Regulation Design in Insurance Markets*

Dhruva Bhaskar†, Andrew McClellan‡, and Evan Sadler§

February 9, 2020

Abstract

Regulators often impose rules that constrain the behavior of market participants. We study optimal regulation design in an insurance market. A regulator restricts the set of contracts a firm is allowed to offer. The firm offers a permitted menu to each agent, and the agent chooses a contract from the menu. The regulator seeks to maximize agent welfare. We show that under a risk-ordering condition, the regulator can implement her first-best allocation—all agents obtain full insurance at a constant price—using a collection of two-option menus. When the regulator has more limited enforcement power, we show that insurance mandates may harm welfare.

1 Introduction

Many insurance markets are subject to stringent regulation in an attempt to promote better risk sharing and ultimately higher welfare among consumers. Health insurance is perhaps the most salient example. In the United States, seemingly every political campaign at some point centers on proposals to adjust the rules to which insurers and insurance customers are subject. Such rules include making insurers offer policies to all consumers regardless of pre-existing conditions, limits on price differentials across demographic groups, and of course the individual mandate to buy an insurance policy. Health insurance is far from the only example—car and homeowners insurance markets are also subject to various rules about what kind of coverage must be offered and purchased. Regulation is needed

*We would like to thank Alp Atakan, Joyee Deb, Alex Frankel, Jacob Leshno, Elliot Lipnowski, Marco Mariotti, Sujoy Mukerji, David Pearce, Harry Pei, Michael Richter, Maher Said, Ludovic Renou, and Ariel Rubinstein for helpful comments and discussions. Dhruva Bhaskar acknowledges support from a grant from the European Research Council (ERC 681460, Informative Prices). Andrew McClellan acknowledges support from the Liew Family Junior Faculty Fellowship at the University of Chicago Booth School of Business.

†Queen Mary University of London

‡University of Chicago, Booth School of Business

§Columbia University

in these markets to overcome incentive issues that arise from information asymmetry and ensure that risk actually is shared within a large pool.[1]

At the same time, insurance companies have a growing amount of data about consumers. In a 2018 report, The Geneva Association[2] states:

> Advances in big data analytics...are transforming the insurance industry and the role that data plays in insurance. New sources of digital data reveal information about behaviours and lifestyle habits that allow insurers to assess individual risks much better than before. The emergence of big data, however, raises several concerns regarding...personalization of insurance"

The increasing analytical sophistication of firms raises new questions about both the efficacy of existing regulations and the scope for designing better ones.

How effective are insurance regulations, and how should a regulator design them? We propose a framework to assess what contracts a regulator should permit and the effects of insurance mandates. A single firm can offer menus of insurance contracts to a large population of agents, and a regulator can restrict the set of permissible menus. The agents have private information about their own risk, represented as one of finitely many types. The firm has some information, observing the "category" of each agent. The regulator knows only an aggregate distribution. The regulator chooses a set of contract menus that the firm is allowed to offer. The firm then chooses whether to incur a cost to enter the market, and if it enters, what menu to offer each agent. Each agent then chooses a contract from the menu they are offered. We assume agents maximize their own utilities, the firm maximizes its profits, and the regulator seeks to maximize the average utility of the agents.

Our main result shows that, under an ordering condition on agent types, the regulator can implement her first-best outcome using menus that contain just two options. We first demonstrate this result for a simplified setting in which the first-best outcome entails providing full insurance at a uniform price to all agents—the price exactly compensates the firm's entry cost. To implement this, the regulator mandates that all agents buy insurance and specifies a menu for each category of agent. Each menu contains two contracts. One is the intended contract, offering full insurance at the desired price, while the other is a "deviation contract." The deviation contract is never chosen in equilibrium, but its presence provides incentives for the firm to offer the right menu to each category.

Our construction depends on a risk-ordering of the observable categories. When considering a deviation in what menu to offer a particular category of agent, the firm can choose the menu intended for a lower risk category or the menu intended for a higher risk category. If the firm chooses the former, then some agents will choose the deviation contract, and the cost of large payouts to higher risk agents makes this unprofitable for the firm. If the firm chooses the latter, then the payouts from the intended contract are higher, but the premium

[1]For instance, see Einav and Finkelstein [2011] for a survey.

[2]The International Association for the Study of Insurance in Economics

is the same. As a result, the firm prefers to offer each category its intended menu. Agents within a category choose the intended contract because only higher risk types are willing to pay the premium of the deviation contract.

In the current era of mass data collection, there is increasing scrutiny of the potential harm caused by firms possessing detailed information about their customers.[3] We point out that, in the presence of strong regulation, it actually benefits consumers when the firm knows more about them. Our implementation leverages the firm's knowledge of agent categories to select the right contract for each agent. As categories get finer, and we maintain the risk-ordering condition, the optimal regulatory regime becomes *better* for the agents.

In the more general setting, we first study what allocations the regulator can implement using a collection of menus that contain just a single deviation contract—all other contracts are chosen by some types in equilibrium. We view this feature, the use of simple menus, as important for the practical significance of our analysis.[4] An analogous risk-ordering condition on agent types allows us to implement any allocation that is agent incentive compatible within categories. This finding greatly expands the applicability of our framework—with our implementation result in hand, it becomes possible to study regulators with varying objectives. We use our result to derive a sufficient condition for implementability of the regulator's first-best outcome via simple menus.

An important feature of our result is that it requires a regulator with strong enforcement power. The regulator needs to ensure that all agents buy insurance and that the firm offers each agent a particular *menu* of contracts. That is, the regulator must ensure the firm cannot offer the agent a single contract, or some larger menu of contracts, drawn from those appearing in some intended menu. This power is crucial for our implementation. Absent this ability, not only is the first-best outcome no longer feasible, but an individual mandate can harm consumer welfare. Einav and Finkelstein [2011] state "Under the textbook assumptions...mandatory insurance coverage is always a (weakly) welfare-improving policy intervention". We identify conditions under which this is not the case.

Moving beyond monopoly regulation, our framework has something to say about regulation in partially competitive markets as well. The first-best allocation entails significant cross-subsidization of high risk agents by low risk agents, which may unravel if an entrant can cream-skim more profitable types. We identify a second-best allocation that is robust to a no cream-skimming constraint. Our work on implementation implies that the regulator can implement this allocation using three-option menus. This demonstrates how simple regulation can improve consumer welfare, even when cream-skimming might otherwise lead to market unravelling.

Insurance regulation is a widely studied and debated topic. We offer a fresh perspective. Our analysis identifies a natural way to implement the first-best along with new reasons why well-intentioned regulatory regimes might fail to deliver good outcomes. Individual insurance

[3]See, for instance, the Geneva Association's report, Rep [2018] for a discussion of these issues.

[4]There is extensive evidence that consumers suffer from choice overload when choosing from large complex sets of alternatives. See Chernev et al. [2015] for an analysis.

mandates are a crucial ingredient, but the benefit from this depends on how effectively a regulator can control other aspects of the insurance market. More broadly, we believe our approach to modeling regulation can offer a template for other applications (e.g. wage and employment regulation).

After discussing related work, we present our modeling framework along with several more specific examples. We then state our main result in a special case of the model chosen for expositional transparency. After presenting the general version, we study the impact of insurance mandates and competition from entrants. Section 7 discusses additional extensions, and section 8 concludes.

1.1 Related Work

As Laffont [1994] highlights, a crucial feature of the problem regulators face is that firms typically have more information about the environment than their regulators. The closest literature to our work is that on monopoly regulation with privately known costs. The seminal work of Baron and Myerson [1982] studies optimal price regulation and subsidy provision for a monopolist facing an exogenous demand curve.[5] Our setting differs in several significant ways. First, the firm's private information concerns the characteristics of potential customers. Second, the firm's action set is much richer as it contracts with each agent individually. An unregulated firm can use its private information to screen and sell different products to different agents. Most importantly, we model agents' incentives. The interaction between the firm and the agents provides a lever for the regulator to exploit—agent incentives discipline the actions of the firm. In contrast, existing work only considers firm incentives, so the the regulation problem reduces to a principal-agent problem.

One can also place our work within the literature on optimal delegation. This literature studies how a principal can optimally restrict choices for an agent with superior information but misaligned preferences. Our problem is a delegation problem: the regulator delegates to the firm the choice of what menu to offer each agent. The regulator wants the firm to offer higher coverage to higher risk agents, charging a uniform price, while the firm wants to offer the lowest coverage at the highest possible price. In the canonical delegation model, the action space is a subset of the real line, and the agent has a bias for higher actions [e.g. Holmström, 1984, Alonso and Matouschek, 2008, Amador and Bagwell, 2013]. Our delegation set is the space of contract menus—standard techniques are not applicable in our setting. Moreover, the interaction with agents' incentives gives rise to additional restrictions on the firm's actions that do not appear elsewhere in this literature.

The deviation contracts in our policy deter the firm from offering menus intended for lower categories to higher categories. This deterrent works because some agents in higher categories would choose the deviation contract while those in lower categories do not. This is reminiscent of the optimal contract in Galperti [2015], where a principal sells commitment

[5] See also Lewis and Sappington [1988], Laffont and Tirole [1986], Laffont and Tirole [1987], Riordan and Sappington [1986].

devices to agents with privately known discount factors. The principal allows extra flexibility in the contracts for more patient agents to deter impatient agents from buying these contracts—the patient agents never use the additional options, but impatient agents would. As in our case, these extra options are not used in equilibrium.

Policy complexity is an important issue in existing work on regulation. Laffont and Tirole [1986] use an infinite menu of linear contracts to regulate their monopolist. Rogerson [2003], noting how difficult it is to implement such complex menus in practice, constructs an alternative with two contracts that achieves a constant fraction of the optimal welfare. In our setting, a collection of 2-option menus attains the optimal welfare. The simplicity of our policy suggests greater scope for practical implementations.

Our problem embeds the classic monopolistic screening model for insurance markets [Stiglitz, 1977, Chade and Schlee, 2012]. Left unregulated, the firm would screen agents to maximize profits, offering full insurance at a high price to high risk agents and distorting coverage for low risk agents. Our optimal regulatory policy eliminates this screening, ensuring that all agents are fully insured at a constant price.

2 Framework

We study a game between a regulator, a firm, and a unit mass of agents. Each agent belongs to one of finitely many categories—we endow the set of categories X with an order $\prec$. For each agent, there is a finite set Ω of verifiable events upon which the firm and the agent can contract. If an agent in category x experiences event ω and receives a net transfer z from the firm, her utility is $u(z, \omega, x)$. The function $u : (\underline{z}, \infty) \times \Omega \times X \to \mathbb{R}$ is strictly increasing and concave in z, and there is a lower bound $\underline{z}$ on the net transfer such that

$$\lim_{z \to \underline{z}} u(z, \omega, x) = -\infty$$

for every category x and every event ω.[6] The events $\omega \in \Omega$ represent different types of loss an agent can suffer. We assume there is a unique event $\omega_0 \in \Omega$ such that $u(z, \omega_0, x)$ is constant in x and $u(z, \omega_0, x) \geq u(z, \omega, x)$ for all $\omega \in \Omega$—we call ω_0 the no-loss event.

Each agent has a privately observed risk type $\theta \in \Delta(\Omega)$ that describes the risk she faces—for an agent of type θ, event ω occurs with probability $\theta(\omega)$. Agents of category x have types contained in a finite set $\Theta_x \subset \Delta(\Omega)$—the two dimensional type (x, θ) completely characterizes an agent. Write $\mathcal{T}$ for the collection of all pairs (x, θ) with $x \in X$ and $\theta \in \Theta_x$. The distribution of agent types is $\mu \in \Delta(\mathcal{T})$—assume full support—and for each $x \in X$, the conditional distribution of risk types is $\mu_x \in \Delta(\Theta_x)$. These distributions are common knowledge. Each agent observes her own type (x, θ), the firm observes the category x of every agent, and the regulator observes nothing.

[6]If one wants to eliminate the lower bound, we can replace this condition with $\lim_{z \to -\infty} u'(z, \omega, x) = \infty$.

The firm chooses whether to enter the market, incurring cost $k \geq 0$ if it enters. Conditional on entering, the firm offers a menu of contracts M_x to agents in category x. A contract (p, t) comprises

(a) A premium $p \in \mathbb{R}_+$ that the agent pays to the firm, and

(b) A transfer function $t : \Omega \to \mathbb{R}_+$ specifying an amount of compensation to the agent for each loss event.

If an agent in category x with risk type θ buys a contract (p, t) from the firm, she earns expected utility

$$U(x, \theta, (p, t)) := \sum_{\omega \in \Omega} \theta(\omega) u(t(\omega) - p, \omega, x)$$

and the firm earns an expected profit

$$\Pi(x, \theta, (p, t)) := p - \sum_{\omega \in \Omega} \theta(\omega) t(\omega).$$

Our framework embeds several important special cases. We focus mostly on the following two, though later results are more general.

Known Monetary Loss

Suppose all agents begin with the same wealth level w, and the event ω corresponds to a known monetary loss $\ell_x(\omega)$ for agents in category x. The agents share a common value for money $v : \mathbb{R} \to \mathbb{R}$, so given the contract (p, t), we have

$$u(z, \omega, x) := v(w - p + t(\omega) - \ell_x(\omega)).$$

Imagine an agent might suffer some verifiable health problem, such as an illness or an accident, and the cost of treatment can depend on observable attributes (e.g. age, location) or pre-existing health conditions, which we capture through the category $x \in X$. The insurance company can use the information in x to decide what contract to offer, and the payout is a function of the verifiable event.

Unverifiable Monetary Loss

Suppose there is a set of true loss events $\hat{\Omega}$, and a verifiable loss event $\omega \subset \hat{\Omega}$ comprises some subset of these—the verifiable events Ω form a finite partition of $\hat{\Omega}$. Event $\hat{\omega} \in \hat{\Omega}$ corresponds to a monetary loss $\ell_x(\hat{\omega})$ for agents in category x. The distribution over $\hat{\Omega}$ conditional on ω depends on an agent's category and is given by $\nu_x(\hat{\omega}|\omega)$. Assume again that all agents have the same initial wealth level w, and $v : \mathbb{R} \to \mathbb{R}$ describes preferences over wealth. Given the contract (p, t), we have

$$u(z, \omega, x) := \int_{\hat{\omega} \in \omega} v\left(w - p + t(\omega) - \ell_x(\hat{\omega})\right) \mathrm{d}\nu_x(\hat{\omega}|\omega).$$

Imagine that event ω corresponds to the agent suffering from a particular illness, but the exact cost of treatment is not verifiable. Different patients have different needs, and different doctors have different opinions about the best course of action. The firm has some distributional knowledge of costs (e.g. ν_x) and offers coverage based on what it can verify.[7]

2.1 Contract Regulation

The regulator can restrict the contracts that the firm is permitted to offer.

Definition 1. *A **regulatory policy** $\mathcal{R}$ is a finite set of menus, where each menu $M \in \mathcal{R}$ is a finite set of contracts.*

If the firm enters the market it must offer some menu in $\mathcal{R}$ to each agent, and each agent must choose a contract from the offered menu. This is without loss of generality. To represent agents who can opt out of insurance, we could restrict the regulator to policies such that $(0, 0) \in M$ for every $M \in \mathcal{R}$. Similarly, to represent a firm that can exclude customers, we could require that the menu $M = \{(0, 0)\}$ is contained in $\mathcal{R}$. As a baseline, we allow the regulator to choose the policy without restrictions.

The structure of a regulatory policy affects whether the firm and the agents get to make choices. At one extreme, policies of the form

$$\mathcal{R} = \{\{(p_1, t_1), (p_2, t_2), ..., (p_N, t_N)\}\}$$

grant only the agents a choice. The firm must offer the same menu to all agents, and each agent faces the same set of options. At the other extreme, policies of the form

$$\mathcal{R} = \{\{(p_1, t_1)\}, \{(p_2, t_2)\}, ..., \{(p_N, t_N)\}\}$$

grant only the firm a choice. Each menu is a singleton, and agents must accept what is offered.

In general, a regulatory policy allows both the firm and the agents to make decisions. The firm chooses which menu to offer to each agent, and each agent chooses a contract from her menu. Consider a policy

$$\mathcal{R} = \{\{(p_1, t_1), (p_2, t_2)\}, \{(p_3, t_3), (p_4, t_4)\}\}$$

[7] Firms necessarily limit coverage to control costs. Insurance companies and patients/healthcare providers often disagree about what treatments are necessary, particularly with regard to newer more expensive treatments. Pollitz et al. [2019] find that "among issuers offering individual market coverage on healthcare.gov... 18 % of in-network claims were denied by issuers in 2017."

There are two menus, each containing two contracts. Allowing choice by both players creates an opportunity to leverage the interplay between the firm's and the agents' incentives. Suppose the regulator wants category x to get contract (p_1, t_1) and category x' to get contract (p_3, t_3). If the firm likes (p_1, t_1) better for both categories, a menu of singletons is not incentive compatible for the firm, and if agents all prefer (p_3, t_3), then a single menu is not incentive compatible for the agents. If the regulator can find a contract (p_2, t_2) that is unprofitable for the firm, and that category x' likes better than (p_1, t_1), we can get the desired allocation.

2.2 The Regulator's Problem

The timing is as follows:

- The regulator chooses a regulatory policy $\mathcal{R}$.
- The firm chooses whether to enter the market, and if it enters, it offers $M_x \in \mathcal{R}$ to each agent in category x
- Each agent chooses a contract from her menu.
- Events in Ω obtain, and contracts pay contingent transfers.

The regulator's goal is to maximize agent welfare. An **allocation** $\mathbf{a} = \{(p,t)_{x,\theta}\}_{x,\theta \in \mathcal{T}}$ specifies a contract for each type of agent. Welfare from the allocation is

$$W(\mathbf{a}) := \sum_{(x,\theta) \in \mathcal{T}} \mu(x,\theta) U(x, \theta, (p,t)_{x,\theta}).$$

The firm's profit is

$$\pi(\mathbf{a}) := \sum_{(x,\theta) \in \mathcal{T}} \mu(x,\theta) \Pi(x, \theta, (p,t)_{x,\theta}) - k.$$

Without loss of generality, we confine attention to regulatory policies $\mathcal{R} = \{M_x\}_{x \in X}$ that contain exactly one menu for each category of agent.

Definition 2. *A policy $\mathcal{R} = \{M_x\}_{x \in X}$* ***implements*** *the allocation $\mathbf{a} = \{(p,t)_{x,\theta}\}_{x,\theta \in \mathcal{T}}$ if there exists a Perfect Bayesian Equilibrium of the induced game in which:*

(a) The firm enters the market,

(b) The firm offers M_x to agents in category x, and

(c) Agents of type (x,θ) choose contract $(p,t)_{x,\theta}$ from M_x.

An implementable allocation must satisfy participation and incentive compatibility constraints for the firm and incentive compatibility constraints for the agents. Given a regulatory policy $\mathcal{R}$ that implements $\mathbf{a}$, the corresponding PBE must specify a contract $(p,t)^{x'}_{x,\theta}$ that type (x,θ) chooses from menu $M_{x'}$—we necessarily have $(p,t)^{x}_{x,\theta} = (p,t)_{x,\theta}$. Define

$$\pi(x, x') = \sum_{\theta \in \Theta_x} \mu_x(\theta)\Pi(x, \theta, (p,t)^{x'}_{x,\theta})$$

as the firm's expected profit from offering menu $M_{x'}$ to category x agents. The allocation $\mathbf{a}$ is implementable if and only if there exists a collection of contracts $\{(p,t)^{x'}_{x,\theta}\}_{(x,\theta)\in\mathcal{T},x'\in X}$ and corresponding menus $M_x = \{(p,t)^{x}_{x',\theta}\}_{(x',\theta)\in\mathcal{T}}$ such that

$$(a) \quad \pi(\mathbf{a}) \geq 0 \qquad (FPC)$$

$$(b) \quad \pi(x,x) \geq \pi(x,x') \text{ for all } x, x' \in X \qquad (FIC)$$

$$(c) \quad U(x,\theta,(p,t)^{x'}_{x,\theta}) \geq U(x,\theta,(p,t)) \text{ for all } x,x' \in X, \theta \in \Theta_x, \text{ and } (p,t) \in M_{x'} \qquad (AIC)$$

FPC is the firm's participation constraint—the firm is willing to enter the market if it earns a non-negative profit. FIC are the firm's incentive constraints—offering menu M_x to category x agents is optimal. AIC are the agents' incentive constraints—an agent with type (x,θ) is willing to select the contract $(p,t)^{x'}_{x,\theta}$ from menu $M_{x'}$. The regulator's problem is to choose a regulatory policy to implement an allocation that maximizes welfare.

3 Implementing First Best: An Illustration

To render the key ideas in our analysis more transparent, this section presents our main result in the case of known monetary losses—agents have initial wealth w, face monetary losses $\ell_x(\omega)$ depending on their categories and loss events, and the utility from wealth is given by $v : \mathbb{R} \to \mathbb{R}$. First, we assess the first-best outcome from the regulator's perspective. Next, we explicitly construct an implementation for the case with two categories and two loss events. Finally, we provide general conditions under which the designer can implement the first-best outcome.

3.1 The Regulator's First-Best Outcome

As a benchmark, we consider the first-best outcome from the perspective of the regulator. Suppose the regulator can observe all agents' categories and risk types, and if the firm enters the market, the regulator can force it to offer specific contracts to each agent. Formally, we solve the regulator's problem subject only to the firm's participation constraint.[8] Since agents are risk averse, it is optimal to fully insure, guaranteeing constant wealth across states.

[8]This guarantees that first-best is well-defined as the firm has to break even in expectation.

Moreover, concave utility implies the regulator should charge all agents the same price. The optimal price equals the expected costs of the firm—the sum of the entry cost and expected payouts.

Proposition 1. *In the **first-best allocation**, agents in category x buy the contract (p^*, ℓ_x), where*

$$p^* = k + \sum_{x,\theta \in \mathcal{T}} \mu(x, \theta) \sum_{\omega \in \Omega} \theta(\omega) \ell_x(\omega)$$

Proof. Strict concavity of v implies that the first-best allocation must provide the same utility across all categories and loss events. The result follows. □

In the first-best outcome, agents with low expected losses subsidize those with high expected losses. The firm makes losses on some agents and profits on others. Insurance is mandatory—no types are allowed to opt out. Note the insurance mandate may be a binding constraint for some agents. Those with low expected losses may prefer to opt out rather than subsidize the others. We also need to assume that the entry cost is not too high, so the allocation provides higher consumer welfare than autarky.

In the absence of any regulation, the firm can offer an agent any menu, and agents in turn choose an offered contract or choose not to buy insurance. This reduces our model to a standard monopolistic screening problem. The firm separately screens each category x with a menu chosen to maximize profits. If there is only one risk type in a category, the firm fully extracts surplus from those agents. If there are multiple risk types within a category, the firm offers a menu with the standard distortions: low risk types are under-insured in order to extract higher premiums from high risk types, and some agents may be excluded from insurance.

3.2 Two Categories and Two Loss Events

Our main result shows that we can implement the first-best outcome using a simple policy. The main ideas are apparent with just two categories, each with a single associated risk type, and two loss events. Suppose the loss events are $\Omega = \{\omega_0, \omega_1\}$, where $\ell_x(\omega_0) = 0$ and $\ell_x(\omega_1) = \ell_x$, and the agent categories are $X = \{L, H\}$. Assume that $\ell_H > \ell_L$, that $\theta_H = \theta_L = \theta$, and $k = 0$. Agents can have either high or low risk. Both have the same probability of loss, but high risk agents suffer larger losses. A contract here is a pair (p, t), where p is the premium charged and t is the transfer made in event ω_1. The first-best outcome entails category x agents buying the contract (p^*, ℓ_x), where

$$p^* = \theta \left(\mu(H)\ell_H + \mu(L)\ell_L \right).$$

Suppose the regulator allows the firm to offer just one menu to all agents, choosing the policy

$$\mathcal{R}^A = \{\{(p^*, \ell_H), (p^*, \ell_L)\}\}.$$

This clearly fails to implement the desired outcome because all agents will choose (p^*, ℓ_H)—the two contracts charge the same premium, so agents opt for the larger payout. Moreover, since all agents receive the large payout, the resulting allocation violates the firm's participation constraint, so the firm will not enter the market. Suppose instead that the regulator allows the firm to offer singleton menus, choosing the policy

$$\mathcal{R}^F = \{\{(p^*, \ell_H)\}, \{(p^*, \ell_L)\}\}.$$

In this case, the firm prefers to offer $\{p^*, \ell_L)\}$ to all agents because of lower payouts. Again, we fail to implement the desired allocation.

If only one of the contracting parties has a non-trivial choice in our regulatory regime, we cannot implement the first-best outcome. Successful implementation requires leveraging both parties' incentives—we must allow the firm to offer the low payout contract to low risk agents and incentivize the firm to offer the high payout contract to high risk agents. We can do this by introducing a third contract that is never chosen in equilibrium.

Consider a contract $(\bar{p}, \bar{t})$ with $\bar{p} - \theta\bar{t} < p^* - \theta\ell_H$, meaning it is less profitable for the firm than the high risk contract. Since u is strictly concave, we can choose $\bar{p}$ and $\bar{t}$ satisfying

$$v(w - p^*) > \theta v(w - \bar{p} + \bar{t} - \ell_L) + (1 - \theta)v(w - \bar{p}), \quad \text{and}$$

$$\theta v(w - \bar{p} + \bar{t} - \ell_H) + (1 - \theta)v(w - \bar{p}) > \theta v(w - p^* + \ell_L - \ell_H) + (1 - \theta)v(w - p^*).$$

The first condition says that a low risk agent prefers the intended contract (p^*, ℓ_L). The second says that a high risk agent prefers the contract $(\bar{p}, \bar{t})$ to the low type contract (p^*, ℓ_L). The contract offers a higher payout $\bar{t} > \ell_L$ at a higher price $\bar{p} > p^*$. We refer to this as the "deviation contract."

If the regulator chooses the policy

$$\mathcal{R}^* = \{\{(p^*, \ell_H)\}, \{(p^*, \ell_L), (\bar{p}, \bar{t})\}\},$$

we obtain the first-best outcome in equilibrium. We allow the firm to choose which of the two intended contracts to offer, but if the firm offers (p^*, ℓ_L), it must also offer $(\bar{p}, \bar{t})$. The firm offers the intended contract to low risk agents because they choose it, and this is the most profitable contract. The firm offers the intended contract to high risk agents because otherwise they choose the less profitable deviation contract. Though never chosen in equilibrium, the deviation contract provides a crucial deterrent.

This example illustrates a more general principle that allows the regulator to implement the first-best allocation through simple menus. Our main result in the next section shows that, as long as the categories are suitably ordered in terms of risk, a collection of two-contract menus suffices for implementation. Each menu contains an intended contract for the given category and a single deviation contract. The deviation contract is chosen by some agents in any higher risk category and is less profitable than the intended contracts for higher risk agents.

3.3 First-Best Implementation

We first need to formalize what "suitably ordered" means. Define

$$m(x, x') = \sum_{\theta \in \Theta_x} \mu_x(\theta) \sum_{\omega \in \Omega} \theta(\omega) \ell_{x'}(\omega).$$

This is an expected loss for category x agents, but we replace the loss values $\ell_x(\omega)$ with the corresponding values for category x'—note $m(x, x)$ is the expected loss for category x agents. If a category x agent obtains the contract intended for a category x' agent, the value $m(x, x')$ is the firm's expected payout. Define also

$$\overline{r}_x(\omega) = \max_{\theta \in \Theta_x} \frac{\theta(\omega)}{\theta(\omega_0)}, \quad \overline{\theta}_x(\omega) = \arg\max_{\theta \in \Theta_x} \frac{\theta(\omega)}{\theta(\omega_0)}.$$

Within a category x, we look at the risk type that maximizes the ratio $\frac{\theta(\omega)}{\theta(\omega_0)}$—the maximum value is $\overline{r}_x(\omega)$ and the maximizing type is $\overline{\theta}_x(\omega)$. Type $\overline{\theta}_x(\omega)$ has the highest relative risk of loss event ω among category x agents.

Assumption 1. *The ordering $\prec$ on X is such that*

(a) Expected payouts are weakly increasing in the intended category of a contract: whenever $x \prec x'$, we have $m(x, x) \leq m(x, x')$.

(b) There exists $\omega^* \in \Omega$ such that $\ell_x(\omega^*)$ and $\overline{r}_x(\omega^*) > 0$ are weakly increasing in x: whenever $x \prec x'$, we have $\ell_x(\omega^*) \leq \ell_{x'}(\omega^*)$ and $0 < \overline{r}_x(\omega^*) \leq \overline{r}_{x'}(\omega^*)$.

According to Assumption 1, there are two distinct senses in which higher categories correspond to higher risk. Condition (a) implies that higher categories tend to suffer higher losses on average, so the contract $(p^*, \ell_{x'})$ is more expensive for the firm than the contract (p^*, ℓ_x) whenever $x' \succ x$. As stated, the condition is somewhat cumbersome to check, but here are two simple sufficient conditions:

- We have $\Theta_x = \Theta$ and $\mu_x = \hat{\mu}$ for all x, and $m(x, x)$ is weakly increasing in x, or
- We have that $\ell_x(\omega)$ is weakly increasing in x for all $\omega \in \Omega$.

Condition (b) implies that there is a particular loss event ω^* in which losses are increasing in x, and the maximum relative likelihood of this event is increasing in x. Note this condition is relatively weak—higher risk categories need not suffer higher losses in all loss events. We only need one event for which the categories are well-ordered.

Definition 3. *A* ***2-option regulatory policy*** *is a collection of menus $\{M_x\}_{x \in X}$ in which each menu $M_x = \{(p_x, t_x), (\overline{p}_x, \overline{t}_x)\}$ consists of two choices.*

Our main result shows that under assumption 1, a 2-option regulatory policy is sufficient to implement the regulator's first-best allocation.

Theorem 1. *Under assumption 1, there exists a 2-option regulatory policy $\mathcal{R} = \{M_x\}_{x \in X}$, with $M_x = \{(p^*, \ell_x), (\overline{p}_x, \overline{t}_x)\}$, that implements the regulator's first-best allocation.*

Proof. Each menu M_x contains the intended contract (p^*, ℓ_x). We construct deviation contracts $(\overline{p}_x, \overline{t}_x)$ with the following properties:

- For every $x' \preceq x$ and $\theta \in \Theta_{x'}$, an agent (x', θ) prefers (p^*, ℓ_x) over $(\overline{p}_x, \overline{t}_x)$
- For every $x' \succ x$, there exists $\theta \in \Theta_{x'}$ such that an agent (x', θ) prefers $(\overline{p}_x, \overline{t}_x)$ over (p^*, ℓ_x).

For $x' \succ x$, the condition $m(x, x) \leq m(x, x')$ ensures the firm prefers to offer M_x to category x agents rather than $M_{x'}$. Consequently, we only need to show that our deviation contracts are such that the firm prefers to offer M_x to category x rather than $M_{x'}$ whenever $x' \prec x$.

Define the contract $(\overline{p}_x, \overline{t}_x)$ so that

$$\overline{t}_x(\omega) - \overline{p}_x = \begin{cases} \ell_x(\omega) - p^* & \text{if } \omega \notin \{\omega_0, \omega^*\} \\ -\overline{p}_x & \text{if } \omega = \omega_0 \\ \hat{t}_x - \overline{p}_x & \text{if } \omega = \omega^*. \end{cases}$$

We are left with two free parameters: the premium $\overline{p}_x$, and the transfer $\hat{t}_x$ in event ω^*. Consider a category x' agent with type θ. This agent prefers $(\overline{p}_x, \overline{t}_x)$ to (p^*, ℓ_x) if and only if

$$\theta(\omega_0)u(-\overline{p}_x) + \theta(\omega^*)u\left(\hat{t}_x - \ell_{x'}(\omega^*) - \overline{p}_x\right) \geq \theta(\omega_0)u(-p^*) + \theta(\omega^*)u\left(\ell_x(\omega^*) - \ell_{x'}(\omega^*) - p^*\right),$$

or equivalently

$$\frac{\theta(\omega^*)}{\theta(\omega_0)}\left(u\left(\hat{t}_x - \ell_{x'}(\omega^*) - \overline{p}_x\right) - u\left(\ell_x(\omega^*) - \ell_{x'}(\omega^*) - p^*\right)\right) \geq u(-p^*) - u(-\overline{p}_x). \tag{1}$$

The firm's profit from this contract is

$$\overline{p}_x - \sum_{\omega \in \Omega} \theta(\omega)\overline{t}_x(\omega) \leq \overline{p}_x - \theta(\omega^*)\hat{t}_x \tag{2}$$

Notice that if (1) holds for $x' = x$ and $\theta = \overline{\theta}_x(\omega^*)$, then it also holds for every $x' \succ x$ and $\theta = \overline{\theta}_{x'}(\omega^*)$—this follows from concavity of u and the second condition in assumption 1. If $x' = x$ and $\theta = \overline{\theta}_x(\omega^*)$, then (1) reduces to

$$\frac{\theta(\omega^*)}{\theta(\omega_0)}\left(u\left(\hat{t}_x - \ell_x(\omega^*) - \overline{p}_x\right) - u\left(-p^*\right)\right) \geq u(-p^*) - u(-\overline{p}_x).$$

Fix any $\overline{p}_x > p^*$ and define $\hat{t}_x(\overline{p}_x)$ as the value of $\hat{t}_x$ that satisfies this with equality—we know this is possible because the right hand side is positive, while for $\hat{t}_x = 0$ the left hand side is negative.

Assumption 1 ensures that any agent of category $x' \preceq x$ weakly prefers the contract (p^*, ℓ_x). For $x' \succ x$, there is a positive probability of type $\overline{\theta}_{x'}(\omega^*)$, and these agents prefer the deviation contract. Since $u(-\overline{p}_x)$ approaches $-\infty$ as $-\overline{p}_x \to \underline{z}$, we know that $\hat{t}_x(\overline{p}_x)$ must approach ∞, making the bound in (2) arbitrarily negative for this type. Consequently, for large enough $\overline{p}_x$, the corresponding deviation contract is sufficiently unprofitable that the firm never offers M_x to an agent of category $x' \succ x$. $\square$

The proof largely follows the example in 3.2: the menu for category x contains the intended contract and a deviation contract. The deviation contract acts as a deterrent, preventing the firm from offering lower risk menus to higher risk agents. The deviation contract takes a simple form. It charges a high price, and the payout differs from the intended contract only in state ω^*—in this state, it offers a higher payout.

Having a distribution of risk types in each category complicates our analysis. Unlike in the example of 3.2, not all agents in a higher category choose the deviation contract from a lower category's menu. Instead, the deviation contract targets a specific type with relatively high risk—the type that maximizes the ratio $\frac{\theta(\omega^*)}{\theta(\omega_0)}$. Agents in a higher category face a higher loss in event ω^*, so they place a higher value on payouts in this event, and there are types for which this event is relatively more frequent. This allows us to separate some agents in higher risk categories without attracting any lower risk agents to the deviation contract. By offering a high payout in event ω^*, and charging an appropriate price, we can make the deviation contract arbitrarily unprofitable for the firm, thereby providing incentives to offer the intended menu.

4 Implementation in the General Case

This section presents our results for the general model of section 2. Recall agent utilities are $u(z, \omega, x)$, where u is strictly increasing and strictly concave in the net transfer z, and there is a lower bound $\underline{z}$ such that

$$\lim_{z \to \underline{z}} u(z, \omega, x) = -\infty$$

for every category x and every event ω. Additionally, there exists a no-loss event $\omega_0 \in \Omega$ such that $u(z, \omega_0, x)$ is constant in x and $u(z, \omega_0, x) \geq u(z, \omega, x)$ for all $\omega \in \Omega$. We first establish an implementation result for arbitrary allocations. Given a particular allocation **a**, we ask: when can the regulator implement **a** using "simple" menus? We subsequently use this characterization to derive conditions under which the regulator can implement her first-best allocation.

4.1 Implementable Allocations

Recall an allocation $\mathbf{a} = \{(p,t)_{x,\theta}\}_{x,\theta\in\mathcal{T}}$ specifies a contract for each type of agent. We restrict attention to allocations such that $p < -\underline{z}$ in every contract—no agent earns utility $-\infty$. In any implementation of $\mathbf{a}$, we necessarily have $M_x^{\mathbf{a}} := \{(p,t)_{x,\theta}\}_{\theta\in\Theta_x} \subset M_x$ for each $x \in X$—the set $M_x^{\mathbf{a}}$ is the minimal set of contracts the firm must make available to category x so that agents can feasibly choose the intended allocation. One necessary condition for implementation is that an agent of type (x,θ) is willing to choose $(p,t)_{x,\theta}$ from $M_x^{\mathbf{a}}$.

Definition 4. *The allocation* $\mathbf{a} = \{(p,t)_{x,\theta}\}_{x,\theta\in\mathcal{T}}$ *is* ***agent incentive compatible*** *if*

$$U\left(x,\theta,(p,t)_{x,\theta}\right) \geq U\left(x,\theta,(p,t)\right)$$

for every $(x,\theta)\in\mathcal{T}$ *and every* $(p,t)\in M_x^{\mathbf{a}}$.

Note that if $M_x^{\mathbf{a}}$ is a singleton for every category x, the allocation is trivially agent incentive compatible.

We provide sufficient conditions for implementability based on a risk-ordering analogous to assumption 1. Define

$$C_{x,\theta}^{x'}(\mathbf{a}) = \arg\max_{(p,t)\in M_{x'}^{\mathbf{a}}} U\left(x,\theta,(p,t)\right).$$

This is the set of contracts an agent of type (x,θ) might choose from the menu $M_{x'}^{\mathbf{a}}$. Now define

$$\underline{\pi}^{\mathbf{a}}(x,x') = \sum_{\theta\in\Theta_x} \mu_x(\theta)\left(\min_{(p,t)\in C_{x,\theta}^{x'}(\mathbf{a})} p - \sum_{\omega\in\Omega}\theta(\omega)t(\omega)\right).$$

This is the minimum profit the firm obtains from offering menu $M_{x'}^{\mathbf{a}}$ to agents in category x.

Definition 5. *The allocation* $\mathbf{a} = \{(p,t)_{x,\theta}\}_{x,\theta\in\mathcal{T}}$ *is* ***upward incentive compatible*** *if* $\sum_{\theta\in\Theta_x}\mu_x(\theta)\Pi\left(x,\theta,(p,t)_{x,\theta}\right) \geq \underline{\pi}^{\mathbf{a}}(x,x')$ *whenever* $x \prec x'$.

Upward incentive compatibility plays the role of condition (a) in assumption 1. If $\mathbf{a}$ is upward incentive compatible, then the firm weakly prefers offering menu $M_x^{\mathbf{a}}$ to category x instead of menu $M_{x'}^{\mathbf{a}}$ whenever $x \prec x'$—higher categories receive more expensive contracts on average.

To define the analog of condition (b) in Assumption 1, we need a sense in which agents in higher categories suffer higher losses. The natural assumption is that, at least in some loss event ω^*, the marginal utility an agent gains from transfers is increasing in x. We also need ω^* to be relatively more likely for agents in higher categories. Recall the notation

$$\overline{r}_x(\omega) = \max_{\theta\in\Theta_x}\frac{\theta(\omega)}{\theta(\omega_0)}, \quad \overline{\theta}_x = \arg\max_{\theta\in\Theta_x}\frac{\theta(\omega)}{\theta(\omega_0)}.$$

Our analog to condition (b) requires that there exist some ω^* such that $\overline{r}_x(\omega^*)$ is weakly increasing in x, and $u(z,\omega^*,x)$ is supermodular in z and x.

Assumption 2. *There exists an event* $\omega^* \in \Omega$ *such that* $u(z, \omega^*, x)$ *is supermodular in* z *and* x *and* $\overline{r}_x(\omega^*) > 0$ *is weakly increasing in* x.

Depending on the allocation we want to implement, we need an additional assumption with no prior analog because $M_x^{\mathbf{a}}$ can contain multiple contracts. We need to ensure that for $x' \succ x$, agents of type $\left(x', \overline{\theta}_{x'}(\omega^*)\right)$ prefer a deviation contract to *all* elements of $M_x^{\mathbf{a}}$, not just to $(p,t)_{x,\overline{\theta}_x(\omega^*)}$. One sufficient condition is the following.

Assumption 3. *The value* $\overline{r}_x(\omega^*)$ *is strictly increasing in* x, *and* $\lim_{z\to\infty} u(z, w^*, x) = \infty$ *for every* x.

This assumption ensures that by scaling up the transfer in state ω^*, we can always make a type $\left(x', \overline{\theta}_{x'}(\omega^*)\right)$ agent prefer the deviation contract. Alternatively, if the allocation is such that agents of type $\left(x', \overline{\theta}_{x'}(\omega^*)\right)$ prefer the contract $(p,t)_{x,\overline{\theta}_x(\omega^*)}$ over any other in $M_x^{\mathbf{a}}$—that is, $(p,t)_{x,\overline{\theta}_x(\omega^*)} \in C^x_{x',\overline{\theta}_{x'}(\omega^*)}$ for every $x' \succ x$—then we need not make further assumptions about agent preferences.

Our first result establishes conditions under which we can implement $\mathbf{a}$ by adding a single deviation contract to each menu $M_x^{\mathbf{a}}$. The argument mirrors that in Theorem 1. We construct a deviation contract that agrees with the intended contract for type $\overline{\theta}_x(\omega^*)$ except in event ω^*, and we leverage that some agents in higher categories have a higher marginal value transfers in event ω^*. Some higher risk types choose the deviation contract over any other contract in $M_x^{\mathbf{a}}$, and this is too unprofitable for the firm to offer a lower category's menu to agents in a higher category.

Lemma 1. *Suppose Assumption 2 holds, and* $\mathbf{a}$ *is both agent incentive compatible and upward incentive compatible. If either*

(a) Assumption 3 holds, or

(b) $(p,t)_{x,\overline{\theta}_x(\omega^*)} \in C^x_{x',\overline{\theta}_{x'}(\omega^*)}$ *for every* $x' \succ x$,

then there exists a collection of contracts $\{(\overline{p}_x, \overline{t}_x)\}_{x\in X}$ *such that the regulatory policy*

$$\mathcal{R} = \left\{ M_x^{\mathbf{a}} \cup \{(\overline{p}_x, \overline{t}_x)\} \right\}_{x\in X}$$

implements the allocation $\mathbf{a}$.

Proof. See Appendix. □

Though upward incentive compatibility is a relatively weak condition, one can imagine situations in which it fails. Even with a clear risk ordering of categories, it might be that different categories suffer drastically different losses in the same verifiable event. Nevertheless, we can still implement an agent incentive compatible allocation if we augment Assumption

2 and add a second deviation contract to each menu. While the first deviation contract makes it unprofitable to offer menus intended for low risk categories to higher risk categories, the second deviation contract makes it unprofitable to offer menus intended for high risk categories to lower risk categories. To ensure this is possible, we need an analogous order condition for "minimal" types within each category. Define

$$\underline{r}_x(\omega) = \min_{\theta \in \Theta_x} \frac{\theta(\omega)}{\theta(\omega_0)}, \quad \underline{\theta}_x = \arg\min_{\theta \in \Theta_x} \frac{\theta(\omega)}{\theta(\omega_0)}.$$

The following assumption allows us to eliminate the upward incentive compatibility condition.

Assumption 4. *There exists an event $\omega^* \in \Omega$ such that $u(z, \omega^*, x)$ is supermodular in z and x and $\overline{r}_x(\omega^*), \underline{r}_x(\omega^*) > 0$ are both weakly increasing in x.*

Analogous to Assumption 3, we may need a condition to ensure that for $x' \succ x$, agents of type $\left(x', \overline{\theta}_{x'}(\omega^*)\right)$ prefer a deviation contract to *all* elements of $M_x^{\mathbf{a}}$, and for $x' \prec x$, agents of type $(x', \underline{\theta}_{x'}(\omega^*))$ prefer a deviation contract to *all* elements of $M_x^{\mathbf{a}}$.

Assumption 5. *The values $\overline{r}_x(\omega^*)$ and $\underline{r}_x(\omega^*)$ are strictly increasing in x, and $\lim_{z \to \infty} u(z, w^*, x) = \infty$ for every x.*

Alternatively, if $(p,t)_{x, \overline{\theta}_x(\omega^*)} \in C^x_{x', \overline{\theta}_{x'}(\omega^*)}$ for every $x' \succ x$, and $(p,t)_{x, \underline{\theta}_x(\omega^*)} \in C^x_{x', \underline{\theta}_{x'}(\omega^*)}$ for every $x' \prec x$, we need not impose this assumption. These conditions allow us eliminate upward incentive compatibility at the cost of one additional deviation contract in each menu.

Lemma 2. *Suppose Assumption 4 holds, and* $\mathbf{a}$ *is agent incentive compatible. If either*

(a) Assumption 5 holds, or

(b) $(p,t)_{x, \overline{\theta}_x(\omega^*)} \in C^x_{x', \overline{\theta}_{x'}(\omega^*)}$ for every $x' \succ x$ and $(p,t)_{x, \underline{\theta}_x(\omega^*)} \in C^x_{x', \underline{\theta}_{x'}(\omega^*)}$ for every $x' \prec x$,

then there exists a collection of contracts $\{(\overline{p}_x, \overline{t}_x), (\underline{p}_x, \underline{t}_x)\}_{x \in X}$ such that the regulatory policy

$$\mathcal{R} = \left\{ M_x^{\mathbf{a}} \cup \{(\overline{p}_x, \overline{t}_x), (\underline{p}_x, \underline{t}_x)\} \right\}_{x \in X}$$

implements the allocation $\mathbf{a}$.

Proof. See Appendix. □

Each Lemma provides a simple risk-ordering condition that is sufficient to implement an agent incentive compatible allocation using relatively small menus. The most crucial assumption is that there exists some loss event that is more likely, relative to the no-loss event, for higher categories. While condition (b) in the two lemmas can be cumbersome to check, in the most natural applications we seek to pool agents within each category—in such cases, the condition is trivially satisfied. In the next subsection, we apply this implementation result to show that a regulator can implement her first-best outcome in a much broader range of problems.

4.2 Implementing First-Best

We first derive a key property of the regulator's first-best allocation: under our maintained assumptions on u, all agents in the same category receive the same contract. Recall that our definition of the first-best allocation means solving the regulator's problem subject only to the firm's participation constraint $\pi(\mathbf{a}) \geq 0$. This feature of Proposition 1 extends easily to the general case. Because u is strictly concave in the net transfer, we can always construct an improvement if two types within the same category receive different contracts. Moreover, first-best requires that we equalize the marginal utility of transfers across loss events categories.

Proposition 2. *In the first-best allocation, agents in category x buy the contract (p^*, t_x^*) where $t_x^*(\omega)$ is set so that the marginal utility of wealth is equal across all ω, x.*

Proof. See Appendix. □

As in the model of section 3, the first-best allocation gives the same contract to all agents within the same category, and it charges a uniform price p^*. This feature greatly simplifies the application of our implementation results. First, since $M_x^{\mathbf{a}}$ contains a single contract in the first-best allocation $\mathbf{a}$, we trivially satisfy agent incentive compatibility. Moreover, Assumptions 3 and 5 are automatically satisfied via condition (b). Hence, one need only check the risk ordering condition in Assumption 2 or 4 in order to apply one of our lemmas. In the latter case, implementation requires menus with 3 contracts, rather than 2.

Definition 6. *A* ***3-option regulatory policy*** *is a collection of menus $\{M_x\}_{x \in X}$ in which each menu $M_x = \{(p_x, t_x), (\overline{p}_x, \overline{t}_x), (\underline{p}_x, \underline{t}_x)\}$ consists of three choices.*

Theorem 2. *Suppose Assumption 2 holds, and the first-best allocation* $\mathbf{a}$ *is upward incentive compatible. There exists a 2-option regulatory policy that implements* $\mathbf{a}$*.*

Suppose Assumption 4 holds. There exists a 3-option regulatory policy that implements the first-best allocation $\mathbf{a}$*.*

Proof. This is immediate from Lemmas 1 and 2. □

Even in the general model, a straightforward risk ordering condition on categories is sufficient to implement the regulator's first-best allocation using simple menus. Upward incentive compatibility is potentially hard to check, but there is a simple sufficient condition: If $u(z, \omega, x)$ is supermodular in x, z for all x and ω, then equalizing marginal utilities across categories entails higher transfers for higher x in every event ω. Note this condition is stronger than what Assumptions 2 and 4 require—we need supermodularity for every event ω, not just in event ω^*.

5 Constrained Regulation

Two important policy questions are whether we should allow firms to exclude customers and whether we should mandate buying insurance. Our implementation crucially depends both on preventing exclusion and on an individual mandate—the firm offers insurance to all agents, and all agents buy insurance. Optimal policy requires both features, but regulators may not always have the ability to enforce them. This section explores policies when the regulator faces more constraints.

We can capture many natural restrictions through constraints on the space of permissible policies. Here are a few examples:

- All agents must have the same options. In this case, the regulator is constrained to policies $\mathcal{R}$ that contain a single menu M. Such a constraint might arise as the result of anti-discrimination laws, or strong fairness norms. The regulator's problem is then to optimize within this restricted class of policies

- The regulator can enforce contracts but not menus. If a regulator is unable verify whether the firm includes particular contracts in the menus it privately offers to agents, then the firm may be able to construct its own menus from some grand set of permissible contracts. In this case, the regulator is constrained to policies of the form $\mathcal{R} = 2^M \setminus \{\emptyset\}$ for some menu M.

- Firms can refuse service. In the absence of strong laws mandating that the firm must offer insurance to all agents, the regulator faces the constraint that $\{(0,0)\} \in \mathcal{R}$ in any policy: the firm retains the right to offer the null menu.

- Agents can opt out. In the absence of an individual insurance mandate, the regulator faces the constraint that $(0,0) \in M$ for all $M \in \mathcal{R}$.

This last constraint is particularly relevant to recent debates in the United States. The next subsection highlights how insurance mandates might harm agent welfare if the regulator is unable to ensure that the firm offers specific menus.

5.1 Insurance Mandates: A Double-Edged Sword

To implement our optimal regulatory policy, the regulator must be able to ensure that whenever the firm offers a particular contract, it also offers the corresponding deviation contract. Absent this enforcement power, the firm could construct its own menus by picking and choosing contracts that appear in some permitted menu. It should be clear that this would undermine the regulator's goals. In the model of section 3, the firm would offer the lowest coverage contract to all categories, leading to positive profits and lower welfare. This section explores how a lack of enforcement power interacts with an individual mandate to buy insurance.

Definition 7. *A **regulatory policy with unenforceable menus** comprises a finite set of contracts Y. If there is an **insurance mandate**, the corresponding regulatory policy is $\mathcal{R} = 2^Y \setminus \{\emptyset\}$. If there is **no insurance mandate**, the corresponding regulatory policy is $\mathcal{R} = \{M \cup \{(0,0)\} : M \in 2^Y \setminus \{\emptyset\}\}$.*

With unenforceable menus, once a contract is permitted in some menu, the regulator cannot prevent the firm from including it, or excluding it, in any offer it makes to an agent. In general, the firm's problem becomes complicated because the set of possible menus is large, and the firm may want to screen within each category. To simplify the analysis, for the rest of this section we assume that there is a single risk type within each category (i.e. $|\Theta_x| = 1$ for every $x \in X$). Within this simplified setting, we show that an insurance mandate can harm agent welfare. Intuitively, the ability of agents to opt out imposes some discipline on what the firm offers, and this can more than offset the benefit of universal coverage.

Consider the setting of section 3 in which agents may suffer known monetary losses. Suppose there are two categories $x \in \{L, H\}$ and two loss events $\omega \in \{\omega_0, \omega_1\}$. In event ω_1, category x agents suffer a loss ℓ_x, where $\ell_H > \ell_L$, with probability θ_x, where $\theta_H > \theta_L$. A fraction μ of agents are category L. We first consider the case in which *insurance is mandatory.* Since there is no need for screening, without loss the firm offers a single contract to each agent, and the universe of permissible contracts is

$$Y = \{(p_H, t_H), (p_L, t_L)\}.$$

Notice that agents have no choice here—they must accept whatever contract the firm offers. The only incentive constraints are those of the firm. The regulator chooses Y to solve

$$\begin{aligned} \max \quad & \mu U(L, \theta_L, p_L, t_L) + (1-\mu) U(H, \theta_H, p_H, \theta_H) & \\ s.t. \quad & \mu(p_L - \theta_L t_L) + (1-\mu)(p_H - \theta_H t_H) \geq k & (FPC) \\ & p_L - \theta_L t_L \geq p_H - \theta_L t_H & (FIC-L) \\ & p_H - \theta_H t_H \geq p_L - \theta_H t_L & (FIC-H) \end{aligned}$$

The regulator must satisfy the firm's participation constraint (FPC) and incentive constraints to offer the correct contract to each type (FIC-L and FIC-H).

Under these assumptions, Whenever insurance is mandatory, the firm offers the same contract to both categories. The firm IC constraints imply

$$\theta_L(t_H - t_L) \geq p_H - p_L \geq \theta_H(t_H - t_L).$$

The regulator wants the firm to offer higher coverage to category H, but this only happens if both categories get the same contract.

Proposition 3. *In the above known monetary loss model with two types and two loss events, any optimal regulatory policy in which insurance is mandatory involves the firm offering the same contract to every agent: $p_H = p_L = p$ and $t_H = t_L = t$.*

Proof. See Appendix. □

If every agent receives the same contract, then at least one category is under-insured or over-insured. The regulator chooses a single contract (p,t) to maximize expected welfare. Since the firm's participation constraint binds, this amounts to choosing $t \in [\ell_L, \ell_H]$ and setting

$$p = t(\mu\theta_L + (1-\mu)\theta_H) + k.$$

This result highlights a problem with insurance mandates under a weak regulator. If the regulator cannot force the firm to include deviation contracts in the menus, then the firm can hold agents hostage, offering only the most expensive contract or the lowest coverage level. The firm will not offer higher coverage to category H, even at a higher price, because if doing so is profitable, then it is even more profitable to offer the expensive contract to category L.

Allowing agents to opt out of insurance can help because it allows us to target different contracts to different categories—if the contract intended for category H is too expensive for category L, then the firm is willing to offer a lower cost or lower coverage option to those agents. However, this entails a trade-off as there is less cross-subsidization across agents. Which effect is more important depends on the particular parameters. A numerical example shows that allowing agents to opt out can improve welfare.

Example: Optimality of Optional Insurance

Suppose

$$u(z,\omega,x) = (100 - p + t(\omega) - \ell_x(\omega))^{1/2}.$$

Assume that the two categories L and H are equally prevalent, that $\theta_L = 0.5$ and $\theta_H = 0.6$, and that $\ell_L = 30$ and $\ell_H = 90$. If insurance is mandatory, the optimal policy prescribes a single contract $(p^*, t^*) \approx (38.5, 70)$ for all agents. Category L is over-insured, category H is under-insured, and the average agent utility is approximately 7.98.

If insurance is optional, the regulator can do better by allowing the contracts $(15, 30)$ and $(54, 90)$. Category H prefers either one to the null contract, but the former is unprofitable since $\theta_H > \frac{1}{2}$, so the firm offers $(54, 90)$. Category L on the other hand would rather go uninsured than pay the high premium, and the firm is still willing to offer $(15, 30)$. All agents are fully insured, and the average agent utility is approximately 8.61.

6 Market Structure and the Role of Competition

In real insurance markets, firms typically face some degree of competition. From an efficiency perspective, a regulator should prefer a single firm due to fixed costs. One can extend our analysis to a market with many firms under the assumption that all entrants split the market equally. In this case, the optimal regulatory policy is unchanged from the one-firm case, and only one firm enters the market in equilibrium—if more firms enter, some firm must make a negative expected profit. While this is reassuring, the assumption of equal splitting may not be reasonable in practice. In a decentralized market, an entrant can selectively target agents through advertising. This may facilitate cream-skimming, even if the entrant is legally required to offer insurance to all agents. Cream-skimming is a serious concern because this possibility can undermine incentives for the first firm to enter, or force incumbents out of the market.[9]

Assume our monopolist (the incumbent firm) must serve all consumers, but there is a potential entrant who may target advertisements to particular categories. Though the entrant must also serve all consumers, we suppose that an agent defaults to the incumbent unless she sees an advertisement for the entrant. The entrant offers the same menus as the incumbent firm, and agents select the same option from a given menu regardless of what firm they choose. Suppose there is a fixed cost to enter the market κ, and an additional constant marginal cost c per agent served—this cost is separate from contract payouts, capturing things like administrative expenses and capital requirements. Hence, we can decompose the fixed cost of a single firm serving the entire market as $k = \kappa + c$. If the entrant captures a market share $\alpha \in [0, 1]$ it incurs costs $\kappa + \alpha c$, while the incumbent incurs costs $\kappa + (1 - \alpha)c$.

Cream-skimming can render our first-best outcome infeasible. Consider again the known monetary loss model with two categories $x \in \{H, L\}$ and two loss events $\omega \in \{\omega_0, \omega_1\}$—an agent of category x suffers loss ℓ_x in event ω_1, which occurs with probability θ_x. In the first-best allocation, both types get full insurance at a common price p^*, and the incumbent's participation constraint binds. Selling to all consumers never covers the entrant's fixed costs, but targeting only category L, where $\ell_L \leq \ell_H$ and $\theta_L \leq \theta_H$, can be profitable. If the entrant claims a share $\alpha \in [0, 1]$ of category L agents, then entry is strictly profitable if

$$\alpha \mu(L, \theta_L)(p^* - \theta_L \ell_L - c) > \kappa.$$

If fixed costs are sufficiently small, or the difference between types is sufficiently large, entry is profitable, and the first-best allocation cannot be an equilibrium.

How well can the regulator do when entrants might cream-skim? Let $X_c \subset X$ denote the set of categories the entrant chooses to serve, and suppose the entrant can claim a share $\alpha_x \in [0, 1]$ of category x. The entrant can then earn a profit of

$$\pi_e(\mathbf{a}) = \max_{X_c \subset X} \sum_{x \in X_c} \alpha_x \sum_{\theta \in \Theta} \mu(x, \theta)(\Pi(x, \theta, (p, t)_{x,\theta}) - c) - \kappa.$$

Entry is not a problem as long as the allocation satisfies a no cream-skimming constraint:

$$\pi_e(\mathbf{a}) \leq 0 \qquad (FCS).$$

[9] This issue has been noted since at least the seminal work of Rothschild and Stiglitz [1976].

The natural analog to our first-best allocation, given the no cream-skimming constraint, is the solution to

$$\begin{aligned} \max_{\mathbf{a}} \quad & W(\mathbf{a}) \\ s.t. \quad & \pi(\mathbf{a}) \geq 0 \\ & \pi_e(\mathbf{a}) \leq 0 \end{aligned} \tag{3}$$

We solve the regulator's problem subject only to the incumbent firm's participation constraint $\pi(\mathbf{a}) \geq 0$ and the entrant firm's no cream-skimming constraint $\pi_e(\mathbf{a}) \leq 0$.

In this version of a second-best allocation, the regulator targets a single contract to each category, offering full insurance, but the price can differ across categories. Offering lower prices to lower risk agents is necessary to prevent cream-skimming, but this also limits cross-subsidzation.

Proposition 4. *The solution to* (3) *provides a contract* (p_x, ℓ_x) *to all types in category* x*. If Assumption 4 holds, there exists a 3-option regulatory policy that implements this allocation.*

Proof. The argument that the second-best allocation provides the same contract to all agents in the same category is analogous to Proposition 2, and we omit it. Implementability follows from Lemma 2.[10] □

Our second-best allocation (3) implicitly assumes the regulator wants to prevent entry. If there are many potential entrants, and each has the same cream-skimming ability, this is without loss—regardless of how many firms enter, the allocation must still satisfy the cream-skimming constraint, and every firm after the first only tightens the participation constraint. With only one potential entrant, the regulator might allow entry in some situations. However, if fixed costs are high, or the entrant has a high ability to cream-skim, doing so is never optimal.

7 Discussion

Many natural extensions to our framework are straightforward. Little changes if we assume compact—rather than finite—sets of categories and types, or if the regulator puts some positive weight on the firm's profits. Moreover, one can further extend the implementation result to allow even weaker order conditions on agent categories if one is willing to include additional deviation contracts in each menu.

Our model assumes that agents have more information about their own risks than the firm does. Given firms' increasing sophistication with regard to consumer data, questioning

[10] Applying Lemma 1 is delicate in this setting because charging different prices to different categories makes it difficult to check upward incentive compatibility.

this assumption seems reasonable. In principle, an individual who pays attention to her own health and conducts research into risk factors should be more informed than a firm. However, few invest significant effort in acquiring such information. Nevertheless, as long as *some* individuals exert this effort, our implementation result is robust. Imagine that some agents are uninformed, knowing neither their categories nor their risk types, but some agents, comprising a positive fraction of the population, are. We can still implement first-best using two-option menus—we simply need to make the deviation contract even more costly for the firm. In the corresponding equilibrium, uninformed agents choose the intended contract from any menu, while informed agents choose the contract that is optimal for them.

Another possible objection is that our construction relies on extremely high payouts in a single, and potentially rare, loss event—this feature might create problems for implementation in practice. There are at least two reasons why such an objection is misplaced. First, the construction in our proof is by no means unique. Depending on the exact losses and distributions, there may be many ways for a deviation contract to satisfy incentives. If higher risk categories have higher losses in many events, then a deviation contract could offer larger but not extreme payouts in each of these events. Second, the better the firm can distinguish risk levels, the less meaningful this objection becomes. As firms gather more informative data about their customers, it should get easier to design appropriate deviation contracts without extreme payouts.

This last point calls attention to a potentially surprising implication: regulation is easier and more effective when the firm has more information about agents. The regulator can leverage the firm's knowledge by including more carefully tailored contracts—if categories allow better discrimination between loss amounts the agents can suffer, then the first-best allocation entails better risk-sharing. Notice also that the firm has incentives to obtain *less* information. If the firm could commit not to use some information about the agents, it would pool the categories in a way that violates our order assumptions. Without this commitment power, accurate information about risk becomes a liability for the firm. This suggests that in a strong regulatory regime, the regulator should take account of firm incentives to gather information about consumers, balancing the value of information against the surplus the firm is allowed to extract.

8 Final Remarks

Regulation in insurance markets has large welfare implications, and the increasing prevalence of big data provokes new questions for regulatory policy. We introduce a novel framework to study regulation, and our analysis highlights important features of successful regulatory regimes. Under mild conditions, a simple policy that combines an individual insurance mandate, a ban on exclusion from insurance, and two-contract menus can implement a socially optimal allocation. All three features are crucial. In particular, if the regulator cannot ensure that the firm offers the desired *menus*, the other pieces can become bad for welfare.

Firms' use of consumer data raises concerns about exploitation. In an unregulated market, information allows a firm to better extract surplus from consumers and can lead to distortions of coverage. Our results show that, properly regulated, a firm with access to more data can improve overall consumer welfare—consumers obtain coverage that is tailored to their needs, and low risk consumers subsidize those who face higher risks. Without data that identifies higher risk agents, this first-best outcome is not achievable. Optimal regulation leverages the firm's information for the benefit of consumers while simultaneously leveraging consumers' incentives to discipline the firm's actions.

Our results suggest several promising avenues for further work. Section 7 alludes to potential issues related to firm incentives for gathering information. An analysis of the trade-off between providing such incentives and the social value of information thereby obtained could prove enlightening. We have emphasized at several points the practical importance of simple menus and simple policies. However, firms may have detailed data distinguishing many different categories, which according to our optimal policy necessitates a large set of menus carefully tailored to each. Natural questions include how closely the regulator can approximate the optimal policy using a limited number of menus, and how well regulation can perform when the regulator is unsure how much information the firm has.

Beyond these immediate questions, our framework highlights a new approach to studying the regulation of mechanisms, a relatively unexplored subject. Outside the literature on price regulation and subsidization of monopolists, there is little work on how regulators might constrain the space of permitted mechanisms. In many other domains, authorities restrict the contracts into which parties can enter. For instance, employers are subject to wage and anti-discrimination regulations. Taxi services and hotels are subject to certification and insurance requirements. Exploring the interplay between contracting parties' incentives, and how contract restrictions can take advantage of this, may lead to new and better regulatory approaches.

References

Big data and insurance: Implications for innovation, competition and privacy. Technical report, The Geneva Association, March 2018.

Ricardo Alonso and Niko Matouschek. Optimal Delegation. *Review of Economic Studies*, 75(1):259–293, 2008.

Manuel Amador and Kyle Bagwell. The Theory of Optimal Delegation With an Application to Tariff Caps. *Econometrica*, 81(4):1541–1599, 2013.

David P. Baron and Roger B. Myerson. Regulating a Monopolist with Unknown Costs. *Econometrica*, 50(4):911–930, 1982.

Hector Chade and Edward Schlee. Optimal Insurance with Adverse Selection. *Theoretical Economics*, 7(3):571–607, 2012.

Alexander Chernev, Ulf Böckenholt, and Joseph Goodman. Choice Overload: A Conceptual Review and Meta-Analysis. *Journal of Consumer Psychology*, 25(2):333–358, 2015.

Liran Einav and Amy Finkelstein. Selection in Insurance Markets: Theory and Empirics in Pictures. *Journal of Economic Perspectives*, 25(1):115–138, 2011.

Simone Galperti. Commitment, Flexibility, and Optimal Screening of Time Inconsistency. *Econometrica*, 83(4):1425–1465, 2015.

Bengt Holmström. On the theory of delegation. In M. Boyer and R. Kihlstrom, editors, *Bayesian Models in Economic Theory*, pages 115–141. New York: North-Holland, 1984.

Jean-Jacques Laffont. The New Economics of Regulation Ten Years After. *Econometrica*, 62(3):507–537, 1994.

Jean-Jacques Laffont and Jean Tirole. Using Cost Observation to Regulate Firms. *Journal of Political Economy*, 94(3):614–641, 1986.

Jean-Jacques Laffont and Jean Tirole. Auctioning Incentive Contracts. *Journal of Political Economy*, 95(5):921–937, 1987.

Tracy R. Lewis and David E. M. Sappington. Regulating a Monopolost with Unknown Demand. *American Economic Review*, 78(5):986–998, 1988.

Karen Pollitz, Cynthia Cox, and Rachel Fehr. Claims Denials and Appeals in ACA Marketplace Plans. *Kaiser Family Foundation*, 2019. URL `https://www.kff.org/private-insurance/issue-brief/claims-denials-and-appeals-in-aca-ma`

Michael H. Riordan and David E. M. Sappington. Awarding Monopoly Franchises. *American Economic Review*, 77(3):375–387, 1986.

William P. Rogerson. Simple Menus of Contracts in Cost-Based Procurement and Regulation. *American Economic Review*, 93(3):919–926, 2003.

Michael Rothschild and Joseph Stiglitz. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, 90 (4):629–649, 1976.

Joseph E. Stiglitz. Monopoly, Non-linear Pricing and Imperfect Information: The Insurance Market. *Review of Economic Studies*, 44(3):407–430, 1977.

A Appendix

Proof of Lemmas 1 and 2

Throughout this proof we suppress the dependence of $\overline{\theta}_x(\omega^*)$ and $\overline{r}_x(\omega^*)$ on the event ω^*, writing simply $\overline{\theta}_x$ and $\overline{r}_x$. We start with Lemma 1. Following the proof of Theorem 1, we construct a deviation contract $(\overline{p}_x, \overline{t}_x)$ so that type $\left(x, \overline{\theta}_x\right)$ is indifferent between the intended contract $(p,t)_{x,\overline{\theta}_x}$ and the deviation contract. This means that

$$\sum_{\omega\in\Omega} u(t_{x,\overline{\theta}_x}(\omega) - p_{x,\overline{\theta}_x}, \omega, x)\overline{\theta}_x(\omega) = \sum_{\omega\in\Omega} u(\overline{t}_x(\omega) - \overline{p}_x, \omega, x)\overline{\theta}_x(\omega).$$

We choose $(\overline{p}_x, \overline{t}_x)$ so that

$$\overline{t}_x(\omega) - \overline{p}_x = \begin{cases} t_{x,\overline{\theta}_x}(\omega) - p_{x,\overline{\theta}_x} & \text{if } \omega \notin \{\omega_0, \omega^*\} \\ -\overline{p}_x & \text{if } \omega = \omega_0 \\ \hat{t}_x - \overline{p}_x & \text{if } \omega = \omega^*. \end{cases}$$

Hence, the indifference condition reduces to

$$\overline{r}_x \left(u\left(\hat{t}_x - \overline{p}_x, \omega^*, x\right) - u\left(t_{x,\overline{\theta}_x}(\omega^*) - p_{x,\overline{\theta}_x}, \omega^*, x\right)\right) = u\left(-p_{x,\overline{\theta}_x}, \omega_0, x\right) - u\left(-\overline{p}_x, \omega_0, x\right).$$

As long as $\overline{p}_x$ is not so large that higher transfers in event ω^* cannot compensate, this equation defines a unique $\hat{t}_x(\overline{p}_x)$. Moreover, it should be clear that $\hat{t}_x(\overline{p}_x) \to \infty$ as $\overline{p}_x \to \underline{z}$.

To complete the proof, we show that for sufficiently high $\overline{p}_x$, and all $x' \succ x$, we have

(a) Agents of type $(x', \overline{\theta}_{x'})$ choose $(\overline{p}_x, \overline{t}_x)$ over any contract in $M_x^{\mathbf{a}}$,

(b) All category x agents prefer their intended contract to the deviation contract, and

(c) Contract $(\overline{p}_x, \overline{t}_x)$ is sufficiently unprofitable that the firm prefers the intended allocation over allowing type $(x', \overline{\theta}_{x'})$ to choose the deviation contract.

These three properties are sufficient for implementation as upward incentive compatibility ensures the firm does not want to offer $M_x^{\mathbf{a}}$ to categories $x' \prec x$, and agent incentive compatibility ensure agents in category x want to select their intended contracts from $M_x^{\mathbf{a}}$.

For property (a), the definition of $\hat{t}_x(\overline{p}_x)$ together with Assumption 2 ensures that type $(x', \overline{\theta}_{x'})$ prefers the deviation contract to the contract $(p,t)_{x,\overline{\theta}_x}$. If Assumption 3 part (b) holds, then we are done with the first property. If Assumption 3 part (a) holds, then we can increase $\overline{p}_x$ to ensure that type $(x', \overline{\theta}_{x'})$ prefers the deviation contract to any other $(p,t) \in M_x^{\mathbf{a}}$. Since $u(z, \omega^*, x')$ is unbounded, and $\overline{r}_{x'} > \overline{r}_x$, the payoff difference between the deviation contract and $(p,t)_{x,\overline{\theta}_x}$ becomes unboundedly large. Since the payoff difference between $(p,t)_{x,\overline{\theta}_x}$ and any other contract in $M_x^{\mathbf{a}}$ is finite, we can clearly choose $\overline{p}_x$ high enough that type $(x', \overline{\theta}_{x'})$ will choose the deviation contract.

Property (b) comes for free because the definition of $\overline{r}_x$ ensures that all category x agents weakly prefer $(p,t)_{x,\overline{\theta}_x}$ to the deviation contract, and we assume that **a** is agent incentive compatible. Finally, for property (c), we bound the profit from offering menu $M_x^{\mathbf{a}} \cup (\overline{p}_x, \overline{t}_x)$ to agents in category x'. This profit is at most some constant plus

$$\mu_{x'}(\overline{\theta}_{x'}) \left(\overline{p}_x - \overline{\theta}_{x'}(\omega^*)\hat{t}_x(\overline{p}_x)\right).$$

Since $\hat{t}_x(\overline{p}_x)$ approaches ∞ as $\overline{p}_x$ approaches $\underline{z}$, this becomes arbitrarily negative for sufficiently large $\overline{p}_x$.

The proof of Lemma 2 is substantively identical. The construction of the deviation contract $(\overline{p}_x, \overline{t}_x)$ is exactly the same. As we no longer assume upward incentive compatibility, we need a second deviation contract in each menu $(\underline{p}_x, \underline{t}_x)$ that appeals to type $(x', \underline{\theta}_{x'})$ for $x' \prec x$. The additional conditions in Assumptions 4 and 5 allow us to carry out the analogous construction. $\square$

Proof of Proposition 2

Let $u_z(z,\omega,x) := \frac{\partial u(z,\omega,x)}{\partial z}$ be the marginal utility of wealth in state ω for type x. Suppose that for some ω, ω' and $(x,\theta_x), (x',\theta_{x'})$, allowing either $\omega' = \omega'$ or $(x,\theta_x) = (x',\theta_{x'})$, we have

$$u_z(t^x_{x,\theta}(\omega) - p^x_{x,\theta}, \omega, x) > u_z(t^{x'}_{x',\theta'}(\omega') - p^{x'}_{x',\theta'}, \omega', x').$$

We show that it is possible to increase the average utility of the agents while holding the principal's profit constant.

Define the set of triples

$$A = \underset{(x,\theta_x,\omega)}{\arg\max}\ u_z(t^x_{x,\theta}(\omega) - p^x_{x,\theta}, \omega, x),$$

and let

$$Q = \sum_{(x,\theta,\omega)\in A} \theta(\omega)\mu(x,\theta_x)$$

denote the total probability of all such triples (x,θ,ω). Define an alternative allocation with the same prices, but different transfers

$$\hat{t}^x_{x,\theta}(\omega) = \begin{cases} t^x_{x,\theta_x}(\omega) - \frac{\epsilon}{1-Q} & \text{if } (x,\theta_x,\omega) \notin A \\ t^{x'}_{x',\theta'}(\omega') + \frac{\epsilon}{Q} & \text{if } (x,\theta_x,\omega) \in A. \end{cases}$$

By construction, this new allocation gives the same profit to the principal.

We now verify that for sufficiently small ϵ, the new allocation yields a strict improvement for the regulator. The new allocation gives welfare

$$
\begin{aligned}
W = & \sum_{(x,\theta,\omega)\in A} u\left(t^x_{x,\theta}(\omega) - p^x_{x,\theta} + \frac{\epsilon}{Q}, \omega, x\right)\mu(x,\theta)\theta(\omega) \\
& + \sum_{(x,\theta,\omega)\notin A} u\left(t^x_{x,\theta}(\omega) - p^x_{x,\theta} - \frac{\epsilon}{1-Q}, \omega, x\right)\mu(x,\theta)\theta(\omega)
\end{aligned}
$$

At $\epsilon = 0$, this is equal to the welfare from the original allocation. Taking the derivative with respect to ϵ at $\epsilon = 0$ gives

$$
\begin{aligned}
\left.\frac{\mathrm{d}W}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = & \frac{1}{Q} \sum_{(x,\theta,\omega)\in A} u_z\left(t^x_{x,\theta}(\omega) - p^x_{x,\theta}, \omega, x\right)\mu(x,\theta)\theta(\omega) \\
& - \frac{1}{1-Q} \sum_{(x,\theta,\omega)\notin A} u_z\left(t^x_{x,\theta}(\omega) - p^x_{x,\theta}, \omega, x\right)\mu(x,\theta)\theta(\omega),
\end{aligned}
$$

which is strictly positive because every marginal utility in the first sum is strictly larger than every marginal utility in the second. Hence, the original allocation was not optimal, and we conclude that the marginal utilities must be equal in an optimal allocation. Since utility in the no-loss state ω_0 is the same for all types, this implies that all types must pay the same price p^*.

Proof of Proposition 3

Let $(p_H, t_H), (p_L, t_L)$ be an optimal policy. The firm's participation constraint clearly binds. The firm's IC constraints require that

$$\theta_L(t_H - t_L) \geq p_H - p_L \geq \theta_H(t_H - t_L).$$

If $t_H \geq t_L$, this can only be true if the two contracts are identical since $\theta_H > \theta_L$. We show that $t_H \geq t_L$ in any optimal allocation.

Suppose $t_L > t_H$. At most one of the two IC constraints can bind. Suppose

$$p_H - p_L = \theta_L(t_H - t_L) > \theta_H(t_H - t_L),$$

meaning that FIC-H is slack. The Lagrangian for the regulator's problem is

$$
\begin{aligned}
\mathcal{L} = & \mu\left[(1-\theta_L)u(w - p_L) + \theta_L u(w - p_H + t_L - \ell_L)\right] \\
& + (1-\mu)\left[(1-\theta_H)u(w - p_H) + \theta_H u(w - p_H + t_H - \ell_H)\right] \\
& + \lambda\left[\mu(p_L - \theta_L t_L) + (1-\mu)(p_H - \theta_H t_H)\right] + \gamma(p_L - \theta_L t_L - p_H + \theta_L t_H),
\end{aligned}
$$

where $\lambda \geq 0$ is the multiplier for the participation constraint, and $\gamma \geq 0$ is the multiplier for the lone IC constraint. The necessary first-order conditions with respect to t_L and t_H are

$$u'(w - p_L + t_L - \ell_L) = \lambda + \frac{\gamma}{\mu}, \quad u'(w - p_H + t_H - \ell_H) = \lambda - \frac{\gamma}{1-\mu}\frac{\theta_L}{\theta_H}.$$

Since $p_H - p_L > \theta_H(t_H - t_L) > t_H - t_L$, we have

$$w - p_H + t_H - \ell_H < w - p_L + t_L - \ell_H < w - p_L + t_L - \ell_L.$$

This implies that marginal utility marginal utility in the loss event is higher for type H than for type L. From the first order conditions, this implies

$$\frac{\gamma}{\mu} < -\frac{\gamma}{1-\mu}\frac{\theta_L}{\theta_H},$$

which is impossible. We conclude that FIC-L is slack: $p_H - p_L < \theta_L(t_H - t_L)$.

We next show that $t_H = \ell_H$. If $t_H > \ell_H$, then H has higher wealth in the loss event than in the no-loss event. Construct an alternative policy (p_L, t_L), (p'_H, t'_H) with $p'_H = p_H - \epsilon$ and $t'_H = t_H - \frac{\epsilon}{\theta_H}$. By construction, this leaves the firm's profit unchanged, and FIC-H still holds. FIC-L also holds because

$$p_L - \theta_L t_L \geq p_H - \theta_L t_H > p_H - \theta_L p_H - \epsilon + \frac{\theta_L}{\theta_H}\epsilon.$$

Concavity of u implies this allocation yields a strict welfare improvement. If $t_H < \ell_H$, then H has higher wealth in the no-loss event than in the loss event. Construct an alternative policy (p_L, t_L), (p'_H, t'_H) with $p'_H = p_H + \epsilon$ and $t'_H = t_H + \frac{\epsilon}{\theta_H}$. By construction, this leaves the firm's profit unchanged, and FIC-H still holds. Because the constraint was slack, FIC-L also holds for sufficiently small ϵ. Concavity of u again implies this allocation yields a strict welfare improvement. We conclude that $t_H = \ell_H$. Moreover, this implies that $t_L > \ell_H > \ell_L$, so L has higher wealth in the loss event than in the no-loss event.

To complete the proof, we consider two cases. First, suppose FIC-H is slack. Construct an alternative policy (p'_L, t'_L), (p_H, t_H) with $p'_L = p_L - \epsilon$ and $t'_L = t_L - \frac{\epsilon}{\theta_L}$. By construction, this leaves the firm's profit unchanged, and both FIC-L and FIC-H hold for small enough ϵ since both constraints were slack. Concavity of u implies this allocation yields a strict improvement, so the original allocation was not optimal.

Now suppose that FIC-H binds. The corresponding Lagrangian is

$$\begin{aligned}\mathcal{L} =& \mu\left[(1-\theta_L)u(w-p_L) + \theta_L u(w - p_H + t_L - \ell_L)\right] \\ &+ (1-\mu)\left[(1-\theta_H)u(w-p_H) + \theta_H u(w - p_H + t_H - \ell_H)\right] \\ &+ \lambda\left[\mu(p_L - \theta_L t_L) + (1-\mu)(p_H - \theta_H t_H)\right] + \gamma(p_H - \theta_H t_H - p_L + \theta_H t_L),\end{aligned}$$

where $\lambda \geq 0$ is the multiplier for the participation constraint, and $\gamma \geq 0$ is the multiplier for the lone IC constraint. The necessary first-order conditions with respect to p_L, t_L, and t_H are

$$(1-\theta_L)u'(w-p_L) + \theta_L u'(w - p_L + t_L - \ell_L) = \lambda - \frac{\gamma}{\mu},$$

$$u'(w - p_L + t_L - \ell_L) = \lambda - \frac{\gamma}{\mu}\frac{\theta_H}{\theta_L}, \quad \text{and} \quad u'(w - p_H + t_H - \ell_H) = \lambda + \frac{\gamma}{1-\mu}.$$

Substituting the second into the first gives

$$u'(w - p_L) = \lambda - \frac{\gamma}{\mu}\frac{1-\theta_H}{1-\theta_L}.$$

Note that

$$w - p_L < w - p_H = w - p_H + t_H - \ell_H < w - p_L + t_L - \ell_L,$$

which implies $u'(w - p_L) > u'(w - p_H + t_H - \ell_H)$. This means

$$-\frac{\gamma}{\mu}\frac{1-\theta_H}{1-\theta_L} > \frac{\gamma}{1-\mu},$$

which is impossible. Therefore the necessary conditions for optimality cannot be satisfied. We conclude that $t_H \geq t_L$ as desired. □