Dynamic Competitive Insurance*

Vitor Farinha Luz†

November 27, 2013

Abstract

I analyze long-term contracting in insurance markets with asymmetric information and a finite or infinite horizon. Risk neutral firms compete in the provision of long-term insurance contracts for risk-averse buyers. A risk-relevant characteristic is privately observed by the buyer and evolves stochastically over time following a Markov process. A long-term contract specifies the premium and coverage obtained by the buyer, possibly contingent on the history of type reports and accidents.

Optimal contracts offer a choice between a partial coverage policy and a perpetual complete coverage policy in each period. The analysis of the incentive constraints shows that contracts display a dynamic pricing scheme: the complete-coverage premium decreases with the number of periods the buyer chooses partial insurance, which serves as a signal of his initial type. Allocative inefficiency decreases along all histories. With an infinite horizon, all agents receive full coverage eventually, and the long-run distribution of premia is described by the length of partial coverage spells. I establish uniqueness of equilibrium with firms making zero profits, regardless of the initial type of the buyer. I also extend the qualitative properties of the contracts to the monopoly setting and provide necessary and sufficient conditions for equilibrium existence.

`Keywords:` Asymmetric Information; Mechanism Design; Insurance

`JEL Classification:` D43, D82, G22

*I would like to thank Dirk Bergemann, Johannes Hörner and Larry Samuelson encouragement and feedback on this project. I would also like to thank Brian Baisa, Eduardo Souza Rodrigues, Fernanda Ave, Jaqueline Oliveira, Bruno Badia, Sofia Moroni, Marina Carvalho, Marcelo Sant'anna, Sabyasachi Das and the participants of the Yale Micro lunch and Summer workshop for helpful comments.

†Department of Economics, Yale University. 28 Hillhouse Ave., New Haven, CT 06511, USA. E-mail: vitor.farinhaluz@yale.edu, website: http://vitorfarinhaluz.commons.yale.edu/.

1 Introduction

A vast majority of insurance contracts cover risks that are present over many periods (e.g. car, health and crop insurance). When choosing an insurance provider, buyers care not only about the available options of premium and coverage but also about the evolution of the policies available along his future interactions with the provider. In car insurance, firms explicitly advertise their driving rewards programs and offer "rate protection" plans that guarantee a fixed premium in future periods. In health insurance, reclassification caused by health episodes is a key concern in the design of regulation.

This paper studies the dynamic behavior of coverage and premium in long-term relationships between insurance buyers and providers using optimal contracts. From a firm's perspective, the repeated interaction with a customer generates valuable information that can be used in subsequent offers of premium and coverage, defined as dynamic pricing. The existing literature has considered the use of accident history (or experience rating) as a useful device for screening buyers with different risk levels (see Cooper and Hayes (1987) and Cohen (2005)). A new source of gains from dynamic pricing is presented, namely, the use of coverage choices as an additional information source. The coverage choice provides valuable information in the presence of asymmetric information. If buyers are better informed about their risk characteristics than insurance companies, buyers with positive information about their risk are interested in buying less coverage because their expected losses are lower. Firms interested in attracting agents that have (ex-ante) positive information are willing to reward the purchase of low coverage, by offering more attractive subsequent insurance contracts.

In this paper, I characterize the optimal contract under repeated interactions and the effects of such pricing for equilibrium welfare and coverage. The model considered is a dynamic extension of Rothschild and Stiglitz (1976)'s competitive insurance setting. I present a novel analysis of adverse selection allowing for the evolution of private risk characteristics through time, the presence of risk aversion and competition.

The literature on dynamic insurance contracts in the presence of asymmetric information focuses on the case of permanent risk characteristics.[1] The evolution of types is a natural assumption in many insurance examples, such as one's medical status (in health insurance) or expected driving needs (in car insurance). In addition, the insuree might obtain incremental information from events that do not lead to insurance claims, such as the occurrence of small accidents in car insurance. From a mechanism design perspective, the case of permanent

[1] See Kunreuther and Pauly (1985), Cooper and Hayes (1987) and Dionne and Doherty (1994).

types is special in that no choices by the customer reveal additional information beyond the initial period of interaction.

The evolution of private information in adverse selection has been considered in the case of linear consumption utility and monopoly (see Battaglini (2005), Pavan, Segal, and Toikka (2013) and Eső and Szentes (2013)). The presence of risk aversion (concavity of the utility function) is a crucial element of insurance analysis, and introduces new features to the analysis of efficient incentive provision. Customers with different characteristics can only be screened through the use of inefficient allocations, represented by partial insurance. In the presence of risk aversion the inefficiency cost of separating customers with different types depends on the consumption level and hence the dynamics of optimal distortions is directly related to the efficient intertemporal allocation of consumption.

The presence of competition is also a relevant feature of insurance markets. An analysis of the welfare consequences of long-term contracts is impossible in the absence of existence and uniqueness results for competitive equilibria. I provide a methodological contribution in extending the static characterization of competitive equilibria from Rothschild and Stiglitz (1976) to a dynamic setting with long-term contracts.

In the model considered, firms commit to a schedule of insurance policies over several periods. This commitment is captured by explicit long-term contracts. A flow insurance contract describes the coverage for each specific loss as well as the premium. In a long-term contract, a menu of flow contracts with different coverage levels is available to the customer in each period. Offers in a given period may depend on previous choices of coverage made by the buyer. Firms maximize the expected discounted net transfers received from the customer, accounting for premia received and potential benefits paid out.

The customer is risk averse and might incur incidental losses in each period. The potential losses are the source of risk that justifies the demand for insurance. Some examples are damages from a car accident (in car insurance) or expenditures from medical procedures (in health insurance). The losses are publicly observed at the end of each period and contractible, which allows firms to design and enforce contingent payments like indemnification.

Information is asymmetric between firms and customers. At the start of each period, the customer possesses risk-relevant information that allows him to better assess the distribution of possible losses from accidents. This information is not available to the insurance firms.[2] The private information of the customer is represented in the model by his type, which

[2]This could be the case because such information is privately observed, or because regulatory restrictions keep firms from using available information in pricing contracts.

evolves over time. Specifically, types follow a two-state Markov chain. The buyer is informed about the initial state of the process at the contracting stage. The high type (low type) represents higher (lower) expected consumption in the absence of insurance. The type process is assumed to be persistent, i.e., the realization of a high type in the current period increases the probability of a high type in the next period.

Many identical firms compete, once and for all, to attract the buyer in the first period. After the contracting stage, in each period the customer chooses among the set of available insurance contracts the selected firm is committed to offer. In the decision of which flow contract to choose, the customer takes into account the costs and benefits of different coverage levels in the current period, as well as the impact that his current choice has on the future flow contracts he has access to. It is assumed that customers are tied to a single firm after the contracting stage. I discuss the relevance of this assumption and provide special cases in which this assumption is not restrictive in Section 9.1.

Long-term contracts involve the voluntary revelation of the private information possessed by the customer. The choice of flow contract in each period is represented by different announcements by the customer. Contracts allow the flow allocation to depend on the complete history of type announcements.[3]

I provide a characterization of equilibrium contracts by analyzing the profit maximization problem faced by each firm. I solve for the most profitable long-term contract that delivers a specified expected utility for each initial type. Since insurance buyers privately observe new information in each period, the contract design problem has to account explicitly for incentive constraints in all possible contingencies. I simplify this problem by showing that one can restrict attention to a small subset of incentive constraints. Buyers with initial high type are rewarded with higher expected utility, which in equilibrium is a consequence of firms' incentives to attract such buyers. Because of the correlation in types over time, the efficient way to reward high-type buyers (without inducing misreports by low-type agents) is by rewarding the customer whenever he remains a high type. Using this characterization of the incentive constraints, I show that continuation contracts are efficient whenever the customer experiences a first "bad shock." In this environment, efficiency is equivalent to complete-coverage insurance with a fixed premium. As a consequence, in the infinite horizon model all agents eventually purchase complete insurance, and their premium is determined by the arrival of the first "bad shock."

[3]In the case of permanent types, previously considered in the literature, long-term contracts do not involve any choices beyond the first period because there is no additional information revelation (Cooper and Hayes (1987)).

A characterization of the dynamics of distortions (amount of residual risk) in this setting is difficult because the timing of consumption (or premia) affects the cost of screening across periods, and the flow allocation problem is potentially multidimensional since a flow policy specifies coverage for each possible loss. This problem is solved by considering an auxiliary (static) problem that determines the cost of separating different types, while delivering a given flow utility to a high-type buyer. Under a mild condition on the utility function (that includes CARA utility), the cost of separating types is increasing in the level of consumption. Efficient intertemporal allocation of consumption implies that the customer is rewarded with higher flow utility for consecutive "good" announcements, increasing the cost of separating different types in later periods. This implies that the distortions in the contract decrease through time; the difference in the flow utility obtained by each type decreases because utility becomes less responsive to the occurrence of losses. As the flow utility obtained by high-type agents increases and distortions decrease with time, the continuation utility obtained by a low-type also increases with time. Hence, the complete coverage premium decreases with the length of time with partial insurance.

As a result, in equilibrium firms engage in a simple dynamic pricing scheme: in every period the customer is offered a complete coverage insurance policy with a perpetual fixed premium and a short-term limited coverage insurance policy. The choice of limited-coverage insurance policies serves as a costly signal of good information held by the customer, leading to better offers in the next period: the complete coverage policy has a lower premium and the partial coverage policy delivers higher flow utility.

The analysis of competition is relevant for welfare analysis. I provide a novel characterization of competitive equilibria in markets with adverse selection and access to long-term contracts. I separate the firm's best response problem into two parts: (i) the design of contracts to minimize costs, for a fixed utility profile to be delivered to each initial type of customer and (ii) the choice of the attractiveness of its offer, i.e., the utility expected by each customer type. It is shown that equilibrium is uniquely determined by a zero profit condition. This result is established by showing that any offers generating nonnegative profits and featuring cross-subsidization (generating positive profits on one type of customer and losses on another) can be undermined by "cream skimming" offers that attract a single customer type. Offers that are not vulnerable to such deviations necessarily generate utility vectors on the boundary of the set of attainable utilities and are shown to generate negative profits. The zero-profits result implies that the access to the richer set of possible contracts leads to higher expected utility to consumers, compared to repeated purchase in spot (anonymous)

markets. Customers with initial low types receive their full information allocation, with perfect consumption smoothing.

In Section 8, I provide necessary and sufficient conditions for the existence of equilibrium. This condition extends the results of Rothschild and Stiglitz (1976). Equilibrium exists whenever the proportion of customers with initially low types is sufficiently high.

The next section discusses the related literature. The paper proceeds with a description of the model in Section 3. Section 4 presents the complete information case, which is useful to provide a benchmark for our analysis. Section 5 defines the profit maximization problem faced by each firm and the set of incentive-feasible utility profiles. The optimal allocation is characterized in Section 6. Section 7 characterizes competitive equilibrium, and conditions for existence are provided in Section 8. Section 9 discusses the extensions of one-sided commitment and monopoly. Finally, Section 10 concludes.

2 Related literature

This paper contributes to the literature on competitive screening, initiated by the seminal contributions of Rothschild and Stiglitz (1976) and Wilson (1977). These papers focus on situations in which private information leads to inefficiencies in competitive markets. Rothschild and Stiglitz (1976) considers insurance markets in which customers have private information regarding their risk characteristics. It shows that competition leads to a unique equilibrium in which high-type agents, who have lower accident probabilities, are screened by the choice of partial insurance at better premium rates (per unit of coverage).

In a related analysis, Cooper and Hayes (1987) considers a multi-period insurance model in which agents have fixed risk types and full commitment. The focus of the paper is on the relevance of experience rating as an efficient sorting device. Equilibrium existence and the dynamics of coverage and premium are not discussed. The optimal contract features a single type announcement in the first period; customers make no further choices.

My analysis is also related to the dynamic mechanism design literature focusing on the monopolist's revenue maximization problem with asymmetric information in dynamic settings (see Courty and Li (2000), Battaglini (2005), Eső and Szentes (2007), Pavan, Segal, and Toikka (2013), Eső and Szentes (2013)). My model differs from this body of research because of risk aversion and competition. The most related paper is Battaglini (2005), which considers the design of dynamic selling mechanisms in a monopoly setting. The customer's valuation follows a two-state Markov chain. In the optimal mechanism, production becomes

efficient when the customer obtains his first high valuation and converges to the efficient level along the history path of consecutive low valuations. Even though the productive allocation is fully characterized, the pricing scheme is not unique. This is a consequence of the linear structure of the model.

Pavan, Segal, and Toikka (2013) and Eső and Szentes (2013) consider continuous private information, while allowing for more general choice problems. They show that impulse response functions of the private type process drive the extent of allocative distortions. In the case of mean-reverting autoregressive processes allocative distortions disappear with time.

In the case of symmetric information, the benefits of long-term insurance contracts and the role of commitment in achieving those gains have been discussed extensively. Cochrane (1995) recognizes the inability of insuring long-term illnesses through short term contracts. It is shown that long-term contracts can be substituted by the use of short-term insurance contracts with respect to future premium uncertainty. Hendel and Lizzeri (2003) shows how customer commitment is important in providing insurance with respect to evolving characteristics. The use of front-loaded payments is used as a way to lock in customers within insurance companies.

An established literature on dynamic contracts considers the problem of efficient provision of incentives under asymmetric information, characterizing cost minimizing mechanisms. The seminal contributions in this literature are Green (1987), Thomas and Worrall (1990) and Atkeson and Lucas (1992). The focus of these papers is on the case of i.i.d. private types,[4] with the assumption of symmetric information at the time of contracting. The presence of taste shocks implies the efficiency of type-dependent consumption levels. The provision of incentives through future expected utility generates an ever-increasing dispersion in consumption. In the model considered here the discrimination of different types is useful as a tool for screening customer with different initial information. As the correlation between initial and future types diminishes with time, the dispersion in consumption within periods disappears.

The issue of potential non-existence of (pure strategy) equilibrium in adverse selection models appears first in Rothschild and Stiglitz (1976). This has lead to a large literature that puts forward alternative equilibrium concepts. Wilson (1977) and Riley (1979) consider reactive equilibrium concepts and show that they exist in the static insurance setting. Allowing

[4]A notable exception is Fernandes and Phelan (2000), which provides a recursive characterization of efficient mechanisms with two state Markovian types.

for the possibility of mixed strategies for the insurance offers, Dasgupta and Maskin (1986) shows that equilibrium always exists. More recently Farinha Luz (2012) shows that there is a unique symmetric mixed equilibrium. This equilibrium coincides with the Rothschild-Stiglitz equilibrium outcome whenever the latter exists. In the dynamic setting considered here, our conditions for equilibrium existence extend the ones provided in Rothschild and Stiglitz (1976). It seems plausible that existence can be guaranteed while allowing for stochastic offers.[5]

3 Model

A single agent lives for T periods, T being finite or infinite. At the beginning of each period, the agent (privately) observes his type $\theta_t \in \Theta$, which determines a probability distribution over realized income $y_t \in Y$, publicly observed at the end of the period and contractible. The occurrence of higher losses or damages is represented by a lower level of final income.[6] I assume that Y is a finite subset of $\mathbb{R}_+$. There are two possible types

$$\Theta \equiv \{\theta_l, \theta_h\}.$$

For a given type $\theta \in \Theta$, income y is distributed according to full-support probability $p_\theta \in \Delta Y$.[7] I assume that the θ_h-type agent, also referred to as the high type, has higher expected income, i.e.,

$$\sum_{y \in Y} p_{\theta_l}(y)\, y < \sum_{y \in Y} p_{\theta_h}(y)\, y.$$

I assume that types follow a Markov chain (μ, π). The distribution $\mu = (\mu_{\theta_l}, \mu_{\theta_h})$ determines the distribution of the initial type θ_1. The time-invariant two-by-two transition matrix $\pi = (\pi_{ij})_{i,j=l,h}$, where

$$\pi_{ij} = \mathbb{P}(\theta_{t+1} = j \mid \theta_t = i),$$

captures the evolution of types. I assume that types are positively correlated through time. More formally, a high-type buyer in period t has a higher probability of being a high type

[5] I leave this extension for future work.

[6] If the customer has fixed per period flow income $y_0 > 0$ and potential losses ℓ, the realized income in case of a loss is $y = y_0 - \ell$.

[7] This means that, for any $y \in Y$, $p_\theta(y) \in (0,1)$ and $\sum_{y \in Y} p_\theta(y) = 1$, for $\theta \in \{\theta_l, \theta_h\}$.

in period $t+1$ than a low-type buyer, i.e.,

$$\pi_{hh} > \pi_{lh}.$$

The complementary condition $\pi_{ll} > \pi_{hl}$ holds as a consequence. Note that $\pi_{ll} = \pi_{hh} = 1$ is included as a special case.

Preferences over final consumption are determined by Bernoulli utility

$$u : \mathbb{R}_+ \to \mathbb{R},$$

assumed to be strictly concave, increasing and C^2.[8] The agent discounts the future according to factor $\delta \in (0,1)$. The utility obtained from deterministic consumption stream $(c_1, \dots c_T)$ is given by

$$\sum_{t=1}^{T} \delta^{t-1} u\left(c_t\right).$$

An agent with initial type $\theta \in \{\theta_l, \theta_h\}$ has an outside option $\underline{V}_\theta \in \mathbb{R}$, which is restricted later in the analysis.

Firms are risk-neutral and also discount the future according to discount factor δ. The payoff (or profit) obtained by a firm is determined by its net payments made to the agent, which only occurs if the contract offered by this firm is accepted by the agent. The net transfer made to the firm, denoted f_t, is determined by the difference between the realized income y_t and the final consumption by the customer, c_t. Hence the payoff obtained by a firm is given by

$$\sum_{t=1}^{T} \delta^{t-1} f_t = \sum_{t=1}^{T} \delta^{t-1} \left(y_t - c_t\right).$$

Flow contracts. In a static environment, a contract determines the insurance premium to be paid, as well as the coverage level for all possible realizations of losses, leading to different income levels. I directly identify such contracts with the *final* amount of consumption the customer obtains. This should be understood as his realized income level *plus* net transfers received from the insurance firm. Formally the set of flow contracts, or flow allocations, is defined as

$$C \equiv \left\{c \mid c : Y \to \mathbb{R}_+\right\}.$$

[8] Note that this formulation rules out logarithmic utility, as its domain excludes zero. At least for finite T, there is no difficulty in extending my results to the case of diverging utility at zero, i.e., $\lim_{c \downarrow 0} u\left(c\right) = -\infty$.

Generic flow allocations are denoted as $c \in C$.

Long-term contracts. A long-term contract specifies from the outset of the model the flow contracts available to the customer in each period. In each period the customer sends a message $r \in R$ (R finite, with at least two elements) that leads to a specific flow contract. Different messages can be interpreted as choosing directly among the set of available flow contracts available in each period. The specification of premium and coverage in a given period depends explicitly on the (observed) outcomes from all past interactions between the firm and the customer. These are composed of previous accident realizations and messages chosen by the customer.

Formally, a history of reports is denoted as $r^t \equiv (r_1, \ldots, r_t) \in R^t$. A history of income realizations is denoted by $h_y^t = (y_1, \ldots, y_t) \in H_y^t \equiv Y^t$. A history is defined as $h^t \in H \equiv R^t \times H_y^t$. A long-term contract, $M \equiv (c_t)_{t=1}^T$, described the flow contract obtained by the agent in each circumstance:

$$c_t : R^t \times H_y^{t-1} \to C.$$

I denote the set of long-term contracts (or mechanisms) as $\mathcal{M}$. Notice that the period t flow contract, $c_t\left(r^t, h_y^{t-1}\right)$ depends on the current message choice. This allows the customer to communicate his current type, which is learned at the beginning of the period.[9]

Realization-independent contracts. As mentioned before, the use of dynamic pricing allows firms to use information from previous contract choices and income realizations in determining subsequent contracts available to the customer. It is my goal to highlight how the contract choice of the customer can be used as a screening device. As a consequence, in the characterization of optimal contracts I restrict attention mostly to a subset of contracts referred to as realization-independent. Such contracts only depend on the previous history through the previous announcements by the agent, and not on the history of past income realizations. I define the set of such contracts as $\underline{\mathcal{M}}$. Formally, $M = (c_t)_{t=1}^T$ is realization independent if for any $t = 1, \ldots, T$, $r^t \in R^t$ and $h_y^{t-1}, \tilde{h}_y^{t-1} \in H_y^{t-1}$

$$c_t\left(r^t, h_y^{t-1}\right) = c_t\left(r^t, \tilde{h}_y^{t-1}\right).$$

In Section 6 I consider the problem of revenue maximization restricted to this set of contracts and discuss how the properties extend to the case of general contracts.

The use of past accident realizations to screen agents is already characterized in the

[9]The definition of contracts does not allow for randomized transfers. The use of randomization does not lead to welfare improvements because uncertainty in transfers, conditional on observed income and reports, increases expected payments and does not affect incentives.

literature (see Cooper and Hayes (1987)). Some of my results can be extended to optimal contracts within the general class of dynamic contracts, and the relevance of this restriction is discussed in the next sections.

The timing in the model is as follows. All $N \geq 2$ firms simultaneously offer long-term contracts to the agent. The agent observes his initial type θ_1 and decides which firm's contract to accept, if any. If the buyer does not accept any offer, he receives payoff $\underline{V}_\theta$. If the agent accepts a contract, in each period he observes type $\theta_t \in \Theta$, then announces a message to the chosen firm. At the end of the period the income realization y_t is observed and the customer receives (or pays) transfers from the firm according to the long-term contract. The timeline is illustrated in Figure 1.

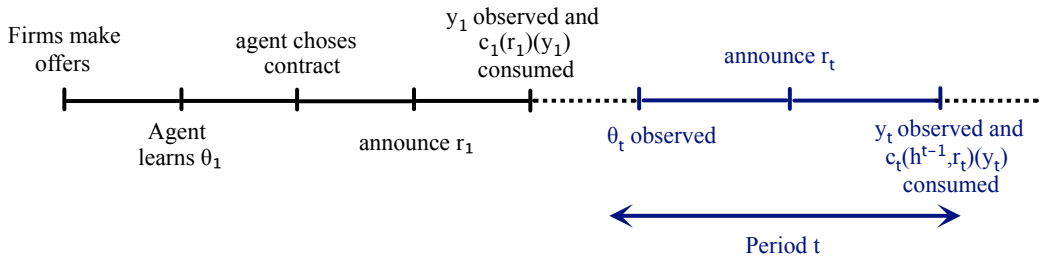

Figure 1: Timeline

A history of types is denoted as $h_\theta^t \in H_\theta^t \equiv \Theta^t$, for $t \leq T$. A reporting strategy is a function $\tilde{r} : H_\theta^T \times H_y^{t-1} \to R^T$ with the restriction that

$$\tilde{r}^t\left(h_\theta^T, h_y^{T-1}\right) \equiv \left(\tilde{r}_1\left(h_\theta^T, h_y^{T-1}\right), \ldots, \tilde{r}_t\left(h_\theta^T, h_y^{T-1}\right)\right)$$

is $\left(h_\theta^t, h_y^{t-1}\right)$-measurable for any $t = 1, \ldots, T$. A strategy profile consists of a contract proposal by each firm, a choice by the agent among the available contracts and the outside option, and the buyer's reporting strategy for any possible contract chosen. I characterize pure strategy[10] Perfect Bayesian Equilibria of this game.[11] This means that (i) for any

[10]The use of mixed reporting strategies does not change the results, but the use of pure contract offers is known to be with loss even in the static case (see Dasgupta and Maskin (1986) and Farinha Luz (2012)).

[11]Following the simultaneous contract offers (which conveys no information), the extensive form only involves moves by nature and the buyer and hence subgame perfection can also be applied.

contract accepted, customers follow an optimal reporting strategy, (ii) for any set of offers, the customer chooses the available offer delivering highest expected utility and (iii) given the expected offers by other firms and the customer strategies (for contract choice and reporting), firms' offers maximize profits.

The revelation principle. The revelation principle (Myerson (1986)) has no direct generalization to the case of competing principals. The main new issues introduced by competition are: (i) the use of contracts that (explicitly or implicitly) condition the allocation on the competing offers (see Epstein and Peters (1999), Peters (2001) and Martimort and Stole (2002)), and (ii) the use of latent contracts in the presence of non-exclusive contracting (see Arnott and Stiglitz (1991) and Attar, Mariotti, and Salanié (2013)).

The extensive form considered restricts attention to contracts that are exclusive and non-responsive. This means that buyers can only receive non-trivial transfers from a single firm (the chosen one) beyond the contracting stage and that the set of transfers the customer can obtain only depend on the contract offered by the chosen firm. This allows us to solve the model by studying the single firm revenue maximization problem, assuming the firm contracts with both types of agents and has to deliver specific type-dependent utility levels. The solution to this single-firm problem can be achieved by restricting attention to direct mechanisms (that only use two of the available messages) with truthful revelation. Hence the definition of incentive compatibility is still crucial to the analysis.

Incentive Compatibility. When focusing on the single firm profit maximization problem, we restrict attention to direct mechanisms with truthful revelation.[12] First define as $U_t^M\left(h^{t-1},\theta\right)$ the expected continuation utility assuming truth-telling after any arbitrary history $\left(\left(h_\theta^{t-1},\theta\right),h_y^{t-1}\right)$, where $h^{t-1}=\left(h_\theta^{t-1},h_y^{t-1}\right)$. Similarly, I define as $\Pi_t^M\left(h_\theta^t,h_y^{t-1}\right)$ the expected continuation profit obtained by a firm with contract M, assuming the agent tells the truth.

A contract M is said to be incentive compatible if

$$U_t^M\left(h^{t-1},\theta\right)\geq\sum_{y\in Y}p_\theta\left(y\right)u\left(c_t\left(h^{t-1},\theta'\right)\left(y\right)\right)$$
$$+\sum_{y,\tilde{\theta}}p_\theta\left(y\right)\mathbb{P}\left(\tilde{\theta}\mid\theta\right)U_{t+1}^M\left(\left(h_\theta^{t-1},\theta'\right),\left(h_y^{t-1},y\right),\tilde{\theta}\right),\quad\textbf{(IC-}(h^t,\theta)\textbf{)}$$

for all $t=1,\ldots,T$, $h^t=\left(\left(h_\theta^{t-1},\theta\right),h_y^{t-1}\right)\in H_\theta^t\times H_y^{t-1}$ and $\theta,\theta'\in\Theta$, and $r^t\notin\Theta^t\Rightarrow$

[12]For concreteness, assume that $R=\left(r^1,r^2,\ldots,r^{\#R}\right)$. Then identify r_1 with θ_l and r_2 with θ_h.

$c_t\left(r^t, h_y^{t-1}\right) = 0$.[13]

This condition states that the customer has no incentive to misreport his type in any given period. The left-hand side contains, by definition, the continuation utility obtained by an agent with history (h^{t-1}, θ). The right-hand side contains the obtained utility if the agent reports his type to be $\theta' \in \Theta$ in the current period, and reports his types truthfully from then on. In period t, the customer receives flow contract $c_t\left(h^{t-1}, \theta'\right) \in C$, which delivers final consumption $c_t\left(h^{t-1}, \theta'\right)(y)$ when income realization is $y \in Y$. From period $t+1$ on the discounted expected utility only depends on the realized state $\theta_{t+1} = \tilde{\theta}$.

In Markovian environments, the absence of profitable deviations involving single misreports on the equilibrium path (which is guaranteed by incentive compatibility) implies that there exists no profitable deviation from truth-telling that involves a finite number of periods. The absence of profitable deviations of finite length is sufficient to guarantee truth-telling is an optimal strategy if flow utility is bounded.

The set of incentive compatible contracts is denoted by $\mathcal{M}^{IC} \subseteq \mathcal{M}$. Similarly, define $\mathcal{M}^{IC_o}$ as the set of contracts that satisfy all the incentive constraints with strict inequality. Finally, define the set of expected utility vectors delivered to an agent be incentive compatible contracts, for initial types θ_l and θ_h, as $\mathcal{V}^{IC} \subseteq \mathbb{R}^2$:

$$\mathcal{V}^{IC} \equiv \left\{\left(U_1^M\left(\theta_l\right), U_1^M\left(\theta_h\right)\right) \mid M \in \mathcal{M}^{IC}\right\}.$$

4 Symmetric information benchmark

The assumption that firms and agents have distinct risk preferences creates gains from trade. The goal of this paper is to understand how the privacy of information affects efficient trade in this market. Therefore it is natural to start by analyzing the "perfect" scenario of competition under symmetric information, i.e., in which the risk types of the customer is publicly observed by all firms. This means that firms can directly contract on this information, deeming incentive concerns irrelevant.

In this case, firms separately compete for both agent types (low and high types), eliminating any possibility for profits. The resulting allocations are efficient, meaning that all the risk involved in the income process is resolved in the initial period, with the agent receiving exactly his expected future income. Formally, c_θ^{FB} denotes the full information outcome obtained by a buyer with initial type $\theta \in \Theta$:

[13]The last requirement guarantees that messages that are not used with direct revelation are indeed unattractive.

$$c_\theta^{FB} \equiv \mathbb{E}\left[\sum_{t\geq 1}\left(1-\delta\right)\delta^{t-1}y_t \mid \theta_1 = \theta\right].$$

Notice that the assumption about the induced distributions over income imply that $c_h^{FB} > c_l^{FB}$. This is the only outcome of symmetric information competition.

Proposition 1. *(Full information allocation)*
The unique symmetric information outcome is $M^{FB} = \left(c_t^{FB}\right)_{t\geq 1}$ *given by*

$$c_t^{FB}\left(h_\theta^t, h_y^{t-1}\right) = c_\theta^{FB},$$

if $h_\theta^t = (\theta, \theta_2, \ldots, \theta_t)$.

For future reference, I define the symmetric information utility obtained by a given type $\theta \in \{\theta_l, \theta_h\}$ as

$$V_l^{FB} = U_1^{M^{FB}}\left(\theta_l\right),$$

and similarly for θ_h.

The observability of the whole history h_θ^T is not necessary to generate such outcomes in a competitive model. As long as firms can observe (and contract) on the initial private information of agents, i.e., $\theta_1 \in \{\theta_l, \theta_h\}$, the equilibrium outcome would be exactly the same. This occurs because the efficient outcome (riskless consumption) does not depend on the contemporaneous type of the agent beyond the first period, and so does not require the elicitation of such information.[14]

5 Profit maximizing contracts

In this section, I define the firm's problem of maximizing profits, subject to providing a specific initial type-dependent utility $V = (V_l, V_h) \in \mathcal{V}^{IC}$. This problem is important to the analysis because, ignoring indifferences, the probability of attracting each type of buyer in the initial period depends solely on the utility provided to that type in the contract. The

[14]This result is in line with results in the dynamic mechanism design literature about irrelevance of private information that arises after the signature of a contract (see Eső and Szentes (2013)), focusing on quasi-linear monopolistic environments. However the reason for such result is quite different. In Eső and Szentes (2013), the efficient allocation depends on future type realizations, however the quasi-linearity (paired with the assumption of unrestricted liability) gives substantial leeway to the designer in constructing transfers that elicit information, while maintaining efficiency and extracting information rents from the agent.

firms' best response problem in the game consists of choosing the attractiveness of its offer to each type (or $V \in \mathcal{V}^{IC}$) and designing a contract that delivers this utility at minimal cost.

I proceed by proving existence and uniqueness of a solution to the problem and showing that weak incentives can be made strict with arbitrarily small cost. This virtual strong implementation result allows firms to break indifferences whenever profitable.

For each $V \in \mathcal{V}^{IC}$, define:

$$\Pi^*(V) \equiv \sup_{M \in \mathcal{M}} \Pi_0^M$$

subject to

$$M \in \mathcal{M}^{IC},$$

$$U_1^M(\theta_h) = V_h,$$

and

$$U_1^M(\theta_l) = V_l.$$

The solution $\Pi^*(V)$ maximizes the ex-ante profit obtained by a firm, choosing among long-term contracts that deliver utility vector $V = (V_l, V_h)$ to each type of agent and satisfy all the incentive compatibility constraints.

The solution of the problem is said to be essentially unique if any two solutions to $\Pi^*(V)$ are equal with probability one.[15] In the following lemma, I show that there is an essentially unique contract, denoted $M(V)$, that solves this problem, which I will call the efficient contract that delivers utility $V \in \mathcal{V}^{IC}$. The analysis considers the product topology on $\mathcal{M}$.

Proposition 2. *(Existence and uniqueness of optimal contracts)*
For each $V \in \mathcal{V}^{IC}$, there exists an essentially unique contract $M(V)$ that achieves $\Pi^*(V)$.

Proof. First assume that $\pi_{ll}, \pi_{hh} \in (0,1)$. Since $(V_l, V_h) \in \mathcal{V}^{IC}$, the feasible set is non-empty. Existence of a solution follows from upper semi-continuity of the objective function and compactness of the set

$$\left\{ M \in \mathcal{M} \mid \begin{array}{c} U^M(\theta_h) = V_{\theta_h}, \\ U^M(\theta_l) = V_{\theta_l}, \\ \Pi_1^M(\theta) \geq \Pi_1^{\overline{M}^\theta}(\theta). \end{array} \right\}.$$

[15] $M = (c_t)_{t=1}^T$ and $M' = (c'_t)_{t=1}^T$ are equal with probability one if

$$\mathbb{P}\left(\left\{\left(\theta^T, y^{t-1}\right) \in \Theta^t \times Y^{t-1} \mid c_s\left(\theta^s, y^{s-1}\right) = c'_s\left(\theta^s, y^{s-1}\right), \forall s \leq T\right\}\right) = 1.$$

Uniqueness follows from strict concavity of $u(\cdot)$. One can rewrite the problem with utility levels $u(c_t(h^t, \theta_t))$ as choice variables, in which case the problem involves the maximization of a concave function subject to a set of linear inequalities.

If $\pi_{ll} = 1$, then it is without loss to consider mechanisms $M = (c_t)_{t=1}^T$ such that $c_t\left(\theta_l, \theta^{t-1}, h_y^{t-1}\right)$ does not depend on $\theta^{t-1} \in \Theta^{t-1}$. An analogous argument holds for the case $\pi_{hh} = 1$. ∎

The characterization of equilibrium depends on the use of "cream skimming" deviations that only attract the buyer with a single initial type. Hence, the relevant firm payoff is given by interim profits, conditioning on the initial type of the buyer. For any $\theta \in \Theta$, define interim optimal profits as

$$\Pi^*(V \mid \theta) \equiv \Pi_1^{M(V)}(\theta).$$

As discussed earlier, the problem $\Pi^*(V)$ uses weak incentive constraints. In order to guarantee that $\Pi^*(V)$ can be achieved in equilibrium, I show that contracts in $\mathcal{M}^{IC}$ can be approximated by contracts with strict incentives (in $\mathcal{M}^{IC_0}$) with arbitrarily small profit loss. This implies that a firm can virtually obtain the same profits while providing strict incentives to the agent.

Lemma 1. *(Virtual full implementation)*
Consider a mechanism $M \in \mathcal{M}^{IC}$. There exists a sequence $\{M_n\}_{n\geq 0}$ in $\mathcal{M}^{IC_o}$ such that, for each $\theta \in \Theta$:

$$\begin{aligned} U^{M_n}(\theta) &\to U^M(\theta), \\ \Pi^{M_n}(\theta) &\to \Pi^M(\theta). \end{aligned}$$

Proof. In the appendix. ∎

6 Structure of the optimal contract

In this section, I provide a characterization of profit maximizing contracts. I start by proving a strong "distortion at the top" result for optimal contracts, which states that the flow contracts present inefficiency (partial coverage) solely on the history nodes that involve high type realizations. Then I provide a characterization of the intertemporal allocation of consumption and distortions. I show that, following subsequent high type realizations, flow contracts become more efficient and generate higher flow utility.

I restrict attention to realization-independent contracts. As discussed in Section 3, this class of contracts allows for current policy offer to depend on the history of announcements, while ignoring the previous realizations of income. I am interested in understanding the use of previous contract choices by the customer in determining future policy coverage and prices, hence I focus attention on contracts with such characteristics. The results in Section 6.1 can be generalized to the original problem and are discussed in the end of the section. The results in Section 6.2 have not been extended so far.

6.1 Allocative distortions

In this section, I show that revenue maximizing contracts involve complete coverage, except at history nodes that only involve consecutive high-type announcements. This result is a consequence of the set of binding incentive constraints in the profit maximization problem. I consider long-term contracts that deliver higher expected utility to agents with an initially high type. Hence the continuation contract obtained by customers with initially high types displays distortions (partial insurance) in order to prevent low type agents from misreporting. The time correlation in the type process implies that this "high-type reward" structure is propagated to future periods. The reason is that consecutive announcements of high type serve as a signal of an initial period high type, when compared to an initially low type customer that misreported his type.

The proof strategy is to consider a relaxed problem that only retains a small subset of incentive constraints. More precisely, it only considers the misreporting incentives faced by customers at their first low type realization.

The main structure theorem states that the crucial statistic used to screen agents is the number of consecutive periods before the (initially) high type agent has a first low shock ($\theta_t = \theta_l$). Therefore, I define the following sets of histories that share a common number of such periods: for any $n \geq 0$, define

$$\mathcal{H}_n = \left\{ h_\theta^\tau \in \cup_t H_\theta^t \mid h_\theta^\tau \succeq (\theta_h^n, \theta_l) \right\},$$

where $h_\theta^\tau \succeq h_\theta^t$, for $\tau \geq t$, implies $h_\theta^\tau = (h_\theta^t, \theta_{t+1}, \ldots, \theta_\tau)$, and

$$\mathcal{H}_\emptyset = \left\{ \left\{ \theta_h^t \right\} \mid t = 1, \ldots, T \right\}.$$

Denote as $\mathcal{H} \equiv \{\mathcal{H}_n \mid n \geq 0\} \cup \mathcal{H}_\emptyset$.

In order to discus efficiency, I define explicitly what is meant by efficient allocations. Given the risk aversion of agents and risk neutrality of firms, efficiency is equivalent to full coverage insurance: constant consumption, independent of income realization. Hence, I define the set of efficient allocations, or full coverage (static) insurance policies as

$$C^* = \{c \in C \mid c(y) \text{ is constant}\}.$$

I am interested in characterizing the solution to the following problem: for $V = (V_l, V_h) \in \mathcal{V}^{IC}$ with $V_h > V_l$,

$$\Pi^*(V) \equiv \max_{M \in \underline{\mathcal{M}}} \Pi_0^M$$

subject to

$$M \in \mathcal{M}^{IC},$$

$$U_1^M(\theta_h) = V_h,$$

and

$$U_1^M(\theta_l) = V_l.$$

The problem amounts to choosing contracts that maximize revenue subject to the usual incentive constraints and the restrictions of providing specified utility levels for each initial possible type.

Below I show that the "distortion at the top" result from the static model generalizes to the dynamic model in a very strong form: the allocation is efficient except at history nodes that only involve high shocks. Additionally, the allocation only depends on how many periods of consecutive high types the agent has announced.

Proposition 3. *(Presence of distortions)*
Any optimal contract is $\mathcal{H}$-measurable and is efficient except at nodes in $\mathcal{H}_\emptyset$, i.e.,

$$c_t(h) \notin C^* \Longleftrightarrow h \in \mathcal{H}_\emptyset.$$

This proposition is proved in the remainder of this section.

Relaxed Problem. Following the common approach in contract design theory, I consider the relaxed problem that only takes into some of the relevant incentive constraints. I define $\mathcal{M}^{IC*}$ as the set of contracts that satisfy all the incentive constraints for nodes in $\mathcal{H}_\emptyset$. We focus on the incentive constraints at the first low type realization of the buyer. Since

contracts reward high-type realizations, the binding incentive constraints always involve a low-type realization. Additionally, since allocation following a low-type announcement is not the target of relevant misreports, it is efficient, corresponding to a constant consumption. Hence incentive constraints following a low-type realization are irrelevant as well. Formally, the set of constraints in the relaxed problem is

$$\mathcal{M}^{IC*} = \left\{ M \in \mathcal{M}_S \mid \begin{array}{c} (\theta_h^\tau, \theta_l)\text{-IC is satisfied,} \\ \text{for any } \tau \geq 1, h_Y^\tau \in H_Y^\tau \end{array} \right\}.$$

The potential deviations from truth-telling considered in the constraint set $\mathcal{M}^{IC*}$ are graphically represented as arrows in Figure 2.

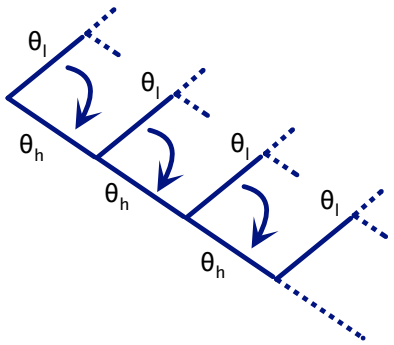

Figure 2: Binding incentive constraints.
Potential deviations from truth-telling considered in the constraint set $\mathcal{M}^{IC*}$ are represented as arrows.

The relaxed problem is

$$\Pi^* (V) \equiv \max_{M \in \underline{\mathcal{M}}} \Pi_0^M$$

subject to

$$M \in \mathcal{M}^{IC*},$$

$$U_1^M (\theta_h) = V_h,$$

and

$$U_1^M (\theta_l) = V_l.$$

The next proposition shows that the relaxed problem solution indeed has the properties described above. It also shows that the relaxed constraint set is minimal, in that all constraints hold as equalities in the solution.

Proposition 4. *(Distortions in the relaxed problem)*
Any solution to the relaxed problem is $\mathcal{H}$-measurable and also is efficient except at nodes in $\mathcal{H}_\emptyset$, i.e.,

$$c_t(h) \notin C^* \Longleftrightarrow h \in \mathcal{H}_\emptyset.$$

Moreover, all incentive constraints hold as equalities.

Proof. A necessary condition for optimality is existence of $(\lambda_t)_{t=1}^T, \mu, \kappa \geq 0$ such that: for $h^{t+1} = (\theta_h)^t$, optimality of $c_t(h^{t+1})(y)$ implies

$$\begin{array}{c} -\kappa + \mu u'(\cdot) \\ -\sum_{\tau<t} \lambda_\tau \frac{\mathbb{P}(\theta_h|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^\tau} u'(\cdot) - \lambda_t \frac{p_{\theta_l}(y)}{\mathbb{P}(h^{t+1}|\theta_1=\theta_h) p_{\theta_h}(y)} u'(\cdot) \end{array} \quad \geq \quad 0, \tag{1}$$

and for any h^{t+1} following $((\theta_h)^s, \theta_l)$ satisfies, optimality of $c_t(h^{t+1})(y)$ implies

$$\begin{array}{c} -\kappa + \mu u'(\cdot) \\ -\sum_{\tau<s} \lambda_\tau \frac{\mathbb{P}(\theta_h|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^{\tau-1}} u'(\cdot) - \sum_{\tau=s} \lambda_\tau \frac{\mathbb{P}(\theta_l|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^{s-1}\mathbb{P}(\theta_l|\theta_h)} u'(\cdot) \\ +\lambda_{s+1} \frac{1}{\mathbb{P}(\theta_h|\theta_h)^{s-1}\mathbb{P}(\theta_l|\theta_h)} u'(\cdot) \end{array} \quad \geq 0.$$

Finally, suppose that the period $t \geq 1$ incentive constraint holds strictly, i.e.,

$$U^M\left(\theta_h^t, \theta_l\right) > U\left(c_t\left(\theta_h^t, \theta_h\right) \mid \theta_l\right) + \sum_{\tilde{\theta}} \mathbb{P}\left(\tilde{\theta} \mid \theta_l\right) U_{t+1}^M\left(\left(h_\theta^t, \theta_h\right), \tilde{\theta}\right).$$

In this case consider the following mechanism, defined as $M^{[t]} = \left(c_\tau^{[t]}\right)$:

$$c_\tau^{[t]}(h^\tau, \theta) = \begin{cases} u^{-1}\left(U^M\left(\theta_h^t, \theta_l\right)\right) & \text{, if } (h^\tau, \theta_\tau) \succeq \left(\theta_h^t, \theta_l\right); \\ u^{-1}\left(U^M\left(\theta_h^t, \theta_h\right)\right) & \text{, if } (h^\tau, \theta_\tau) \succeq \left(\theta_h^t, \theta_h\right), \end{cases}$$

and $c_\tau^{[t]} = c_\tau$ otherwise.

$M^{[t]}$ satisfies all of the incentive constraints, except for period t. Additionally, $M^{[t]}$ is necessarily cheaper than M (strictly so if they differ). Therefore, considering the mechanism

$\tilde{M}$ defined by

$$u\left(\tilde{c}_\tau\left(h^\tau,\theta_\tau\right)(y)\right) \equiv (1-\varepsilon)\, u\left(c_\tau\left(h^\tau,\theta_\tau\right)(y)\right) + u\left(c_\tau^{[t]}\left(h^\tau,\theta_\tau\right)(y)\right),$$

for $\varepsilon > 0$ small enough satisfies all of the incentive constraints (because of the linearity of the incentive constraints in utility levels and the slack on the period t incentive constraint) and it is a strict improvement in terms of profits, a contradiction. Therefore it follows that $M = M^{[t]}$. Therefore, period t incentive constraints imply that consumption following θ_h^t-history is constant. Therefore, all incentive constraints for $\tau \geq t$ hold as equalities trivially. ■

Now I characterize sufficient conditions under which a solution to the relaxed problem satisfies the original incentive constraints. In the standard quasi-linear auction setting (as Mussa and Rosen (1978)), it is known that downward incentive constraints and monotonicity of the allocation (quantity increasing with the valuation) are sufficient to obtain global incentive constraints. In this setting, a similar result holds. However the monotonicity constraint becomes a requirement that contracts and preferences be aligned. As an example, in the static case, the condition reduces to

$$\mathbb{E}\left[u\left(c\left(\theta_h\right)\right) \mid \theta_h\right] - \mathbb{E}\left[u\left(c\left(\theta_l\right)\right) \mid \theta_h\right] \geq \mathbb{E}\left[u\left(c\left(\theta_h\right)\right) \mid \theta_l\right] - \mathbb{E}\left[u\left(c\left(\theta_l\right)\right) \mid \theta_l\right],$$

which is equal to

$$\sum_y \left[p_{\theta_h}(y) - p_{\theta_l}(y)\right]\left[u\left(c\left(\theta_h\right)(y)\right) - u\left(c\left(\theta_l\right)(y)\right)\right] \geq 0.$$

In the dynamic case, I require that contracts and and announcements are aligned, so that θ_h-annoucement leads to more attractive contracts for a θ_h-type agent. This is required for any announcement up to period t.

Proposition 5. *(Monotonicity condition)*
Consider a contract $M \in \underline{\mathcal{M}}$ such that, for any t and $h^t \in H_\theta^t$, the upward incentive constraint binds, i.e.,

$$U^M\left(h^t,\theta_l\right) = U\left(c_t\left(h^t,\theta_h\right) \mid \theta_l\right) + \sum_\theta \mathbb{P}\left(\theta \mid \theta_l\right) U^M\left(h^t,\theta_l,\theta\right),$$

and, for all $t \geq 0$, $\tau = 0,\ldots,T-t$, $\theta \in \Theta$ and $\theta' \neq \theta$, the following monotonicity constraint

holds

$$U\left(c_{t+\tau}\left(h^t,\theta,(\theta_h)^\tau\right)\mid\theta\right)\geq U\left(c_{t+\tau}\left(h^t,\theta,(\theta_h)^\tau\right)\mid\theta'\right), \tag{2}$$

then $M\in\mathcal{M}^{IC}$.

Finally, I prove that the solution to the relaxed problem satisfies the monotonicity condition. In order to prevent θ_l-type agents from announcing θ_h, the allocation obtained by high-type agents is such that events that are relatively more likely for θ_h-type agents lead to rewards (higher consumption), and events that are relatively less likely lead to punishment (low consumption). The allocation is otherwise constant across income realizations, so that all the required inequalities hold as equalities. This guarantees the strong necessary monotonicity constraint.

Proposition 6. *(Sufficiency of relaxed constraints)*
Any solution to the relaxed problem satisfies (2), therefore, it is optimal.

Proof. Fix $t\geq 1$ and $\tau\in\{0,\ldots,T-t\}$. The condition holds as an equality if (h^t,θ) contains at least one θ_l realization, since then it follows that

$$c_{t+\tau}\left(h^t,\theta,(\theta_h)^\tau\right)\in C^*.$$

Now, focus on $h=(\theta_h,\theta_h,\ldots,\theta_h)$ and $\theta=\theta_h$. From the necessary condition (1) it follows that

$$c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right)(y)\geq c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right)(y'),$$

if and only if

$$\frac{p_{\theta_h}(y)}{p_{\theta_l}(y)}\geq\frac{p_{\theta_h}(y')}{p_{\theta_l}(y')},$$

which implies that

$$U\left(c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right)\mid\theta_h\right)\geq U\left(c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right)\mid\theta_l\right).$$

■

General contracts. In the characterization of revenue maximizing contracts, I have focused so far on realization-independent contracts for simplicity. However, the main qualitative features of the solution are extended to the general case.

First, the "distortion at the top" result is directly extended: in each period the agent chooses between an inefficient partial insurance continuation contract and an efficient con-

tinuation contract, which involves a fixed consumption level that does not depend on time, income realizations and type realizations. This leads to strong characterization of inefficiency: the optimal allocation only features inefficient outcomes (partial insurance) in the event of consecutive high-type histories, which only involve θ_h-type realizations.

Second, in the previous section I have shown that consumption "outside the top" only depends on the number of consecutive high-type shocks at the beginning of the contract. A similar result remains true in the general case: the allocation depends on the history of income shocks occurring within consecutive high-type history at the beginning of the contract.

In order to define precisely these properties, I need to define the length of consecutive high-type histories,

$$\tau_H\left(h_\theta^T\right) \equiv \sup\left\{t \in \{1,\ldots,T\} \mid \theta_1 = \ldots = \theta_t = \theta_h\right\},$$

and set $\tau_H\left(\theta_l, \ldots \theta_l\right) = 0$.

Now the extended result is formally stated. The proof is completely described in the appendix. It follows the same steps as the proof for realization-independent contracts. I focus on a relaxed problem, and provide sufficient conditions for the solution to the relaxed problem to be feasible in the original problem.

Proposition 7. *(Distortions in general contracts)*
If $M = (c_t)_t$ is an optimal contract, then the function $\left(t, h^T\right) \mapsto c_t\left(h^T\right)$ is measurable with respect to $\left(\mathbf{1}\left\{\tau_H < t\right\}, h_y^{\tau_H \wedge t}\right)$ and

$$t > \tau_H\left(h_\theta^T\right) \Rightarrow c_t\left(h^T\right) \in C^*.$$

6.2 Consumption path

In the previous section, I have shown that the level of consumption in the full insurance continuation contract obtained by low type agents, or equivalently the per period premium charged by the firm, depends on how long the agent has waited to opt into the most comprehensive insurance contract. In equilibrium, this is equivalent to the first period in which the agent is part of the low type group, i.e., $\theta_t = \theta_l$. An important question, that still remains unanswered so far, is whether distortions are decreasing along all paths.

I show that, under a mild assumption on utility function u, flow utility obtained along the θ_h^T path is increasing through time and the flow contract features diminishing distortions.

This implies that the constant consumption level received by agents with low types increases in the number of periods before the first low type realization. In other words, agents are screened by the length of time they wait until buying full insurance. Consumers with later purchase dates receive less expensive insurance (equivalent to higher consumption level).

In order to analyze how distortions are spread through periods, I will use an auxiliary optimization problem, that determines the cost of introducing a wedge in the utility obtained from both agent's types for the same allocation. This cost is due to the distortion necessary to separate agents of different types. In this section I show how the relationship between the marginal distortion cost and wealth level play an important role in the determination of the optimal allocation. For the analyses in this section we assume that $\pi_{hh} < 1$. If types are constant the optimal long-term contract features repetition of the optimal static contracts (the ones obtained from the case $T = 1$).

I define the cost of delivering utility $v \in u(\mathbb{R})$, while maintaining a wedge Δ between the two types of agents as the solution of the problem

$$\chi(v, \Delta) \equiv \inf_{c:Y \to \mathbb{R}_+} \sum_y p_h(y)\, c(y),$$

subject to

$$\sum_y p_h(y)\, u(c(y)) = v,$$

and

$$\sum_y p_l(y)\, u(c(y)) = v - \Delta.$$

The domain of $\chi(\cdot)$ is given by the set of points for which the equalities considered has at least one solution:

$$A \equiv \left\{ (v, \Delta) \in \mathbb{R}^2 \mid \exists z \in [u(\mathbb{R}_+)]^Y \text{ s.t.: } \begin{bmatrix} \sum_y p_{\theta_h}(y)\, z(y) \\ \sum_y p_{\theta_l}(y)\, z(y) \end{bmatrix} = \begin{bmatrix} v \\ v - \Delta \end{bmatrix} \right\}.$$

In the appendix (Lemma 13), I show that $\chi(\cdot)$ is continuously differentiable in its interior, hence I define $\chi_v \equiv \frac{\partial \chi}{\partial v}$ and $\chi_\Delta \equiv \frac{\partial \chi}{\partial \Delta}$.

The connection between this static optimization problem and profit maximization is as follows. For a contract $M = (c_t)_t$, define as $v_t^M(h^{t-1}, \theta)$ the flow utility obtained in period t by a truth-telling agent with history (h^{t-1}, θ). This is the expected utility from the flow

contract $c_t\left(h^{t-1},\theta\right)$. Formally $v^M=\left(v_t^M\right)_{t=0}^T$, with $v_t^M: H^{t-1}\times\Theta\to\mathbb{R}$, is defined by by

$$v_t^M\left(h^{t-1},\theta\right)\equiv\sum_{y\in Y}p_\theta\left(y\right)u\left(c_t\left(h^{t-1},\theta\right)\left(y\right)\right).$$

The variable v_t^M represents the instantaneous utility derived by an agent in contract M. The set of feasible instantaneous utilities is given by

$$W\equiv\left\{\left(v_t\right)_t\mid\left(v_t\left(h_\theta^t\right),\Delta_t^v\left(h_\theta^t\right)\right)\in A,\,\forall t\geq 1\right\},$$

in which

$$\Delta_t^v\left(\theta_h^t\right)\equiv\left[\begin{array}{c} v_t\left(\theta_h^t\right)+\mathbb{E}^{\theta_t=\theta_l}\left[\sum_{\tau>t}\delta^{\tau-t}v_\tau\left(\theta_h^t,\widetilde{\theta}^{\tau-t}\right)\right] \\ -\left\{v_t\left(\theta_h^{t-1},\theta_l\right)+\mathbb{E}^{\theta_t=\theta_l}\left[\sum_{\tau>t}\delta^{\tau-t}v_\tau\left(\theta_h^{t-1},\theta_l,\widetilde{\theta}^{\tau-t}\right)\right]\right\}\end{array}\right],$$

and $\Delta_1^v\left(\theta_h\right)=v_1\left(\theta_h\right)+\mathbb{E}^{\theta_1=\theta_l}\left[\sum_{\tau>1}\delta^{\tau-t}v_\tau\left(\theta_h,\widetilde{\theta}^{\tau-1}\right)\right]-V_l$. Also define $\Delta_t^v\left(h_\theta^t\right)=0$, if $h^t\neq\theta_h^t$.

For a given mechanism M, $\Delta_t^v\left(\theta_h^t\right)$ is equal to the flow punishment a low-type buyer needs (relative to the flow utility obtained by a high-type buyer) to receive in period t, node θ_h^t, so that he is willing to truthfully report his type.

If $M\in\underline{\mathcal{M}}$ solves problem $\Pi^*\left(V\right)$, then v^M solves the following problem

$$\max_{v\in W}-\mathbb{E}\left[\sum_t\delta^{t-1}\chi\left(v_t\left(h^t,\theta_t\right),\Delta_t^v\left(h^t,\theta_t\right)\right)\mid\theta_1=\theta_h\right],$$

subject to: for $\theta\in\Theta$,

$$\mathbb{E}^{\theta_1=\theta}\left[\sum_{t>1}\delta^{\tau-t}v_\tau\left(h^\tau,\theta_\tau\right)\right]\geq V_\theta.$$

However, to efficiently spread distortions through time the marginal gain of delivering a single utility at time t and time $t+1$ must be equalized. Providing higher instantaneous utility at node θ_h^t is costly because it requires higher payments on average, and also because requires extra distortions in the allocation, so that incentive constraints are still satisfied. However, increasing the instantaneous utility delivered to the agent in period $t+1$, at history θ_h^{t+1},

also interferes with period t incentive constraint. Formally, optimality implies the following:

$$\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)+\chi_\Delta\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)=$$
$$\chi_v\left(v_{t+1}\left(\theta_h^t\right),\Delta_{t+1}\left(\theta_h^t\right)\right)+\chi_\Delta\left(v_{t+1}\left(\theta_h^t\right),\Delta_{t+1}\left(\theta_h^t\right)\right)+\frac{\pi_{lh}}{\pi_{hh}}\chi_\Delta\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right).$$

As shown in the proof of Proposition 5, the agent is rewarded for every extra θ_h announcement with higher (average) consumption, i.e., $v_t\left(\theta_h^t\right)<v_{t+1}\left(\theta_h^t\right)$. This means that the marginal cost of utility is higher in period $t+1$:

$$\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)<\chi_v\left(v_{t+1}\left(\theta_h^t\right),\Delta_{t+1}\left(\theta_h^t\right)\right).$$

Hence, indifference requires that the marginal cost of further distorting the allocation is smaller in period $t+1$, when compared to period t:

$$\chi_\Delta\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)>\chi_\Delta\left(v_{t+1}\left(\theta_h^t\right),\Delta_{t+1}\left(\theta_h^t\right)\right).$$

Concavity implies that the costs of distortions are increasing, i.e., $\frac{\partial^2\chi}{\partial\Delta\partial v}>0$. If the costs of distortion are increasing in utility level as well, i.e. $\frac{\partial^2\chi}{\partial\Delta\partial v}>0$, the wedge in utilities is decreasing:

$$\Delta_t\left(\theta_h^t\right)>\Delta_{t+1}\left(\theta_h^t\right).$$

Proposition 8. *(Flow utility and distortion path)*
Suppose $\frac{\partial\chi(v,\Delta)}{\partial\Delta}$ *is increasing in* v *and contract* M *is interior, then*

$$v_t\left(\theta_h^t\right)<v_{t+1}\left(\theta_h^{t+1}\right),$$

and

$$\Delta_t\left(\theta_h^t\right)>\Delta_{t+1}\left(\theta_h^{t+1}\right).$$

Proof. See Appendix. ∎

This statement means that, along the θ_h^T-path, continuation contracts always involve higher expected utility and lower distortion. Hence the efficient contract option becomes more attractive with more high type realizations. Since the complete coverage insurance policy available at time t is such that the θ_l-type agents are indifferent between their offer contract and obtaining the θ_h-contract. Hence full coverage offer also increases through time as long as the expected discounted continuation risk of a θ_l-type agent is constant through

time. This is true in the infinite horizon model ($T = \infty$), since the distribution over future types is stable, and in the finite case with an absorbing low type (case $T < \infty$ and $\pi_{ll} = 1$).

Corollary 1. *(Continuation utility)*
Suppose that $\frac{\partial \chi(v,\Delta)}{\partial \Delta}$ *is increasing in* v*, contract* M *is interior and either (i)* $T = \infty$ *or (ii)* $T < \infty$ *and* $\pi_{ll} = 1$*. Then*

$$U_t^M\left(\theta_h^t\right) < U_{t+1}^M\left(\theta_h^{t+1}\right), \tag{3}$$

and

$$c_t\left(\theta_h^{t-1}, \theta_l\right) < c_{t+1}\left(\theta_h^t, \theta_l\right). \tag{4}$$

Proof. Notice that

$$\begin{aligned} U_t^M\left(\theta_h^t\right) &= (1-\delta)\sum_{s\geq t}\delta^{s-t}\left(v_s\left(\theta_h^s\right) - \mathbb{P}\left(\theta_s = \theta_l \mid \theta_t = \theta_h\right)\Delta_s\left(\theta_h^s\right)\right), \\ &< (1-\delta)\sum_{s\geq t+1}\delta^{s-t}\left(v_{s+1}\left(\theta_h^{s+1}\right) - \mathbb{P}\left(\theta_s = \theta_l \mid \theta_t = \theta_h\right)\Delta_s\left(\theta_h^{s+1}\right)\right) \\ &= U_t^M\left(\theta_h^t\right). \end{aligned}$$

Similarly, the second part of the statement follows from $U_t^M\left(\theta_h^{t-1}, \theta_l\right) = u\left(c_t\left(\theta_h^t, \theta_l\right)\right)$ and

$$\begin{aligned} U_t^M\left(\theta_h^{t-1}, \theta_l\right) &= (1-\delta)\sum_{s\geq t}\delta^{s-t}\left(v_s\left(\theta_h^s\right) - \mathbb{P}\left(\theta_s = \theta_l \mid \theta_t = \theta_l\right)\Delta_s\left(\theta_h^s\right)\right), \\ &< (1-\delta)\sum_{s\geq t+1}\delta^{s-t}\left(v_{s+1}\left(\theta_h^{s+1}\right) - \mathbb{P}\left(\theta_s = \theta_l \mid \theta_t = \theta_l\right)\Delta_s\left(\theta_h^{s+1}\right)\right) \\ &= U_t^M\left(\theta_h^t, \theta_l\right). \end{aligned}$$

A similar argument holds for the case $T < \infty$ and $\pi_{ll} = 1$. ∎

The assumption about the cost function χ requires that introducing risk into the allocation of the agent is more costly for higher income levels. The requirement is that absolute risk aversion does not decrease too much with consumption. This is guaranteed if $u(\cdot)$ has non-decreasing absolute risk aversion (IARA). Recall that absolute risk aversion, for utility u, is defined as:

$$r_u(x) \equiv -\frac{u''(x)}{u'(x)}.$$

Proposition 9. *(Cost single crossing property)*
Assume $r : \mathbb{R}_+ \to \mathbb{R}_+$ is differentiable, then $\frac{\partial \chi(v,\Delta)}{\partial \Delta}$ is increasing in v if, and only if,

$$r'(x)\, u'(x) + 2\,[r(x)]^2 \geq 0.$$

Proof. In the appendix. ∎

7 Equilibrium

In this section, I show that competition leads to zero profits available in equilibrium. Since agents have private information about their initial types, this allows for the possibility of cross-subsidization. In order to show that this is not possible, I demonstrate that firms are always able to "cream skim" by making offers that attract only the type of agents for which there is a profitable offer to be made. The argument depends crucially on the fact that equilibrium utility provided to the agents lies in the interior of $\mathcal{V}^{IC}$, so that firms have leeway to change contracts in a way that attracts a single type of agent while being unattractive to the alternative type.

I proceed by showing that, whenever offers are on the boundary of $\mathcal{V}^{IC}$, they will necessarily generate negative profits. This result is important because it allows us to show that equilibrium utility vectors always lie in the interior of the feasible set. Therefore cross-subsidization among (initial) types is impossible in equilibrium.

Lemma 2. *(Negative boundary profits)*
For any $V_l \geq V_l^{FB}$, $\inf\left\{\Pi_h(V_l, V) \mid (V_l, V) \in \mathcal{V}^{IC}\right\} < 0$.

Proof. See Appendix. ∎

Now I am in position to show that, whenever firms are making profits out of the θ_h-type agents, the equilibrium utility obtained by agents is necessarily in the interior of the feasible set. This result, as discussed above, allows us to construct "cream-skimming" offers that only attract the more profitable high-type buyers.

Corollary 2. *(Interiority of firm-rational offers)*
If $V_l \geq V_l^{FB}$ and $\Pi_h(V_l, V_h) \geq 0$, then $V = (V_l, V_h) \in int\left(\mathcal{V}^{IC}\right)$.

Proof. From Lemma 2, it follows that $(V_l, V_h + \varepsilon) \in \mathcal{V}^{IC}$. The utility vectors $(V_l - \varepsilon, V_l - \varepsilon)$ and $(V_h + \varepsilon, V_h + \varepsilon)$ are also contained in $\mathcal{V}^{IC}$, but then (V_l, V_h) is in the interior of

$$co\left(\left\{\left(V_l, V_h + \varepsilon\right), \left(V_l - \varepsilon, V_l - \varepsilon\right), \left(V_h + \varepsilon, V_h + \varepsilon\right)\right\}\right) \subseteq \mathcal{V}^{IC}.$$

■

Notice that delivering first best utility profile $\left(V_l^{FB}, V_h^{FB}\right)$ without distortions is impossible, since the low type agent would benefit from pretending to be a high type. Therefore, the equilibrium allocation will feature risk on the continuation contract delivered to the high agent, so that it still gives him higher utility while deterring deviations by the low type.

I assume that the outside option of the agents is not high enough so that the market collapses.

Assumption: $\underline{V}_{\theta_l} < V_l^{FB}$ and $\underline{V}_{\theta_h} < \max\left\{u \in \mathbb{R} \mid \left(V_l^{FB}, u\right) \in \mathcal{V}^{IC} \text{ and } \Pi_h\left(V_l^{FB}, u\right) \geq 0\right\}$.

In this section I show that there exists a unique equilibrium allocation. The result is derived from the fact that, in equilibrium, there will be zero profit opportunities from each initial type of agent. This result generalizes the no cross-subsidization result of Rothschild and Stiglitz (1976) to a dynamic setting.

Lemma 3. *(Zero conditional profits)*
Any equilibrium utility vector V is contained in $\mathcal{V}^{IC}$ and

$$\Pi_\theta(V) = 0, \text{ for } \theta \in \Theta.$$

Proof. I start by showing that $V = (V_l, V_h) \in \mathcal{V}^{IC}$. If both agent types accept a mechanism with positive probability, then $V \in \mathcal{V}^{IC}$ since the allocation derived from any optimal pure strategy for the agents (that involves accepting a contract) is an incentive compatible mechanism that delivers utility vector V. Now assume agent with (initial) type θ_h does not accept any mechanism and has utility $V_h = \underline{V}_h$, while θ_l agent accepts some mechanism with positive probability. Nonnegativity of profits implies that $V_l \leq V_L^{FB}$. If $V_l < V_l^{FB}$, a firm that has its offer accepted by the θ_l agent with probability lower than $\frac{1}{2}$ benefits from offering contract

$$c_t\left(h^t, \theta\right) = u^{-1}\left(u_l + \varepsilon\right),$$

for $\varepsilon > 0$ small, since it attracts all the θ_l type agents with probability one. If $V_l = V_l^{FB}$, consider mechanisms $\{M_n\}_n$, with strict incentives, that approach $M\left(V_l^{FB}, \underline{V}_{\theta_h} + \varepsilon\right)$, for $\varepsilon > 0$ small. Any firm benefits from offering contract M_n, for n sufficiently large, since

it attracts type θ_h with probability one and $\Pi_h\left(V_l^{FB}, \underline{V}_{\theta_h} + \varepsilon\right) > 0$, for $\varepsilon > 0$ small. A similar argument eliminates the possibility that θ_l agents do not accept any mechanisms. If no agent accepts a mechanism with positive probability, the deviating offer designed above would determine a profitable deviation.

Now I assume that $V = (V_l, V_h) \in \mathcal{V}^{IC}$ and proceed by cases.

A) $\Pi_l(V) > 0$:

A.1) $\Pi_l(V) > 0$ and $\Pi_h(V) \geq 0$. Consider offer $\{M_n\}_n$ approaching $M(V)$. The profit obtained from offer M_n approaches

$$\mu_l \Pi_l(V) + \mu_h \Pi_h(V).$$

Therefore, any firm which has its mechanism accepted by the θ_l agent with probability below $\frac{1}{2}$ can profitably deviate to M_n, for n sufficiently large.

A.2) $\Pi_l(V) > 0$ and $\Pi_h(V) < 0$.

$\Pi_l(V) > 0$ implies that $V_l < V_l^{FB}$. Suppose firm i has its mechanism accepted by the θ_l type agent with probability less than $\frac{1}{2}$. This firm has a profitable deviation to the mechanism

$$c_t\left(h^t, \theta\right) = u^{-1}(u_l + \varepsilon),$$

for some $\varepsilon > 0$, since it attracts the θ_l type agents with probability one and does not make a loss from the θ_h type agents.

B) $\Pi_l(V) < 0$:

B.1) $\Pi_l(V) < 0$ and $\Pi_h(V) \leq 0$.

In this case the firm that offers a mechanism that is accepted by the θ_l type agent makes negative profits. A contradiction.

B.2) $\Pi_l(V) < 0$ and $\Pi_h(V) > 0$.

Suppose firm i has its mechanism accepted by type θ_l with positive probability, this means the mechanism delivers utility u_l to θ_l-type agent. Then a firm $j \neq i$ must attract a θ_h-type agent with probability one. Otherwise, consider offer $\{M_n^{\varepsilon}\}_n$, for n sufficiently large, approaching mechanism $M(V_l - \varepsilon, V_h + \varepsilon)$ for $\varepsilon > 0$ sufficiently small (so that $(V_l - \varepsilon, V_h + \varepsilon) \in \mathcal{V}^{IC}$).

The argument for case $B.2$) also eliminates the possibility that $\Pi_l(V) = 0$ and $\Pi_h(V) > 0$. The argument for case $B.1$) also deals with the case $\Pi_l(V) = 0$ and $\Pi_h(V) < 0$. ∎

Lemma 3 implies that, in equilibrium, there are no profit opportunities available from any of the two types of agents. An immediate consequence is that firms earn zero profits

in equilibrium. The fact that firms make zero profits implies that the low type agent must receive at least his first best utility level in equilibrium. This follows from the fact that, were this not the case, firms offering a contract slightly worse than the first best contract of the low type agent, which pays a fixed stream equal to his expected discounted income stream, would yield strictly positive profits. If the high type agents decided to consume this contract, it would generate even more profits. I state this result in the following lemma.

Lemma 4. *(Low type equilibrium utility)*
In any equilibrium,

$$V_l = V_l^{FB}.$$

Proof. Consider firm i that has its contract accepted with positive probability by a low type agent. Consider the outcome function $c = \{c_t\}_t$, with $c_t : h^t \to X$ denoting the outcome that follows if an agent accepts firm i's contract (if an agent follows a mixed reporting strategy, consider a specific pure reporting strategy in its support). This outcome function defines a direct mechanism M that is incentive compatible, and satisfies

$$U^M(\theta_l) \geq V_l,$$

$$U^M(\theta_h) \leq V_h.$$

This means that the profit obtained by firm i, from a low type agent, is at most

$$\Pi_l(V) = 0.$$

A similar argument follows to show that no firm makes profit out of the high type agent.

If $V_l < V_l^{FB}$, then a firm can make positive profits by offering a no risk contract that pays $\varepsilon > 0$ less than the first best contract of the low type agent.

If $V_l > V_l^{FB}$, than any firm that has its contract accepted with positive probability by the low type agent incurs negative total profits. This firm has negative profits in case the agent is a low type, and has at most zero profits if the agent is a high type (by the argument above).

Regarding the allocation, the solution to the problem

$$\begin{aligned} &\max_{M \in \mathcal{M}^{IC}} \Pi^M(\theta_l) \\ &\text{s.t. } U^M(\theta_l) \geq V_l^{FB}, \end{aligned}$$

is unique (given by $c_t = c_l^{FB}$). Therefore, if the equilibrium consumption of the low type agent differs from c_l^{FB}, it generates negative profits to the firm whose contract is accepted by the low type agent. ∎

However, the allocation obtained by the high type agent has to entail inefficiencies. This is due to the fact that firms have an incentive to separate both types of agents in order to attract solely the high type agents to their offer. The only way to separate agents in this setting is by introducing risk into the contract offered, since both agents have different perceptions about the relative probability of future events.

Lemma 5. *(High type equilibrium utility)*
In any equilibrium,

$$V_h^* \equiv \max\left\{u \in \mathbb{R} \mid \Pi_h\left(V_l^{FB}, u\right) \geq 0\right\},$$

moreover, $V_h^{FB} > V_h^* > V_l^{FB}$. *Additionally, the equilibrium allocation is given by* $M(V)$.

Proof. We know that $\Pi_h(\cdot)$ is strictly decreasing in V_h, therefore there is (at most) a unique utility level $\overline{V}_h$ such that $\Pi_h\left(V_l^{FB}, \overline{V}_h\right) = 0$, which is defined in the lemma above. The second part of the lemma follows from the fact that $\Pi_h\left(V_l^{FB}, V_l^{FB}\right) > 0$, since a firm offering the mechanism with fixed consumption c_l^{FB} generates strictly positive profits, and $\Pi_h\left(V_l^{FB}, V_h^{FB}\right) < 0$ since any non-constant contract delivering utility V_h^{FB} generates losses.

Suppose that the equilibrium allocation is not given by $M(V)$. Then there is an initial type θ whose continuation allocation is different from $M(V)$. Consider firm i whose contract is accepted with positive probability by type θ agent. Consider the outcome function $c = \{c_t\}_t$, in which $c_t : h^t \to X$ denotes the outcome that follows if an agent accepts firm i's contract (if an agent follows a mixed reporting strategy, consider a specific pure reporting strategy in its support). This outcome function defines a direct mechanism $\overline{M}$ that is incentive compatible and satisfies

$$U^{\overline{M}}(\theta_l) \geq V_l,$$

$$U^{\overline{M}}(\theta_h) \leq V_h,$$

and is not equal to $M(V)$ following an initial θ announcement. Therefore it generates strictly less profits than $\Pi_\theta(V) = 0$. ∎

8 Existence

This paper extends the analysis adverse selection in insurance markets by Rothschild and Stiglitz (1976) to settings with long-term contracts. However, as largely discussed in the literature of competitive screening (Rothschild and Stiglitz (1976)), existence of pure strategy equilibria is not guaranteed in static models. The same result extends to the dynamic setting considered here. Characterizing equilibrium existence reduces to checking whether the contract described in Section 7 leaves open the possibility of profit by a single deviation.

There are two possible ways in which a firm might be interested in deviating. It might offer a new contract that only attracts the high type agents, without attracting the low type agents. This means that it delivers a utility below V_l^{FB}, the one obtained in the status quo contract. The contract described in Section 7 is designed such that the high type agent receives the highest possible ex-ante utility without generating losses, or inducing misreporting by the low type agent. Therefore, the only possible way to generate positive profits is to offer a new contract, namely M', that attracts both types of agents and still generated positive profits on average.

By offering a more attractive contract to customers with initially low types, given by $M\left(V_l^{FB}+\varepsilon, V_h^*\right)$, a firm makes losses on the initially low-type buyers, since $\Pi_{\theta_l}\left(V_l^{FB}+\varepsilon, V_h^*\right) < 0$. However, the change generates slack on the incentive constraints of the problem, allowing the firm to offer the same utility level to a high-type agent with lower cost, i.e.,

$$\Pi_{\theta_h}\left(V_l^{FB}+\varepsilon, V_h^*\right) > 0 = \Pi_{\theta_h}\left(V_l^{FB}, V_h^*\right).$$

The local deviation considered is profitable if the gain, obtained from initially high-type buyers, outweighs the loss on the initially low-type buyers. Hence, if the share of high-type buyers in the population (or the prior) is small, the proposed candidate is an equilibrium. The focus on local deviations is warranted by the concavity of $\Pi_\theta(\cdot)$. The condition guaranteeing that the local deviation is not profitable is presented in the next proposition. We use the notation $\frac{\partial_+ f}{\partial x}$ to denote the right partial derivative of a (convex) function f with respect to x.

Proposition 10. *(Existence conditions)*
Define V_h such that:

$$\Pi_h\left(V_l^{FB}, V_h\right) = 0.$$

Each firm offering contract $M(V)$ constitutes (part of) an equilibrium if, and only if,

$$\frac{\mu_h}{\mu_l} < \frac{1}{u'\left[u^{-1}\left(\left(1-\delta\right)V_l^{FB}\right)\right]}\frac{1}{\frac{\partial_+\Pi_h(V)}{\partial V_l}}. \tag{5}$$

Proof. The fact that V_h is well defined follows from Lemma 2. By definition of V, for any (V_l', V_h') with $V_l' \leq V_l$ and $V_h' \geq V_h$, it follows that $\Pi_h\left(V_l', V_h'\right) \leq 0$.

By way of contradiction, suppose there is $\mathcal{V}^{IC} \ni V' = (V_l', V_h') \geq V$ (with $V_l' \leq V_h'$) such that the offer $M\left(V'\right)$ generates positive expected, i.e.,

$$-\mu_l \frac{1}{1-\delta} u^{-1}\left(V_l'\left(1-\delta\right)\right) + \mu_h \Pi_h\left(V'\right) \geq 0 = -\mu_l \frac{1}{1-\delta} u^{-1}\left(V_l^{FB}\left(1-\delta\right)\right) + \mu_h \Pi_h\left(V\right).$$

Concavity of the function Π_h implies that

$$\begin{aligned} 0 &< \left[-\mu_l \frac{1}{u'\left[u^{-1}\left(u_l^{FB}\left(1-\delta\right)\right)\right]} + \mu_h \frac{\partial \Pi_h\left(\mathbf{u}\right)}{\partial u_l}\right]\left(u_l' - u_l^{FB}\right) + \mu_h \frac{\partial \Pi_h\left(\mathbf{u}\right)}{\partial u_h}\left[u_h' - u_h\right] \\ &\leq \left[-\mu_l \frac{1}{u'\left[u^{-1}\left(u_l^{FB}\left(1-\delta\right)\right)\right]} + \mu_h \frac{\partial \Pi_h\left(\mathbf{u}\right)}{\partial u_l}\right]\left(u_l' - u_l^{FB}\right) \leq 0. \end{aligned}$$

■

The right-hand side in equation (5) depends on the preferences, type and income distribution, but it does not depend on the prior μ_h. Hence it follows that it imposes a threshold $\overline{\mu}_h \in (0,1)$ such that equilibrium exists if, and only if, $\mu_h \leq \overline{\mu}_h$.

9 Extensions

9.1 One-sided commitment

In the model, customers are able to commit to long-term contracts with insurance firms. The customer commitment assumption is important in obtaining the result of efficiency after a low-type realization. Hence the complexity of the contract design problem is reduced to characterizing distortions and flow utility along the θ_h^T-path. In this section we discuss in which situations this assumption is irrelevant, which occurs if the buyer has no incentive to abandon the contract.

As discussed in the literature (Hendel and Lizzeri (2003), Pavan, Segal, and Toikka (2013)), customer commitment can be achieved through the use of bond posting schemes or contract termination fees, which impose a cost to the agent for leaving the contract. On the same direction, customer commitment can be interpreted as the limiting case of a model

with high search costs. The relevance of search costs in insurance markets has been discussed in Dahlby and West (1986) and Mathewson (1983).

The main result shows that, as the contract unravels, distortions decrease and contracts become more beneficial (in the infinite horizon model). In fact, the highest gain from reneging on the contract is available to customers that purchase complete insurance in the first period, but become high-types in afterwards. Formally, in equilibrium this would correspond to a buyer with type history $h_\theta^2 = (\theta_l, \theta_h)$. As a consequence, if buyers have sufficiently high reneging costs so that it is not optimal to abandon the contract in the first two periods, then there are no incentives to abandon the contract in subsequent periods.

The result that the customer commitment problem becomes less severe as time goes by is at odds with the literature on dynamic contracts (see Atkeson and Lucas (1992)), in which dispersion in the distribution of utility continually increases. This is due to the fact that, in my model, efficiency implies no consumption dispersion at all. Ex-post dispersion in consumption is inefficient and is only introduced as a way for firms to screen agents with respect to their initial private information. As the process considered has a mean reversion feature, future types become less correlated with initial conditions as the horizon increases, leading to the decreasing distortions result. This is at odds with the model considered in Atkeson and Lucas (1992), in which efficiency and incentives require utility dispersion within each period.

Absorbing state. An informative benchmark is the case of absorbing bad state, i.e., $\pi_{ll} = 1$. In the case of health insurance, this can be understood as the revelation of a chronic disease. In this specific case, the continuation utility within the contract necessarily increases with time and hence customers never have incentives to renege on the optimal long-term contract. This occurs because initially low-type agents benefit from the use of long-term contracts by anticipating consumption from future periods, in the case he becomes a high-type. This direction of intertemporal consumption reallocation generates incentives for reneging. Once the agent becomes a high-type, the contract imposes a debt from consumption in previous periods, which can be ignored by leaving the contract. In the case of an absorbing bad state, the permanence of low-types removes this source of gains.

On the other hand, initially high-type customers benefit from receiving lower premium if they become low-type in the future, which means postponing consumption. This direction of intertemporal consumption reallocation reduces the gain from reneging on the contract in future periods.

Proposition 11. *(Absorbing state)*

Consider the case $\pi_{ll} = 1$. *If* (3) *and* (4) *hold, then for any* $t \geq 1$ *and for (almost) all* $(h^{t-1}, \theta) \in H_\theta^t$*:*

$$U_t^M\left(h^{t-1}, \theta\right) \geq U_1^M\left(\theta\right).$$

The proposition above shows that continuation utility within the contract never goes below the initial expected utility from the contract. This implies that, if customers outside option is given by a fixed level $\underline{V}\left(\theta_t\right)$, it would never be in their interest to renege on the contract. This would naturally be the case for $T = \infty$, in which the best available option to customers would be to obtain another contract, i.e., $\underline{V}\left(\theta_t\right) = U_1^M\left(\theta_t\right)$.

9.2 Monopoly

So far, I have focused on the analysis of competitive insurance markets. However, the main characterization results follow directly from the single firm profit maximization problem. The main difference between the competitive and monopoly settings is the determination of initial utility vector $V \in \mathcal{V}^{IC}$. In the competitive case, initial utilities can be understood as equilibrium variables, that are adjusted so that firms have no profit opportunities when optimizing over possible long-term contracts. In the monopolist problem, the initial utility vector is freely determined by the seller, subject to the participation constraints. In this section I assume that higher types have more attractive outside offers, i.e., $\underline{V}_{\theta_l} \geq \underline{V}_{\theta_h}$.

For simplicity, I assume that it is always profitable for the monopolist to sell insurance to both types of agents. This is guaranteed by assuming $\left(\underline{V}_{\theta_l}, \underline{V}_{\theta_h}\right) \in \mathcal{V}^{IC}$ and

$$\Pi_{\theta_h}\left(\underline{V}_{\theta_l}, \underline{V}_{\theta_h}\right) \geq 0.$$

In this case, the monopolist's optimal offer, denoted as $M^{monopoly} \in \mathcal{M}$, solves the same cost minimization problem as the competitive equilibrium offers. The most important feature of the equilibrium utility level is that, as in the competitive model, initially high-type agents receive better continuation contracts. Hence the characterization results presented in Propositions 3 and 8 apply directly.

The next proposition shows that the monopolist's optimal contract solves $\Pi^*\left(V_l, \underline{V}_{\theta_h}\right)$, with $V_l < \underline{V}_{\theta_h}$. The participation constraint for the initially high-type buyer necessarily binds. However, the low-type buyer's participation constraint might be slack. This can be optimal because increasing the utility offered to the low-type buyer relaxes the incentive constraint binding in the profit maximization problem, leading to higher profits from the high-type buyers ($\Pi_{\theta_h}(\cdot)$ decreases).

Proposition 12. *The optimal offer is* $M^{monopoly} \equiv M\left(V_l, \underline{V}_{\theta_h}\right)$*, where* $V_l \in \left[\underline{V}_{\theta_l}, \underline{V}_{\theta_h}\right)$ *solves the problem*

$$\max_{V_l \geq \underline{V}_{\theta_l}} \mu_l \Pi_l\left(V_l, \underline{V}_{\theta_h}\right) + \mu_h \Pi_h\left(V_l, \underline{V}_{\theta_h}\right).$$

Proof. First notice that, if agents ex-ante utility is given by (V_l, V_h), then the offered mechanism is necessarily $M(V_l, V_h)$. The optimal mechanism necessarily satisfies

$$U_1^M(\theta_h) \geq U_1^M(\theta_l),$$

otherwise offering $M\left(U_1^M(\theta_l) - \varepsilon, U_1^M(\theta_l)\right)$ would be a strict improvement for the seller. It is feasible since $\underline{V}_{\theta_l} \leq \underline{V}_{\theta_h} \leq U_1^M(\theta_h) < U_1^M(\theta_l)$ and $\left(U_1^M(\theta_l) - \varepsilon, U_1^M(\theta_l)\right) \in \mathcal{V}^{IC}$ (by convexity of $\mathcal{V}^{IC}$).

Now assume that the optimal mechanism satisfies $U_1^M(\theta_h) > \underline{V}_{\theta_h}$, then by similar reasoning, offering $M\left(U_1^M(\theta_l), U_1^M(\theta_l) - \varepsilon\right)$ leads to a strict improvement. Therefore

$$\underline{V}_{\theta_h} = U_1^M(\theta_h) \geq U_1^M(\theta_h) \geq \underline{V}_{\theta_l}.$$

■

10 Conclusion

In this paper I provide a characterization of optimal dynamic pricing and delineate the welfare effects of the use of such pricing schemes. A deeper understanding of the role of long-term contracts is relevant to the debate of regulation of the pricing policies by insurance firms. This discussion was recently raised in the United States due to the implementation of new Patient Protection and Affordable Care Act (see Handel, Hendel, and Whinston (2013)). A goal of this policy is to restrict how firms can utilize previous information from customers in insurance pricing.

Part of the empirical literature considers the testable implications from adverse selection in insurance markets, assuming the dynamic structure of contracts observed is efficient (such as Dionne and Doherty (1994) and Hendel and Lizzeri (2003)). The characterization of optimal dynamic pricing presented here displays new implications of asymmetric information for the dynamics of coverage and premium[16]: (i) the purchase of partial insurance, serving

[16]The empirical relevance of these predictions is subject to the restrictions imposed by regulation in each market. In the U.S., health insurance regulation has moved in the direction of stricter pricing regulation

as a costly signal of a good type, leads to lower premium per unit of coverage in subsequent periods and (ii) agents with longer firm relationships obtain more coverage on average.

The extension of the competitive equilibrium characterization of Rothschild and Stiglitz (1976) to the dynamic setting considered here could be extended to other settings with adverse selection and long-term relationships. The commitment assumption effectively reduces competition to a static problem. The technical difficulties amount to characterizing profits for offers delivering extreme payoffs.

The commitment assumption on the side of firms does not require the signature of explicit long-term contracts with customers. As long as customers understand the dynamic pricing scheme used by firms, it is optimal for the firms to publicly commit to a sequence of contracts offers. In the case of car insurance, firms offer policies that explicitly involve commitment regarding future premia, an example being policies that include "accident forgiveness" clauses and explicit bonus for good driving.[17]

However, commitment by the consumer might not be reasonable and the one-sided commitment problem deserves further consideration. The characterization of the extreme case of an absorbing state presented here provides a class of examples in which customer commitment is irrelevant.

Another important avenue for research is how consumer's option to maintain a longer relationship with their current firms, by simply renewing policies, affects spot market liquidity and market risk-distribution.

and federal overhaul (Austin and Hungerford (2009)). The regulation of automobile insurance rating is done at state-level, with wide variation in the level of restrictions imposed. Witt and Urrutia (1983) propose a classification of the regulatory rules applied in each state in the U.S., considering 17 states as "competitive" and 33 as "non-competitive". "Competitive" states have more lenient requirement on underwriting. A more recent discussion of the different regulations across states is presented in Grace and Klein (2009).

[17]Most major U.S. car insurance companies offer some version of "accident forgiveness", with the details changing substantially.

References

Arnott, R., and J. Stiglitz (1991): "Equilibrium in competitive insurance markets with moral hazard," Discussion paper, National Bureau of Economic Research.

Atkeson, A., and R. E. Lucas (1992): "On efficient distribution with private information," *The Review of Economic Studies*, 59(3), 427–453.

Attar, A., T. Mariotti, and F. Salanié (2013): "Nonexclusive Competition under Adverse Selection," *Theoretical Economics*.

Austin, D. A., and T. L. Hungerford (2009): "The market structure of the health insurance industry," *Congressional Research Service R40834, Library of Congress*.

Battaglini, M. (2005): "Long-term contracting with Markovian consumers," *American Economic Review*, pp. 637–658.

Cochrane, J. H. (1995): "Time-consistent health insurance," *Journal of Political Economy*, pp. 445–473.

Cohen, A. (2005): "Asymmetric information and learning: Evidence from the automobile insurance market," *Review of Economics and Statistics*, 87(2), 197–207.

Cooper, R., and B. Hayes (1987): "Multi-period insurance contracts," *International Journal of Industrial Organization*, 5(2), 211–231.

Courty, P., and H. Li (2000): "Sequential screening," *The Review of Economic Studies*, 67(4), 697–717.

Dahlby, B., and D. S. West (1986): "Price Dispersion in an Automobile Insurance Market," *Journal of Political Economy*, 94(2), pp. 418–438.

Dasgupta, P., and E. Maskin (1986): "The Existence of Equilibrium in Discontinuous Economic Games, II: Applications," *The Review of Economic Studies*, 53(1), 27–41.

Dionne, G., and N. A. Doherty (1994): "Adverse selection, commitment, and renegotiation: Extension to and evidence from insurance markets," *Journal of Political Economy*, pp. 209–235.

Epstein, L. G., and M. Peters (1999): "A revelation principle for competing mechanisms," *Journal of Economic Theory*, 88(1), 119–160.

Eső, P., and B. Szentes (2007): "Optimal information disclosure in auctions and the handicap auction," *The Review of Economic Studies*, 74(3), 705–731.

Eső, P., and B. Szentes (2013): "Dynamic Contracting with Adverse Selection: An Irrelevance Result," *Working paper*.

Farinha Luz, V. (2012): "Characterization and uniqueness of equilibrium in competitive insurance," *Working paper*.

Fernandes, A., and C. Phelan (2000): "A recursive formulation for repeated agency with history dependence," *Journal of Economic Theory*, 91(2), 223–247.

Grace, M. F., and R. W. Klein (2009): "Insurance regulation: The need for policy reform," *The Future of Insurance Regulation*.

Green, E. (1987): "Lending and the smoothing of uninsurable income," *Contractual arrangements for intertemporal trade*, 1, 3–25.

Handel, B. R., I. Hendel, and M. D. Whinston (2013): "Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk," *NBER Working Paper No. 19399*.

Hendel, I., and A. Lizzeri (2003): "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *The Quarterly Journal of Economics*, 118(1), 299–328.

Kunreuther, H., and M. Pauly (1985): "Market equilibrium with private knowledge: An insurance example," *Journal of Public Economics*, 26(3), 269–288.

Martimort, D., and L. Stole (2002): "The revelation and delegation principles in common agency games," *Econometrica*, 70(4), 1659–1673.

Mathewson, G. F. (1983): "Information, Search, and Price Variability of Individual Life Insurance Contracts," *The Journal of Industrial Economics*, 32(2), pp. 131–148.

Mussa, M., and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic Theory*, 18(2), 301–317.

Myerson, R. B. (1986): "Multistage games with communication," *Econometrica: Journal of the Econometric Society*, pp. 323–358.

Pavan, A., I. Segal, and J. Toikka (2013): "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*.

Peters, M. (2001): "Common agency and the revelation principle," *Econometrica*, 69(5), 1349–1372.

Riley, J. G. (1979): "Informational Equilibrium," *Econometrica*, 47(2), 331–359.

Rothschild, M., and J. Stiglitz (1976): "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information," *The Quarterly Journal of Economics*, pp. 629–649.

Thomas, J., and T. Worrall (1990): "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem," *Journal of Economic Theory*, 51(2), 367–390.

Wilson, C. (1977): "A model of insurance markets with incomplete information," *Journal of Economic Theory*, 16(2), 167–207.

Witt, R. C., and J. Urrutia (1983): "Price competition, regulation, and systematic underwriting risk in automobile insurance," *Geneva Papers on Risk and insurance*, 8(29), 403–429.

11 Appendix A - Proofs

Proof of Lemma 1. Choose arbitrary income realization $y^{\theta} = y$ such that

$$\frac{p_{\theta_h}\left(y^h\right)}{p_{\theta_h}\left(y^l\right)} > \frac{p_{\theta_l}\left(y^h\right)}{p_{\theta_l}\left(y^l\right)}.$$

Income realization y^{θ} is relatively more likely to happen for agent θ, therefore it can be used to effectively separate both types of agents. For each t and $\varepsilon > 0$, define $c_t^{\varepsilon}\left(h^t, \theta\right)$ such that

$$u\left(c_t^{\varepsilon}\left(h^t, \theta_h\right)\left(y_h\right)\right) = u\left(c_t\left(h^t, \theta_h\right)\left(y_h\right)\right) + \varepsilon$$
$$u\left(c_t^{\varepsilon}\left(h^t, \theta_l\right)\left(y_l\right)\right) = u\left(c_t\left(h^t, \theta_l\right)\left(y_l\right)\right) + \gamma,$$

for γ so that

$$\varepsilon\frac{p_{\theta_h}\left(y^h\right)}{p_{\theta_h}\left(y^l\right)} > \gamma > \varepsilon\frac{p_{\theta_l}\left(y^h\right)}{p_{\theta_l}\left(y^l\right)},$$

which can be done by setting $\gamma = \frac{\varepsilon}{2}\left(\frac{p_{\theta_h}\left(y^h\right)}{p_{\theta_h}\left(y^l\right)} + \frac{p_{\theta_l}\left(y^h\right)}{p_{\theta_l}\left(y^l\right)}\right)$. This perturbation is well defined for ε sufficiently small. And also notice that, as $\varepsilon \to 0$,

$$\begin{aligned}\max_{h^t,\theta}\left\|u\left(c_t^{\varepsilon}\left(h^t,\theta\right) \mid \theta\right) - u\left(c_t\left(h^t,\theta\right) \mid \theta\right)\right\| &\to 0;\\ \max_{h^t,\theta}\left\|\Pi\left(c_t^{\varepsilon}\left(h^t,\theta\right) \mid \theta\right) - \Pi\left(c_t\left(h^t,\theta\right) \mid \theta\right)\right\| &\to 0.\end{aligned}$$

Notice that by perturbing the original mechanism M at period t to c_t^{ε}, no incentive constraints are changed for $t' \neq t$ (since I have only added a constant to both sides of the constraint). The period t incentive constraint is satisfied due to the construction of $\gamma > 0$. For a fixed $n > 0$ and any $t = 1, \ldots, n$, define $\varepsilon_t(n) > 0$ so that $c_t^{\varepsilon_t}$ is well defined and

$$\max\left\{\begin{array}{c}\max_{h^t,\theta}\left\|\Pi\left(c_t^{\varepsilon_t}\left(h^t,\theta\right) \mid \theta\right) - \Pi\left(c_t\left(h^t,\theta\right) \mid \theta\right)\right\|,\\ \max_{h^t,\theta}\left\|u\left(c_t^{\varepsilon_t}\left(h^t,\theta\right) \mid \theta\right) - u\left(c_t\left(h^t,\theta\right) \mid \theta\right)\right\|\end{array}\right\} < \frac{1}{n}.$$

The sequence defined by $M_n = \left(c_t^{\varepsilon_t(n)}\right)$ satisfies the required limits. ∎

Lemma (2)**.** *For any* $V_l \geq V_l^{FB}$, $\inf\left\{\Pi_h\left(V_l, V\right) \mid \left(V_l, V\right) \in \mathcal{V}^{IC}\right\} < 0$.

Proof. Let, without loss of generality, the image of $u(\cdot)$ be $[0, \overline{v})$, with $\overline{v} \leq \infty$.

For simplicity, I define a mechanism by the instantaneous utility it generates in each possible contingency. Formally, define a u-mechanism as $\upsilon = \left(\upsilon_t\right)_{t=1}^{T}$ where $\upsilon_t : H_{\theta}^t \times$

$H_Y^t \to [0, \overline{v})$. The set of such mechanisms is denoted as Υ. The set of incentive compatible mechanisms is also denoted as Υ^{IC}.

If $T < \infty$, then $\inf\left\{\Pi_h\left(V_L^{FB}, V\right) \mid \left(V_L^{FB}, V\right) \in U^{IC}\right\} < 0$.

Part A) $\left(V_l^{FB}, V_h\right) \notin V^{IC}$:

In this case, it follows that $\inf\left\{\Pi_h\left(v_L^{FB}, v\right) \mid \left(v_L^{FB}, v\right) \in U^{IC}\right\} < 0$.

Assume not, i.e., $\inf\left\{\Pi_h\left(v_L^{FB}, v\right) \mid \left(v_L^{FB}, v\right) \in U^{IC}\right\} > 0$. In this case I consider the sequence $V^n \nearrow V_h$, and M^n to be the solution to $\Pi_h\left(V_l^{FB}, V^n\right)$. Then $M^n \in \left\{M \in \mathcal{M} \mid \Pi_1^M\left(\theta_h\right) \geq 0\right\}$, which is a compact set. Let $\overline{M} = \lim_n M^n$. $\overline{M}$ is feasible and $\left(U_1^{\overline{M}}\left(\theta_l\right), U_1^{\overline{M}}\left(\theta_h\right)\right) = \left(V_l^{FB}, V_h\right)$.

Part B) $\left(V_l^{FB}, V_h\right) \in V^{IC}$:

Consider $T < \infty$. The mechanism that solves $\Pi_h\left(V_l^{FB}, V_h\right)$ solves the following problem:

$$\max_{v \in \Upsilon} \mathbb{E}\left[\sum_{t=1}^{T} \delta^t v_t\left(h_t\right) \mid \theta_1 = \theta_h\right]$$

subject to incentive constraints and

$$\mathbb{E}\left[\sum_{t=1}^{T} \delta^t v_t\left(h_t\right) \mid \theta_1 = \theta_l\right] \leq V_l.$$

The necessary conditions for optimality are: $\exists \left(\lambda_t\right)_{t=0}^{T}$ with $\lambda_t : \Theta^{t-1} \times Y^{t-1} \to \mathbb{R}_+$ such that for any $t = 0, \ldots, T$ and $h_{t-1} = \left(\theta_h^{t-1}, h_y^{t-1}\right) \in H_{t-1}$,

$$\begin{aligned}\mathbb{P}\left(\left([h_\theta]_2^{t-1}, \theta_h\right), \left([h_Y]_{t=1}^{t-1}, y\right) \mid \theta_h\right) - \sum_{s=0}^{t-1} \lambda_s\left([h_\theta]_2^{s-1}, [h_y]_1^{s-1}\right) \mathbb{P}\left([h_\theta]_{s+1}^{t-1}, \theta_h, [h_y]_s^t, y \mid \theta_s = \theta_l\right) &\leq 0, \\ -\lambda_t\left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l\left(y\right)&\end{aligned}$$

and equal to zero if $v_t\left(h_{t-1}, \theta_h, y\right) > 0$; and also

$$\begin{aligned}\mathbb{P}\left(\left([h_\theta]_2^{t-1}, \theta_l\right), \left([h_y]_{t=1}^{t-1}, y\right) \mid \theta_h\right) - \sum_{s=0}^{t-1} \lambda_s\left([h_\theta]_2^{s}, [h_y]_1^{s-1}\right) \mathbb{P}\left([h_\theta]_{s+1}^{t-1}, \theta_l, [h_y]_s^t, y \mid \theta_s = \theta_l\right) &\leq 0 \\ +\lambda_t\left([h_\theta]_2^{t-1}, \theta_l, [h_y]_1^{t-1}\right) p_l\left(y\right)&\end{aligned}$$

and equal to zero if $v_t\left(h_{t-1}, \theta_l, y\right) > 0$.

B.1) $v_t\left(\theta_h^t, h_Y^T\right) > 0$ implies $t = T$ and $h_y^T = \left(y_h, \ldots, y_h\right)$.

Suppose that, for some $t < T$and $h_{t-1} = \left(h_\theta^{t-1}, h_y^{t-1}\right) \in H_{t-1}$ and $y \in Y$:

$$\begin{aligned}\mathbb{P}\left(\left([h_\theta]_2^{t-1}, \theta_h\right), \left([h_Y]_{t=1}^{t-1}, y\right) \mid \theta_h\right) - \sum_{s=0}^{t-1} \lambda_s \left([h_\theta]_2^{s-1}, [h_y]_1^{s-1}\right) \mathbb{P}\left([h_\theta]_{s+1}^{t-1}, \theta_h, [h_y]_s^t, y \mid \theta_s = \theta_l\right) &= 0, \\ -\lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y)&\end{aligned}$$

then the first order conditions at $h_{t+1} = (h_t, \theta_h, y_h)$ and $h'_{t+1} = (h_t, \theta_l, y_l)$ imply that

$$\begin{aligned}K \frac{\pi_{hh} p_h(y_h)}{p_l(y_h)} - \lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y) \pi_{lh} \frac{p_h(y_h)}{p_l(y_h)} &\leq \lambda_{t+1}\left([h_\theta]_2^t, [h_y]_1^t\right) \\ &\leq \lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y) \pi_{ll} - K\pi_{hl},\end{aligned}$$

in which

$$K = \mathbb{P}\left(\left([h_\theta]_2^{t-1}, \theta_h\right), \left([h_Y]_{t=1}^{t-1}, y\right) \mid \theta_h\right) - \sum_{s=0}^{t-1} \lambda_s \left([h_\theta]_2^{s-1}, [h_y]_1^{s-1}\right) \mathbb{P}\left([h_\theta]_{s+1}^{t-1}, \theta_h, [h_y]_s^t, y \mid \theta_s = \theta_l\right).$$

So:

$$K\left[\pi_{hh} \frac{p_h(y_h)}{p_l(y_h)} + \pi_{hl}\right] - \lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y) \left[\pi_{lh} \frac{p_h(y_h)}{p_l(y_h)} - \pi_{ll}\right] \leq 0,$$

which is impossible since $K - \lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y) = 0$ and $\lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) > 0$.

However, since the first order condition for $\left(\theta_h^t, \theta_l, h_Y^{t+1}\right)$ always holds with an equality (since consumption is always positive), I have that

$$\lambda_{t+1}\left([h_\theta]_2^t, [h_y]_1^t\right) = -\pi_{hl} \begin{bmatrix} \lambda_t \left([h_\theta]_2^{t-1}, [h_y]_1^{t-1}\right) p_l(y) \pi_{ll} \\ \mathbb{P}\left(\left([h_\theta]_2^{t-1}, \theta_h\right), \left([h_y]_{t=1}^{t-1}, y\right) \mid \theta_h\right) \\ -\sum_{s=0}^{t-1} \lambda_s \left([h_\theta]_2^{s-1}, [h_y]_1^{s-1}\right) \mathbb{P}\left([h_\theta]_{s+1}^{t-1}, \theta_h, [h_y]_s^t, y \mid \theta_s = \theta_l\right) \end{bmatrix}.$$

Therefore, it follows that $\lambda_t \left(\theta_h^{t-2}, [h_y]_1^{t-1}\right)$ does not depend on $[h_y]_1^{t-1}$.

Therefore, first order condition at node $\left(\theta_h^T, h_Y^T\right)$, with $h_y^T = (y_1, \ldots, y_T)$ is given by

$$\begin{aligned}(\pi_{hh})^{T-1}\prod_{t=1}^{T} p_h(y_t) - \sum_{s=1}^{T-1}\lambda_s\left(\theta_h^{t-2}\right)\pi_{lh}(\pi_{hh})^{T-s+1} p_l(y_s)\prod_{s'>s}^{T} p_h(y_{s'}) \\ \lambda_{T-1}\left(\theta_h^{T-3}\right)\pi_{ll}p_l(y_{T-1})\,p_l(y_T) \quad \leq 0. \\ -\lambda_T\left(\theta_h^{T-2}\right)p_l(y_T)\end{aligned}$$

This is summation is maximal at $h_y^T = (y_h, \ldots y_h)$.

Now consider the case $T = \infty$ and consider the mechanism that solves $\Pi_h\left(V_l^{FB}, V_h\right)$. The truncation of this mechanism to $T' < \infty$ periods satisfies the same optimality conditions above. This implies that, for any $t \geq 1$, $c_t\left(\theta_h^t, h_y^t\right) = 0$, which implies $V_h \leq V_l^{FB}$, contradicting the definition of V_h.

B.2) $\Pi_h\left(V_l^{FB}, V_h\right) < 0$.

From now on I focus on the case $T < \infty$.

At all nodes following $\left(\theta_h^{t-1}, \theta_l, y_h^{t-1}\right)$, the agent receives consumption κ_t such that

$$\frac{\left(1-\delta^{T-t+1}\right)}{(1-\delta)}u(\kappa_t) = \mathbb{P}(y_t = \ldots = y_T \mid \theta_t = \theta_l)\,u\left(c_T\left(\theta_h^T, y_h^T\right)\right)\delta^{T-t}.$$

Therefore all binding incentive constraints imply that

$$u\left(c_T\left(\theta_h^T, y_h^T\right)\right) = \frac{V_l}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right)\delta^{T-1}},$$

and (using $\psi = u^{-1}$)

$$k_t = \psi\left[\frac{\mathbb{P}(y_t = \ldots = y_T \mid \theta_t = \theta_l)(1-\delta)\,\delta^{T-t}}{1-\delta^{T-t+1}}\frac{V_l}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right)\delta^{T-1}}\right].$$

But notice that

$$\begin{aligned}\frac{V_l}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right)\delta^{T-1}} &= \left[\begin{array}{c}\frac{V_h^{FB}}{\mathbb{P}\left(y_h^T|\theta_1=\theta_l\right)\delta^{T-1}} \\ +\frac{V_l}{\mathbb{P}\left(y_h^T|\theta_1=\theta_h\right)}\left(\frac{\mathbb{P}\left(y_h^T|\theta_1=\theta_h\right)}{\mathbb{P}\left(y_h^T|\theta_1=\theta_l\right)} - \frac{V_h^{FB}}{V_l}\right)\end{array}\right] \\ &\geq \frac{V_h^{FB}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right)\delta^{T-1}},\end{aligned}$$

because

$$\frac{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_h\right)}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right)} \geq \frac{V_h^{FB}}{V_l^{FB}} \geq \frac{V_h^{FB}}{V_l}.$$

Therefore the total profits from this allocation are

$$\begin{aligned}
&\frac{\left(1-\delta^T\right)}{(1-\delta)} x_h^{FB} - \sum_{t=2}^{T} \delta^{t-1} \mathbb{P}\left(\left(\theta_h^{t-1}, \theta_l\right), y_h^{t-1} \mid \theta_1 = \theta_h\right) \left[\frac{\left(1-\delta^{T-t+1}\right)}{(1-\delta)} k_t\right] \\
&\qquad - \delta^{T-1} \mathbb{P}\left(\theta_h^T, y_h^T \mid \theta_1 = \theta_h\right) \delta^{T-1} \psi\left(\frac{V_l}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right) \delta^{T-1}}\right) \leq \\
&\frac{\left(1-\delta^T\right)}{(1-\delta)} x_h^{FB} - \sum_{t=2}^{T} \delta^{t-1} \mathbb{P}\left(\left(\theta_h^{t-1}, \theta_l\right), y_h^{t-1} \mid \theta_1 = \theta_h\right) \left[\frac{\left(1-\delta^{T-t+1}\right)}{(1-\delta)} \hat{k}_t\right] \\
&\qquad - \delta^{T-1} \mathbb{P}\left(\theta_h^T, y_h^T \mid \theta_1 = \theta_h\right) \delta^{T-1} \psi\left(\frac{V_h^{FB}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_h\right) \delta^{T-1}}\right) \equiv K,
\end{aligned}$$

where $\hat{k}_t$ is defined by (notice that $\hat{k}_t < k_t$):

$$\frac{\left(1-\delta^{T-t+1}\right)}{(1-\delta)} u\left(\hat{k}_t\right) = \mathbb{P}\left(y_t = \ldots = y_T \mid \theta_t = \theta_l\right) \frac{V_h^{FB}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right) \delta^{T-1}} \delta^{T-t}.$$

$$\frac{\left(1-\delta^{T-t+1}\right)}{(1-\delta)} u\left(\hat{k}_t\right) = \frac{(1-\delta)\,\mathbb{P}\left(y_t = \ldots = y_T \mid \theta_t = \theta_l\right)}{\left(1-\delta^{T-t+1}\right)} \frac{V_h^{FB}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right) \delta^{T-1}} \delta^{T-t}.$$

However, convexity of ψ implies $K < 0$, since

$$\begin{aligned}
u\left(x_h^{FB}\right) - \sum_{t=2}^{T} \delta^{t-1} \mathbb{P}\left(\left(\theta_h^{t-1}, \theta_l\right), y_h^{t-1} \mid \theta_1 = \theta_h\right) \left[\mathbb{P}\left(y_t = \ldots = y_T \mid \theta_t = \theta_l\right) \frac{V_h^{FB} \delta^{T-t}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_l\right) \delta^{T-1}}\right] \\
- \delta^{T-1} \mathbb{P}\left(\theta_h^T, y_h^T \mid \theta_1 = \theta_h\right) \delta^{T-1} \frac{V_h^{FB}}{\mathbb{P}\left(y_h^T \mid \theta_1 = \theta_h\right) \delta^{T-1}} = 0.
\end{aligned}$$

■

Proof of Proposition 5. Just notice that continuation utility is given by

$$U^M\left(h^t, \theta' \mid \theta\right) = U\left(c_t\left(h^t, \theta'\right) \mid \theta\right) + \sum_{\tau} \sum_{\theta} \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta\right) U\left(c_{t+\tau}\left(h^t, \theta', (\theta_h)^\tau\right) \mid \theta\right).$$

Hence, form the binding upward incentive constraint, I have that

$$\begin{aligned}
&U^M\left(h^t,\theta_h \mid \theta_l\right) - U^M\left(h^t,\theta_l \mid \theta_l\right) = \\
&\quad U\left(c_t\left(h^t,\theta_h\right) \mid \theta_l\right) + \sum_\tau \sum_\theta \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_l\right) U\left(x_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right) \mid \theta\right) \\
&\quad - U\left(c_t\left(h^t,\theta_l\right) \mid \theta_l\right) - \sum_\tau \sum_\theta \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_l\right) U\left(x_{t+\tau}\left(h^t,\theta_l,(\theta_h)^\tau\right) \mid \theta\right).
\end{aligned}$$

Using this, I can rewrite the downward incentive constraint as:

$$\begin{aligned}
&U^M\left(h^t,\theta_h \mid \theta_h\right) - U^M\left(h^t,\theta_l \mid \theta_h\right) = \\
&\quad U\left(c_t\left(h^t,\theta_h\right) \mid \theta_h\right) + \sum_\tau \sum_\theta \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_h\right) U\left(x_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right) \mid \theta\right) \\
&\quad - U\left(c_t\left(h^t,\theta_l\right) \mid \theta_h\right) - \sum_\tau \sum_\theta \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_h\right) U\left(x_{t+\tau}\left(h^t,\theta_l,(\theta_h)^\tau\right) \mid \theta\right) \\
&\quad = U^M\left(h^t,\theta_h \mid \theta_l\right) - U^M\left(h^t,\theta_l \mid \theta_l\right) \\
&\quad + U\left(c_t\left(h^t,\theta_h\right) \mid \theta_h\right) - U\left(c_t\left(h^t,\theta_l\right) \mid \theta_h\right) - \left[U\left(c_t\left(h^t,\theta_h\right) \mid \theta_l\right) - U\left(c_t\left(h^t,\theta_l\right) \mid \theta_l\right)\right] \\
&\quad \sum_\tau \sum_\theta \left[\mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_h\right) - \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_l\right)\right] U\left(c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right) \mid \theta\right) \\
&\quad - \sum_\tau \sum_\theta \left[\mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_h\right) - \mathbb{P}\left(\theta_{t+\tau} = \theta \mid \theta_t = \theta_l\right)\right] U\left(c_{t+\tau}\left(h^t,\theta_l,(\theta_h)^\tau\right) \mid \theta\right) \\
&\quad = U^M\left(h^t,\theta_h \mid \theta_l\right) - U^M\left(h^t,\theta_l \mid \theta_l\right) \\
&\quad + \sum_{\tau \geq 0} \Delta_{t,\tau}\left[U\left(c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right) \mid \theta_h\right) - U\left(c_{t+\tau}\left(h^t,\theta_h,(\theta_h)^\tau\right) \mid \theta_l\right)\right] \\
&\quad - \sum_{\tau \geq 0} \Delta_{t,\tau}\left[U\left(c_{t+\tau}\left(h^t,\theta_l,(\theta_h)^\tau\right) \mid \theta_h\right) - U\left(c_{t+\tau}\left(h^t,\theta_l,(\theta_h)^\tau\right) \mid \theta_l\right)\right],
\end{aligned}$$

where $\Delta_{t,\tau} \equiv \mathbb{P}\left(\theta_{t+\tau} = \theta_h \mid \theta_t = \theta_h\right) - \mathbb{P}\left(\theta_{t+\tau} = \theta_l \mid \theta_t = \theta_l\right)$. Therefore I know that

$$U^M\left(h^t,\theta_h \mid \theta_h\right) - U^M\left(h^t,\theta_l \mid \theta_h\right) \geq 0.$$

■

Path of Prices

For any $(v,\Delta) \in A$, $\chi(v,\Delta)$ is the solution to the problem of delivering instantaneous utility v to an agent of type θ_h, while reducing the attractiveness of the allocation by Δ from the

point of view of an agent of type θ_l. In the appendix, for simplicity I write the problem in terms of utility levels, instead of consumption levels. Therefore it is useful to define $\psi : u(\mathbb{R}) \to \mathbb{R}_+$ as the inverse function of $u(\cdot)$, i.e., $\psi(u(c)) = c$. Formally, the problem is defined as:

$$\chi(v, \Delta) = \inf_{z:Y \to u(\mathbb{R}_+)} \sum_y p_h(y) \psi[z(y)],$$

subject to

$$\sum_y p_h(y) z(y) = v,$$

and

$$\sum_y p_l(y) z(y) = v - \Delta.$$

Since the constraints in the problem are linear, the choice set is convex and and objective function is convex it follows that χ is convex.

The following statement deals with the existence and uniqueness of solution and differentiability.

Proposition 13. *$\chi(\cdot)$ has a unique solution. If $\psi'' > 0$ and the solution at $(v, \Delta) \in A$, with $\Delta > 0$, is in the interior of $[u(\mathbb{R}_+)]^Y$, then it is twice continuously differentiable at an open neighborhood of (v, Δ), sign $\left(\frac{\partial \chi(v,\Delta)}{\partial \Delta}\right) = sign(\Delta)$. Additionally, sign $\left(\frac{\partial^2 \chi}{\partial v \partial \Delta}\right) > 0$ (sign $\left(\frac{\partial^2 \chi}{\partial v \partial \Delta}\right) > 0$) if $\psi''(\cdot)$ is increasing (decreasing).*

Proof. Existence of solution follows from the fact that

$$\left\{ z \in [u(\mathbb{R}_+)]^Y \mid \sum_y p_h(y) \psi[z(y)] \leq K \right\}$$

is compact, for any $K \in \mathbb{R}_+$. Uniqueness follows from the convexity of ψ.

Necessary and sufficient conditions for $z^{w,\Delta}$ interior to be a solution is: $\exists \lambda, \mu \in \mathbb{R}$ such that

$$\psi'(z(y)) + \lambda + \mu \frac{p_l(y)}{p_h(y)} = 0,$$

for all $y \in Y$ and the required constraints are satisfied at z. Therefore, if χ has a solution at (z, Δ), then (by application of the implicit function theorem to the equalities required for optimality), it has a solution (including multipliers) on an open neighborhood $N \subseteq \mathbb{R}$,

denoted as $\left(u^{v,\Delta}, \lambda^{v,\Delta}, \mu^{v,\Delta}\right)$ that are continuously differentiable. Therefore χ is continuously differentiable at N, additionally its derivative is given by

$$\left[\begin{array}{c} \frac{\partial}{\partial v}\chi\left(v,\Delta\right) \\ \frac{\partial}{\partial \Delta}\chi\left(v,\Delta\right) \end{array}\right] = \left[\begin{array}{c} -\left(\mu^{v,\Delta}+\lambda^{v,\Delta}\right) \\ \mu^{v,\Delta} \end{array}\right].$$

Finally, continuous differentiability of $\left(\lambda^{v,\Delta}, \mu^{v,\Delta}\right)$ implies that $\chi\left(\cdot\right)$ is twice continuously differentiable in N. Finally, differentiating the equalities defining $\left(u^{v,\Delta}, \lambda^{v,\Delta}, \mu^{v,\Delta}\right)$ leads to

$$\frac{\partial^2 \chi\left(v,\Delta\right)}{\partial v \partial \Delta} = \frac{\sum_y \frac{p_l(y)}{\psi''(u(y))} - \sum_y \frac{p_h(y)}{\psi''(u(y))}}{\left(\sum_y \frac{p_h(y)}{\psi''(u(y))}\right)\left(\sum_y \frac{[p_l(y)]^2}{p_h(y)\psi''(u(y))}\right) + \left(\sum_y \frac{p_l(y)}{\psi''(u(y))}\right)^2}.$$

Since $sign\left(u^{v,\Delta}\left(y\right) - u^{v,\Delta}\left(y'\right)\right) = sign\left(\frac{p_h(y)}{p_l(y)} - \frac{p_h(y')}{p_l(y')}\right)$, it follows that $\frac{\partial^2 \chi(v,\Delta)}{\partial v \partial \Delta} > 0$ is non-negative as long as $\psi''\left(z\right)$ is increasing. ∎

The characterization of $\chi\left(\cdot\right)$ is crucial to understanding the structure of distortions in efficient mechanisms. The following statement connects variation in marginal cost of utility and distortions to the levels of utility and distortions and will be instrumental in the proof of Proposition 5.

Lemma 6. *Suppose $\chi_v\left(\cdot\right)$ is increasing in Δ, then*

$$\left\{\begin{array}{c} \chi_v\left(v,\Delta\right) \geq \chi_v\left(v',\Delta'\right), \\ \chi_\Delta\left(v,\Delta\right) \leq \chi_\Delta\left(v',\Delta'\right) \end{array}\right\} \Rightarrow \left\{\begin{array}{c} v \geq v', \\ \Delta \geq \Delta' \end{array}\right\}.$$

Proof. Since χ is twice continuously differentiable, I have that

$$\chi_v\left(v,\Delta\right) - \chi_v\left(v',\Delta'\right) = \left(w-w'\right)\int_0^1 \chi_{vv}\left(\iota\left(\alpha\right)\right)d\alpha + \left(\Delta-\Delta'\right)\int_0^1 \chi_{v\Delta}\left(\iota\left(\alpha\right)\right)d\alpha \geq 0,$$

$$\chi_\Delta\left(v,\Delta\right) - \chi_\Delta\left(v',\Delta'\right) = \left(w-w'\right)\int_0^1 \chi_{v\Delta}\left(\iota\left(\alpha\right)\right)d\alpha + \left(\Delta-\Delta'\right)\int_0^1 \chi_{\Delta\Delta}\left(\iota\left(\alpha\right)\right)d\alpha \leq 0,$$

where $\iota\left(\alpha\right) = \alpha\left(w,\Delta\right) + \left(1-\alpha\right)\left(w',\Delta'\right)$, for $\alpha \in [0,1]$.

This implies that

$$\left(\Delta-\Delta'\right)\left[\frac{\int_0^1 \chi_{v\Delta}\left(\iota\left(\alpha\right)\right)d\alpha}{\int_0^1 \chi_{vv}\left(\iota\left(\alpha\right)\right)d\alpha} - \frac{\int_0^1 \chi_{\Delta\Delta}\left(\iota\left(\alpha\right)\right)d\alpha}{\int_0^1 \chi_{v\Delta}\left(\iota\left(\alpha\right)\right)d\alpha}\right] \geq 0.$$

However, convexity of $\chi(\cdot)$ second order continuous differentiability, implies that the function Γ defined over a neighborhood of $(0,0)$, given by $(v_0,\Delta_0)\mapsto\int_0^1\chi\left(f(\alpha)+(v_0,\Delta_0)\right)d\alpha$ is also convex and twice continuously differentiable. Convexity implies that

$$\left|\Gamma''(0,0)\right|=\left(\int_0^1\chi_{vv}\left(\iota(\alpha)\right)d\alpha\right)\left(\int_0^1\chi_{\Delta\Delta}\left(\iota(\alpha)\right)d\alpha\right)-\left(\int_0^1\chi_{v\Delta}\left(\iota(\alpha)\right)d\alpha\right)^2>0,$$

therefore $\Delta<\Delta'$. ∎

Now I present the proof of Proposition of 5.

Proof of Proposition 5. Case $T<\infty$. The following necessary conditions for optimality, which I will use extensively in this proof, are: for all $t\geq 1$,

$$\left[w_t\left(\theta_h^t\right)\right]:\quad\left[\begin{array}{c}\lambda-\chi_w\left(w_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\\-\chi_\Delta\left(w_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\\-\sum_{\tau<t}\frac{\pi_{lh}}{\pi_{hh}}\chi_\Delta\left(w_\tau\left(\theta_h^\tau\right),\Delta_\tau\left(\theta_h^\tau\right)\right)\end{array}\right]=0, \tag{6}$$

$$\left[w_t\left(\theta_h^{t-1},\theta_l\right)\right]:\quad\left[\begin{array}{c}\lambda-\chi_w\left(w_t\left(\theta_h^{t-1},\theta_l\right),0\right)\\+\frac{\pi_{hh}}{\pi_{hl}}\chi_\Delta\left(w_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\\-\chi_\Delta\left(w_{t-1}\left(\theta_h^{t-1}\right),\Delta_{t-1}\left(\theta_h^{t-1}\right)\right)\frac{\pi_{ll}}{\pi_{hl}}-\sum_{\tau<t-1}\frac{\pi_{lh}}{\pi_{hh}}\chi_\Delta\left(w_\tau\left(\theta_h^\tau\right),\Delta_\tau\left(\theta_h^\tau\right)\right)\end{array}\right]=0.$$

If $\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\geq\chi_v\left(v_t\left(\theta_h^{t-1},\theta_l\right),0\right)$, then from

$$\chi_v\left(v_{t-1}\left(\theta_h^{t-1}\right),\Delta_{t-1}\left(\theta_h^{t-1}\right)\right)=\pi_{hh}\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)+\pi_{hl}\chi_v\left(v_t\left(\theta_h^{t-1},\theta_l\right),0\right),$$

I have that

$$\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\geq\chi_v\left(v_{t-1}\left(\theta_h^{t-1}\right),\Delta_{t-1}\left(\theta_h^{t-1}\right)\right),$$

and

$$\begin{aligned}\chi_v\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)+\chi_\Delta\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)\leq\\ \chi_v\left(v_{t+1}\left(\theta_h^{t+1}\right),\Delta_{t+1}\left(\theta_h^{t+1}\right)\right)+\chi_\Delta\left(v_{t+1}\left(\theta_h^{t+1}\right),\Delta_t\left(\theta_h^{t+1}\right)\right).\end{aligned}$$

Therefore, from Lemma 6 I know that $v_t\left(\theta_h^t\right)<v_{t+1}\left(\theta_h^{t+1}\right)$ and $\Delta_t\left(\theta_h^t\right)>\Delta_{t+1}\left(\theta_h^{t+1}\right)$.

Now, suppose that for all $s\geq t+1$, I have that $\chi_s\left(v_s\left(\theta_h^s\right),\Delta_s\left(\theta_h^s\right)\right)\geq\chi_v\left(v_s\left(\theta_h^{s-1},\theta_l\right),0\right)$ (and hence $\left\{\Delta_s\left(\theta_h^s\right)\right\}_s$ decreasing) and $\chi_t\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right)<\chi_v\left(v_t\left(\theta_h^{t-1},\theta_l\right),0\right)$. This immediately implies that $v_t\left(\theta_h^{t-1},\theta_l\right)>v_t\left(\theta_h^t\right)$. Also, notice that $v_t\left(\theta_h^{t-1},\theta_l\right)>v_{t+1}\left(\theta_h^t,\theta_l\right)$

since

$$\chi_v\left(v_t\left(\theta_h^{t-1},\theta_l\right),0\right) > \chi_t\left(v_t\left(\theta_h^t\right),\Delta_t\left(\theta_h^t\right)\right) \geq \chi_{t+1}\left(v_{t+1}\left(\theta_h^t,\theta_l\right),0\right).$$

Then I have that

$$\begin{aligned}
\Delta_t\left(\theta_h^t\right) = \sum_{s=t}^{T}\delta^{s-t}v_s\left(\theta_h^s\right) - \sum_{s=t+1}^{T}\delta^{s-t}\mathbb{P}\left(\theta_s=\theta_l \mid \theta_t=\theta_l\right)\Delta_s \\
- \sum_{s=t}^{T}\delta^{s-t}v_t\left(\theta_h^{t-1},\theta_l\right) \\
\leq \sum_{s=t+1}^{T}\delta^{s-t}v_s\left(\theta_h^s\right) - \sum_{s=t+1}^{T}\delta^{s-t}\mathbb{P}\left(\theta_s=\theta_l \mid \theta_t=\theta_l\right)\Delta_s \\
- \sum_{s=t+1}^{T}\delta^{s-t}v_t\left(\theta_h^{t-1},\theta_l\right) \\
\leq \sum_{s=t+1}^{T}\delta^{s-t}v_s\left(\theta_h^s\right) - \sum_{s=t+1}^{T}\delta^{s-t}\mathbb{P}\left(\theta_s=\theta_l \mid \theta_t=\theta_l\right)\Delta_s \\
- \sum_{s=t+1}^{T}\delta^{s-t}v_{t+1}\left(\theta_h^t,\theta_l\right).
\end{aligned}$$

However, using the definition of $\Delta_{t+1}\left(\theta_h^{t+1}\right)$ I have that

$$\begin{aligned}
\Delta_t\left(\theta_h^t\right) &\leq \Delta_{t+1}\left(\theta_h^{t+1}\right) + \delta^2\left(1-\pi_{ll}\right)\left(\pi_{ll}-\pi_{hl}\right)\sum_{s\geq t+2}\delta^{s-t-2}\left(\pi_{ll}-\pi_{hl}\right)^{s-t-2}\Delta_s \\
&\quad - \left(1-\delta\left(1-\pi_{ll}\right)\right)\Delta_{t+1} \\
&\leq \Delta_{t+1}\left(\theta_h^{t+1}\right) + \frac{\delta^2\left(1-\pi_{ll}\right)\left(\pi_{ll}-\pi_{hl}\right)}{1-\delta\left(\pi_{ll}-\pi_{hl}\right)}\Delta_{t+1} - \left(1-\delta\left(1-\pi_{ll}\right)\right)\Delta_{t+1} \\
&\leq \Delta_{t+1}\left(\theta_h^{t+1}\right).
\end{aligned}$$

This contradicts $\chi_{t+1}\left(v_{t+1}\left(\theta_h^{t+1}\right),\Delta_{t+1}\left(\theta_h^{t+1}\right)\right) \geq \chi_v\left(v_{t+1}\left(\theta_h^t,\theta_l\right),0\right)$.

Case $T=\infty$. Define as $\Pi^K(V)$ the value obtained from the optimal contract, with the extra restriction that the agent consumes zero beyond period $K+1$, and equal to $-\infty$ if V is not feasible with such contracts.

Fix $K\in\mathbb{N}$. Consider mechanism $M^\varepsilon = \left(c_t^\varepsilon\right)_t \in \underline{\mathcal{M}}\cap\mathcal{M}^{IC_0}$ such that $\left\|M^\varepsilon - M(V)\right\|_\infty < \varepsilon$

and $\left(U_1^{M^\varepsilon}(\theta_l), U_1^{M^\varepsilon}(\theta_h)\right) = V$. Finally define mechanism $\tilde{M}^K = \left(\widetilde{c}_t^K\right)_t$ defined by

$$u\left(\tilde{c}_t^K\right) = \left(\frac{1+\delta}{2}\right)^K u\left(c_t^\varepsilon\right) + \left[1 - \left(\frac{1+\delta}{2}\right)^K\right] u\left(c_t\right).$$

Consider mechanism $\overline{M}^K$ to be the mechanism that is identical to $\tilde{M}$ up to period K, and has allocation rule $c_t\left(h_\theta^t\right) = 0$, for $t \geq K+1$.

Consider σ as an optimal pure reporting strategy, for an agent facing mechanism $\overline{M}^K$ and define M^K by

$$c_t^K\left(h_\theta^t\right) = c_t\left(\sigma\left(h_\theta^t\right)\right).$$

Let V^K be the interim utility obtained by both types in mechanism M^K (notice that $V^k \to V$). It follows that M^K is feasible in $\Pi^K(V_K)$. And let $V^K = \left(V_l^K, V_h^K\right)$ be the expected utility obtained by both initial types θ_l and θ_h, respectively. Boundedness of the the utility function implies that $M^K \to M(V)$ in the product topology. This occurs because the gains from any strategy following misreports, from period t and history h_θ^t on, is negative for large K. Let $\overline{v}_0, \overline{v}'$ be instantaneous utility from truth-telling and following a misreport strategy at $M(V)$ and v_0, v' be instantaneous utility from truth-telling and following a misreport strategy at M^K similarly let v_{0+} and v'_+ denote continuation utility of both strategies under M^K and $\overline{v}_0, \overline{v}'$ denote the continuation utility from both strategies on $M(V)$. It follows that

$$v_0 - v' \geq \overline{v}_0 - \overline{v}' + \left(\frac{1+\delta}{2}\right)^K,$$

$$v'_+ - v_{0+} - \left(\overline{v}'_+ - \overline{v}_{0+}\right) \leq \delta^K \frac{\overline{u}}{1-\delta}.$$

Which means that

$$v_0 - v' + v_{0+} - v'_+ \geq \overline{v}_0 + \overline{v}_{0+} - \overline{v}' - \overline{v}'_+ + \left(\frac{1+\delta}{2}\right)^K - \delta^K \frac{\overline{u}}{1-\delta},$$

which is positive for K sufficiently large.

Since the allocation in $M(V)$ is bounded, it follows that

$$\Pi_0^{M^K} \nearrow_{k\to\infty} \Pi^{M(V)},$$

and it also follows that

$$\Pi^*(V_k) \geq \Pi^K\left(V^k\right) \geq \Pi_0^{M^K}.$$

Taking limits and using continuity of Π^*, we have that

$$\lim_{k\to\infty} \Pi^K\left(V^K\right) = \Pi^*\left(V\right).$$

Hence it follows that $M\left(V\right)$ equals the limit of the optimal mechanisms in the problem $\Pi^K\left(V^k\right)$, each satisfying the properties characterized for the finite horizon case. ∎

Lemma 7. *Function $\psi''\left(\cdot\right)$ is increasing if, and only if,*

$$r'\left(x\right)u'\left(x\right) + 2\left[r\left(x\right)\right]^2 \geq 0.$$

The second order derivative of $\psi = u^{-1}$ is given by:

$$\psi''\left(z\right) = \left(-\frac{u''\left(\psi\left(z\right)\right)}{u'\left(\psi\left(z\right)\right)}\right)\frac{1}{u'\left(\psi\left(z\right)\right)^2}.$$

The absolute risk aversion of utility $u\left(\cdot\right)$ at $c \in \mathbb{R}_+$ is given by

$$r\left(c\right) \equiv -\frac{u''\left(c\right)}{u'\left(c\right)}.$$

Therefore direct derivation implies that

$$\psi'''\left(z\right) = \frac{r'\left(x\right)u'\left(x\right) + 2\left[r\left(x\right)\right]^2}{\left[u'\left(x\right)\right]^5},$$

where $z = u\left(x\right)$.

Proposition 9 follows directly from Lemma 7 and Proposition 13.

12 Appendix B - General mechanisms

The goal of this section is to extend the characterization result in section 6 to revenue maximizing mechanisms within the lager choice set $\mathcal{M}$. The main structure theorem states that the crucial statistic used to screen agents is the number of consecutive periods before the (initially) high type agent has a first low shock ($\theta_t = \theta_l$). I now consider contracts that can also use income realization history y^t to screen agents, however these are used in a limited way: consumption only depends on the history of income realizations until the first low shock.

I am interested in characterizing the solution to the following problem, for $V_h \geq V_l^{FB}$,

$$\begin{aligned} &\max_{M \in \mathcal{M}^{IC}} \Pi^M(\theta_h) \\ &\text{s.t. } c_t\left(h^t\right) = c_l^{FB}, \text{ for all } h^t \in \{\tau = 0\}; \\ &U^M(\theta_h) \geq V_h. \end{aligned}$$

The solution to this problem is defined as an optimal mechanism.

Below I show that the "distortion at the top" result from the static model generalizes to the dynamic model in a very strong form, the allocation is efficient except at history nodes that only involve high shocks. Additionally, the allocation only depends on how many periods of consecutive high types the agent has announced, and the realized income history within such periods.

Proposition 14. *(Distortions in general mechanisms)*
If $M = (c_t)_t$ is an optimal mechanism, then the function $\left(t, h^T\right) \mapsto c_t\left(h^T\right)$ is measurable with respect to $\left(\mathbf{1}\{\tau < t\}, h_Y^{\tau \wedge t}\right)$ and

$$c_t\left(h^T\right) \in C^* \iff t > \tau\left(h_\theta^T\right).$$

Relaxed Problem

As in the case of realization-independent mechanisms, I consider the relaxed problem that only takes into some of the relevant incentive constraints. I define $\mathcal{M}^{IC*}$ as the set of mechanisms that satisfy all the incentive constraints for nodes with consecutive high-type histories, i.e., $h_\theta^{t-1} = (\theta_h)^{t-1}$. Formally, this means that

$$\mathcal{M}^{IC*} = \left\{ M \in \mathcal{M} \mid \begin{array}{c} \left[\left(\theta_h^\tau, h_y^\tau\right), \theta_l\right] \text{-IC is satisfied,} \\ \text{for any } \tau \geq 1, h_y^\tau \in H_y^\tau \end{array} \right\}.$$

The relaxed problem is

$$\begin{aligned} &\max_{M \in \mathcal{M}^{IC*}} \Pi^M(\theta_h) \\ &\text{s.t. } c_t\left(h^t\right) = c_l^{FB}, \text{ for all } h^t \in \mathcal{H}_0 \\ &U^M(\theta_h) \geq V_h. \end{aligned}$$

Proposition 15. *(Distortions in the relaxed problem)*

If $M = (c_t)_t$ solves the relaxed problem, then the function $\left(t, h^T\right) \mapsto c_t\left(h^T\right)$ is $(\mathbf{1}\{\tau < t\}, h_Y^{\tau \wedge t})$-measurable and

$$c_t\left(h^T\right) \in C^* \Longleftrightarrow t > \tau\left(h_\theta^T\right).$$

Moreover, all incentive constraints hold as equalities.

Proof. A necessary condition for optimality is existence of $(\lambda_t)_{t=1}^T, \mu, \kappa \geq 0$ such that: $\lambda_t : H^t \to \mathbb{R}_+$for $h_\theta^t = (\theta_h)^t$ and $h_y^{t-1} = (y^{t-1})$

$$\left[c_t\left(h^{t-1}, \theta_h\right)(y)\right] : \begin{array}{c} -\kappa + \mu u'(\cdot) \\ -\sum_{\tau<t} \lambda_\tau \frac{\mathbb{P}(\theta_h|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^\tau} u'(\cdot) - \lambda_t \frac{p^l(y)}{\mathbb{P}(h^{t+1}|\theta_1=\theta_h)p^h(y)} u'(\cdot) \end{array} \geq 0, \tag{7}$$

and for any h_θ^t following $\left((\theta_h)^s, \theta_l\right)$, satisfies

$$\left[c_t\left(h^{t-1}, \theta\right)(y)\right] : \begin{array}{c} -\kappa + \mu u'(\cdot) \\ -\sum_{\tau<s} \lambda_\tau \frac{\mathbb{P}(\theta_h|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^{\tau-1}} u'(\cdot) - \sum_{\tau=s} \lambda_\tau \frac{\mathbb{P}(\theta_l|\theta_l)}{\mathbb{P}(\theta_h|\theta_h)^{s-1}\mathbb{P}(\theta_l|\theta_h)} u'(\cdot) \\ +\lambda_{s+1} \frac{1}{\mathbb{P}(\theta_h|\theta_h)^{s-1}\mathbb{P}(\theta_l|\theta_h)} u'(\cdot) \end{array} \geq 0.$$

Finally, suppose that for some period $t \geq 1$ and history node $\left(h_\theta^t, h_y^{t-1}\right) = (\theta_h^t, y^{t-1})$, the incentive constraint holds strictly, i.e.,

$$U^M\left(\left(\theta_h^{t-1}, y^{t-1}\right), \theta_l\right) > U\left(c_t\left(\left(\theta_h^{t-1}, y^{t-1}\right), \theta_h\right) \mid \theta_l\right) + \sum_{\tilde{\theta}, y_t} \mathbb{P}\left(\tilde{\theta} \mid \theta_l\right) U_{t+1}^M\left(\left(\left(\theta_h^t\right), \left(y^{t-1}, y_t\right)\right), \tilde{\theta}\right).$$

In this case consider the following mechanism, defined as $M^{\left[y^{t-1}\right]} = \left(c_\tau^{[t]}\right)$:

$$c_\tau^{[t]}(h^\tau, \theta) = \begin{cases} u^{-1}\left(U^M\left(\theta_h^t, y^t, \theta_l\right)\right) & \text{, if } (h^\tau, \theta_\tau) \succeq (\theta_h^t, \theta_l); \\ u^{-1}\left(U^M\left(\theta_h^t, y^t, \theta_h\right)\right) & \text{, if } (h^\tau, \theta_\tau) \succeq (\theta_h^t, \theta_h), \end{cases}$$

and $c_\tau^{[t]} = c_\tau$ otherwise.

$M^{[t]}$ satisfies all of the incentive constraints, except for period t with history $\left(\left(\theta_h^{t-1}, \theta_l\right), y^{t-1}\right)$. Additionally, $M^{[t]}$ is necessarily cheaper than M (strictly so if they differ). Therefore, considering the mechanism $\tilde{M}$ defined by

$$u\left(\tilde{c}_\tau\left(h^{\tau-1}, \theta_\tau\right)(y)\right) \equiv (1-\varepsilon)\, u\left(c_\tau\left(h^{\tau-1}, \theta_\tau\right)(y)\right) + u\left(c_\tau^{[t]}\left(h^{\tau-1}, \theta_\tau\right)(y)\right),$$

for $\varepsilon > 0$ small enough satisfies all of the incentive constraints (because of the linearity of the

IC constraints in utility levels and the slack on the period t IC) and it is a strict improvement in terms of profits, a contradiction. Therefore it follows that $M = M^{[t]}$. Therefore, period t incentive constraints imply that consumption following θ_h^t-history is constant. Therefore, all incentive constraints for $\tau \geq t$ hold as equalities trivially. ∎

I know characterize sufficient conditions under which a solution to the relaxed problem satisfies IC.

Proposition 16. *(Sufficiency of relaxed constraints)*
Consider a mechanism $M \in \mathcal{M}$ such that, for any t and $h^{t-1} \in H_\theta^{t-1}$ and $\theta \in \Theta$:

$$U^M\left(h^{t-1},\theta_l\right) = U\left(c_t\left(h^{t-1},\theta_h\right) \mid \theta_l\right) + \sum_{\theta,y} p^l(y)\,\mathbb{P}(\theta \mid \theta_l)\,U^M\left(\left(h_\theta^{t-1},\theta_h\right),\left(h_Y^{t-1},y\right),\theta\right),$$

and, for all $t \geq 0$, $\tau = 0,\ldots,T-t$, $\theta \in \Theta$ and $\theta' \neq \theta$

$$U\left(c_{t+\tau}\left(h_\theta^t,\theta,(\theta_h)^\tau,h_Y^{t+\tau}\right) \mid \theta\right) \geq U\left(c_{t+\tau}\left(h^t,\theta,(\theta_h)^\tau,h_Y^{t+\tau}\right) \mid \theta'\right), \tag{8}$$

then $M \in \mathcal{M}^{IC}$.

Proof. Just notice that

$$U^M\left(h^t,\theta' \mid \theta\right) = U\left(c_t\left(h^t,\theta'\right) \mid \theta\right) + \sum_\tau \sum_{y^\tau} \sum_\theta \mathbb{P}(\theta_{t+\tau} = \theta \mid \theta_t = \theta)\,U\left(x_{t+\tau}\left(h^t,\theta',(\theta_h)^\tau\right) \mid \theta\right)$$

therefore

$$\begin{aligned}
&U^M\left(h^t,\theta_h \mid \theta_h\right) - U^M\left(h^t,\theta_l \mid \theta_h\right) = \\
&U\left(c_t\left(h^t,\theta_h\right) \mid \theta_h\right) + \sum_\tau \sum_{y^\tau} \sum_\theta \mathbb{P}(y^\tau,\theta_{t+\tau} = \theta \mid \theta_t = \theta_h)\,U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_h,(\theta_h)^\tau\right) \mid \theta\right) \\
&-U\left(c_t\left(h^t,\theta_l\right) \mid \theta_h\right) - \sum_\tau \sum_{y^\tau} \sum_\theta \mathbb{P}(y^\tau,\theta_{t+\tau} = \theta \mid \theta_t = \theta_h)\,U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_l,(\theta_h)^\tau\right) \mid \theta\right) \\
&= U^M\left(h^t,\theta_h \mid \theta_l\right) - U^M\left(h^t,\theta_l \mid \theta_l\right) \\
&+ \sum_{\tau\geq 0} \sum_{y^\tau} \Delta_{t,\tau}(y^\tau)\left[U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_h,(\theta_h)^\tau\right) \mid \theta_h\right) - U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_h,(\theta_h)^\tau\right) \mid \theta_l\right)\right] \\
&- \sum_{\tau\geq 0} \sum_{y^\tau} \Delta_{t,\tau}(y^\tau)\left[U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_l,(\theta_h)^\tau\right) \mid \theta_h\right) - U\left(c_{t+\tau}\left(h^t,y^\tau,\theta_l,(\theta_h)^\tau\right) \mid \theta_l\right)\right],
\end{aligned}$$

where $\Delta_{t,\tau}(y^\tau) \equiv \mathbb{P}(y^\tau,\theta_{t+\tau} = \theta_h \mid \theta_t = \theta_h) - \mathbb{P}(y^\tau,\theta_{t+\tau} = \theta_l \mid \theta_t = \theta_l) > 0$. ∎

Finally, I prove that the solution to the relaxed problem satisfies the monotonicity condition.

Proposition 17. *(Equivalence of relaxed and original problem)*
Any solution to the relaxed problem satisfies (8), therefore, it is optimal.

Proof. Fix $t \geq 1$ and $\tau \in \{0, \ldots, T-t\}$. The condition holds as an equality if h_θ^t contains at least one θ_l realization, since then it follows that

$$c_{t+\tau}\left(h^t, \theta, (\theta_h)^\tau\right) \in C^*.$$

Now I focus on $h = (\theta_h, \theta_h, \ldots, \theta_h)$ and $\theta = \theta_h$. From the necessary condition (7) it follows that

$$c_{t+\tau}\left(h^t, \theta_h, (\theta_h)^\tau, h_y^{t+\tau}\right)(y) \geq c_{t+\tau}\left(h^t, \theta_h, (\theta_h)^\tau, h_y^{t+\tau}\right)(y'),$$

if and only if

$$\frac{p^h(y)}{p^l(y)} \geq \frac{p^h(y')}{p^l(y')},$$

which implies that

$$U\left(c_{t+\tau}\left(h^t, \theta_h, (\theta_h)^\tau\right) \mid \theta\right) \geq U\left(c_{t+\tau}\left(h^t, \theta_h, (\theta_h)^\tau\right) \mid \theta'\right).$$

■