of two parts. We first prove that any I.C. policy $\gamma(\cdot)$ such that $\gamma(\bar{\theta}) = \xi$ is determined by a type θ_0 and a non-increasing policy $\tilde{x}(\cdot)$ defined in $[\theta_0, \bar{\theta}]$. Then we prove that the characterization given by Ψ yields the optimal policy.

Step1

- Since $v_{x\theta}(\xi, \bar{\theta}) > 0$ and $v_{x\theta}(\xi, \hat{\theta}) < 0$, then the policy $\gamma(\cdot)$ must pass through $x_0(\cdot)$, then there exists θ_0 such that $\gamma(\theta_0) = x_0(\theta_0)$.
- $x_0(\cdot)$ is decreasing and $\gamma(\cdot)$ must be non-increasing in CS_+, thus θ_0 is uniquely defined and $\gamma(\cdot)$ lies in CS_+ for all $\theta > \theta_0$.
- We will denote by $\tilde{x}(\cdot)$ the section of $\gamma(\cdot)$ that lies in CS_+ $(\theta > \theta_0)$.
- It is easy to note that for all θ in $[\theta_0, \bar{\theta}]$ the policy $\tilde{x}(\cdot)$ must have a reflective type. To see this note that $\hat{\theta}(\xi, \theta)$ is continuous in both arguments, $\hat{\theta}(\tilde{x}(\theta_0), \theta_0) = \theta_0$, $\hat{\theta}(\tilde{x}(\bar{\theta}), \bar{\theta}) = \hat{\theta}$. Moreover, for any type $\theta' \in [\hat{\theta}, \theta_0)$ the policy $\gamma(\theta')$ can be found as a reflection of a type $\theta \in (\theta_0, \bar{\theta}]$, that is:

$$\Big(\forall \theta' \in [\hat{\theta}, \theta_0]\Big)\Big(\exists \theta \in (\theta_0, \bar{\theta}]\Big)\Big(\exists \xi \in \tilde{x}(\theta)\Big) \quad \text{such that} \quad \theta' = \hat{\theta}(\xi, \theta)$$

- Therefore, the policy $\gamma(\cdot)$ is uniquely defined by a type θ_0 and a non-increasing policy $\tilde{x}(\cdot)$ in CS_+. Moreover, since $\frac{\partial \hat{\theta}(\theta,\xi)}{\partial \xi} < 0$ and $\frac{\partial \hat{\theta}(\theta,\xi)}{\partial \theta} < 0$, $\tilde{x}(\cdot)$ non-decreasing implies that it's reflection is non-increasing, and thus $\gamma(\cdot)$ satisfies the local I.C.C.
- By construction pooling types have the same transfers. Using lemma 19 it is easy to see that global I.C.C. are satisfied for types greater than θ_0, and thus we only need to show that global I.C.C. are satisfied for types smaller than θ_0. Let's take $\theta' < \theta'' < \theta_0$, in this case it is easy to see that the local I.C.C. are a sufficient condition for $\Phi^\gamma(\theta', \theta'') > 0$, on the other hand we have that:

$$\Phi^\gamma(\theta'', \theta') = \int_{\theta'}^{\theta''} \int_{\gamma(\theta')}^{\gamma(y)} v_{x\theta}(y, z) dz dy = \int_{\gamma(\theta')}^{\gamma(\theta'')} \int_{\gamma^{-1}(z)}^{\theta''} v_{x\theta}(z, y) dy dz$$

Noting that:

$$\int_{\gamma^{-1}(z)}^{\hat{\theta}(z, \gamma^{-1}(z))} v_{x\theta}(z, y) dy = 0 \text{ (By definition of } \hat{\theta}(\cdot, \cdot))$$

we get,

$$\int_{\gamma^{-1}(z)}^{\theta''} v_{x\theta}(z,y)dy = \underbrace{\int_{\gamma^{-1}(z)}^{\hat{\theta}(z,\gamma^{-1}(z))} v_{x\theta}(z,y)dy}_{=0} - \int_{\theta''}^{\hat{\theta}(z,\gamma^{-1}(z))} v_{x\theta}(z,y)dy$$

$$= -\int_{\theta''}^{\hat{\theta}(\gamma(\theta''),\theta'')} v_{x\theta}(z,y)dy - \int_{\hat{\theta}(\gamma(\theta''),\theta'')}^{\hat{\theta}(z,\gamma^{-1}(z))} v_{x\theta}(z,y)dy$$

Note that $\int_{\theta''}^{\hat{\theta}(\gamma(\theta''),\theta'')} v_{x\theta}(\gamma(\theta''),y)dy = 0$, and since $v_{xx\theta} > 0$ we know that $\int_{\theta''}^{\hat{\theta}(\gamma(\theta''),\theta'')} v_{x\theta}(z,y)dy > 0$ (remember $z > \gamma(\theta'')$), moreover it is easy to see that $\int_{\hat{\theta}(\gamma(\theta''),\theta'')}^{\hat{\theta}(z,\gamma^{-1}(z))} v_{x\theta}(z,y)dy > 0$, therefore,

$$\forall z \quad \int_{\gamma^{-1}(z)}^{\theta''} v_{x\theta}(z,y)dy < 0 \Rightarrow \int_{\gamma(\theta'')}^{\gamma(\theta')} \left(-\int_{\gamma^{-1}(z)}^{\theta''} v_{x\theta}(z,y)dy \right) dz > 0 \Rightarrow \Phi^{\gamma}(\theta'',\theta') > 0$$

Step 2

Now we will prove that the characterization given for θ_0 and $\tilde{x}(\cdot)$ yields the optimal policy.

- Note that by subtracting a constant to the maximization problem we can rewrite the problem as follows:

$$\Lambda^{\xi}(\cdot) \in \underset{\text{s.t. I.C.C.}}{argmax_{\gamma(\cdot)}} \int_{\hat{\theta}}^{\bar{\theta}} f(\gamma(\theta),\theta)p(\theta)d\theta \iff \Lambda^{\xi}(\cdot) \in \underset{\text{s.t. I.C.C.}}{argmax_{\gamma(\cdot)}} \int_{\hat{\theta}}^{\bar{\theta}} f(\gamma(\theta),\theta)p(\theta)d\theta - \int_{\hat{\theta}}^{\bar{\theta}} f(\xi,\theta)p(\theta)d\theta$$

$$\iff \Lambda^{\xi}(\cdot) \in \underset{\text{s.t. I.C.C.}}{argmax_{\gamma(\cdot)}} \int_{\hat{\theta}}^{\bar{\theta}} \int_{\xi}^{\gamma(\theta)} f_x(\gamma(z),\theta)p(\theta)dzd\theta$$

- Rewriting the term $\int_{\hat{\theta}}^{\bar{\theta}} \int_{\xi}^{\gamma(\theta)} f_x(\gamma(z),\theta)dzd\theta$ conveniently.
  Let's denote θ_0' implicitly by $\gamma(\theta_0') = x_0(\theta_0')$, note that $\gamma(\theta_0')$ is the lowest assignments of policy $\gamma(\cdot)$. Note that for all $\xi \in (\gamma(\theta_0'), \xi]$ the policy $\gamma(\cdot)$ assigns ξ to at least two types. Let's define $\gamma_+^{-1}(\cdot)$ and $\gamma_-^{-1}(\cdot)$ as follows:

$$\gamma_+^{-1}(\xi) = \min\{\theta \in (x_0^{-1}(\xi), \bar{\theta}) | \gamma(\theta) = \xi\}$$

$$\gamma_-^{-1}(\xi) = \max\{\theta \in (\hat{\theta}, x_0^{-1}(\xi)) | \gamma(\theta) = \xi\}$$

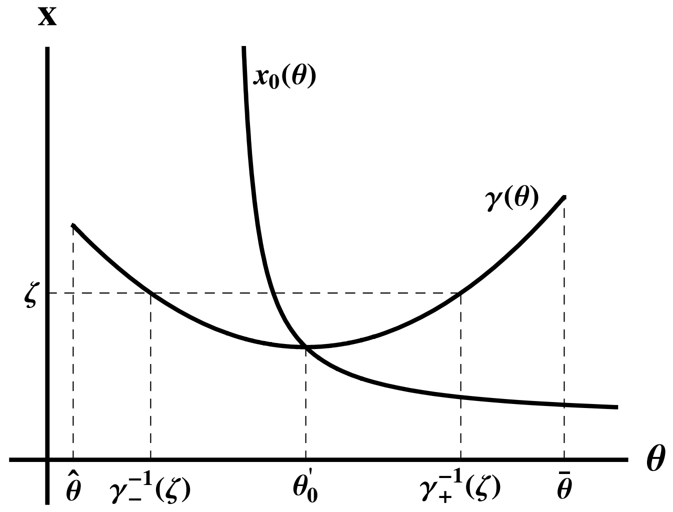

Figure 23: U-Shape policy II

Rewriting the integrals:

$$\int_{\hat{\theta}}^{\bar{\theta}}\int_{\gamma(y)}^{\xi} f_x(z,y)p(y)dzdy = \int_{\gamma(\theta_0')}^{\xi}\int_{\gamma_-^{-1}(z)}^{\gamma_+^{-1}(z)} f_x(z,y)p(y)dydz$$

$$\int_{\gamma(\theta_0')}^{\xi}\int_{\gamma_-^{-1}(z)}^{x_0^{-1}} f_x(z,y)p(y)dydz + \int_{\gamma(\theta_0')}^{\xi}\int_{x_0^{-1}(z)}^{\gamma_+^{-1}(z)} f_x(z,y)p(y)dydz$$

Note that for all $z \in [\gamma(\theta_0'), \xi]$ $\quad \hat{\theta}(z, \gamma_+^{-1}(z)) = \gamma_-^{-1}(z)$. Using the following change of variable for the last integral:

$$v_x(z,y) = v_x(z,y') \Rightarrow v_{x\theta}(z,y)dy = v_{x\theta}(z,y')dy'$$

but by definition $y = \hat{\theta}(z,y')$, thus:

$$\Rightarrow \int_{\gamma(\theta_0')}^{\xi}\int_{\gamma_-^{-1}(z)}^{x_0^{-1}} f_x(z,y)p(y)dydz = \int_{\gamma(\theta_0')}^{\xi}\int_{\gamma_+^{-1}(z)}^{x_0^{-1}} f_x(z,\hat{\theta}(z,y'))p(\hat{\theta}(z,y'))\frac{v_{x\theta}(z,y')}{v_{x\theta}(z,\hat{\theta}(z,y'))}dy'dz$$

thus,

$$\int_{\hat{\theta}}^{\bar{\theta}}\int_{\gamma(y)}^{\xi} f_x(z,y)p(y)dzdy = \int_{\gamma(\theta_0')}^{\xi}\int_{x_0^{-1}(z)}^{\gamma_+^{-1}(z)} f_x(z,y)p(y) - f_x(z,\hat{\theta}(z,y))p(\hat{\theta}(z,y))\frac{v_{x\theta}(z,y)}{v_{x\theta}(z,\hat{\theta}(z,y))}dydz$$

We have that $\gamma_+^{-1}(z) = \tilde{x}^{-1}(\cdot)$, and as we previously showed $\tilde{x}(\cdot)$ non-decreasing is sufficient for implementability. Thus we get the following:

$$\tilde{x}(\cdot) \in \underset{\text{s.t. } x(\cdot) \text{ non-decreasing}}{argmin_{x(\cdot)} \int_{\gamma(\theta_0')}^{\xi}\int_{x_0^{-1}(z)}^{x^{-1}(z)} f_x(z,y)p(y) - f_x(z,\hat{\theta}(z,y))p(\hat{\theta}(z,y))\frac{v_{x\theta}(z,y)}{v_{x\theta}(z,\hat{\theta}(z,y))}dydz}$$

$$\iff \Lambda^{\xi}(\cdot) \in \underset{\text{s.t. I.C.C.}}{argmax_{\gamma(\cdot)} \int_{\hat{\theta}}^{\bar{\theta}} f(\gamma(\theta),\theta)d\theta}$$

. ◻

Lemma 20 turns what may seem a very difficult problem into a very tractable one. There are several cases that can easily be solved using this approach, we will give a couple of examples.

Example 21. *We will now show we can recover the case shown in Araujo Moreira [1]. This is the case in which there is a unique non-decreasing curve $x_u(\cdot)$, such that $x_u(\cdot)$ divides the space CS_+ in two, such that $f_x(x,\theta)p(\theta) - f_x(x,\hat{\theta}(x,\theta))p(\hat{\theta}(x,\theta))\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(x,\hat{\theta}(x,\theta))}$ is negative "above" curve $x_u(\cdot)$ and positive "below" $x_u(\cdot)$*[10]

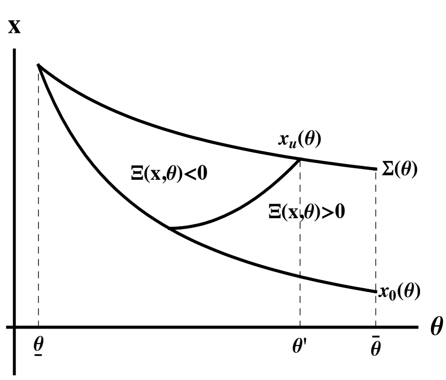

Figure 24: U-Shape example I

[10] In the following figures we use the definition $\Xi(x,\theta) = f_x(x,\theta)p(\theta) - f_x(x,\hat{\theta}(x,\theta))p(\hat{\theta}(x,\theta))\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(x,\hat{\theta}(x,\theta))}$.

In this case, depending if ξ is in zone B or C of the procedure, it is easy to see we recover both U-Shaped parts described by Araujo Moreira [1].

If $\xi < x_u(\underline{\theta})$

In this case it is easy to notice that θ_0 solves the equation $x_0(\theta_0) = x_u(\theta_0)$ and $\tilde{x}(\cdot)$ is given by:

$$\tilde{x}(\theta) = \begin{cases} x_u(\theta) & x_u^{-1}(\xi) > \theta > \theta_0 \\ \xi & \bar{\theta} > \theta > x_u^{-1}(\xi) \end{cases}$$

and therefore Λ^{ξ} is given by:

$$\Lambda^{\xi}(\theta) = \begin{cases} \tilde{x}(\theta)\theta > \theta_0 \\ \tilde{x}(\hat{\theta}(\Lambda^{\xi}(\theta), \theta)) & \theta < \theta_0 \end{cases}$$

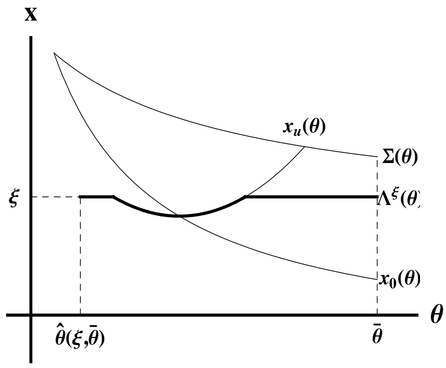

Figure 25: Optimal U-Shape example I

If $\xi > x_u(\underline{\theta})$

In this case it is easy to notice that θ_0 also solves the equation $x_0(\theta_0) = x_u(\theta_0)$ and $\tilde{x}(\cdot)$ is given by:

$$\tilde{x}(\theta) = \begin{cases} x_u(\theta) & \Sigma^{-1}(\xi) > \theta > \theta_0 \\ [x_u(\Sigma^{-1}(\xi)), \xi] & \theta = \Sigma(\xi) \end{cases}$$

and therefore Λ^{ξ} is given by:

$$\Lambda^{\xi}(\theta) = \begin{cases} \tilde{x}(\theta) & \theta > \theta_0 \\ \tilde{x}(\hat{\theta}(\Lambda^{\xi}(\theta), \theta)) & \theta < \theta_0 \end{cases}$$

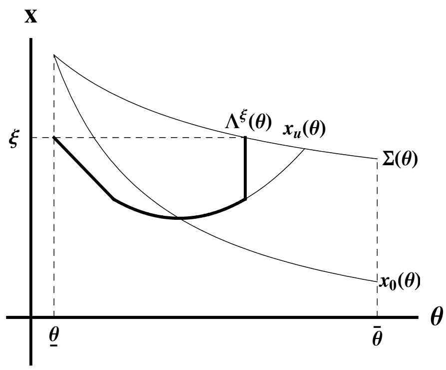

Figure 26: Optimal U-Shape example II

Example 22. *Let's consider the the case in which there is a unique non-increasing curve $x_u(\cdot)$, such that $x_u(\cdot)$ divides the space CS_+ in two, such that $f_x(x,\theta)p(\theta) - f_x(x,\hat{\theta}(x,\theta))p(\hat{\theta}(x,\theta))\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(x,\hat{\theta}(x,\theta))}$ is negative "below" curve $x_u(\cdot)$ and negative "above" $x_u(\cdot)$*

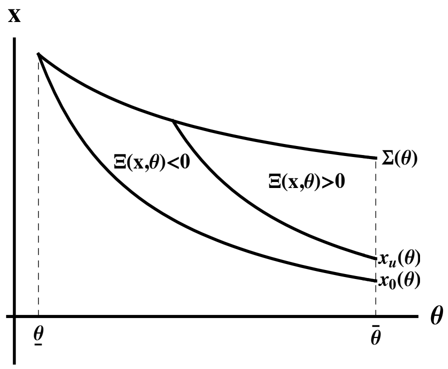

Figure 27: U-Shape example II

In this case we have that $\tilde{x}(\cdot)$ must have the following form:

$$\tilde{x}(\theta) = \begin{cases} \xi & \theta \in [\theta_0, \bar{\theta}] \\ [x_0(\theta_0), \xi] & \theta = \theta_0 \end{cases}$$

and therefore Λ^ξ is given by:

$$\Lambda^\xi(\theta) = \begin{cases} \tilde{x}(\theta) & \theta > \theta_0 \\ \tilde{x}(\hat{\theta}(\tilde{x}(\theta), \theta)) & \theta < \theta_0 \end{cases}$$

Moreover, if $f_x(x,\theta)p(\theta) - f_x(x,\hat{\theta}(x,\theta))p(\hat{\theta}(x,\theta))\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(x,\hat{\theta}(x,\theta))}$ is non-increasing in θ we have that θ_0 satisfies the following equation:

$$\int_{x_0(\theta_0)}^{\xi} f_x(x,\theta)p(\theta) - f_x(x,\hat{\theta}(x,\theta))p(\hat{\theta}(x,\theta))\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(x,\hat{\theta}(x,\theta))}dx = 0$$

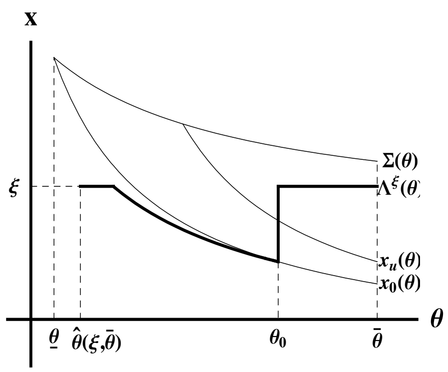

Figure 28: Optimal U-Shape example III

6.5 Optimal Decreasing Policy

In this section we will find the optimal decreasing policy for the zone B of the ironing. Therefore, we will find the optimal policy $\Lambda^\xi(\cdot)$ in $[\underline{\theta}, \hat{\theta}(\bar{\theta}, \xi)]$, such that $\Lambda^\xi(\bar{\theta}) = \xi$. We will disregard the case in which the maximum between $x_1(\cdot)$ and ξ is implementable since this case the problem is trivial. That is, if the policy $y(\cdot)$ given by:

$$y(\theta) = \begin{cases} \xi & \theta > \hat{\theta}(\bar{\theta}, \xi) \\ \max\{x_1(\theta), \xi\} & \theta \leq \hat{\theta}(\bar{\theta}, \xi)\} \end{cases}$$

is implementable, then the trivially we have that $\forall \theta < \hat{\theta}(\bar{\theta}, \xi) \quad \Lambda^{\xi}(\theta) = y(\theta)$.

In what follows we will denote $\Lambda^{\xi}(\cdot)$ by $\Lambda(\cdot)$ and $\hat{\theta}(\bar{\theta}, \xi)$ simply by $\hat{\theta}$. As we explain in subsection 6.2 this problem can be simplified by consider $\Lambda(\theta) = \xi$ for all types in $[\hat{\theta}, \bar{\theta}]$. The decreasing policy in zone A of the ironing can be found in a similar way, although it is possible to make some simplifications which we explain in remark 39.

Before discussing the optimal decreasing policy we will go back to some intuitions concerning the I.C.C. of decreasing policies

Lemma 23. *Let $x'(\cdot)$ and $x''(\cdot)$ be two policies such that for all θ in interval $[\hat{\theta}, \bar{\theta}]$ $x'(\theta) = x''(\theta) = \xi$ and for all $\theta \in \Theta$ $x'(\theta) \geq x''(\theta)$. Then the following condition holds:*

if x'($\cdot$) is an I.C. policy then x"($\cdot$) is also an I.C. policy

Intuitively, a policy that makes higher assignments in the decreasing part is closer to curve $x_0(\theta)$ and thus it is more likely that some global I.C.C. is broken (see fig 15)

Proof. By looking at fig 15 we can see that the global I.C.C. between θ_1 and θ_2 depends on the shaded are, and thus the closer the policy is to $x_0(\theta)$ the smaller this area is. □

Definition 24. *Since the ratio $\frac{f_x \cdot p}{v_{x\theta}}$ will play a crucial role on the optimal policy we will make the following definition:*

$$\Gamma(x, \theta) = \frac{f_x(x, \theta) p(\theta)}{v_{x\theta}(x, \theta)}$$

We will refer to $\Gamma(\cdot, \cdot)$ as the critical ration

Now we will make a couple of technical assumptions.

Assumption 25. *The following two assumptions are made in what follows of this subsection :*

$$\frac{\partial \Gamma(x, \theta)}{\partial x} > 0 \quad \textit{For all } (x, \theta) \textit{ in } CS_-$$

$$\frac{\partial \Gamma(x, \theta)}{\partial \theta} > 0 \quad \textit{For all } (x, \theta) \textit{ in } CS_-$$

The domain of the assumption is slightly stronger than we need. Since in the decreasing policies it is never optimal for the policy to be above $x_1(\cdot)$ we could restrict the domain to all (x, θ) such that $x \leq x_1(\theta)$. In this case we would need to recurrently make special mentions in the proofs so to avoid unnecessary difficulties we make a stronger assumption, the method for the weaker assumption is the same and should become apparent to the reader. Finally, note these two assumptions are natural when $x_1(\cdot)$ is completely contained in CS_- but it is not implementable.

From the assumptions 25 a first straight forward conclusion can be obtained about the shape of the optimal decreasing policy

Lemma 26. *The optimal decreasing policy consists of bunching part in an interval $[\dot{\theta}, \hat{\theta}]$, a continuous and strictly decreasing part in $[\ddot{\theta}, \dot{\theta}]$ and a bunching part in $[\underline{\theta}, \ddot{\theta}]$, where $\dot{\theta} \in (\underline{\theta}, \hat{\theta}]$ and $\ddot{\theta} \in [\underline{\theta}, \dot{\theta})$[11] . In what follows we will refer to $\dot{\theta}$ as the point where the policy starts being strictly decreasing, and $\dot{\theta}^x$ as the point $\dot{\theta}$ corresponding to policy $x(\cdot)$, likewise we will refer to $\ddot{\theta}$ as the point where the policy stops being strictly decreasing, and $\ddot{\theta}^x$ as the point $\ddot{\theta}$ corresponding to policy $x(\cdot)$.*

[11] Only if $\Lambda(\underline{\theta}) \geq \Sigma(\bar{\theta})$ we can have that $\ddot{\theta}^{\Lambda} > \underline{\theta}$, otherwise it is easy to see we can apply the same reasoning as in the proof of this lemma to discard the bunching zone in $[\underline{\theta}, \ddot{\theta}^{\Lambda}]$

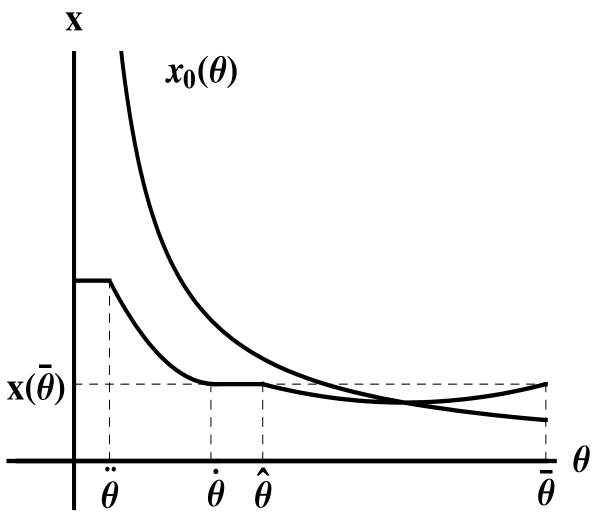

Figure 29: Optimal Decreasing Policy I

Proof. The reasoning is the same as in lemma 28, but using the assumptions 25. It is easy to convince oneself in any bunching zone, which is encompassed by two decreasing zones, can have the value of the critical ratio smoothen by increasing the value of the policy in the beginning of the bunching zone and decreasing the value of policy at the end of the bunching zone (see fig 30):

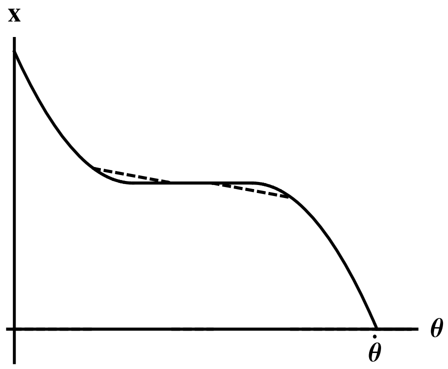

Figure 30: Bunching Policy I

Likewise, smoothening a discontinuous jump by increasing the policy at the right of the discontinuous jump and decreasing it at left of the jump effectively smoothens the value of the value of the critical ratio in the policy (see fig 31):

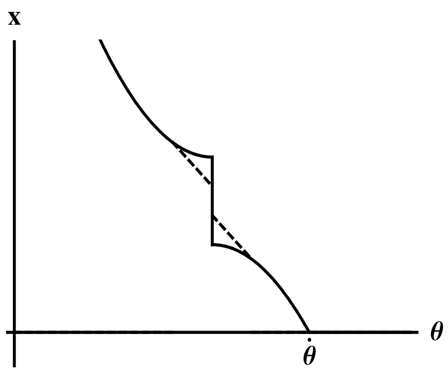

Figure 31: Discontinuous Policy I

The techincal details can be found in the appendix. □

The next lemma (which is a direct consequence of theorem 11) states that the optimal policy tends to "evens up" the value of the critical ratio across the policy, which only changes when the I.C.C. are binding.

Lemma 27.

Let θ_1 be such that $\Gamma(\Lambda(\theta),\theta)$ is strictly increasing in θ_1, then θ_1 must be indifferent between its assignment and the assignment given to some type $\theta_2 < \theta_1$. On the other hand, if θ_1 is such that $\Gamma(\Lambda(\theta),\theta)$ is strictly decreasing in θ_1, then some type $\theta_2 > \theta_1$ is indifferent between its assignment and the assignment given to θ_1. That is to say:

- *If $\theta_1 \in (\underline{\theta}, \hat{\theta}]$ is such that $\frac{d\Gamma(\Lambda(\theta_1),\theta_1)}{d\theta_1} > 0$, then exists a $\theta_2 < \theta_1$ such that $\Phi^\Lambda(\theta_1,\theta_2) = 0$*
- *If $\theta_1 \in (\underline{\theta}, \hat{\theta}]$ is such that $\frac{d\Gamma(\Lambda(\theta_1),\theta_1)}{d\theta_1} < 0$, then exists a $\theta_2 > \theta_1$ such that $\Phi^\Lambda(\theta_2,\theta_1) = 0$*

Proof. It is a reformulation of theorem 11 □

The previous lemma states that the variations of the value of the critical ratio must be as small as possible. The next lemma states two properties on types that are binding. The first, which is a direct consequence of lemma 15, says that binding types can always be jointly varied, thus it is never optimal for them to have different value of the critical ratio The second property says that types in-between a pair of types that are binding must have lower value of the critical ratio than the binding types on the edges, otherwise this would also leaves slack to have lower variations in the value of the critical ratio by averaging the changes in the middle with the edges.

Lemma 28.

Consider $\theta_2, \theta_1 \in \Theta$ such that $\theta_1 < \theta_2$ and θ_2 is indifferent between its assignment and the assignment given to θ_1, that is $\Phi^\Lambda(\theta_2,\theta_1) = 0$.

1. *If $\theta_1, \theta_2 \in (\ddot{\theta}^\Lambda, \dot{\theta}^\Lambda)$,then the value of the critical ratio must be the same for both types, moreover all types in-between must have a lower value of the critical ratio. That is to say:*

$$\Gamma(\Lambda(\theta_1),\theta_1) = \Gamma(\Lambda(\theta_2),\theta_2) \quad \wedge \quad \Gamma(\Lambda(\theta_1),\theta_1) \geq \Gamma(\Lambda(\theta),\theta) \quad \forall \theta \in [\theta_1,\theta_2]$$

2. *If $\theta_2 \in (\ddot{\theta}^\Lambda, \dot{\theta}^\Lambda)$ and $\theta_1 \leq \ddot{\theta}^\Lambda$ all types smaller than θ_2 must have a smaller value of the critical ratio than θ_2. That is to say:*

$$\Gamma(\Lambda(\theta_2),\theta_2) \geq \Gamma(\Lambda(\theta),\theta) \quad \forall \theta \in [\underline{\theta},\theta_2]$$

3. *If $\theta_1 \in (\ddot{\theta}^\Lambda, \dot{\theta}^\Lambda)$ and $\theta_2 \geq \dot{\theta}^\Lambda$ all types in between θ_1 and $\dot{\theta}^\Lambda$ must have a smaller value of the critical ratio than θ_1. That is to say:*

$$\Gamma(\Lambda(\theta_1),\theta_1) \geq \Gamma(\Lambda(\theta),\theta) \quad \forall \theta \in [\theta_1,\dot{\theta}^\Lambda]$$

Proof. We will start by proving item 1. The first part of 1 is direct from theorem 15, the second part can be proved by contradiction.
Suppose there exists $\theta' \in (\theta_1,\theta_2)$ such that $\Gamma(\Lambda(\theta'),\theta') > \Gamma(\Lambda(\theta_1),\theta_1)$

- Let θ'' be the biggest type in $[\theta',\theta_2]$ such that $\Gamma(\Lambda(\theta''),\theta'') = \Gamma(\Lambda(\theta'),\theta')$. That is to say, $\theta'' = \max\{\theta \in [\theta',\theta_2] | \Gamma(\Lambda(\theta),\theta) = \Gamma(\Lambda(\theta'),\theta')\}$

- From lemma 26 we have that discontinuous jumps are never optimal, and thus we know that $\theta'' < \theta_2$, moreover by continuity $\frac{d\Gamma(\Lambda(\theta),\theta)}{d\theta}|_{\theta=\theta''} < 0$.
- If $\frac{d\Gamma(\Lambda(\theta),\theta)}{d\theta}|_{\theta=\theta''} < 0.$, then theorem 11 there must exists a $\theta''' > \theta''$ such that $\Phi^{\Lambda}(\theta''',\theta'') = 0$, but by theorem 9 we know that $\theta''' \in [\theta_1, \theta_2]$.
- From theorem 15 we know that $\Gamma(\Lambda(\theta'''),\theta''') = \Gamma(\Lambda(\theta''),\theta'') = \Gamma(\Lambda(\theta'),\theta')$.
- Thus we arrive to a contradiction

The proof of item number 2 is similar as before.

- For all $\theta \in (\ddot{\theta}, \theta_2)$ we must have that $\Gamma(\Lambda(\theta_2),\theta_2) \geq \Gamma(\Lambda(\theta),\theta)$ (the proof can be done exactly as before).
- By continuity in $(\ddot{\theta}, \theta_2)$ we know that $\Gamma(\Lambda(\theta_2),\theta_2) \geq \Gamma(\Lambda(\ddot{\theta}),\ddot{\theta})$

The proof of item number 3 is similar as before.

- For all $\theta \in (\theta_1, \dot{\theta})$ we must have that $\Gamma(\Lambda(\theta_1),\theta_1) \geq \Gamma(\Lambda(\theta),\theta)$ (the proof can be done exactly as before).
- By continuity in $(\theta_1, \dot{\theta})$ we know that $\Gamma(\Lambda(\theta_1),\theta_1) \geq \Gamma(\Lambda(\dot{\theta}),\dot{\theta})$ □

So far we have that lemma 26 gives us a general overview of how the optimal policy looks. On the other hand, lemmas 28 and 27 gives us an intuition on how the optimal policy looks like in the strictly monotonic zone, which minimizes the variations of the value of the critical ratio. Using both of these insights we define a family of policies which have the general form given by lemma 26 and keeps the value of the critical ratio constant on the strictly decreasing part.

Definition 29. *Let $\gamma[x,\theta'](\theta)$ be the isocurve that passes throught (x,θ') and keeps $\Gamma(x,\theta)$ constant. That is :*

$$\Gamma(\gamma[x,\theta'](\theta),\theta) = \Gamma(x,\theta') \quad \forall\theta$$

which is uniquely defined and decreasing in θ because of assumption 25.

As we previously explained, because of lemma 27 it is natural to define a function that is equal to ξ in an interval $[\dot{\theta},\bar{\theta}]$, keeps $\Gamma(\cdot,\cdot)$ constant for all $\theta \in [\ddot{\theta},\dot{\theta}]$ and is bunching in some interval $[\underline{\theta},\ddot{\theta}]$. Let $x[\kappa,\theta'](\cdot)$ be a function that is bunching in $[\min\{\hat{\theta},\theta'\},\bar{\theta}]$, keeps Γ constant and is truncated by κ . That is to say:

$$x[\kappa,\theta'](\theta) = \begin{cases} \xi & \theta > \min\{\hat{\theta},\theta'\} \\ Min[\kappa,\gamma[\xi,\theta'](\theta)] & \theta \leq \min\{\hat{\theta},\theta'\} \end{cases}$$

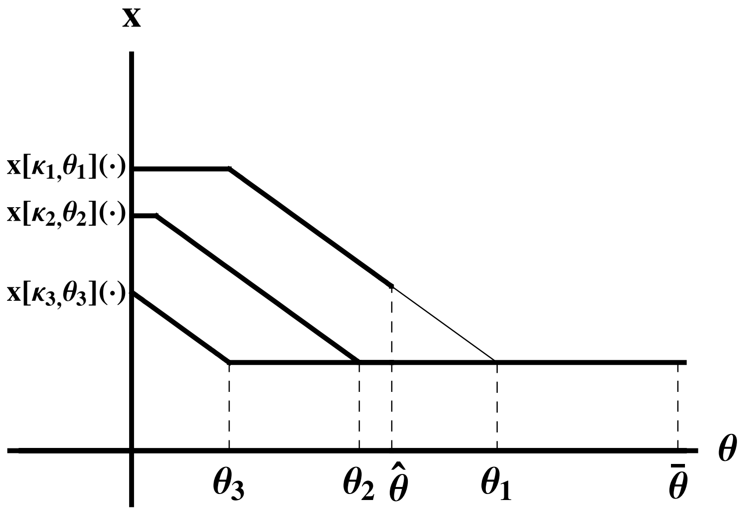

Figure 32: Function $x[\kappa,\theta](\cdot)$

The method will consist in taking a fixed assignment for $\underline{\theta}$ and characterizing the optimal policy for the given assignment, then we will maximize over all possible assignments of $\underline{\theta}$.

Definition 30. *Let $\Lambda[\kappa](\cdot)$ be the optimal policy $x(\cdot) \in \chi$ such that $x(\underline{\theta}) = \kappa$*[12]

The optimal policy will be characterized using the family of policies $x[\kappa,\theta](\cdot)$, and finding all the intersections between the functions $\Lambda[\kappa](\cdot)$ and $x[\kappa,\theta](\cdot)$, for all θ. That is to say, for each function $x[\kappa,\theta](\cdot)$ we will find a collection of types that have the same value in functions $\Lambda[\kappa](\cdot)$ and $x[\kappa,\theta](\cdot)$. To identify these critical collection of types we make the following definition:

Definition 31. *For a policy $z(\cdot)$ consider the following problem*[13]*:*

$$Q^z := \min_{n^z \in N} \left\{ \min_{\substack{\theta^z_{11} < \theta^z_{21} < \ldots < \theta^z_{1n} < \theta^z_{2n} \\ \theta^z_{21} \ldots \theta^z_{1n} \in [\ddot{\theta}^z, \dot{\theta}^z] \\ \theta^z_{11} \in [\ddot{\theta}^z, \dot{\theta}^z] \cup \{\underline{\theta}\} \\ \theta^z_{2n} \in [\ddot{\theta}^z, \dot{\theta}^z] \cup \{\bar{\theta}\}}} \sum_{j=1}^{n} \Phi^z(\theta_{2j}, \theta_{1j}) \right\}$$

We call the solutions of Q^z generalized minimum.

Remark 32.

- *Note that if Q^z is attained at a unique collection of types then $\Phi^z(\theta^z_{1j}, \theta^z_{2j}) < 0 \quad \forall j \in \{1,..,n^z\}$*
- *If $z(\cdot)$ is I.C. then $n = 0$ is a solution of Q^z and $Q^z = 0$.*
- *The solution of $Q^{x[\kappa,\theta]}$ is uniquely defined for almost for every θ (almost en el sentido discreto de verdad).*
- *Let $\{\theta^{x[\kappa,\theta']}_{11}, \ldots, \theta^{x[\kappa,\theta']}_{2n'}\}$ and $\{\theta^{x[\kappa,\theta'']}_{11}, \ldots, \theta^{x[\kappa,\theta'']}_{2n''}\}$ be solutions of problems $Q^{x[\kappa,\theta']}$ and $Q^{x[\kappa,\theta'']}$ respectively. If $\theta' < \theta''$ then $\{\theta^{x[\kappa,\theta']}_{11}, \ldots, \theta^{x[\kappa,\theta']}_{2n'}\} \backslash \{\underline{\theta}, \bar{\theta}\} \subset \bigcup_{l=1}^{n''} (\theta^{x[\kappa,\theta'']}_{1l}, \theta^{x[\kappa,\theta'']}_{2l})$. $\Big($In this case we will say $\{\theta^{y[\theta']}_{11}, \ldots, \theta^{y[\theta']}_{2n'}\}$ and $\{\theta^{x[\kappa,\theta'']}_{11}, \ldots, \theta^{x[\kappa,\theta'']}_{2n''}\}$ are nested, and we will denote $\{\theta^{x[\kappa,\theta']}_{11}, \ldots, \theta^{x[\kappa,\theta']}_{2n'}\} \prec \{\theta^{x[\kappa,\theta'']}_{11}, \ldots, \theta^{x[\kappa,\theta'']}_{2n''}\}\Big)$*

Problem Q^z gives a basic measure on how much extra rent is a policy leaving for agents to attain by lying, or in other words how much incentives are there to deviate for the agents. If the solution of Q^z is 0 then the policy $z(\cdot)$ is I.C., and thus no agent can get more than the informational rents, therefore there are no incentives for deviation. On the other hand if the solution of Q^z is strictly positive then there exists an agent that has incentives to deviate, moreover the incentives to deviate for any agent are bounded by Q^z.

Definition 33. *We will define $\vec{\vartheta}^{x[\kappa,\theta']}$ as the difference of the interior between the right and left limit of $Q^{x[\kappa,\theta]}$ at θ'*[14]*. That is to say:*

$$\vec{\vartheta}^{x[\kappa,\theta']} = \bigcup_{j=1}^{n'+} [\theta^{x[\kappa,\theta']+}_{1j}, \theta^{x[\kappa,\theta']+}_{2j}] \Big\backslash \bigcup_{j=1}^{n'-} (\theta^{x[\kappa,\theta']-}_{1j}, \theta^{x[\kappa,\theta']-}_{2j})$$

where $\{\theta^{x[\kappa,\theta']-}_{1j}, \ldots, \theta^{x[\kappa,\theta']-}_{2n'-},\}$ and $\{\theta^{x[\kappa,\theta']+}_{1j}, \ldots, \theta^{x[\kappa,\theta']+}_{2n'+},\}$ are the left and right limits respectively of $Q^{x[\kappa,\theta']}$. That is:

$$\{\theta^{x[\kappa,\theta']-}_{1j}, \ldots, \theta^{x[\kappa,\theta']-}_{2n'-},\} = \lim_{\theta \to \theta'-} \{\theta^{x[\kappa,\theta]}_{1j}, \ldots, \theta^{x[\kappa,\theta]}_{2n},\}$$

[12] Remember there is a superscript ξ in $\Lambda(\cdot)$ we are already omitting

[13] The complex domain in the maximization problem is to avoid degeneracies in the solution arising from pooling types

[14] Using the properties of $Q^{x[\kappa,\theta]}$ described in 32 it is easy to see that we can find the right and left limit of the solution at any θ

$$\{\theta_{1j}^{x[\kappa,\theta']+},...,\theta_{2n'+}^{x[\kappa,\theta']+},\} = \lim_{\theta\to\theta'+}\{\theta_{1j}^{x[\kappa,\theta]},...,\theta_{2n}^{x[\kappa,\theta]},\}$$

Note that if $Q^{x[\kappa,\theta']}$ has a unique solution, then $\vec{\vartheta}^{x[\kappa,\theta']}$ is the given solution. Note that if $Q^{x[\kappa,\theta']}$ does not have a unique solution, then $\vec{\vartheta}^{x[\kappa,\theta']}$ is the smallest subset of Θ that contains all of $Q^{x[\kappa,\theta']}$ solutions. For this reason we call $\vec{\vartheta}^{x[\kappa,\theta]}$ the extended generalized minimum of $x[\kappa,\theta]$

Lemma 34. *For all $\theta' \in (\underline{\theta},\bar{\theta})$ there exists a unique θ'' such that $\theta \in \vec{\vartheta}^{x[\kappa,\theta'']}$*

Proof.

Let's take some $\theta' \in (\underline{\theta},\bar{\theta})$, we will show there exists a unique θ'' such that $\theta' \in \vec{\varphi}^{x[\kappa,\theta'']}$.

- Note we can find a θ big enough such that $\{\underline{\theta},\bar{\theta}\}$ is a unique solution of $Q^{x[\kappa,\theta]}$.
- Note we can find a θ small enough such that $n^{x[\kappa,\theta]}=0$ is a unique solution of $Q^{x[\kappa,\theta]}$.
- Thus, we can find a θ'' and a ε small enough such that for all $\theta \in (\theta'',\theta''+\varepsilon)$ $\quad \theta' \in \bigcup_{j=1}^{n^-}[\theta_{1j}^{x[\kappa,\theta]-},\theta_{2j}^{x[\kappa,\theta]-}]$ and $\theta \in (\theta''-\varepsilon,\theta'')$ $\quad \theta' \in \bigcup_{j=1}^{n^+}[\theta_{1j}^{x[\kappa,\theta]+},\theta_{2j}^{x[\kappa,\theta]+}]$.
- Using the upper hemi-continuity of $Q^{x[\kappa,\theta]}$ we know that $\theta' \in \bigcup_{j=1}^{n''+}[\theta_{1j}^{x[\kappa,\theta'']+},\theta_{2j}^{x[\kappa,\theta'']+}]$ and $\theta' \notin \bigcup_{j=1}^{n''-}(\theta_{1j}'^{x[\kappa,\theta'']-},\theta_{2j}^{x[\kappa,\theta'']-})$
- Thus by definition $\theta' \in \vec{\vartheta}^{x[\kappa,\theta'']}$.
- Since the solutions of $Q^{x[\kappa,\theta]}$ are nested it is easy to see that θ'' is unique. $\square$

The next lemma will be the main result that will allow us to characterize the optimal policy.

Lemma 35. *If $\Lambda[\kappa](\theta) > x[\kappa,\theta'](\theta)$ for some $\theta \in (\underline{\theta},\hat{\theta})$, then*

1. $(\forall\theta \in \vec{\vartheta}^{x[\kappa,\theta']})\backslash\{\underline{\theta},\bar{\theta}\}$ $\quad \Lambda[\kappa](\theta) = x[\kappa,\theta'](\theta)$
2. $\forall\theta \in \bigcup_{j=1}^{n'}(\theta_{1j}^-,\theta_{2j}^-)\Big\backslash\{\underline{\theta},\bar{\theta}\}$ $\quad \Lambda[\kappa](\theta) < x[\kappa,\theta'](\theta)$
3. $\forall\theta \notin \bigcup_{j=1}^{n'}[\theta_{1j}^+,\theta_{2j}^+]$ $\quad \Lambda[\kappa](\theta) \geq x[\kappa,\theta'](\theta)$

Proof.

We will prove this lemma in two steps, first for the case in which $Q^{x[\kappa,\theta']}$ is uniquely defined and then the case in which $Q^{x[\kappa,\theta']}$ is not uniquely defined.

Taking the case in which $n^{x[\kappa,\theta']} > 0$ and $Q^{x[\kappa,\theta']}$ is uniquely defined:

- Since $\Lambda[\kappa](\theta) > x[\kappa,\theta'](\theta)$ for some $\theta \in (\underline{\theta},\hat{\theta})$, then it is easy to notice that exists some $\theta \in (\ddot{\theta}^{x[\kappa,\theta']},\dot{\theta}^{x[\kappa,\theta']})$ such that $\Lambda[\kappa](\theta) > x[\kappa,\theta'](\theta)$
- But, $n^{x[\kappa,\theta']} > 0$ implies that $x[\kappa,\theta'](\cdot)$ is not I.C. Since $\Lambda[\kappa](\cdot)$ is I.C., the I.C.C. constraint implies that there must exists some $\theta \in (\dot{\theta}^{x[\kappa,\theta']},\ddot{\theta}^{x[\kappa,\theta']})$ such that $\Lambda[\kappa](\theta) < x[\kappa,\theta'](\theta)$
- We can define the collection of types $\vec{\Pi} = \{\pi_{11},\pi_{21},...,\pi_{1m},\pi_{2m}\}$ such that

  $$\forall\theta \in \bigcup_{j=1}^{m}(\pi_{1j},\pi_{2j}) \quad \Lambda[\kappa](\theta) < x[\kappa,\theta'](\theta) \quad \wedge \quad \forall\theta \notin \bigcup_{j=1}^{m}(\pi_{1j},\pi_{2j}) \cup \{\underline{\theta},\bar{\theta}\} \quad \Lambda[\kappa](\theta) \geq x[\kappa,\theta'](\theta)$$

  Note $\forall\pi \in \vec{\Pi}\backslash\{\pi_{11},\pi_{2m}\}$ $\quad \pi \in (\ddot{\theta}^{x[\kappa,\theta']},\dot{\theta}^{x[\kappa,\theta']})$
- Using lemma 27 we know that $\Phi^{\Lambda[\kappa]}(\pi_{1j},\pi_{2j}) = 0 \quad \forall j \in \{1,2....,m\}$ (in case $\pi_{2m} = \dot{\theta}^{x[\kappa,\theta']}$ we have that $\Phi^{\Lambda[\kappa]}(\pi_{1m},\bar{\theta}) = 0$ and if $\pi_{11} \leq \ddot{\theta}^{x[\kappa,\theta']}$ we have that $\Phi^{\Lambda[\kappa]}(\pi_{21},\underline{\theta}) = 0$)

- Remember that the I.C.C. are given by computing some areas weighted by $v_{x\theta}$. Therefore, the sum of these areas between types $\vec{\vartheta}$ in policy $\Lambda[\kappa]$ can be at most the areas between types $\vec{\vartheta}$ in policy $x[\kappa,\theta']$, plus the areas added in between $\vec{\pi}$. But, from the previous item we know that the areas added in between $\vec{\pi}$ are given by $\sum_{j=1}^{m} \Phi^{x[\kappa,\theta']}(\pi_{1j},\pi_{2j})$. therefore, we have the following inequality:

$$\sum_{j=1}^{n'} \Phi^{\Lambda[\kappa]}(\theta_{1j}^{x[\kappa,\theta']},\theta_{2j}^{x[\kappa,\theta']}) \leq \sum_{j=1}^{n'} \Phi^{x[\kappa,\theta']}(\theta_{1j}^{x[\kappa,\theta']},\theta_{2j}^{x[\kappa,\theta']}) + \sum_{j=1}^{m} \Phi^{x[\kappa,\theta']}(\pi_{1j},\pi_{2j})$$

  But, if $\vec{\Pi} \neq \vec{\vartheta}^*(\theta')$ we would have the following inequality (remember $Q^{x[\kappa,\theta']}$ is uniquely defined):

$$\sum_{j=1}^{n'} \Phi^{\Lambda[\kappa]}(\theta_{1j}^{x[\kappa,\theta']},\theta_{2j}^{x[\kappa,\theta']}) \leq \sum_{j=1}^{n'} \Phi^{x[\kappa,\theta']}(\theta_{1j}^{x[\kappa,\theta']},\theta_{2j}^{x[\kappa,\theta']}) + \sum_{j=1}^{m} \Phi^{x[\kappa,\theta']}(\pi_{1j},\pi_{2j}) < 0$$

  So there must exist a pair of θ that are not I.C.
- Therefore we must have that $\vec{\pi} = \vec{\vartheta}^*(\theta')$, moreover by continuity we can see that $\Lambda[\kappa](\theta) = x[\kappa,\theta'](\theta) \quad \forall\theta \in (\theta_{1j}^{x[\kappa,\theta']},\theta_{2j}^{x[\kappa,\theta']}) \quad \forall j \in \{1,2....,n^*\}$
- The last point is trivial by the construction of the proof.

Now analyzing the case $n^{x[\kappa,\theta']} = 0$ and $Q^{x[\kappa,\theta']}$ is uniquely defined

- If there exists a type $\theta_1 < \dot{\theta}^{x[\kappa,\theta']}$ such that $\Lambda[\kappa](\theta_1) = x[\kappa,\theta'](\theta_1)$, then by lemma 27 there must exist a type θ_2 such that $\Lambda[\kappa](\theta_2) = x[\kappa,\theta'](\theta_2)$ and $\Phi^{\Lambda[\kappa]}(\theta_1,\theta_2) = 0$ or $\Phi^{\Lambda[\kappa]}(\theta_2,\theta_1) = 0$
- By lemma 11 we know that $\Phi^{\Lambda[\kappa]}(\theta_2,\theta_1) > \Phi^{x[\kappa,\theta']}(\theta_2,\theta_1)$
- But, if $n^{x[\kappa,\theta']} = 0$ and uniquely defined, then $\forall\theta_1,\theta_2 \in \Theta \quad \Phi^{x[\kappa,\theta']}(\theta_2,\theta_1) > 0$
- Thus, we arrive to a contradiction.

Now taking the case in which $Q^{x[\kappa,\theta']}$ is not uniquely defined:

- From lemma 32 we can find $\theta'' < \theta'$ such that $\vec{\vartheta}^{x[\kappa,\theta'']}$ is uniquely defined and $\vec{\vartheta}^{x[\kappa,\theta'']}$ is arbitrarily close to $\{\theta_{11}^{x[\kappa,\theta']-},...,\theta_{2n}^{x[\kappa,\theta']-}\}$. Using the previous result for when $\vec{\vartheta}^{x[\kappa,\theta']}$is uniquely defined, $\forall\theta \notin \bigcup_{j=1}^{n'^-}[\theta_{1j}^{x[\kappa,\theta']-},\theta_{2j}^{x[\kappa,\theta']-}] \quad \Lambda[\kappa](\theta) \geq x[\kappa,\theta'](\theta)$.
- From lemma 32 we can find $\theta'' > \theta'$ such that $\vec{\vartheta}^{x[\kappa,\theta'']}$ is uniquely defined and $\vec{\vartheta}^{x[\kappa,\theta'']}$ is arbitrarily close to $\{\theta_{11}^{x[\kappa,\theta']+},...,\theta_{2n}^{x[\kappa,\theta']+}\}$. Using the previous result for when $\vec{\vartheta}^{x[\kappa,\theta']}$is uniquely defined, $\forall\theta \in \bigcup_{j=1}^{n'^+}[\theta_{1j}^{x[\kappa,\theta']+},\theta_{2j}^{x[\kappa,\theta']+}] \quad \Lambda[\kappa](\theta) \leq x[\kappa,\theta'](\theta)$
- Thus, $\forall\theta \in \bigcup_{j=1}^{n'^+}[\theta_{1j}^{x[\kappa,\theta']+},\theta_{2j}^{x[\kappa,\theta']+}] \Big\backslash \bigcup_{j=1}^{n'^-}[\theta_{1j}^{x[\kappa,\theta']-},\theta_{2j}^{x[\kappa,\theta']-}] \quad \Lambda[\kappa](\theta) = x[\kappa,\theta'](\theta)$ □

.

Definition 36. *We define function $\Upsilon(\cdot) : \Theta \to \Theta$ implicitly as follows:*

$$\theta \in \vec{\vartheta}^{x[\kappa,\Upsilon(\theta)]}$$

The next theorem is the final characterization of the optimal policy. Given lemma 28 and 27 we can see that for all types $\Lambda[\kappa](\cdot)$ is bounded by $x[\kappa,\Upsilon(\theta)](\cdot)$, and $\Lambda[\kappa](\theta) = x[\kappa,\Upsilon(\theta)](\theta)$ for all types is I.C.C. Thus, it is clear that the optimal policy $\Lambda[\kappa](\cdot)$ will be $x[\kappa,\Upsilon(\theta)](\theta)$ whenever the $x_1(\cdot)$ is not achievable, therefore we have the following theorem:

Lemma 37. $\Lambda[\kappa](\cdot)$ *is given by:*

$$\Lambda[\kappa](\theta) = Min\{x[\kappa, \Upsilon(\theta)](\theta)\}, x_1(\theta)\} \quad \forall\theta \in [\underline{\theta}, \hat{\theta}]$$

Proof.

- Let's denote by $\tilde{\Gamma}$ the maximum of $\Gamma(\cdot,\cdot)$ in policy $\Lambda[\kappa]$ in $[\underline{\theta}, \hat{\theta}]$, and $\tilde{\theta}$ the respective type such that $\Gamma(\Lambda[\kappa](\tilde{\theta}), \tilde{\theta}) = \tilde{\Gamma}$. That is:

$$\tilde{\Gamma} = max_{\theta \in [\underline{\theta}, \hat{\theta}]} \Gamma(\Lambda[\kappa](\theta), \theta)$$

- If $\tilde{\Gamma} = 0$ ($\Gamma(\Lambda[\kappa](\theta) = 0$ for a given θ is equivalent to $\Lambda[\kappa](\theta) = x_1(\theta)$), then using lemma 35 it is easy to see that the optimal policy would be given by the characterization of lemma 37. It is also easy to see that $\tilde{\Gamma} \leq 0$, so we now take the case $\tilde{\Gamma} < 0$.
- If $\tilde{\Gamma} < 0$ and $\{\underline{\theta}, \bar{\theta}\} \in \vec{\vartheta}^{x[\kappa, \tilde{\theta}]}$, then for all $\theta \in [\underline{\theta}, \hat{\theta}]$ we have that $\Upsilon(\theta) \leq \tilde{\theta}$, and thus we can use directly lemma 35 to see that the optimal policy would be given by the characterization of lemma 37.
- If $\tilde{\Gamma} < 0$ and $\{\underline{\theta}, \bar{\theta}\} \notin \vec{\vartheta}^{x[\kappa, \tilde{\theta}]}$ then:
  - we can find an interval $[\theta_1, \theta_2]$ that doesn't intersect $\bigcup_{j=1}^{n'+} [\theta_{1j}^{x[\kappa,\theta']+}, \theta_{2j}^{x[\kappa,\theta']+}]$. That is to say:

$$[\theta_1, \theta_2] \bigcap \left(\bigcup_{j=1}^{n'+} [\theta_{1j}^{x[\kappa,\theta']+}, \theta_{2j}^{x[\kappa,\theta']+}] \right) = \phi$$

  - In this case, using lemma 35 it is easy to see that $\forall\theta \in [\theta_1, \theta_2] \quad \Lambda[\kappa](\theta) = x[\kappa, \tilde{\theta}], (\theta)$
  - By construction of lemma 35 it is easy that for any $(\theta' < \theta'') \in \Theta$ such that $[\theta', \theta''] \bigcap [\theta_1, \theta_2] \neq \phi$, $\Phi^{\Lambda[\kappa]}(\theta'', \theta') > 0$.
  - Thus we can find a function $h(\cdot) > 0$ defined in $[\theta_1, \theta_2]$ and "small enough" such that function $\Lambda[\kappa](\cdot) + h(\cdot)$ is I.C.
  - Since $\tilde{\Gamma} < 0$ for a $h(\cdot)$ "small enough" it is easy to see that it is an improvement on $\Lambda[\kappa](\cdot)$.
  - Thus, we arrive to a contradiction, therefore it is never possible to have $\tilde{\Gamma} < 0$ and $\{\underline{\theta}, \bar{\theta}\} \notin \vec{\vartheta}^{x[\kappa, \tilde{\theta}]}$. □

.

As we previously explained, the optimal policy can be found by maximizing $\Lambda[\kappa](\cdot)$ over all possible κ.

Theorem 38. *The optimal policy* $\Lambda(\cdot)$ *is given by:*

$$\Lambda(\cdot) = \max_{\kappa} \Lambda[\kappa](\cdot)$$

Remark 39. *It is easy to see that in case we are solving for the zone A of the ironing the same procedure could be done just considering* $\hat{\theta} = \bar{\theta}$ *and repeating it the same way. Nevertheless, there is a simplification that can be done for this case. It is easy to see that lemma 26 can be extended, and the optimal policy in zone A of the procedure, if* $x(\bar{\theta}) > \underline{x}$ *consists in a bunching zone* $[\underline{\theta}, \ddot{\theta}^{\Lambda[\kappa]}]$ *and a strictly decreasing zone* $[\ddot{\theta}^{\Lambda[\kappa]}, \bar{\theta}]$*. Thus, the optimal policy in zone A of the ironing can be done just by repeating the previous algorithm for the case* $x(\bar{\theta}) = \underline{x}$ *(if in the optimal policy* $x(\bar{\theta})$ *were to be greater than* $\underline{x}$ *then there will be a discontinuous jump at* $\bar{\theta}$ *which is irrelevant from the optimization point of view)*

Summing up, the previous method is based in finding the generalized minimums for the family of functions $x[\kappa,\theta](\cdot)$. All types that are part of the generalized minimum for some function $x[\kappa,\theta](\cdot)$ have the same value in the optimal policy than in the respective function $x[\kappa,\theta](\cdot)$. All types that do not belong to the generalized minimum of any function $x[\kappa,\theta](\cdot)$ are found from the closest type that is part of the generalized minimum by staying on the respective function $x[\kappa,\theta](\cdot)$. We proceed to show a pictographic example on how the optimal policy is found.

Consider the following figure in which we consider the functions $x[\kappa,\theta_1](\cdot), x[\kappa,\theta_2](\cdot), x[\kappa,\theta_3](\cdot), x[\kappa,\theta_4](\cdot)$, $x[\kappa,\theta_5](\cdot)$ and $x[\kappa,\theta_6](\cdot)$ which will help us characterize the optimal policy. We identify the solutions of $Q^{x[\kappa,\theta_i]}$ with the circles, and we use black circles if the solution consist of one pair of types, and white circles if the solution consists of two pair of types[15].

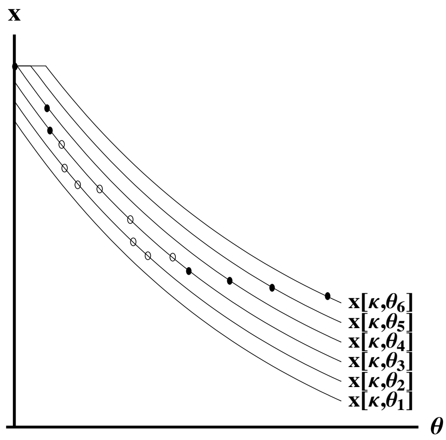

Figure 33: Example Generalized Minimum

There are some characteristics we would like to highlight:

- Note that the solution of $Q^{x[\kappa,\theta_3]}$ is not unique
- Note that $\underline{\theta}$ is part of the solution of $Q^{x[\kappa,\theta_6]}$ and $Q^{x[\kappa,\theta_5]}$.
- To exemplify we will assume θ_5 is the smallest type such that $\underline{\theta}$ belongs to it's generalized minimum and $Q^{x[\kappa,\theta_5]}$ has a unique solution.
- To exemplify we will also assume that $n^{x[\kappa,\theta_2]} = 0$ is also a solution of $Q^{x[\kappa,\theta_2]}$
- To exemplify we will assume $x_1(\cdot)$ is above $x[\kappa,\theta_6](\cdot)$ for all θ

Now we will show how the optimal policy would look like based on the solutions of $Q^{x[\kappa,\theta_i]}$. We will explain step by step how the optimal policy would be constructed, we will start by showing the assignments of types that do not belong to the generalized minimum of any function $x[\kappa,\theta](\cdot)$ and then explain how the rest is found by finding the generalized minimum of all functions $x[\kappa,\theta](\cdot)$. As we proceed we will show the optimal policy with a thick line.

- First note that $x[\kappa,\theta_3](\cdot)$ has multiple solutions, therefore the right and left limit of the generalized minimum in this case corresponds to both of it's respective solutions , and therefore all points in between both solutions are in $\vec{\vartheta}^{x[\kappa,\theta_3]}$. So, necessarily the optimal solution passes through the points that are in the difference between the interior of both solutions.

[15]Note that the figure is a pictographic representation that isn't scaled to a situation in which we could have the marked generalized minimums and the necessary hypothesis fulfilled

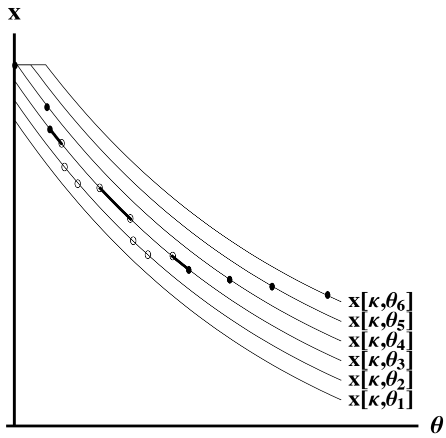

Figure 34: Construction Decreasing Policies I

- The reasoning for $x[\kappa,\theta_2](\cdot)$ is the same as for $x[\kappa,\theta_3](\cdot)$ only the left limit of the generalized minimum is empty and thus the points that are in the difference between both solutions in this case are the same than the points that are in between the only non empty solution.

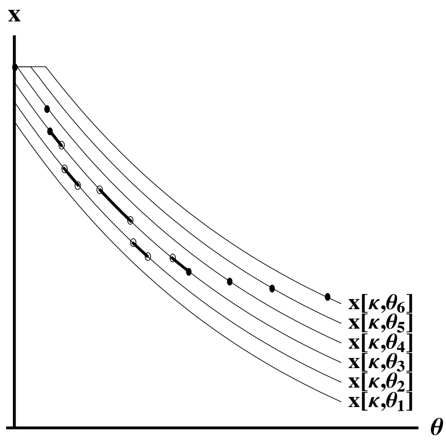

Figure 35: Construction Decreasing Policies II

- Since θ_5 is the smallest type such that $\underline{\theta}$ belongs to it's generalized minimum, we know that $\underline{\theta}$ will be part of the right limit of the generalized minimum but will not be part of the left limit of the solutions, thus $\ddot{\theta}^{x[\kappa,\theta_5]}$ will be the smallest type in the left limit of the generalized minimum. Therefore, all $\theta \in [\underline{\theta}, \ddot{\theta}^{x[\kappa,\theta_5]}]$ will be in $\vec{\vartheta}^{x[\kappa,\theta_5]}$, and thus the optimal solution passes through $x[\kappa,\theta_5](\cdot)$ for all these θ. Therefore for all $\theta \in [\underline{\theta}, \ddot{\theta}^{x[\kappa,\theta_5]}]$ the optimal solution has a value of κ

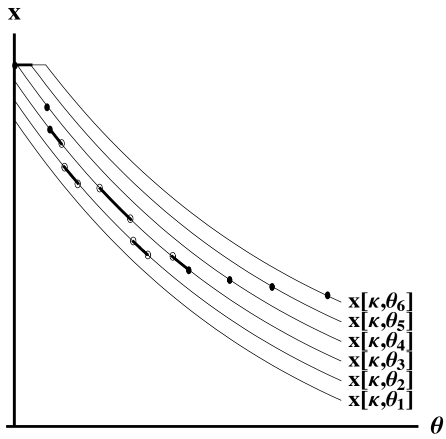

Figure 36: Construction Decreasing Policies III

- All types that have not yet being assigned belong to the generalized minimum of some function $x[\kappa, \theta]$ that are in between the functions shown in the figure. Thus all types that need to be assigned would be found by finding the generalized minimum of all functions $x[\kappa, \theta]$, which would connect all the marked dots. Connecting the dots from $x[\kappa, \theta_6]$ to $x[\kappa, \theta_2]$ we would have the following sequence

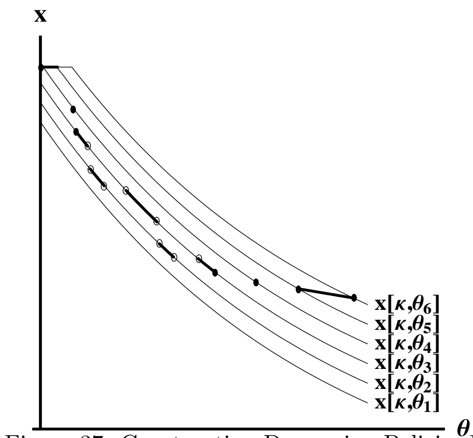

Figure 37: Construction Decreasing Policies IV

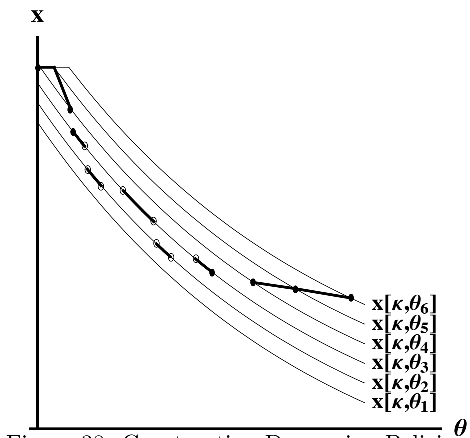

Figure 38: Construction Decreasing Policies V

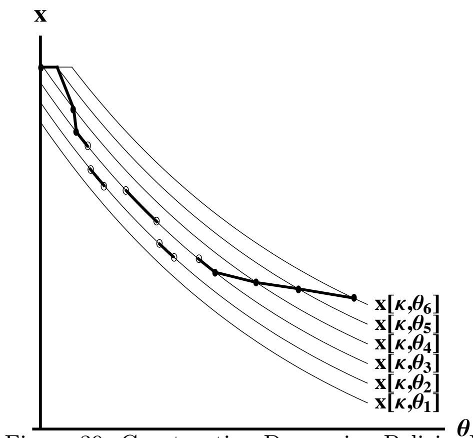

Figure 39: Construction Decreasing Policies VI

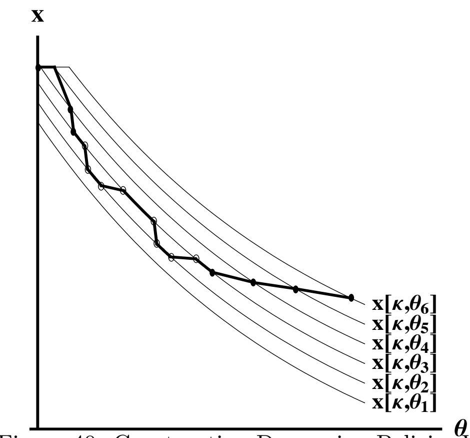

Figure 40: Construction Decreasing Policies VII

Finally we would like to give an overview of the solution of the second example explained in section 3. This problem corresponds to a particular case in which there is a unique type $\tilde{\theta}$ such that $\{\underline{\theta}, \bar{\theta}\}$ and $n^{x[\kappa,\tilde{\theta}]} = 0$ are solutions of $Q^{x[\kappa,\tilde{\theta}]}$, and thus in this case the optimal policy is obviously equal to $x[\kappa, \tilde{\theta}](\cdot)$. This is the easiest possible case to analyze, nevertheless the parameterizations needed to have multiple solutions to $Q^{x[\kappa,\theta]}$ are not simple so, even though it is the easiest case, it is a usefull case to understand.

7 Conclusions

The difficulty in approaching screening problems in which the S.C.P. is not fulfilled has lead to lack of research of this kind of problems. This does not only limit our understanding of particular models that should be studied, but also undermine our ability to understand other complex problems, like multidimensional screening. In this paper we introduce new techniques to tackle screening problems that do not satisfy the S.C.P. To approach this problem, we exploit the structure of the agent's preferences to derive necessary conditions for optimality, and variational methods to derive necessary conditions for optimality.

One of our main contributions is the new necessary conditions for implementability and optimality of a policy. We were able to interpret these conditions economically, which allows to extend the intuitions to other economic problems in which there is incomplete information. These necessary conditions also help understand the limitations that the S.C.P. poses on screening problems, thus helping understand in a better way other problems in which the standard techniques cannot be used, like multidimensional screening.

We also propose a method to find optimal policies in this model which is robust and can be applied almost independently of the principal's objective function. The general optimization problem is reduced to finding optimal U-Shaped forms, decreasing policies and increasing policies. The U-Shaped forms are characterized in the same way as problems in which the S.C.P. is fulfilled, but with a modified objective function. On the other hand, the characterization of decreasing policies is based on the concept of generalized minimums, which is a natural concept in this kind of problems, and thus we believe it could be useful to solve multidimensional screening. We exemplify the method with two examples, both of them being very natural problems, nevertheless yielding complex but interesting solutions.

The first direction in which our work could be extended is by relaxing assumption 25, this assumption guarantee the variational calculus approach is sufficient, and thus relaxing them would require new tools to approach the problem. Another interesting problem is to use this approach to tackle multidimensional screening. Since the main difficulty of multidimensional screening is the lack of the S.C.P., we believe this approach could useful in the characterization of the solution.

8 References

1. A. Araujo, H. Moreira, Adverse selection problems without the Spence-Mirrlees condition, Journal of Economic Theory, 145 (2010) 1113-1141.
2. D. Bernheim, A theory of conformity, J. Polit. Economy 102 (1994) 841-877.
3. L. Bagwell, D. Bernheim, Veblen effects in a theory of conspicuous consumption, Amer. Econ. Rev. 86 (1996) 349-373.
4. J. Quah, B. Strulovici, Aggregating the single crossing property: theory and applications to comparative statics and Bayesian games. Discussion Paper (2010) . Department of Economics (University of Oxford). (Unpublished)
5. B. Jullien, Participation constraints in adverse selection models, J. Econ. Theory 93 (2000) 1-47.
6. A. Araujo, H. Moreira, A general Lagrangian approach for non-concave moral Hazard problem, J. Math. Econ. 35 (2001) 1-23.
7. T. Lewis, D. Sappington, Countervailing incentives in agency problems, J. Econ. Theory 66 (1989) 238-263.
8. G. Maggi, A. Rodrguez-Clare, On countervailing incentives, J. Econ. Theory 66 (1995) 238-263.
9. J.-C. Rochet, P. Chon, Ironing, sweeping and multidimensional screening, Econometrica 66 (1998) 783826.

10. W.P. Rogerson, The rst-order approach to principalagent problems, Econometrica 53 (1985) 13571367.
11. A. Araujo, D. Gottlieb, H. Moreira, A model of mixed signals with applications to countersignalling, The RAND Journal of Economics Vol. 38 (2007) 10201043

9 Appendix

Theorem 8: For any given interval $[\theta_1, \theta_2]$such that $x_\alpha = \hat{x}(\theta_2) = \hat{x}(\theta_1)$, we have that the policy $\hat{x}'$ defined by:

$$\hat{x}'(\theta) = \begin{cases} \hat{x}(\theta) & \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}] \\ x_\alpha & \theta \in [\theta_1, \theta_2] \end{cases}$$

is also I.C.

It is easy to see that the opposite proposition is also true. That is to say, let $\hat{x}(\theta)$ be I.C. and such that the policy is bunching in an interval $[\theta_1, \theta_2]$. For any policy $\tilde{x}(\theta)$ defined in $[\theta_1, \theta_2]$, such that $\tilde{x}$ is I.C. and such that $x_\alpha = \tilde{x}(\theta_2) = \tilde{x}(\theta_1) = \hat{x}(\theta_1)$, then policy $\hat{x}'$ defined by:

$$\hat{x}'(\theta) = \begin{cases} \hat{x}(\theta) & \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}] \\ \tilde{x}(\theta) & \theta \in [\theta_1, \theta_2] \end{cases}$$

is also I.C.

Proof. First, note that $\forall \theta, \theta' \in [\underline{\theta}, \theta_1] \cup (\theta_2, \bar{\theta})$ we have that $\Phi^{\hat{x}}(\theta, \theta') = \Phi^{\hat{x}'}(\theta, \theta')$, thus the I.C.C. are fulfilled. We only need to prove that $\big(\forall \theta \in [\underline{\theta}, \theta_1] \cup (\theta_2, \bar{\theta})\big)\, (\forall \theta' \in (\theta_1, \theta_2))\ v(x_\alpha, \theta') + T_\alpha \geq v(\hat{x}(\theta), \theta') + \hat{T}(\theta)$. Since $\hat{x}(\cdot)$ is incentive compatible we have the following:

$$v(x_\alpha, \theta_1) + T_\alpha \geq v(\hat{x}(\theta), \theta_1) + \hat{T}(\theta) \quad \wedge \quad v(x_\alpha, \theta_2) + T_\alpha \geq v(\hat{x}(\theta), \theta_2) + \hat{T}(\theta)$$

$$\iff \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz \leq T_\alpha - \hat{T}(\theta) \quad \wedge \quad \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz \leq T_\alpha - \hat{T}(\theta)$$

If $\hat{x}(\theta) \geq x_\alpha$ and using $v_{x\theta\theta} \geq 0$ we know that
$\forall \theta' \in (\theta_1, \theta_2) \quad \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta') dz \leq \max\{\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz, \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz\}$
and thus,

$$\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz \leq T_\alpha - \hat{T}(\theta) \quad \wedge \quad \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz \leq T_\alpha - \hat{T}(\theta)$$

$$\Rightarrow \forall \theta' \in (\theta_1, \theta_2) \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta') dz \leq T_\alpha - \hat{T}(\theta) \Rightarrow v(x_\alpha, \theta') + T_\alpha \geq v(\hat{x}(\theta), \theta') + \hat{T}(\theta)$$

If $\hat{x}(\theta) \leq x_\alpha$, then we know that $\max\{\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz, \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz\} \leq \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta) dz$. Again using that $v_{x\theta\theta} \geq 0$ we know that $\forall \theta' \in (\theta_1, \theta_2)$ the function $\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta')$ is monotonic in θ', and thus $\forall \theta' \in (\theta_1, \theta_2) \quad \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta') dz \leq \max\{\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz, \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz\}$
therefore,

$$\int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_1) dz \leq T_\alpha - \hat{T}(\theta) \quad \wedge \quad \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta_2) dz \leq T_\alpha - \hat{T}(\theta)$$

$$\Rightarrow \forall \theta' \in (\theta_1, \theta_2) \int_{x_\alpha}^{\hat{x}(\theta)} v_x(z, \theta') dz \leq T_\alpha - \hat{T}(\theta) \Rightarrow v(x_\alpha, \theta') + T_\alpha \geq v(\hat{x}(\theta), \theta') + \hat{T}(\theta)$$

To prove the second part, note that $\left(\forall\theta' \in [\underline{\theta},\theta_1] \cup (\theta_2,\bar{\theta})\right)(\forall\theta \in (\theta_1,\theta_2))$ the following is true:

$$v(\hat{x}'(\theta),\theta) + \hat{T}'(\theta) \underbrace{\geq}_{\text{This is because } \tilde{x}(\theta) \text{ is I.C.}} v(\theta_\alpha,\theta) + T_\alpha \underbrace{\geq}_{\text{This is because } \hat{x}(\theta) \text{ is I.C.}} v(\hat{x}(\theta'),\theta) + \hat{T}(\theta')$$

therefore $\left(\forall\theta \in [\underline{\theta},\theta_1] \cup (\theta_2,\bar{\theta})\right)(\forall\theta' \in (\theta_1,\theta_2))$ we have that $\Phi^{\hat{x}'}(\theta,\theta') \geq 0$.

Thus, we need to prove that $\left(\forall\theta' \in [\underline{\theta},\theta_1] \cup (\theta_2,\bar{\theta})\right)(\forall\theta \in (\theta_1,\theta_2))$ we have that $v(\hat{x}'(\theta),\theta') + \hat{T}'(\theta) \geq v(\hat{x}(\theta'),\theta) + \hat{T}(\theta')$. Note that since $\tilde{x}(\cdot)$ is I.C. we have that

$\forall\theta \in (\theta_1,\theta_2)$ $\int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta)dz \leq \min\{\int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta_2)dz, \int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta_1)dz\}$, but since $v_{x\theta\theta} > 0$ we must have that $\hat{x}'(\theta) \leq x_\alpha$ which implies

$\forall\theta' \in [\underline{\theta},\theta_1] \cup (\theta_2,\bar{\theta})$ $\int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta')dz \geq \min\{\int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta_2)dz, \int_{\hat{x}'(\theta)}^{x_\alpha} v_x(z,\theta_1)dz\} \geq \hat{T}'(\theta) - T_\alpha$, where the last inequality is given by the implementability of $\hat{x}(\cdot)$. Thus we have the following $\left(\forall\theta' \in [\underline{\theta},\theta_1] \cup (\theta_2,\bar{\theta})\right)(\forall\theta \in (\theta_1,\theta_2))$:

$$v(\hat{x}'(\theta'),\theta') + \hat{T}'(\theta') \underbrace{\geq}_{\text{Using that } \hat{x}(\cdot) \text{ is I.C.}} v(x_\alpha,\theta') + T_\alpha \underbrace{\geq}_{\text{Using that } \int_{x_\alpha}^{\hat{x}'(\theta)} v_x(z,\theta')dz \leq T_\alpha - \hat{T}'(\theta)} v(\hat{x}'(\theta),\theta') + \hat{T}'(\theta)$$

. □

Theorem 9: Let $\hat{x}(\theta)$ be an I.C. mechanism, such that $\hat{x}(\theta)$ is decreasing in some interval $[\theta_\alpha,\theta_\beta]$. For any $(\theta_1 < \theta_2 \leq \theta_3 < \theta_4) \in [\theta_\alpha,\theta_\beta]$ it can never hold true that $\Phi(\theta_4,\theta_2) = 0$ and $\Phi(\theta_3,\theta_1) = 0$.

Proof. We will prove that for given $\theta_1 < \theta_2 \leq \theta_3 < \theta_4$ it cannot hold true that:

$$\Phi^{\hat{x}}(\theta_3,\theta_1) = \Phi^{\hat{x}}(\theta_4,\theta_2) = 0$$

. I will prove it by assuming that there exists $\theta_1 < \theta_2 \leq \theta_3 < \theta_4$ such that $\Phi^{\hat{x}}(\theta_1,\theta_3) = \Phi^{\hat{x}}(\theta_2,\theta_4) = 0$, and showing that the mechanism is not I.C. between 1and 4.

$$\begin{aligned}
\Phi^{\hat{x}}(\theta_4,\theta_1) &= \int_{\theta_1}^{\theta_4}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy \\
&= \int_{\theta_1}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy + \int_{\theta_2}^{\theta_4}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy - \int_{\theta_2}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy \\
&= \int_{\theta_1}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy + \int_{\theta_2}^{\theta_4}\int_{\hat{x}(\theta_2)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy - \int_{\theta_2}^{\theta_3}\int_{\hat{x}(\theta_2)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy \\
&\quad + \int_{\theta_2}^{\theta_4}\int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} v_{x\theta}(z,y)dzdy - \int_{\theta_2}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} v_{x\theta}(z,y)dzdy \\
&= \underbrace{\int_{\theta_1}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy}_{\Phi^{\hat{x}}(\theta_3,\theta_1)=0} + \underbrace{\int_{\theta_2}^{\theta_4}\int_{\hat{x}(\theta_2)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy}_{\Phi^{\hat{x}}(\theta_4,\theta_2)=0} - \underbrace{\int_{\theta_2}^{\theta_3}\int_{\hat{x}(\theta_2)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy}_{=\Phi^{\hat{x}}(\theta_3,\theta_2)\geq 0(\text{By I.C. between 2 and 3})} \\
&\quad + \int_{\theta_3}^{\theta_4}\int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} v_{x\theta}(z,y)dzdy
\end{aligned}$$

So, we have the following:

$$\Phi^{\hat{x}}(\theta_4,\theta_1) \leq \int_{\theta_3}^{\theta_4}\int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} v_{x\theta}(z,y)dzdy = \int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)}\int_{\theta_3}^{\theta_4} v_{x\theta}(z,y)dydz$$

Since we are considering a decreasing policy we have that $\hat{x}(\theta_1) > \hat{x}(\theta_2)$, and using that $v_{x\theta\theta} > 0$ we have that:

$$\Phi^{\hat{x}}(\theta_4,\theta_1) < \int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} \int_{\theta_3}^{\theta_4} v_{x\theta}(z,y)dydz < (\theta_4-\theta_3)\int_{\hat{x}(\theta_1)}^{\hat{x}(\theta_2)} v_{x\theta}(z,\theta_3)dydz = -(\theta_4-\theta_3)\int_{\hat{x}(\theta_2)}^{\hat{x}(\theta_1)} v_{x\theta}(z,\theta_3)dydz$$

But, we know that $\int_{\theta_1}^{\theta_3}\int_{\hat{x}(\theta_1)}^{\hat{x}(y)} v_{x\theta}(z,y)dzdy = 0$, so we must have that $\int_{\hat{x}(\theta_3)}^{\hat{x}(\theta_1)} v_{x\theta}(z,\theta_3)dydz \geq 0$ (with equality if the policy is continuous at θ_3. Now, using that $v_{xx\theta} > 0$ we have that

$$\int_{\hat{x}(\theta_3)}^{\hat{x}(\theta_1)} v_{x\theta}(z,\theta_3)dydz \geq 0 \Rightarrow \int_{\hat{x}(\theta_2)}^{\hat{x}(\theta_1)} v_{x\theta}(z,\theta_3)dydz \geq 0$$

and thus we have that

$$\Phi^{\hat{x}}(\theta_4,\theta_1) < -(\theta_4-\theta_3)\int_{\hat{x}(\theta_2)}^{\hat{x}(\theta_1)} v_{x\theta}(z,\theta_3)dydz < 0$$

So, the mechanism is not I.C. between θ_1 and θ_4. □

Theorem 11: Let $\hat{x}(\theta)$ be the optimal policy.

1. **Let the interval (θ_1,θ_2) be such that $\forall\theta\in(\theta_1,\theta_2)\forall\theta'\in\Theta\quad \Phi^{x*}(\theta,\theta') > 0$, then $\frac{f_x(\hat{x}(\theta),\theta)p(\theta)}{v_{x\theta}}$ must be non-increasing**
2. **Let the interval (θ_1,θ_2) be such that $\forall\theta\in(\theta_1,\theta_2)\forall\theta'\in\Theta\quad \Phi^{x*}(\theta',\theta) > 0$, then $\frac{f_x(\hat{x}(\theta),\theta)p(\theta)}{v_{x\theta}}$ must be non-decreasing**

Proof. We will first prove corollary 13 and then make some considerations to prove Theorem 11.

Let's take a interval (θ_1,θ_2), such that for all θ in this interval the I.C.C. of $x^*(\cdot)$ are not active[16]. That is:

$$\forall\theta\in(\theta_1,\theta_2)\forall\theta\in\Theta\quad \Phi^{x^*}(\theta',\theta) > 0 \quad\wedge\quad \Phi^{x^*}(\theta,\theta') > 0$$

Let's consider a perturbation $h(\cdot)$, such that

- $\forall\theta\in[\underline{\theta},\theta_1]\cup[\theta_2,\bar{\theta}]\ h(\theta) = 0$
- $h(\cdot)$ is continuous in Θ.
- $\int_{\theta_1}^{\theta_2}\int_0^{h(y)} v_{x\theta}(x^*(y)+z,y)dzdy = 0$

Let's define $\tilde{x}(\theta) = x^*(\theta) + \alpha h(\theta)$. We will prove that for a $h(\cdot)$ small enough $\tilde{x}(\cdot)$ is I.C., and thus $h(\cdot)$ is an admissible perturbation

- First note that for a small enough perturbation all I.C.C. concerning types in (θ_1,θ_2) are also fulfilled (including the local I.C.C. in (θ_1,θ_2)). That is:

$$\exists\alpha \text{ such that } \forall\theta\in(\theta_1,\theta_2)\forall\theta\in\Theta\quad \Phi^{\tilde{x}}(\theta',\theta) > 0 \quad\wedge\quad \Phi^{\tilde{x}}(\theta,\theta') > 0$$

- The I.C.C. between types greater than θ_2 or smaller than θ_1 remain unchanged, and thus are also fulfilled. That is:

$$\forall\theta,\theta'\in[\underline{\theta},\theta_1]\quad \Phi^{\tilde{x}}(\theta',\theta) = \Phi^{x^*}(\theta',\theta) \geq 0 \quad\wedge\quad \Phi^{\tilde{x}}(\theta,\theta') = \Phi^{x^*}(\theta,\theta') \geq 0$$

$$\forall\theta,\theta'\in[\theta_2,\bar{\theta}]\quad \Phi^{\tilde{x}}(\theta',\theta) = \Phi^{x^*}(\theta',\theta) \geq 0 \quad\wedge\quad \Phi^{\tilde{x}}(\theta,\theta') = \Phi^{x^*}(\theta,\theta') \geq 0$$

[16]Note that a necessary condition for this is that $x^*(\cdot)$ must be strictly monotonic in (θ_1,θ_2)

- Let's take $\theta > \theta_2$ and $\theta' < \theta_1$, we have that:

$$\begin{aligned}\Phi^{\tilde{x}}(\theta,\theta')&= \int_{\theta'}^{\theta}\int_{\tilde{x}(\theta')}^{\tilde{x}(y)} v_{x\theta}(z,y)dzdy \\ &= \underbrace{\int_{\theta'}^{\theta}\int_{x^*(\theta')}^{x^*(y)} v_{x\theta}(z,y)dzdy}_{=\Phi^{x^*}(\theta,\theta')\geq 0} + \underbrace{\int_{\theta_1}^{\theta_2}\int_{x^*(y)}^{x^*(y)+h(y)} v_{x\theta}(z,y)dzdy}_{=0} \geq 0\end{aligned}$$

$$\begin{aligned}\Phi^{\tilde{x}}(\theta,\theta')&= \int_{\theta}^{\theta'}\int_{\tilde{x}(\theta)}^{\tilde{x}(y)} v_{x\theta}(z,y)dzdy \\ &= \underbrace{\int_{\theta}^{\theta'}\int_{x^*(\theta)}^{x^*(y)} v_{x\theta}(z,y)dzdy}_{=\Phi^{x^*}(\theta',\theta)\geq 0} - \underbrace{\int_{\theta_1}^{\theta_2}\int_{x^*(y)}^{x^*(y)+h(y)} v_{x\theta}(z,y)dzdy}_{=0} \geq 0\end{aligned}$$

Therefore $h(\cdot)$ is an admissible perturbation. So, if $x^*(\cdot)$ is the optimal policy, then we must have that $\int_{\theta_1}^{\theta_2} V(\hat{x}(\theta)+h(\theta),\theta)p(\theta)d\theta$ has it's maximum value at $h(\cdot)=0$. In this case the utility to the principal of policy $\tilde{x}(\cdot)$ is given by:

$$\begin{aligned}U^{\tilde{x}}&= \int_{\underline{\theta}}^{\bar{\theta}} f(\tilde{x}(y),y)p(y)dy \\ &= \int_{\underline{\theta}}^{\bar{\theta}} f(x^*(y),y)p(y)dy + \int_{\theta_1}^{\theta_2} f(\tilde{x}(y),y)p(y)dy - \int_{\theta_1}^{\theta_2} f(x^*(y),y)p(y)dy \\ &= U^{x^*} + \int_{\theta_1}^{\theta_2}\int_{0}^{h(\theta)} f_x(x^*(y)+z,y)p(y)dzdy \\ &= U^{x^*} + \int_{\theta_1}^{\theta_2}\int_{0}^{h(\theta)} \frac{f_x(x^*(y)+z,y)p(y)}{v_{x\theta}(x^*(y)+z,y)} v_{x\theta}(x^*(y)+z,y)dzdy\end{aligned}$$

Optimizing with respect to $h(\cdot)$ we get the following optimization problem:

$$\max_{h(\cdot)} \quad U^{x^*} + \int_{\theta_1}^{\theta_2}\int_{0}^{h(\theta)} f_x(x^*(y)+z,y)p(y)dzdy$$

$$s.t. \int_{\theta_1}^{\theta_2}\int_{0}^{h(y)} v_{x\theta}(x^*(y)+z,y)dzdy = 0$$

this leads to the following pointwise F.O.C.

$$f_x(x^*(y)+h(y),y)p(y) - \lambda v_{x\theta}(x^*(y)+h(y),y) = 0$$

which has it's optimum at $h(y)=0$, and thus we get the following equation:

$$\frac{f_x(x^*(y),y)p(y)}{v_{x\theta}(x^*(y),y)} = \lambda$$

The demonstration of Theorem 11 is basically the same than corollary 13 only that it is necessary to be a more careful with the variation $h(\cdot)$.

Let's consider an interval (θ_1,θ_2) such that $\forall\theta \in (\theta_1,\theta_2)\forall\theta' \in \Theta \quad \Phi^{x^*}(\theta,\theta') > 0$ (condition $\Phi^{x^*}(\theta',\theta) \geq 0$ is guarantee by the I.C.C. of $x^*(\cdot)$).

Let's consider a perturbation $h(\cdot)$, such that

- $\forall\theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, \bar{\theta}]$ $h(\theta) = 0$
- $h(\cdot)$ is continuous in Θ.
- $\int_{\theta_1}^{\theta_2} \int_0^{h(y)} v_{x\theta}(x^*(y) + z, y)dzdy = 0$
- $\forall\theta \in [\theta_1, \theta_2]$ $\quad \int_{\theta_1}^{\theta} \int_0^{h(y)} v_{x\theta}(x^*(y) + z, y)dzdy \leq 0$

Let's define $\tilde{x}(\theta) = x^*(\theta) + \alpha h(\theta)$. We need to prove that for a $h(\cdot)$ small enough $\tilde{x}(\cdot)$ is I.C., and thus $h(\cdot)$ is an admissible perturbation.

- It is easy to see that the cases in which $\theta', \theta \in [\theta_1, \theta_2]$ and $\theta', \theta \in \Theta\backslash[\theta_1, \theta_2]$ can be proved the same way as corollary 13.
- If $\Big(\theta \in (\theta_1, \theta_2)\Big)\Big(\theta' \in \Theta\Big)$ then for a $h(\cdot)$ small enough $\Phi^{x^*}(\theta, \theta') > 0$, which can also be proven the same way as as corollary 13.
- Thus we need to prove that $\Big(\forall\theta' \in (\theta_1, \theta_2)\Big)\Big(\forall\theta \in \Theta\Big)$ exists a $h(\cdot)$ small enough such that $\Phi^{x^*}(\theta, \theta')$.
  If $\theta' \in (\theta_1, \theta_2)\forall\theta \in [\underline{\theta}, \theta_1]$:

$$\begin{aligned}\Phi^{\tilde{x}}(\theta, \theta') &= \int_{\theta'}^{\theta} \int_{\tilde{x}(\theta')}^{\tilde{x}(y)} v_{x\theta}(z, y)dzdy \\ &= \underbrace{\int_{\theta'}^{\theta} \int_{x^*(\theta')}^{x^*(y)} v_{x\theta}(z, y)dzdy}_{=\Phi^{x^*}(\theta', \theta) \geq 0} - \underbrace{\int_{\theta_1}^{\theta'} \int_{x^*(y)}^{x^*(y)+h(y)} v_{x\theta}(z, y)dzdy}_{\leq 0} \geq 0\end{aligned}$$

If $\theta' \in (\theta_1, \theta_2)\forall\theta \in [\theta_2, \bar{\theta}]$:

$$\begin{aligned}\Phi^{\tilde{x}}(\theta, \theta') &= \int_{\theta'}^{\theta} \int_{\tilde{x}(\theta')}^{\tilde{x}(y)} v_{x\theta}(z, y)dzdy \\ &= \underbrace{\int_{\theta'}^{\theta} \int_{x^*(\theta')}^{x^*(y)} v_{x\theta}(z, y)dzdy}_{=\Phi^{x^*}(\theta', \theta) \geq 0} + \underbrace{\int_{\theta_2}^{\theta'} \int_{x^*(y)}^{x^*(y)+h(y)} v_{x\theta}(z, y)dzdy}_{\geq 0} \geq 0\end{aligned}$$

Thus, $h(\cdot)$ is a admissible perturbation. Optimizing with respect to $h(\cdot)$ we get the following optimization problem:

$$\max_{h(\cdot)} \quad U^{x^*} + \int_{\theta_1}^{\theta_2} \int_0^{h(\theta)} f_x(x^*(y) + z, y)p(y)dzdy$$

$$s.t. \begin{array}{c} \int_{\theta_1}^{\theta_2} \int_0^{h(y)} v_{x\theta}(x^*(y) + z, y)dzdy = 0 \\ \forall\theta \in [\theta_1, \theta_2] \quad \int_{\theta_1}^{\theta} \int_0^{h(y)} v_{x\theta}(x^*(y) + z, y)dzdy \leq 0 \end{array}$$

this leads to the following lagrangian:

$$L = f_x(x^*(y) + h(y), y)p(y) + \lambda v_{x\theta}(x^*(y) + h(y), y) + \int_{\theta_1}^{y} \nu(w)dw v_{x\theta}(x^*(y) + h(y), y)$$

with $\nu(w) \geq 0 \forall w \in [\theta_1, \theta_2]$and the pointwise F.O.C. leads to:

$$f_x(x^*(y) + h(y), y)p(y) + \left(\lambda + \int_{\theta_1}^{y} \nu(w)dw\right) v_{x\theta}(x^*(y) + h(y), y) = 0$$

which has it's optimum at $h(y)=0$, and thus we get the following equation:

$$\frac{f_x(x^*(y),y)p(y)}{v_{x\theta}(x^*(y),y)}=-\lambda-\int_{\theta_1}^{y}\nu(w)dw$$

since $\int_{\theta_1}^{y}\nu(w)dw$ is not decreasing we get the result.

The second case can be done the same way. □

Theorem 15: Let $x^*(\cdot)$ be the optimal implementable convex-valued correspondence, and $\theta'',\theta'\in\Theta$ be such that $\Phi^{x^*}(\theta'',\theta')=0$. If $x^*(\cdot)$ is strictly monotonic and continuous at θ' and θ'' , then:

$$\frac{f_x(x^*(\theta''),\theta'')p(\theta'')}{v_{x\theta}(x^*(\theta''),\theta'')}=\frac{f_x(x^*(\theta'),\theta')p(\theta')}{v_{x\theta}(x^*(\theta'),\theta')}$$

Proof. Since the case $x^*(\theta'')=x^*(\theta')$ can be found in Araujo and Moreira [1] we will only analyze the case in which $x^*(\theta'')\neq x^*(\theta')$. It is easy to see from the analysis made in section 6 that if $x^*(\cdot)$ is an implementable convex-valued correspondence the only way we can have $x^*(\theta'')\neq x(\theta')$ and $\Phi^{x^*}(\theta'',\theta')=0$ is if $\theta''>\theta'$ and $x^*(\theta')>x(\theta'')$. Moreover, using the analysis made in section 6 and using theorem 9 to know

$$(\forall\tilde{\theta}\in\Theta\backslash[\theta',\theta''])(\forall\theta\in[\theta',\theta''])\quad\Phi^{x^*}(\tilde{\theta},\theta)>0\quad\wedge\quad\Phi^{x^*}(\theta,\tilde{\theta})>0$$

In what follows we will assume we can find intervals (θ_1,θ_2) and (θ_3,θ_4) such that:

- $\theta''\in(\theta_3,\theta_4)$ and $\theta'\in(\theta_1,\theta_2)$
- $x^*(\cdot)$ is strictly monotonic and continuous in (θ_1,θ_2) and (θ_3,θ_4).
- For all $\theta''\in(\theta_3,\theta_4)$ there exists a θ' such that $\Phi^{x^*}(\theta'',\theta')=0$, and for all $\theta'\in(\theta_1,\theta_2)$ there exists a θ'' such that $\Phi^{x^*}(\theta'',\theta')=0$
- $\Phi^{x^*}(\theta_1,\theta_4)=\Phi^{x^*}(\theta_2,\theta_3)=0$

We can always find intervals (θ_1,θ_2) and (θ_3,θ_4) that satisfy the first two conditions. If we can't satisfy the third condition, then we could prove the theorem with a variation similar to the one made in theorem 11. If the third condition can be satisfied then the fourth condition can also be satisfied, and then we can there is a unique bijection between $\theta'\in(\theta_1,\theta_2)$ and $\theta''\in(\theta_3,\theta_4)$ where each pair is defined by $\Phi(\theta'',\theta')=0$. For each pair the following conditions must be true:

$$v_x(x(\theta'),\theta')=v_x(x(\theta'),\theta'') \tag{2}$$

$$v_\theta(x(\theta'),\theta'')=v_\theta(x(\theta'),\theta'') \tag{3}$$

Where conditions 2 and 3 are given by lemma 10. Now we will make a variation around $x^*(\cdot)$, so $\tilde{x}(\theta)=x^*(\theta)+h(\theta)$. For $\tilde{x}(\theta)$ to be I.C. we will define $h(\theta)$ continuous and arbitrary in (θ_3,θ_4) such that $h(\theta_3)=h(\theta_4)=0$. $\tilde{x}(\theta)$ in (θ_1,θ_2) will be obtained from condition 2 and 3, and $\forall\theta\notin(\theta_1,\theta_2)\cup(\theta_3,\theta_4)\quad h(\theta)=0$. First we need to show $\tilde{x}(\cdot)$ is I.C.:

- $\forall\theta',\theta''\in\Theta\backslash[\theta_1,\theta_4]\quad\Phi^{x^*}(\theta'',\theta')=\Phi^{\tilde{x}}(\theta'',\theta')$
- $\forall\theta',\theta''\in[\theta_2,\theta_3]\quad\Phi^{x^*}(\theta'',\theta')=\Phi^{\tilde{x}}(\theta'',\theta')$
- Using using the analysis made in section 6 and using theorem 9:

$$(\forall\theta'\in(\theta_1,\theta_2)\cup(\theta_3,\theta_4))\,(\forall\theta''\notin(\theta_1,\theta_2)\cup(\theta_3,\theta_4))\quad\Phi^{x^*}(\tilde{\theta},\theta)>0\quad\wedge\quad\Phi^{x^*}(\theta,\tilde{\theta})>0$$

thus, for a $h(\cdot)$ small enough we have that the conditions are also satisfied for $\tilde{x}(\cdot)$

So, with condition 2 and 3 we can define the function $\varphi(x,\theta)$ and $\psi(x,\theta)$ by:

$$v_\theta(x,\theta) = v_\theta(\psi(x,\theta),\theta)$$

$$v_x(\psi(x,\theta),\theta) = v_x(\psi(x,\theta),\varphi(x,\theta))$$

So we have the following:

$$\begin{aligned}
G(h) =& \int_{\theta_3}^{\theta_4} f_x(\tilde{x}(\theta),\theta)p(\theta)d\theta + \int_{\varphi(\tilde{x}(\theta_2),\theta_2)}^{\varphi(\tilde{x}(\theta_1),\theta_1)} f_x(\tilde{x}(\theta),\theta)p(\theta)d\theta \\
=& \int_{\theta_3}^{\theta_4} f_x(\tilde{x}(\theta),\theta)p(\theta)d\theta - \int_{\theta_3}^{\theta_4} f_x(\psi(\tilde{x}(\theta),\theta),\varphi(\tilde{x}(\theta),\theta))p(\varphi(\tilde{x}(\theta),\theta))(\varphi_x(\tilde{x}(\theta),\theta)\tilde{x}_\theta(\theta) + \varphi_\theta(\tilde{x}(\theta),\theta))d\theta \\
=& \int_{\theta_3}^{\theta_4} w(\tilde{x}(\theta),\theta) - w(\psi(\tilde{x}(\theta),\theta),\varphi(\tilde{x}(\theta),\theta))(\varphi_x(\tilde{x}(\theta),\theta)\tilde{x}_\theta(\theta) + \varphi_\theta(\tilde{x}(\theta),\theta))d\theta
\end{aligned}$$

where $f_x(\tilde{x}(\theta),\theta)p(\theta) = w(\tilde{x}(\theta),\theta)$. Now, taking the variation in the direction of h:

$$\begin{aligned}
\delta_h G(h) =& \int_{\theta_3}^{\theta_4} \Big[w_x(x^*(\theta),\theta)h - \hat{w}\left(\varphi_{xx}(x^*(\theta),\theta)x^*_\theta(\theta)h + \varphi_{\theta x}(x^*(\theta),\theta)h + \varphi_x(x^*(\theta),\theta)h_\theta\right) \\
& -h\left(\hat{w}_x\psi_x(x^*(\theta),\theta) + \hat{w}_\theta\varphi_x(x^*(\theta),\theta)\right)\left(\varphi_x(x^*(\theta),\theta)x^*_\theta(\theta) + \varphi_\theta(x^*(\theta),\theta)\right)\Big] d\theta
\end{aligned}$$

Where $\hat{w}$ means that it is evaluated in $(\psi(x^*(\theta),\theta),\varphi(x^*(\theta),\theta))$. Integratating the term with h_θ by parts we get:

$$\begin{aligned}
\delta_h G(h) =& \int_{\theta_3}^{\theta_4} \Big[w_x(x^*(\theta),\theta)h - h\hat{w}\left(\varphi_{xx}(x^*(\theta),\theta)x^*_\theta(\theta) + \varphi_{\theta x}(x^*(\theta),\theta)\right) + h\hat{w}\left(\varphi_{xx}(x^*(\theta),\theta)x^*_\theta(\theta) + \varphi_{\theta x}(x^*(\theta),\theta)\right) \\
& +h\varphi_x(x^*(\theta),\theta)\left(\hat{w}_x(\psi_x(x^*(\theta),\theta)x^*_\theta(\theta) + \psi_\theta(x^*(\theta),\theta)) + \hat{w}_\theta\left((\varphi_x(x^*(\theta),\theta)x^*_\theta(\theta) + \varphi_\theta(x^*(\theta),\theta))\right)\right) \\
& -h\left(\hat{w}_x\psi_x(x^*(\theta),\theta) + \hat{w}_\theta\varphi_x(x^*(\theta),\theta)\right)\left(\varphi_x(x^*(\theta),\theta)x^*_\theta(\theta) + \varphi_\theta(x^*(\theta),\theta)\right)\Big] d\theta
\end{aligned}$$

Simplifying terms...

$$\begin{aligned}
\delta_h G(h) =& \int_{\theta_3}^{\theta_4} \Big[w_x(x^*(\theta),\theta)h \\
& +h\hat{w}_x\left(\varphi_x(x^*(\theta),\theta)\psi_\theta(x^*(\theta),\theta)) - \varphi_\theta(x^*(\theta),\theta)\psi_x(x^*(\theta),\theta))\right)\Big] d\theta
\end{aligned}$$

Since $x^*(\cdot)$ is optimal we must have that $\delta_h G(h) = 0$, and since $h(\theta)$ is arbitrary we have that

$$w_x(x^*(\theta),\theta) + \hat{w}_x\left(\varphi_x(x^*(\theta),\theta)\psi_\theta(x^*(\theta),\theta) - \varphi_\theta(x^*(\theta),\theta)\psi_x(x^*(\theta),\theta)\right) = 0$$

Finally we know that $\varphi(x,\theta)$ and $\psi(x,\theta)$ are defined by:

$$v_\theta(x,\theta) = v_\theta(\psi(x,\theta),\theta)$$

$$v_x(\psi(x,\theta),\varphi(x,\theta)) = v_x(\psi(x,\theta),\theta)$$

So, we have that:

$$\psi_\theta(x,\theta) = \frac{v_{\theta\theta}(x,\theta) - v_{\theta\theta}(\psi(x,\theta),\theta)}{v_{x\theta}(\psi(x,\theta),\theta)}$$

$$\psi_x(x,\theta) = \frac{v_{x\theta}(x,\theta)}{v_{x\theta}(\psi(x,\theta),\theta)}$$

$$\varphi_\theta(x,\theta) = \frac{v_{x\theta}(\psi(x,\theta),\theta) + \psi_\theta(x,\theta)\left(v_{xx}(\psi(x,\theta),\theta) - v_{xx}(\psi(x,\theta),\varphi(x,\theta))\right)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}$$

$$\varphi_x(x,\theta) = \frac{\psi_x(x,\theta)\left(v_{xx}(\psi(x,\theta),\theta) - v_{xx}(\psi(x,\theta),\varphi(x,\theta))\right)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}$$

So, finally we have that:

$$\begin{aligned}\varphi_x(x^*(\theta),\theta)\psi_\theta(x^*(\theta),\theta)) - \varphi_\theta(x^*(\theta),\theta)\psi_x(x^*(\theta),\theta)) = {} & \frac{\psi_x(x,\theta)\left(v_{xx}(\psi(x,\theta),\theta) - v_{xx}(\psi(x,\theta),\varphi(x,\theta))\right)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}\psi_\theta(x,\theta) \\ & - \frac{v_{x\theta}(\psi(x,\theta),\theta) + \psi_\theta(x,\theta)\left(v_{xx}(\psi(x,\theta),\theta) - v_{xx}(\psi(x,\theta),\varphi(x,\theta))\right)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}\psi_x(x,\theta)\end{aligned}$$

$$\Rightarrow \varphi_x(x^*(\theta),\theta)\psi_\theta(x^*(\theta),\theta)) - \varphi_\theta(x^*(\theta),\theta)\psi_x(x^*(\theta),\theta)) = -\frac{v_{x\theta}(\psi(x,\theta),\theta)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}\psi_x(x,\theta) = -\frac{v_{x\theta}(x,\theta)}{v_{x\theta}(\psi(x,\theta),\varphi(x,\theta))}$$

therefore,

$$w_x(x^*(\theta),\theta) - w_x(\psi(x^*(\theta),\theta),\varphi(x^*(\theta),\theta))\frac{v_{x\theta}(x^*(\theta),\theta)}{v_{x\theta}(\psi(x^*(\theta),\theta),\varphi(x^*(\theta),\theta))} = 0$$

$(\forall\theta'' \in (\theta_3,\theta_4))(\forall\theta' \in (\theta_1,\theta_2))$ such that $\Phi^{x^*}(\theta'',\theta') = 0$ we have that $\varphi(x^*(\theta''),\theta'') = \theta'$ and $\psi(x^*(\theta''),\theta'') = x^*(\theta')$. Thus we get the result. $\square$

Lemma 26: The optimal decreasing policy consists of bunching part in an interval $[\dot{\theta},\hat{\theta}]$ and a continuous and strictly decreasing part in $[\underline{\theta},\dot{\theta}]$, where $\dot{\theta} \in (\underline{\theta},\hat{\theta}]$. In what follows we will refer to $\dot{\theta}$ as the point where the policy starts being strictly decreasing, and $\dot{\theta}^x$ as the point $\dot{\theta}$ corresponding to policy $x(\cdot)$.

Proof. Let's first consider the case of a I.C. mechanism $x(\cdot)$ with discontinuous jump in $\theta' < \hat{\theta}$. We will use θ'^{+} and θ'^{-} to denote that a policy is being evaluated at the right and left limit of θ' respectively. Let's consider a variation of the form $h(\cdot)$ such that:

- $h(\cdot)$ is continuous in $[\theta'-\varepsilon,\theta') \cup (\theta',\theta'+\varepsilon]$
- $h'(\cdot) < 0$ This assumption is just to ensure the mechanism $x(\theta) + h(\theta)$ is I.C. if there is a bunching zone near θ'
- $h(\theta) = \begin{cases} < 0 & \forall\theta \in (\theta'-\varepsilon,\theta') \\ > 0 & \forall\theta \in (\theta',\theta'+\varepsilon) \\ = 0 & \forall\theta \notin (\theta'-\varepsilon,\theta'+\varepsilon) \end{cases}$
- $\int_{\theta'-\varepsilon}^{\theta'+\varepsilon}\int_{x(\tilde{\theta})}^{x(\tilde{\theta})+h(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})dzd\tilde{\theta} = 0$
- $x(\theta^-) + h(\theta^-) \geq x(\theta^+) + h(\theta^+)$

We will denote $\rho(\cdot) = x(\cdot) + h(\cdot)$. Note that the I.C.C. between pairs of θ not in $(\theta'-\varepsilon,\theta'+\varepsilon)$ remain unchanged by variation $h(\theta)$, and thus it is only necessary to check that $h(\theta)$ improves the principals utility and the I.C.C. with $\theta \in (\theta'-\varepsilon,\theta'+\varepsilon)$ are not broken. The change in the principals utility is given by:

$$
\begin{aligned}
\Delta U = & \int_{\theta'-\varepsilon}^{\theta'+\varepsilon} f(\rho(\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})d\tilde{\theta} - \int_{\theta'-\varepsilon}^{\theta'+\varepsilon} f(x(\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})d\tilde{\theta} \\
= & \int_{\theta'-\varepsilon}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} f_x(z,\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})dzd\tilde{\theta} \\
= & \int_{\theta'-\varepsilon}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\frac{f_x(z,\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})}{v_{x\theta}(z,\tilde{\theta})}dzd\tilde{\theta} \\
= & \int_{\theta'-\varepsilon}^{\theta'} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\frac{f_x(z,\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})}{v_{x\theta}(z,\tilde{\theta})}dzd\tilde{\theta} + \int_{\theta'}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\frac{f_x(z,\tilde{\theta}),\tilde{\theta})p(\tilde{\theta})}{v_{x\theta}(z,\tilde{\theta})}dzd\tilde{\theta} \\
= & \int_{\theta'-\varepsilon}^{\theta'} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\Gamma(z,\tilde{\theta})dzd\tilde{\theta} + \int_{\theta'}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\Gamma(z,\tilde{\theta})dzd\tilde{\theta}
\end{aligned}
$$

Using that the variation $h(\theta)$ satisfies:

$$
\int_{\theta'-\varepsilon}^{\theta'} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})dzd\tilde{\theta} + \int_{\theta'}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})dzd\tilde{\theta} = 0
$$

Note that the first term is positive and the second one is negative. Using that $\frac{\partial \Gamma(x,\theta)}{\partial x} > 0$ we know that for a $h(\theta)$ small enough we have that

$$
\left(\forall \tilde{\theta} \in (\theta'-\varepsilon,\theta')\right)\left(\forall z \in (\rho(\tilde{\theta}),x(\tilde{\theta}))\right)\left(\forall \tilde{\theta}' \in (\theta',\theta'+\varepsilon)\right)\left(\forall z' \in (x(\tilde{\theta}),\rho(\tilde{\theta}))\right) \quad \Gamma(\tilde{\theta},z) > \Gamma(\tilde{\theta}',z')
$$

and thus,

$$
\Delta U = \int_{\theta'-\varepsilon}^{\theta'} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\Gamma(z,\tilde{\theta})dzd\tilde{\theta} + \int_{\theta'}^{\theta'+\varepsilon} \int_{x(\tilde{\theta})}^{\rho(\tilde{\theta})} v_{x\theta}(z,\tilde{\theta})\Gamma(z,\tilde{\theta})dzd\tilde{\theta} > 0
$$

thus, it is a profitable variation.

Since we have shown that $h(\cdot)$ is a profitable variation we now need to show that it is I.C., for this it is necessary to check the following:

$$
\left(\forall \theta \in (\theta'-\varepsilon,\theta'+\varepsilon)\right)\left(\forall \tilde{\theta} \in \Theta\right) \quad \Phi^{\rho}(\theta,\tilde{\theta}) \geq 0 \quad \wedge \quad \Phi^{\rho}(\tilde{\theta},\theta) \geq 0
$$

First note that if $x(\cdot)$ is I.C. then $\int_{x(\theta'^{+})}^{x(\theta'^{-})} v_{x\theta}(z,\theta')dz < 0$, thus the local I.C.C. guarantee that the I.C.C. are not broken for all $\tilde{\theta} \in (\theta'-\varepsilon,\theta'+\varepsilon)$.

We also know that $\forall \tilde{\theta} \in [\underline{\theta},\theta'-\varepsilon] \quad \Phi^{\rho}(\theta,\tilde{\theta}) \geq \Phi^{x}(\theta,\tilde{\theta}) \geq 0$, and for $\tilde{\theta} \in (\theta'+\varepsilon,\bar{\theta})$ we know that the local I.C.C. guarantee that $\Phi^{\rho}(\theta,\tilde{\theta}) \geq 0$. So, if $x(\cdot)$ is I.C. then the condition $\Phi^{\rho}(\theta,\tilde{\theta}) \geq 0$ is satisfied for all $\tilde{\theta} \in \Theta$

Thus we need to check $\Phi^{\rho}(\tilde{\theta},\theta) \geq 0$, which only needs to be checked for $\tilde{\theta} < \theta$ (likewise for $\tilde{\theta} > \theta$ the local I.C.C. guarantees the global I.C.C.). The proof proceeds as follows:

- Check that $\Phi^{\rho}(\theta'^{+},\tilde{\theta}) \geq 0 \quad \wedge \quad \Phi^{\rho}(\theta'^{-},\tilde{\theta}) \geq 0 \Rightarrow \rho(\cdot)$ is I.C. for all points between θ'^{+} and θ'^{-}
- Check that if $\Phi^{x}(\tilde{\theta},\theta'^{-}) = 0$, then for a ε small enough for all θ in $(\theta'-\varepsilon,\theta']$ $\quad \Phi^{\rho}(\tilde{\theta},\theta) = 0$ (if $\Phi^{\rho}(\theta'^{-},\tilde{\theta}) > 0$ then the implication is trivial)

- Check that if $\Phi^x(\tilde{\theta}, \theta'^+) = 0$, then for a ε small enough for all θ in $[\theta'^-, \theta' + \varepsilon)$ $\quad \Phi^\rho(\tilde{\theta}, \theta) = 0$ (if $\Phi^\rho(\theta'^+, \tilde{\theta}) > 0$ then the implication is trivial)

To check that $\Phi^\rho(\tilde{\theta}, \theta'^+) \geq 0 \quad \wedge \quad \Phi^\rho(\tilde{\theta}, \theta'^-) \geq 0 \Rightarrow$ The mechanism is I.C. for all points between θ'^+ and θ'^- note that the I.C.C. of any point between θ'^+ and θ'^- can be written as follows:

$$\Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy$$

with $\mu \in [0, 1]$, in particular the cases $\mu = 1$ and $\mu = 0$ are $\Phi^\rho(\tilde{\theta}, \theta'^-)$ and $\Phi^\rho(\tilde{\theta}, \theta'^+)$ respectively. So we will show that

$$\Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy \geq 0 \quad \wedge \quad \Phi^\rho(\tilde{\theta}, \theta'^+) \geq 0 \Rightarrow$$
$$\Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy \geq 0 \quad \forall \mu \in (0, 1)$$

Too proove this let"s suppose otherwise, that is to say for some $\mu \in (0, 1)$:

$$\Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy < 0$$

but, if $\Phi^\rho(\tilde{\theta}, \theta'^+) \geq 0$ then,

$$\int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy < 0$$

but, using that $v_{xx\theta}(z, y) > 0$ we can see that

$$\int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy < 0 \Rightarrow \int_{\theta'}^{\tilde{\theta}} v_{x\theta}((1-\mu)\rho(\theta'^+) + \mu\rho(\theta'^-), y) dy > 0$$
$$\Rightarrow \int_{\theta'}^{\tilde{\theta}} \int_{\rho(\theta'^-)}^{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)} v_{x\theta}(z, y) dz dy < 0$$

$$\Rightarrow \Phi^\rho(\tilde{\theta}, \theta'^-) = \Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy$$
$$= \Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy + \int_{\theta'}^{\tilde{\theta}} \int_{\rho(\theta'^-)}^{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)} v_{x\theta}(z, y) dz dy < 0$$

Thus, we arrive to a contradiction

We will now Check that if $\Phi^x(\tilde{\theta}, \theta'^-) = 0$ the variation is I.C. in $(\theta' - \varepsilon, \theta')$. First note that $\Phi^x(\tilde{\theta}, \theta'^-) = 0$ implies that $x(\cdot)$ is bunching in some interval (θ_1, θ'). To check this notice let's assume otherwise.

$$\text{If } v_x(x(\theta'^-), \theta') < v_x(x(\theta'^-), \tilde{\theta}) \Rightarrow \exists \theta < \theta' \text{ such that } \Phi^x(\tilde{\theta}, \theta) < 0$$

$$\text{If } v_x(x(\theta'^-), \theta') \geq v_x(x(\theta'^-), \tilde{\theta}) \Rightarrow \forall \mu \in [0, 1) \quad \Phi^\rho(\tilde{\theta}, \theta'^+) + \int_{\theta'}^{\tilde{\theta}} \int_{(1-\mu)\rho(\theta'^+)+\mu\rho(\theta'^-)}^{\rho(\theta'^+)} v_{x\theta}(z, y) dz dy < 0$$

Thus, the mechanism is not I.C. So, if $\Phi^x(\theta'^-,\tilde{\theta})=0$ implies that $x(\cdot)$ is bunching in some interval (θ_1,θ'). Now we will check that if $\Phi^x(\theta'^-,\tilde{\theta})=0$ for a ε small enough the variation is I.C. for a $\theta\in(\theta'-\varepsilon,\theta']$.

$$
\begin{aligned}
\Phi^{x+h}(\tilde{\theta},\theta) =\ & \Phi^x(\tilde{\theta},\theta'^-)+\int_{\theta'}^{\tilde{\theta}}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy-\int_{\theta}^{\theta'}\int_{x(\theta^-)+h(y)}^{x(\theta^-)+h(\theta)}v_{x\theta}(z,y)dzdy\\
&-\int_{\theta'}^{\theta'+\varepsilon}\int_{x(\theta'^+)+h(y)}^{x(\theta'^+)}v_{x\theta}(z,y)dzdy\\
=\ & \Phi^x(\tilde{\theta},\theta'^-)+\int_{\theta'}^{\tilde{\theta}}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy-\int_{\theta}^{\theta'}\int_{x(\theta^-)+h(y)}^{x(\theta^-)+h(\theta)}v_{x\theta}(z,y)dzdy\\
&-\int_{\theta'}^{\theta'+\varepsilon}\int_{x(\theta'^+)+h(y)}^{x(\theta'^+)}v_{x\theta}(z,y)dzdy\pm\int_{\theta}^{\theta'}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy\pm\int_{\theta-\varepsilon}^{\theta}\int_{x(\theta^-)+h(y)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy\\
=\ & \Phi^x(\tilde{\theta},\theta'^-)+\int_{\theta'}^{\tilde{\theta}}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy+\int_{\theta}^{\theta'}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy\\
&\underbrace{-\int_{\theta-\varepsilon}^{\theta'}\int_{x(\theta^-)+h(y)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy-\int_{\theta'}^{\theta'+\varepsilon}\int_{x(\theta'^+)+h(y)}^{x(\theta'^+)}v_{x\theta}(z,y)dzdy}_{=0}+\int_{\theta-\varepsilon}^{\theta}\int_{x(\theta^-)+h(y)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy\\
=\ & \Phi^x(\tilde{\theta},\theta'^-)+\int_{\theta}^{\tilde{\theta}}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy+\int_{\theta-\varepsilon}^{\theta}\int_{x(\theta^-)+h(y)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy
\end{aligned}
$$

But, since $v_x(x(\theta'^-),\theta')<v_x(x(\theta'^-),\tilde{\theta})$ for a ε and $h(\cdot)$ small enough we have that:

$$\Big(\forall y\in(\theta'-\varepsilon,\theta')\Big)\Big(\forall z\in[0,h(y))]\Big)\quad v_x(x(\theta'^-)+z,y)<v_x(x(\theta'^-)+z,\tilde{\theta})\Rightarrow$$

$$\int_{\theta}^{\tilde{\theta}}\int_{x(\theta^-)+h(\theta)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy+\int_{\theta-\varepsilon}^{\theta}\int_{x(\theta^-)+h(y)}^{x(\theta^-)}v_{x\theta}(z,y)dzdy>0\Rightarrow$$

$$\Phi^{x+h}(\tilde{\theta},\theta)>\Phi^x(\tilde{\theta},\theta'^-)=0$$

The case $\Phi^x(\tilde{\theta},\theta'^+)=0$ can be proved in a similar way.

The bunching case can be proved in the same way as the discontinuity, it is easy to note that a bunching zone is a discontinuity in case the axis are exchanged. The optimizing problem an be thought as finding the optimal type for each assignment, changing the problem into finding the optimal function $\theta(\xi)$. If there is a strictly decreasing part at the edges of a bunching zones, then the same variations done for the discontinuities can be done in the bunching zones, and the problem is completely analog. Nevertheless, if the bunching zone ends at $\underline{\theta}$ the the problems are no longer analog, and as a matter of fact these type of bunching zones may be optimal. □

Lemma 32: The following observation can be made of the solution of $Q^{x[\kappa,\theta']}$.

- **The solution of $Q^{x[\kappa,\theta]}$ is uniquely defined for almost for every θ.**
- **Let $\{\theta_{11}^{x[\kappa,\theta']},...,\theta_{2n'}^{x[\kappa,\theta']}\}$ and $\{\theta_{11}^{x[\kappa,\theta'']},...,\theta_{2n''}^{x[\kappa,\theta'']}\}$ be solutions of problems $Q^{x[\kappa,\theta']}$ and $Q^{x[\kappa,\theta'']}$ respectively. If $\theta'<\theta''$ then $\{\theta_{11}^{x[\kappa,\theta']},...,\theta_{2n'}^{x[\kappa,\theta']}\}\backslash\{\underline{\theta},\bar{\theta}\}\subset\bigcup_{l=1}^{n''}(\theta_{1l}^{x[\kappa,\theta'']},\theta_{2l}^{x[\kappa,\theta'']})$. $\Big($In this case we will say $\{\theta_{11}^{x[\kappa,\theta']},...,\theta_{2n'}^{x[\kappa,\theta']}\}$ and $\{\theta_{11}^{x[\kappa,\theta'']},...,\theta_{2n''}^{x[\kappa,\theta'']}\}$ are nested, and we will denote $\{\theta_{11}^{x[\kappa,\theta']},...,\theta_{2n'}^{x[\kappa,\theta']}\}\prec\{\theta_{11}^{x[\kappa,\theta'']},...,\theta_{2n''}^{x[\kappa,\theta'']}\}$ $\Big)$**

Proof. To avoid excess notation we will denote $x[\kappa,\theta]$ by $y[\theta]$. Let's denote the set of all possible solutions of $Q^{y[\theta']}$ by $S(\theta')$. The proof of lemma 32 will consist of 3 steps.

1. We will show that $S(\cdot)$ is an upper hemicontinuous correspondence.
2. For any given θ' such that $S(\theta')$ is not a singleton we can find $\underline{s},\bar{s} \in S(\theta')$ such that for all $s \in S(\theta')\backslash\{\underline{s},\bar{s}\}$ $\quad \underline{s} \prec s \prec \bar{s}$
3. $S(\cdot)$ is a singleton in a neighborhood around θ', and using a comparative static analysis we will show they are nested.

Note that for all $\theta' < \theta_0$ we can find an ε such that for all $\theta \in [\underline{\theta},\hat{\theta}]$, $y[\theta'+\eta](\theta) < x_0(\theta)$. Thus for all $\theta' < \theta_0$ we can find a bound $\bar{n}$ such that for all θ' $n^{y[\theta']} < \bar{n}$ in some neighborhood around θ'. We will define problem $C(\theta')$ as follows:

$$C(\theta') = \min_{\underline{\theta}\leq\theta^{\theta'}_{11}\leq\theta^{\theta'}_{21}\leq\ldots\leq\theta^{\theta'}_{1\bar{n}}\leq\theta^{\theta'}_{2\bar{n}}\leq\bar{\theta}} \sum_{j=1}^{\bar{n}} \Phi^{y[\theta']}(\theta_{2j},\theta_{1j})$$

It is easy to see that a solution to $C(\theta')$ will consist of all the elements of a solution of problem $Q^{y[\theta']}$ plus elements of the form $\theta^{\theta'}_{1k} = \theta^{\theta'}_{2k}$ for some $k \in \{1,..,\bar{n}\}$. Using Berge's Maximum Theorem we can see that the solutions to problem $C(\theta')$ are upper hemicontinuous in θ'. But, we can take a subset of solutions of $C(\theta')$ in which for all elements such that $\theta^{\theta'}_{1k} = \theta^{\theta'}_{2k}$ we have that $\theta^{\theta'}_{1k} = \theta^{\theta'}_{2k} = \bar{\theta}$, and this subset must also be upper hemicontinuous. Thus, the solutions to problem $Q^{y[\theta']}$ are upper hemicontinuous in θ'. The formal definition of $C(\theta')$ that allows to easily see that satisfies the hypothesis of Berge's Maximum Theorem goes as follows:

Let's define the set $H \subset R^{2n}$, such that $\forall\vec{h} = (h_{11},h_{12},..,h_{1n},h_{2n}) \in H \quad \underline{\theta} \leq h_1 \leq h_2 \leq \ldots \leq h_m \leq \bar{\theta}$, and let's define function $\Psi : H \times [\underline{\theta},\theta_0) \to R$ as follows:

$$\Psi(\vec{h},\theta) = \sum_{i=1}^{n} \Phi^{y[\theta]}(h_{2i},h_{1i})$$

we can see by the maximum theorem that $C^*(\theta) = argmax\{\Psi(\vec{h},\theta)|\vec{h} \in H\}$ is a upper hemicontinuous correspondence in θ.

Now we will show that for any θ' such that $S(\theta')$ is not a singleton, there exists $\underline{s},\bar{s}$. First we will show that for any pair of types $\theta_{1k},\theta_{2k},\varphi_{1j},\varphi_{2j}$ that belong to a solution of $S(\theta')$, then it can never hold true that $\theta_{1k} < \varphi_{1j} \leq \theta_{2k} < \varphi_{2j}$. We will prove this by contradiction

- Let's consider a solution $\{\theta_{11},\theta_{21},...,\theta_{1n},\theta_{2n}\} \in S(\theta')$, and $\varphi_1,\varphi_2 \in s \in S(\theta')$, such that $\theta_{1k} < \varphi_1 \leq \theta_{2k} < \varphi_2$
- There are two subcases that need to be consider, for some $l > k \quad \theta_{2(l-1)} \leq \varphi_2 < \theta_{1l}$ or $\theta_{1l} \leq \varphi_2 < \theta_{2l}$
  - In the former case it can be shown that $\Phi(\varphi_2,\theta_{1k}) < \sum_{j=k}^{l-1}\Phi(\theta_{2j},\theta_1 j)$, thus arrive to a contradiction.
  - In the latter case it can be shown that $\Phi(\theta_{2l},\theta_{1k}) < \sum_{j=k}^{l}\Phi(\theta_{2j},\theta_1 j)$, thus arrive to a contradiction.

We will show how to prove a particular example of the second subcase previously mentioned, the general case can be proved in the same way. Let's take the case in which $S(\theta')$ is not a singleton, and let $\theta_{11},\theta_{21},\theta_{12},\theta_{22},\theta_{13},\theta_{23} \in s \in S(\theta')$ and $\varphi_1,\varphi_2 \in s' \in S(\theta')$ such that $\theta_{11} < \varphi_1 < \theta_{21} < \theta_{12} < \theta_{22} < \theta_{13} < \varphi_2 < \theta_{23}$ (see fig 41)

- Using the same argument as in theorem 9 we can show that areas A,B,C must be strictly negative.

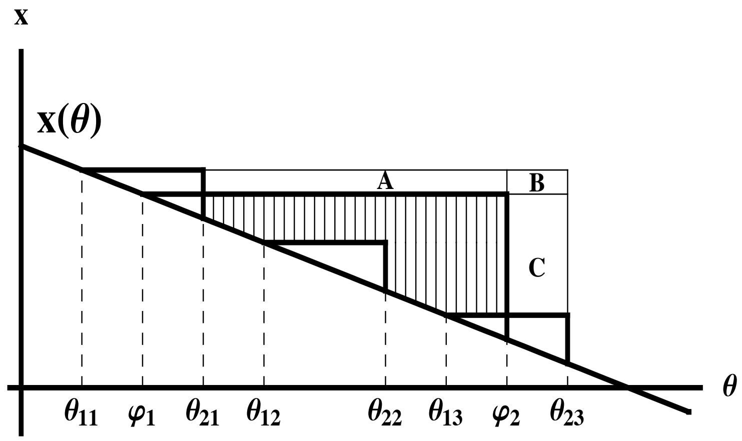

Figure 41: Decreasing Policy IV

- Since φ_1, φ_2 belong to a solution in $S(\theta')$ it must hold true $\Phi(\varphi_2, \varphi_1) \leq \Phi(\theta_{21}, \varphi_1) + \Phi(\theta_{22}, \theta_{12}) + \Phi(\varphi_1, \theta_{13})$. Thus, he shaded area must be less or equal to 0
- Therefore, $\Phi(\theta_{11}, \theta_{23}) < \Phi(\theta_{21}, \theta_{11}) + \Phi(\theta_{22}, \theta_{12}) + \Phi(\theta_{23}, \theta_{13}) + A + B + C +$ shaded area
- Thus we arrive to a contradiction.

Now we will show that for any θ such that $S(\theta)$ is not a singleton $\forall r \in S(\theta) \backslash \{\underline{s}, \bar{s}\}$ $\frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=\underline{s}} > \frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=r} > \frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=\bar{s}}$. Let's take some solution $\{\theta_{11}, \theta_{21}, ..., \theta_{1n}, \theta_{2n}\} = p \in S(\theta)$:

$$\frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=p} = \frac{\partial \sum_{j=1}^{n'_p} \Phi^{x[\theta]}(\theta_{2j}, \theta_{1j})}{\partial \theta} = \sum_{j=1}^{n'_p} \left(\int_{\theta_{1j}}^{\theta_{2j}} v_{x\theta}(y^{[\theta]}(\tilde{\theta}), \tilde{\theta}) \frac{\partial y^{[\theta]}(\tilde{\theta})}{\partial \theta} d\tilde{\theta} - \int_{\theta_{1j}}^{\theta_{2j}} v_{x\theta}(y^{[\theta]}(\theta_{1j}), \tilde{\theta}) d\tilde{\theta} \frac{\partial y^{[\theta]}(\theta_{1j})}{\partial \theta} \right)$$

- $\frac{\partial y^{[\theta]}(\cdot)}{\partial \theta} \geq 0$
- The local I.C.C. guarantees that $v_{x\theta}(y^{[\theta]}(\tilde{\theta}), \tilde{\theta}) < 0$. Thus, the term $\sum_{j=1}^{n'_p} \int_{\theta_{1j}}^{\theta_{2j}} v_{x\theta}(y^{[\theta]}(\tilde{\theta}), \tilde{\theta}) \frac{\partial y^{[\theta]}(\tilde{\theta})}{\partial \theta} d\tilde{\theta}$ is least for $p = \bar{s}$ and greatest for $p = \underline{s}$
- On the other hand, by the first order condition we have that $-\sum_{j=1}^{n'_p} \int_{\theta_{1j}}^{\theta_{2j}} v_{x\theta}(y^{[\theta]}(\theta_{1j}), \tilde{\theta}) d\tilde{\theta} \frac{\partial y^{[\theta]}(\theta_{1j})}{\partial \theta} = -\int_{\theta_{11}}^{\theta_{21}} v_{x\theta}(y^{[\theta]}(\theta_{11}), \tilde{\theta}) d\tilde{\theta} \frac{\partial y^{[\theta]}(\theta_{11})}{\partial \theta}$. Once again by the F.O.C. the term $-\int_{\theta_{11}}^{\theta_{21}} v_{x\theta}(y^{[\theta]}(\theta_{11}), \tilde{\theta}) d\tilde{\theta} \frac{\partial y^{[\theta]}(\theta_{11})}{\partial \theta}$ is different than 0 only if $\theta_{11} = 0$, in which case it is least for $p = \bar{s}$ and greatest for $p = \underline{s}$

Thus we have the following

$$\forall r \in S(\theta) \backslash \{\underline{s}, \bar{s}\} \quad \frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=\underline{s}} > \frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=r} > \frac{\partial Q^{y[\theta]}}{\partial \theta}|_{s=\bar{s}}$$

Using that $S(\theta)$ is upper hemicontiuous we can see that for any θ' such that $S(\theta')$ is not a singleton there exists a neighborhoods such that for all $\theta \in (\theta' - \epsilon, \theta')$ the correspondence $S(\theta)$ is a singleton and is arbitrarily close to $\underline{s}$ and for all $\theta \in (\theta', \theta' + \epsilon)$ the correspondence $S(\theta)$ is a singleton and is arbitrarily close to $\bar{s}$.

Let $\theta', \theta'' \in (\ddot{\theta}^y[K], \dot{\theta}^y[K])$ be such that $\Phi^{y[K]}(\theta', \theta'')$ is a local minimum (for local minimums in which θ'' or θ' are a corner solution the comparative statics is basically the same). We have the following equations (remember $\Phi^{y[K]}(\theta', \theta'') = \int_{\theta''}^{\theta'} \int_{y[K](\theta'')}^{y[K](y)} v_{x\theta'}(z, y) dz dy$):

$$\frac{\partial \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'} = \int_{y[K](\theta'')}^{y[K](\theta')} v_{x\theta'}(\tilde{x}, \theta') d\tilde{x} = v'_\theta(y[K](\theta'), \theta') - v'_\theta(y[K](\theta''), \theta') = 0$$

$$\frac{\partial \Phi^{y[K]}(\theta', \theta'')}{\partial \theta''} = -\int_{\theta''}^{\theta'} v_{x\theta'}(y[K](\theta''), \tilde{\theta}') d\tilde{\theta}' \frac{\partial y[K](\theta'')}{\partial \theta''} = (v_x(y[K](\theta''), \theta'') - v_x(y[K](\theta''), \theta')) \underbrace{\frac{\partial y[K](\theta'')}{\partial \theta''}}_{\neq 0} = 0$$

Now making the comparative statics with respect to K, we derive the previous equations with respect to K and we get the following equations:

$$\begin{pmatrix} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta'} & \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''} \\ \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial \theta'} & \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial \theta''} \end{pmatrix} \begin{pmatrix} \frac{\partial \theta'}{\partial K} \\ \frac{\partial \theta''}{\partial K} \end{pmatrix} + \begin{pmatrix} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial K} \\ \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial K} \end{pmatrix} = 0$$

inverting the matrix we get the following:

$$\begin{pmatrix} \frac{\partial \theta'}{\partial k} \\ \frac{\partial \theta''}{\partial k} \end{pmatrix} = \underbrace{\frac{-1}{\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta'} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial \theta''} - \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''}}}_{<0 (S.O.C.)} \begin{pmatrix} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial \theta''} & -\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''} \\ -\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''} & \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta'} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial K} \\ \frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial K} \end{pmatrix}$$

Calculating the terms:

$$\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta'} > 0 (S.O.C.)$$

$$\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial \theta''} > 0 (S.O.C.)$$

$$\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial \theta''} = -\underbrace{v_{\theta' x}(y[K](\theta''), \theta')}_{>0} \underbrace{\frac{\partial y[K](\theta'')}{\partial \theta'}}_{<0} > 0$$

$$\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta'' \partial K} = \underbrace{(v_{xx}(y[K](\theta''), \theta'') - v_{xx}(y[K](\theta''), \theta'))}_{<0} \underbrace{\frac{\partial y[K](\theta'')}{\partial K}}_{>0} \underbrace{\frac{\partial y[K](\theta'')}{\partial \theta'}}_{<0} > 0$$

$$\frac{\partial^2 \Phi^{y[K]}(\theta', \theta'')}{\partial \theta' \partial K} = v_{x\theta'}(y[K](\theta'), \theta') \underbrace{\frac{\partial y[K](\theta')}{\partial K}}_{>0} - v_{x\theta'}(y[K](\theta''), \theta') \underbrace{\frac{\partial y[K](\theta'')}{\partial K}}_{>0} < 0$$

replacing the signs we have the following:

$$\begin{pmatrix} \frac{\partial \theta'}{\partial k} \\ \frac{\partial \theta''}{\partial k} \end{pmatrix} = -\begin{pmatrix} + & - \\ - & + \end{pmatrix} \cdot \begin{pmatrix} - \\ + \end{pmatrix} = \begin{pmatrix} + \\ - \end{pmatrix}$$

Thus θ'' is decreasing in K and θ' is increasing in K which means the solutions are also nested when they are unique. $\square$